In the Name of God

Journal of
Information Systems & Telecommunication
Vol. 9, No.4, October-December 2021, Serial Number 36

Research Institute for Information and Communication Technology
Iranian Association of Information and Communication Technology
Affiliated to: Academic Center for Education, Culture and Research (ACECR)

Print ISSN: 2322-1437
Online ISSN: 2345-2773
Publication License: 91/13216
Editorial Office Address: No.5, Saeedi Alley, Kalej Intersection., Enghelab Ave., Tehran, Iran,
P.O.Box: 13145-799 Tel: (+9821) 88930150 Fax: (+9821) 88930157
E-mail: info@jist.ir , infojist@gmail.com
URL: www.jist.ir

Indexed by:

- SCOPUS	www. Scopus.com
- Index Copernicus International	www.indexcopernicus.com
- Islamic World Science Citation Center (ISC)	www.isc.gov.ir
- Directory of open Access Journals	www.Doaj.org
- Scientific Information Database (SID)	www.sid.ir
- Regional Information Center for Science and Technology (RICeST)	www.ricest.ac.ir
- Iranian Magazines Databases	www.magiran.com

Publisher:
Iranian Academic Center for Education, Culture and Research **(ACECR)**

This Journal is published under scientific support of
Advanced Information Systems (AIS) Research Group and
Telecommunication Research Group, ICTRC

Acknowledgement

JIST Editorial-Board would like to gratefully appreciate the following distinguished referees for spending their valuable time and expertise in reviewing the manuscripts and their constructive suggestions, which had a great impact on the enhancement of this issue of the JIST Journal.

(A-Z)

- Agahi, Hamed, Islamic Azad University of Shiraz, Iran
- Agarwal, Parul, Jamia Hamdard University, New Delhi, India
- Ahmadizad, Arman, University of Kurdistan, Kurdistan, Iran
- Ahmadifar,Hamidreza, University of Guilan, Rasht, Iran
- Alavi, Seyed Enayatallah, Shahid Chamran University, Ahvaz, Iran
- Alam Tabriz, Akbar, Shahid Beheshti University, Tehran, Iran
- Azizi, Sadoon, University of Kurdistan, Sanandaj, Iran
- Badie, Kambiz, Tehran University, Tehran, Iran
- Babaei, Shahram, Islamic Azad University of Tabriz, Iran
- Bahrepour, Davoud, Islamic Azad University, Mashhad, Iran
- Bayanati, mahmonir, West Tehran Islamic Azad University, Tehran, Iran
- Bhalaji N, Sri Sivasubramaniya Nadar College, Chennai, Tamil Nadu, India.
- Bouyer, Asgar Ali, Shahid Madani University, Azarbaijan, Iran
- Darmani, Yousef, K. N. Toosi University of Technology, Tehran, Iran
- Davarpanah, seyed hashem, Birjand University of Technology, South Khorasan, Iran
- Deldari, Arash, University of Torbat Heydariyeh, Khorasan Razavi, Iran,
- Ebadati, Omid Mahdi, Kharazmi University,Tehran, Iran
- Ebrahimpour, Nader, Islamic Azad University of Mahabad, West Azerbaijan, Iran
- Farsi, Hassan, University of Birjand, South Khorasan, Iran
- Fatemi Khorasgani, Afsaneh, University of Isfahan, Isfahan, Iran
- Fekri Ershad, Shervan, Islamic Azad University, Najafabad, Iran
- Geetha R, Dayananda Sagar Academy of Technology, Karnataka, India
- Ghayoomi, Masood, Institute for Humanities and Cultural Studies, Tehran, Iran
- Ghaffari, Ali, Islamic Azad University, Tabriz Branch, Iran
- Ghasemzadeh, Mohammad, Yazd University, Yazd, Iran
- Haghighi, Hassan, Shahid Beheshti University, Tehran, Iran
- Kheirkhah, Fatemeh, ACECR, Tehran, Iran
- Mirroshandel, Seyed Abolghasem, University of Guilan, Rasht, Iran
- Mirzaei, Abbas, Islamic Azad University, Ardabil, Iran
- Minoofam, Seyed Amir Hadi, Qazvin Islamic Azad University, Qazvin, Iran
- Mir, Ali, Lorestan University, Iran
- Momtazi, Saeedeh, Amir kabir University, Tehran, Iran
- Mohammadi, Mohammad Reza, Iran University of Science and Technology, Tehran, Iran
- Mothku, Sai Krishna, National Institute of Technology Tiruchirappalli, Tamil Nadu, India

- Nangir, Mahdi, University of Tabriz, Tabriz, Iran
- Pahlevanzadeh, Bahareh, ISC, Shiraz, Iran
- Paindavoine, Michle, University of Burgundy, Dijon, France
- Reshadat, Vahideh, Malek-Ashtar University of Technology, Tehran, Iran
- Rezvanian, Alireza, University of Science and Culture, Tehran, Iran
- Sarrafzadeh, Abdolhossein, Unitec Institute of Technology, Auckland, New Zealand
- Sekar, M, Raja, VNR VJIET, Institute of Engineering and Technology, Hyderabad, India
- Shirmarz, Alireza, Ale Taha Institute of Higher Education, Tehran, Iran
- Shadroo, Shabnam, Islamic Azad University Mashad, Iran
- Sandooghdar, Afshin, ACECR, Tehran, Iran
- Sohrabi, Babak, Tehran University, Tehran, Iran
- Solouk, Vahid, Urmia University of technology, Iran
- Vu Duc, Nghia, Chung-Ang University, Seoul, Korea
- Yaghoobi, Kaebeh, Ale Taha Institute of Higher Education, Tehran, Iran
- Zare, Hadi, Tehran University, Tehran, Iran

Table of Contents

A Threshold-based Brain Tumour Segmentation from MR Images using Multi-Objective Particle Swarm Optimization

Katkoori Arun Kumar*
Research Scholar, KLEF, Hyderabad, India.
arun.katkoori@gmail.com
Ravi Boda
Associate Professor, KLEF, Hyderabad, India.
raviou2015@klh.edu.in

Received: 31/Jan/2021 Revised: 13/July/2021 Accepted: 25/Aug/2021

Abstract
The Pareto optimal solution is unique in single objective Particle Swarm Optimization (SO-PSO) problems as the emphasis is on the variable space of the decision. A multi-objective-based optimization technique called Multi-Objective Particle Swarm Optimization (MO-PSO) is introduced in this paper for image segmentation. The multi-objective Particle Swarm Optimization (MO-PSO) technique extends the principle of optimization by facilitating simultaneous optimization of single objectives. It is used in solving various image processing problems like image segmentation, image enhancement, etc. This technique is used to detect the tumour of the human brain on MR images. To get the threshold, the suggested algorithm uses two fitness(objective) functions- Image entropy and Image variance. These two objective functions are distinct from each other and are simultaneously optimized to create a sequence of pareto-optimal solutions. The global best (G_{best}) obtained from MO-PSO is treated as threshold. The MO-PSO technique tested on various MRI images provides its efficiency with experimental findings. In terms of "best, worst, mean, median, standard deviation" parameters, the MO-PSO technique is also contrasted with the existing Single-objective PSO (SO-PSO) technique. Experimental results show that Multi Objective-PSO is 28% advanced than SO-PSO for 'best' parameter with reference to image entropy function and 92% accuracy than Single Objective-PSO with reference to image variance function.

Keywords: Multi-objective optimization; PSO; Median filter; Threshold; Image segmentation.

1- Introduction

Image segmentation is important step in the image processing. It is the process of dividing the image into number of picture elements (generally, called "pixels"). The goal of image segmentation is to gradually change or modify the portrayal of an image, which is important and simple to analyze [1]. The image segmentation's output contains either a lot of forms removed from the picture or a lot of segments that will cover the total image. In a region of image, each pixel is comparative as for certain qualities like intensity, color, or texture [2]. Therefore, many tests and studies have been conducted to develop strategies and methods related to image segmentation. These strategies are classified into different classifications, including threshold-based and clustered-based segmentations. Image thresholding is an important tool in image segmentation, which separates the object distinct from background. This is done by based the on the different gray levels.

Basically, there are two types of thresholding techniques exits, one binary thresholding and another one is multi-thresholding [3][4]. In thresholding methods, it is very difficult to find optimum threshold value suitable for the separation of the target image. One of the best solutions for the above-mentioned problem is Particle Swarm Optimization (PSO) algorithm. PSO is recursive method that optimizes (minimize or maximizes) given problem. It is used in solving various image processing problems like image segmentation, image enhancement, etc. [5]

Brain is typical organ, which contains billion number of neurons. It is comprised of huge cells, and each cell performs some function. Magnetic Resonance Imaging (MRI) is a tool, which gives a high-quality image of various parts of the human body. But when dealing with one of the
sensitive organs of the human body i.e. Brain, a care should be taken. A Brain tumor is nothing, but abnormal cells grew inside the brain. Basically, there are two types of tumors exists. One is benign (Non-cancerous) and

* Corresponding Author

another one is malignant (cancerous). Also, there is another classification exists, tumor that starts inside the brain – primary brain tumor, whereas a tumor which starts in another part of the human body and spreading into brain- secondary brain tumor (also called metastatic) [6].

The main objective of this paper is to segment the lesion from MRI image using Multi-objective Particle Swarm Optimization (MOPSO) algorithm. First, the MRI image is pre-processed, then skull-stripping is performed. Skull-stripping is the process of removing extra-meningeal tissue from the head image to find the boundaries of head and brain. Then the tumor segmentation is done by MOPSO algorithm using two fitness functions called – image entropy and image variance.

The organization of the paper as follows- section 2 gives the related work done by different authors and section 3 about acquisition of the MR images. Section 4 explains the proposed MO-PSO technique and section 5 shows the results of various MRI images. And finally, the conclusion of the paper is given in section 6.

2- Literature Survey

In reference [7], a gray-scale image is segmented using 2-D entropy maximization was proposed. Reference [8] used a technique called variance minimization for image segmentation. And it also introduced local statistics in the formulation of energy function. Reference [9] proposed a method for multilevel threshold selection for image segmentation. In this paper, a multi threshold value was computed by hill-clustering technique, linear minimax approximation algorithm and then golden search minimization algorithms.

Reference [10] proposed a method to implement image segmentation using two entropy functions. Entropy region and entropy layout functions were considered. Reference [11] gives a method called skull stripping to remove the non-cerebral tissue in T1-weighted MRI image. Reference [12] proposed multi-objective particle swarm optimization (MOPSO) for image segmentation using multi-threshold technique. Reference [13] gives a clustered- based MOPSO algorithm to separate the tumour part from MRI images. Here two fitness functions- KFECSB and AWGLAC are used to produce the non-dominated solutions.

Reference [15] proposed a method in which the leader party, self-adaptive criteria, and disruptive operator for balancing convergence and diversity are included. They used an elitist learning technique as the perturbation operator with a Gaussian mutation. Reference [16] proposed a technique to classify the tumors using brain emotional learning. Reference [17] used thresholding, segmentation, and morphological operations to identify the accurate location of brain tumour.

3- MR Images Acquisition

Different kinds of slices of brain of patients are taken for testing the proposed method. Some of the patients are presently taking chemotherapy and some of them are cured. Different types of lesions exist, from vascular lesions to tumors. Using centricity DICOM viewer, the MRI images are taken. Axial T2 view images were considered for our experiments. These images provide the pathology of the disease. Experiments were performed on 1.5-T MRI imaging device. Thickness of slice is 5 mm and resolution are 256 x 256. The gap between two slices is 1.5mm.

4- Proposed Technique

The proposed method contains several steps to segment the tumor from MRI images.

1. **Preprocessing:** In this step, the input MRI image is resized into 256 x 256 size image. Then convert the RGB image into gray-scale image. To remove the noise present in the input image a digital filter is used, called Median filter. It is most popularly used filter for removing the noise in the image. In each image, median filter is applied on pixel by pixel, each pixel is replaced with the median(middle) of neighboring pixels.

2. Skull stripping: In this process, by using Otsu thresholding and morphological operations skull stripping was done. This process is used to remove extra-meningeal tissue from head image.

3. Threshold based Optimization**:** Next step in the proposed algorithm is finding the threshold value using one of the optimization techniques called Multi-objective Particle Swarm Optimization (MOPSO).

In single objective PSO (SO-PSO) algorithm, the model contains ‘n’ number of particles. They communicate with other particles by using gradients (search directions) directly or indirectly. By using single objective (fitness) function, the global threshold (global minima or global maxima) has been calculated.

The SO-PSO was originally introduced by Kennedy and Eberhart, inspired by the social actions of flocking birds and fish schooling, for problems of optimization. Any individual considered to be a potential solution to the problem of optimization of a given swarm will benefit from the previous experiences of all other individuals in the same population [13]. Every particle, through the search process in the solution space, the speed and location will be changed according to their own flying experiences and from the others in the swarm as well.

Consider Mp is the swarm size, any particle that contains the elements N, with a position vector Xi and velocity vector Vi, its own best location, p^{best} find so far, and communicates with neighboring particles via the best location so far, g^{best} has been found in the neighborhood. The optimality of the position is measured using one or more fitness functions described in the relationship to the issue of optimization. Each particle is moved according to the below equations at qth iteration in the search space.

$$Vq+1 = wVq + c1 * r1 * [\boldsymbol{p^{best}}q - Xq] + c2 * r2 * [\boldsymbol{g^{best}}q - Xq] \quad (1)$$

$$Xq+1 = Xq + Vq + 1 \quad (2)$$

Where w is inertia weight,
r1 and r2 are the random variables,
c1 and c2- acceleration coefficients.

But most of the real-world problems contain simultaneous optimization of one or more objective functions. In general, these objective functions belong to different groups and conflicting and competing. The MOO (multi-objective optimization) techniques have such conflicting and competing objective functions which gives optimal solutions set instead of single solution. With respect to all objectives, no solution is better than the other. These optimal solution set is known as Non-dominated solution set. Also called Pareto-solution set.

The MOO has 'n' number of objectives and 'm' number of equality and inequality constants. Mathematically,

Minimize/Maximize

$$F_i(x) = [\, f_1(x), f_2(x), \ldots\ldots, f_N(x)] \quad (3)$$

Subject to

$$r^j(x) = 0;\ j=1, 2, \ldots\ldots J$$
$$s^k(x) \leq 0;\ k=1, 2, \ldots\ldots K \quad (4)$$

where $F_i(x)$- fitness function,
x- decision vector,
J- Equality constant,
K- Inequality constant.

In MOO, multiple fitness functions need to be handled at the same time. Pareto improvement is the process of moving from one solution to other that can make at least one objective function to return a best value and with no other objective function becoming worst. When no further pareto improvements can be happened, those candidate solutions are said to be "Pareto optimal solutions". Still, all of the elements of an ideal pareto set may not be desirable, and although it provides certain space and time constraints, the pareto set may be infinite.[14]

Threshold based image segmentation is one of the popular methods used to extract tumour from MRI images. To segment tumour part from MRI images, two fitness functions are used here: Image entropy and Image variance.

Entropy is the actual ratio of randomness that can be utilized to describe the shape of a gray-scale image. It gives the average information of image can be resolved around from the histogram of the image. Mathematically, the entropy is defined as

$$E = -\sum_L q_l \, log_2 \, q_l \quad (5)$$

Where L represents number of gray levels,
q_l is the probability related to with gray level 'l'.

Another fitness function is variance of the image. It is used to discover how every pixel varies from its neighboring pixel (or centre pixel) and is utilized in characterize into various regions. Mean of the image is nothing but, the average of total pixels of given image. With M X N image size, mean is defined mathematically,

$$\bar{d} = \frac{1}{MN} \sum_m \sum_n d_{mn} \quad (6)$$

The simplified notation of equation (4) is

$$\bar{d} = \frac{1}{K} \sum_K d_k \quad (7)$$

The variance is

$$\sigma^2 = \frac{1}{K} \sum_K (d_k - \bar{d})^2 = \frac{1}{K} \sum_K {d_k}^2 - \bar{d}^2 \quad (8)$$

These are the two objective functions need to be optimized. Image entropy function must be maximized, and image variance function must be minimized.

After finding the threshold using above method, the object and backgrounds are separated as follows,

- Pixel intensity is above the obtained threshold => taken as object (1).
- Pixel intensity is below the obtained threshold => taken as background (0).

4-1- Algorithm:

1. Convert the input RGB image to gray-scale image.
2. Apply median filter to remove the noise.
3. Perform skull stripping process to remove skull from input MRI image.
4. Find the threshold value using MOPSO technique.
   i) Initialize the PSO parameters like c1, c2, inertia weight, initial position and velocity.
   ii) Initialize repository.
   iii) Evaluate the fitness for all particles using two functions- entropy and variance.
   iv) Find the particle $_{best}$ and global $_{best}$.
   v) If the total number of iterations is not satisfied DO FOR each particle:
      - Update velocities and positions
      - Perform mutation
      - Evaluate fitness for all particles
      - Find particle $_{best}$ and global $_{best.}$
5. Update repository.
6. Display the tumor part from input MRI image.

The flowchart of the proposed method is shown in figure 1.

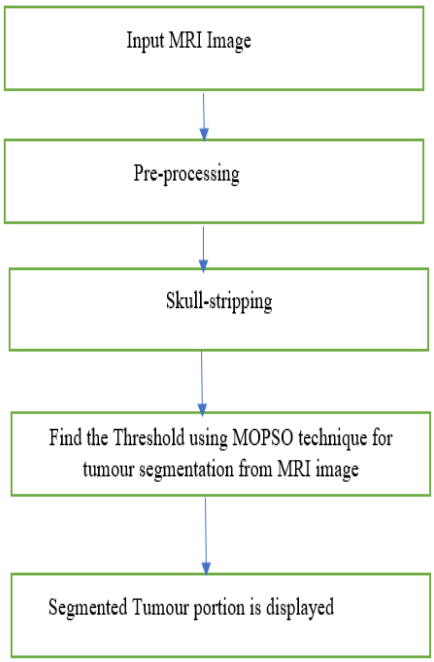


Fig. 1 Flowchart of proposed algorithm

5- Results

We have taken different slices of the human brain and single patient's MRI images are analyzed. The MRI images are taken from different datasets of the same patient at different times and in different view-focuses. These are taken from Kaggle dataset, using the link htttp://www. kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection/.

In this, 240 datasets are considered. Among them, 100 datasets are related to malignant and 140 are related to benign type of tumor. The extensive simulation has been done to validate the significance of the proposed algorithm. These MRI images are T1-weigthed type, which are widely considered in medical diagnosis of tumor.

The experiments are done in MATLAB on PC. The MOPSO parameters are set as follows:
Acceleration constants: c1=2, c2=2,
Minimum and maximum weights:
w^{min}=0.1, w^{max}=0.9
Number of iterations= 100,
Number of particles in population= 100.

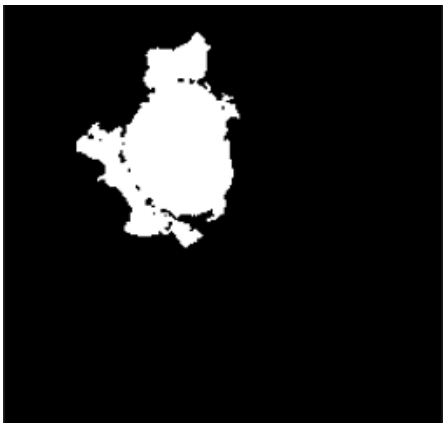


Fig. 2 a) Input MRI image from dataset 1

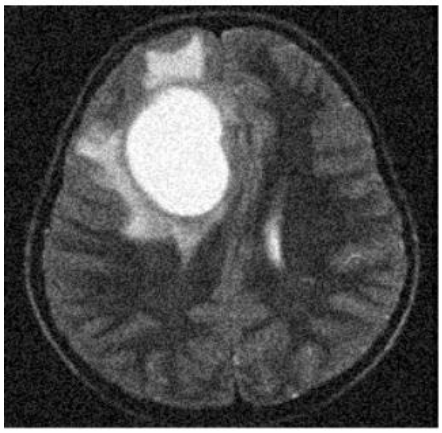


Fig. 2 b) Segmented output image

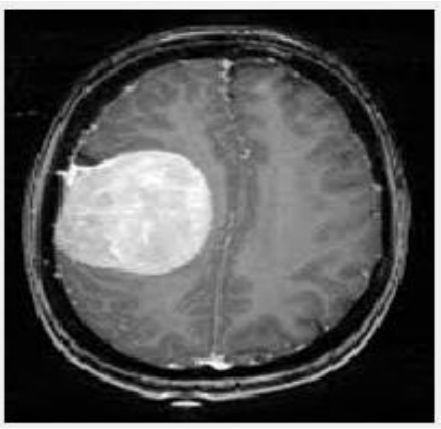


Fig. 3 a) Input MRI image from dataset 2

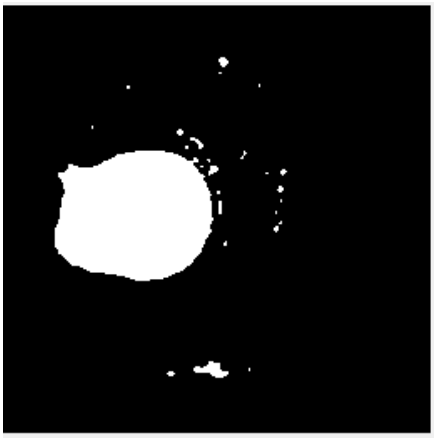


Fig. 3 b) Segmented output image

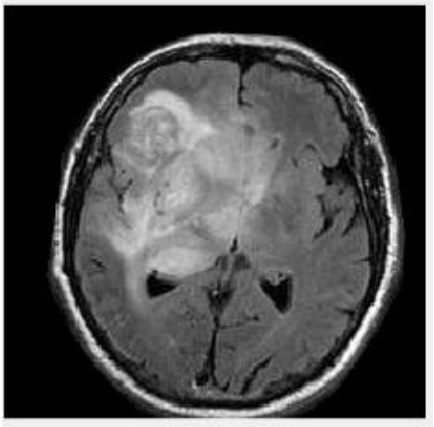


Fig. 4 a) Input MRI image from dataset 3

Fig. 4 b) Segmented output image

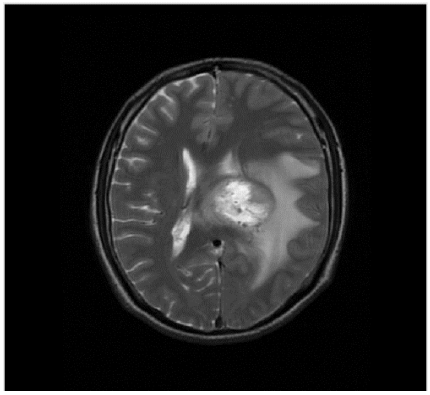


Fig. 5 a) Input MRI image from dataset 4

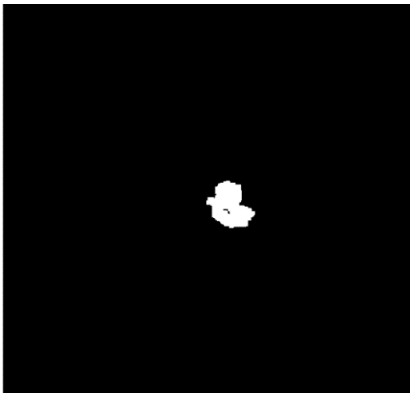


Fig. 5 b) Segmented output image

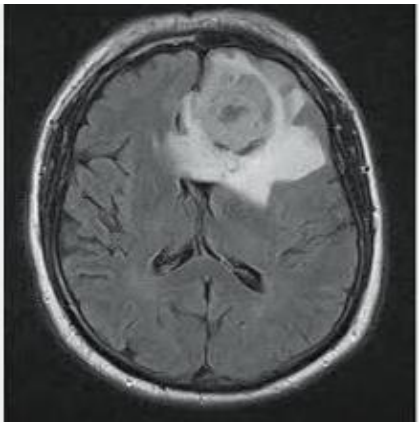


Fig. 6 a) Input MRI image from dataset 5

Fig. 6 b) Segmented output image

Figures 2 (a), 3(a), 4 (a), 5 (a), 6 (a) represents the input MRI images and figures 2 (b), 3 (b), 4 (b), 5 (b), 6 (b) represents their outputs which contain tumors and threshold values are 138,151,144,133,198,135, 128 respectively.

Table I gives the summary of experiments done on MOPSO algorithm. It gives the executed time, best fitness, and threshold of given input MR images.

Table I. Summary of Experimental Results

Input	Time	Best fitness	Threshold
Image 1	16.98 sec	1.009	138
Image 2	18.44 sec	1.004	151
Image 3	16.63 sec	1.008	144
Image 4	17.52 sec	1.006	133
Image 5	18.13 sec	1.004	198
Image 6	16.92 sec	1.007	128

Table II and table III gives the analysis of PSO and MOPSO algorithms with and without preprocessing step.

In the following table II, with reference to entropy and variance functions the performance measures like "best, worst, mean, median, standard deviation" are compared for both PSO and MOPSO techniques before preprocessing step.

Table II.PSO and MOPSO algorithms Analysis without Preprocessing

Parameter	**With reference to Entropy**		**With reference to Variance**	
	PSO	MOPSO	PSO	MOPSO
Best	0.823	0.62	7.82	0.92
Worst	0.00223	0.00219	5333	3394
Mean	0.423	0.391	2244	1549
Median	0.4357	0.412	2196	1449
Standard deviation	0.052	0.062	429	349

By considering entropy function, it is observed that MOPSO is 24% advanced than PSO for 'best' parameter, mean is 7.5% improved than PSO, and median is 5.4% is better than MOPSO.

Table III.PSO and MOPSO algorithms Analysis with Preprocessing

Parameter	**With reference to Entropy**		**With reference to Variance**	
	PSO	MOPSO	PSO	MOPSO
Best	0.8432	0.6051	7.923	0.592
Worst	0.00312	0.00219	5449	3459
Mean	0.4556	0.4123	2139	1297
Median	0.4559	0.4294	2096	1304
Standard deviation	0.0637	0.0874	495	307

By considering entropy function, it is observed that MOPSO is 28% advanced than PSO for 'best' parameter, mean is 9.5% improved than PSO, and median is 6% is better than MOPSO. The analysis is shown graphically in following figure 7.

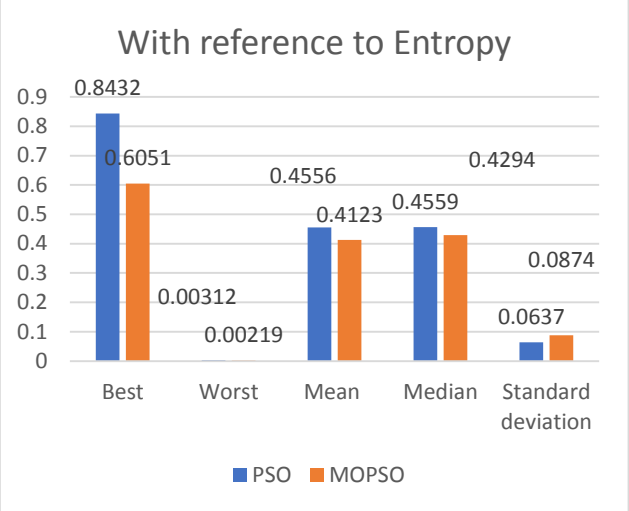


Fig. 7 Analysis of PSO and MOPSO algorithms with reference to Entropy function

These are the following observations from figure 8- for best measure the MOPSO is 92% advanced than PSO, the mean for MOPSO is 39% better than PSO algorithm. Median of MOPSO is 38% advance than PSO with reference to variance function.

From the above analysis, MO-PSO is performed well in lesion segmentation of brain MR images.

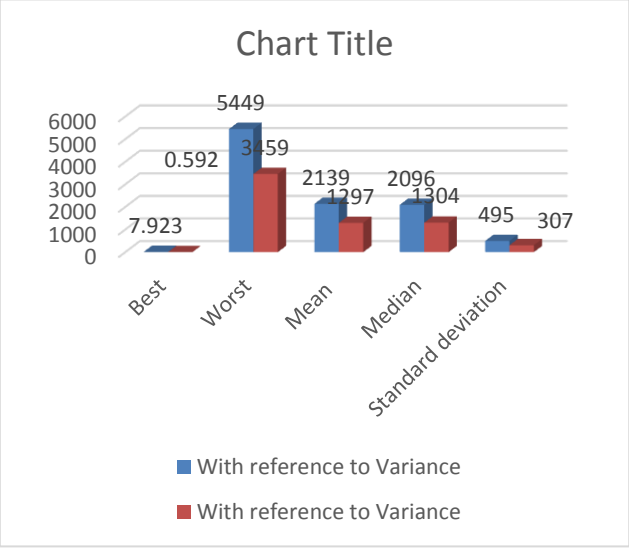


Fig. 8 Analysis of PSO and MOPSO algorithms with reference to Variance function

6- Conclusions

In this paper, MOPSO algorithm has been described to segment the lesion from MRI images. The goal of Multiobjective optimization (MOO) technique is to optimize at least two fitness (quantitative) functions simultaneously. Image entropy and variance functions are used in this work. Entropy is maximized and variance is minimized to get the optimal threshold value. This value is used to segment the lesion from MRI image. The exploratory outcomes give fulfilled and great image segmentations. Also, this technique is better when compared to single objective PSO technique.

Future Scope

In future, we will extract the texture features of segmented tumor and given to the machine learning algorithms to evaluate the performance of the classifier. Finally, the user will know the type of tumour whether it is Benign or Malignant.

References

[1] Linda G. Shapiro and George C. Stockman: "Computer Vision"., New Jersey, Prentice-Hall, 2001, pp 279-325.

[2] R.O Duda and P.E. Hart, " Pattern Classification and Scene Analysis ", John Wiley & Sons, New-York, 1973, vol. 3, pp 731-749.

[3] N. Papamarkos, B. Gatos, "A new approach for multilevel threshold Selection". Graphics Models Image Processing, 1994, vol 56, issue 5, pp 357-370.

[4] S. Boukharouba, J.M. Rebordao, P.L. Wendel, "An amplitude segmentation method based on the distribution function of an image". Computer Vision Graphics Image Processing, 1985, vol. 29 issue 1, pp 47-59.

[5] Y.-T. Kao, E. Zahara and I-W. Kao, "A Hybridized Approach to Data Clustering," Expert Systems with Applications, 2008, Vol. 34, No. 3, pp.1754-1762.

[6] "The essential guide to brain tumors", National brain tumor society (NBTS), 2007.

[7] Molka DHIEB, Sabeur Masmoudi, Mohamed Ben Messa OUD, and Faten Ben Arfia "2-D Entropy Image Segmentation on Thresholding Based on Particle Swarm Optimization (PSO)", in 1st International Conference on Advanced Technologies for Signal and Image processing (ATSIP), 2014, pp. 143-147, doi: 10.1109/ATSIP.2014.6834594.

[8] Bo Chen , Qing-hua Zou, Yan Li "A new image segmentation model with local statistical characters based on variance minimization", Applied Mathematical Modelling, 2014, vol. 39, issue 12, pp 3227-3235.

[9] Mala, C., Sridevi, M. "Multilevel threshold selection for image segmentation using soft computing techniques", Soft Computing, 2016, vol. 20, issue 5, pp 1793–1810.

[10] H. Zhang, J. Fritts, and S. Goldman, "An Entropy-based objective evaluation method for image segmentation", in Proc. SPIE-Storage and Retrieval Methods and Application for Watershed Transform Multimedia, 2004, pp 38-49.

[11] Horst K. Hahn et al," The Skull stripping Problem in MRI solved by a Single 3D", Proc. MICCAI,LNCS, Springer, Berlin, 2000, pp 134-143.

[12] Habba Maryam, Ameur Mustapha and Jabrane Younes , "A multilevel thresholding method for image segmentation based on multiobjective particle swarm optimization", International Conference on wireless technologies, Embedded and Intelligent systems (WITS), 2017, pp. 1-6, doi: 10.1109/WITS.2017.7934620.

[13] Pham TX, Siarry P, Oulhadj H. "A multi-objective optimization approach for brain MRI segmentation using fuzzy entropy clustering and region-based active contour methods." Magnetic Resonance Imaging, 2019, vol 61, pp 41-65.

[14] Coello Coello CA, González Brambila S, Figueroa Gamboa J, Castillo Tapia MG, Hernández Gómez R, "Evolutionary multiobjective optimization: open research areas and some challenges lying ahead", Complex & Intelligent Systems, 2020, vol.6, pp 221-236.

[15] Hu W, Yen GG, "Adaptive multiobjective particle swarm optimization based on parallel cell coordinate system", IEEE Transactions on Evolutionary Computation, 2015, vol 19, issue 1, pp 1–18.

[16] S. B. Emami, N. Nourafza, S. Fekri–Ershad, "A method for diagnosing of Alzheimer's disease using the brain emotional learning algorithm and wavelet feature", Journal of Intelligent Procedures in Electrical Technology, vol. 13, no. 52, pp. 1-15, 2021.

[17] G. Mahesh Kumar, K. Arun Kumar, P. Rajashekar Reddy, J. Tarun Kumar, "A Novel Approach of Tumor Detection in Brain using MRI Scan Images", Research J. Pharm. and Tech., 2020, vol 13, issue 12, pp 5914-5918.

Katkoori Arun Kumar received the B.Tech. degree in Department of Electronics & Communication Engineering from JNTU, Hyderabad, India in 2008, and M.Tech. degree in Embedded Systems from JNTU, Hyderabad, India in 2012. Currently he is Research Scholar in KLEF, Hyderabad branch, India. Her research interests include Medical image processing, Deep learning.

Ravi Boda is an Associate Professor of Electronics & Communication Engineering at KLEF, Hyderabad branch, Telangana, India. His research is focused on signal processing and Image processing, Medical Image processing using ML.

Predicting Student Performance for Early Intervention using Classification Algorithms in Machine Learning

Kalaivani K*
Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies, India
kalai.se@velsuniv.ac.in
Ulagapriya K
Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies, India
upriya.se@velsuniv.ac.in
Saritha A
Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies, India
saritha.se@velsuniv.ac.in
Ashutosh Kumar
Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies, India
ak214651@gmail.com

Received: 20/Aug/2020 Revised: 20/July/2021 Accepted: 25/Aug/2021

Abstract

Predicting Student's Performance System is to find students who may require early intervention before they fail to graduate. It is generally meant for the teaching faculty members to analyze Student's Performance and Results. It stores Student Details in a database and uses Machine Learning Model using i. Python Data Analysis tools like Pandas and ii. Data Visualization tools like Seaborn to analyze the overall Performance of the Class. The proposed system suggests student performance prediction through Machine Learning Algorithms and Data Mining Techniques. The Data Mining technique used here is classification, which classifies the students based on student's attributes. The Front end of the application is made using React JS Library with Data Visualization Charts and connected to a backend Database where all student's records are stored in MongoDB and the Machine Learning model is trained and deployed through Flask. In this process, the machine learning algorithm is trained using a dataset to create a model and predict the output on the basis of that model. Three different types of data used in Machine Learning are continuous, categorical and binary. In this study, a brief description and comparative analysis of various classification techniques is done using student performance dataset. The six different machine learning Classification algorithms, which have been compared, are Logistic Regression, Decision Tree, K-Nearest Neighbor, Naïve Bayes, Support Vector Machine and Random Forest. The results of Naïve Bayes classifier are comparatively higher than other techniques in terms of metrics such as precision, recall and F1 score. The values of precision, recall and F1 score are 0.93, 0.92 and 0.92 respectively.

Keywords: Machine Learning; Classification; Supervised Machine Learning; Data Analysis; Naïve Bayes.

1- Introduction

In India there are a large number of Universities which use the traditional methods to analyze the student's performance and find it difficult to manage hundreds of students and make use of their skills. There is a huge amount of data generated with student details and performance which can be used to improve the Education System like Identifying Problems, predict the performance, find out who needs an intervention etc.

Machine Learning is a subcategory of Artificial Intelligence, where a machine learns on its own from a given dataset without being programmed to make predictions. Machine Learning has become one of the most popular field of study today with a wide variety of applications in different domains. Machine Learning Algorithms can be used in recommendation system for customers, predicting stock prices or housing prices and clustering of customers etc., With the amount of data available today and the increasing computation speed of computers, the machine learning algorithms are able to tackle a variety of problems of high dimensional space. One of the best features of a machine learning algorithm is the ability to continually learn on its own and gradually increase its accuracy with time. If the prediction is not as expected, then the algorithm is re-trained multiple number of times until the desired output is found.

The Machine Learning process first requires a selection of algorithm and then a training data is used as input for

* Corresponding Author

the selected algorithm. If the output is unknown, it is called unsupervised learning and when the output is given it is called supervised learning. In supervised learning the algorithm takes both input and output for training and a model is generated which maps the inputs with the desired output. The model is then tested with a new set of input data where the predicted results are checked with the desired output.

Supervised learning algorithm can be further divided into two types: Regression and Classification. Both the algorithms can predict the output value from the given labeled input dataset with the only difference being that in regression the output variable is numerical and categorical in case of classification. Example for regression problem could be a situation where the output value is continuous such as salary, height or weight and age etc. In classification problems the output variable is a category such as yes or no, pass or fail etc. Some applications of classification algorithm are: Filtering of spam or not spam emails, student is going to pass the exam or not etc. [1].

H. Al-Shehri et al., used the popular dataset which consists of 395 data samples collected from the University of Minho in Portugal. They proposed that support vector machine-based prediction model accurately predicts the students' performance than K Nearest Neighbor [2]. S. Hossain et al., recommends that the Belief Rule based knowledge and Evidential Reasoning approach predicts thestudent performance with considering the personal and institutional parameters, they also stated that continuous performance analysis is necessary to find the skill and efficiency of students at various levels [3].E. S. Bhutto et al.,proposed that sequential minimal optimization algorithm greatly predicts the behavior of students by achieving improved accuracy than logistic regression. Their proposed system also suggests the measures to reduce the student's dropout ratio [4].

Ahammad, Khalil et al., stated that students who are at higher risk can be recognized by the use of machine learning models. They used the results of SSC exam and performed a comparative study by employing different machine learning techniques. They suggested that Multi-Layer Perceptron achieved higher accuracy for their chosen dataset and also showed that all other techniques were also yielded satisfactory accuracy in predicting the student's performance [5]. M. B. Shah et al., considered academic and other input variables like interests, attributes and opinions in predicting the performance. They explored various machine learning, deep learning modelsand basic exploratory data analysisto understand the correlations of student's performance using psychographic attributes [6].

Nurafifah Mohammad Suhaimiet al., proposed a model for the academic assessment to predict the student graduation time using Neural network and Support Vector Machine [7]. Fan Yang, Frederick W.B. Li, suggested that the Back Propagation based Neural Network outperforms and they collected data from 60 schools to design the prediction model [8].Reynold A. Rustia et al., focused on a classification model using data mining techniques for predicting the probability of a student to pass the Licensure Examination for Teachers (LET) [9].

2- Methodology

Major Predicting students' performance from the current academic records is critical for adopting necessary pedagogical measures to make the students graduate on time and for satisfactory results. There are several challenges in predicting and to intervene for better performance. Some of the challenges are

- Students are from different backgrounds
- Lack of student evolving progress in making prediction.

The existing system is outdated and follows a traditional way of monitoring and generating student's reports

- Marks-based evaluation
- Exams have become "mugging-up" and memory tests
- Total reliance on pen and paper test
- Monitoring Hundreds of Students is difficult
- A weaker bond between assessment and learning outcomes

The drawbacks of the existing system are

- Focus was on rote learning and exams
- Missing of quick and timely Feedback on assessed work
- Not assessing Progress of Students or Analyzing Results or Scope of Improvement
- Project-based experimental learning is essential

To overcome these drawbacks, we make use of Classification Algorithms in Machine Learning to generate a model, which predicts the student's performance. There are a number of classification models and this paper describes and compares six different classification techniques with their advantages and disadvantages to analyse the student performance.

2-1- Logistic Regression

Logistic Regression can be used only when the output is categorical. It is similar to linear regression with a threshold. Based on the threshold value the classification is performed. Since Linear Regression cannot be used to solve classification problems, therefore in Logistic Regression, an activation function like sigmoid function to the Linear Regression model, which makes the value range from 0 to 1, can be added [10].

Advantage of Logistic Regression is that a threshold can be set once a value between 0 to 1 is arrived. For example, if there is a dataset to predict cancer is malignant or benign

based on its size and the predicted continuous value is 0.4 and the threshold value is 0.5, the data point will be classified as not malignant, which can lead to serious consequence. With Logistic Regression, a threshold value can be set as 0.6 or 0.7 accordingly. Logistic regression usually states where the boundary between the classes exists [1].

2-2- Naïve Bayes Classifier

Naïve Bayes is a classification technique based on Bayes Theorem. The classifier makes two assumptions: firstly, the attributes or features present in the dataset are independent and second is that each feature is given the same weight to predict the outcome. Naive Bayes model is easy to implement and works efficiently with both smaller and larger datasets. The dataset is divided into two parts, firstly input data which consists of the dependent features whose conditional probability is to be calculated based on the output class and the output data which contains the value of the class variable [11].

2-3- K-Nearest Neighbor

K-Nearest Neighbor is used to solve both classification and regression problems. KNN being a non-parametric technique, widely used in statistical estimation and pattern recognition. It is a lazy learning model, which is one of the easiest Machine Learning Classification techniques to implement. It first selects a group of labeled points and uses them to label other points. To predict a new point, it makes use of all the data points that are currently closest to that new point which are called the nearest neighbors and has those neighbors vote, here the "k" is the number of neighbors it checks. So whichever label has the most number of votes is the label for the new point [11].

K-Nearest Neighbor is also called a case-based learning method, where all the training data is used for classification. It is used in number of applications such as dynamic web mining, recommender systems etc. Further its efficiency can be improved with the help of some representatives to characterize the whole data i.e. implementing an inductive learning model from the training dataset and using that model for classification [12].

2-4- Support Vector Machine

A Support Vector Machine (SVM) is a classifier, which makes use of a separating hyperplane to classify points in higher dimensional space. In other words, when a labeled training dataset is given, the algorithm categorizes new points using an optimal hyperplane. In two-dimensional space the hyperplane can be visualized as a simple line, which divides the plane into two parts. Support Vector Machine is very effective and can handle dataset with high dimensional spaces efficiently. It is used to classify non-linearly separable classes and this algorithm allows you to avoid overfitting due to its regularization parameter [13].

2-5- Decision Tree

A decision tree is a scenario-based tree structure, where the internal nodes represents a test on the attribute and each branch represents an outcome of the test, and each leaf node has a class label.

A Decision tree is easy to implement and the domain knowledge is not required for constructing decision tree classifiers. This classification algorithm can easily handle multidimensional data. The learning and classification phases of decision tree algorithms are simple and fast with promising accuracy. An attribute selection method is used to split the data based on the attribute selected; it divides the rows into distinct classes.

An attribute selection measure is a heuristic in choosing the splitting criterion to separate a given data partition. It suggests a ranking for each attribute describing the given training tuples. The splitting attribute for the given list of tuples attribute is selected based on the highest score. If the splitting attribute is continuous-valued or if it is restricted to binary trees, then a split point or a splitting subset must also be found respectively as a part of the splitting criterion. In the process of forming decision tress, most of the branches might expose noise or outliers in the training data. Tree pruning is done to find and remove those branches, in order to improve classification accuracy on unseen data.

One example of attribute selection measure is Information gain. Information Gain (IG) is normally used to measure the amount of information a feature provides about the class and can be calculated using the Eq. (1). While constructing a Decision Tree, Information gain is mainly used to find the key attribute, which splits the dataset into classes. Decision Tree algorithms constantly try to maximize the IG and an attribute with the highest IG will split first [2].

Information gain = entropy (parent) – [weights average] * entropy (children) (1)

2-6- Random Forest

Random forest algorithm is an ensemble learning method for classification. This algorithm can be used to implement both classification and regression tasks. It uses various decision trees with different attributes as primary node. It is generally said to be a collection of decision trees and makes a mean prediction to the output. Random forests are better than decision trees as they avoid of over fitting to the training data and are very fast to train. Random Forest

algorithm involves two steps: first step is the creation of random forest and the second step is to make a prediction from the classifier [10][13].

3- Evaluation measures for Machine Learning Algorithms

There are number of measures or metrics [14][15] to evaluate the performance of the classification algorithms and they are described as follows:

Mean, Median and Mode

- Mean can be obtained by calculating the ratio between the sum of all values and the total number of observations.
- Mode is obtained by evaluating the most occurring value in the sample.
- Median is obtained by first sorting the numbers in increasing order and finding the middle value.

Variance

Variance is used to show the dispersion of the values around the mean. To find the variance, the first step is to calculate the mean and then sum the square difference between each value and finally divide the total number of observations.

Standard Deviation

Standard deviation gives us the variation of the values. It is obtained by taking the Square root of the variance.

Correlation

Correlation is used to find how the attributes are related to each other and how they contribute to the outcome. It is a measure of the relationship between attributes. Correlation value ranges between -1 and 1 where -1 specifies that the variables are negatively correlated and +1 confirms that the variables are positively correlated. 0 specifies that there is no correlation midst the target variables.

R-Squared

R-Squared is used to measure explained variation over total variation. Formula to calculate R squared is :
R squared = 1 - (Sum of Squared Residuals / Total Sum of Squares)

Confusion Matrix

Confusion matrix is in the form of a table that contains the results of classification algorithm. It is formed after prediction is performed on the test set and the actual true values are known.

Precision and Recall

Precision is used to identify the relevant instances from the model and can be calculated as Eq. (2). It is given by:
Precision = true positives / (true positives + false positives) (2)

Recall is used to identify all the relevant cases within a dataset and can be calculated as Eq. (3). It is given by:
Recall = true positives / (true positives + false negatives) (3)

F1 Score

F1 uses both recall and precision in other words it provides a single metric that combines recall and precision using the harmonic mean. The results range between 1 and 0 where the values that are closer to 1 are considered the best whereas those, which lie towards 0, are considered as the worst.

Mean Absolute Error

Mean Absolute Error is used to find the average of the difference between the actual values and the predicted values. It is used to measure the distance between the predictions and the actual output [15].

Mean Squared Error

Mean Squared Error is used to find the average of the square of the dissimilarity between the original and the predicted values. It is similar to Mean Absolute Error but here the squared values are considered due to which the computation of gradient is easier [15].

Table 1: Advantages & Disadvantages of various Classification Algorithms

Machine Learning Algorithm	***Advantages***	***Disadvantages***	***Applications***
Logistic Regression	• Interpretability • Little tuning	• Correlated features may affect performance • Extensive feature engineering required	• Risk Assessment • Fraud Prevention
Naïve Bayes Classifier	• It performs well with categorical variables • Good when dataset contains several features	• Correlated features affect performance	• Sentiment Analysis • Document categorization
K-Nearest Neighbor	• Good performance	• Slow when predicting • Susceptible to high dimension (lots of features)	• Gene expression • Protein-protein interaction
Support Vector Machine (SVM)	• Good for datasets with more variables than observations • Good performance	• Long training time required • Tuning is required to determine which kernel is optimal for non-linear SVMs	• Stock market forecasting • Value at risk determination
Decision Trees & Random Forests	• Render feature importance • Less data pre-processing required • Does not overfit (in theory) • Good performance and accuracy • Robust to noise	• It does not predict beyond the range of the response values in the training data • Biased towards categorical variables with several categories	• Credit Risk Assessment • Assess probability of developing a chronic disease (healthcare)

Table 2: Comparison of various classification algorithms on training speed, feature scaling, missing data and outliers

Machine Learning Algorithm	***Average Predictive Accuracy***	***Training Speed***	***Feature Scaling***	***Missing Data***	***Outliers***
Logistic Regression	Medium	High	YES	Sensitive	Potentially sensitive
Naïve Bayes	Medium	High	NO	Can handle missing data	Robust to outliers
KNN	Medium	Low	YES	Sensitive	Sensitive
SVM	High	High	YES	Sensitive	Robust to outliers
Decision Tree	Medium	Medium	NO	Can handle missing data	Sensitive
Random Forest	High	Low	NO	Can handle missing data	Robust to outliers

	sex	failures	paid	activities	absences	G1	G2	G3
0	F	0	no	no	4	5	5	6
1	F	3	yes	no	10	7	8	10
2	F	0	yes	yes	2	15	14	15
3	F	0	yes	no	4	6	10	10
4	M	0	yes	yes	10	15	15	15
5	M	0	no	no	0	12	12	11
6	F	0	no	no	6	6	5	6
7	M	0	yes	no	0	16	18	19
8	M	0	yes	yes	0	14	15	15
9	F	0	yes	no	0	10	8	9
10	F	0	no	yes	4	10	12	12
11	M	0	yes	yes	2	14	14	14
12	M	0	yes	no	2	10	10	11
13	M	0	no	no	0	14	16	16
14	F	0	no	no	4	14	14	14
15	F	0	yes	yes	6	13	14	14

Fig. 1: Student Performance Dataset

4- Results and Discussion

The advantages, disadvantages and appropriate applications of different machine learning algorithm is shown in table 1. An analysis in terms of training speed, feature scaling, missing data and outliers is shown in table 2. Six different classification algorithms were employed on the Student performance dataset, taken fromUCI Machine Learning [16] – [21]

to analyze the performance. The dataset includes the student achievement in secondary education of two Portuguese schools. The data attributes include student grades, demographic, social and school related features) and it was collected by using school reports and questionnaires. The output is binary i.e. Pass (1) or Fail (0). The dataset contains the following information:

- Total number of students: 395
- Number of students who passed: 265
- Number of students who failed: 130

A comparison is done on student performance dataset with following 8 feature columns:

- sex: Male or Female (binary)

- failures: number of failures in previous classes or semesters (numeric)
- paid: enrolled in extra paid online courses (binary: yes, or no)
- activities: participating in extra-curricular activities (binary: yes, or no)
- absences: number of days absent in class (numeric: from 0 to 93)
- G1 G2 G3: Internal marks of students (numeric)
- passed: final result of exam (binary: yes, or no)

precision, recall and f1-score were chosen to test the accuracy of the model as these measures works best on classification problems. These metrics clearly shows how effectively the model has predicted or classified the test data [22 – 26].

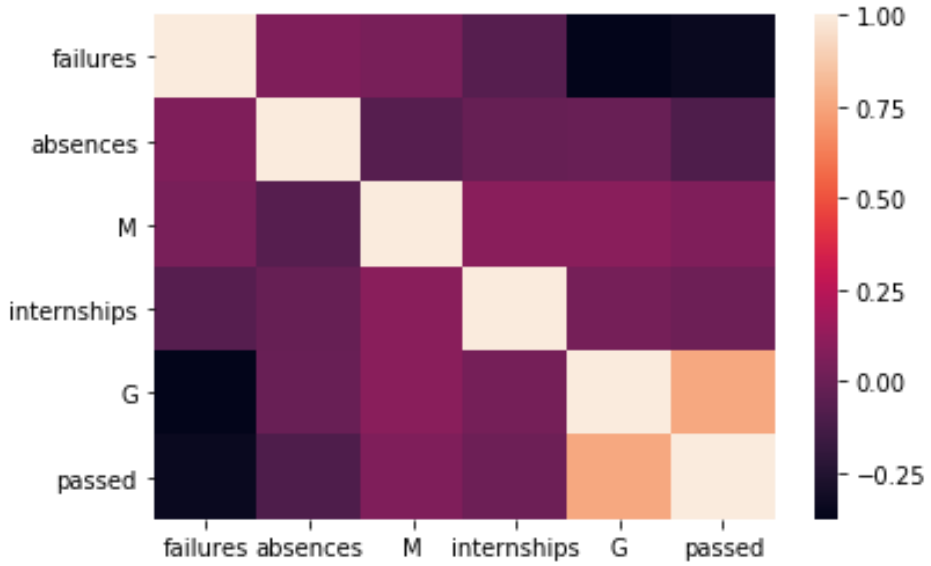


Fig. 2: Heat map of the Correlation Function

Table 3: Test Results of the Machine Learning Model

Machine Learning Algorithm	***Precision***	***Recall***	***F1-Score***
Logistic Regression	0.78	0.78	0.77
Naïve Bayes	**0.93**	**0.92**	**0.92**
KNN	0.81	0.73	0.65
SVM	0.91	0.91	0.91
Decision Tree	0.88	0.88	0.88
Random Forest	0.92	**0.92**	**0.92**

```
print(confusion_matrix(y_test,y_pred))
print('\n')
print(classification_report(y_test,y_pred))
```

```
[[19  6]
 [ 0 54]]


             precision    recall  f1-score   support

          0       1.00      0.76      0.86        25
          1       0.90      1.00      0.95        54

avg / total       0.93      0.92      0.92        79
```

Fig. 3: Test Results of Naïve Bayes Classifier

From the performance of various classification algorithms and the results achieved on Student performance dataset, it is evident that the performance of random forest and Naïve Bayes classifier is same in terms of recall and f1 score. The precision value of Naïve Bayes is slightly better than random forest. The Naïve Bayes and Random Forests are good at handling missing data and outliers.

5- Conclusions

In this paper, an analysis on student performance using different Classification Algorithms is done, to show whether the student will pass or fail based on the chosen attributes. The input data was trained on six classification algorithms and their test results were compared using metrics such as accuracy, precision, recall and F1 score. These measures were used to evaluate the accuracy of the classification models and as a result Naïve Bayes algorithm was found to be more effective. The study clearly shows that classification algorithm behaves differently with different attributes. Decision tree Algorithm shows high precision when the attributes are of binary type and not continuous and are prone to overfitting of data. Tree pruning process needs to be performed to avoid overfitting of data. Support Vector Machine, Naïve Bayes and Random Forests algorithms perform with high accuracy and precision regardless of the number of attributes.

The analysis was done on the merits and demerits of different algorithms at varied situations to understand their efficiency. After a better understanding of these algorithms for future prospects, how two algorithms can be combined together with their strengths and weaknesses should be investigated.

References

[1] F.Y. Osisanwo, J.E.T.Akinsola, O. Awodele, J.O. Hinmikaiye, O. Olakanmi, and J. Akinjobi, "Supervised Machine Learning Algorithms: Classification and Comparison", International Journal of Computer Trends and Technology, Vol. 48, No. 3, 2017, pp. 128-138.

[2] H. Al-Shehri et al., "Student performance prediction using Support Vector Machine and K-Nearest Neighbor," in IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), 2017, pp. 1-4.

[3] S. Hossain, D. Sarma, F. Tuj-Johora, J. Bushra, S. Sen and M. Taher, "A Belief Rule Based Expert System to Predict Student Performance under Uncertainty", in 22nd International Conference on Computer and Information Technology (ICCIT), 2019, pp. 1-6.

[4] E. S. Bhutto, I. F. Siddiqui, Q. A. Arain and M. Anwar, "Predicting Students' Academic Performance Through Supervised Machine Learning",in International Conference on Information Science and Communication Technology (ICISCT), 2020, pp. 1-6.

[5] Khalil Ahammad, Partha Chakraborty, Evana Akter, Umme Fomey, and Saifur Rahman, "A Comparative Study of Different Machine Learning Techniques to Predict the Result of an Individual Student Using Previous Performances", International Journal of Computer Science and Information Security, Vol. 18, No. 1, 2021, pp. 5-10.

[6] M. B. Shah, M. Kaistha and Y. Gupta, "Student Performance Assessment and Prediction System using Machine Learning", in 4th International Conference on Information Systems and Computer Networks (ISCON), 2019, pp. 386-390.

[7] Nurafifah Mohammad Suhaimi, ShahAlam, Selangor, Shuzlina Abdul-Rahman, Sofianita Mutalib, Nurzeatul Hamimah Abdul Hamid, Ariff Md Ab Malik, "Review on Predicting Students', Graduation Time Using Machine Learning Algorithms", International Journal of Modern Education and Computer Science, Vol. 7, 2019, pp. 1-13.

[8] Fan Yang, Frederick W.B. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining", Computers & Education, Vol. 123, 2018, pp. 97-108,

[9] Reynold A. Rustia, Ma. Melanie A. Cruz, Michael Angelo P. Burac, Thelma D. Palaoag, "Predicting Student's Board Examination Performance using Classification Algorithms", in 7th International Conference on Software and Computer Applications, 2018, pp. 233–237.

[10] Jiawei Han, Micheline Kamber, Jian Pei, "Data Mining Concepts and Techniques", USA: Morgan Kaufmann Publishers Elsevier, 2012.

[11] Saishruthi Swaminathan, "Logistic Regression – Detailed Overview", Canada: Towards Data Science – Medium Publication, 2018.

[12] Mandy Sidana, "Intro to types of Classification Algorithms in Machine Learning", USA: Medium Publication, 2017.

[13] Gongde Guo, Hui Wang, David Bell, Yaxin Bi, Kieran Greer, "KNN Model-Based Approach in Classification", Lecture Notes in Computer Science, Vol. 2888, 2003, pp. 986-996.

[14] Simon Tong, Daphne Koller, "Support Vector Machine Active Learning with Applications to Text Classification", Journal of Machine Learning Research, Vol. 2, No.1, 2001, pp. 45-66.

[15] H Zhang, J Zimmerman, D Nettleton, DJ Nordman, "Random forest prediction intervals", The American Statistician, Vol.74, Issue 4, 2019, pp. 392-406.

[16] Farhad Malik, "Must Know Mathematical Measures for Every Data Scientist", USA: Medium Publication, 2018.

[17] R.J. Howarth, "r2 (R Squared)", Dictionary of Mathematical Geosciences: With Historical Notes, Switzerland: Springer, 2017, pp. 503-527

[18] Aditya Mishra, "Metrics to Evaluate your Machine Learning Algorithm", Canada: Towards Data Science, Medium Publication, 2018.

[19] Haiyi Zhang, Di Li, "Naïve Bayes Text Classifier", in IEEE International Conference on Granular Computing, 2007, pp. 708.

[20] Nawzat Sadiq Ahmed, Mohammed Hikmat Sadiq, "Clarify of the Random Forest Algorithm in an Educational Field", in IEEE International Conference on Advanced Science and Engineering, 2018, pp.179-184.

[21] Dataset,https://archive.ics.uci.edu/ml/datasets/Student+Performance

[22] Sotiris Kotsiantis, "Supervised Machine Learning: a review of classification techniques", Informatica, Vol. 31, No. 3, 2007, pp. 249-268.

[23] S.B.Kotsiantis, I.D. Zaharakis, P.E.Pintelas, "Machine Learning: A Review of Classification Techniques and Combining Techniques", Artificial Intelligence Review, Vol. 26, No. 3, 2006, pp.159-190.

[24] R. Choudhary and H. K. Gianey, "Comprehensive Review on Supervised Machine Learning Algorithms," in IEEE International Conference on Machine Learning and Data Science (MLDS), 2017, pp. 37-43.

[25] C. A. Ul Hassan, M. S. Khan and M. A. Shah, "Comparison of Machine Learning Algorithms in Data classification," in 24th International Conference on Automation and Computing (ICAC), 2018, pp. 1-6.

[26] O. Obulesu, M. Mahendra and M. Thrilok Reddy, "Machine Learning Techniques and Tools: A Survey," in International Conference on Inventive Research in Computing Applications (ICIRCA), 2018, pp. 605-611.

Kalaivani K is an Assistant Professor in the Department of Computer Science and Engineering, Vels Institute of Science, Technology and Advanced Studies, Chennai, India. She received her B.E degree in Computer Science and Engineering from VIT University, India in 2001, and her M.E degree in Computer Science and Engineering from Vels Institute of Science, Technology and Advanced Studies, Chennai, India in 2012. She has teaching experience of more than 10 years and her research interests include Image Processing, Machine Learning and Big Data Analytics.

Ulagapriya K is an Assistant Professor in Department of Computer Science and Engineering in the School of Engineering at Vels Institute of Science, Technology and Advanced Studies, Chennai, India. She is having 12 years of IT experience and 8 years of teaching experience. Her research interests include Machine learning and Data Analytics.

Saritha A is an Assistant-Professor in Department of Computer Science and Engineering in the School of Engineering at Vels Institute of Science, Technology and Advanced Studies, Chennai, India. She is in teaching for more than 8 years and currently doing research in the area of Mobile Adhoc Networks.

Ashutosh Kumar is a UG scholar of Computer Science and Engineering in the School of Engineering at Vels Institute of Science, Technology and Advanced Studies, Chennai, India. His interests include Web Application and Machine Learning.

Evaluating the Cultural Anthropology of Artefacts of Computer Mediated Communication: A Case of Law Enforcement Agencies

ChukwuNonso H. Nwokoye*
Nigerian Correctional Service
chinonsonwokoye@gmail.com
Njideka Mbeledogu
Faculty of Physical Sciences, Nnamdi Azikiwe University Awka, Nigeria
njidembeledogu@yahoo.com
Chikwe Umeugoji
Faculty of Physical Sciences, Nnamdi Azikiwe University Awka, Nigeria
Chikw@live.ca

Received: 10/Mar/2020 Revised: 24/Jun/2021 Accepted: 06/Jul/2021

Abstract
The renowned orientations of cultural models proposed by Hall and Hofstede has been the subject of criticisms. This is due to the weak, inflexible and old-fashioned nature of some designs resulting from them. In addition, is the ever-changing, formless and undefined nature of culture and globalization. Consequently, these vituperations have resulted in better clarifications when assessing the cultural anthropology of websites. Based on these later clarifications and other additions, we seek to evaluate the cultural heuristics of websites owned by agencies of the Nigerian government. Note that this is verily necessary because older models did not include Africa in their analyses. Specifically, we employed the online survey method by distributing questionnaires to different groups of experts drawn from the various regions of Nigeria. The experts employed methods such as manual inspection and use of automated tools to reach conclusions. Afterwards, the results were assembled and using the choice of a simple majority, we decided whether a design parameter is either high or low context. Findings show that websites developers tend to favor low context styles when choosing design parameters. The paper attempts to situate Africa in Hall's continuum; therein, Nigeria (Africa) may fall within French Canadian and Scandinavian and/or within Latin and Scandinavian for the left hand and right hand side diagram respectively. In future, we would study the cultural anthropology of African websites employing the design parameters proposed by Alexander, *et al.*

Keywords: Website; Culture Contexts; E-governance; Cultural Anthropology; Human Computer Interaction.

1- Introduction

The burgeoning use of the internet as well as the flourishing number of application developers has productively influenced business organizations in exhibiting and fulfilling their goals on the World Wide Web using websites. Most marketing industries quickly adopted this approach as it offered them an ideal platform to present their products and services to local and foreign customers. Not long ago, with the exploration and incorporation of multimedia forms such as flash, sound and video, web site design have moved beyond simple static text-based formats to dynamic and interactive approaches. By that means, encouraging a towering measure of standard in terms of superior interaction using the web. Nevertheless, to construct a cross-culturally adequate website from a marketing viewpoint, web developers mostly addresses the concerns that border a "culture-specific color connotations, preferences in layout, animation, sounds, and other effects that are characteristic of today's generation of Web sites" [1]. This perspective is also sustained to a huge length for websites of e-government which fundamentally endeavor to supply a citizen-centric device for effectual delivery of service by government agencies. Investigating the user conduct and temperament of design constituents is to extract most of the hints a developer should utilize in order to make certain that the "values and behavior indoctrinated through cultural influences are reflected in design practices" [1].

Researchers at some quarters has criticized the renowned orientations of cultural models proposed by Hall [2] and Hofstede [3] due to the inflexible and old-fashioned nature of some designs originating from them. This is hugely due to the fact that culture and globalization are ever-changing, formless [4] and undefined by geographical boundaries existent in-between nations. Notwithstanding the vituperations, empirical studies on e-government websites are still based on the conceptions of Hall and Hofstede. Several variables were proposed by Hall to support the placement of cultures in a continuum that spans bi-

* Corresponding Author

directionally from high-content/low-content to the low-context/high-content. These variables include Nonverbal communication, Directness vs Indirectness, Time perception (monochronic vs polychronic) and Message speed. In summary, "the classifications of Hall [2] for either high or low context culture is mostly based on the degree to which meaning comes from the words that are being exchanged in a communication or from the setting (context) in which the communication is taking place".

Hofstede also proposed cultural variables which include the dimensions of Collectivism versus Individualism. Culture differences have impacted the evolution and the utilization of information and communication technologies. Myers and Tan [5] maintained that, "…by far the majority of studies concerned with various cultural aspects of the development, implementation, use and management of IT have relied on Hofstede's model of national culture". Even though cultural anthropology has been applied to management and advertising, this study presents its application to one of the forms of computer mediated communication. Note that cultural anthropology is a field that investigates cultures of humans, their belief systems, customs, principles, objectives, technologies, economies and other dimensions of social and cognitive organization. Though culture context analyses are essential in defining some aspects of a society, a nation, a culture or even a continent, it is not easy to say if a country is either high or low-context in nature. Depending on the values and attributes of a particular society, people can operate at both ends, in the middle or at one end of the continuum. However, it is clear from Fig. 1 below, that Africa (Nigeria) is absent in the cultural models of Hall [2].

Fig. 1 High/Low Context Culture [2]

Consequently, in the light of the above, this study is aimed at evaluating the culture context dimensions of websites of law enforcement e-government agencies (LEEGA) in Nigeria using online survey method. This is to situate Africa (and Nigeria) in Hall's bidirectional continuum for analyzing high or low context cultures. The government agencies include Nigeria Air Force (NAF), Nigeria Navy (NN), Nigeria Army (NA), Nigeria Police Force (NPF), Nigeria Prisons Service (NPS), Nigeria Immigration Service (NIS), Nigeria Security and Civil Defense Corps (NSCDC), Nigeria Customs Service (NCS), Federal Road Safety Corps (FRSC), Federal Fire Service (FFS), Defense Intelligence Agency (DIA), National Drug Law Enforcement Agency (NDLEA), Independent Corrupt Practices and Other Related Offences Commission (ICPC), Economic and Financial Crimes Commission (EFCC). Thirteen (13) design parameters to be used for the analyses include Animation (Multiple use of images/banners) (AIB), Promotion of values (PV), Individuals separate or together with the product (IS), Level of transparency (LT), Navigation (linear vs parallel) (NLP), Multiple use of links (ML), Sell approach (hard/soft) (SA), Use of aesthetics (UA), Color used (CU), Searching (SB), Polite/Direct (PD), Chronemics/Time (CT) and Use of flash features (UF).

2- Headings

By employing Hall's dimensions, Wurtz [1] carried out a quantitative evaluation in order to discover and elucidate the variations within website samples of both high context and low context cultures. To support his analyses, various nations were first assessed as either High Context (HC) or Low Context (LC). Particularly, Japan, China, and Korea were classified under HC while United States, Sweden, Germany, Norway, Finland and Denmark were classified as LC. In Wurtz's study, navigation was based on the attributes of the websites. On the other hand, Wurtz used the analytical model proposed by Thorlacius [6] and dwelt on the characteristics of the site which include animation, images, photographs and layout. These websites characteristics are relevant to our study in the sense that they are amongst the list of design parameters (DP) to be analyzed. Gygi and Spyridakis [7] proposed a cultural model that depended on the present demographic data and values so as to assess a school's website in Uzbek. This is based on the assumption that "…web sites designed by local producers for local users would embody and exhibit identifiable cultural markers". Additionally, they assessed the impact of language (English or Russian) in the Uzbek secondary school website. In the light of Ref [7], our study analyzed websites built for local users, though the LEEGA websites are all in English and not in any local Nigerian language.

To dictate the cultural context profile of the website of the South Africa (SA), Yeratziotis and Greunen [8] used the cultural model developed by Hall, to determine if the site's design and implementation is in accordance with the United Kingdom guidelines of e-government. They presented an evaluation of the SA website on a three point

scale compliancy level. It was however discovered in the study that the SA government site is not in accordance to the UK guidelines of website development. The website only furnishes a fundamental level of fulfilment for service delivery. Herselman and Greunen [9] carried out a global survey on cultural contexts variations which is aimed at supporting the government in making sure that their websites address the specific needs of the users. Their argument is on their contribution to the websites' usability and effectiveness. The study was done using questionnaire methods applied on selected populations. Ten participants were identified through purposive sampling and divided into two groups (5 persons) in low-context culture and (5 persons) in high context culture. Results of the study went contrary to literature and has it that high-context participants preferred more low-context styles when using government websites. Though this is a global survey, it informed our research method as well as the attributes of high and low context cultures presented in Table 1 and Table 2. The work also presented two continuums (Fig. 2) for high and low context cultures; wherein it is evident that Africa (or Nigeria) is still absent in discussions concerning website design and analysis.

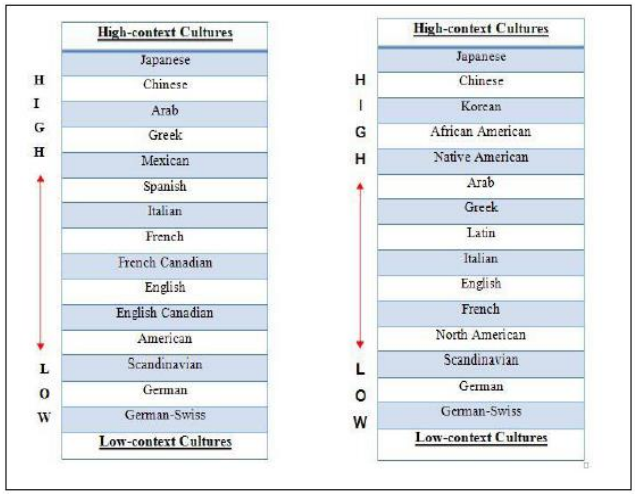


Fig. 2 High and Low-context Nationalities Scale according to Culture [9]

Vatrapu and Pérez-Quiñones [10] found out that the main hindrance to the development of interfaces is the insufficiency of empirical proofs showing the impact of culture on usability engineering approaches. Therefore, they presented a controlled experiment in order to study the impact of culture on usability testing. After two independent Indian groups tested a website for usability, the result generated shows that culture remarkably influences the success of structured interviews. The absence of a comprehensive review on the function of culture in web development spurred Moura, et al. [11] to compile and synthesize results of 90 prior studies. The review was structured into three major strongholds namely, the manifestation of cultural values in websites, the investigation of cultural markers in websites, and the influence of culture on users' perceptions of online stimuli. Finally, they suggested that the reviewed studies presents a convincing case of considering culture while adapting web features to meet expectations of the user. Refs [10,11] shows the need for performing a study such as ours, wherein the impact of culture context heuristics on websites are evaluated. At other quarters, it was discovered that designers hardly undertake the time consuming and costly task of creating culturally tailored forms of websites. This issue was traced to the absence of standard best practices or guidelines of web design and development. Having previously identified the relationship that exist between design features, cultural and HCI factors, Alexander, et al. [12] aimed at developing a set of guidelines for designing cross-cultural websites. They proposed 5 steps of development, which include evaluation of the usage of design elements between cultures, identification of prominent design elements, organization of cultural factors, organization of HCI factors, and finally development the cross-cultural design guidelines.

Umeh, et al. [13] evaluated the cultural profiles of websites to some Nigerian institutions, they include Cooperate Affairs Commission of Nigeria, Federal Inland Revenue Service of Nigeria, National Agency for Food and Drug Administration Commission of Nigeria, National Pension Commission of Nigeria, Nigeria Export Promotion Commission, Nigerian Tourism Development Corporation, National Emergency Management Agency of Nigeria. This is the first that investigates the cultural dimensions of Nigerian websites. The analyses performed using online survey method showed that websites of these institutions are high context styles in terms of animation/images, value promotion, transparency level, navigation, use of links, color, search boxes and polite/direct approach. Oyefolahan, et al. [14] assessed the usability and accessibility of websites owned by Nigerian commercial banks using automated tools and the manual inspection method. This was aimed at determining their conformance with the Web Accessibility Initiative. Findings of the study showed that significant enhancements are still needed in most of the websites. Although, ref [14] analyzed Nigerian websites like we did in our study, its analyses was absent insights on cultural context analyses. However, from the study we learnt of the existence of automated tools for website analyses and by searching the internet, we discovered SEO SiteCheckup [15] and Smallseotools [16], which was also used by our respondents.

Hofstede's cultural dimensions for Canadian immigration website was analyzed using Chinese, Indian, and Nigerian and French-speaking international students

[17]. The five dimensions include; Individualism/Collectivism, Power Distance, Uncertainty Avoidance, Time Orientation and Masculinity/Femininity. More so, analyses of other websites were for the following objectives; healthcare in European countries [18], cues for taking risks [19], e-commerce usability [20], and cross-cultural translation for Australia, China, and Saudi Arabia [21]. On a different perspective i.e. technology adoption, ref [22] e-government learning services. As Vatrapu and Pérez-Quiñones [10] puts it, "Cultural background influences the design methods employed in building interfaces". Based on these premises, we aim to evaluate the culture context heuristics of websites of law enforcement e-government agencies (LEEGA) in Nigeria using several design parameters. Refs [8,9] analyzed the culture context heuristics using the parameters gleaned from Table 1 (observations of the characteristics of high and low cultures websites) and Table 2 (present the other perspectives for values and features that are necessary when designing software products).

Table 1: Observations for High/Low Contexts in Web Design [1]

Parameter	***HC Cultures***	***LC Cultures***
Animation	High use of animation, especially in connection with images of moving people.	Lower use of animation, mainly reserved for highlighting effects (e.g., of text, active links, logos)
Promotion of Values	Images promote values characteristic of collectivist societies (e.g., being in good physical shape, spending time with family and friends)	Images promote values characteristic of individualistic societies (e.g. individuals are portrayed being in a more relaxed situations, such as holiday or listening to music – value personal time)
Individuals separate or together with the product	Featured images depict products and merchandise in use by individuals	Images portray lifestyles of individuals, with or without a direct emphasis on the use of products or merchandise
Level of transparency	Links promote an exploratory approach to navigation on the website; process-oriented	Clear and redundant cues in connection with navigation on a website; goal-oriented
Linear vs. parallel navigation on the Web site	More of a montage/layer-upon layer approach. Many sidebars and menus, opening of new browser windows for each new page	Few sidebars and menus, constant opening in same browser window

Table 2. High- and Low-context features [9]

High-context features	***Low context features***
Polychronic aspects of time	Monochronic aspects of time
Multiple use of images and/or banners	Less use of images and/or banners
Multiple use of links (external links promote a collectivist nature, working together)	Less use of links
Use of Flash features	Little use of Flash features
Being polite and indirect	Being direct and even confrontational
Create a friendly relationship with the customer (soft-sell approach)	Sales orientation (hard-sell approach
Use of aesthetics to elicit emotion (harmony, beauty, nature, art, designs)	Direct communication (focus on rank and prestige, superlatives, terms and conditions)

3- Headings

The research method for the study involves the following; identification of participants, (online) dissemination of the online survey to participants, analysis of survey data (questionnaires), preparation of survey report and presentation of results. Note that analyses here implies a two-fold evaluation which involves first, the analysis performed by the survey participants and second, the analysis we performed after we collated responses of the participants. Pertinent literature such as cultural anthropology, e-government, web design and computer-mediated communication were reviewed. As a result, we came up with a questionnaire that will allow the easy and accurate assessment of the law enforcement agencies' websites through the online survey method. A sample size of experts in web development and information technology were drawn from the six geopolitical regions of Nigeria contributed to the survey findings. Six groups (3 persons from each geo-political zone), making it a total of 18 persons. The questionnaire was designed to have biographical information, the design parameters (shown in Table 1 and Table 2, with some instructions of Table 3) and a section for open ended answers. Note that 13 design

parameters was used for the analyses. Aside using manual inspection method as well as some automated tools [15, 16] to reach conclusions while analyzing the LEEGA websites, we simply instructed the respondents to choose from the following 3 points namely High, Low and Not Applicable (NA). On the need for providing some guidance for the respondents, we established in Table 3 some rules to further guide the respondents. Since the tables culled from [1,9] were not explicit on how many

images/banner constitute high/low, how many colors constitute high/low, how many links constitute high/low and what time actually means on the sites etc. For chronemics (time), we instructed the respondents to judge context heuristics using these factors which include emphasis on relationship building, wide array of content, more exploration time, greater depth and breadth of content and busier design. On the promotion of values we instructed the respondents to look out for images that depict values such as unity, team work, cooperation, togetherness, and a cordial relationship with the masses. For aesthetics, they were instructed to use beauty (nature, art or design), emotion and sensation as opposed to pure intellectuality. Other specific parameters for judging site features are listed in Table 3.

Table 3. Rules for Further Guidance

Parameters	***High***	***Low***
AIB	4 & above	3 & below
CU	4 & above	3 & below
ML	Above 100	100 and below
UA	2 out of the 3 parameters	Less than 2 of the parameters
SB	1 or more search boxes	No search box
CT	3 & above factors for time	2 or less factors for time

Additionally, to automatically check whether a site has flash technology we instructed the respondents to use SEO SiteCheckup [15] or do a manual check by right clicking anywhere on the website and by checking for the presence of "Zoom in" and "About Adobe Flash Player" at the top and bottom of the box that appears. Additionally, we also instructed them to check the number of links on the sites using Smallseotools [16].

4- Results and Discussion

Recall that we simply instructed the respondents to choose from the following 3 points namely High, Low and Not Applicable. This simple three point method is suitable for evaluating the above-listed heuristics of e-governance websites. The collection of the responses took 8 days and afterwards we analyzed them to form the results of this study. Since the questionnaire involved a section for open ended answers, we also collated and analyzed the responses. Specifically, we assembled the results of all the experts and we used the choice of the majority to decide whether a design parameter (DP) is either high or low context. The results collation and analyses are listed below (in Table 4 and Table 5). Under each LEEGA webpage, we presented and discussed the reasons why respondents described a particular heuristic as either high or low context (as guided in Table 3, Table 4 and Table 5). Beside each discussion, we italicized and placed in parenthesis the matching culture context.

Table 4. Context Results for Agencies I

DP	**NAF**	**NN**	**NA**	**NPF**	**NPS**	**NIS**	**NSCD**
AIB	High	Low	High	Low	High	High	High
PV	High	High	Low	High	High	Low	High
IS	N/A	N/A	N/A	Low	High	Low	Low
LT	High	High	High	High	High	High	High
NLP	Low	Low	High	Low	Low	Low	Low
ML	High	High	High	Low	Low	High	High
SA	Low	Low	Low	Low	High	Low	Low
UA	High	Low	High	Low	High	Low	High
CU	High	Low	High	High	Low	Low	High
SB	High	High	High	Low	Low	Low	High
PD	High	Low	Low	Low	High	High	Low
CT	High	Low	High	High	High	Low	High
UF	Low	Low	Low	Low	Low	Low	Low

Table 5. Context Results for Agencies II

DP	**NCS**	**FRSC**	**FFS**	**DIA**	**NDLE**	**ICPC**	**EFCC**
AIB	High	High	High	High	High	Low	Low
PV	High	High	High	High	High	Low	High
IS	Low	High	N/A	Low	High	Low	N/A
LT	High	High	Low	High	High	High	High
NLP	Low	High	Low	Low	High	Low	High
ML	Low	High	Low	Low	Low	High	Low
SA	Low	High	Low	Low	High	Low	Low
UA	Low	High	High	Low	Low	Low	Low
CU	High	Low	High	High	High	Low	Low
SB	High	High	Low	High	High	Low	High
PD	Low	High	Low	Low	Low	Low	Low
CT	Low	High	High	Low	High	Low	Low
UF	Low	Low	Low	Low	Low	Low	Low

4-1- Nigerian Air Force (NAF)

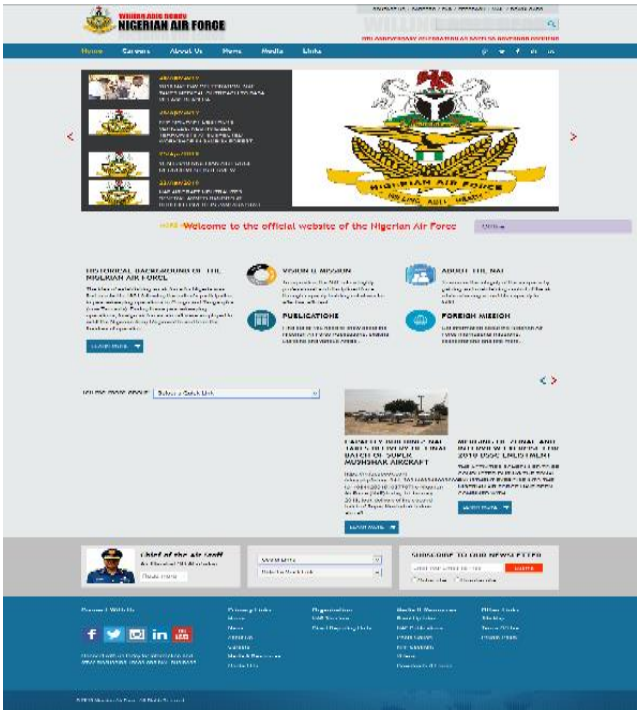


Fig. 3 Welcome Page of the Nigerian Air force [23]

The colors on the NAF website (Fig. 3) are light red, blue, white, yellow and black (high) while the alternating images/banners are 10 in number (high). On PV, some of the values promoted with photographs include togetherness, cooperate responsibility to the society and unity (high). More so, the picture where the force personnel were sharing food depicts unity and cooperative responsibility. Most of the respondents chose N/A for IS on the website and there reasons was that it promotes service delivery and not product sales. On level of transparency, they maintained the site is exploratory (high), while, for navigation they chose (low) because the site has few side bars and open in the same browser. On the use of aesthetics, most respondents said the site is beautiful and has pure emotion (high) as well as a friendly sell approach (low). The site has one search box (high) while for chronemics, the site displays some attempt at relationship building (high). The site has no flash features (low) and the links on the site are 115 in number with 110 internal and 5 external links (high).

4-2- Nigerian Navy (NN)

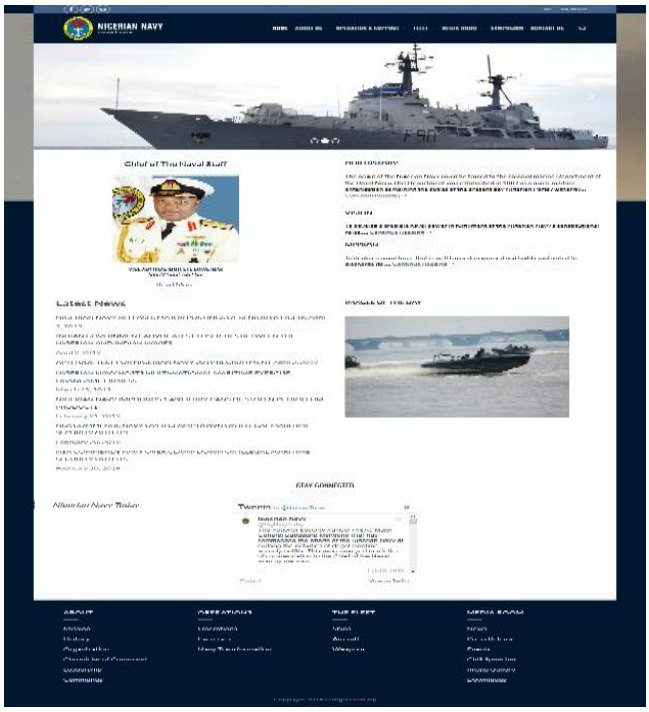


Fig. 4 Welcome Page of the Nigerian Navy [24]

Here (Fig. 4), the colors used for the website design are yellow, blue and white (low) while the alternating images/banners are 3 in number (low). On promotion values, the picture of personnel standing together depicts unity (high). The approach of sell here is that of (low) context, this is because the images present on the site makes no effort to relate with the public (N/A). The images does not suggest politeness but rather suggests a direct and confrontational position, this is largely because of the picture of the warship and that of the soldiers stationed to shoot (low). On navigation they chose (low) because the site has few side bars and open in the same browser. On level of transparency, they maintained the site is exploratory (high), while for navigation they chose (low) because the site has few side bars and open in the same browser. On the use of aesthetics, most respondents agreed that the site is simple yet beautiful but does not involve art, nature and designs (low). The site has no flash features (low) and the links on the site are 115(high) in number with 103 internal and 12 external links.

4-3- Nigerian Army (NA)

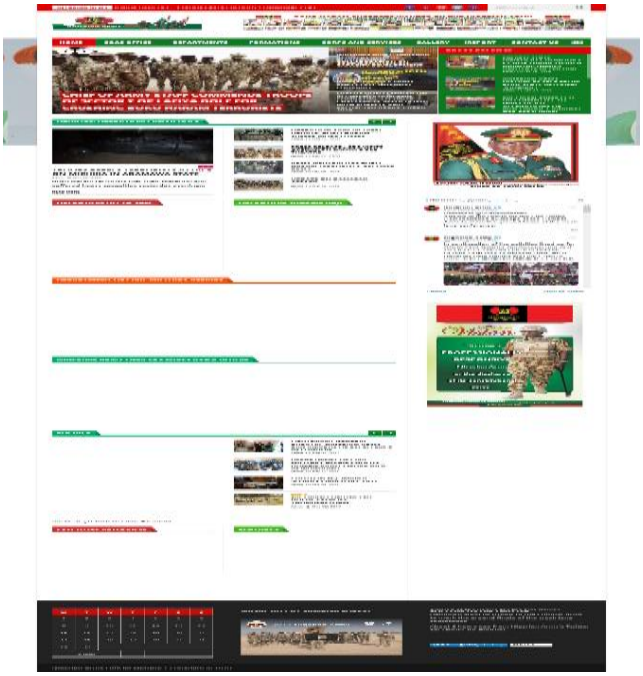


Fig. 5 Welcome Page of the Nigerian Army [25]

On the NA website (Fig. 5), two alternating banners and other images were scattered all over the home page (*high*). The colors used for the website design include green, red, white and black (*high*). On PV, most of the respondents agreed that the website did not promote values such as cordial relationship with the populace (*low*). The site maintains a direct sell approach (*low*); owing to images of guns, bullets and no relationship with the society. This is traceable to their aim of protecting the nation from external threats through the use of force. For IS, they chose *N/A*, because the heuristic is not applicable for the website. For NLP, the respondents agreed that the NA site is linear due to the many side bars and menus (*high*). At least a search box can be found on the site (*low*). On level of transparency, they maintained the site is exploratory (*high*). On the use of aesthetics, most respondents agreed that the site is beautiful and shows some form of nature and design (*high*). For chronemics, they agreed on more exploration time, greater depth and breadth of content and busier design (*high*). The site has no flash features (*low*) and the links on the site are 238 (*high*), 210 internal and 28 external links.

4-4- Nigerian Police Force (NPF)

Fig. 6 Welcome Page of the Nigerian Police Force [26]

The NPF home page (Fig. 6) has 3 alternating images (*low*) and the colors used for the website design include white, light, navy blue and dark blue (*high*). On the promotion of values, most of the respondents agreed that the NPF website promote values such as togetherness (*high*); this is expressly shown with 2 images of police personnel on a parade. On the level of transparency, the site used an exploratory and process-oriented approach (*high*). On the navigation type, the site has few side bars, many menus and it opens in same browser window (*low*). The sell approach is of *low* context i.e. the site employed a direct method. On the use of aesthetics (*low*), most respondents agreed that the site is just simple and cannot be described as beautiful. Additionally, it does not involve art, nature, designs or any form of emotion. For chronemics, they agreed on more exploration time, greater depth and breadth of content but did not agree on busier design (*high*). The site has no search box (*low*), no flash features (*low*) and the links on the site are 77 (*low*) with 73 internal and 4 external links.

4-5- Nigerian Prisons Service (NPS)

Fig. 7 Welcome Page of the Nigerian Prisons Service [27]

The NPS home page (Fig. 7) has 9 alternating images/banner (*high*) and the colors used for the website design include white, green and black (*low*). On the promotion of values, most of the respondents agreed that the NPS website promote values such as industry (on the part of the inmates), unity and perseverance (*high*). These values are clearly shown through the images of the inmates and the officers as well as the pictures of graduating inmate-students. These pictures also depicts IS i.e. the individuals together with the product (*high*). In this case the product(s) are the inmates that are in need for reformation. On the navigation type, the site has few side bars and menus (*low*). On the level of transparency, the site used an exploratory and process-oriented approach (*high*). The sell approach is of (*high*) context because the site has a desire to build relationships as well as a wide array of content. On the use of aesthetics, most respondents agreed that the site cannot be described as beautiful though it shows emotion and sensation as against pure intellectuality (*high*). The site depicts a polite approach (high) and displays some attempt at relationship building (*high*) for chronemics. The site has no search box (*low*), no flash features (*low*) and the links on the site are 50 (*low*), with 43 internal and 7 external links.

4-6- Nigerian Immigration Service (NIS)

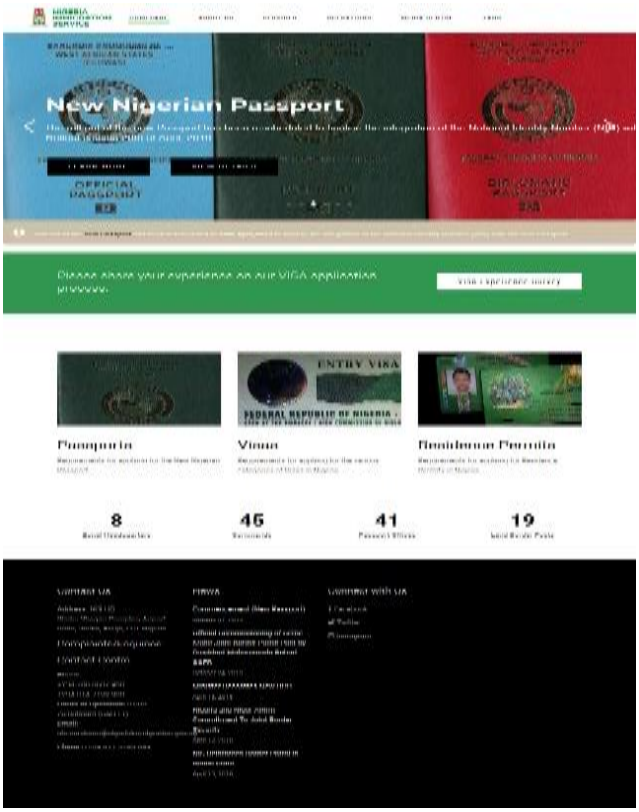


Fig. 8 Welcome Page of the Nigerian Immigration Service [28]

The NIS home page (Fig. 8) has 5 alternating images/banner (*high*) and the colors used for the website design include white, green and black (*low*). On promotion of values, most of the respondents agreed that the NIS website does not promote any value through the use of pictures (*low*). On IS, the site has no images together with any product (*low*). On the level of transparency, the site used an exploratory and process-oriented approach (*high*). On the navigation type, the site has few side bars and menus and opening in the same browser (*low*). The sell approach is of (*low*) context because the site employed a direct method. On the use of aesthetics, most respondents agreed that the site is not beautiful, does not display any form of emotion or sensation (*low*). The site depicts a polite approach (*high*) and but does not displays any attempt at relationship building (*high*) for chronemics. The site has no search box (*low*), no flash features (*low*) and the links on the site are 103 (*high*) with 97 internal and 6 external links.

4-7- Nigeria Security Civil Defense Corps (NSCDC)

Fig. 9 Welcome Page of the Nigeria Security Civil Defense Corps [29]

The NSCDC home page (Fig. 9) has 5 alternating images/banner (*high*) and the colors used for the website design include red, blue, white, black, green, and purple (*high*). On PV, most of the respondents agreed that the NSCDC website promote most especially the defense for the helpless (*high*). Another value promoted here is cooperation (with other agencies); a picture where the NSCDC personnel are standing together with Army officers. On IS, their commandant general can be seen sitting alone with an indifferent look (*low*). On the level of transparency, the site used an exploratory and process-oriented approach (*high*). On the navigation type, the site has few side bars and opening in the same browser (*low*). The sell approach is of (*low*) context because it does not depict any form of friendliness with the masses. The site is also confrontational and strongly maintains an impersonal atmosphere (*low*). On the use of aesthetics, the respondents agreed that the mixture of red and blue (or how blue gives way to red) makes the site very beautiful (*high*). This site according to the respondents suggests some form of emotion and sensation. For chronemics, it is clearly evident that the site has a wide array of content, more exploration time, greater depth and breadth of content and busier design (*high*). The site has a search box (*high*), no flash features (*low*) and the links on the site are 170 (*high*) with 164 internal and 6 external links.

4-8- Nigeria Customs Service (NCS)

Fig. 10 Welcome Page of the Nigerian Customs Service [30]

The NCS home page (Fig. 10) has 3 images and the alternating images of their past controllers (high). The colors used for the website design include green, ash, red, white and black (*high*). The image on the lower left hand side of the site promotes values such as cooperation and team work (*high*). On IS, the site has no images together with any product (*low*). On the navigation type, the site has few side bars and menus and opening in the same browser (*low*). On the level of transparency, they agreed that the site is exploratory and process-oriented (*high*). The sell approach is of (*low*) context because the site reveals no form of friendship with anyone. On the use of aesthetics, the respondents agreed that there is nothing visually appealing about the site (*low*). It is bland and without nature, art, design or any form of emotion/ sensation. The site has two search boxes (*high*) on the homepage for CET Tariff/Approved Form M. The CET page on the website has two search boxes too; for HS codes and for keywords. The NCS site is the only site with more than two search boxes. The atmosphere of the site is direct and confrontational (*low*). For chronemics, it is obvious that the site does not have a wide array of content, nor does it requires more exploration time, for site of content. In fact, the design is less busy (*low*). Additionally, the site has no flash features (*low*) and the links on the site are 95 (*low*) with 77 internal and 18 external links.

4-9- Federal Road Safety Corps (FRSC)

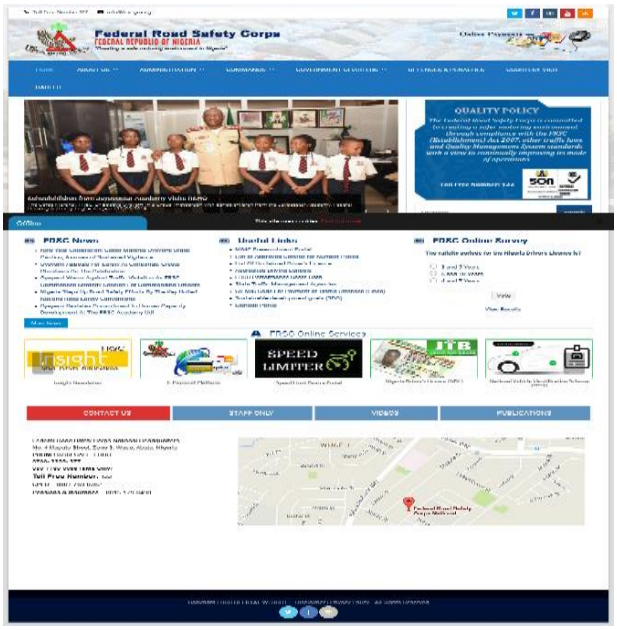


Fig. 11 Welcome Page of the Nigerian Federal Road Safety Corps [31]

The FRSC home page (Fig. 11) has 11 images (*high*) and the colors used for the website design include blue, red and white (*low*). The site promotes values such as cooperation, industry and team work through pictures of group leaning (high). For IS, the picture where the officers are with the children also shows corporate responsibility (*high*). On the level of transparency, the participants agreed that the site is exploratory and process-oriented (*high*). On the navigation type, the site has many side bars and opens in the same browser (*high*). The sell approach is of high context because of the friendliness between the officers and the children (high). This site was described as beautiful even though it is somewhat simple (*high*). Additionally, the site displays some form of emotion and involves a polite approach (*high*). For chronemics, it is clearly evident that the site encourages relationship building, has a wide array of content and requires more exploration time (high). The site has at least a search box (*high*), no flash features (*low*) and the links on the site are 186 (*high*) with 143 internal and 43 external links.

4-10- Federal Fire Service (FFS)

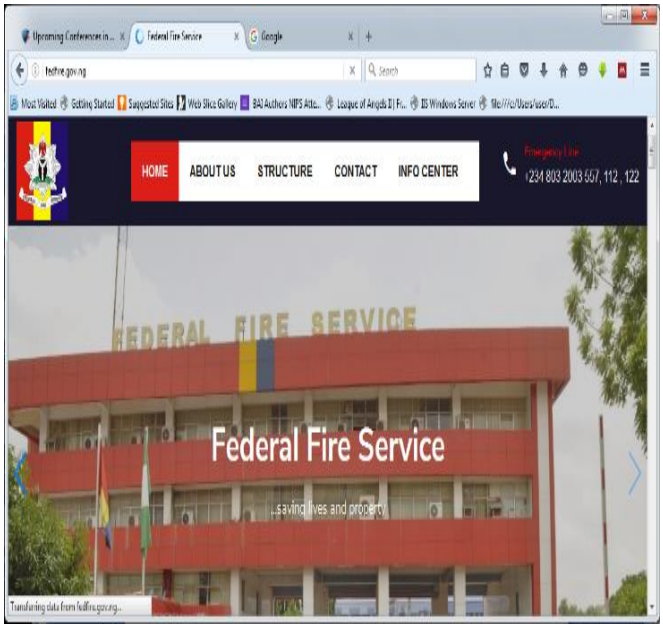


Fig. 12 Welcome Page of the Federal Fire Service [32]

The FFS home page (Fig. 12) has 5 alternating images/banner (*high*) and several other images placed on the gallery section. The colors used for the website design include white, red, green and black (*high*). On PV, the respondents agreed that the FFS website promotes safety (i.e. saving lives and property) as well as cooperation and teamwork using the pictures where their personnel are on parade (*high*). On IS, the site has no images of individuals together with any product (*low*). On the level of transparency, the participants agreed that the site involves clear and redundant cues in connection with navigation (*low*). On the navigation type, the site has few side bars/menus and opens in the same browser (*low*). The sell approach is of (*low*) context because the site employed a direct method. On the use of aesthetics, most respondents agreed that the site is beautiful but it doesn't display any emotion (*high*). The exudes a direct approach (*low*) and for chronemics, it is clearly evident that the site has a wide array of content, more exploration time, greater depth and breadth of content and busier design (*high*). The site has no search box (*low*), flash features (*low*) and the links on the site are 73 (*low*) with 61 internal and 12 external links.

4-11- Defense Intelligence Agency (DIA)

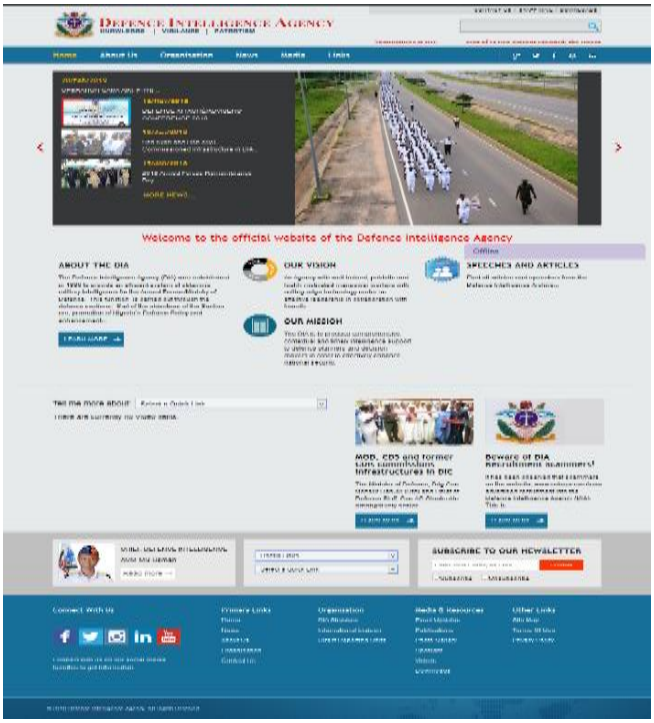


Fig. 13 Welcome Page of the Defense Intelligence Agency [33]

The DIA home page (Fig. 13) has 8 alternating images/banner (*high*) and the colors used for the website design include red, blue, white and yellow (*high*). On PV, the respondents agreed that the FFS website promotes fitness and being in good shape. The picture of officers matching on the road also depicts team work, cooperation, unity and professionalism (high). On IS, the site has no images together with any product (*low*). On the level of transparency, the site used an exploratory and process-oriented approach (*high*). On the navigation type, the site has few side bars and menus and opens in the same browser (*low*). The respondents confirms that the site is direct and confrontational due to the picture of the war helicopter (*low*). The site has three search boxes (*high*); one search box can be found at the top of the home page, while two search boxes are found below that same page. For chronemics, it displays a far less attempt at relationship building and does not have a wide array of content (*low*). On the use of aesthetics, most respondents agreed that the site is beautiful but it doesn't display any form of emotion, nature, art or design (*low*). The site has no flash features (*low*) and the links on the site are 95 (*low*) with 94 internal and 1 external link.

4-12- National Drug Law Enforcement Agency (NDLEA)

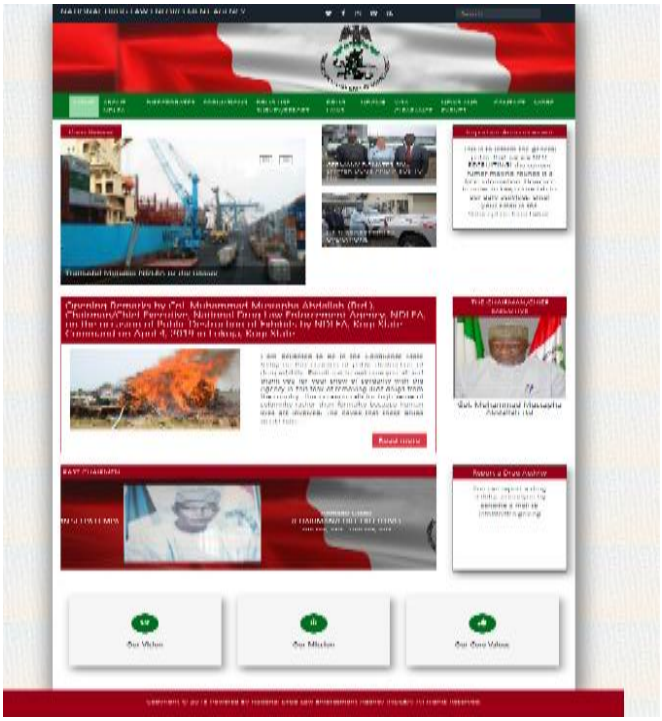


Fig. 14 Welcome Page of National Drug Law Enforcement Agency [34]

The NDLEA home page (Fig. 14) has 5 alternating images/banner and other images scattered on the site (*high*). The colors used for the website design include white, green, red (pure and dark) (*high*). On PV, the respondents agreed that the FFS website promotes deterrence and zero tolerance for drug crimes (*high*). The picture of officers from different agencies also shows team work and cooperation. The burning drugs also shows *high* context for IS. On the level of transparency, the site is exploratory and process-oriented (*high*). On the navigation type, the site has few side bars and opens in the same browser (*high*). On the sell approach, the collaboration with other agencies shows friendliness that suggests *high* context. The site has one search box (*high*) while for chronemics, the site displays some attempt at relationship building though it has no external links (*high*). On the use of aesthetics, the experts agreed that the site is neither beautiful nor does it display any form of emotion (*low*). The site exudes a direct atmosphere; this is expressly shown in the picture of the burning drugs – a zero tolerance for drug peddling. The site has no flash features (*low*) and the links on the site are 45 (*low*) with 45 internal and 0 external links. Additionally, the site the lowest number of external links in this study.

4-13- Independent Corrupt Practices and Other Related Offences Commission (ICPC)

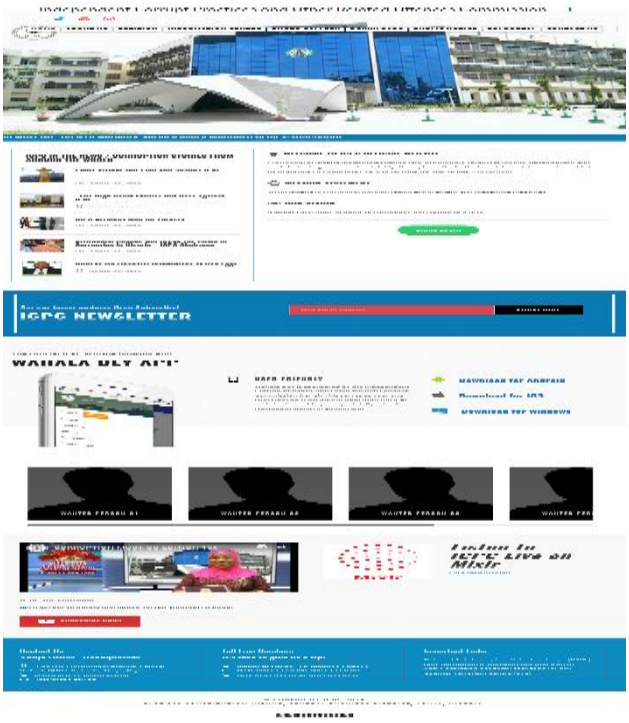


Fig. 15 Welcome Page of the Corrupt Practices and Other Related Offences Commission [35]

The ICPC home page (Fig. 15) has 3 alternating images/banner (*low*) and the colors used for the website design include white, blue and black (*low*). For PV and IS, no picture on the site promotes any form of values (*low*). On the level of transparency, the site is exploratory and process-oriented (*high*). On the navigation type, the site has few side bars and opening in the same browser (*low*). On the sell approach, the site suggest a hard sell orientation due to its cold and indifferent stance (*low*). The site has no search boxes (*low*) while for chronemics, the site makes no attempt at relationship building nor would it be adjudged to involve busier designs, therefore it depicts low culture context (*low*). On the other hand it has a wide array of content and requires more exploration time. On the use of aesthetics, the experts agreed that the site is beautiful but it does not have any special visual or animation effect (*low*). The site exudes a direct atmosphere (low). In addition, the site has no flash features (*low*) and the links on the site are 159 (*high*) with 148 internal and 11 external links.

4-14- Economic and Financial Crimes Commission (EFCC)

Fig. 16 Welcome Page of the Economic and Financial Crimes Commission [36]

The EFCC home page (Fig. 16) has 3 alternating images (*low*) and the colors used for the website include white, coffee brown and red (*low*). On PV, the picture of a handcuffed guy promotes deterrence (*high*). On the level of transparency, the site is exploratory and process-oriented (*high*). On the navigation type, the site has many side bars and opens in the same browser (*high*). On SA, the site suggest a hard sell orientation (*low*); this is because the organization doesn't exist to favor anyone. The site has one search box (*high*) while for chronemics, the site makes no attempt at relationship building nor would it be said to include busier designs, though it has greater depth. For aesthetics, the experts agreed that the site is beautiful but it does not have any special visual /animation effect, nature or art (*low*). Also, the site doesn't show any form of emotion/sensation. The site has no flash features (*low*) and the links on the site are 69 (*low*) with 63 internal and 6 external links.

From the responses collated, we confirmed that the reviewed sites were mostly arranged in the same manner when it comes to the level of transparency. Here, the main menu of the sites are linearly placed (close to the top) and by clicking on them, it displays horizontal submenus. Additionally, several information are placed horizontally as side bars making most of the sites seem exploratory and process-oriented. Most of the respondents agreed that the sites do not possess flash technology. Note that flash is an outdated technology that is used to deliver rich multimedia content. This technology poses some difficulty for web crawlers that may intend to interpret. More so it has some hitch-prone issues when assessed using mobile devices.

Table 6 presents the grand summary of collated results for the agencies i.e. it shows the cumulative choices made by respondents when assessing LEEGA websites. From this table it is clearly evident that these websites are of low context, however, one should note that the margin is not so wide. The second grand summary (Table 7) for the study presents the design parameters and the agency that are either high or low context.

Table 6. Grand Summary of Cultural-Context Result

S/N	***Agency***	***High***	***Low***	***N/A***
1	NAF	9	3	1
2	NN	4	8	1
3	NA	8	4	1
4	NPF	4	9	-
5	NPS	8	5	-
6	NIS	4	9	-
7	NSCDC	8	5	-
8	NCS	5	8	-
9	FRSC	11	2	-
10	FFS	5	7	1
11	DIA	5	8	-
12	NDLEA	9	4	-
13	ICPC	2	11	-
14	EFCC	4	8	1
	TOTAL	86	91	5

Consequent upon these tables, one can conclude that web designers for LEEGA follow the high context styles for animation, promotion of values, level of transparency, color used, searching and chronemics/time. On the other hand, the designers follow the low context style in terms of individuals separate or together with the product, navigation, sell approach, use of aesthetics, polite/direct and use of flash features. From the two tables, 5 responses were N/A; this was traced to IS.

Table 7. Grand Summary for Design Parameters

S/N	Parameter	High	Low	N/A
1	AIB	10	5	-
2	PV	11	4	-
3	IS	3	6	5
4	LT	13	1	-
5	NLP	5	9	-
6	ML	7	7	-
7	SA	4	10	-
8	UA	6	8	-
9	CU	8	6	-
10	SB	9	5	-
11	PD	4	10	-
12	CT	8	6	-
13	UF	0	14	-
	TOTAL	88	91	5

The rationale for this choice at certain websites was the fact that LEEGA basically offer services as opposed to the sale of products, which is usually the case for shopping websites.

On our instructions regarding the number of links that constituted high/low context cultures, the rule of thumb is to keep the links at 100 per website; this was taken from Google sources. Cutts [37] much later in a blog post quoted the Google guidelines saying, "Keep the links on a given page to a reasonable number (fewer than 100)". This is traceable to reasons such as bandwidth utilization, fitness and processing limits. The consequences according to SEO and Matt, is that "the more links a page has, the less internal PageRank each of those links passes" and in Matt's words, "…at any rate, you're dividing the PageRank of that page between hundreds of links, so each link is only going to pass along a minuscule amount of PageRank anyway". While reviewing the number of links used on the sites, we discovered that the NA site has the highest number of links (on a total) i.e. 238 as well as the highest number of internal links i.e. 210. This is followed by the FRSC site with 186 total links. On internal links, NA is followed by the ICPC site which has 148 links. For external links, the FRSC site has the highest number i.e. 43, and this is followed closely by the NA site which has 28. While the DIA site has only one external link, the NDLEA site has no external link.

5- Conclusions

In this study, we employed the online survey approach which is aimed at evaluating the cultural anthropology/profile of law enforcement agencies in Nigeria. Interestingly, the derived conclusions confirms the uniqueness of the efforts expanded in this research. This is because the majority of websites' usability studies in HCI and informatics make use of financial institution and shopping websites. Questionnaires that represent the features of the above-mentioned proponents of cultural models were distributed to a group of experts and their answers were collected, analyzed and presented in the tables above. The results show that for law enforcement agencies in Nigeria, the culture context profile is predominantly low (91 out of 182 expected statuses on Table 6 and 91 out of 184 expected statuses on Table 7). This is against the high context results, which are 86 out of 182 expected statuses on Table 6 and 88 out of 184 expected statuses on Table 7.

It is noteworthy to reiterate here that the margin in between (cumulative high and low choices/responses) is not wide. First, this result is in contradiction to Umeh, *et al*. [13], wherein the context culture is high (42 out of 77 expected statuses) for institutional websites in Nigeria. Additionally, our result contradicts the conclusion drawn

by Köszegi, *et al*. [38], which goes thus, "If countries from Africa were to be positioned on the diagram (Fig. 17), they would reside somewhere within the red circle". Although this diagram performed a global evaluation of cultural dimensions of websites from several nationalities, the bias is on communication patterns. More so, it represents the first attempt to situate African nations in a culture context continuum.

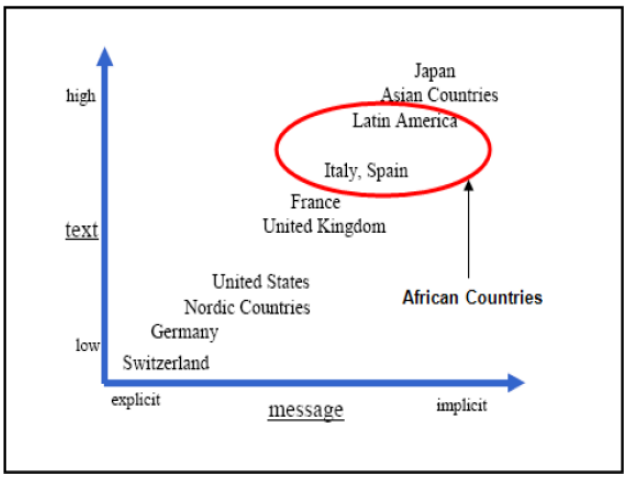


Fig. 17. Communication Patterns of Different Countries [38]

If we are to use Fig. 17, considering the direction of our study and its results, it would be wrong to situate Africa within the red circle (i.e. Latin America and Italy, Spain). The position of Africa would go lower on the continuum; within United Kingdom and Germany (Fig. 18).

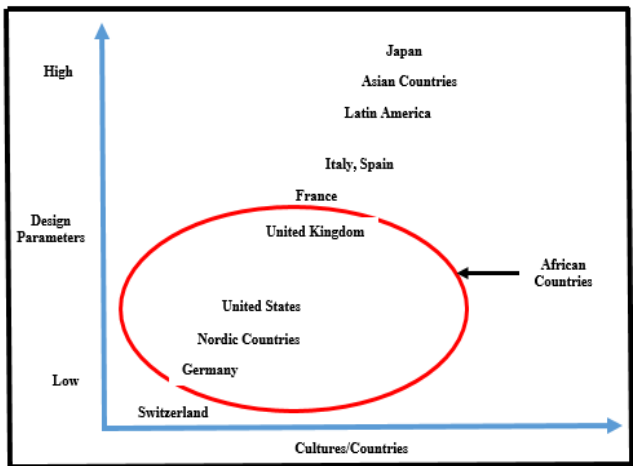


Fig. 18. Proposed Culture Contexts Continuum involving Africa

Considering Fig. 2 above alongside the aim of our study herein, Nigeria (Africa) would also come down to fall within French Canadian and Scandinavian (for the diagram on the left hand side) and within Latin and Scandinavian (for the diagram on the right hand side). However, since all these are speculations with little proof (except perhaps our study), there is need for a holistic study on the cultural dimensions of African nations. This would properly place Africa in any of the listed continuums. On the other hand, we noticed that none of the LEEGA websites used an indigenous language, even though Nigeria is a country with several indigenous languages. This is quite pitiable if one considers websites from Uzbekistan, China, Germany etc. Most times these sites are first displayed in their national language but with the help of an embedded language translator the visitor can use another language he/she feels comfortable in. And as Gygi and Spyridakis [7] puts it, therein, "the national language becomes a cultural marker". For a multilingual nation like Nigeria, using one national language might be a problem, which is why most websites use English – Nigeria's official language. Consequently, website developers can add the language translating capability for the three major languages in Nigeria i.e. Igbo, Hausa and Yoruba.

In the future, we hope to further investigate the cultural anthropology of websites designed for Africans using the design guidelines developed by Alexander, *et al*. [12, 21] as well as the Hofstede cultural dimension. The current study can also be extended by evaluating the accessibility and usability of websites owned by law-enforcement agencies in Nigeria.

Acknowledgments

We acknowledge the experts (web developers) scattered across the various regions of Nigeria for bringing out time from their busy schedules to attend to our questionnaires. The success of the study was largely because of their efforts. Additionally, we thank the anonymous reviewers that painstakingly ensured the accuracy of our work.

References

[1] E. Wurtz, "Intercultural communication on web sites: A cross-cultural analysis of web sites from high-context cultures and low-context cultures", *J. of Computer-Mediated Communication*, vol. 11, 2006, pp. 274–299.

[2] E. T Hall, *The Silent Language*, New York: Doubleday. 1976.

[3] G. Hofstede, *Culture's consequences: International Differences in Work-Related Values*. UK: Sage, 1984.

[4] D. Morley and K. Robins, *Spaces of Identity: Global Media, Electronic Landscapes and Cultural Boundaries*. US: Routledge, 1995.

[5] M. D. Myers, and F. B. Tan, "Beyond models of national culture in information systems research", *Journal of Global Information Management*, vol. 10, no. 2, pp. 23-56, 2002.

[6] L. Thorlacius, "Visuel Kommunikation pa Web sites," Ph.D. dissertation, Roskilde Universite, 2002.

[7] K. F. Gygi, and J. H. Spyridakis, "Developing a cultural model to support web site localization: a case study of Uzbek school websites," in *Annual Conference Society for Technical Communication*, vol. 54, 2007, pp. 1-20.

[8] A. Yeratziotis, and D. Greunen, "E-government – putting service at your fingertips," in *IST-Africa 2009 Conference & Exhibition*, 2009, pp. 1–15.

[9] M. Herselman, and D. Greunen, "Global survey on culture differences and context in using e-government systems: a pilot study," in *Second International ICST Conference, AFRICOM*, 2010, pp. 46–68.

[10] R. Vatrapu, and M. A. Pérez-Quiñones, "Culture and international usability testing: The effects of culture in structured interviews", *Journal of Usability Studies*, vol. 1, no. 4, pp.156-170, 2006.

[11] F. T. Moura, N. Singh and W. Chun, "The influence of culture in website design and users' perceptions: Three systematic reviews," *Journal of Electronic Commerce Research*, vol. 17, no. 4, pp. 312-339, 2016.

[12] R. Alexander, D. Murray, and N. Thompson, "Cross-cultural web design guidelines," *W4A 2017*, 2017, pp. 1-16.

[13] I. Umeh and C. H. Nwokoye, "Culture Context Profiles: A Case of Institutional Websites in Nigeria", *Global Journal of Computer Science and Technology: G Interdisciplinary*, vol. 17, pp. 1-17, 2017.

[14] I. O. Oyefolahan, A. A. Sule, S. A. Adepoju, and F. Babakano, "Keeping with the global trends: An evaluation of accessibility and usability of Nigerian banks website," *I. J. Information Engineering and Electronic Business*, vol. 2, pp. 44-53, 2019.

[15] SEO, Assessed on: July 24, 2020. [Online]. Available from https://seositecheckup.com/seo-audit/flash-test/

[16] Smallseotools, Assessed on: Aug 19, 2020. [Online]. Available from https://smallseotools.com/website-link-analyzer-tool/

[17] J. Ouellet, J. Kang and A. Girouard, "Cultural and linguistic variables in usability testing: A Canadian evaluation of international students," *Journal of Usability Studies*, vol. 16, pp. 121–147, 2021.

[18] C. Sik-Lanyi and E. Orbán-Mihálykó, "Analyses of websites was done for healthcare in European countries," *Arabian Journal for Science and Engineering*, vol. 44, pp. 9171–9190, 2019.

[19] Z. Nong and S. M. Gainsbury, "Website design features: exploring how social cues present in the online environment may impact risk taking," *Human behavior and emerging technologies*, vol. 2, no. 1, pp. 1 – 35, 2019.

[20] N. Al-Shammari, "E-Commerce in Saudi Arabia: characteristics of a trustworthy usable e-commerce websites," *International Journal of Information Science*, vol. 9, pp. 6-15, 2019.

[21] R. Alexander, N. Thompson and D. Murray, "Towards cultural translation of websites: a large-scale study of Australian, Chinese, and Saudi Arabian design preferences," *Behaviour & Information Technology*, vol. 36, pp. 351 – 363, 2017.

[22] L. Al-Omairi, H. Al-Samarraie, A. I. Alzahrani and N. Alalwan, "Students' intention to adopt e-government learning services: a developing country perspective," *Library Hi Tech*, vol. 39, pp. 308-334, 2020.

[23] Nigeria Airforce, January 2020. Assessed on: June 20, 2021. [Online]. Available from https://airforce.mil.ng/

[24] Nigeria Navy, March 2020. Assessed on: June 20, 2021. [Online]. Available from http://www.navy.mil.ng/

[25] Nigeria Army, January 2019. Assessed on: June 20, 2021. [Online]. Available from https://www.army.mil.ng/

[26] Nigeria Police Force, January 2019. Assessed on: June 20, 2021. [Online]. Available from https://www.npf.gov.ng/

[27] Nigeria Prisons Service, January 2019. Assessed on: May 5, 2019. [Online]. Available from http://www.prisons.gov.ng/

[28] Nigeria Immigration Service, January 2019. Assessed on: June 20, 2021. [Online]. Available from https://immigration.gov.ng/

[29] Nigeria Security and Civil Defense Corps, January 2019. Assessed on: June 20, 2021. [Online]. Available from http://nscdc.gov.ng/

[30] Nigeria Customs Service, January 2019. Assessed on: June 20, 2021. [Online]. Available from https://www.customs.gov.ng/

[31] Federal Road Safety Corps, January 2019. Assessed on: June 20, 2021. [Online]. Available from https://frsc.gov.ng/

[32] Federal Fire Service, January 2019. Assessed on: June 20, 2021. [Online]. Available from http://fedfire.gov.ng/

[33] Defense Intelligence Agency, January 2019. Assessed on: June 20, 2021. [Online]. Available from http://www.dia.gov.ng/

[34] National Drug Law Enforcement Agency, January 2019. Assessed on: June 20, 2021. [Online]. Available from https://ndlea.gov.ng/

[35] Independent Corrupt Practices and Other Related Offences Commission, January 2019. Assessed on: June 20, 2021. [Online]. Available from http://icpc.gov.ng/

[36] Economic and Financial Crimes Commission, January 2020. Assessed on: June 20, 2021. [Online]. Available from https://efccnigeria.org/efcc/

[37] M. Cutts, *How many links per page*. Gadgets, Google, and SEO, 2009, January 2020. Assessed on: January 30, 2020. [Online]. Available from https://www.mattcutts.com/blog/how-many-links-per-page/

[38] S. Köszegi, R. Vetschera, and G. E. Kersten, "National cultural differences in the use and perception of Internet-based NSS: Does high or low context matter?" Assessed on: January 30, 2020. [Online]. Available from interneg.concordia.ca/views/bodyfiles/paper/2003/09.pdf

ChukwuNonso H. Nwokoye obtained a BSc and PhD degrees in Computer Science. He is two-time ACM SIGCHI Gary Marsden Student Award recipient. His interests include simulation and modeling of complex systems, semantic web, agent-based modeling, wireless sensor networks and network security, human computer interaction (social computing, computer supported cooperative work) and computer mediated communication. He is currently on evaluating the accessibility and usability of websites owned by government agencies in other African countries.

Njideka N. Mbeledogu is a Senior Lecturer at the Department of Computer Science, Nnamdi Azikiwe University, Awka, Nigeria. She holds Bachelor of Science (BSc), Master of Science (MSc) and PhD degrees in Computer Science. She is a member of the Computer Professional of Nigeria (CPN) and the Nigerian Computer Society (NCS). Her interests include network security, neural networks, semantic web, human computer interaction and fuzzy logic

Chikwe Umeugoji obtained a BSc in Computer Science from Nnamdi Azikiwe University, Awka, Nigeria. She is currently a postgraduate student at the same institution, wherein she is currently working on design and implementation of mobile and desktop applications for diverse disciplines. Her other disciplines include network security, neural networks, semantic web, human computer interaction and fuzzy logic.

Proposing Real-time Parking System for Smart Cities using Two Cameras

Phat Nguyen Huu*
School of Electronics and Telecommunications, Hanoi University of Science and Technology, Hanoi, Vietnam
phat.nguyenhuu@hust.edu.vn
Loc Hoang Bao
School of Electronics and Telecommunications, Hanoi University of Science and Technology, Hanoi, Vietnam
Loc.HB202786M@sis.hust.edu.vn

Received: 03/Oct./2020 Revised: 18/March/2021 Accepted: 18/June/2021

Abstract

Today, cars are becoming a popular means of life. This rapid development has resulted in an increasing demand for private parking. Therefore, finding a parking space in urban areas is extremely difficult for drivers. Another serious problem is that parking on the roadway has serious consequences like traffic congestion. As a result, various solutions are proposed to solve basic functions such as detecting a space or determining the position of the parking to orient the driver. In this paper, we propose a system that not only detects the space but also identifies the vehicle's identity based on their respective license plate. Our proposal system includes two cameras with two independent functions, Skyeye and LPR cameras, respectively. Skyeye module has function to detect and track vehicles while automatic license plate recognition system (ALPR) module detects and identifies license plates. Therefore, the system not only helps drivers to find suitable parking space but also manages and controls vehicles effectively for street parking. Besides, it is possible to detect offending vehicles parking on the roadway based on its identity. We also collect a set of data that correctly distributes for the context in order to increase the system's performance. The accuracy of proposal system is 99.48% that shows the feasibility of applying into real environments.

Keywords: object detection; single shot detector; multi-view cameras; automatic license plate recognition; ALPR.

1- Introduction

With the development of the economy today, cars are becoming a popular means of life. This rapid development has resulted in an ever-increasing demand for private parking beyond the available supply. Therefore, the situation of not meeting the demand for parking is increasingly serious. Not only in terms of available parking space that does not meet but also the main common management method that depends on human communication with physical space entities. This leads to the waste of manpower as well as efficiency of the parking space. The main difficulty comes from not knowing about the status of parking spaces that are empty for a given period of time. Most of the parking always have space. However, the parking efficiency is low due to inefficient storage.

Finding a suitable parking space in an urban city without knowing the location of the parking lot is extremely difficult for the driver. That's why driving around in search of a parking space is a frustrating experience. The problems lead to unsafe driving, time consuming, fuel consumption, and increasing emissions that pollute the environment and cause traffic congestion. Besides, illegal parking is more popular. To solve this problem, parking information and guidance systems (PGI) were developed. The PGI system requires real-time accuracy and responsive updates on the status of a parking space since it can provide users with vacancy and optimize management.

The outstanding advantages of PGI systems that use cameras comparing to other existing systems include: first, no additional infrastructure are required; secondly, camera-based systems provide exact location of empty parking spaces; thirdly, the method of using cameras are highly applicable to the parking spaces on the roads and residential areas.

Image-based parking occupancy detection is basically related to the problem of vehicle detection in the parking space. One of the most important issues with the detection is whether the training, development and test data share the actual distribution. Therefore, we have collected a new data set with the correct distribution for the context. We then use transition learning to increase system performance.

To manage the parking space and information about its status for drivers is an important issue. Therefore, a proposal system not only detects the total amount of free

* Corresponding Author

space but also identifies the identity of each respective vehicle in the paper. Besides, sharing information about the status of the parking space helps to reduce the time to search for a parking location. Vehicle identification is based on a license plate that can be paid online, a parking time notification for the driver, and a vehicle offender. Those are our main contributions.

2- Related Work

Smart parking is one of the most applied and fastest growing solutions worldwide. Due to the inadequacies leading to serious consequences such as time, traffic congestion, and environmental pollution, smart parking has become an attractive field for many researchers. By using low-cost sensors, real-time data and response allows the user to monitor parking spaces. The goal is to automate and reduce the search time for drivers. Besides, we also propose solutions that include a complete set of services such as online payment, parking time notification, and driver orientation.

Existing PGI systems are divided into 4 categories: 1) counter-based systems, 2) wired sensor-based systems, 3) wireless magnetic sensor-based systems, and 4) camera or imaging based system. The counter-based system is dependent on the sensors at the parking. This system can only provide information about the total number of seats instead of guiding the driver to the exact location of the parking space. Therefore, it is not applicable for on-road parking. Wired sensor-based systems and wireless magnetic sensor-based systems depend on ultrasonic sensors, infrared light or wireless magnetic sensors which are installed on each parking space location. Both systems have been adopted in real commercial and used for large shopping malls. However, these methods require the installation of wireless technology sensors, processors, and transceivers. Although sensor-based systems are highly reliable, their high installation and maintenance costs limit for a wide range of applications. To compare with systems using sensors, systems using cameras are cost effective since both monitoring and blank detection functions can be performed simultaneously.

Various methods and techniques have been proposed to deal with parking problems in urban areas. In many modern car parks, popular systems such as ground sensors are used to determine the status of parking locations. The methods use range from using RFID radio frequency identification [1] [2], or infrared sensor [3] to ultrasonic sensor. One of these is the RFID-based smart parking system developed by [4] that implements an automatic process of entry control and exit control using an RFID. Another system is proposed by [5] based on the effect of magnetic on sensors. It is used to determine the state of the parking space. Besides, authors propose smart parking service systems using short message services (SMS) or a global system for mobile devices (GSM) or microcontroller to enhance security by [6]. ZigBee technique is used together with GSM module to manage and reserve parking by [7]. These systems achieve the function of which is primarily required by blank detection. However, it is affected by factors such as changing weather conditions or the presence of electromagnetic interference.

Node-to-node implementations require a lot of time. For these systems, it is required to set up, install, and maintain sensors in each parking space. Therefore, the cost is very expensive especially in parking spaces with a large number of locations.

Through camera-based methods, empty parking positions can be identified. The data being processed and the generated result will precisely determine the specific number and location of the parking space. Zhang Bin et al. proposed a vision-based parking detection method that is easy to install, low cost. Besides, the detector can be easily adjusted according to requirements. Furthermore, image data is increasingly plentiful and diverse. However, the disadvantage of these methods is that the accuracy is highly dependent on the camera position. H. Ichihashi et al. has proposed a parking space detection system based on a vision that is affected mainly by weather and lighting conditions such as raindrops falling on the lens of camera. For these reasons, cameras are predominantly used for detecting vehicles in indoor parking spaces [8].

In [18] the authors used the wide-angle camera as a sensor. It detects only empty parking spaces and records them. Post-processing information is used to designate a parking space for the newly arrived driver. The Intelligent Traffic System (ITS) and Electronic Toll Collection (ETC) use character recognition (OCR) to create a record for all incoming vehicles. This produces less entry tags for all vehicles in the parking but it does not assign a position for the new driver. A generalized character recognition algorithm is not possible that makes it difficult to create the records shown in [17]. Another proposing system by image processing is proposed in [9]. In this article, an image processing technique is demonstrated to capture the brown circle drawn on the parking area and processed it to detect empty space. In [10], an image of vehicle is saved as a reference and the others are matched with the reference using edge detection and information about the positions and displays. Besides, several methods have been proposed to extract features from images such as [11-15]. With the great progress and achievements gained from deep learning, the authors have come up with other methods by using Artificial Intelligence (AI) to process input images from cameras. Advances of multicore architecture allow to use of convolutional deep neural network architectures (CNN) to detect and classify objects [16,17]. Mauri et al. [22] mentioned the object detection

methods that are divided into two categories: (i) single-stage object detector and (ii) two-stage object detection detector. Several popular single-stage object detection algorithms are Single Shot Detector (SSD) [32], versions of You Only Look Once (YOLO) [18,19]; Two-phase object detection algorithms such as RCNN (Region proposal CNN) [20,21]. By using the proposal algorithms, [23] shows that the model for the media detection problem has the ability to extract good features and high results.

As analysis above, we found that current methods focus on detecting space and identifying license plates. In this article, we therefore propose an automatic vehicle management system for street car parks by applying deep learning. Besides, our system can also identify each vehicle since it

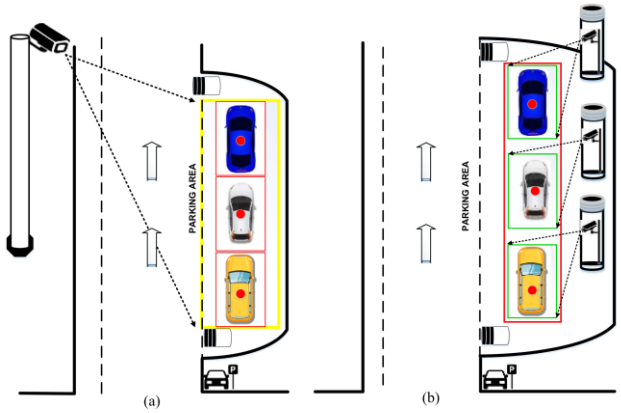


Fig. 1 System using (a) one camera, (b) multiple cameras

can be automatically charged electronically. Experimental results have high accuracy when comparing with existing methods. Therefore, it shows that the feasibility of the application system in practice is very high.

The proposed parking model is discussed in detail in section 3. In part 3, the proposed algorithm of a parking system is discussed in detail. Part 4 presents experimental results achieved and evaluated. Finally, conclusions will be drawn in section 5.

3- Proposed System

Realistic requirements for the system include:

- Accurately detecting empty space;
- Identifying the vehicle based on the license plate.

During the analysis and design, the following cases occur: 1) a system using one camera and 2) a system using multiple cameras are shown in Fig. 1. For a system using one camera, the camera performs both functions. The system is low cost and easy to use. However, the license plate recognition function has low accuracy since it is not possible to recognize the number plate for cars with a long distance.

For a system using multiple cameras, each parking position will be installed with a camera with a low angle in order to be able to observe the license plate when the vehicle enters the parking. The advantage of system is its high accuracy. Its disadvantage is high energy consumption and cost.

Therefore, we propose to use two cameras to perform two separate functions in the paper. Specifically, a camera has a high resolution and is mounted in a low position for the purpose of detecting and identifying the license plate most accurately. The other camera will have a wide angle to maximize the parking area to detect and track objects entering the parking.

3-1- Proposal System

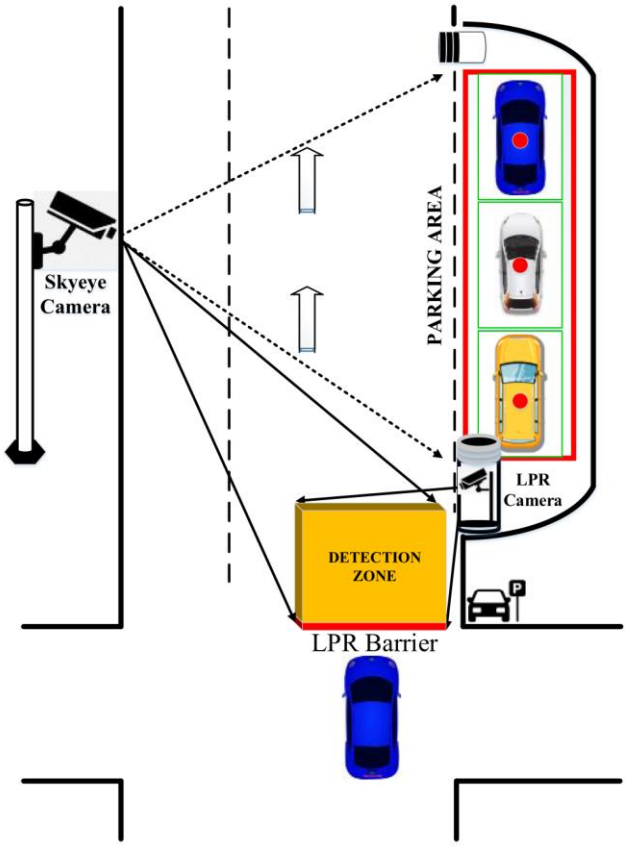


Fig. 2 Diagram of proposal system

Due to the close relationship of object detection to video and image analysis, the proposed and implemented object detection methods have achieved certain achievements recently. Traditional object detection methods are built on image characteristics drawn from experience. The performance of the methods is easily affected by the construction of complex clusters that combine low-level image characteristics and high-level features. With the rapid development of deep learning, as

well as computer science bring real-time response systems with very high accuracy to solve problems that exist in traditional methods.

The initial requirement of problem is to accurately detect the space position of car parking and control the vehicle entering by determining its identity. Our proposed system uses two cameras with two independent functions, Skyeye and LPR cameras, respectively. We compare with real-world feasible systems as follows:

For an example, a smart car park for a condominium or planned parking lot that includes:

Objects are apartment car or planned parks. The system recognizes the license plate number correctly since the vehicle must stop in front of the barrier for procedures. From there, easily identify the vehicle's identity and collect the toll. However, this system is only suitable for planned parking lots and available facilities.

Other case is an image of a smart parking space with wireless sensor networks. Subjects are street parking lots. The system only detects whether the parking lot has space or not. However, the system does not recognize the license plate since it will not be able to identify the vehicle. Deploying the system will be expensive since it depends on the number of slots corresponding to the number of sensors needed. Besides, the system is affected by environmental conditions.

Therefore, we design the two respective main modules as follows:

- Skyeye module has function to detect and track vehicles.
- ALPR module detects and identifies license plates.

Our proposal system is shown in Fig. 2. It includes:

- Skyeye camera is a wide angle camera that covers the car parking.
- LPR camera is a high resolution camera to detect the license plate of vehicle.
- Detection zone is the area that detects when new vehicles enter.
- LPR barrier is a virtual barrier to detect the license plate of vehicle when passing the barrier.

With the management of information and locating the assistive vehicle for the operator, we use images to capture from a wide-angle camera (Skyeye camera) as input to the Skyeye module that allows to determine the positions. Therefore, the system also detects the number of vehicles in the parking area. Furthermore, the Skyeye camera also detects and tracks new objects entering the car parking.

With the function of identifying number plates to determine the identity of each vehicle when entering the car parking, LPR camera has a high resolution and long zoom capability.

There are two main difficulties of proposal system:

- One issue related to the accuracy of the LPR module is camera placement. When the vehicle has entered the parking spaces, the license plate is likely to be obscured by the front and rear objects. Therefore, we have installed a license plate detection camera as shown in Fig. 2 to detect it before the vehicle enters the car parking to increase the performance of ALPR module.
- The second is a matter of synchronizing two modules in order to know how to be the same object. Our solution is to process each module sequentially.

The system operation is shown in Fig. 3 as follows:

- Skyeye module is always active with the input of images collecting from the Skyeye camera. After processing, information about parking status will be shared to users. In the case of a new vehicle entries, the Skyeye module detects the vehicle in the detection zone and sends trigger signal to the LPR module.
- LPR module is activated to detect and recognize number plates. The results are the letters of the number plate.
- Assigning #ID to the newly detected vehicle with license plate characters and tracking the vehicle using centroid tracking algorithm based on [24] and checking if the vehicle enters the parking.
- Updating and displaying information.

Above is an overview of the proposal system. In next section, we will present more details about the modules.

3-2- Skyeye Module

The Skyeye module consists of two main blocks: 1) Vehicle detection and 2) vehicle tracking. Vehicle detection is one of the important applications of object detection problem in ITS. It is intended to extract vehicle-specific information from images or videos containing the vehicle. To solve current vehicle detection problems, such as vehicle classification, low detection accuracy and real-time non-response, we perform evaluation and comparison of state-of-the-art detection algorithms. Using the network architectures is to extract features to detect the object with one-time detectors such as SSD, YOLO on data sets such as ImageNet, COCO [30].

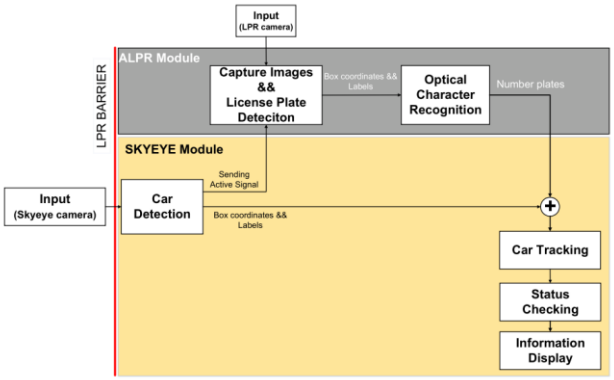


Fig. 3 The proposed system's operational modules

Vehicles are one of the basic objects present in many basic identification and detection datasets, such as PASCAL-VOC [25], ImageNet [26], and COCO [30]. Therefore, we decided to select a trained model for vehicle detection after comparing the performance of different network architectures as shown in Tab. 1.

Table 1: Object detection results on data set PASCAL VOC 2007 and MS-COCO

CNN	*VOC07*		*COCO*	
	FLOPs	***mAP***	***FLOPs***	***mAP***
SSD-512 [32]	90.2 B	**74.9**	99.5 B	**26.8**
SSD-300 [32]	31.3 B	72.4	35.2 B	23.2
YOLOv2 [33]	6.8 B	69.0	17.5 B	21.6
MobileNetv1-320 [34]	-	-	1.3 B	22.2
MobileNetv2-320 [35]	-	-	**0.8 B**	22,1

We use a model that has been trained on COCO data set to save time and initialize initial parameters to learn features faster. The using featured extractor is MobileNet_v2 paired with an SSD single-phase detector [32]. Overall, Single Shot Multibox Detectior (SSD) [32] is used to reduce model size and complexity. It works by using multiple feature maps along a network. Therefore, the network can use this information to predict large objects through deeper layers, as well as predict small objects using shallow layers. We do not make any changes or refinements to the SSD- MobileNetv2 and only use it as a black box to merge outputs related only to one vehicle being a car and ignore the remaining layers. The process of implementing consists of two main phases: one is the training phase and the other is the test phase to evaluate the performance of the model as shown in Fig. 4. After detecting the object, it will track using the centroid tracking algorithm to check if the vehicle enters the car parking.

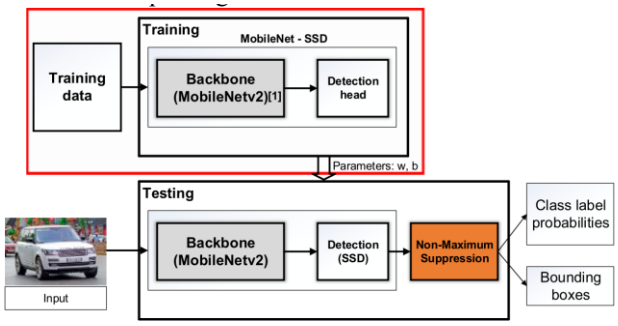


Fig. 4 Diagram of vehicle detection block implementation

Training Objectives: The training target of SSD is based on MultiBox [36,37]. It is extended to handle a wide variety of objects. Setting x_{ij}^{p} = {1, 0} is specified for the first default box to the j[th] truth of type *p*. The total loss function is the weighting of the localization loss (loc) and the confidence loss (conf) defining as follows:

$$L(x,c,l,g)=\frac{1}{N}\left(\boldsymbol{L}_{conf}(x,c)+\alpha\boldsymbol{L}_{loc}(x,l,g)\right), \quad (1)$$

where *N* is the appropriate number of default boxes. If *N* = 0, the loss function is 0. The localization loss function is the L_1 loss smooth function [38] between the predicted box (l) and the truth box (g) as

$$L_{loc}(x,l,g)=\sum_{i\in Pos}^{N}\sum_{m\in\{cx,cy,w,h\}}x_{ij}^{k}smooth_{L1}\left(l_i^m-\hat{g}_j^m\right), \quad (2)$$

$$\hat{g}_j^{cx}=\left(g_j^{cx}-d_i^{cx}\right)/d_i^w;\ \hat{g}_j^{cy}=\left(g_j^{cy}-d_i^{cy}\right)/d_i^h,$$
$$\hat{g}_j^w=\log\left(\frac{g_j^w}{d_j^w}\right);\ \hat{g}_j^h=\log\left(\frac{g_j^h}{d_j^h}\right); \quad (3)$$

The reliability loss function is the softmax loss over many layers (c) as

$$L_{conf}(x,c)=-\sum_{i\in Pos}^{N}x_{ij}^{k}\log\left(\hat{c}_i^p\right)-\sum_{i\in Neg}\log\left(\hat{c}_i^0\right), \quad (4)$$

Where $\hat{c}_i^p=\frac{\exp\left(c_i^p\right)}{\sum_p\exp\left(c_i^p\right)}$. (5)

Training Details: For the training of SSD_Mobinetv2, we created a data set with 18000 images collecting from locations where the camera angle was installed according to the recommended system. The aim is to create a properly distributed data set to increase system performance. For each image, we manually labeled the objects of the car. From the selection of transformations to enhance the input, we then train the model using Tensorflow [39]. We trained the network with 300,000 mini-batch loops of 16 size. We used a standard optimization algorithm RMSPropOptimizer with momentum and decay values 0.9 and 0.9, respectively. Besides, using batch normalization technique after each

layer and the standard decay weight is set to 0.0001. The results will be presented in the following section.

3-3- ALPR Module

In order to be able to identify each vehicle based on its license plate, we use an automatic license plate recognition system (ALPR). In the module, we use an ALPR system that operates on a variety of scenarios [27]. One of the main advantages of system is its ability to detect number plates in a variety of contexts that allows a process of aligning the number plate before character recognition. Therefore, the system has the flexibility to detect and identify highly accurate number plates in independent test datasets using the same system parameters. ALPR has two main tasks: finding and identifying number plates in input images. Typically, it is divided into four missions, vehicle detection, number plate detection, character segmentation and character recognition. In [27], they combined the last two missions, OCR. The method implemented consists of three main blocks: vehicle detection, license plate display method, and OCR. With high-resolution LPR camera input image data, the module will first detect the media in the image. For each region, the curved plane object detection network (WPOD-NET) searches for the license plate containers and regress an affine transformation. This allows the license plate area to be edited into a rectangle like the one from the front view. These edited findings are sent to the OCR network for character recognition. The following is the block diagram of the ALPR module implementation as shown in Fig. 5.

Fig. 5 Block diagram for detecting and identifying license plate [27]

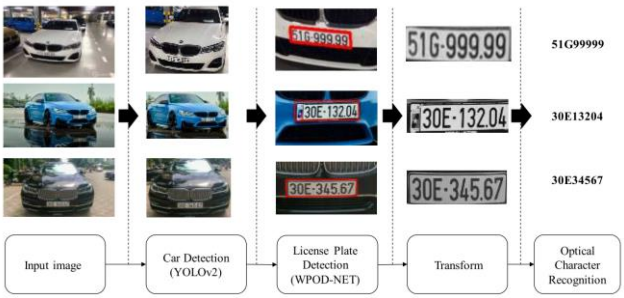


Fig. 6 Demonstration of the module implementation steps

Car Detection: Since vehicles are one of the basic objects that present in many basic detection and identification datasets (PASCAL-VOC [25], ImageNet [26], and COCO [30]), we do not decide to train a detector from scratch. Besides, we choose a trained model to perform vehicle detection by considering several evaluation criteria such as AP, recall, and precision. Based on [27], we use the YOLOv2 network because of its fast execution speed, high precision and recall (76.8% mAP on the PASCAL-VOC dataset) [25]. The detected zones are then resized before inclusion in WPOD-NET to detect possible areas of number plates. As a rule of thumb, large sized input images allow detecting smaller objects. However, the downside is the increase in computation costs. Experiments have shown that if the license plate is taken from a front or front angle, the ratio between the size of the license plate and the vehicle's limit box will be high. However, this ratio tends to be smaller for the case of number plates taken from an oblique angle. Therefore, oblique frames should be changed to a larger size since the number plate can still be detected. More specifically, the coefficient of resizing is calculated as follows [27]:

$$f_{sc} = \frac{1}{\min\{W_v, H_v\}} \min\left\{D_{\min} \frac{\max(W_v, H_v)}{\min(W_v, H_v)}, D_{\max}\right\}, \quad (6)$$

where W_v, H_v are the width and length of the limited box of the vehicle respectively with D_{max} = 608; D_{min}=208.

License plate detection: License plates are rectangular and flat objects that are attached to each vehicle for the purpose of identifying it. In the module, we use the WPOD-NET network [27] to learn how to detect number plates in a variety of contexts and the regression coefficients of affine transformations to convert the curved plate into a rectangle. The detection process using WPOD-NET is illustrated in Fig. 7. Besides, the network input is resized by the vehicle detection unit. The result after being forwarded to the WPOD-NET network is a feature map with 8 channels - encoding the object/no-object probability and affine conversion parameters. To extract the warped license plate, the author first considers a fictional square of fixed size around the center position of a cell (m, n). If the probability of this object exceeds a certain detection threshold, a part of regression parameters will be used to construct an affine matrix that transforms the fictional square into number plate. The WPOD-NET network model architecture consists of a total of 21 convolution layers with 14 residual blocks [31]. All convolution sizes are 3x3. The ReLU trigger function is used on the entire network except for block detection. There are 4 layers of Max pooling of size 2x2 and stride 2 that will reduce the input size by a factor of 16. Finally, the detect block consists of two parallel convolution layers: (i) for calculating probabilities and triggering by the function. softmax, and (ii) for affine matrix regression with no trigger function.

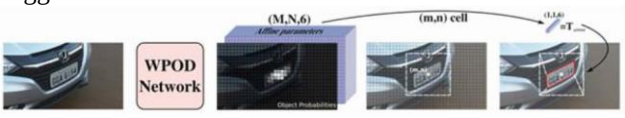


Fig. 7 Detection of license plates using WPOD-NET [27]

The loss function of the lattice is composed of two components [27]: the first part looks at the error between a deforming version of normal square and normalizing points of license plate; the second part handles the probability that there is or not object at position (m, n). Combining the two components above, the loss function is defined as follows:

$$loss = \sum_{m=1}^{M}\sum_{n=1}^{N}\left[\mathrm{I}_{obj}\, f_{affine}(m,n) + f_{affine}(m,n) \right], \quad (7)$$

where

$$f_{affine}(m,n) = \sum_{i=1}^{4} \left\| T_{mn}(\mathbf{q}_i) - A_{mn}(\mathbf{p}_i) \right\|_1 \quad (8)$$

$$f_{probs}(m,n) = \log loss(\mathrm{I}_{obj}, v_1) + \log loss(1-\mathrm{I}_{obj}, v_2) \quad (9)$$

In the context of the system, we propose the solution to implement ALPR module as shown in Figs. 8 and 9.

Fig. 8 Block diagram of the proposed ALPR module implementation

Fig 9: Illustration of ALPR module result

In the system, the image is collected from LPR camera that only contains the area with the license plate of the vehicle. Therefore, we skipped the vehicle detection step before detecting the license plate and adding a preprocessing block to increase the accuracy of character recognition. One of the factors that affect the character recognition block accuracy is the quality of input image. Therefore, we have used homomorphic filtering to minimize the problem. This technique uses a light reflection model. This model considers an image based on two components: 1) illumination on the field being viewed L (x, y) (low frequency); 2) the reflective composition of the objects on the scene R (x, y) (high frequency) is defined as follows:

$$I(x,y) = L(x,y).R(x,y) \quad (10)$$

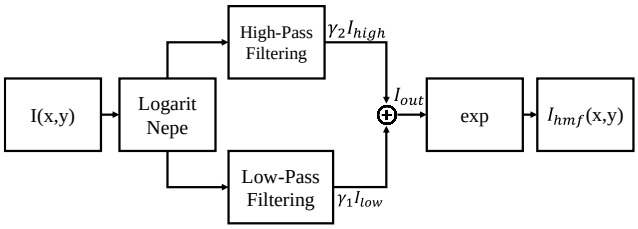


Fig. 10 Block diagram of preprocessing

The input image of length and width are H, W respectively. By creating a Gaussian filter of size (M, N) (where M = 2.H + 1, N = 2.W + 1) to reduce the jagged effect, low pass filters (H_{low}, H_{high}) are defined as follows:

$$H_{low}(x,y) = e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}}, \quad (11)$$

$$H_{high}(x,y) = 1 - H_{low} \quad . \quad (12)$$

After passing two filters, we get I_{low} and I_{high}. To add these two results, we have

$$I_{out} = \gamma_1 I_{low} + \gamma_2 I_{high} \quad (13)$$

where $\gamma_1 = 0.3,\ \gamma_2 = 1.5$.

Finally, the output image is defined as follows:

$$I_{hmf} = e^{I_{out}} \quad (14)$$

Fig. 11 Image before and after pretreatment

4- Experiment and Results

After the analysis and design process, we build a test scenario for the system and evaluate the achieved results to its feasibility. Evaluation method consists of two main stages: one is to evaluate the results achieved on each module, other is to evaluate the system.

4-1- Experiment Scenario

4.1.1 Realistic Background

Parking is located on a one-way lane as shown in Fig. 8.

Fig. 12 Actual parking lot

4.1.2 Circumstances Occurred

We divided into two cases:

- A car enters
- Many vehicles go in sequence.

4-2- Build the Data Set

With the proposal system, Skyeye module retrieves image data from camera mounted on the overhead. We built a realistic reflection dataset to increase the performance of the module. Table 2 shows the results of the collecting data.

Table 2: Statistical table of collected data sets

Place	***Features***	***Number of images***	***Number of objects***
Tan Trieu K Hospital	Taken from camera phone, high angle view.	5400	6292
Tunnel of Thang Long	Taken from camera phone, high angle view	7800	9475
Street	Taken from camera phone, high angle view	3600	4425
Hanoi University of Science and Technology	Taken from camera phone, high angle view	1200	4000
Total		18000	24192

Fig. 13 Illustration of collecting data samples

4-3- Skyeye Module Results

Table 3: The results obtained after performing the Skyeye module

Criteria / **Method**	**AP@ 05 (%)**	**Precision (%)**	**Recall (%)**	**Processing time (second)**	
				416 × 416	960× 540
YOLOv2 [3], [5]	99.95	99.20	99.32	0.533	1.274
Skyeye module	99.48	99.28	99.18	0.273	0.829

Fig. 14 Illustration of a test example of the Skyeye module

After labeling the self-collected data set, we retrained the vehicle detection model. The results are presented in Tab. 3.

The parameters are defined as follows:

- True Positive (TP): A correct detection. Detection with IOU ≥ threshold
- False Positive (FP): A wrong detection. Detection with IOU < threshold
- False Negative (FN): A ground truth not detected
- True Negative (TN): We do not apply. It would represent a corrected misdetection. In the object detection task, there are many possible bounding boxes that should not be detected within an image. TN would be all possible bounding boxes that were corrrectly not detected (many possible boxes within an image). Therefore, it is not used by the metrics

The number of objects is 12226 and threshold is 0.5 in our paper. The parameters are shown in Tab. 4.

Table 4: Parameter setup

	Predicted as Positive	***Predicted as Negative***
Actual: Positive	TP = 12126	FN = 100
Actual: Negative	FP = 88	

From the results achieved, it shows that the Skyeye module has a very high accuracy. Although the model's accuracy is lower than that of YOLOv2 [23], the processing time is only half.

In this article, we focus on research and propose a smart parking management system solution using multiple cameras. Therefore, when a solution is proposed, we accelerate the process of implementing the solution by installing cameras and collecting datasets that share the

actual context. On the other hand, the performance of deep learning application detection and identification models in particular comes from the fact that the quality of data and the amount of data that can cover all the contexts will be decisive for model training. Currently, we collect data in environmental conditions including: sunny, cloudy, rainy, etc. Our system achieved in hitting performance evaluation over the availability of data collected including the aforementioned contexts.

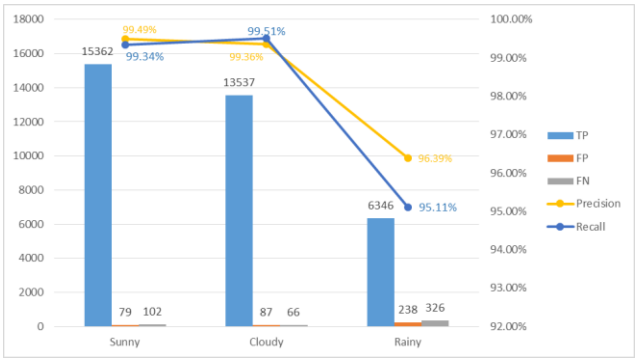


Fig. 15 Result of sunny, cloudy, and rainy environment

Our system still achieves high accuracy through the Precision and Recall evaluation criteria shown in Fig. 15. However, our system has not yet handled all the actual cases. In the future, we are implementing to improve system performance and work well in the following contexts, namely sunny, rainy, afternoon and evening.

4-4- ALPR Module Results

To evaluate the proposed method, we used two test datasets with different characteristics: 1) a car_long tuple and 2) a self-collection dataset called VN_LP, which is presented in Tab. 3.

Character Recognition: To segment and recognize characters on license plate after pre-processing, we use pre-trained YOLO network [27] to evaluate the effectiveness. The results are presented in Tabs. 5 and 6 and Fig. 16.

Table 5: Test data set

Data set	***Characteristic***	***Quantity***
Car long [28]	Number plate images from the front with fixing angle	998 (images)
VN_LP	Images of number plates from many different angles	1000 (images)

The following results are shown in Tab. 6.

Table 6: Test results of ALPR modules

Criteria / ***Method***	***Car long [29]***		***VN_LP***	
	Processing Time (s)	Accuracy	Processing Time (s)	Accuracy
YOLOv2 [29]	0.218	86.34%	0.481	**70%**
Module ALPR	**0.142**	**94%**	0.366	60%

The results show that the result of proposal solution has lower processing time and higher accuracy for our dataset. However, the accuracy is significantly lower for other datasets.

Fig. 16 Illustration of ALPR module test result

4-5- Proposed System Results

To evaluate the system's performance based on the results obtained from the Skyeye and ALPR modules, we give a simple formula that is the average of two-module results. The following results are achieved by proposal system as shown in Tab. 7 and Figs. 17 and 18. The results show that the system has high accuracy for our dataset. Therefore, the system is highly feasible when applying in practice.

Table 7: Results achieved by the proposal system

Method	***Acuracy of Skyeye module (%)***	***Accuracy of ALPR module (%)***	***System***	***Processing Time (second)***	
				416 x416	960 x540
YOLO v2 [23][27]	**99.95**	86.34	89.83	0.533	1.274
Proposal system	94.48	**94**	**96.74**	**0.273**	**0.829**

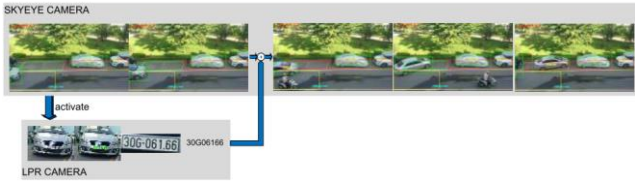


Fig. 17 Image of a vehicle entering a parking space

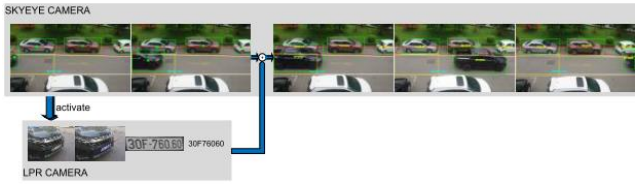


Fig. 18 Image of a vehicle not entering a parking space

5- Conclusion

The image-based proposal system in the article not only detects the location of parking space but also determines the vehicle's identity based on its respective license plate to share status information for user. Therefore, we can control and manage the parking space effectively. In addition to the application of the system to street parking spaces, the system can also be used for other purposes such as detecting illegal parking on the roadway as well as expanding functions. It helps to automatically collect electronic fees based on the license plate of vehicle that is linked to the owner's e-wallet. The system achieves high accuracy of 96.74% in actual setting. It points out the feasibility of the system into practice. However, there are many challenges that affect the performance of the system such as input image quality, detection algorithms, and identification. Therefore, we will try to improve the performance of the system by solving real-world problems in the future.

References

[1] Z. Pala and N. Inanc, "Smart Parking Applications Using RFID Technology," *2007 1st Annual RFID Eurasia*, 2007, pp. 1-3.

[2] J. P. Benson et al., "Car-Park Management using Wireless Sensor Networks," Proceedings. 2006 31st IEEE Conference on Local Computer Networks, 2006, pp. 588-595.

[3] A. Kianpisheh, N. Mustaffa, P. Limtrairut, and P. Keikhosrokiani, "Smart Parking System (SPS) architecture using ultrasonic detector," Int. J. Softw. Eng. its Appl., vol. 6, no. 3, pp. 51–58, 2012.

[4] G. Ostojic, S. Stankovski, M. Lazarevic and V. Jovanovic, "Implementation of RFID Technology in Parking Lot Access Control System," 2007 1st Annual RFID Eurasia, 2007, pp. 1-5.

[5] J. Wolff, T. Heuer, Haibin Gao, M. Weinmann, S. Voit and U. Hartmann, "Parking monitor system based on magnetic field senso," 2006 IEEE Intelligent Transportation Systems Conference, 2006, pp. 1275-1279.

[6] N. H. H. Mohamad Hanif, Mohd Hafiz Badiozaman and H. Daud, "Smart parking reservation system using short message services (SMS)," 2010 International Conference on Intelligent and Advanced Systems, 2010, pp. 1-5.

[7] A. Sayeeraman and P. S. Ramesh, "Zigbee and gsm based secure vehicle parking management and reservation system," J. Theor. Appl. Inf. Technol., vol. 37, no. 2, pp. 199–203, 2012.

[8] H. Ichihashi, T. Katada, M. Fujiyoshi, A. Notsu and K. Honda, "Improvement in the performance of camera based vehicle detector for parking lot," International Conference on Fuzzy Systems, 2010, pp. 1-7.

[9] R. Yusnita, F. Norbaya, and N. Basharuddin. "Intelligent parking space detection system based on image processing," International Journal of Innovation, Management and Technology, vol. 3, no. 3, pp. 232-235, 2012.

[10] S. Banerjee, P. Choudekar and M. K. Muju, "Real time car parking system using image processing," 2011 3rd International Conference on Electronics Computer Technology, 2011, pp. 99-103.

[11] S. Khaled, and H. Tounsi, "Parking space detection system using video images," Transportation Research Record: J. of the Transportation Research Board 2537, no. 1, pp. 137-147, 2015.

[12] S. H. Pratap, O. P. Uniyal, and K. Joshi, "An approach to implement cost efficient space detection technology with lower complexity for smart parking system," Indonesian Journal of Electrical Engineering and Computer Science, vol. 15, no. 3, pp. 415-419, 2015.

[13] S. Saleh Al-Amri, N. V. Kalyankar, and Khamitkar S, "Image segmentation by using thershold techniques," Journal of Computing. vol. 2, no. 5, 2010.

[14] S. Li, H. Dawood and P. Guo, "Comparison of linear dimensionality reduction methods in image annotation," 2015 Seventh International Conference on Advanced Computational Intelligence (ICACI), 2015, pp. 355-360.

[15] R. Mehmood, R. Bie, H. Dawood and H. Ahmad, "Fuzzy clustering by fast search and find of density peaks," 2015 International Conference on Identification, Information, and Knowledge in the Internet of Things (IIKI), 2015, pp. 258-261.

[16] A. R. Pathak, M. Pandey,S. Rautaray, "Application of deep learning for object detection," Proceding Comput. Sci, vol. 132, pp. 1706–1717, 2018.

[17] X. Wang, "Deep learning in object recognition, detection, and segmentation," Trends Signal Process, vol. 8, pp. 217–382, 2016.

[18] M. Shafiee, B. Chywl, F. Li, A. Wong, "Fast YOLO: A fast you only look once system for real-time embedded object detection in video," arXiv:1709.05943, 2017.

[19] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: unified, real-time object detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

[20] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 580-587.

[21] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 2017.

[22] A. Mauri, R. Khemmar, B. Decoux, N. Ragot, R. Rossi, R. Trabelsi, R. Boutteau, R. Ertaud, J. Savatier, "Deep learning for real-time 3D multi-object detection, localisation, and tracking: application to smart mobility," Sensors, vol. 20, no. 2, pp. 1-15, 2020.

[23] J. Sang, Z. Wu, P. Guo P, et al, "An improved YOLOv2 for vehicle detection," Sensors (Basel), vol 18, no. 12, pp. 1-15, 2018.

[24] J. C. Nascimento, A. J. Abrantes and J. S. Marques, "An algorithm for centroid-based tracking of moving objects," 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258), 1999, pp. 3305-3308, vol.6.

[25] M. Everingham, L. V. Gool, C.K.I. Williams, J. Winn, A. Zisserman, "The pascal visual object classes (voc) challenge," International J. of Computer Vision, vol. 88, no. 2, pp. 303–338, 2010.

[26] G. R. Gonalves, S.P.G. Silva,D. Menotti, W.R. Schwartz, "Benchmark for license plate character segmentation," J. of Electronic Imaging, vol. 25, no. 5, pp. 1–5, 2016.

[27] S. M. Silva and C.R. Jung, "License plate detection and recognition in unconstrained scenarios," Lecture Notes in Computer Science, vol 11216, Springer, Cham, pp. 593–609, 2018.

[28] P.N. Huu, L.H. Bao, and H.L. The, "Proposing an image enhancement algorithm using CNN for applications of face recognition system," J. Adv. Math. Comput. Sci., vol. 34, no. 6, pp. 1–14, 2020.

[29] Computer Vision, "Aggregate data", Online: https://thigiacmaytinh.com/tai-nguyen-xu-ly-anh/tong-hop-data-xu-ly-anh/ (accessed Jun. 30, 2020).

[30] T.Y. Lin, et al, "Microsoft COCO: Common objects in context," in ECCV 2014: Computer Vision – ECCV 2014, vol. 8693, pp. 740-755, 2014.

[31] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.

[32] W. Liu, et al., "SSD: Single shot multiBox detector," In ECCV 2016: Computer Vision – ECCV 2016, vol. 9905, pp. 21-37, 2016.

[33] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 6517-6525.

[34] A. G. Howard, et. al, "Mobilenets: efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, pp. 1-9, 2017.

[35] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520.

[36] D. Erhan, C. Szegedy, A. Toshev and D. Anguelov, "Scalable object detection using deep neural networks," 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2155-2162.

[37] C. Szegedy, S. Reed, D. Erhan, D. Anguelov, "Scalable, high-quality object detection," arXiv:1412.1441 v3, pp. 1-10, 2015.

[38] R. Girshick, "Fast R-CNN," 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440-1448.

[39] M. Abadi, et. al, "TensorFlow: a system for large-scale machine learning," In Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation (OSDI'16). USENIX Association, USA, pp. 265–283, 2016.

Phat Nguyen Huu received his B.E. (2003), M.S. (2005) degrees in Electronics and Telecommunications at Hanoi University of Science and Technology (HUST), Vietnam, and Ph.D. degree (2012) in Computer Science at Shibaura Institute of Technology, Japan. Currently, he lecturer at School of Electronics and Telecommunications, HUST Vietnam. His research interests include digital image and video processing, wireless networks, ad hoc and sensor network, and intelligent traffic system (ITS) and internet of thing (IoT). He received the best conference paper award in SoftCOM (2011), best student grant award in APNOMS (2011), hisayoshi yanai honorary award by Shibaura Institute of Technology, Japan in 2012.

Loc Hoang Bao received his B.E. (2020) degrees in Electronics and Telecommunications at Hanoi University of Science and Technology (HUST), Vietnam. Currently, he is the master student at School of Electronics and Telecommunications, HUST, Vietnam. His research interests include digital image, video processing, and embedded system.

Word Sense Induction in Persian and English: A Comparative Study

Masood Ghayoomi
Faculty of Linguistics, Institute for Humanities and Cultural Studies, Tehran, Iran
m.ghayoomi@ihcs.ac.ir

Received: 05/Jan/2021 Revised: 20/Jun/2021 Accepted: 20/Jul/2021

Abstract
Words in the natural language have forms and meanings, and there might not always be a one-to-one match between them. This property of the language causes words to have more than one meaning; as a result, a text processing system faces challenges to determine the precise meaning of the target word in a sentence. Using lexical resources or lexical databases, such as WordNet, might be a help, but due to their manual development, they become outdated by passage of time and language change. Moreover, the lexical resources might be domain dependent which are unusable for open domain natural language processing tasks. These drawbacks are a strong motivation to use unsupervised machine learning approaches to induce word senses from the natural data. To reach the goal, the clustering approach can be utilized such that each cluster resembles a sense. In this paper, we study the performance of a word sense induction model by using three variables: a) the target language: in our experiments, we run the induction process on Persian and English; b) the type of the clustering algorithm: both parametric clustering algorithms, including hierarchical and partitioning, and non-parametric clustering algorithms, including probabilistic and density-based, are utilized to induce senses; c) the context of the target words to capture the information in vectors created for clustering: for the input of the clustering algorithms, the vectors are created either based on the whole sentence in which the target word is located; or based on the limited surrounding words of the target word. We evaluate the clustering performance externally. Moreover, we introduce a normalized, joint evaluation metric to compare the models. The experimental results for both Persian and English test data showed that the window-based partitioningK-means algorithm obtained the best performance.

Keywords*:* Corpus Linguistics; Word Sense Induction; Clustering; Word Embedding; Sense Embedding; Parametric Clustering; Non-parametric Clustering; Joint Evaluation Metric.

1- Introduction

Language, as a means of communication between human beings, is composed of two components [1]: form, and meaning. The 'form' can be represented either via an audio signal transmitted through a voice channel from a speaker to a recipient, or via an orthographic form through the writing system and the alphabetical set of the language. In text processing, the orthographic form of the language is taken into consideration. Ambiguity is a property of a natural language that causes challenges in text processing. There exist two types of ambiguities: a) syntactic ambiguity, and b) lexical ambiguity. The sentence 'I saw the man with a telescope.', for instance, is a sample of syntactic ambiguity to either mean 'I used a telescope to see the man' or 'I saw the man who carried a telescope'.

There are two reasons to cause lexical ambiguity [2, p: 146]: (a) polysemy where a word has more than one meaning, such as /rošan/ (light/bright) in /ran ge rošan/ (light color) and /ʔotāqe rošan/ (bright room) in Persian or 'plane' in 'fly by plane' and 'cut by plane' in English; and (b) homonymy where the word is both homograph and homophone, such as /rox/ (rook/face/roc) in /mohreye rox/ (the rook piece [in chess]), /roxe zibāye ʔu/ (her beautiful face), and /parandeye rox/ (the roc bird) in Persian or 'bank' (financial place/side of river) in English. In Example (1)a-f, the sentences that contain the target word 'bank' are grouped (clustered) in Figure 1. Based on the semantic similarity of the target word 'bank' in the sentences, one group belongs to the concept 'financial place' ($bank_1$) and the other group belongs to the concept 'side of river' ($bank_2$).

(1) a. He cashed a check at the bank.
b. She sat on the bank of the river and watched the currents.
c. They detected frauds in the bank.
d. I saw a deer near the river bank.
e. That bank holds the mortgage on my home.
f. They pulled the canoe up on the bank.

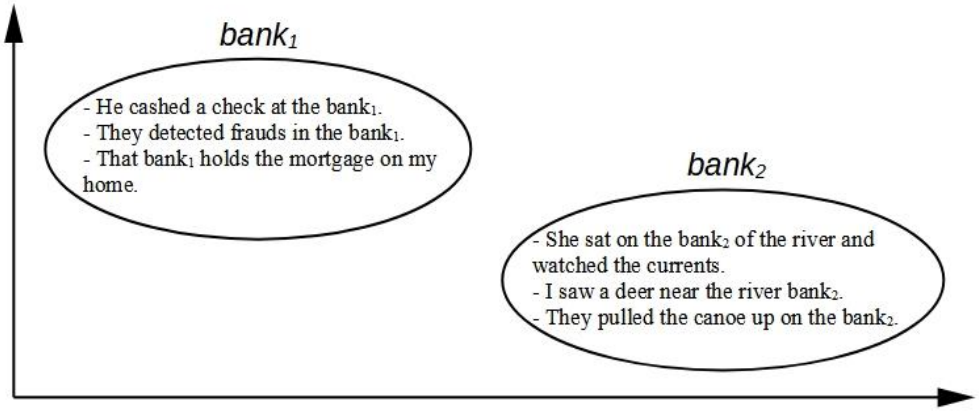


Figure 1: Clustering result of instances for the target word "bank"

The lexical ambiguity in text processing is more pronounced in languages that use the Arabic script in their writing system, such as Persian, due to avoiding writing short vowels than languages that use the phonemic orthography, such as English. In text processing, both polysemy and homonymy are recognized as one problem. The context of the target ambiguous word plays a very important role to determine and to disambiguate the meaning.

The Word Sense Induction (WSI) task means that the machine has to induce word senses from the natural data automatically without prior knowledge. This task uses an unsupervised machine learning approach and it can be defined as a clustering task. The example in Figure 1 represents the idea of how clustering can identify the senses of a word. One property of this task is that no initial training data is required.

This paper focuses on WSI and aims at inducing the meaning of both polysemous and homonymous Persian and English words from their local contexts and comparing the performance of the clustering algorithms. One additional contribution of this paper is introducing a normalized, joint, external evaluation metric to be able to compare the models more accurately against the naïve baselines.

The construction of the paper is as follows: after the introduction, in Section 2, we describe the semantic representation methods to be used for the clustering task. Section 3 reviews the related works on WSI. In Section 4, our models for both Persian and English are proposed. The obtained results as well as our proposed, joint evaluation metric are discussed in Section 5; and finally, the paper is concluded in Section 6.

2- Semantic Representation

2-1- Distributional Semantics

Ambiguity is one of the properties of the natural language. According to the idea proposed by Wittgenstein [3], the meaning of a word can be determined by its usage in the language. Following this idea, Harris [4] proposed an idea in the framework of 'distributional semantics' such that the words which are used in the same local contexts intend to have a similar meaning. Based on this idea, the 'distributional hypothesis' was proposed, and Firth [5] emphasized that "the local context of the word plays an important role in determining words' senses". Miller and Charles [6] proposed 'strong contextual hypothesis' such that two words are to some extent semantically similar if they have similar contexts. Based on this hypothesis, the words 'year', 'date', and 'Wednesday' in Example (2) are semantically similar.

(2) a. I go to the cinema this year.
b. I go to the cinema on this date.
c. I go to the cinema this Wednesday.

Since the context plays a very important role to capture the meaning of a word, precise encoding of the word's context information is required. To this end, Peirsman and Geeraerts [7] introduced three types of linguistic contexts to be extracted from a large corpus: a) document-based model where the words in the same paragraph or in the same documents are used as the context [8, 9]; b) syntax-based model where words are compared according to their syntactic relations, from dependency relations [10, 11, 12, 13] to the combinatory categorial grammar [14]; and c) word-based model where word-word co-occurrence statistics are extracted from a

fixed window size. These word co-occurrences resemble the 'bag-of-words' model [9].

Song et al. [15] introduced two general approaches to represent context information in 'distributional semantics': a) using the Bayesian model utilized in topic modeling [16], and b) using a feature-based model to represent the semantic information as a vector. The latter model uses a vector space model to represent the vectorized semantic information of words. The vectors can be used in the clustering task to induce words' meanings. The advantage of using a vector space model is compressing the information about the words and their contexts, called 'word embedding'. Computing the geometric distance between the vectors makes it possible to decide how two words intend to be similar. Euclidean distance and Cosine distance are two well-known methods for computing the geometric distance between the vectors [17]. However, there are studies that try to better represent the distributional semantics by combining word embeddings with the knowledge-bases known as 'knowledge embedding model' [18], enriching word embeddings with ontologies [19], and utilizing a contextualized knowledge embedding model as a joint model where word embedding and sense embedding (sense representations of the words in the local context from corpora that are sense tagged) are combined with knowledge-bases [20].

2-2- Modeling Methods

To use word embedding methods for capturing the local context information of a word and compressing the information to be represented in a vector, two methods can be utilized: a) using the matrix decomposition techniques, and b) using the neural network-based techniques. The Global Vector (GloVe) representation [21] uses the matrix decomposition technique to provide the distributional representation of words. Continuous Skip gram (Skip-gram) and Continuous Bag Of Words (CBOW) models [22] use the neural network-based technique to represent the contextual information of a word in a vector. In this paper, we use the Skip-gram model for capturing contextual information of the target word in a vector.

2-3- Context Clustering

There are two major clustering algorithms in terms of defining the number of clusters: parametric and non-parametric. The two well-known parametric clustering algorithms commonly used in natural language processing applications are partitioning and hierarchical. Partitioning clustering uses a centroid-based clustering and computes the distance of individual vectors to the centroid, such as the K-means algorithm [23]. The hierarchical clustering uses a statistical criterion to compute the clusters' distance. This algorithm is either agglomerative (bottom-up) or divisive (top-down). We use the divisive clustering algorithm for the WSI task.

The common property of parametric algorithms is that they require a pre-defined number of clusters. Therefore, the State-Of-The-Art (SOTA) techniques in the field have performed their experiments on a pre-defined number of clusters; e.g., the proposed model by used the K-means algorithm with 3 clusters, i.e. $k=3$. To have a better estimation on the number of clusters, Ghayoomi [25] utilized the silhouette score [26] in Equation (1) as a metric to define the number of clusters. Using this method to identify the number of clusters outperformed the SOTA results.

(1)

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

where $a(i)$ is the average dissimilarity of element i with other elements in the same cluster computed by Equation (2); and $b(i)$ is the minimum distance between an element of a cluster with all other elements in the rest of clusters, computed by Equation (3).

(2)

$$a(i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \neq j} d(i, j)$$

(3)

$$b(i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j)$$

where i and j are two elements in cluster C and $d(i,j)$ is the distance between i and j, C_i is the cluster in which element i belongs to and j is another element of this cluster, and C_k is cluster that element i is not its member.

In this research, we use the silhouette score as a metric for each cluster to decide about the best number of clusters: the higher the score, the better the clustering result.

Non-parametric clustering algorithms are another approach for the WSI task. The number of senses (clusters) is unknown in advance and the algorithms should try to find the senses. Chinese Restaurant Processing (CRP) [27] and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [28] are two non-parametric clustering algorithms that we use for this goal. CRP models the behavior of Chinese when they go to a restaurant: either to sit on a table that one has already sat on, or to take a new seat. The algorithm uses a probabilistic Bayesian model. The DBSCAN uses a density-based model to find the best partitioning of clusters.

In this paper, we compare the performance of both parametric and non-parametric algorithms for the WSI task

in Persian and English. The study of the algorithms themselves and their properties are out of the scope of this paper.

3- Related Works on WSI

Clustering the context to distinguish senses of the target polysemous or homonymous word is one of the main approaches in WSI. In this approach, the number of clusters indicates the number of the target word's senses. Huang et al. [24] used the K-means algorithm with word embedding to cluster word contexts. Neelakantan et al. [29] predicted each sense of a word as a context cluster assignment. Their model worked based on the K-means algorithm. In these two researches, a fixed number of clusters, namely 3 clusters, was defined to run the K-means clustering. Li and Jurafsky [30] proposed using CRP as a non-parametric model to capture the senses dynamically. In their approach, the model decided either to generate a new sense for each context or to assign the context to an already generated sense. Wang et al. [31] proposed a model to use weighted topic modeling for sense induction. Amrami and Goldberg [32] utilized the BiML model, a bidirectional recurrent neural network model, proposed by Peters et al. [33] for WSI and extended the model such that predicted word probabilities were used in the language model. Alagić et al. [34] used the lexical substitution model to induce word senses. Therefore, words which belonged to a cluster should be able to be substituted in an appropriate context. The proposed model was compared against manual substitution along with other clustering evaluation metrics. Corrêa and Amancio [35] proposed a model to capture the structural relationship among contexts. To this end, they used the complex network proposed by Perozzi et al. [36] for context embedding. Tallo [37] used sentence embedding for WSI and investigated the encoding of linguistic properties of words in the embedding. Dong and Wang [38] used WSI in the medical domain to enhance sense inventories. They evaluated four models, namely using context clustering, two types of word clustering, and sparse coding in word vector space. Among them, the sparse coding model proposed by Arora et al. [39] outperformed the other models to discover more complete word senses.

As reported by Song et al. [15], the K-means parametric model used by Neelakantan et al. [29] outperforms the CRP algorithm proposed by Li and Jurafsky [30] based on the SemEval2010 WSI task [40]. As Song et al. [15] stated, the main reason for obtaining such results is the poor performance of CRP in making a decision to assign a word to a new cluster. In the results of the two models, the K-means algorithm used 3 clusters as the predefined, fixed number of clusters, while CRP ended to a lesser number of clusters on average than the best average number of clusters for both noun and verb categories in the SemEval2010 WSI task. This indicates that relaxing the predefined number of clusters in K-means can further improve the performance of the task.

4- Architecture of the Proposed Model

The clustering model we proposed in our research is represented in Figure 2. As can be seen in the figure, the model is constructed of three modules and datasets which are described below.

4-1- Major Modules of the Model

The model contains three modules: vectorization, clustering, and evaluation. In vectorization, first the words' vectors based on the big corpus of a language described in Section 4.2 are created. In vectorization of words, three parameters should be taken into consideration in advance: a) the number of dimensions of each vector; b) the number of the surrounding words of the target word in the local context; c) the information to be considered in vectorization which is the word forms in our case. The setting of the parameters is described in Section 5.3. The vector of the instances that contain the target word is created in two modes: a) in the first mode, thereafter called the 'SentContext' mode, the weighted vectors of the words in a sentence are summed up to build the vector of each instance that includes the target word. Then, this score is normalized based on the sentence length. In the second mode, thereafter called the 'WinContext', the limited surrounding context of the target word is used to build the sentence vector.

It has to be mentioned that not all words in a sentence are content words and there exists a closed list of functional words frequently used, such as preposition, conjunctions, coordinators, etc. These words can be considered as stop words. We use a weighting method to increase the impact of content words, and reduce the impact of functional words. To this end, we use TF-IDF[1] [41] to assign a weight to the words.

In the next step, all instances of the target word are clustered based on their vector representation. We assume that each cluster shows one sense of the word. In the clustering module, we utilize both parametric and non-parametric clustering algorithms described in Section 2.3. The parametric algorithms are run based on the two context modes. For clustering, the data should be reformatted from word forms to a vector space model described above. More precise vectors result in better clustering performance.

It should be added that a two step embedding process

[1] Term Frequency-Inverse Document Frequency

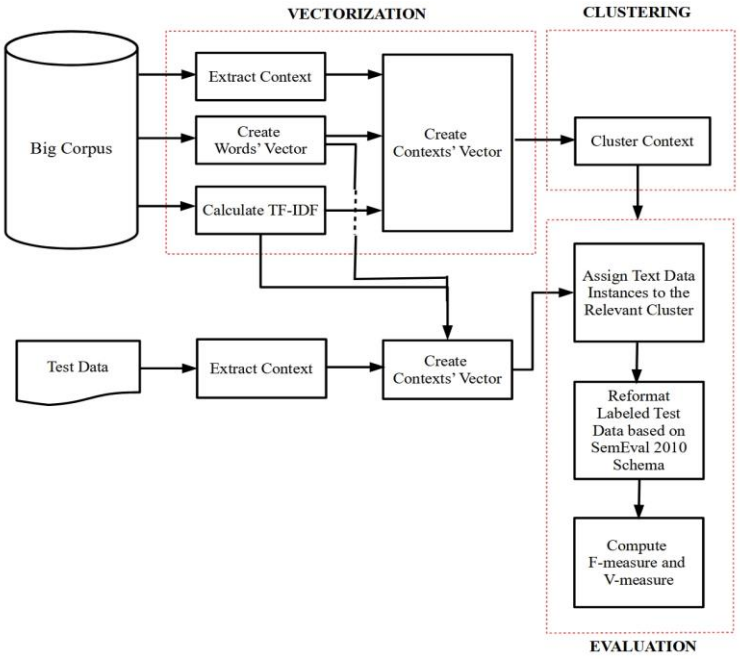


Figure 2: Architecture of our proposed model for WSI

is used in this model. The word embedding is first done based on the semantic distribution of words in a language. And after the clustering step, sense embedding is done for semantic distribution of the target word with respect to its meaning in the local context.

In the evaluation module, two evaluation criteria, namely F-measure and V-measure in addition to a joint metric, are used. These metrics are explained in more detail in Section 5.2. In the evaluation process, the instances of the test data are added to the data pool to be clustered and the induction results of the test data are compared with the corresponding gold standard labels. To this end, we used the toolkit developed in SemEval2010 WSI task [40] that does this mapping.[1]

4-2- Datasets

To run our experiments, we require three datasets for Persian and English: a big corpus, data pool, and test data. The big corpus is used for training word embedding as well as sense embedding to identify the senses of the target words based on the clustering output. The data pool is used for clustering the target words based on their context; and the test data is used for evaluating the models.

The big corpus that we use for creating the Persian words embedding contains over 538 million word tokens developed by Ghayoomi [42]. This corpus is a composition of several other corpora, including a) The Pesian Linguistic DataBase [43] which is a balanced Persian corpus containing both historical and contemporary Persian. In this research, we only use the contemporary dataset; b) The Newspaper Corpus which is a collection of news crawled from the online archive of several Persian newspapers; c) The Hamshahri Corpus [44] which is also another news corpus collected from the online archive of the Hamshahri Newspaper; d) The Bijankhan Corpus [45] which is a fraction of Peykare [46], the Persian Text Corpus; and e) The Persian Wikipedia corpus which contains 361,479 articles downloaded from the dump of Persian Wikipedia articles in July 2016.[2]

The big corpus that we use for creating English word vectors is the Westbury Lab Wikipedia Corpus developed by Shaoul and Westbury [47]. This corpus, which is freely available, is collected from the dump of English Wikipedia articles in April 2010. The corpus contains almost 990 million word tokens of the general domain and it has been used for similar tasks as reported in the literature [24, 29]. It should be mentioned that the documents with less than 2000 characters long are excluded from the corpus.

To evaluate the clustering results of the Persian WSI experiments, we use the test data developed by Ghayoomi [42]. This dataset is standardized based on the SemEval2010 framework. In this dataset, 20 Persian words which are either polysemous or homonymous, are selected from Farsnet [48], the Persian Wordnet. For each target word, 100 sentences are manually annotated; as a result,

[1] https://www.cs.york.ac.uk/semeval2010_WSI/files/evaluation.zip

[2] https://archive.org/details/fawiki-20160720

the test dataset contains 2000 instances in total. Moreover, 279,567 unannotated sentences which contain any of the target words are selected from the big corpus as the data pool.

To evaluate the clustering results of the English WSI experiments, we use the SemEval2010 dataset for the WSI task [40] that is mostly from the news domain. In total, 100 words (50 verbs and 50 nouns) are the target words in this dataset. This dataset contains 8,915 instances as test data with sense annotation and 888,722 unannotated sentences in the data pool. Table 1 summarizes the statistical information of the data pool, the test data, and the size of the big corpus for Persian and English.

Table 1: Statistical information of test and train datasets for Persian and English

Language			Persian	English
Data	Pool	Instance (sentence)	279,567	888,722
	Test	Target Words	20	100
		Instance (sentence)	2,000	8,915
		Average Sense	6.15	5.04
		Average Instance	100	89.15
	Big Corpus for Embedding	Word Token	538 million	990 million

5- Experimental Results

5-1- Baselines

To evaluate the performance of the clustering algorithms, we use two naïve baselines introduced in SemEval2010 [40]: a) the Most Frequent Sense (MFS) where all instances are assigned to a single cluster that contains the most frequent sense; b) one sense per cluster, thereafter called 1S1C, where each instance is assigned to an individual cluster; therefore the number of clusters is equal to the number of instances.

In addition, there are two SOTA results reported in the literature: a) the CRP algorithm utilized by Li and Jurafsky [30] for non-parametric clustering; and b) the K-means algorithm proposed by Neelakantan et al. [29] for parametric clustering. In this K-means algorithm, there is no optimization on the number of clusters and 3 senses are assumed as the pre-defined number of senses for each English word. Thereafter, we call this model K-means-3.

All of the basic baselines and the SOTA models are performed with the Persian data to compare the clustering performance, disregarding the dependency of the algorithm to the data.

5-2- Evaluation Metrics

To evaluate the performance of the clustering results, we utilize two known external evaluation metrics which are commonly used for WSI, namely F-measure [49] and V-measure [50]. In addition, we propose a new normalized, joint evaluation metric, called J-measure, for a fair evaluation of the models.

5.2.1 F-Measure

F-measure computes the accuracy of information retrieval as in Equation (4).

$$F - measure = \frac{(1+\beta) \times P \times R}{(\beta \times P) + R} \tag{4}$$

where P is precision, R is recall, and β is a weighting parameter. If $\beta > 1$, more weight is assigned to recall, and in case $\beta < 1$, more weight is assigned to precision. If $\beta = 1$, precision and recall are considered equally. Equations (5) and (6) compute precision and recall, respectively. In all equations, K is the CLUSTER set, which is the hypothesized clusters from the clustering output and C is the CLASS set, which is the correct partitioning of the data; i.e., for a target dataset with N elements, we have two partitions: the guess partition K, and the gold partition C.

$$P = \frac{n_{ij}}{|k_i|} \tag{5}$$

$$R = \frac{n_{ij}}{|c_i|} \tag{6}$$

where n_{ij} is the number of members of class $ci \in C$ that is the element of cluster $kj \in K$.

5.2.2 V-Measure

Another alternative to evaluate clustering is an entropy-based approach proposed by Rosenberg and Hirschberg [50]. Different entropy-based evaluation metrics have been proposed for clustering so far [51, 52]. Among them, the V-measure metric proposed by Rosenberg and Hirschberg [50] is the most popular one. V-measure computes the harmonic mean of homogeneity, h, and completeness, c, of clustering as stated in Equation (7).

$$V - measure = \frac{(1+\beta) \times h \times c}{(\beta \times h) + c} \tag{7}$$

Homogeneity means that in each CLUSTER, there are a few numbers of CLASSes. The best mode of homogeneity is when a cluster consists of only samples of one class. Completeness, which is the reverse of homogeneity, means that each CLASS is appeared in a few numbers of CLUSTERs. The best mode of completeness is when all samples of the same class are within a single cluster.

As Rosenberg and Hirschberg [50] explained, homogeneity and completeness are formally defined in (8) and (9):

(8)

$$h = \begin{cases} 1 & if \quad H(C,K) = 0 \\ 1 - \frac{H(C|K)}{H(C)} & else \end{cases}$$

where

$$H(C|K) = -\sum_{k=1}^{|K|}\sum_{c=1}^{|C|}\frac{a_{ck}}{N} log \frac{a_{ck}}{\sum_{c=1}^{|C|} a_{ck}}$$

$$H(C) = -\sum_{c=1}^{|C|}\frac{\sum_{k=1}^{|K|} a_{ck}}{N} log \frac{\sum_{k=1}^{|K|} a_{ck}}{N}$$

(9)

$$c = \begin{cases} 1 & \text{if } H(K) = 0 \\ 1 - \frac{H(K|C)}{H(K)} & \text{else} \end{cases}$$

where

$$H(K|C) = -\sum_{k=1}^{|K|}\sum_{c=1}^{|C|}\frac{a_{ck}}{} log \frac{a_{ck}}{}$$

$$H(K) = -\sum_{k=1}^{|K|}\frac{\sum_{c=1}^{|C|} a_{ck}}{N} log \frac{\sum_{c=1}^{|C|} a_{ck}}{N}$$

$C=\{c_i \mid i = 1, \ldots, n\}$ is the set of CLASS, $K = \{k_i \mid 1, \ldots, m\}$ is the set of CLUSTER, and N is the number of data points in the data set, and a_{ck} is the number of elements of class c in cluster k.

5.2.3 The Proposed Evaluation Metric to Evaluate the Clustering Performance

The advantage of V-measure over F-measure is that in the evaluation, completeness as well as homogeneity are taken into consideration, while in F-measure only the distribution of classes in clusters, i.e. homogeneity in the clustering, is considered and it does not care about whether in each cluster the number of classes are minimized. This difference indicates that V-measure is more reliable than F-measure. On the other hand, V-measure alone dedicates a high score to the partitioning with one instance per cluster, because in such partitioning the number of classes in each cluster is perfectly minimized. This indicates that despite the advantages of V-measure, it is not a reliable metric. Therefore, to accurately evaluate the performance of the clustering result, we need to consider both metrics.

The results of the two metrics represent two extremes such that there is a trade-off between them, i.e. in most of the cases if V-measure is high, F-measure is low, and vice versa. For instance, if the SOTA scores based on V- and F-measures are compared against naïve baselines in the WSI task, it can be determined that the naïve baselines, namely 1S1C and MFS, obtain better scores than the advanced SOTA clustering algorithms and the SOTA models are not able to beat the simple baselines. This determines that V- and F-measures in Equations (4) and (7) are not perfect to compare the clustering performance accurately. As a result, we propose a normalized, joint metric, called J-measure in Equation (10) which is the harmonic mean of V- and F-measures. The obtained score is uniformed such that both homogeneity and completeness are included.

(10)

$$J - measure = \frac{(1+\beta) \times F \times V}{(\beta \times F) + V}$$

where F is F-measure and it obtains the result from Equation (4), V is V-measure and it obtains the result from Equation (7), and β is the weighting parameter. If $\beta > 1$, then more weight is assigned to F-measure; therefore only homogeneity in clustering is considered. In case $\beta < 1$, then more weight is assigned to V-measure to consider both homogeneity and completeness. If $\beta = 1$, then there is a uniform distribution over F- and V-measure.

If $\beta = 1$ in Equations (4), (7) and (10), then Equation (10) can be rewritten as Equation (11) to show how precision, recall, homogeneity, and completeness can relate to each other:

(11)

$$\begin{aligned} J - measure &= \frac{2 \times \frac{2PR}{P+R} \times \frac{2HC}{H+C}}{(\frac{2PR}{P+R}) + (\frac{2HC}{H+C})} \\ &= \frac{8PRHC}{(2PRH) + (2PRC) + (2PHC) + (2RHC)} \\ &= \frac{8PRHC}{2(PRH + PRC + PHC + RHC)} \\ &= \frac{4PRHC}{PR(H+C) + HC(P+R)} \end{aligned}$$

Table 2: Results of the baselines, SOTAs, and the experimented models for Persian according to V-measure (V), F-measure (F) and J-measure (J) criteria

	Model	V (%)	H (%)	C (%)	F (%)	P (%)	R (%)	J (%)
Baseline	1S1C	**37.33**	100	23.57	00.07	00.70	00.04	00.13
	MFS	00.07	00.04	96.06	**59.51**	44.65	99.82	00.14
SOTA	CRP: SentContext	12.92	14.30	28.86	50.67	51.45	59.58	20.59
	Kmeans-3: SentContext	26.70	21.91	34.18	51.84	42.62	66.16	**35.25**
Models	DBSCAN: SentContext	02.36	1.36	45.94	**59.26**	44.82	97.83	04.53
	Kmeans-3: WinContext	31.97	30.29	35.61	56.09	54.14	58.18	40.72
	Divisive: SentContext	24.00	21.95	27.39	56.79	55.61	60.23	33.73
	Divisive: WinContext	29.63	26.49	36.41	59.94	56.21	66.56	39.43
	Kmeans- silhouette: SentContext	26.20	22.08	32.23	42.56	33.56	58.17	32.43
	Kmeans- silhouette: WinContext	**34.61**	37.95	34.64	51.95	61.73	50.27	**41.54**

Table 3: Results of the baselines, SOTAs, and the experimented models for English according to V-measure (V), F-measure (F) and J-measure (J) criteria

	Model	V (%)	H (%)	C (%)	F (%)	P (%)	R (%)	J (%)
Baseline	1S1C	**31.70**	23.51	48.66	00.09	00.05	00.50	**17.95**
	MFS	00.00	00.00	00.00	**63.40**	47.55	95.23	00.00
SOTA	CRP: SentContext	05.70	03.13	31.88	55.30	42.41	79.46	10.35
	Kmeans-3: SentContext	09.80	05.37	56.18	55.10	41.94	80.31	16.64
Models	DBSCAN: SentContext	04.66	03.20	44.78	**61.20**	46.84	93.99	08.66
	Kmeans-3: WinContext	15.63	16.22	18.12	49.30	49.96	52.81	23.74
	Divisive: SentContext	14.87	14.13	21.12	37.42	36.71	43.81	21.28
	Divisive: WinContext	16.15	12.57	43.51	53.77	48.69	65.61	24.84
	Kmeans- silhouette: SentContext	18.36	20.77	17.91	47.75	51.94	47.65	26.52
	Kmeans- silhouette: WinContext	**19.74**	23.90	18.09	43.03	52.00	39.21	**27.06**

5-3- Setup of Experiments

In this study, we experimentally compare the performance of several clustering algorithms to induce Persian and English word senses. The clustering algorithms require vector representation of the data. To this end, the Gensim Python[1] library is utilized to create the words' vectors according to this setups: a) employing the skip-gram model to capture the context of words; b) setting 8 words (4 words before and 4 words after the target word) similar to Huang et al. [24] to extract the information of the words' local contexts; c) setting the vector size to 300 dimensions similar to Neelakantan et al. [29]; and d) using the words with frequency 5 and above to build words' vector. In the next step, the weighted average of words' vector is created from the context vectors. Then, we compute TF-IDF of each word based on the idea proposed by Neelakantan et al. [29] and use it as a weighting value in each vector to compute the context vector.

The partitioning and hierarchy-based clustering algorithms are run in two modes, SentContext and WinContext modes, described in Section 4.1. In the WinContext mode, the context is set to 8 words to be similar to the context to build the words' vector. As a result, we perform our experiments by considering 4 words before and 4 words after the target word.

We also compute the two-tailed *t*-test to compare the performance of the models and study how statistically significant the difference between the models is.

5-4- Results and Discussion

Tables 2 and 3 summarize the obtained results of using various algorithms for inducing Persian and English words' senses. Among the basic baselines, the 1S1C has obtained a higher score for V-measure than the MSF baseline, but the score of F-measure is the lowest. The obtained results for the MFS baseline are vice-versa. Although the 1S1C baseline considers homogeneity and completeness properties, the MFS baseline takes only homogeneity into consideration.

[1] https://radimrehurek.com/gensim/index.html

Among the two clustering approaches used for the SOTA models, the parametric clustering algorithm implemented in the Kmeans-3 model obtained a higher result than the CRP model based on the J-measure criterion. The difference between the models based on the J-measure was statistically significant ($p < 0.05$). It has to be mentioned that the F-measure results for both models are almost the same. This showed that in terms of homogeneity, the models behaved the same; but considering the completeness property, the advantage of the Kmeans-3 model over the CRP model was highlighted.

In addition to the SOTA techniques, we utilized different parametric and non-parametric methods in our study. We utilized DBSCAN model, as a non-parametric algorithm, for inducing word senses. The model could not beat the CRP model as a baseline according to the J-measure results for both Persian and English. We further observed that the performance of the DBSCAN model was very similar to the MFS baseline since it had a high score for F-measure which means that this clustering algorithm ends up to one single cluster in most of the cases and only homogeneity was taken into consideration.

As mentioned, we used two modes in our experiments, SentContext and WinContext. To have a fair comparison between the modes, we ran the WinContext mode based on the Kmeans-3 model for both Persian and English to be able to compare the results with the SentContext mode of Kmeans-3 as one of the SOTA models.

According to the results, the WinContext mode of the Kmeans-3 model for both Persian and English had beaten the Kmeans-3 model in SentContext mode based on V-measure. The difference between the modes of the Kmeans-3 model was statistically significant ($p < 0.01$). The superiority of the Kmeans-3 model in WinContect mode was reflected in the J-measure. This result indicated that the surrounding words of the target word in the local context have a major impact on determining the meaning of the target word, and all of the words in the sentence are not effective. Comparing the results based on F-measure, the WinContext mode obtained a higher result than the SentContext for Persian; however, SentContext achieved better F-measure than the WinContext for English.

Comparing the proposed models of parametric clustering in either Win- or SentContext mode with the baselines indicated that none of the models had been able to beat the two naïve baselines: the 1S1C baseline based on V-measure, and the MFS baseline based on F-measure. Therefore, it was not possible to compare and to rank the models fairly. J-measure, however, filled the gap. According to the results of the proposed evaluation metric, i.e. J-measure, the proposed WSI models outperformed the naïve baselines. The score of the joint metric has made it possible to compare the proposed models with the SOTA models as well.

We further utilized two parametric algorithms to induce word senses. First, we used the divisive algorithm in SentContext and WinContext modes for both Persian and English. According to the V-measure results, the divisive algorithm had beaten the MFS baseline as well as CRF.

As can be seen, the WinContext mode of the divisive algorithm for both Persian and English obtained a higher result than the SentContext mode. This showed that the divisive algorithm required a narrow context to determine the meaning of words. The differences between the modes were statistically significant ($p < 0.05$). It had to be mentioned that neither of the modes of the divisive clustering algorithm for Persian were able to beat the respective mode of the Kmeans-3 model according to J-measure. While WinContext mode of the divisive clustering for English dataset had been able to beat the respective mode of the Kmeans-3 model based on V-measure which was also reflected in J-measure.

In addition to the divisive algorithm, we used the K-means algorithm enhanced with the silhouette score, thereafter called Kmeans-silhouette, for finding the best number of clusters in the two modes for both Persian and English. According to the results of V-measure, the WinContext mode of this algorithm for both datasets had beaten the SentContext mode. The difference between the modes of this clustering algorithm for the Persian data was statistically significant ($p < 0.05$) but not for the English dataset. Comparing this clustering algorithm to the Kmeans-3, as the SOTA baseline, it had to be mentioned that the WinContext mode of Kmeans-silhouette model for both datasets was able to beat the respective mode of the Kmeans-3 model according to V-measure with statistically significant difference ($p < 0.05$). This shows that the surrounding words in the local context are important for K-means clustering to induce word senses. The SentContext mode of the English data had beaten the SentContext mode of the Kmeans-3 model with statistically significant difference ($p < 0.05$), but not the Persian data, where a slightly poor performance of the SentContext mode was obtained. Comparing the Kmeans-silhouette model to the divisive algorithm for both modes of the two languages, the Kmeans-silhouette model had beaten the divisive clustering algorithm with statistically significant difference ($p < 0.05$).

We further ranked the models and found the best model which has a reasonable good performance based on J-measure. In general, the WinContext mode of the Kmeans-silhouette model for both Persian and English performed the best. This determined that the surrounding words in the context play a significant role in determining the meaning of the word and all of the words in the sentence do not play a major role. This achievement results in reduction of the computation time to produce words' vectors and perform clustering.

6- Conclusion

In this paper, we studied the performance of various clustering algorithms, from parametric to non-parametric, to induce words' senses automatically. The algorithms were run by using Persian or English datasets. Furthermore, two modes, WinContext or SentContext, were used to build words' vectors. Finally, we utilized two evaluation criteria, namely V- and F-measure. There is always a trade-off between these metrics and a model evaluated with these metrics cannot beat a naïve baseline. Therefore, we contributed to propose J-measure as a harmonic mean of V- and F-measure to ease comparing the models. The results were compared with two basic baselines, 1S1C and MFS, and two SOTA models, CRP and Kmeans-3. By comparing the experimental results, we concluded that the parametric clustering algorithm performs better than the non-parametric clustering algorithm for inducing word senses. Among the parametric clustering algorithms, the Kmeans-silhouette clustering model in WinContext performed the best to induce senses of both Persian and English words. This result indicated that the surrounding words of the local context are highly effective in determining the meaning of words than other words in the sentence.

Devlin et al. [53] proposed a model for language representation known as the Bidirectional Encoder Representations from Transformers (BERT) model. This model is currently the SOTA model. One direction of this study as the future work is using the BERT embedding model for the WSI task and comparing the results with the Word2Vec-based embedding model.

References

[1] F. de Saussure, *Cours de linguistique générale*, C. Bally, A. Sechehaye, and A. Riedlinger, Eds. Lausanne, Paris: Payot, 1916.

[2] J. Lyons, *Language and Linguistics: An Introduction*. Cambridge, UK: Cambridge University Press, 1981.

[3] L. Wittgenstein, *Philosophical Investigations*. Oxford, UK: Blackwell Publishing Ltd, 1953.

[4] Z. S. Harris, "Distributional structure," *Word*, vol. 23, no. 10, pp. 146–162, 1954.

[5] J. R. Firth, "A synopsis of linguistic theory 1930-1955," *Studies in Linguistic Analysis* (special volume of the Philological Society), pp. 1–32, 1957.

[6] G. A. Miller and W. G. Charles, "Contextual correlates of semantic similarity," *Language and Cognitive Processes*, vol. 6, no. 1, pp. 1–28, 1991.

[7] Y. Peirsman and D. Geeraerts, "Predicting strong associations on the basis of corpus data," in *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2009, pp. 648–656.

[8] T. K. Landauer and S. T. Dumais, "A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge," *Psychological Review*, vol. 104, no. 2, pp. 211–240, 1997.

[9] M. Sahlgren, *The Word-space Model: Using Distributional Analysis to Represent Syntagmatic and Paradigmatic Relations Between Words in High-dimensional Vector Spaces*. Ph.D. dissertation, Stockholm University, Stockholm, Sweden, 2006.

[10] Z. S. Harris, *A Theory of Language and Information: A Mathematical Approach*. Oxford, England: Oxford University Press, 1991.

[11] D. Lin, "Automatic retrieval and clustering of similar words," in *Proceedings of the 17th international conference on Computational linguistics*. Morristown, NJ, USA: Association for Computational Linguistics, 1998, pp. 768–774.

[12] S. Padó and M. Lapata, "Dependency-based construction of semantic space models," *Computational Linguistics*, vol. 33, no. 2, pp. 161–199, June 2007.

[13] O. Levy and Y. Goldberg, "Dependency-based word embeddings," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, vol. 2. Baltimore, Maryland: Association for Computational Linguistics, June 2014, pp. 302–308.

[14] K. M. Hermann and P. Blunsom, "The role of syntax in vector space models of compositional semantics," in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, vol. 1, Sofia, Bulgaria, 2013, pp. 894–904.

[15] L. Song, Z. Wang, H. Mi, and D. Gildea, "Sense embedding learning for word sense induction," in *Proceedings of the 5th Joint Conference on Lexical and Computational Semantics. The *SEM 2016 Organizing Committee*, 2016, pp. 85–90.

[16] D. M. Blei, A. Ng, and M. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[17] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 2020, https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf.

[18] S. K. Jauhar, C. Dyer, and E. Hovy, "Ontologically grounded multi-sense representation learning for semantic vector space models," in *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Denver, Colorado: Association for Computational Linguistics, May 2015, pp. 683–693.

[19] S. Rothe and H. Schütze, "AutoExtend: Extending word embeddings to embeddings for synsets and lexemes," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, vol. 1. Beijing, China: Association for Computational Linguistics, July 2015, pp. 1793–1803.

[20] S. Ramprasad and J. Maddox, "CoKE: Word sense induction using contextualized knowledge embeddings," in *Proceedings of the AAAI 2019 Spring Symposium on Combining Machine Learning with Knowledge Engineering*, 2019.

[21] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for word representation," in *Proceedings of*

the 2014 Conference on Empirical Methods in Natural Language Processing, vol. 14, 2014, pp. 1532–1543.

[22] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proceedings of the 26th International Conference on Neural Information Processing Systems*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 3111–3119.

[23] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1. Berkeley, California: University of California Press, 1967, pp. 281–297.

[24] E. Huang, R. Socher, C. D. Manning, and A. Ng, "Improving word representations via global context and multiple word prototypes," in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, vol. 1. Jeju Island, Korea: Association for Computational Linguistics, July 2012, pp. 873–882.

[25] M. Ghayoomi, "Density measure in context clustering for distributional semantics of word sense induction," *Journal of Information Systems and Telecommunication*, vol. 8, no. 1, pp. 15–24, 2020.

[26] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, no. 1, pp. 53–65, November 1987.

[27] D. M. Blei, M. I. Jordan, T. L. Griffiths, and J. B. Tenenbaum, "Hierarchical topic models and the nested Chinese restaurant process," in *Proceedings of the 16th International Conference on Neural Information Processing Systems*. MIT Press, 2003, pp. 17–24.

[28] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, E. Simoudis, J. Han, and U. M. Fayyad, Eds. AAAI Press, 1996, pp. 226–231.

[29] A. Neelakantan, J. Shankar, A. Passos, and A. McCallum, "Efficient nonparametric estimation of multiple embeddings per word in vector space," in *Processing of the Conference on Empirical Methods in Natural Language*. Doha, Qatar: Association for Computational Linguistics, 2014.

[30] J. Li and D. Jurafsky, "Do multi-sense embeddings improve natural language understanding?" in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2015, pp. 1722–1732.

[31] J. Wang, M. Bansal, K. Gimpel, B. D. Ziebart, and C. T. Yu, "A sensetopic model for word sense induction with unsupervised data enrichment," *Transactions of the Association for Computational Linguistics*, vol. 3, pp. 59–71, 2015.

[32] A. Amrami and Y. Goldberg, "Word sense induction with neural biLM and symmetric patterns," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics, 2018, pp. 4860–4867.

[33] M. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1. New Orleans, Louisiana: Association for Computational Linguistics, 2018, pp. 2227–2237.

[34] D. Alagić, J. Šnajder, and S. Padó, "Leveraging lexical substitutes for unsupervised word sense induction," in *Proceedings of the 32nd Conference of the Association for the Advancement of Artificial Intelligence*. New Orleans, LA, 2018.

[35] E. A. Corrêa and D. R. Amancio, "Word sense induction using word embeddings and community detection in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 523, pp. 180–190, 2019.

[36] B. Perozzi, R. Al-Rfou', V. Kulkarni, and S. Skiena, "Inducing language networks from continuous space word representations," in *Complex Networks*, P. Contucci, R. Menezes, A. Omicini, and J. Poncela-Casasnovas, Eds. Cham: Springer International Publishing, 2014, pp. 261–273.

[37] P. T. Tallo, *Using Sentence Embeddings for Word Sense Induction*. Master's thesis, Electrical Engineering and Computer Science, University of Cincinnati, Ohio, USA, 2020.

[38] Q. Dong and Y. Wang, "Enhancing medical word sense inventories using word sense induction: A preliminary study," in *Proceedings of the 6th International Workshop on Data Management and Analytics for Medicine and Healthcare, in conjunction with the 46th International Conference on Very Large Data Bases*, 2020, pp. 151–167.

[39] S. Arora, Y. Li, Y. Liang, T. Ma, and A. Risteski, "Linear algebraic structure of word senses, with applications to polysemy," *Transactions of the Association for Computational Linguistics*, vol. 6, pp. 483–495, 2018.

[40] S. Manandhar, I. P. Klapaftis, D. Dligach, and S. S. Pradhan, "Semeval-2010 task 14: Word sense induction & disambiguation," in *Proceedings of the 5th International Workshop on Semantic Evaluation*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 63–68.

[41] G. M. Salton, A. Wong, and C.-S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, November 1975.

[42] M. Ghayoomi, "Finding the meaning of Persian words automatically using word embedding," *Iranian Journal of Information Processing & Management*, vol. 35, no. 1, pp. 25–50, 2019.

[43] S. Assi, "Farsi linguistic database (FLDB)," *International Journal of Lexicography*, vol. 10, no. 3, p. 5, 1997.

[44] A. AleAhmad, H. Amiri, E. Darrudi, M. Rahgozar, and F. Oroumchian, "Hamshahri: A standard Persian text collection," *Knowledge-Based Systems*, vol. 22, no. 5, pp. 382–387, 2009.

[45] M. Bijankhan, "naqše peykarehāye zabāni dar neveštane dasture zabān: mo'arrefiye yek narmafzāre rāyāneyi ["The role of corpora in writing a grammar: Introducing a software"]," *Journal of Linguistics*, vol. 19, no. 2, pp. 48–67, 2004.

[46] M. Bijankhan, J. Sheykhzadegan, M. Bahrani, and M. Ghayoomi, "Lessons from building a Persian written corpus: Peykare," *Language Resources and Evaluation*, vol. 45, no. 2, pp. 143–164, 2011.

[47] C. Shaoul and C. Westbury, "The Westbury Lab Wikipedia Corpus," 2010,

http://www.psych.ualberta.ca/~westburylab/downloads/westburylab.wikicorp.download.html.

[48] M. Shamsfard, H. S. Jafari, and M. Ilbeygi, "STeP-1: A set of fundamental tools for Persian text processing," in *Proceedings of the 7th International Conference on Language Resources and Evaluation*, N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, and D. Tapias, Eds. Valletta, Malta: European Language Resources Association (ELRA), May 19–21 2010, pp. 859–865.

[49] C. J. V. Rijsbergen, *Information Retrieval*, 2nd ed. Newton, MA, USA: Butterworth-Heinemann, 1979.

[50] A. Rosenberg and J. Hirschberg, "V-measure: A conditional entropy-based external cluster evaluation measure," in *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Prague, Czech Republic: Association for Computational Linguistics, June 2007, pp. 410–420.

[51] B. E. Dom, *An Information-theoretic External Cluster-validity Measure*. IBM, Tech. Rep., 2001.

[52] M. Meilă, "Comparing clusterings – an information based distance," *Journal of Multivariate Analysis*, vol. 98, no. 5, pp. 873–895, May 2007.

[53] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, Minnesota: Association for Computational Linguistics*, 2019, pp. 4171–4186.

Masood Ghayoomi received his PhD degree in Computational Linguistics from Berlin Freie University, Berlin, Germany in 2014, and M.S. degree in Computational Linguistics from Nancy2 University, Nancy, France and Saarland University, Saarbrücken, Germany, in 2009. Currently he is a faculty member at the Institute for Humanities and Cultural Studies. His research interests include Computational Linguistics, Natural Language Processing, Machine Learning, Corpus Linguistics, Syntax and Lexical Semantics.

Digital Transformation Model, Based on Grounded Theory

Mohammadali Mirfallah Lialestani
Department of Technology Management, Roodehen Branch
Islamic Azad University, Roodehen, Iran
Ali.mirfallah@gmail.com
Abbas Khamseh*
Department of Industrial Management, Karaj branch
Islamic Azad University, Karaj, Iran
abbas.khamseh@kiau.ac.ir
Reza Radfar
Department of Industrial Management, Science and Research Branch
Islamic Azad University, Tehran, Iran
radfar@gmail.com

Received: 26/ Jul /2020 Revised: 25/Jul/2021 Accepted: 25/Aug/2021

Abstract
Given the emergence of Digital Transformation from Industry 4.0 and the rapid dissemination of technological innovations as well as their impact as a strong driving force in new businesses, efforts should be made to identify the dimensions of this core factor as rapidly as possible. Providing a comprehensive overview of all aspects of the model .The purpose of this article is to provide insights into the state of the art of digital transformation in the last years and suggest ways for future research. This analysis is like a mapping of the subject literature into categories, so that with the help of a number of experts the evolutionary trends can be identified and further researched .In this way, with a deeper understanding of the subject, we have attempted to identify existing gaps . The findings suggest that organizations of all sizes must adapt their business strategy to the realities of digital transformation. This will largely lead to changing business processes as well as managing operations in a new and more intelligent tool-based way. Based on this research, organizations will evolve not just on their own, but on the whole value chain, and this will clearly change the way they produce and deliver value. Organizations can develop their digital ecosystem by creating and developing innovation centers and using open innovation strategy, and as a result, link their digital business to a value chain. Also in this article, we have identified the main categories and subcategories by examining the sources and using the grounded theory approach, as well as determining the relationship between them. Finally, we completed the work by identifying the digital transformation model as the central phenomenon of research.

Keywords: Digital Transformation; Business Strategy; Organization Change; Technology Management.

1- Introduction

The main research question started from this issue: What is the model of digital transformation for organizations, based on what factors and how? Therefore, while identifying indicators and components, we created a model to show the interaction of influential components in the process of digital transformation for organizations. Today, the world is realizing a big change, a change centered on evolving innovations and technologies that will have a profound impact on people's lives, the structure and thinking of organizations, and even the interactions of countries. Understanding this change and timely accompanying this change seems to be essential for all organizations of all sizes. The advent of new and powerful digital technologies, digital platforms and digital infrastructure has transformed innovation and entrepreneurship in significant ways. Beyond opening up new opportunities for innovators and entrepreneurs, digital technologies have wider implications for value creation and development [1]. But digital transformation (DT) is a process by which people can adapt themselves to modern technology. As digital technology becomes more prevalent (automation, cameras, sensors, touch screens, artificial intelligence, etc.), there will be more pressure on companies to make more profit [2]. Digital transformation in the field of industry (also known as industry 4.0 or smart manufacturing) at both the professional and academic levels has increased interest in and interest in production, but is still in its infancy and in-depth research Needs more and more. Even with the current and potential benefits of digital production that are very significant, in terms of improved productivity, sustainability, customization and flexibility, only a limited number of companies have been able to adopt interim strategies to achieve superior performance and utilize this and formulate their position [3].

* Corresponding Author

We know that research to understand digital transformation requires combining multiple themes and cross-level analysis, embracing ideas and concepts from multiple contexts and disciplines [1]. The main purpose of this study is to identify the dimensions of digital transformation and to examine the different domains of its impact and ultimately to prepare a Grounded theory model to provide an appropriate analysis of the digital transformation process for future studies. The innovation of this article in preparing a comprehensive model has included all dimensions for digital transformation in which 9 categories, 24 sub-categories and 152 open coding have been used to determine the more accurate relationship between the components of the model. In the past, a model with these dimensions was not provided.

2- Background Research

We know that the disruption caused by digital conditions is transforming all industries, leading to new business models based on new technologies. Industry 4.0 is therefore a model for understanding and using digital disruption and interference. Industry 4.0 offers include vertical and horizontal integration of value chain, digital services, digital transformation of products, digital transformation of production equipment, digital transformation of factories, and digital transformation of supply chains [4]. The "Industry 4.0" factories have machines that are powered by wireless connectivity and sensors, connected to a smart system that can visualize and decide on the entire production line. In essence, industry 4.0 is a trend towards automation and data exchange in technologies and production processes that include physical and cyber systems, IoT, IIoT, cloud computing, cognitive computing and artificial intelligence [5]. Systematic risks are also essential to be able to execute Industry 4.0 comprehensively. The economic, environmental and social risks posed by the implementation of Industry 4.0 in the field of SMEs need to be addressed. In addition, the technical, IT and political or legal risks arising from the concept are very important [6].

Digital transformation includes new elements that deserve sufficient attention and interesting challenges for future research. Specifically, when managers need to adapt their business strategy to digital realities by integrating new technologies into their business models [7]. The understanding of digital transformation can be distinguished by focusing on the following: technology first, strategy second, technology and strategy third. Organizations in the third group will understand the importance of integrating the different aspects of digital transformation, so that new technology is implemented and associated organizational change is complemented by a separate digital strategy through transformation strategy [8]. On the other hand, the digital transformation and innovation of the resulting business model fundamentally alters consumer expectations and behaviors and disrupts the performance of many small markets [9]. In addition, the intellectual property system, along with the immense potentials available as digital tools, includes data and knowledge on complementary capital, labor, natural resources and other processes that are constantly changing, leading to restructuring of supply chains [10].

In the business models sector, digital transformation is an ongoing process of applying new digital technologies to the everyday life of an organization that is agile as a key mechanism for strategic reshaping of business models in parts of (1) the organization, (2) a collaborative approach, and finally (3) will bring culture [11]. In recent years, companies in almost all industries have undertaken a number of initiatives to discover the benefits of digital and to use technologies and exploit their benefits. These changes have often affected the evolution of key business functions, products, processes, and organizational changes. Corporate governance structures and concepts need new management approaches to manage these collections. Transformations are an important approach to formulating digital transformation strategies. The potential benefits of digitization are enormous and include increased sales or productivity, innovation in value creation and other forms of customer interaction. As a result of this change, entire business models can be transformed or replaced [12]. Digital transformation is described as a new business model or as a transformation. A digital transformation project involves implementing digital capabilities to support the business model. The changes resulting from this change affect the entire organization, especially the operational processes, resources, and internal operations, which are in fact coordinated and based on the intensive cooperation and interaction of external users in optimizing work habits and methods [13]. And in order to digitize the product, you need to analyze the size of the company or its core size. Similarly, it is interesting that the patterns of B2C companies are quite different from those of B2B models for success [14]. The four main tasks that the industry faces are: enhancing customer experience, improving business processes, delivering new products, and preparing to compete with other industries [15]. Digital transmission is accelerating by two key drivers in the industry: cloud services and resource virtualization, as key building blocks in cyber systems that integrate IT-OT principles, models, platforms and integration requirements Digitally produced, they focus on the concept of Industry 4.0, with a focus on the "future industry" [16].

The institutional perspective is also a very applicable lens for the study of digital innovation and transformation. In digital transformation, we mean the effects of a combination of several digital innovations that make new actors, structures, practices, values and beliefs that change,

replace or complement the existing rules of play in organizations and contexts. From this perspective, three types of new formalities for digital transformation have been identified that include: digital organizational forms, digital institutional infrastructures, and digital building blocks [17]. Despite the increasing importance of digital transformation and the concept of malicious innovation, strategy literature still lacks a more complete picture of how organizations committed to their business models after this disruption [18]. However, scientific research and innovation management practices have emphasized the important role of individual competencies in addressing the challenges of digital transformation. However, this sector still lacks sufficient empirical studies, and preliminary results show that individuals' high growth in cognitive and metacognitive skills enhances a company's digital transformation processes. But surprisingly, social competences have only a small effect [19].

Organizations are using digital technologies to change the paths of value creation and change the conditions of competition. To this end, they must implement structural change and overcome the obstacles that hinder their efforts to evolve. Because they have more information with digital technologies, their computing, communication and connectivity creates new forms of collaboration between actors and distributed networks. In doing so, they also create dependencies among actors whose interests may not be fully aligned. Digital technologies influence the strategic implications of digital transformation and the dynamic interaction that occurs between companies and their environments [20]. Digital transformation can also come from the integration of end-to-end systems that are separate from one another in the traditional value chain or in the digital ecosystem of the larger IT industry and through IoT. Digital transformation will potentially affect the workforce. It is challenging to build the skill and pay gap between skilled digital workers and workers in their more traditional role in industry, especially in emerging economies [21]. Changing the industrial production paradigm, digitizing business processes, reconfigures every aspect of the organizational and operational activities throughout the value chain, and manufacturing companies need to take a systematic approach by mapping digital roadmaps to address job opportunities throughout the value chain [3]. However, the application of technology alone is not enough, but to gain the benefits of digitalization, it requires business model innovation such as making changes to advanced jobs and service models. Specifically, the challenges of creating value, delivering value and the components of a business innovation value capture model need to be understood more fully as well as how to align these components in creating a sustainable industry [22]. Successful implementation of innovation practices that ensure effective value creation throughout the supply chain includes: (1) changing the mind and developing an innovation strategy and communicating it to all members of the supply chain; (2) seeing evolution as a long-term process and evolutionary innovation as a cycle, which will be implemented after several tests (3). Functional and inter-organizational [23]. Process improvement is the most added value in the Business Process Management (BPM) cycle. With the mature knowledge, many approaches have failed due to lack of guidance on how to make the process better. Given the diversity of emerging digital technologies, organizations are not only facing the black box of process improvement, but there is also uncertainty about digital technologies [24]. Many customers are also facing a digital transformation process that will lead to partial or fully virtual processes, compliant enterprise structures and digital business models. Virtualization promises innovative opportunities for optimum performance and service delivery, thereby strengthening the competitive position [25]. Digital transformation calls for a redefinition of economics, labor, and democracy for humanity. AI-based devices may take over key areas of human work, reorganize supply chains, induce platform economics, and alter the participation of economic actors in the value chain. Digital transformation defines knowledge and data as the main variables of economic, capital, labor and natural resources. Digital data and technologies will produce a major capital and fuel in the subsequent profitability process, and traditional democratic processes can be (intentionally or unintentionally) replaced by digital technologies [10].

Digitalization promises to change tax management faster than the tax law itself will change. These changes include systems analysis, big data, and ongoing process automation. Although digital transformation will be challenging for taxpayers, the benefits are also significant. In this way, examining legal, ethical issues is important together. With the development of big data technology, automation, artificial intelligence, security and blockchain, all of these changes will affect tax management [26]. In reporting from traditional business to government (B2G), can use the conceptual lens of the institutional function of examining how traditional business reporting to government and how digital reporting is to replace it and try to reduce it [27]. The very dynamic success of the global expansion of digital multinationals is largely due to the widespread use of platform-based business models. Customer behavior patterns and expectations have become more sophisticated as the boundary between traditional industries fades in favor of digital leaders and consumers. In this environment, dynamic multinational activities are currently on the path to digital transformation [28]. In the economic field, banks are still distant from digital banking because they still lack the key jobs and technical bases to implement digital banking. Likewise, digital culture is a lost asset in banking [29]. We know that fin techs are a new segment of the financial market that is made up of a

combination of technology and financial services. This section focuses on financial services and innovation. Innovations come under the heading of research, blockchain and security with a strong emphasis on this area, and represent the most sensitive aspects of the current global issue of digital transformation [30]. The digitalization of banking services based on new technology empowers banks to respond to new customer demands, and the banking sector has undergone major changes [31].

With the growth of technology, organizations are also experiencing massive changes in the design and leadership of their work. Change in life - work and health, use of information and communication technology, performance and management of talent and organizational hierarchy. In addition, two dimensions of macro-level change have evolved with a focus on work structure and leadership [32]. In the area of senior management, despite the importance of healthy partnerships between CEOs and CIOs in organizations for effective business and information technology alignment, we still have little information on how to compare aspects of mutual understanding between them and how they can collaborate and build a unique perspective [33]. In the field of health in the healthcare ecosystem, the digital transformation of health services requires more advanced information technology competence that integrates directly with service users, residents, patients and their relatives in providing care services and creating value [34]. Collection, analysis and management of clinical data with electronic applications has already been widely used. But digitizing medical records along with the principles of electronic medical data management, in addition to enhancing efficiency and reducing treatment costs, ensures clinical effectiveness across all medical institutions and provides a good opportunity for structural transformation efforts in the field of health with digital transformation [35].

With the exchange of information, humans and intelligent objects are able to make common decisions on a broader, higher quality field [36]. And smart, interconnected products are transforming the industry. Live smart products must be (1) integrated at different levels of organizational strategy, including policy, intelligence, control, performance and communication channels, (2) modeled (3) integrated at different levels of organizational strategy. And digital transformation should not lead to separate investments in smart technologies [37]. Future human scenarios and subsequent work and robotics life highlights the need for a robotic roadmap that covers key aspects of industrial and service robotics, including three important areas: the future of robotic technology, digitization and technology. Analyzes ICTs in the key economic, social and political challenges of digital transformation [38]. In the domain of identity, the virtual identity model is also presented as a multi-dimensional concept involving several levels. Virtual identity building is based on a variety of macro- or community-based factors, including narrative scripts, virtual intimacy, virtual community, and virtual instrument culture [39]. In the field of security, a growing synergy between IoT and social technologies is helping to advance the physical cyber-social systems. The integration of new technologies faces key challenges related to information security and privacy [40].

Knowledge sharing in networks also seems to have not received much attention so far. Therefore, integrating the vision of key areas of knowledge, strategy, and innovation and information security management with the aim of identifying the requirements of knowledge protection in the era of digital transformation has been considered. In this regard, both (1) the threat of leakage and exploitation by unauthorized persons, and (2) the threat of unavailability and destruction, are significant challenges for internal and external threats. What makes these threats more difficult to address is the digital transformation that constantly changes the operating landscape of organizations and stimulates the development of complex networks in which organizations participate [41]. There is a significant relationship between variables of organizational culture, knowledge sharing and organizational innovation [42]. Also, in knowledge management of the three dimensions of effectiveness (knowledge acquisition effectiveness, knowledge sharing effectiveness, and knowledge application effectiveness), knowledge acquisition effectiveness has a significant positive effect on innovation [43].

3- Methodology

In data collection, this paper follows a systematic literature review approach that is closely adhered to. A set of scientific approaches that aim to limit systematic error (bias), mainly by trying to identify, evaluate and combine all studies in this field [44]. To reduce the impact of systematic error, we adopted two different approaches. A quantitative approach focusing on content analysis from the research literature and database of selected articles from WOS and a second approach involving qualitative data-based analysis of data-driven methodologies that consisted of articles, interviews, and expert and expert analysis in the field. (Table 1) Data search was conducted on 4 to 12 September 2019. In the data analysis and analysis of the data using the Grounded theory method, open coding was used for key points and themes and consequently the classification of concepts.

The purpose of the Grounded theory method is to produce a theory in the form of a set of related hypotheses obtained through continuous data comparison. In fact, researchers who use the data theory method identify the main concerns of participants and express how they address them. In the Grounded theory method, instead of using data to test the

theory, data is used to create the theory. This method involves a coherent action that leads to the emergence of conceptual categories. These interrelated categories provide a theoretical explanation of the actions that are addressing the main concerns of the participants in the core area of study. Grounded theory analysis involves the transfer of concepts to theories and theories being explored. And coding is optional.

Grounded theory analysis involves the transfer of concepts to theories and theories being discovered. This process includes, open coding, axial coding Such as determining the Core category or phenomenon and determining the Causal conditions, determining the Strategies, determining the Context and Intervening conditions, determining the Consequences and selective coding. Foundation data sampling is a non-random and purposeful sampling and sampling continues until saturation and includes taking notes and discovering gaps at the beginning of research and first interviews, discovering comparisons and designing new questions and questions that are saturated Theoretically, it continues. The details of the research method and the steps to be followed are as shown in Figure 1.

Table 1. Research methodology

method	Description	Details
Quantitative method	Using literature review and analysis of selected past articles from the WOS database containing 1197 articles.	Use specialized vocabulary of the research subject. Pay attention to the most cited articles. Analysis of collected data based on selected parameters.
Qualitative method	Data analysis and analysis using the Foundation data method with expert interviewing (12 people)	Coding and linking to article content, extracting definitions, and common and applied approaches along with expert opinion.

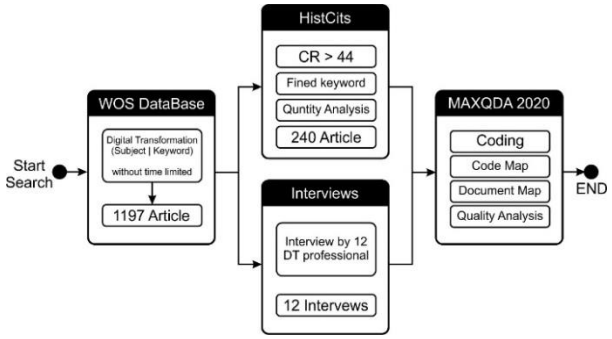


Fig 1. Details of the research method

To evaluate the Grounded theory, four parameters have been used adaptability, comprehensibility, generalizability and control. For audit validation (reliability and validity) of the theory and the final model, the research audit method has been used and the criteria of researcher sensitivity, sampling coherence, sample appropriateness, simultaneous data collection and analysis, have been the most important audit strategies.

4- Findings

4-1- Quantitative Analysis

Using keywords including digital transformation, business strategy, economics, business model, digital innovation, innovative services, technology and digital technology frameworks and primary refinement, the WOS suite of 1197 articles served as the main database .The following article was selected and initial analysis based on them. This number of articles was prepared by 3322 people and included a range of documentation from 1968 to September 2019 and the growing trend of articles from 2014 until now has accelerated. (Figure 2)

Schmidt R, Zimmermann A and Mohring M had the highest number of researches related to the subject in question. In terms of the number of internal citations to articles, four authors (Hess T, Benlian A, Matt C, and Wiesbock F) had the most internal citations in the research database and were the most influential individuals in the research database.

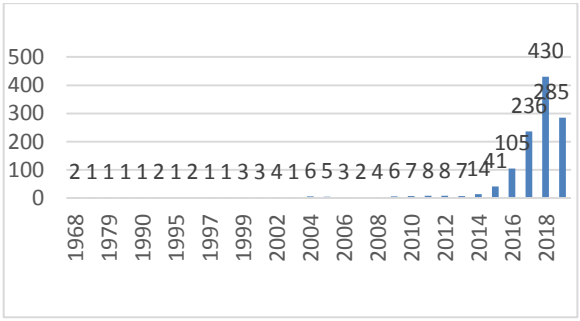


Fig 2. Year and number of articles published

Since 1968, 87 countries in the world have been writing in this field, ranking the number 25 in Germany with 252 and the United States with 146 second and Russia, or 111 among the most active countries in scientific development the subject of digital transformation in the world. Also the most commonly used keywords in all the selected databases of this research are vocabulary such as digital transformation, industry 4.0, data, business.

After preliminary statistical analysis, using HistCite software (Bibliometric Analysis and Visualization Software), 240 articles with the highest CR (Cited references> 44) were selected for content analysis using the Grounded theory method. After this step Source all 240 selected articles were prepared and injected into MAXQDA Analytics pro 2020 software for qualitative analysis.

4-2- Qualitative Analysis

Quantitative findings information was also used as the topic of the main discussions in the interviews with experts in order to examine the broader relationships between them in a more precise and accurate manner. We

also coded the content of 240 articles and, with the help of an expert team, identified 73 codes as variables and categorized them into 9 main code categories (Figure 3). In total, 32,146 codes were identified from the total number of articles and interviews and were marked in the articles and interviews.

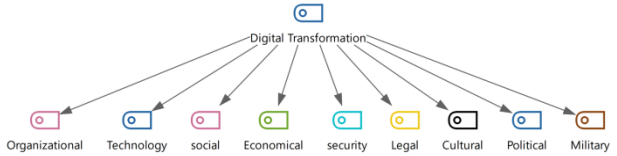


Fig 3. Classification of the main codes

Using the analysis tools available in MAXQDA software, we were able to analyze the data and their results as follows.

4-2-1- In Analyzing and Comparing the use of Codes in Both Categories

In the legal category, the most frequently cited word in the section on Governance is in the Interviews section on Regulations. In the interviews section, the importance of Legal cases is the same as that of Regulations (Figure 4).

	Articles	Interview	
Digital Transformation			
Legal			
Legal	19.0%	40.0%	19.1%
Claims	3.8%		3.8%
Intellectual Property	6.7%		6.7%
Governance	40.9%	20.0%	40.8%
Ownership	14.4%		14.3%
Regulations	15.3%	40.0%	15.4%
SUM	100.0%	100.0%	100.0%
N = Documents	128 (91.4%)	12 (8.6%)	140 (100.0%)

Fig 4. Level of use of legal subcodes in both categories

In both categories of articles and expert interviews, attention to economic issues is estimated to be twice that of the digital economy, and this could indicate the direct impact of digital transformation on the entire economy in addition to complementing the digital economy. In the security category, the overall topic of security has received the most attention in both categories, while the terms protection and cyber security are next in the articles. In the interviews section, cyber security was ranked higher than protection. In the technology category, both focus on the importance of the term digital transformation, with articles in the later stages of IT, followed by ICT, Digitalization and Big Data. But in the interviews, Cloud Computing, Automation and IoT are next. In the organizational category, in the article group, topics related to management, innovation, human resources, knowledge and strategic planning are the most referenced and in the interview group leadership, management, innovation, strategic plan and customer experience are more important. In the military category, only 24 articles referred to the use of digital transformation in the military industry, and no interviews were mentioned in the interview group. In the social category, subjects assigned to social issues had the highest number of referrals in both groups, followed by attention to health and ethics, respectively. It is noteworthy that in the interviews section there is no mention of ethical issues. In general, the study groups in the articles group had the highest level of exposure to the field of technology and then to technology, but in the interviews, the first was in the field of technology and the second was in the field of organization.

4-2-2- In the Rrelational Analysis Section of Codes and Information Content:

In the analysis of the relationship between the codes used in the legal category and the two groups of articles and interviews in general, although the emphasis was on the word "Governance", the most significant relationship was determined between "Regulation" and "Legal" and later between Legal and Governance. In this category, the terms IP and Claims are the most distant from the other codes in this category, indicating a less similarity with the rest of this code. In this domain, only one group with high similarity has been identified, consisting of 4 subjects: Governance, Legal, Regulation, Ownership.

Based on the software output (Figure 5) in the analysis of the relationship between the codes used in the technology category and the two groups of articles and interviews in total, the strongest links between the "Digital Transformation - ICT " codes and the subsequent steps between the "Digital Transformation - IT" and Digital codes "Transformation - Digitalization", "Cloud computing - Big data" and "ICT-IT". In this study, 4 areas with the most dependency were identified. In general, digital technologies, digital transformation, cloud computing, and industry-independent domain 4 can be mentioned.

Based on the software output, the analysis of the relationship between the codes used in the organizational category and the two groups of articles and interviews are the strongest links between the codes of “Management – Innovation”, “management – knowledge”, “Innovation – Knowledge”. The codes for Leadership and Digital skills are the most distant from other codes and are less similar to other codes in this category. In this study, two areas with the most interdependence were identified, the first consisting of management, knowledge and innovation, and the second area of customer experience, business process, business model, digital service.

Based on the software output in the analysis section of the codes used in the social category and in the two groups of articles and interviews, the strongest links were found between the Social - Health codes and in this study one area with the most dependency was identified which included Social and Health.

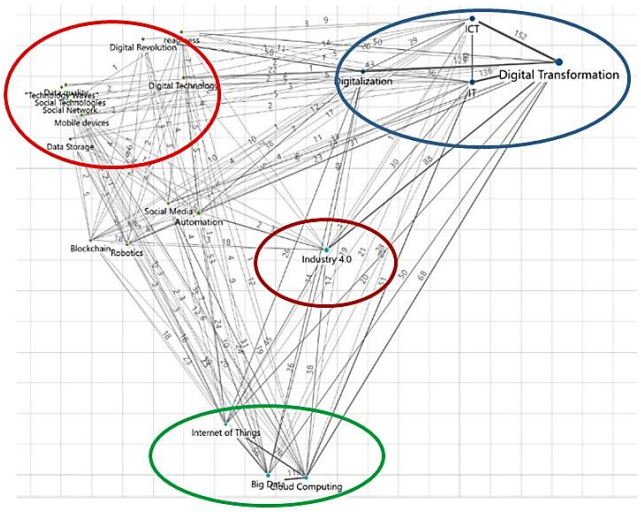


Fig 5. Relational analysis of subcodes and content of information in technology categories by MAXQDA 2020

In the cultural, political and military categories, for the existence of only one code, communication is not meaningful, and also based on the software output in the analysis of the relationship between the codes used in all three cultural, political and military categories and the sum of two groups of articles. And interviews were the most correlated between the Politic - Culture code, and the military with the most distance from the other two had the least relevance in the sources used to other categories in the category. Therefore, one area with the most dependency can be identified in this study, which is the same as Culture and Politics.

Comprehensive links between all the codes used in the research are shown in. In this figure, different batch codes are generated with different colors, the overall scattering of the identified code sets, and that in the paper research design. They have been used to illustrate the overarching consistency, similarity, and overall relevance of the topics used in the research project, based on a series of articles and interviews with experts. As can be seen in this design, the highest level of communication intensity is contained in the category of organizational and technology codes.

Overall, with Merge all the codes in each category can also identify the areas of influence and extent of communication between the main codes. The greater the linkage and the stronger the line between the two codes, the greater the degree of interdependence and influence these issues have on each other, while distinguishing two distinct domains, the degree of compression between the two categories of technology and organization can also be clearly seen.

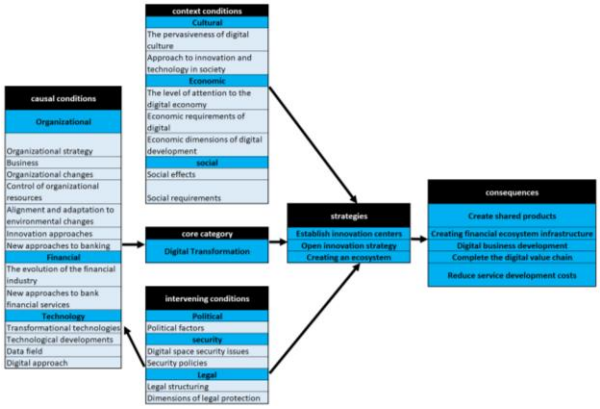


Fig 6. Digital Transformation visual coding paradigm (Model)

All open codes used in the research were categorized based on the characteristics defined for each category and their relationship was determined by creating categories and subcategories related to the central theme of the research (in axial coding). It continued to determine the causal conditions of digital transformation in the organizational, financial and technological sectors. For intervening conditions that affect causal conditions and strategies, security, political, and legal issues were used. In the field of contextual factors that directly affect the strategies, the cultural, economic and social parts of the research were used. (Figure 6)

Then, three main strategies in the face of digital transformation were identified as desirable. It has included the establishment of an innovation center in the organization, the use of an open innovation strategy in the organization, as well as the creation of an efficient ecosystem in the development of digital transformation.

5- Concluding Remarks and Future Perspectives

Our review of the digital transformation model began with the use of WOS research on the phenomenon of digital transformation and utilized the rich literature available in these articles, and at each stage of our expert analysis, collaboration, and viewpoints on dimensions, requirements, Benefits as well as the challenges associated with digital transformation at different levels. In completing this paper, we attempted to analyze the key dimensions of digital transformation by identifying 73 codes in 9 categories and modeling the relationships between them.

Our findings emphasize that in terms of thematic correlation, the technology and organizational domains have the strongest relationship to the final model, constituting the first important and common domain of the model, and the economic, social, and cultural domains, while forming the next common domain with distance. There are more issues in the first area, and issues related to security and legal issues are more separate and more

distant than the two previous ones. After analyzing the bits of code in each batch, we obtain the final cumulative model of Merge interconnection of all batch codes so that we can visualize inter-batch relationships within a model. In the subcodes, the most general impact and use of vocabulary was related to management, innovation, digital transformation, community, economics, human resources and IT, respectively. The digital transformation code comes third after management and innovation, indicating the importance of the first two issues and the extent to which digital transformation is related to management and innovation. Likewise, the topic of cloud computing ranks 24, culture 26, business model 29, and value chain rank 37 out of the 73 codes examined in the article.

Given the underpinnings of Industry 4.0 and the growth of the digital transformation research and development process, it owes more to the 14th position. The maturity of industries and the entry of other areas of business, especially the service sector into digital transformation, seems to have shifted its focus from purely industrial, technical and robotics to human resources such as innovation, knowledge, and strategy. We also find in the literature that digital technologies and their technological innovations are still recognized as the beating heart of digital transformation. The role of technologies such as IoT, blockchain, virtual identity, Big data, and AI in driving this development is undeniable. But organizational approaches, such as changing business strategies and synchronizing the competitive position of companies with these developments, have had a broad role in the evolution of the digital transformation process and its organizational adoption. On the other hand, it can be argued that many of the articles studied in this study, which were also part of the research database, had a specialized, case-by-case look at digital transformation in the particular industry. An overview of digital transformation with a cumulative approach and moving from component to concept has received little attention.

According to the final model, establish innovation centers, Open innovation strategy, creating ecosystem strategies are useful in digital transformation. Also based on this research, create shared products, creating financial ecosystem infrastructure, Digital business development, Complete the digital value chain, reduce service development costs The most important consequences of digital transformation are the implementation of the introduced strategies.

In the practical suggestions section, we can use new strategies to implement digital transformation in organizations, such as creating innovation centers for the development of digital innovations, using open innovation in organizations to attract ideas and partner in product development with other actors. also mentioned the creation of a digital ecosystem for greater interaction with members of the value chain. All of this will help create a successful digital transformation process by developing shared products, digital business development, digital value chain development, as well as reducing the risk and cost of service development.

For further research, we believe that although we have made many efforts to better understand the field of digital transformation, this section still needs to be explored in a more comprehensive way. Also, the study of the functional interaction between each category or code related to one category with other categories and its positive or negative impact on the process of digital transformation in large-scale and small-scale systems has not yet been addressed.

In-depth research is also needed to develop digital business strategies and combine them with new technologies such as artificial intelligence, the Internet of Things and the blockchain. In this regard, coordination between the capabilities of the organization, upstream laws and coordination in digital ecosystems can be a comprehensive research topic.

References

[1] S. Nambisan, M. Wright and M. Feldman, "The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes," RESEARCH POLICY 48 (8): Art. No. 103773, 2019 OCT.

[2] Young and P. Rogers, "A Review of Digital Transformation in Mining," MINING METALLURGY & EXPLORATION 36, pp. 683-699, AUG 2019.

[3] Savastano, C. Amendola, F. Bellini and F. D'Ascenzo, "Contextual Impacts on Industrial Processes Brought by the Digital Transformation of Manufacturing: A Systematic Review," Sustainability 2019, 11, 891, 9 February 2019.

[4] Cozmiuc and . I. Petrisor, "Industrie 4.0 by Siemens: Steps Made Today," JOURNAL OF CASES ON INFORMATION TECHNOLOGY 20 (2): Art. No. UNSP 3, 2018 APR-JUN.

[5] Wikipedia, "Industry_4.0," 2016. [Online]. Available: https://en.wikipedia.org/wiki/Industry_4.0.

[6] H. S. Birkel, J. Veile, J. M. Müller and E. Hartmann, "Development of a Risk Framework for Industry 4.0 in the Context of Sustainability for Established Manufacturers," Sustainability 2019, 11, 384, 2019.

[7] J. Reis, M. Amorim, N. Melão and P. Matos, "Digital Transformation: A Literature Review and Guidelines for Future Research," Springer International Publishing AG, part of Springer Nature , p. 411–421, 2018.

[8] V. Bosilj Vuksic, L. Invancic and D. Susa Vugec, "A preliminary Litereture Review of Digital Transformation Case Studies," in Conference Proceedings, Rome Italy Sep 17-18, Rome Italy, 2018.

[9] P. C. Verhoefa, T. Broekhuizena and Y. Bartb, "Digital transformation: A multidisciplinary reflection and research agenda," Journal of Business Research, 2019.

[10] Scholz, Roland W; Bartelsman, Eric J; Diefenbach, Sarah, "Unintended Side Effects of the Digital Transition:European Scientists' Messages from a Proposition-Based Expert Round Table," Sustainability 2018, 10, 2001, 2018.

[11] K. S. R. Warner and M. Waeger, "Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal," LONG RANGE PLANNING 52 (3), pp. 326-349, 2019 JUN.

[12] L. Downes and P. Nunes , "Big-bang disruption," Harv Bus Rev91(3), p. 44–56, 2013.

[13] E. Henriette, M. Feki and I. Boughzala, "The Shape of Digital Transformation: A Systematic Literature Review," in Mediterranean Conference on Information Systems MCIS 2015 Proceedings, 2015.

[14] C. Matt, T. Hess and A. Benlian, "Digital Transformation Strategies," Bus Inf Syst Eng 57(5), p. 339–343, 2015.

[15] M. Eling and . M. Lehmann, "The Impact of Digitalization on the Insurance Value Chain and the Insurability of Risks," The International Association for the Study of Insurance Economics, 2017 .

[16] T. Borangiu,. D. Trentesaux, A. Thomas, P. Leitao and . J. Barata, "Digital transformation of manufacturing through cloud services and resource virtualization," COMPUTERS IN INDUSTRY 108, pp. 150-162, 2019 JUN.

[17] B. Hinings,. T. Gegenhuber and R. Greenwood, "Digital innovation and transformation: An institutional perspective," INFORMATION AND ORGANIZATION 28 (1), pp. 52-61, 2018 MAR.

[18] A. Cozzolino, . G. Verona and Frank T. R, "Unpacking the Disruption Process: New Technology,Business Models, and Incumbent Adaptation," Journal of Management Studies 55:7, November 2018.

[19] . J. Butschan, . S. Heidenreich, B. Weber and . T. Kraemer, "TACKLING HURDLES TO DIGITAL TRANSFORMATION - THE ROLE OF COMPETENCIES FOR SUCCESSFUL INDUSTRIAL INTERNET OF THINGS (IIoT) IMPLEMENTATION," INTERNATIONAL JOURNAL OF INNOVATION MANAGEMENT 23 (4): Art. No. UNSP 1950036, 2019 MAY.

[20] G. Vial, "Understanding digital transformation: A review and a research agenda," Journal of Strategic Information Systems, 2019.

[21] Accenture, "Digital Transformation Initiative Mining and Metals Industry," In collaboration with Accenture, Cologny/Geneva, January 2017.

[22] D. S. a. W. R. Vinit Parida, "Reviewing Literature on Digitalization, Business Model Innovation, and Sustainable Industry: Reviewing Literature on Digitalization, Business," Licensee MDPI, Basel, Switzerland., 2019.

[23] Y. Sabria, G. J. Michelia and . C. Nuurb, "Exploring the impact of innovation implementation on supply chain configuration," Journal of Engineering and Technology Management, 2018.

[24] M.-S. Denner , L. C. Pu¨schel and M. Ro¨glinger, "How to Exploit the Digitalization Potential of Business Processes," Springer Fachmedien Wiesbaden GmbH, part of Springer Nature , 2017.

[25] V. Nissen, "Digital Transformation of the Consulting Industry-Introduction and Overview," DIGITAL TRANSFORMATION OF THE CONSULTING INDUSTRY: EXTENDING THE TRADITIONAL DELIVERY MODEL , pp. 1-58, 2018.

[26] Bentley, Duncan, "Timeless principles of taxpayer protection: how they adapt to digital disruption," EJOURNAL OF TAX RESEARCH 16 (3), pp. 679-713, 2019.

[27] I. Troshania, M. Janssenb, A. Lymerc and L. D. Parkerd, "Digital transformation of business-to-government reporting: An institutional work perspective," International Journal of Accounting Information Systems, 2018.

[28] B. Alexey V, "MULTINATIONAL BUSINESS IN THE ERA OF GLOBAL DIGITAL REVOLUTION," MIROVAYA EKONOMIKA I MEZHDUNARODNYE OTNOSHENIYA 62 (9), pp. 5-17, 2018 SEP.

[29] N. Pourebrahimi , A. Kordnaeij, K. H. Hamid and A. Azar, "Developing A Digital Banking Framework in the Iranian Banks: Prerequisites and Facilitators," International Journal of E-Business Research (IJEBR) 14(4), pp. 65-77, 2018.

[30] E. Z. Milian, M. d. M. Spinola and M. M. de Carvalho, "Fintechs: A literature review and research agenda," ELECTRONIC COMMERCE RESEARCH AND APPLICATIONS 34: Art. No. 100833, pp. Milian EZ (Milian, Eduardo Z.); Spinola MD (Spinola, Mauro de M.); de Carvalho MM (de Carvalho, Marly M.), 2019 MAR-APR.

[31] B. J. Drasch, A. Schweizer and N. Urbach, "Integrating the 'Troublemakers': A taxonomy for cooperation between banks and fintechs," JOURNAL OF ECONOMICS AND BUSINESS 100, pp. 26-42, 2018 NOV-DEC.

[32] T. Schwarzmüller, P. Brosi, . D. Duman and I. M. .Welpe, "How Does the Digital Transformation Affect Organizations? Key Themes of Change in Work Design and Leadership," mrev, 29 (2), pp. 114-138, 2017.

[33] A. Benlian and . I. Haffke, "Does mutuality matter? Examining the bilateral nature and effects of CEO–CIO mutual understanding," Journal of Strategic Information Systems, p. 104–126, 2017.

[34] J. Dugstad, T. Eide, E. R. Nilsen and H. Eide, "Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia," BMC Health Services Research, 2019.

[35] M. P. Schoenermark, "Medical Data Management - Approach, Concepts, Strategic and Operative Implications," LARYNGO-RHINO-OTOLOGIE 98, pp. S220-S236, 2019 MAR.

[36] A. Zimmermann, . R. Schmidt, . K. Sandkuhl and Others, "Leveraging Analytics for Digital Transformation of Enterprise Services and Architectures," Springer International Publishing Switzerland , 2016.

[37] J. Barata and P. Rupino da Cunha, "The Viable Smart Product Model: Designing Products that Undergo Disruptive Transformations," Cybernetics and Systems, 2019.

[38] J. Kaivo-oja, S. Roth and L. Westerlund, "Futures of Robotics. Human Work in Digital Transformation," International Journal of Technology Management, March 2017.

[39] P. Nagy and B. Koles, "The digital transformation of human identity: Towards a conceptual model of virtual identity in virtual worlds," The International Journal of Research into New Media Technologies Vol. 20(3), p. 276–292, 2014.

[40] S. Mendhurwar and R. Mishra, "Integration of social and IoT technologies:architectural framework for digital transformation and cyber security challenges," Enterprise Information Systems, 2019.

[41] I. Ilvone, . S. Thalmann, . M. Manhart and C. Sillaber, "Reconciling digital transformation and knowledg protection: a research agenda," Knowledge Management Research & Practice, 2018.
[42] F. Kowsari and A. Darush, "Investigating the Impact of Organizational Culture on Innovation with Regard to the Mediating Role of Knowledge Sharing," Journal of Management Systems, vol. 3, no. 4, pp. 22-44, 2017.
[43] S. M. Seyedi, N. Shahidi and S. Hadadnia, "Studying the Relationship Between Knowledge Management Effectiveness and Administrative Innovation Among Shiraz Manufacturing Firms," Journal of Management System, vol. 1, no. 1, pp. 45-63, 2015.
[44] M. Petticrew and H. Roberts, Systematic Reviews in the Social Sciences: A Practical Guide, wiley, 2006.

Mohammad Ali Mirfallah Lialestani is Ph.D. Candidate in Azad University, Roodehen Branch, Iran. His research interests include Information Technology, Innovation and Open Innovation management, Digital Transformation and Digital Ecosystem Management.

Abbas Khamseh is an associate professor of Technology Management in the Department of Industrial Management, Karaj branch, Islamic Azad University, Karaj, Iran. His research interests include Technology Management, R&D Management, Innovation Management, and Entrepreneurship.

Reza Radfar is a Professor of System Management in the Department of Industrial Management, Science and Research Branch, Islamic Azad University, Tehran, Iran. His research interests include System management, Complex systems modeling, Technology Management, Innovation Management.

An Automatic Thresholding Approach to Gravitation-Based Edge Detection in Grey-Scale Images

Kimia Rezaei
Department of Electrical and Electronic Engineering, University College Cork, Ireland
kimia_Rezaei@yahoo.com
Hamed Agahi*
Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran.
agahi@iaushiraz.ac.ir

Received: 14/Nov/2020 Revised: 25/Jul/2021 Accepted: 25/Aug/2021

Abstract
This paper presents an optimal auto-thresholding approach for the gravitational edge detection method in grey-scale images. The goal of this approach is to enhance the performance measures of the edge detector in clean and noisy conditions. To this aim, an optimal threshold is automatically found, according to which the proposed method dichotomizes the pixels to the edges and non-edges. First, some pre-processing operations are applied to the image. Then, the vector sum of the gravitational forces applied to each pixel by its neighbors is computed according to the universal law of gravitation. Afterwards, the force magnitude is mapped to a new characteristic called the force feature. Following this, the histogram representation of this feature is determined, for which an optimal threshold is aimed to be discovered. Three thresholding techniques are proposed, two of which contain iterative processes. The parameters of the formulation used in these techniques are adjusted by means of the metaheuristic grasshopper optimization algorithm. To evaluate the proposed system, two standard databases were used and multiple qualitative and quantitative measures were utilized. The results confirmed that the methodology of our work outperformed some conventional and recent detectors, achieving the average precision of 0.894 on the BSDS500 dataset. Moreover, the outputs had high similarity to the ideal edge maps.

Keywords: Auto-thresholding; edge detection; force feature; the law of universal gravity; the grasshopper optimization algorithm.

1- Introduction

Edge detection is a fundamental step in the image-understanding, which aims to locate the edge points in an image. An edge, is a point at which image brightness changes sharply. In other words, edges are boundary pixels connecting two sections with different amplitude attributes. Edges play a key role in image processing field and contain important information of objects in images [1]. Via edge detection, it is possible to find the boundaries and separate objects from the background. Successful edge detection can valuably simplify the analysis of the information contents in high-level processing tasks such as image segmentation, feature extraction and object identification [2]. However, correct edge recognition in images, depends on using the appropriate edge detector in the existing condition. Therefore, it is required to create an algorithm which discovers most accurate edges.

Several edge detectors have been proposed by researchers. Some of the earliest works include the methods of Sobel [3], Prewitt [4], Canny [5], Roberts [6] and the Laplacian of Gaussian (LoG) [7]. Romani et al. [8] used an iterative method based on interpolation with variably scaled kernels for edge detection. Eser and Derya [9] designed a method based on neutrosophy set using maximum norm entropy for the edge detection. To this aim, they made use of various entropy types. Yuesheng and Lionel [10] proposed an edge detection system based on the physical law of diffusion. Their algorithm dealt with the edge detection as a character of an energy diffusion in media space. Finally, the energy information could be extracted. Bhogal and Agrawal [11] evaluated three algorithms for the image edge detection. They used Sobel, type-1 and interval type-2 fuzzy logic detectors. Their experimental result demonstrated that type-2 fuzzy edge detector achieved better output compared with the other two. Banharnsakun [12] proposed an enhancing edge detection technique via artificial bee colony (ABC) algorithm. In this method, first, an optimal edge filter was found; finally, the threshold value was optimized. Verma and Parihar [13] used a fuzzy system for the edge detection. They applied the 'Uni-value Segment Assimilating Nucleus' (USAN) area calculation. Then, the edge map was fuzzified and the bacterial foraging algorithm (BFA) was used for the membership functions

* Corresponding Author

optimization. Ansari et al. [14] developed a new technique based on intuitionistic fuzzy divergence and entropy measures for the edge detection problem. Then, the results were compared to some other methods using mean square error (MSE) and peak signal-to-noise ratio (PSNR) measures. Shen et al. [15], the authors utilized some hierarchal features learned by Convolutional Neural Networks. (CNN) to attain multi-scale representations. Xie et al. [16] introduced an end-to-end detection model that leveraged the outcomes from different intermediate layers with skip-connections. Xu et al. [17] presented a hierarchical deep model to compute multi-scale features and a gated conditional random field (CRF) to combine them.

An effective approach to handle some complex engineering problems is to utilize nature-inspired algorithms. Newton's law of universal gravitation [18] is one of the typical physical models, based on which, several image-processing algorithms have been proposed for applications including segmentation [19], multi-level thresholding [20] and color texture classification [21]. The edge detection method we present in this note is based on the original use of gravitational forces by Sun et. al. [22]. We will refer to this method as the gravitational approach. In this approach every pixel is considered as a celestial body, which has relationships with other neighboring pixels in the image. Afterward, the vector sum of all gravitational forces of a pixel applied on by pixels in a limited neighborhood is computed. Finally, a threshold for the force magnitude is considered, due to which pixels are separated to edges and non-edges. Several variations of this method are presented in the literature, due to the advantages of the gravity field. Lopez-Molina et al. [23] analyzed the effect of the substitution of the product operation by other triangular norms when computing the gravitational forces. They treated edge points as fuzzy sets and the membership degrees were computed from the resulting gravitational force on each pixel. Verma and Sharma [24] presented an approach for the edge detection using the universal law of gravity and ant colony optimization (ACO). They computed a heuristic function using law of universal gravity as a way to food sources for the ants to detect the edge pixels. Verma et al. [25] proposed an edge detector which combined the universal law of gravity and the gravitational search algorithm (GSA). They treated the edges as masses for which using the law of universal gravity and the GSA, movement of edges were computed. Deregeh and Nezamabadi-Pour [26] assumed image pixels to be celestial objects. Further, the edge points were found using a number of moving agents in the image space. The movement of agents were determined via the forces of objects located in their neighborhood region. Sun et al. [27] proposed an edge detector for hyperspectral images (HSI) using the gravitational theory. Each spatial-spectral vector pixel in an HSI was considered as a celestial body to which forces were exerted by its neighboring pixels. Thus, each body traveled until it reaches a stable equilibrium. Finally, the edges were distinguished by computing the gravitational potential energy.

Wang et al. [28] proposed an edge detector based on the NAGD and MGMF methods. They decomposed each color image into six components in the RGB and HSV model and found the gradient amplitude of the image edge by the CLAHE. Then they constructed an NAGD to extract the edge map of the color image and finally, used the SVD to fuse each channel component to improve the accuracy of the edge detection. Hi et al. [29] proposed a bi-directional cascade network for the edge detection of objects at different scales. In this method, a layer is supervised by labeled edges at its scale, rather than applying the supervision to the network outputs. To enrich the multi-scale representations, a module used the dilated convolution to generate multi-scale features. Li et al. [30] proposed a novel technique to resolve the balance between the fast training and accurate testing aspects of multi-scale representation in the edge detection task. According to multi-stream structures and the image pyramid principle, they constructed two pyramid networks to enrich the multi-scale representation. Then using a backbone network, the overall bi-directional pyramid network (BDP-Net) architecture was constructed. Lu et al. [31] proposed a vector co-occurrence morphological operator for the edge detection, which considers both the pixel and boundary information. This method resists the influence of the noise points and detects the edges from the color image rather than the gray image. Anand and Sangeethapriya [32] presented an isotropic Gaussian modulated hyperbolic tangent high-pass filter to remove the high frequency components in noisy images, as a pre-process in the edge detection task. Their filter achieved better directional selectivity and offered less noise sensitivity along with regularization by the least square error design.

The main focus of this paper is to find an optimal threshold for the force magnitudes, according to which, the pixels are dichotomized to the edge and non-edge classes. First, a normalizing map is applied to the intensity values from the domain $[0,255]$ to the range $(0,1]$. Through applying this mapping, even zero-intensity pixels could be considered as edges. Afterwards, similar to the original work, the gravitational force exerted by neighboring pixels on each pixel is computed. Following this, a fuzzy membership function is applied to the overall force magnitude in order to control the uncertainties due to discretization and noises in the original image. The output of this step is called the *force feature*, for which an optimal threshold is desired to be found. Auto-thresholding is the main novelty of this paper. To this aim, the histogram of the force feature is computed. Then, an iterative process is applied to tune the threshold, starting from the mean value

of the histogram. Once the threshold value is converged, its final value is chosen as the optimal threshold for the force feature. Selecting the initial point for the iterative process has a substantial effect on the speed of convergence and accuracy of the classification. Thus, we present a new formulation for $T_{initial}$ which contains the mean and the standard deviation of the force feature distribution together with two constant regulating parameters. This formulation includes both information from the image under study and general information provided from all the images in the dataset, hidden in the constant parameters. To find the optimal values of these parameters and one parameter in the fuzzy membership function of the previous step, we perform a metaheuristic optimization algorithm. For this goal, the grasshopper optimization algorithm (GOA), proposed by Saremi et al. in 2017 [33] is chosen. This algorithm is inspired from the life of grasshoppers in large swarms and their effort for finding new sources of food. This algorithm works well in both the exploration and exploitation phases in complex uni-modal and multi-modal problems and accurately discovers the global optimum solution. This solution yields the optimal values of the required parameters.

The proposed algorithm is tested on two standard image datasets with and without the contamination of noise. The first one is the USC-SIPI and the second dataset is the BSDS500. The results are compared with the ideal edge maps to evaluate the performance of the system. Some qualitative and quantitative measures are employed for this purpose, including visual comparison, Pratt's figure of merit, Shannon Entropy and the Average Precision. Comparisons are also performed with some conventional edge detectors and also some recent methods from the literature on the mentioned datasets, to show the superiority of the proposed approach. Four cases are considered according to the noise type and density, comprising of clean image, two Gaussian noise with different variances and the salt and pepper noise. The results are also compared with some conventional edge detectors. Moreover, four scenarios are studied for the thresholding phase, containing: (*i*) the Otsu method, (*ii*) auto-thresholding starting from T_{mean}, (*iii*) direct finding the optimal threshold using the GOA, and (*iv*) auto-thresholding initiating from the formulation of $T_{ini.opt.}$ in (16) whose parameters are determined by the GOA. The latter three scenarios are the proposed thresholding approaches of this paper. The fourth scenario shows to be competitive, obtaining better results than different published methods and also other proposed scenarios.

The remainder of this paper is structured as follows. Section 2 describes the background of the Newton's law of gravy. The proposed approach is presented in Section 3. In this section, first the basic gravitational edge detector is reviewed. Then the proposed thresholding methods are introduced. At last, the section is extended by laying out the steps of computing an optimal threshold value by the GOA. Section 4 is devoted to the evaluating the proposed method and comparing it with some others, while applying to two standard databases. In the end, Section 5 presents a brief summary and includes a conclusion and future lines of research into this area.

2- Background of the Universal Law of Gravitation

Based on the Newton's law of universal gravitation [18], objects attract and being attracted by other nearby objects. The mutual attractive force is directly proportional to the product of the masses and inversely to the square of their distance, as illustrated in Fig. 1. The mathematical formula describing the gravitational force can be written as follows:

$$\vec{f}_{1,2} = \frac{G\, m_1\, m_2\, \vec{r}}{\|\vec{r}\|^3} \quad (1)$$

Where, $\vec{f}_{1,2}$ represents the force vector applied by mass #2 on mass #1. The universal gravitational constant is represented by G while the objects gravitational masses are shown by m_1 and m_2 respectively. The mass is a measure of the strength of the gravitational field on a particular object. The gravitational field of a body has a direct proportional relation with its gravitational mass. Also $\vec{r}$ stands for the distance vector connecting the centers of the masses.

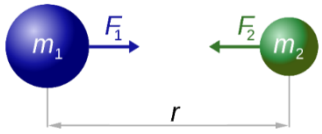


Fig. 1 Model of Newton's law of universal gravitation for masses m_1 *and* m_2

The gravitational force obeys the superposition principle. The force exerted on object k by its neighbor objects in the region Ω is found by the vector sum of forces exerted by each object.

$$\overrightarrow{F_k} = \sum_{i\in\Omega} \vec{f}_{k,i} \quad (2)$$

3- The Auto-thersholded Gravity-based Edge Detection

3-1- Overview

Based on the universal law of gravitation, several edge detection systems are presented in the literature. In this section, the proposed approach for auto-thresholding the edge detection system based on the gravitational approach is presented. The general platform of our approach consists of two phases, as Fig. 2 shows. In the first phase,

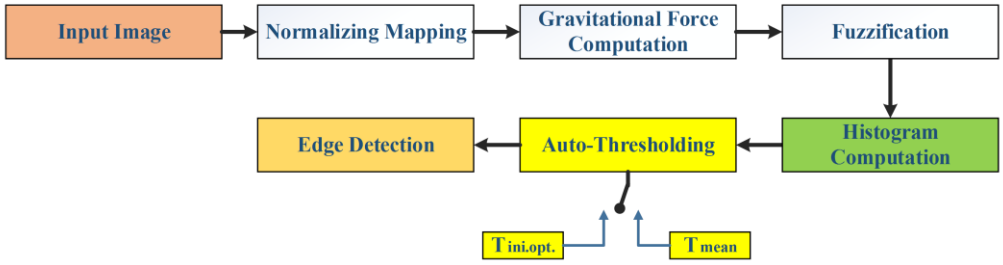


Fig. 2 The block diagram of the proposed edge detection system

the gravitational edge detector is generated to determine the force exerted on each pixel. Then, the force magnitude is mapped to the force feature using a fuzzy membership function; and the histogram of this feature for the whole image is created. In the second phase, an optimal threshold is found for that histogram with respect to which, the pixels are classified to edges and non-edges. The next sub-sections describe the details of these steps.

3-2- Generation of the Gravitational Edge Detector

In order to construct an edge detector, it is supposed that every pixel in the image is a celestial body, which affects and being affected by other neighboring pixels through gravitational forces [22]. For each pixel, the forces of pixels beyond a pre-specified distance are assumed to be zero. In other words, far-away points do mutually not exert nor receive any forces. The magnitude and the direction of the vector sum of all gravitational forces that each pixel receives from its neighbors include significant information about the intensity changes and the existence of an edge in that point.

To start with, we denote a pixel at location (i,j) by $x_{i,j}$ and its gray level intensity value by $I(i,j)$. As the next relations will show, the force exerted on zero-intensity pixels will be zero. Thus, they will never be considered as edge points, even when their neighbors' intensities change sharply. To avoid this problem, a mapping of intensities from the original domain to the range (0,1] is implemented. Suppose that the maximum intensity value in the image is I_M. The following operation are performed with $\eta \ll 1$ [23].

$$G(i,j) = I(i,j)/I_M\,, \quad g(i,j) = \frac{G(i,j)+\eta}{1+\eta} \tag{3}$$

This mapping enables the gravitational edge detector to identify the possible edge pixels with zero intensities. The edge detection technique using the law of universal gravitation is described below [22]:

i. An $m \times n$ neighborhood Ω with pixels $x_{k,l} \in \Omega$ and $(k,l) \neq (i,j)$ is considered for each image point $x_{i,j}$. With respect to (1), the gravitational force exerted by each neighbor $x_{k,l}$ on $x_{i,j}$ is computed as follows:

$$\vec{f}_{i.j;k.l} = \frac{G\, m_{i,j}\, m_{k,l}\, \vec{r}}{\|\vec{r}\|^3} \tag{4}$$

Here, $\vec{f}_{i,j;k,l}$ is the applied gravity force and $m_{i,j}$ and $m_{k,l}$ are the mapped intensity values of pixels, i.e., $g(i,j)$ and $g(k,l)$. Vector $\vec{r}$, represents the distance vector from $x_{k,l}$ to $x_{i,j}$, whose magnitude is computed by (4):

$$\|\vec{r}\| = \sqrt{(k-i)^2 + (l-j)^2} \tag{5}$$

ii. Considering an image as a 2D plane, the vectors of the gravitational forces in the horizontal and vertical directions (i.e., $\hat{x}$ and $\hat{y}$) can be computed as below:

$$f^x_{i,j;k,l} = \left|\vec{f}_{i,j;k,l}\right| \sin\theta \;= \frac{G\, m_{i,j}\, m_{k,l}}{\|\vec{r}\|^2}\frac{k-i}{\|\vec{r}\|} = \frac{G\, m_{i,j}\, m_{k,l}\,(k-i)}{\|\vec{r}\|^3},$$
$$f^y_{i,j;k,l} = \left|\vec{f}_{i,j:k,l}\right| \cos\theta \;= \frac{G\, m_{i,j}\, m_{k,l}}{\|\vec{r}\|^2}\frac{l-j}{\|\vec{r}\|} = \frac{G\, m_{i,j}\, m_{k,l}\,(l-j)}{\|\vec{r}\|^3} \tag{6}$$

Hence, the vector $\vec{f}_{i,j;k,l}$ is represented as below:

$$\vec{f}_{i,j;k,l} = f^x_{i,j;k,l}\hat{x} + f^y_{i,j;k,l}\hat{y} \tag{7}$$

The vector sum of all gravitational forces applied by neighbors on $x_{i,j}$ is demonstrated as:

$$\vec{F}_{i,j} = \textstyle\sum_\Omega F^x\,\hat{x} + F^y\hat{y},\ (k,l) \in \Omega\ \&\ (k,l) \neq (i,j) \tag{8}$$

Where,

$$F^x = \textstyle\sum f^x_{i,j;k,l}\ , F^y = \sum f^y_{i,j;k,l}\ ,\quad (k,l) \in \Omega\ \&\ (k,l) \neq (i,j) \tag{9}$$

iii. The magnitude and direction of the vector $\vec{F}_{i,j}$ are determined as in (10):

$$F = \sqrt{{F_x}^2 + {F_y}^2}, \qquad \theta = \tan^{-1}\left(F_x \big/ F_y\right) \tag{10}$$

iv. Finally, the edge map is produced by setting an appropriate threshold on the force magnitude histogram.

For simplicity, $G.m_{i,j}$ is replaced with a constant C. Base on experiments, we set $C = 1$ in clean images and $C = sig(g(i,j))$ in noisy images, where $sig(.)$ is a sigmoid function. The effect of noise on edge detection may also be reduced if more neighboring pixels are considered. In this paper, the neighborhood is represented by a square of size 5×5.

Digital images have some inherent uncertainties due to discretization and noise. Hence, the fuzzy logic is proved to be an appropriate tool for representing and processing

of the edges. Since the magnitude of the force characterizes the difference in intensities of the neighboring points, we generate a membership function to map this feature to a new one which will be used for finding the edge image. In order to do so, the Gaussian membership function is considered to act on the force magnitude of each pixel [34]. The function M_G, given by (11) and shown in Fig. 3, has only a single parameter δ which represents the standard deviation of the distribution. The maximum of the force magnitude F is represented by γ. We mapped the force magnitude from $[0, \gamma]$ to [0,1], so that, $M_G(0) = 0$ and $M_G(\gamma) = 1$. Hence, the function is:

$$q = M_G(F) = \exp(-(\ln(F) - \ln(\gamma))^2 / 2\delta^2), \quad (11)$$

We call q the 'force feature'. This operation can improve the results of the edge detection method. The parameter δ is one of the three parameters which will be adjusted through an optimization technique in next steps.

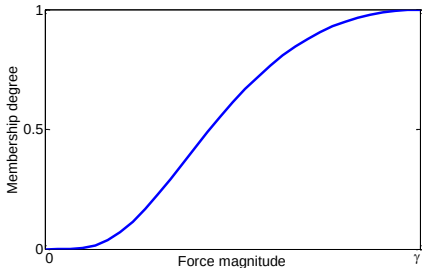


Fig. 3 Gaussian membership function for the conversion of force magnitudes into membership degrees

3-3- Thresholding

Thresholding is a non-linear operation used for edge detection and segmentation [35]. This operation converts a grayscale image into a binary image in which, two levels are assigned to pixels that are below or above a specified threshold value. Selecting an optimal threshold value is the most important step in every image thresholding algorithm. This paper mainly focuses on computing the optimal value for the threshold by which, the edge points of the images are well distinguished. Using this method, finding weak edges of noisy and clean images will be attainable. First, the histogram of distribution of the force feature q is determined. The goal is to find an optimal threshold for this histogram. Those pixels whose force feature is above this threshold are classified as edges, while the rest will be non-edges. To do this, an iterative approach based on the averages of force features is used as follow:

1. Compute the average force feature of all pixels in the image, to be used as the initial threshold value:

$$T = T_{initial} = T_{mean} = mean(q) \quad (12)$$

2. Dichotomize the histogram according to the threshold T:

$$Q_1 = \{q | \, q \leq T\}, \qquad Q_2 = \{q | \, q > T\}, \quad (13)$$

3. Compute the means of Q_1 and Q_2:

$$M_1 = mean(Q_1), M_2 = mean(Q_2) \quad (14)$$

4. The new threshold value is obtained by computing mean of M_1 and M_2:

$$T_{new} = {}^1/_2 (M_1 + M_2) \quad (15)$$

5. If $T_{new} - T < \epsilon$, stop and choose T_{new} as the optimal threshold value. Otherwise, set $T = T_{new}$ and go to step 2.

As it will be seen in the next section, this approach yields accurate results, both in clean and noisy images with higher indices, compared with other existing methods. However, we tried to still improve our technique by finding an appropriate initial threshold value for the first step. Smart selection of $T_{initial}$ helps in the fast convergence of the iterative process and more robustness against different noises, as the simulation results validate in the simulation section. Hence, we redefine $T_{initial}$ as in (16).

$$T_{initial} = T_{ini.opt.} = \frac{\mu + \alpha.\sigma}{\mu + \beta.\sigma} \quad (16)$$

In which, μ and σ represent the mean and the standard deviation of q respectively, and α and β are two regulating parameters. It should be noted that, $T_{initial}$ in (16) is not fixed for all images, but it depends on some constants α and β together with two characteristics of the image under study (i.e., μ and σ). In order to find α, β and also δ in (11), we make use of an optimization technique. The grasshopper optimization algorithm, reviewed in the next sub-section, is run on several images and their ideal edge maps to find the optimal values of these parameters. The position of each grasshopper is represented by a vector of length three in the form $\boldsymbol{x} = [\alpha \; \beta \; \delta]^T$. Figure of Merit (FOM), defined in (24), is used as the objective function for parameter optimization. The optimization problem can be stated as maximizing FOM subjected to $\alpha, \beta, \delta > 0$. Once the optimal valued of these parameters are found, $T_{initial}$ in (16) will be used as the initial point of the iterative process (12-15) for the test images.

3-4- Overview of the Grasshopper Optimization Algorithm

The grasshopper optimization algorithm (GOA), proposed by Saremi et al. in 2017 [33], is a novel metaheuristic optimization algorithm that models and mimics the swarm behavior of grasshoppers. As notorious insects for damaging the farms, grasshoppers move as large swarms to seek the food sources. In the nature-inspired GOA, grasshoppers can globally search the given space, discover the regions of answers and move locally in the final steps in order to find the optimal solutions. GOA, models the mutual forces exerted by the grasshoppers for the repulsion and attraction. While the repulsion forces encourage the grasshoppers to explore the search space, the attraction forces allow them to exploit the local regions. Both exploitation and exploration of this

algorithm are satisfactory for uni- and multi-modal test functions and these two tendencies are balanced in the GOA. The objective function is improved over the course of iterations to provide a more accurate approximation of the global optimum solution. This algorithm has been applied to many engineering problems and shown a remarkable performance in many fields, such as feature selection [36], vibration analysis [37] and more. The movement of a grasshopper is subjected to three factors: (*i*) social interaction, (*ii*) gravity force and (*iii*) wind advection. In general, the position of the i-th grasshopper in a swarm can be modeled as follows:

$$X_i = S_i + G_i + A_i \quad (17)$$

In which, S_i, G_i *and* A_i represent the social interaction, the gravity force and the wind advection for the grasshopper, respectively. The social interaction for the i-th grasshopper is defined as follows:

$$S_i = \sum_{\substack{j=1 \\ j\neq i}}^{N} s(d_{ij})\,\hat{d}_{ij} \quad (18)$$

In which, d_{ij} is the Euclidian distance between the two grasshoppers i and j. Also, $\hat{d}_{ij} = (x_j - x_i)/d_{ij}$ is the unit vector directing from the location of the i-th grasshopper (x_i) to that of the j-th one (x_j). The number of grasshoppers is shown by N. The s function indicates the social force, as defined in (19).

$$s(r) = f.\exp(-r/ls) - \exp(-r) \quad (19)$$

Where, f is the attraction intensity and ls represents the attractive length scale. Additionally, the gravity component is computed as follows:

$$G_i = -g.\hat{e}_g \quad (20)$$

In (20), the parameter g is the gravity constant and $\hat{e}_g$ is the unit vector towards the center of the earth. As well, the wind advection component is defined in (21).

$$A_i = u.\hat{e}_w \quad (21)$$

Where, u is the constant drift and $\hat{e}_w$ is the unit vector which represents the wind direction. To solve optimization problems in D dimension, an improved version of (17) was presented in [33]:

$$X_i^d = c\left(\sum_{\substack{j=1 \\ j\neq i}}^{N} c\,\frac{ub_d - lb_d}{2}\,s\left(\left|x_j^d - x_i^d\right|\right)\frac{(x_j - x_i)}{d_{ij}}\right) + \hat{T}_d, \quad d = 1, \dots, D \quad (22)$$

In which, ub_d and lb_d are the upper and lower bounds in the d-th dimension and $\hat{T}_d$ indicates the value of the d-th dimension in the best solution found so far. In (22), the gravity coefficient is ignored and the wind advection component is set equal to $\hat{T}_d$. Additionally, c is a multiplying coefficient which decreases the exploration while increases the exploitation proportional to the number of iterations:

$$c = c_{max} - l\,\frac{c_{max} - c_{min}}{L} \quad (23)$$

Where c_{max} and c_{min} indicate the maximum and minimum values. Also l and L are the current and maximum number of iterations, respectively. In this paper, we set: $f = 0.5$, $ls = 1.5$, $c_{max} = 1$ and $c_{min} = 1e-5$.

4- Experimental Results

This section compares the performance results of the proposed edge detection system in different scenarios to those produced by some classic and some recent methods from the literature. First, the datasets for evaluating the methods are introduced, then the methodology of the experiments is described. Here, both qualitative and quantitative measures are utilized. Finally, the results are given and compared for clean and noisy images.

4-1- The Datasets

In this experiment, two datasets are used for evaluating the methods. The first one is the USC-SIPI (Signal and Image Processing Institute, USC) [38] which offers a collection of digitized images with different sizes. In this paper, the images of size 256×256 are used (Fig. 4a-b). The second dataset is the BSDS500 (Berkeley Segmentation Dataset, Computer Vision group, UCB) [39]; which consists of 500 natural images and ground-truth annotations (Fig. 4c-e). Each image in this dataset is manually annotated by multiple annotators; whose outcomes are averaged to give the final ground-truth for that image. The final annotations are used as the ideal edge maps, which are solutions for the edge detection problem. Moreover, the results for the cameraman image (Fig. 4f) are also given. The 10-fold cross-validation method on the datasets is used for evaluating the methods.

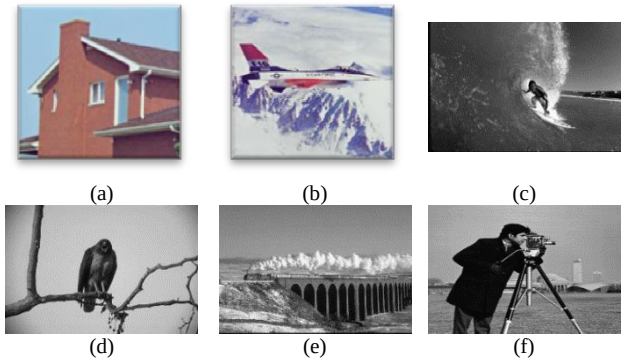


Fig. 4 Samples of the datasets images. (a-b) USC-SIPI, (c-e) BSDS500, (f) the cameraman

4-2- The Measures

In this section, by the *proposed method* we mean computing the histogram of the feature force and applying the auto-thresholding iterative process initiating from $T_{ini.opt.}$ whose optimal parameters are found via the GOA. To evaluate the proposed method applied to the mentioned

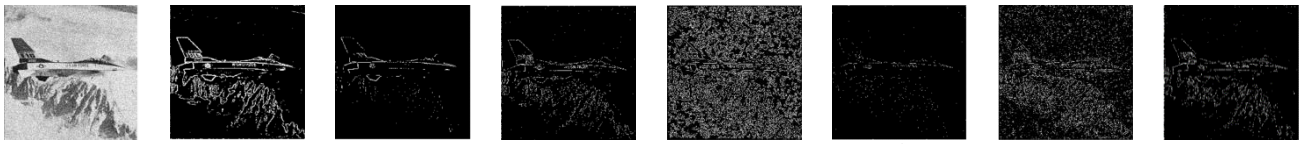
datasets, two sets of measures are used. The first set contains the qualitative measure, while the second one comprises the quantitative measures. Four experimental cases were considered according to the noise type and density. The first case contains clean images while, in Cases II-III, images are corrupted with two zero-mean Gaussian noises with $var = 0.01$ and 0.05, respectively. In the latter cases, a 5×5 mask is used to decrease the Gaussian noise effect. Also, the images in Case IV are contaminated with the salt and pepper noise with density of 0.05. In this case, a median filter is applied to the images in pre-processing step, in order to moderate the effect of the noise and enhance the performance. Besides, notice that while generating the gravitational edge detector, we set $C = 1$ in clean images and $C = sig(g(i,j))$ in noisy images.

4-2-1- The qualitative Measure

The most common and simplest way of evaluating an edge detector is the visual comparison of output images which could be done in a range of different ways. For instance, image gradients can be compared visually where an edge image is assessed by a group of individuals and the average score may serve as the quality index. Fig. 5 shows the output images of various classic detectors with *clean image* input (Case I). It is obvious that, most of the edge structures are well preserved by the proposed detector compared with others such as Canny, Sobel or Robert. Therefore, the proposed technique is more efficient than others for the edge detection of these images. When noise is added to the image, finding an optimal threshold value becomes more important, since the aims are to discover true edges and ignore those made due to the noise. Results on images corrupted with Gaussian noise (Case II) is shown in Fig. 6. It should be noted that, we used Case III only in evaluating the system by quantitative measures. Visual comparison clearly reveals the superiority of the proposed approach compared to the others. Not only sharp edges are successfully found, but also more artefacts and irrelevant information are removed using the proposed detector. The results of Case IV are shown in Fig. 7. According to the visual differences, it goes without saying that, the noise density in the edge map resulted from the proposed method is considerably lower than that of the output images attained by other methods. Additionally, sharp edges are successfully found.

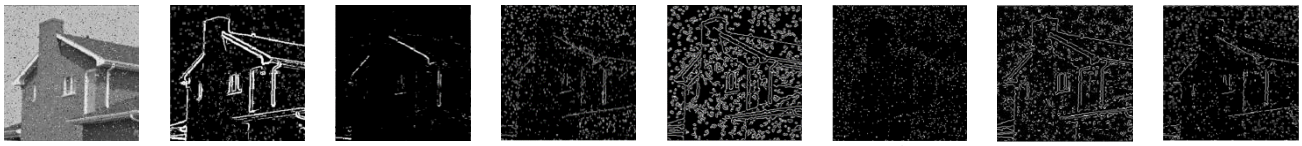
(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 5 Edge Detection of a clean image (Case I): (a) the original image, (b) the proposed edge detector, (c) Otsu using 5×5 mask, (d) Sobel, (e) Canny, (f) Robert, (g) LOG, (h) Prewitt.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 6 Edge detection of images corrupted with Gaussian noise (Case II). The sub-figure captions are same as those in Fig. 5.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 7 Edge detection of images contaminated with the salt and pepper noise (Case IV). The sub-figure captions are same as those in Fig. 5.

4-2-2- The quantitative Measure

To perform quantitative comparison, we use three measures: the 'figure of merit' (FOM), the 'Shannon Entropy' and the Average Precision (AP). With respect to these measures, a more accurate evaluation on the performance of the proposed and existing methods in different conditions is provided.

Figure of Merit (FOM)

Pratt's figure of merit is a popular measure for evaluating the performance of an edge detector system. The miss-detection errors can be classified as (*i*) missing valid edge pixels, (*ii*) detecting noises and artefacts as valid edges, and (*iii*) smearing of edges [40]. The Pratt's FOM is defined as Follows:

$$F = \frac{1}{\max(N_l, N_D)} \sum_{i=1}^{N_D} \frac{1}{1+\varepsilon d_i^2} \quad (24)$$

In (24), number of the edge pixels in the detected and ideal edge images are indicated by N_D and N_l respectively. The Euclidean distance between the location of the i-th detected edge pixel to the nearest edge pixel in the ideal edge image is indicated by d_i. Also ε is a scaling constant chosen to be $\frac{1}{9}$ based on the Pratt's work. F is an index in the range of [0,1], which measures the accuracy of the edge localizations. The larger the value of F, the better the performance is. An edge detector with $F = 1$ is perfect. For some images in Fig. 4, the FOM of the proposed and some existing methods are reported in Table 1. The proposed algorithm found meaningful edges in cleanest images (Case I) in comparison with other algorithms. In fact, it kept the main objects in the images and provided the most similarity with the ideal edge maps. Also, the proposed method significantly decreased the effect of Gaussian noise in Case II; while the performances of other detectors were considerably lower. When the density of noise was increased (Case III), our method yielded higher similarity between detected and ideal edges compared with other detectors. Moreover, the performance of the proposed technique on images corrupted by salt and pepper noise (Case IV), was still higher compared to other detectors. However, it was less than those for Cases I-III. As a matter of fact, the salt and pepper noise existed in the output image beside the edge points, yet with a lower density compared with other methods.

The Shannon Entropy Measure

The Shannon entropy presents a measure for estimating the average of a random flux of information in an image I, which is formulated as follows [41]:

$$H(I) = -\sum_{k=0}^{K-1} p_k . \log(p_k) \quad (25)$$

Where K is the number of gray levels (here, $K = 256$) and p_k is the probability associated with the gray level k. In general, high Shannon entropy means high disorder in the image data. To put it another way, the edge pixels which represent the discontinuity are better identified in an image with higher entropy. Therefore, Shannon entropy is an efficient method for evaluating the performance of an edge detector. However, a high value of Shannon entropy is not necessarily equal to rich information, due to the effect of uncontrolled noise on the entropy value. Tables 2-3 compared the performance of the proposed and some classic edge detectors in term of the Shannon entropy for all images in Fig. 4.

To attain a fair evaluation, the entropy and FOM measures should be considered at once. That is to say, if the entropies of two methods are close to each other, their FOMs should be compared in order to determine the more accurate one. For example, although the canny detector resulted in high entropy values, extra and undesirable edges were detected, leading to output image distortion. To illustrate, a visual comparison of the outputs of the canny and the proposed approaches is shown in Fig. 8. The former method detected some undesirable points as edges, while the latter approach attained a better outcome. For all clean images (except Fig. 4c), the entropy of the proposed method was higher than others which generally demonstrates its higher efficiency. Moreover, higher FOM results achieved by the proposed approach verified its effectiveness.

If the variations in the entropy of a system in clean and noisy conditions are small, that system is robust against the noise existence and the effect of noise is controlled in the system output. Due to the non-gradient structure of the gravitational edge detection method, such robustness is expected. Comparing the variations in Cases I-III of Tables 3-4 with the existence of Gaussian noises, it is obvious that, our approach not only has high entropy values (i.e., remaining the edges), but also reveals less entropy variations with respect to the noise-free condition, compared to other methods. Results in Case IV (image with salt and pepper noise), also show that the entropy variation is more acceptable than the other techniques.

Table 1: Comparison of the edge detection techniques based on the FOM measure. 'PrM.' stands for the 'Proposed Method'.

Detector	Case I			Case II			Case III			Case IV		
	Fig. 4c	Fig. 4d	Fig. 4e	Fig. 4c	Fig. 4d	Fig. 4e	Fig. 4c	Fig. 4d	Fig. 4e	Fig.4c	Fig.4d	Fig.4e
Sobel	0.3948	0.2843	0.3901	0.2951	0.2592	0.3010	0.2523	0.2394	0.2401	0.1735	0.3464	0.3691
Canny	0.2247	0.4091	0.7589	0.1053	0.1889	0.7463	0.1021	0.1818	0.7255	0.1270	0.2314	0.7286
Roberts	0.3604	0.3308	0.2754	0.1278	0.1688	0.0957	0.0619	0.0868	0.0475	0.1427	0.1902	0.1636
LoG	0.2516	0.3335	0.5895	0.1459	0.3678	0.6561	0.1277	0.2874	0.7285	0.1429	0.3063	0.7110
Prewitt	0.3947	0.2847	0.3857	0.2993	0.2618	0.3039	0.2543	0.2439	0.2432	0.1784	0.3544	0.3627
PrM.	0.5164	0.7663	0.8312	0.4547	0.6248	0.7718	0.4561	0.6781	0.7431	0.3319	0.5311	0.7992

Table 2: Comparison of the edge detection techniques based on the Shannon Entropy measure for images in Fig. 4a-c.

Detector	Case I			Case II			Case III			Case IV		
	Fig. 4a	Fig. 4b	Fig. 4c	Fig. 4a	Fig. 4b	Fig. 4c	Fig. 4a	Fig. 4b	Fig. 4c	Fig. 4a	Fig. 4b	Fig. 4c
Sobel	0.2198	0.2243	0.1943	*0.1832*	*0.1894*	*0.1494*	*0.1645*	*0.1751*	*0.1488*	0.2955	0.2828	0.2564
Canny	0.3842	0.4062	0.4664	*0.7180*	*0.7136*	*0.8117*	*0.7696*	*0.7953*	*0.8393*	0.4770	0.5111	0.5569
Roberts	0.2150	0.2190	0.1606	*0.0700*	*0.0729*	*0.0648*	*0.0437*	*0.0468*	*0.0411*	0.1898	0.1687	0.1643
LoG	0.2967	0.3197	0.3364	*0.4291*	*0.4264*	*0.5612*	*0.5201*	*0.5080*	*0.6176*	0.3459	0.3779	0.4412
Prewitt	0.2187	0.2236	0.1944	*0.1857*	*0.1910*	*0.1505*	*0.1672*	*0.1764*	*0.1484*	0.3215	0.3318	0.3084
PrM.	0.4914	0.4765	0.3918	*0.4423*	*0.4542*	*0.4391*	*0.4835*	*0.4711*	*0.4516*	0.5258	0.5958	0.4726

Table 3: Comparison of the edge detection techniques based on the Shannon Entropy measure for images in Fig. 4d-f.

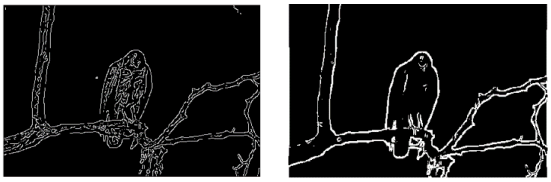
Detector	Case I			Case II			Case III			Case IV		
	Fig. 4d	Fig. 4e	Fig. 4f	Fig. 4d	Fig. 4e	Fig. 4f	Fig. 4d	Fig. 4e	Fig. 4f	Fig. 4d	Fig. 4e	Fig. 4f
Sobel	0.1651	0.3043	0.2339	0.1533	0.2519	0.2174	0.1553	0.2098	0.2102	0.2461	0.2786	0.2341
Canny	0.2714	0.5007	0.4713	0.8142	0.5910	0.7270	0.8333	0.7535	0.7673	0.4327	0.5552	0.8143
Roberts	0.1795	0.2327	0.2223	0.1005	0.1038	0.1314	0.0623	0.0567	0.0873	0.1923	0.2329	0.1187
LoG	0.2219	0.4278	0.3223	0.4064	0.4606	0.3992	0.5069	0.5281	0.4784	0.3263	0.3543	0.4285
Prewitt	0.1648	0.3012	0.2344	0.1541	0.2512	0.2177	0.1543	0.2139	0.2098	0.2899	0.2822	0.2014
PrM.	0.4758	0.7471	0.5725	0.3819	0.6888	0.5622	0.4511	0.7691	0.5418	0.5161	0.6214	0.6072

Table 4: Comparison of the thresholding scenarios in four mentioned cases based on the quantitative measures for images in Fig. 4. Each value states the average of each measure for all seven images.

	Scenario #1		Scenario #2		Scenario #3		Scenario #4	
	FOM	*Entropy*	*FOM*	*Entropy*	*FOM*	*Entropy*	*FOM*	*Entropy*
Case I	0.4688	0.2668	0.6819	0.4976	0.6772	0.4734	0.7046	0.5258
Case II	0.4689	0.2718	0.6214	0.4530	0.5663	0.4450	0.6171	0.4948
Case III	0.4407	0.2772	0.6099	0.4817	0.6148	0.4754	0.6258	0.5280
Case IV	0.4892	0.2916	0.5295	0.5264	0.5081	0.4764	0.5541	0.5566

Table 5: Comparison of the Average Precision with some other methods from the literature on BSDS500 dataset.

SCG [42]	0.773	DeepContour [15]	0.800	COB [43]	0.859	Deep Boundary [44]	0.789
PMI [45]	0.799	HFL [46]	0.795	DCD [47]	0.849	CED [48]	0.847
OEF [49]	0.820	HED [16]	0.840	AMH-Net [17]	0.869	BDCN [29]	0.796
MEDM [31]	0.77	NAGK-MF [28]	0.72	BDP-Net [30]	0.847	**PrM.**	**0.894**



(a) (b)

Fig. 8 The outputs of (a) the Canny and (b) the proposed approaches for Fig. 4d. The entropy of the first method is higher than that of the second method. Though, the output of the latter method seems more appropriate.

Four Scenarios were also considered and compared for the thresholding of the force feature histogram. In the first scenario, we applied the Otsu thresholding method to the histogram of the force feature to find the optimal threshold according to which, we classified the pixels to edges and non-edges. The second scenario found the optimal threshold by starting from the mean of the histogram and following the iterative process (12-15). The third scenario directly considered (16) as the optimal threshold of the histogram. The parameters of the model were found using the GOA. Finally, the fourth scenario (i.e., the proposed method) made use of the model in (16) as $T_{initial}$ whose optimal parameters were found using the GOA. Then, the iterative process (12-15) was performed to find the optimal threshold. It should be noticed that, the optimal parameters for Scenarios #3 and #4 are computed independently, by the GOA. The results of applying these four scenarios to four previously mentioned cases evaluated by the average of FOM and Entropy measures are reported in Table 4. It should be noticed that the results of Scenario #4 were

already presented in Tables 1-3. The results established that, the effectiveness of the proposed method is the highest one among the four scenarios in all the cases. The Scenarios #2 and #3 are generally positioned in the second and the third ranks, while the Otsu thresholding (Scenarios #1) yields the lowest rate.

Average Precision

To compare the performance of the proposed approach with some recent edge-detection methods according to the Average Precision (AP) measure, we used the BSDS500 benchmark with clean images; results are summarized in Table 5. It should be noticed that the APs of the other methods are directly reported from the literature. As shown in the results, the proposed method achieves the AP of 0.894 which outperforms all of these competing methods.

5- Conclusion

Various heuristic and nature-inspired techniques for the image-understanding have been proposed in the last decade. In this paper, a new thresholding approach for the gravitational edge detector is designed to obtain an optimal edge image in clean and noisy conditions. The quantitative and qualitative measures confirmed that the performance of the proposed method was higher than a variety of other classic and recent detectors. On the BSDS500 dataset, the proposed method obtained the Average Precision value of 0.894. Experiments on a number of images proved that, this method was reliable and efficient for the edge detection, also it achieved more acceptable results in comparison with other techniques. Furthermore, the similarity of the output images and the ideal edge maps were far more, compared with other detectors. That is to say, the extracted edges using the proposed method had a close Euclidean distance with the ideal edges; as validated by high FOM. Selecting an optimal threshold value in the developed edge detector resulted in covering the weak edges of the image and obtaining the maximum information (entropy) in the output. The main disadvantage of this approach is its sensitivity to noise which can be solved by the pre-processing operations. Using a 5×5 mask decreased the sensitivity to the Gaussian and salt and pepper noises. This, demonstrated the method robustness in noisy conditions, as entropy slightly changes. Furthermore, the higher average precision of our method validated its dominance over some recently published methods. The results of the quantitative and qualitative evaluations of the detected edges showed that the conducted method had merit in detecting edges in noisy conditions among all tested detectors. For further research, it is suggested to use a multi-modal thresholding approach to enhance the segmentation performance in multi-class images. Also, the evaluation of the method of this work in color images can be considered. Additionally, the effect of applying non-square neighboring area will be studied in a further work.

References

[1] R. C. Gonzalez and R. E. Woods, "Digital image processing," ed: Prentice hall Upper Saddle River, 2002.

[2] D. Marmanis, K. Schindler, J. D. Wegner, S. Galliani, M. Datcu, and U. Stilla, "Classification with an edge: Improving semantic image segmentation with boundary detection," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 135, pp. 158-172, 2018.

[3] I. Sobel, "Camera models and machine perception [PhD Dissertation]," ed: Stanford, CA: Stanford University, 1970.

[4] J. M. Prewitt, "Object enhancement and extraction," Picture processing and Psychopictorics, vol. 10, no. 1, pp. 15-19, 1970.

[5] J. Canny, "A computational approach to edge detection," IEEE Transactions on pattern analysis and machine intelligence, no. 6, pp. 679-698, 1986.

[6] L. S. Davis, "A survey of edge detection techniques," Computer graphics and image processing, vol. 4, no. 3, pp. 248-270, 1975.

[7] L. Bin and M. S. Yeganeh, "Comparison for image edge detection algorithms," IOSR Journal of Computer Engineering, vol. 2, no. 6, pp. 1-4, 2012.

[8] L. Romani, M. Rossini, and D. Schenone, "Edge detection methods based on RBF interpolation," Journal of Computational and Applied Mathematics, vol. 349, pp. 532-547, 2019.

[9] S. Eser and A. Derya, "A new edge detection approach via neutrosophy based on maximum norm entropy," Expert Systems with Applications, vol. 115, pp. 499-511, 2019.

[10] Y. He and L. M. Ni, "A novel scheme based on the diffusion to edge detection," IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 1613-1624, 2018.

[11] R. K. Bhogal and A. Agrawal, "Image Edge Detection Techniques Using Sobel, T1FLS, and IT2FLS," in Information and Communication Technology for Intelligent Systems: Springer, 2019, pp. 303-317.

[12] A. Banharnsakun, "Artificial bee colony algorithm for enhancing image edge detection," Evolving Systems, vol. 10, no. 4, pp. 679-687, 2019.

[13] O. P. Verma and A. S. Parihar, "An optimal fuzzy system for edge detection in color images using bacterial foraging algorithm," IEEE Transactions on Fuzzy Systems, vol. 25, no. 1, pp. 114-127, 2016.

[14] M. D. Ansari, A. R. Mishra, and F. T. Ansari, "New divergence and entropy measures for intuitionistic fuzzy sets on edge detection," International Journal of Fuzzy Systems, vol. 20, no. 2, pp. 474-487, 2018.

[15] W. Shen, X. Wang, Y. Wang, X. Bai, and Z. Zhang, "Deepcontour: A deep convolutional feature learned by positive-sharing loss for contour detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3982-3991.

[16] S. Xie and Z. Tu, "Holistically-nested edge detection," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1395-1403.
[17] D. Xu, W. Ouyang, X. Alameda-Pineda, E. Ricci, X. Wang, and N. Sebe, "Learning deep structured multi-scale features using attention-gated crfs for contour prediction," in Advances in Neural Information Processing Systems, 2017, pp. 3961-3970.
[18] R. P. Olenick, T. M. Apostol, D. L. Goodstein, and A. Arons, "The Mechanical Universe: Introduction to Mechanics and Heat," ed: AAPT, 1986.
[19] E. Rashedi and H. Nezamabadi-Pour, "A stochastic gravitational approach to feature based color image segmentation," Engineering Applications of Artificial Intelligence, vol. 26, no. 4, pp. 1322-1332, 2013.
[20] G. Sun, A. Zhang, Y. Yao, and Z. Wang, "A novel hybrid algorithm of gravitational search algorithm with genetic algorithm for multi-level thresholding," Applied Soft Computing, vol. 46, pp. 703-730, 2016.
[21] J. J. D. M. S. Junior, A. R. Backes, and P. C. Cortez, "Color texture classification based on gravitational collapse," Pattern Recognition, vol. 46, no. 6, pp. 1628-1637, 2013.
[22] G. Sun, Q. Liu, Q. Liu, C. Ji, and X. Li, "A novel approach for edge detection based on the theory of universal gravity," Pattern Recognition, vol. 40, no. 10, pp. 2766-2775, 2007.
[23] C. Lopez-Molina, H. Bustince, J. Fernández, P. Couto, and B. De Baets, "A gravitational approach to edge detection based on triangular norms," Pattern Recognition, vol. 43, no. 11, pp. 3730-3741, 2010.
[24] O. P. Verma and R. Sharma, "An optimal edge detection using universal law of gravity and ant colony algorithm," in 2011 World Congress on Information and Communication Technologies, 2011: IEEE, pp. 507-511.
[25] O. P. Verma, R. Sharma, M. Kumar, and N. Agrawal, "An optimal edge detection using gravitational search algorithm," Lecture Notes on Software Engineering, vol. 1, no. 2, p. 148, 2013.
[26] F. Deregeh and H. Nezamabadi-Pour, "A new gravitational image edge detection method using edge explorer agents," Natural Computing, vol. 13, no. 1, pp. 65-78, 2014.
[27] G. Sun et al., "Gravitation-based edge detection in hyperspectral images," Remote Sensing, vol. 9, no. 6, p. 592, 2017.
[28] D. Wang, J. Yin, C. Tang, X. Cheng, and B. Ge, "Color edge detection using the normalization anisotropic Gaussian kernel and multichannel fusion," IEEE Access, vol. 8, pp. 228277-228288, 2020.
[29] J. He, S. Zhang, M. Yang, Y. Shan, and T. Huang, "Bi-directional cascade network for perceptual edge detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 3828-3837.
[30] K. Li, Y. Tian, B. Wang, Z. Qi, and Q. Wang, "Bi-Directional Pyramid Network for Edge Detection," Electronics, vol. 10, no. 3, p. 329, 2021.
[31] Y. Lu, C. He, Y. F. Yu, G. Xu, H. Zhu, and L. Deng, "Vector co- occurrence morphological edge detection for colour image," IET Image Processing, 2021.
[32] S. Anand and G. Sangeethapriya, "Gaussian modulated hyperbolic tangent high pass filter for edge detection in noisy images," arXiv preprint arXiv:2005.11432, 2020.
[33] S. Saremi, S. Mirjalili, and A. Lewis, "Grasshopper optimisation algorithm: theory and application," Advances in Engineering Software, vol. 105, pp. 30-47, 2017.
[34] C. Lopez-Molina, B. De Baets, and H. Bustince, "Generating fuzzy edge images from gradient magnitudes," Computer Vision and Image Understanding, vol. 115, no. 11, pp. 1571-1580, 2011.
[35] N. Otsu, "A threshold selection method from gray-level histograms," IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1, pp. 62-66, 1979.
[36] M. Mafarja, I. Aljarah, H. Faris, A. I. Hammouri, A.-Z. Ala'M, and S. Mirjalili, "Binary grasshopper optimisation algorithm approaches for feature selection problems," Expert Systems with Applications, vol. 117, pp. 267-286, 2019.
[37] X. Zhang, Q. Miao, H. Zhang, and L. Wang, "A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery," Mechanical Systems and Signal Processing, vol. 108, pp. 58-72, 2018.
[38] U. SIPI, "The usc-sipi image database," ed, 2016.
[39] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 5, pp. 898-916, 2010.
[40] S. E. Umbaugh, Computer imaging: digital image analysis and processing. CRC press, 2005.
[41] C. E. Shannon, "A mathematical theory of communication," Bell system technical journal, vol. 27, no. 3, pp. 379-423, 1948.
[42] R. Xiaofeng and L. Bo, "Discriminatively trained sparse code gradients for contour detection," in Advances in neural information processing systems, 2012, pp. 584-592.
[43] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Convolutional oriented boundaries," in European Conference on Computer Vision, 2016: Springer, pp. 580-596.
[44] I. Kokkinos, "Pushing the boundaries of boundary detection using deep learning," arXiv preprint arXiv:1511.07386, 2015.
[45] P. Isola, D. Zoran, D. Krishnan, and E. H. Adelson, "Crisp boundary detection using pointwise mutual information," in European Conference on Computer Vision, 2014: Springer, pp. 799-814.
[46] G. Bertasius, J. Shi, and L. Torresani, "High-for-low and low-for-high: Efficient boundary detection from deep object features and its applications to high-level vision," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 504-512.
[47] Y. Liao, S. Fu, X. Lu, C. Zhang, and Z. Tang, "Deep-learning-based object-level contour detection with CCG and CRF optimization," in 2017 IEEE International Conference on Multimedia and Expo (ICME), 2017: IEEE, pp. 859-864.
[48] W. Yupei, Z. Xin, and K. Huang, "Deep crisp boundaries," 2017: CVPR.
[49] S. Hallman and C. C. Fowlkes, "Oriented edge forests for boundary detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1732-1740.

Kimia Rezaei received her BSc and MSc from Islamic Azad University, Fasa and Shiraz Branches, Fars, Iran. She is currently a PhD student at Department of Electrical and Electronic Engineering, University College Cork, Ireland. Her research interests are pattern recognition and Image processing for practical applications.

Hamed Agahi received B.Sc. M.Sc. and Ph.D. degrees in Electrical Engineering from University of Shiraz, Amirkabir University of Technology and University of Tehran, Iran, in 2005, 2008 and 2013, respectively. From 2009, he was with the Islamic Azad University, Shiraz Branch, Shiraz, Iran. His research interests include pattern recognition, image and signal processing, and fault detection and diagnosis applications.