International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Facial Recognition with the Super Pixel Entropy Estimation with the Virtual Assistance System

Tanuja Dhope[1], Dr. Rajneesh Tyagi[2], Dr. Vidya Nitin Patil[3], Dr. Veeramani Vijaya Raghavan[4], Dr. Tarun K. Sharma[5], Dr. Usha C. Pawar[6]

Submitted: 28/08/2022 **Accepted:** 20/11/2022

Abstract: Face recognition system is the developing technology for the recent years. However, the traditional facial recognition system subjected to vast range of challenges due to variation in the structural features, color, and variation in the face. Traditionally, the facial recognition system comprises of the facial features with the landmark estimation. Bur those technique are not sufficient enough to extract the appropriate pixel in the identification of the features. This paper developed a Facial Super Pixel Entropy (FSPE) integrated with the virtual assistance for the identification of faces. The proposed FSPE model uses the segmentation of the super pixel in the facial images with the consideration of the entropy estimation. The performance of the proposed FSPE model is evaluated for the consideration of the different feature model for the processing, The analysis expressed that the proposed FSPE model achieves the segmentation accuracy of 99% and the accuracy is achieved as the 99%. This implies that propsoed FSPE model is effective for the facial recognition system with the virtual assistant.

Keywords: *Facial Recognition, Super pixel, Virtual Assistance, Machine Learning, Entrophy*

1. Introduction

Recognition System is a computer technology related to Image processing that detects and recognizes objects in images. Many recent advancements in technology like surveillance system [1], Face Detection system, Medical Imaging [2] and Autonomous vehicles are depending heavily on accuracy of Recognition System. These systems classify, localize and recognize every object in the image. These systems use existing classification models as backbone for Recognition System. Recognition System [3] is the method of identifying instances of real world artifacts such as human beings, animals and conveyances in images or videos [4]. It identifies the object's class by assigning a label and performs the process called Object Localization. Region-Based Convolutional Neural Networks (R-CNNs), are one of the most widely used networks for Object Localization.

Profound Learning Models are mostly utilized in Recognition System Algorithms due to their exact Image Recognition ability. Progression in Deep Learning is because of the headway of Convolution Neural Network Architectures and the execution of new calculations. Profound Learning is a sub-part of Machine Learning that gains highlights from the information [5]. The productivity of Deep Learning calculations surpassed the Machine Learning calculations as how much data is developing. "Geoffrey Hinton" gave a forward leap by effectively preparing these organizations. The Deep Learning based Recognition System approach has advantages, for example, decrease in preparing time because of the utilization of the current pre-prepared grouping network as the base organization and the improvement of a total start to finish Deep Learning-based object locator [6]. Acknowledgment System utilizing Deep Learning Framework at first chooses a structure from the accessible writing specifically Faster Region based Convolution Neural Network (R-CNN) [7], Single Shot Detector (SSD), You Only Look Once (YOLO) and so forth and these systems are executed on a reasonable Base Network.

Recognition System's future has tremendous potential across large number of areas. The need for precise Recognition System has thus become an important area of study. While the literature includes several Recognition System algorithms, none of them addresses the optimal combination of Hyper-Parameters to achieve maximum Recognition System accuracy [8]. This present research work focuses primarily on the optimal range of Hyper-Parameters to detect an object with much better accuracy than the stateof-the-art algorithms. Optimal Recognition System accuracy zones are identified in such a way that the accuracy reaches its maximum value in that region, by combining different parameters. The proposed model focused on computation of the images super pixels for the segmentation and extraction of the features. Within this work, the Fine-Tuning techniques are applied to many Datasets, and the relationship between Dataset Sizes and other Hyper-Parameters is also established.

[1] Electronics and Communication, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India.
Email: tanuja_dhope@yahoo.com, 0000-0003-2907-8509
[2] Dean, Department of Agriculture, Sanskriti University, Mathura, Uttar Pradesh, India. Email: dean.soa@sanskriti.edu.in. 0000-0002-8571-6221
[3] Professor, Department of Civil Engineering, AISSMS COE PUNE, India vnpatil@aissmscoe.com
4 ECE, Vignan's Foundation for Science, Technology and Research, Guntur, India. Email: vijayaraghavan123@gmail.com
[5] Professor & Dean, Department of Computer Science, Shobhit Institute of Engineering & Technology (Deemed to-be University), Meerut, India. Email: tarun.sharma@shobhituniversity.ac.in. 0000-0002-9043-8641
[6] Assistant Professor, Mechanical Engineering, Datta Meghe College Of Engineering, Maharashtra, Navi Mumbai, India.
Email:usha.pawar@dmce.ac.in, 0000-0002-6516-5354

2. Related Works

In different fields there is a need to identify the objective item and furthermore track them really while dealing with impediments and other included intricacies. Numerous analysts endeavored for different methodologies in object following. The idea of the strategies generally relies upon the application space. A portion of the exploration works which made the development to proposed work in the field of item following are portrayed as follows. Coming up next is a rundown of a few papers to underscore the exploration works completed on Recognition System utilizing Deep Learning. In [9] proposed an original technique for vehicle occasion division, which requires distinguishing, at the pixel level, where the vehicles show up and connecting every pixel with an actual occurrence of a vehicle. This work fostered a brought together perform multiple tasks learning network that learns two reciprocal errands, which are fragmenting vehicle districts and identifying semantic limits. Likewise, another Dataset for vehicle case division, to be specific, Busy Parking Lot Unmanned Aerial Vehicle Video is produced for future exploration purposes.

In [10] proposed a pixel characterization strategy which utilizes division procedure to distinguish fixed objects in the picture. These items are then followed utilizing another method called versatile edge direction. This strategy recognizes objects with a precision of 95% which is an improvement over best in class techniques. In [11] overviewed different Image change location calculations and talked about the standards for contrasting execution of progress recognition calculations.

In [12] examines different Deep Learning-based Recognition System structures. This work centers around refinements that can be applied on Recognition System structures to further develop their exhibition further. Likewise, this work gave a few promising headings for planning a superior Object Detector. Ross In [13] proposed a Fast Region-based Convolutional Neural Network (Fast R-CNN) for Recognition System. This strategy became famous on account of its speed when contrasted with the cutting edge procedures. Quick RCNN trains VGG-16 organization multiple times quicker than RCNN and tests similar organization multiple times quicker than RCNN. In [14] introduced a preparation methodology that utilizes information expansion to utilize the accessible pictures. This system gave best outcomes even with not many preparation pictures and beat the first best on the ISBI challenge for division of neuronal designs in electron tiny stacks.

In [15] proposed a revolution invariant technique for identifying geospatial objects from satellite pictures. Three procedures were utilized particularly in the proposed strategy. Super pixel division procedure, right off the bat, was proposed to create non-repetitive patches and in the second stage a multi-facet profound component age model was created to produce significant level element portrayals of patches and in the last stage a bunch of multi-scale Hough backwoods with implanted fix directions was built to project turn invariant decisions in favor of assessing object centroids. This work gave generally excellent outcomes and beat the cutting edge strategies.

In [16] gave another methodology for identifying objects by layout matching from the huge data set assortment. The methodology was reasonable for multi-scale contrast foundations and utilized a variety spatial strategy to distinguish objects. This approach neglected to distinguish different items in a given client situation and furthermore fizzled when the articles were in non-direct movement. This issue of bombed identifications was really beaten in the proposed framework via preparing the framework to distinguish the articles through a powerful framework learning strategy and the items in non-direct movement are followed utilizing the proposed molecule gathering approach. In[17] proposed an original way to deal with become familiar with a RotationInvariant CNN (RICNN) model for propelling the presentation of Recognition System in VHR Optical Remote Sensing Images. This errand is accomplished by presenting another pivot invariant layer. RICNN model is prepared by enhancing another goal capability which expressly authorizes the element portrayals of the preparation tests when turning to be planned near one another, thus accomplishing pivot invariance. The proposed work is tried on a freely accessible ten-class Recognition System Dataset and gave good outcomes.

3. Facial SuperPixel Entropy (FSPE) Recognition System Framework

The proposed FSPE Layer Properties has the information about layer name, layer type and layer connection structure. FSPE communicates everything in the form of blobs. A blob is a unified memory interface and is a standard array. It is an N dimensional array which is used for processing data. It provides the synchronization between CPU and GPU. The blob gives details about how the information is stored and communicated between layers and nets. Blobs hide the computational and mental overhead of CPU/GPU operation by 42 synchronizing between the CPU host to the GPU device using ‘SyncedMem’ class. Examples of blobs are batches of images, model parameters, and derivatives for optimization. Layer Parameters are the various convolution layer parameters such as Kernal_size, Stride etc. These parameters are necessary for learning features of an image when it is passed through the FSPE Recognition System Framework. The figure 1 provides the tensor flow graph of the proposed FSPE model is presented.

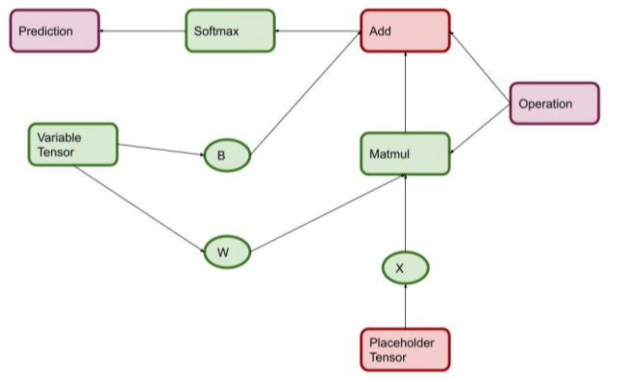

Figure 1: Tensor Flow in FSPET

Tensorflow framework helps in building Neural Network models to classify images. These networks are made up of Convolutional Neural Networks (CNN). Tensorflow Recognition System procedure makes use of a series of steps. They are

1. Data Pre-Processing
2. Reshape input if necessary to match the convolutional layer that is intended to be built.
3. Convolutional layer is created using tf.nn.conv1d() method
4. After convolution operation pooling operation is applied. Pooling layer is created using tf.nn.maxpool()
5. . Steps 3 and 4 are repeated as more and more convolution and pooling layers are added.
6. Outputs from convolution and pooling layers are reshaped and they are applied as input to flatten layer and the output

from the flatten layer is passed through the fully connected layer.

7. Fully Connected Layer (FC) is created using tf.matmul() function and an activation function such as ReLu is added using tf.nn.relu() and finally, a Dropout operation is also applied using tf.nn.dropout() method.
8. . tf.matmul() method is used for final class prediction
9. Lastly, weights and biases are stored by the Tensorflow variables.

Algorithm 1: FSPE model for the facial recognition
Initially set Epoch rate ER = 1, PAavg = 0, PAsum = 0, a
// a is an array
while ER<=4 //ER – Epoch rate
compute PA% and store the value in array 'a'
Increment ER
end while // PA% - % of Prediction Accuracy
do //PAavg – % of Average of PA
increment ER // PAsum – Sum of PA
compute PA% and store it in array 'a'
compute the average of last four PA% - PAavg
Subtract newly computed PA% from PAavg
While difference > 3%
end while
print PA%

3.1 Segmentation of the Super Pixel

The proposed FSPE model uses the pixel wise distributions $\widehat{a_k}$ at Superpixel k with the computation of the feature vectors F_i using a ReNet network. The equation (1) - (3) provides the prediction feature calculation

$$y_i = w_2 tanh(w_1 F_i + b_i) \quad (1)$$

$$\widehat{a_{i,k}} = \frac{e^{yi,a}}{\sum_{b\in classes} e^{yi,a}}$$

$$L_{cat} = -\sum_{i\in pixels} \sum_{a\in classes} a_{i,d} ln\, \widehat{a_{i,k}} \quad (2)$$

$$\widehat{a_{i,k}} = \frac{1}{s(k)} \sum_{i\in k} \widehat{a_{i,k}}$$

where a_i is the ground truth distribution at location i, and $s(k)$serves as the surface of the component k. The class features of the proposed FSPE model is presented in equation (3)

$$l_k = argmax_{a\in classes}\, \widehat{a_{i,k}} \quad (3)$$

The l_i local variable classifier independently at each site i, provided the filter outputs xi are the pixel centre as shown in Equation (4)

$$p_c\left(\frac{L}{x}, \lambda\right) = \pi_x P_c\left(\frac{l_i}{x_i}, \lambda\right) \quad (4)$$

where λ denotes the classifier parameter.
The differential features in the facila point features are computed using the equation (5)

$$P_R(L, f) \propto \exp\{\textstyle\sum_{r,a} f_{x.a}\, w_a^T i_r\} \quad (5)$$

where $f = \{f_{r,a}\}$represents the hidden region binary variable, $w_a = [w_a, 1 \ldots\ldots, w_a, J, \alpha_a]$, $Ir = [lr, 1, \ldots\ldots, lr, J, 1]$and $\alpha_a$represents a bias term. The patches of the label field are represented as a coarse aspects of the label field. These patches are non-overlapping patches $p_m, m \in \{1, \ldots\ldots, M\}$ and for each hidden global variable g_b is presented in equation (6)

$$P_g(L, g) \propto exp\{\textstyle\sum_b g_b u_b^T L\} \quad (6)$$

The integration of the global features of the proposed FSPE mode is presented in equation (7)

$$P\big((L|X;\theta)\big) = \frac{1}{z}\pi_i P_c^{\gamma}(l_i|X_i, \lambda) X \pi_{r,a}[1 + exp(w_a^T l_r)]X[1 + exp(u_b^T L)] \quad (7)$$

where $\theta = \{\lambda, \{w_a\}, \{u_b\}, \gamma\}$is the set of parameters in the model.
The depth in the proposed model is computed based on the equatin $y_j = softmax\left(F(h_j; W_{label})\right)$ where $y_j$is the predicted geometric features by the j^{th} pixel, and $W_{label}$is the network parameter. The transformation function is represented as $F(.)$.
The proposed FSPE model uses the MSE variable calculated using the equation (8)

$$MSE = \sum_{y=1}^{M} \sum_{x=1}^{N} [I(x.y) - I'(x,y)]^2 \quad (8)$$

where, $I(x.y)$ is image values, $I'(x,y)$noise in an images M, N are the dimension of the error images.
The PSNR value is calculated using the Equation (9)

$$PSNR = 10log_{10}(MAXi^2/MSE) \quad (9)$$

where, MAXi reperesents the maximal pixel values.

4. Experimental Results

The performance of the proposed FSPE model the facial recognition dataset is considered for the varying dataset size of Dataset 1, 2, 3 and 4. With the evaluation of the training and testing epochs the features are computed based on the consideration of the hyper parameters. The table 1 provides the attributes considered for the proposed FSPE dataset evaluation is presented. The figure 2 and figreu 3 provides the training and testing process in dataset is presented.

Table 1: Dataset Attributes

Dataset	Training Dataset Size	Testing Dataset Size	Object Classes
Dataset-1	163	37	9
Dataset-2	437	83	9
Dataset-3	1363	286	9
Dataset-4	6974	876	53

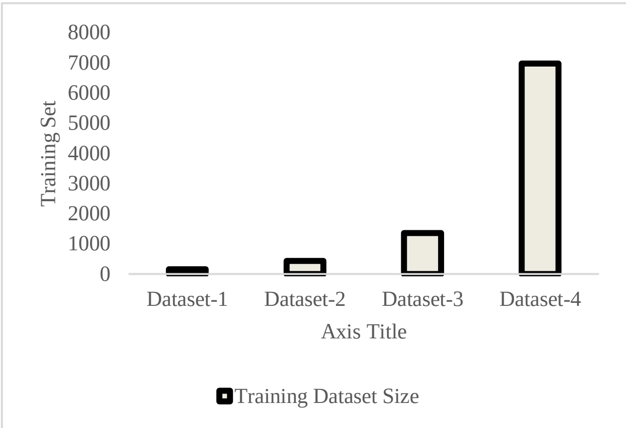

Figure 2: Training Dataset

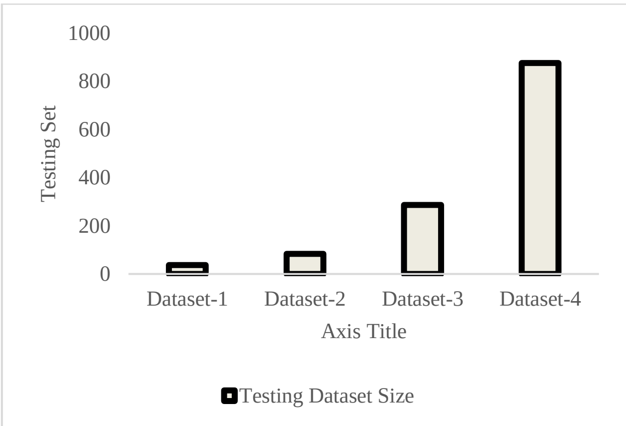

Figure 3: Testing Dataset

Based on the consideration of the attributes the hyper parameters for the computation of the variables with the proposed FSPE mode is presented in table 2.

Table 2: Hyper parameters

Hyper-Parameters	Values
Learning Rate	10-4
Drop-out rate	0.5
Activation	Relu
Epoch count	9
Training Layers	Top-4 Layers

The performance of the proposed FSPE mode is evaluated for the consideration of the different facial images for the analysis. The performance emetrics for the proposed model is evaluated and examined with the consideration of the different variables presented in table 3.

Table 3: Comparison of Performance

Image	MSE	PSNR	Specificity	Classifier accuracy	Pixel accuracy
Image 1	1.78	46.75	99.24	98.65	99.24
Image 2	1.45	46.83	99.46	98.76	98.67
Image 3	1.29	46.53	99.87	97.63	99.25
Image 4	1.34	46.38	98.36	99.67	98.75
Image 5	1.28	46.28	99.87	99.34	99.56

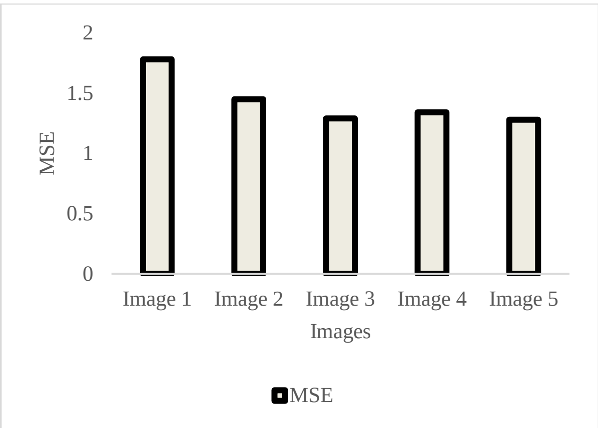

Figure 4: Computation of MSE

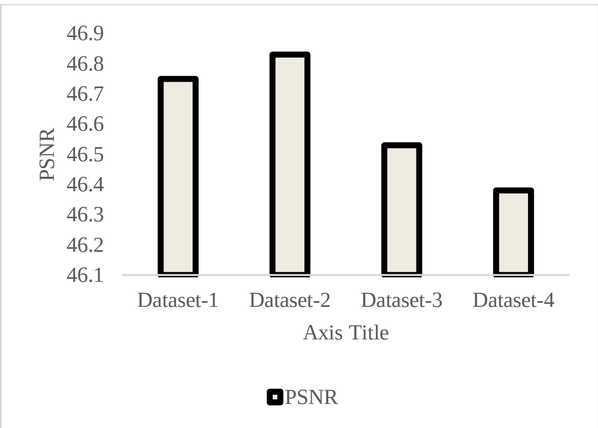

Figure 5: Measured PSNR

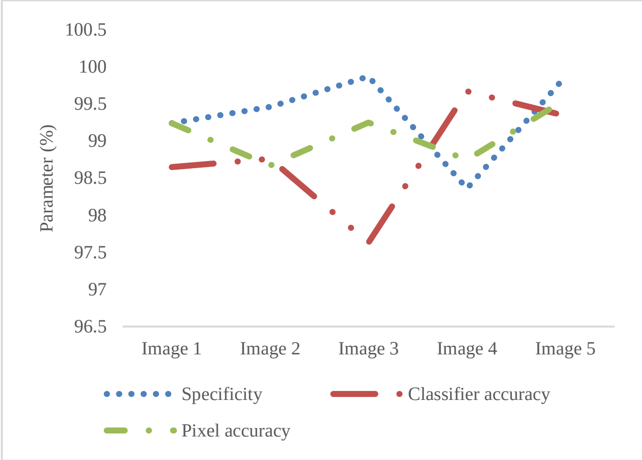

Figure 6: Comparison of Parameters

The figure 4 – 6 the performance of the proposed FSPE for the different variable in the images are presented. The estimation of the facial recognition is evaluated with the virtual system for the consideration of te three different datasets as presented in table 4.

Table 4: Comparison of Pixel Accuracy

NNM	Dataset-1		Dataset-2		Dataset-3	
	PA%	WP	PA%	WP	PA%	WP
VGG 16	93.45	11	82.45	19	97.67	2
VGG 19	95.67	8	81.56	21	95.74	5
InceptionV3	98.67	6	89.56	8	98.78	2
Xception	99.56	1	97.87	2	99.78	0

In table 4 the propsoed FSPE model pixel accuracy is evaluated that the proposed FSPE model perform significantly with the Xception layer compared with the VGG 16, VGG 19 and Inception V3 mdoel. With the Xception model the proposed FPSE achieves the maximal value of 99.78% which is signficnatly higher that the other models.

5. Conclusion

This paper presented a virtual assistant facial recognition system with the super pixel computation in the features. The proposed FSPE model computes the super pixel in the images and computes the entrophy. Upon the estimated superpixel the features are computed as processed. The propsoed FSPE model performance is evaluated for the varying architecture model. The comparative analysis expressed that proposed FSPE model achieves the accuracy of 99% of the accuracy and the segmentation accuracy of 99%. The peroposed FSPE model perform effectively with the Xception layer archieetcture model.

References

[1] Xia, J., Yuan, X., Li, J., Lu, S., Cui, X., Li, S., & Fernández-Ramírez, L. M. (2020). Foreign Recognition System for electric vehicle wireless charging. Electronics, 9(5), 805.

[2] Coe, J., & Atay, M. (2021). Evaluating impact of race in facial recognition across machine learning and deep learning algorithms. Computers, 10(9), 113.

[3] Raju, K., Chinna Rao, B., Saikumar, K., & Lakshman Pratap, N. (2022). An Optimal Hybrid Solution to Local and Global Facial Recognition Through Machine Learning. In A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems (pp. 203-226). Springer, Cham.

[4] Singhal, P., Srivastava, P. K., Tiwari, A. K., & Shukla, R. K. (2022). A Survey: Approaches to facial detection and recognition with

machine learning techniques. In Proceedings of Second Doctoral Symposium on Computational Intelligence (pp. 103-125). Springer, Singapore.

[5] Salama AbdELminaam, D., Almansori, A. M., Taha, M., & Badr, E. (2020). A deep facial recognition system using computational intelligent algorithms. Plos one, 15(12), e0242269.

[6] Asaithambi, S. P. R., Venkatraman, S., & Venkatraman, R. (2021). Proposed big data architecture for facial recognition using machine learning. AIMS Electronics and Electrical Engineering, 5(1), 68-92.

[7] Finizola, J. S., Targino, J. M., Teodoro, F. G., & Lima, C. A. (2019, July). Comparative study between deep face, autoencoder and traditional machine learning techniques aiming at biometric facial recognition. In 2019 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

[8] Boulos, M. M. (2021). Facial recognition and face mask detection using machine learning techniques.

[9] Akter, T., Ali, M. H., Khan, M., Satu, M., Uddin, M., Alyami, S. A., ... & Moni, M. A. (2021). Improved transfer-learning-based facial recognition framework to detect autistic children at an early stage. Brain Sciences, 11(6), 734.

[10] Van Noorden, R. (2020). The ethical questions that haunt facial-recognition research. Nature, 587(7834), 354-359.

[11] Raji, I. D., & Fried, G. (2021). About face: A survey of facial recognition evaluation. arXiv preprint arXiv:2102.00813.

[12] Sarkar, E., Benkraouda, H., & Maniatakos, M. (2020). FaceHack: Triggering backdoored facial recognition systems using facial characteristics. arXiv preprint arXiv:2006.11623.

[13] Nachmani, O., Saun, T., Huynh, M., Forrest, C. R., & McRae, M. (2022). "Facekit"—Toward an Automated Facial Analysis App Using a Machine Learning–Derived Facial Recognition Algorithm. Plastic Surgery, 22925503211073843.

[14] Fábián, I., & Gulyás, G. G. (2020). De-anonymizing facial recognition embeddings. Infocommunications Journal, 12(2), 50-56.

[15] Singh, N. S., Hariharan, S., & Gupta, M. (2020). Facial recognition using deep learning. In Advances in Data Sciences, Security and Applications (pp. 375-382). Springer, Singapore.

[16] Omoyiola, B. O. (2018). Overview of biometric and facial recognition techniques. IOSR journal of computer engineering (IOSRJCE), 20(4), 1-5.

[17] Finizola, J. S., Targino, J. M., Teodoro, F. G. S., & Moraes Lima, C. A. D. (2018, November). A comparative study between deep learning and traditional machine learning techniques for facial biometric recognition. In Ibero-American Conference on Artificial Intelligence (pp. 217-228). Springer, Cham.