International Journal of Engineering Research and Sustainable Technologies
Volume 2, No.2, June 2024, P 26-35
ISSN: 2584-1394 (Online version)

FACE RECOGNITION USING CNN

D.Gayathry[1], R Latha[2]
[1]Research Scholar, [2]Professor and Head, Department of Computer Science and Applications
[1,2] St. Peters Institute of Higher Education and Research, India
*** Corresponding author email address**: gaya200689@gmail.com

https://doi.org/000000/000000/

Abstract

The use of face recognition is growing at an astounding rate these days. Currently, researchers are developing many methods for how a facial recognition system functions. People often find themselves alone in their families in situations such as normal disasters, kidnappings, accidents, missing persons cases, and many more situations. To reach the family of those refugees and ensure their safety and support, it is imperative to identify their relatives. Every day, police departments enroll missing cases. Some of these enlisted cases are resolved, but not all of them are resolved by employing the labor-intensive manual approach. This study aims to address the time delay caused by current police examination procedures by leveraging the latest innovations. Therefore, we implement a framework that makes use of the CNN (Convolutional Neural Network) technique and the VGG16 architecture. We begin with our input dataset, which consists of 84 photos taken from 21 different households. The final dataset, obtained by applying the enhancement method, consists of 1512 photos, of which 80% are utilized for training and 20% are for testing. This framework offers a quick and easy method for locating a refugee's personal and helps validate an individual's identity by using their picture and family information with related models that are more accurate.

Keywords: Face recognition, CNN, VGG16, Deep learning

1. Introduction

It is important to recognize that a person may become a refugee due to a variety of uncontrollable conditions. Those who have lost family members to various causes and, in addition, lack a legal means of gathering their belongings and traveling to rejoin them are considered refugees. Police departments are receiving a large number of refugee cases on a daily basis. Ten years ago, these cases were handled by police using a fairly manual approach that involved gathering family information first and then conducting investigations to locate the family. Most instances were not resolved during the lengthy manual process.

Later, as technology advanced, a plethora of new techniques arose to address issues arising from the manual approach. Of all the methods currently in use for identifying people and resolving refugee situations, face one of the most practical often used technologies is that of detection and well-thought-out methods. A face recognition system is a piece of technology that can compare a digital image or video frame containing a human face to a database of faces [5]. Accurate face identification is still a difficulty, and face acknowledgement systems differ in their capacity to identify individuals. The interest in face acknowledgement is growing for a number of reasons. The primary fields that have contributed [11]to the development of facial recognition systems are pattern recognition, machine learning, and deep learning [1]. One type of cutting-edge technique that can provide excellent facial recognition performance is deep learning. CNN (Convolutional Neural Network)-based face recognition technology has emerged as the industry standard in the field of face recognition thanks to the advancements in deep learning.

[19] CNNs are a subclass of [15] deep neural networks that are mostly used for visual vision analysis. [9] They are composed of three layers: the convolutional layer, the pooling layer, and the fully associated layer. For feature extraction, the pooling layer and convolutional layer are used. For grouping, the fully associated layer is used. Experts and researchers require industry or company data in order to do classical visual recognition. With the help of Convolutional Neural Networks (CNNs)[14,17], automated feature engineering replaces the laborious process of human feature extraction. Because CNN has a high accuracy rate, we used it to identify the family members of the refugees using photos as our input data.

2. Related Works

Our method is comparable to other current studies, The face recognition field was dominated in the mid-1990s and early 2000s by nearly high quality methodology and all-encompassing learning approaches, which produced subpar outcomes for unrestricted face changes and in stimulating conditions such as low light, poor image quality, and ideal viewing angle since these methods employed several descriptors of feature. The 2012 ImageNet Large Scale Visual [1,15] Recognition Challenge (ILSVRC) was won by AlexNet, by employing a deep learning strategy to reduce the top-5 error from 26% to 15.3% on ImageNet [8], resulting in a top-1 error rate of 37.5% [15]. One kind of deep learning technique is the convolutional neural network (CNN), which extracts and transforms features using many layers of feature descriptors [2,5]. The fundamental traits of the face are extracted by the early layers, and the finer details are extracted by the subsequent layers. Another approach, termed DeepID, was proposed in 2012[4,6]. It made use of an [11,1] ensemble of smaller and shallower deep convolutional networks than DeepFace. [11] This method was regarded as the first to obtain high accuracy, approximately 90%, on the LFW (Labeled Face in-the-Wild) dataset. FaceNet, a CNN-based method, was proposed in 2015. It used triplets of roughly aligned matching/nonmatching face patches using an online [14] triplet mining method. It achieved state-of-the-art face recognition performance and achieved [14,4]99.63% on LFW and 95.12% on YouTube Faces DB using 128 embedding's preface. In order to evaluate facial recognition performance, ResNet50 [5] trained on [11] VGGFace2, MSCeleb-1M, and their union using residual blocks and residual connections. 2018 saw VGG16 reach excellent performance and accuracy in real-time face recognition, while also being effective on embedded devices.

Table 1. Literature survey table

Sno.	Paper Referred	Published year	Dataset	Method	Accuracy	Review
1	1	2014	Kin Face V2 & Family 101	LBP&KNN	90	Accuracy decreases with change in resolution
2	2	2017	ORL	CAFFE	99.82	Change in Posture Reduces
			AR Face		99.78	Accuracy
3	3	2015	LFW	SVM	72.9	Less Accuracy for many Classes
4	4	2018	CK+	INCEPTION V3	76.53	Over fitting and underfitting issues
				VGG19	86.20	
				VGG Face	91.37	
5	5	2017	GEORGIA TECH FACE	CNN	98.08	Better Computational Speed and Feature extraction

Table 2. Some existing government portals worldwide

Sno	Portal Name	Established year	Dataset	Comments
1	IDMC	1998	Collects information through NGO's Research	This gives only information regarding the disaster and its counts
2	NCMEC	2018	Collects information regarding children	Ment only for children
3	National Tracking system for Missing & Vulnerable Children	2009-10	Checks for child welfare	Checks for child details in police complaints only

3. System Methodology

The [1]26% to 15.3% on ImageNet, achieving a top-1 error rate of 37.5% using a deep learning technique.One kind of deep learning technique is the [17,8]convolutional neural network (CNN), which extracts and transforms features using many layers of feature descriptors [2,5]. The fundamental traits of the face are extracted by the early layers, and the finer details are extracted by the subsequent layers. Another approach, termed DeepID, was proposed in 2012[4,6]. It made use of an collective of smaller and shallower [11,1] deep convolutional networks than DeepFace. [11] This method was regarded as the first to obtain high precision, approximately 90%, on the LFW dataset. FaceNet, a CNN-based method, was proposed in 2015. It [3] used triplets of roughly aligned matching/nonmatching [27] face patches using an online triplet mining method. It achieved state-of-the-art face recognition performance and achieved [14,4]99.63% on LFW and 95.12% on YouTube Faces DB using 128 embedding's preface. In order to evaluate facial recognition performance, ResNet50 [5] trained on [11] VGGFace2, MSCeleb-1M, and their union using residual blocks and residual connections. 2018 saw VGG16 reach excellent performance and accuracy in real-time face recognition, while also being effective on embedded devices.

3.1 SIFT

The use of features in face recognition is widespread. We have downsized all face photos to 64 × 64, as per [16], and we have [1] set the block size to 16 x 16 with an 8 stride. Each image thus has a total of 49 blocks, resulting in a feature vector of 128 × 49 = 6, 272D.

3.2 LBP

Because it captures an image's appearance in a tiny, nearby community surrounding a pixel, [1]it is often used for texture analysis and face identification.

3.3 VGG-Face CNN

Employs a "Very Deep" architecture with convolutional stride (one pixel) and very small convolutional kernels (three by three). The CNN model's [1,7] second-to-last fully-connected layer [11], or fc7, received approximately 2.6 million pre-trained photos of 2, 622 celebrities. Each of the face picture was then scaled to 224 × 224 and fed forward, creating 4, 096D feature vectors [1].

3.4 IDMC

The most reliable origin of information and examination of internal displacement worldwide is [1]the Internal Displacement Monitoring Centre (IDMC). Since its founding as a division of the [11]Norwegian Refugee Council in 1998, IDMC has provided the global community with a thorough, dependable, and independent service.

3.5 National Tracking System For Missing & VulnerableChildren

Nationwide monitoring program for children who are missing or in danger. The goal of [11,1]the centrally funded Integrated Child Protection Scheme[1] is to help children in challenging situations become well. The Indian government's Ministry of Women and Child Development[24] is in charge of carrying out the program..

4. Proposed System

Two steps can be used to construct the facial recognition system. The technique of extracting or picking up the face features is the first phase, and pattern classification is the second. Figure 1 depicts our facial recognition system and details every step of the procedure.

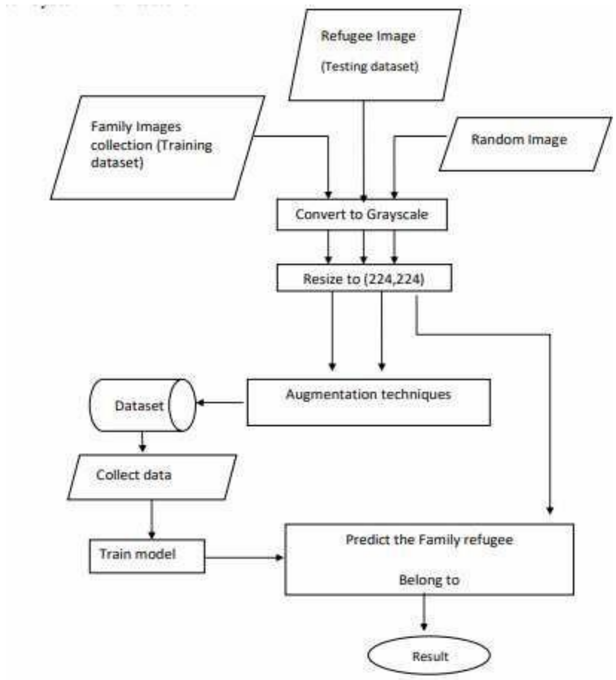

Fig 1. System Architecture

4.1 Dataset

Our information is a raw information set made up of 84 images that we increased to 1512 by applying an augmentation method. Out of this dataset,[11] 80% is used for training and 20% is used for testing. The photographs are of each family member, along [1]with their names, contact information, and address, and were collected from 21 families[1].

```
ip=input("enter name: ")
img=cv2.imread("/content/gdrive/My Drive/MAJOR_PROJECT/RESIZE_FAM
/"+a+"_family/"+a+"_family"+str(j)+".jpeg",0)
gray_image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
image=cv2.resize(image,(224,224))
```

4.2 Data Pre-Processing

Every family has a unique folder established. Converting RGB to Greyscale and Dipping

[11,1] The Python code for data pre-processing is shown in the above code snippet. Every family photo is saved as a separate folder, the name of the folder is saved as the family name that is entered, and each and Each picture in the folder has a caption and overwritten after processing. After reading the image, the first step is to use cv2 to convert the RGB image to grayscale (Figure 2 and Figure 3). Built-in Cvt Color functionality.

Fig 2. RGB to Grayscale

Fig 3. Shifted Image

The image is then adjusted to have a dimension of (224,224) in the second stage. [11,1] The input image dimension must be (224,224) since the VGG16 model requires a constant image size. Image augmentation is the method used to create the dataset.

A method called "image data augmentation" can be used to make altered versions of the [6] dataset's photos in order to artificially increase the size of a training dataset.

More data can be used to train [6] deep learning neural network models, which can provide more sophisticated models. Additionally, augmentation techniques can produce different picture versions, which can enhance the fit [13] models' capacity to apply what they have learnt to new images. Through the use of the [13,10] Image Data Generator class, the Keras deep learning neural network framework offers the ability to fit models using augmented image data.

Shift: Enhancement of Straight and Perpendicular Shift A shift to an image is just moving all of its pixels in one direction, such as vertically or horizontally, while maintaining the image's dimensions.

Rotation: Augmenting Random Rotation An image is rotated right-handed by a stated number of degrees, ranging from 0 to 360, at random via a rotation augmentation using Figure 4.

It is probable that the rotation will cause pixels to be rotated out of the picture frame, leaving blank spaces in the frame that need to be filled in.

Zoom: Random Zoom Augmentation: This technique involves randomly enlarging the image, interpolating pixel values, or adding additional pixel values around the image[19]. The zoom_range parameter to the Image Data Generator constructor can be used to arrange image zooming by Figure 5.

Brightness: Random Enhancement of Brightness. Here, images that are either randomly brightening, darkening, or both are added to the image. The idea is to enable a model to make generalizations across pictures that were trained in varying lighting conditions.

Shear: One of the following can be used to specify the [11]range of flat shear that is applied to the input image. Shear is expressed as an angle with a range of (–90, 90) degrees. Two-element vector of numbers.

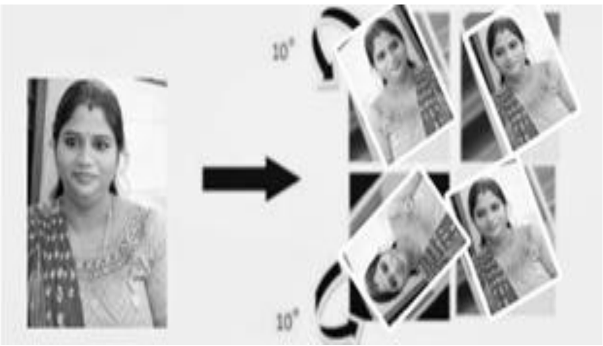

Fig 4. Rotated Image

Fig 5. Zoom Image

4.3 VGG16 Model

The convolution layers in the above Figure 4 and Figure 5 are represented by all of the blue rectangles[11,1]along with the non-linear instigation role, which is a rectified linear unit, or ReLU[1]. Thirteen blue and five red rectangles, or thirteen convolution layers and five max-pooling layers, are visible in the figure. Three green rectangles, which stand for three fully connected layers, are present in addition to these. Thus, the name VGG-16 (Figure 6)refers to the total of [11]16 layers with tunable parameters, of which 13 are convolution layers and 3 are completely connected layers. A softmax layer with 1000 outputs for each [1]image category in the image Net dataset is present at the output.

The channel size in this architecture was first set very low at 64 and was then progressively enlarged [1]by a factor of 2 after each max-pooling layer until it reached 512.[11].

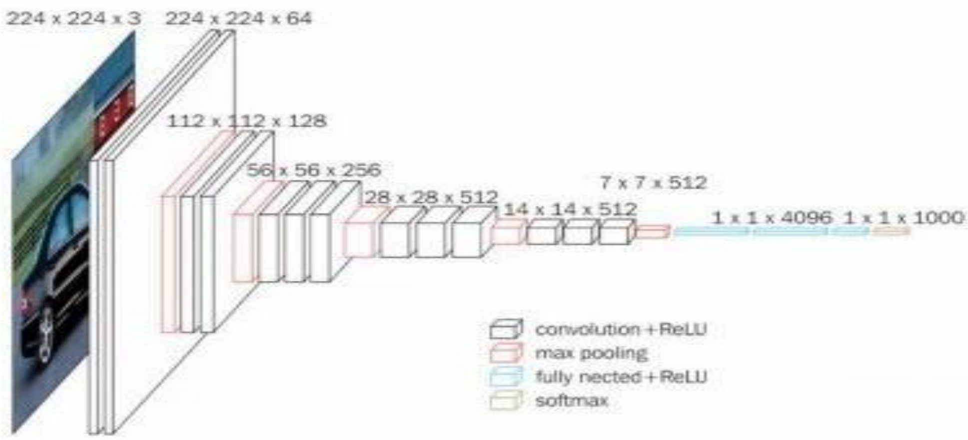

Fig 6. VGG16 Architecture

It has a fairly straightforward architecture. After max-pooling, it consists of two contiguous blocks of two convolution layers, three consecutive blocks of three convolution layers, and three thick layers at the conclusion. The depth of the last three convolution layers varies based on the architecture using Figure 7. This analysis's most important finding is that the size reduces by 50% with each max-pooling.

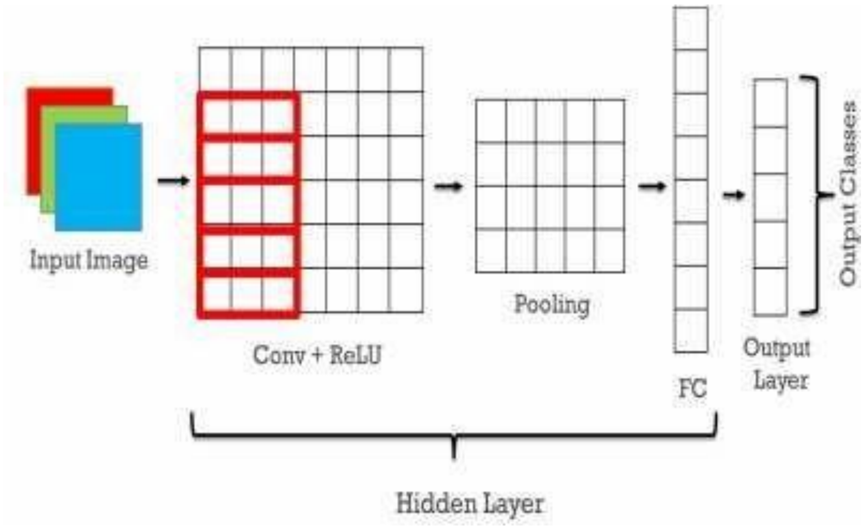

Fig 7. CNN layers

Features of VGG16 network:

Input Layer: It takes 224 by 224 color photos with three channels—Red, Green, and Blue—as input.

Convolution Layer: The images are passed through a series of convolution layers, each of which has a stride of one and a very narrow receptive field of 3 x 3.

Max pooling: It is carried out over a 2 x 2 max-pool window with a stride of 2[18], meaning that the max-pool windows in this case do not overlap. A max pool layer does not always follow a convolution layer; in certain instances, one convolution layer follows alternative convolution layer without a max-pool layer in between.

4.4 Implementations

In the beginning, we gather data from 21 families, gathering information such as contact and residence details, family member pictures, names, and relationships between the members. Each family gets its own folder, in which we file all the information gathered, labelling the folder with the family name. The RGB image must next be converted before being resized to 224 by 224 pixels. Then, more photos are added by employing augmentation techniques like rotation, flipping, shearing, and shifting.

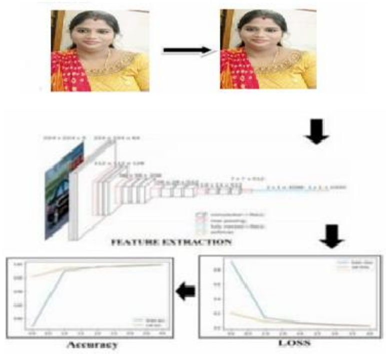

Fig 8. Steps of Implementation

As seen in Figure 8, [1] each image in the dataset is subjected to 18 random rotations, each of which rotates 10 degrees. Eighty percent of the gathered [11]data is used as training data after augmentation, and the model is trained. Throughout training, the model pulls information from several convolution layers. There are five maxpooling layers, each with size (2,2), and thirteen convolution layers, each with filter size (2,2). [11] In the convolution layer, ReLU activation function is affixed to every neuron. The convolution layer maps the features and summarizes the information in the input image. The instigation function F(x)= max (0,x) that determines the value of each negative activation to zero is the ReLU layer. In the suggested method, we use maxpooling, which computes the maximum value for each batch of feature maps. The pooling layer works on each feature map independently in order to build a new set of the equal numbers of pooled feature maps. This is followed by a dense layer, then in the output layer, a softmax classifier that predicts and classes the family. The two provided graphs show the variation in accuracy and loss throughout training and validation. As we continue to train the model, the accuracy rises and the loss falls. After five epochs of training, the model will be saved as a h5 file.

5. Results and Discussion

The figure displays the outcome of the project. The family name appears as the principal name on the outcome, together with the names, addresses, and mobile numbers of the family members. As seen in the figure, the pictures of the family relations are also on exhibit. The framework is educated and evaluated for 21 family datasets using three distinct strategies using Tensor Flow.

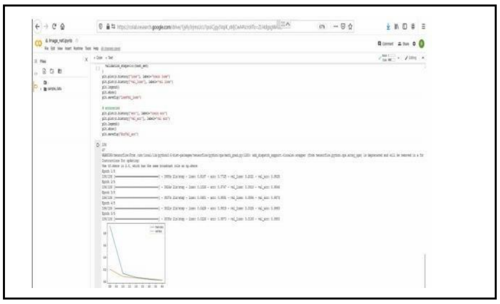

Fig 9. Displaying Family Name and Details

Fig10. Google Colab Platform

[11,1] The model is created in Google Colab, an online platform provided by Google for the creation of sophisticated artificial intelligence models. Figure 9 and Figure 10 illustrates how a model is prepared during the colab stage. Placing of courtesy model and Figure 11 illustrates the loading of the image and the image tools that are input. Figure 12 illustrates the train loss versus validation loss of VGG16 architecture, Figure 13 illustrates the accuracy of train and accuracy of VGG16 architecture.

Fig 11. Prediction Slide

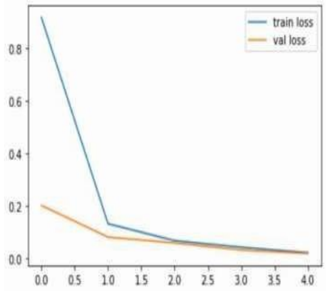

Fig 12. Train Loss vs Validation Loss of VGG16 Architecture

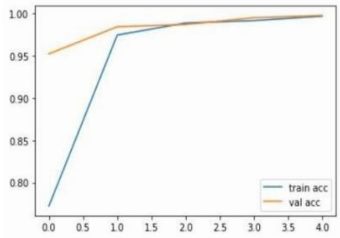

Fig 13. Train accuracy vs validation accuracy of VGG16 Architecture

6. Conclusion

The suggested framework validates a person's character by utilizing their face and familial quirks. In order to calculate value and single value order, this structure uses CNN calculation with VGG16 engineering, which is more accurate than some other calculations. The actualization of the organized model on the equipment stage, the intensify in accuracy, and the reduces in the false expectation rate set this work apart. This work differs from the earlier research described in the literature review in that it applies the prepared model to a hardware platform, increasing accuracy and lowering the false prediction rate. Additionally, the suggested system effectively addresses the national and local refugee crises.

References

1. R. Dandekar and M. S. Nimbarte, "Verification of family relation from parents and child facial images," 2014 International Conference on Power, Automation and Communication (INPAC), Amravati, 2014, pp. 157-162, doi: 10.1109/INPAC.2014.6981146.
2. "Face recognition based on convolution neural network,"36th Chinese Control Conference (CCC), Dalian, 2017, pp. 4077-4081, doi: 10.23919/ChiCC.2017.8027997, 2017.
3. D. K. Thara, B. G. Premasudha, V. R. Ram and R. Suma, "Impact of big data in healthcare: A survey," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, 2016, pp. 729-735, doi: 10.1109/IC3I.2016.7918057., 2016.
4. Q. Dai, P. Carr, L. Sigal and D. Hoiem, "Family Member Identification from Photo Collections," 2015 IEEE Winter Conference on Applications of Computer Vision, Waikoloa, HI, 2015, pp. 982-989, doi: 10.1109/WACV.2015.136., 2015.
5. D. K. Thara, B. G. Premasudha, Ramesh Sunder Nayak, T. V. Murthy,G. Ananth Prabhu & Naeem Hanoon,"Electroencephalogram for epileptic seizure detection using stacked bidirectional LSTM_GAP neural network",Evolutionary Intelligence,2020 Springer,https://doi.org/10.1007/s12065-020-00459-9.
6. Sajjanhar, Atul& Wu, ZhaoQi& Wen, Quan., "Deep Learning Models for Facial Expression Recognition". 1-6. 10.1109/DICTA.2018.8615843., 2018.
7. M. Coşkun, A. Uçar, Ö. Yildirim and Y. Demir, "Face recognition based on convolutional neural network," 2017 International Conference on Modern Electrical and Energy Systems (MEES), Kremenchuk, 2017, pp. 376-379, doi: 10.1109/MEES.2017.8248937. 2017.
8. Thara D.K., PremaSudha B.G., Fan Xiong,"Epileptic seizure detection and prediction using stacked bidirectional long short term memory",Pattern Recognition Letters,Volume 128, pp. 529- 535,ISSN 0167-8655,https://doi.org/10.1016/j.patrec.2019.10.034., 2019.
9. Deep face recognition using imperfect facial data Ali Elmahmudi,Hassan Ugail Centre for Visual Computing, Faculty of Engineering and Informatics, University of Bradford, Bradford BD7 1DP, UK.
10. Tolba, A.S., El-Baz, A.H. and El-Harby, A.A.,'Face Recognition: A Literature Review'. International Journal of Signal Processing,2005.
11. Bindushree S, Rakshitha AN, "Face Recognition using Deep learning'IJASI,Vol 01, 2020.www.ijasi.org

12. Dr. Abhinandan N, Karthika Ravi. "A comprehensive study on the role of forensic accounting in detecting financial frauds with reference to karnataka" , mLAC Journal for Arts, Commerce and Sciences (m-JACS) ISSN: 2584- 1920, 2024.
13. Utkarsh Kushwaha, Puja Gupta, Sonu Airen, Megha Kuliha. "Analysis of CNN Model with Traditional Approach and Cloud AI based Approach" , 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), 2022
14. Lalit Mohan Goyal, Tanzila Saba, Amjad Rehman, Souad Larabi-Marie-Sainte. "Artificial Intelligence and Internet of Things - Applications in Smart Healthcare" , CRC Press, 2021
15. Shambala Rameshwar Mahokar, Vishal V. Patil. "Understanding The Link Between Weather, Climate Variability, And Farmer's Financial Planning: A Study Of Khamgaon Region", mLACJournal for Arts, Commerce and Sciences (mJACS) ISSN: 2584-1920, 2024
16. Dr. Abhinandan N, Karthika Ravi. "A comprehensive study on the role of forensic accounting in detecting financial frauds with reference to karnataka" , mLAC Journal for Arts, Commerce and Sciences (m-JACS) ISSN: 2584- 1920, 2024.
17. Maryam Asadzadeh Kaljahi, Palaiahnakote Shivakumara, Tianping Hu, Hamid A. Jalab et al. "A geometric and fractional entropy-based method for family photo classification" , Expert Systems with Applications: X, 2019
18. Ahmed Husham Al-Badri, Nor Azman Ismail, Khamael Al-Dulaimi, Amjad Rehman, Ibrahim Abunadi, Saeed Ali Bahaj. " Hybrid CNN Model for Classification of in Grassland " , IEEE Access, 2022
19. Pawan Singh Mehra, Dhirendra Kumar Shukla. "Artificial Intelligence, Blockchain, Computing and Security - Volume 2" , CRC Press, 2023