Face Recognition Using Deep Learning

Bindushree S[#1], Rakshitha A N[*2]

UG Student ,Department of Information Science and Engineering,

Visveswaraya Technological University , Channabasaveshwara Institute Of Technology, Gubbi, Tuumkur, Karnataka, India

bindumjs1998@gmail.com, athi.rakshithaan27@gmail.com

Abstract -- Today face recognition and its usage are developing at a remarkable rate. Researches are at present building up different strategies in which facial recognition framework works. In circumstances like accidents, normal disasters, missing cases, clashes between nations, kidnappings and numerous different circumstances individuals are regularly isolated by their families. Recognizing the relatives of those refugees is essential to arrive at their family for refugee's security and backing. Everyday polices are enrolling with missing cases, a portion of those enlisted cases are getting tackled and some are definitely not by using the manual method where it takes more time. The goal of this paper is to provide a solution to overcome time delay from existing strategies for police examination utilizing most recent innovation. Hence we adopt a framework which utilizes CNN (Convolutional Neural Network) technique with VGG16 architecture where we use our raw dataset which contains 84 images collected from 21 families data, after applying augmentation method the image count in final dataset is increased to 1512, then from this dataset 80% of data is used for training data and 20% is used for testing data. This framework helps to verify an individual's trait using their face and family subtleties with related model with increased accuracy, and gives a effective solution for identifying refugee's family.

Keywords — Face recognition, Deep learning, CNN, VGG16.

I. INTRODUCTION

It should be appreciated that a person becomes a refugee because of various circumstances which are beyond that person control. Refugees are the ones who have lost their families members due to various circumstances who furthermore regardless don't have a legitimate situation to gather their resources and reach their family. Everyday polices are enrolling with many refugee cases, a decade before police use to solve the cases in a very traditional manual method by first collecting the information about their family and investigating based on that, to find their family. This manual method took very long time where most of the cases were not solved.

Later on as technology evolved many new methods emerged to solve the problem faced due to manual method.

Face Recognition technology is one of the most convenient and coherent technique of all the existing approaches for human identification and solving the refugee cases. Face Recognition system is a technology capable of matching a human face from a digital image or a video frame against a database of faces. However face recognition systems vary in their ability to identify people and accurate face recognition is still a challenge.

Several factors have contributed to the increasing interest in face recognition. Pattern recognition, machine learning and deep learning are the main contributors to the development of facial recognition systems. Deep learning is a kind of state-of-art method which can give high performance on face recognition. With the development of deep learning, face recognition technology based on CNN (Convolutional Neural Network) has become the main method adopted in the field of face recognition.

CNN is a class of deep neural networks, most commonly applied to analyzing visual imagery, it comprises of three layers which are convolutional layer, pooling layer and fully associated layer. Convolutional layer and pooling layer are utilized for feature extraction. Fully associated layer is utilized for grouping. Expert and researcher need area or business information for visual recognition in traditional manner. Convolutional Neural Networks (CNNs)[1][3] take out the manual feature extraction and it appear as a form of automated feature engineering.

As we use images as input data we adopted CNN technique for identifying family members of the refugees because of its high accuracy.

II. RELATED WORK

Our approach is similar to other recent works, In mid 1990's and 2000's, all comprehensive learning approach and nearly high quality methodology dominated face recognition area which gave poor results for unconstrained facial changes and under challenging conditions such as poor lighting, low quality image resolution and suboptimal angle of view since this approach used more than two feature descriptors. In 2012 ImageNet Large Scale Visual Recognition Challenge(ILSVRC) AlexNet won the ImageNet competition by reducing the top-5 error from 26% to 15.3% on ImageNet, achieving a top-1 error rate of 37.5% using a deep learning technique.

Convolutional neural network (CNN) is a type of deep learning methods [2][5] which uses multiple layers feature descriptors for feature extraction and transformation. The early layers extract the basic features of face and later layers extract the detail features of face. In 2012 there was another framework proposed called DeepID which utilized an ensemble of shallower and smaller deep convolutional networks than DeepFace[4][6]. This approach was considered as first approach that achieved high accuracy on LFW (Labeled Face in-the-Wild) dataset around 90%. In 2015 a CNN-based approach called FaceNet was proposed which used triplets of roughly aligned matching/nonmatching face patches using an online triplet mining method, and achieved 99.63% on LFW and95.12% on YouTube Faces DB and state-of-art face recognition performance and use 128 embedding's preface. In 2017, ResNet50 [5] use residual blocks and residual connections and train on VGGFace2, on MSCeleb-1M and on their union to assess face recognition performance. In 2018, VGG16 achieves high accuracy with real-time high-performance face recognition and efficient on embedded devices.

SIFT: features have been widely applied in face recognition. As done in [16], we have resized all facial images to 64×64, and set the block size to 16×16 with a stride of 8. Thus, there were a total of 49 blocks for each image, yielding a feature vector $128 \times 49 = 6,272D$.

LBP: it is frequently used for texture analysis and face recognition, as it describes the appearance of an image in a small, local neighborhood around a pixel.

VGG-Face CNN: uses a "Very Deep" architecture with very small convolutional kernels (i.e., 3×3) and convolutional stride (i.e., 1 pixel). The model was pre-trained on over 2.6 million images of 2,622 celebrities and each face image was resized to 224×224 and then fed-forward to the second-to-last fully-connected layer (i.e., fc7) of the CNN model, producing a 4,096D feature vectors.

A. IDMC
The Internal Displacement Monitoring Centre (IDMC)[7][8] is the world's authoritative source of data and analysis on internal displacement. Since its establishment in 1998 as part of the Norwegian Refugee Council, IDMC has offered a rigorous, independent and trusted service to the international community.

SL. NO	PAPER REFERRED	PUBLISHED YEAR	DATASET	METHOD USED	ACCURACY	REVIEW
1	[1]	2014	Kin Face V2& Family 101	LBP&KNN	90.00	Accuracy Decreases with change in Resolution.
2	[2]	2017	ORL	CAFFE	99.82	Change in posture reduces accuracy.
			AR Face		99.78	
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
					86.20	
				VGG19		
					91.37	
				VGGFACE		
5	[5]	2017	GEORGIA TECH FACE	CNN	98.08	Better computational speed and feature extraction

Fig 1: Literature Survey Table

III. Existing System

SL.NO	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 CLIDREN	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

Fig 2: Some existing Government Portals worldwide.

B. NCMEC
The National Centre for Missing & Exploited Children (NCMEC)[9][10] is a private, non-profit organization established in 1984 by the United States Congress.

C. *National Tracking System For Missing & Vulnerable Children*

National tracking system for missing and vulnerable children[11]. The centrally sponsored Integrated Child Protection Scheme aims at contributing to the improvement of the children who are in difficult circumstances. The scheme is being implemented by Ministry of Women and Child Development, Government of India.

IV. PROPOSED SYSTEM

A. *System Architecture*
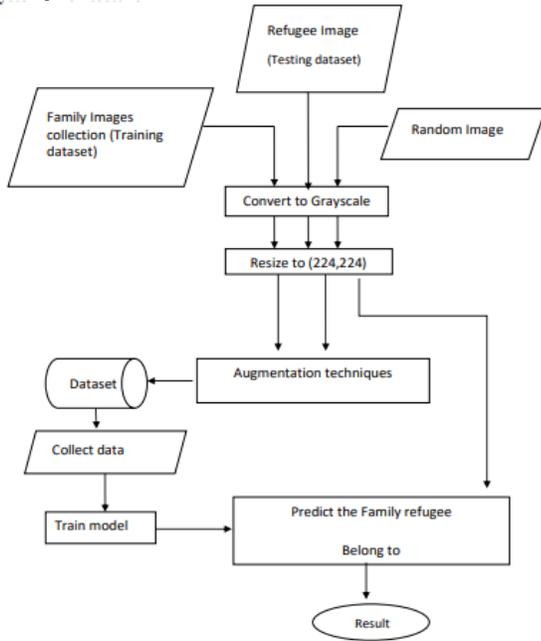
Fig 3: System Architecture

The facial recognition system can be built using two steps. The first step is a process through which the facial features are picked up or extracted, and the second step is pattern classification. The fig.3.1. Represents our face recognition system which shows all process and steps involved .

Steps involved[12]:
- Collecting the Dataset
- Importing Libraries and Splitting the Dataset
- Data Augmentation
- Building the CNN
- Full Connection
- Training our Network
- Testing

B. *Dataset*

Our dataset is a raw dataset which contains photographs of each member in the family with their names contact detail and address collected from 21 families, the dataset consisted of 84 images after applying augmentation method the image count in final dataset is increased to 1512, then by(from) this dataset 80% of data is used for training data and 20% is used for testing data[13].

C. *Data Pre-Processing*
- Individual folders are created for each family.
- RGB to Grey Conversion.
- Resizing.

```
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))
```

Above code snippet represents the python code for data pre-processing, each family image is stored as individual folders, the folder name is saved as its family name which is taken as input and every single image in the folder is read and rewritten after processing. The 1st step after reading the image is, Converting the RGB image to Grayscale by cv2 (using)[14][15]. Cvt Color built-in function.
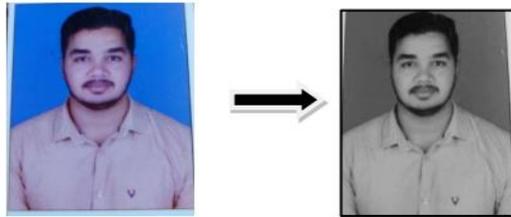
Fig 4: RGB to Grayscale.

Then in the 2nd step the image is resized to (224,224) in dimension. As the image size is fixed in the VGG16 model, that is the input image dimension must be (224,224)[16][17].The techniques used to generate the dataset is Image Augmentation.

What is Image Augmentation?
Image data augmentation is a technique that can be used to artificially expand the size of a training dataset by creating modified versions of images in the dataset.

Training deep learning neural network models on more data can result in more skillful models, and the augmentation techniques can create variations of the images that can improve the ability of the fit models to generalize what they have learned to new images[18]. The Keras deep learning neural network library provides the capability to fit models using image data augmentation via the *ImageDataGenerator* class.

A few methods utilized here are:
• Shift
• Rotation
• Zoom
• Brightness
• Shear

Shift : Horizontal and Vertical Shift Augmentation, A shift to an image means keeping the image dimensions same and just moving all pixels of the image in one direction, such as horizontally or vertically.
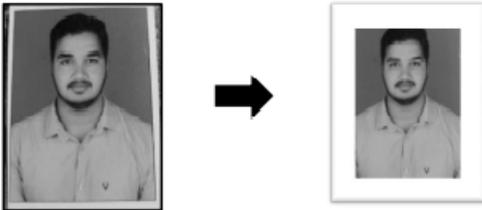
Fig 5: Shifted Image

Rotation : Random Rotation Augmentation, A rotation augmentation randomly rotates the image clockwise by a given number of degrees from 0 to 360.
The rotation will likely rotate pixels out of the image frame and leave areas of the frame with no pixel data that must be filled in.
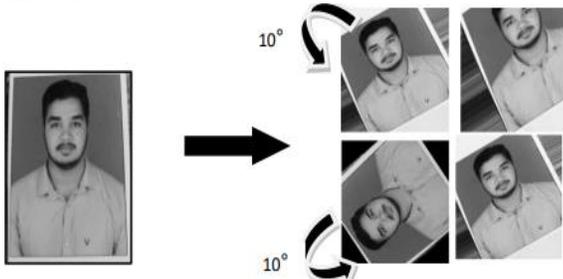
Fig 6: Rotated Image

Zoom : Random Zoom Augmentation, In this the image is zoomed in randomly or adds new pixel values around the image or interpolates pixel values respectively[19]. Image zooming can be arranged by the zoom_range contention to the Image DataGenerator constructor.
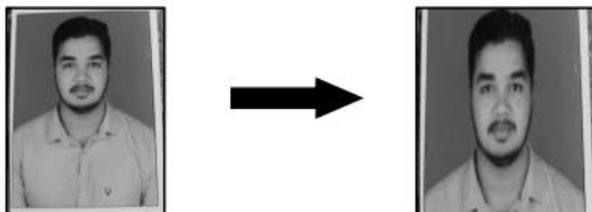
Fig 7: Zoom Image

Brightness : Random Brightness Augmentation. In this the image is augmented by either randomly darkening images, brightening images or both. The intent is to allow a model to generalize across images trained on different lighting levels.

Shear : Range of horizontal shear applied to the input image, specified as one of the following. Shear is measured as an angle in degrees, and is in the range (–90, 90). 2-element numeric vector.

D. VGG16 Model

VGG16 is a convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition". The model achieves 92.7% top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes.
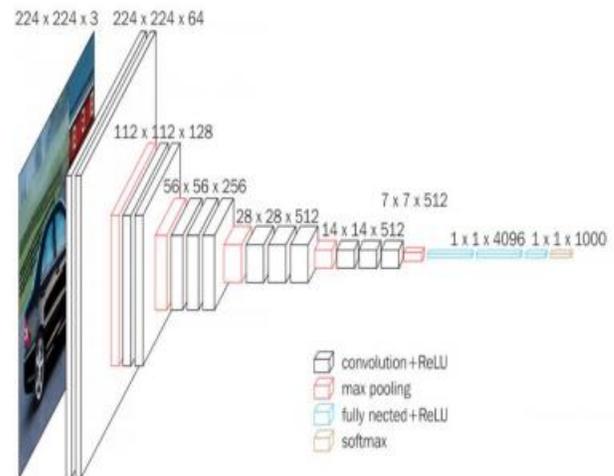
Fig 8: VGG16 Architecture

In the figure above, all the blue rectangles represent the convolution layers along with the non-linear activation function which is a rectified linear unit (or ReLU). As can be seen from the figure that there are 13 blue and 5 red rectangles i.e there are 13 convolution layers and 5 max-pooling layers. Along with these, there are 3 green rectangles representing 3 fully connected layers. So, the total number of layers having tunable parameters is 16 of which 13 is for convolution layers and 3 for fully connected layers, thus the name is given as VGG-16. At the output, we have a softmax layer having 1000 outputs per image category in the imageNet dataset. In this architecture, we have started with a very low channel size of 64 and then gradually increased by a factor of 2 after each max-pooling layers, until it reaches 512.

The architecture is very simple. It has got 2 contiguous blocks of 2 convolution layers followed by a max-pooling, then it has 3 contiguous blocks of 3 convolution layers followed by max-pooling, and at last, we have 3 dense layers. The last 3 convolution layers have different depths in different architectures. The important

thing to analyze here is that after every max-pooling the size is getting half.

Features of VGG-16 network:
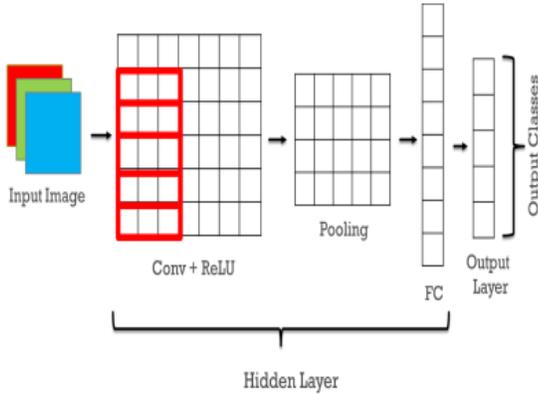
Fig 9: CNN Layers

Input Layer: It accepts color images as an input with the size 224 x 224 and 3 channels i.e. Red, Green, and Blue.

Convolution Layer: The images pass through a stack of convolution layers where every convolution filter has a very small receptive field of 3 x 3 and stride of 1.

Max pooling: It is performed over a max-pool window of size 2 x 2 with stride equals to 2, which means here max pool windows are non-overlapping windows.Not every convolution layer is followed by a max pool layer as at some places a convolution layer is following another convolution layer without the max-pool layer in between.

FC Layer: The first two fully connected (FC) layers[19][20] have 4096 channels each and the third fully connected layer which is also the output layer have 1000 channels, one for each category of images in the imageNet database.The hidden layers have ReLU as their activation function.

E. Implementation Steps

The steps that each image in the dataset will undergo is represented in the fig.4.8.
Initially we do data collection where we have collected data from 21 families which includes photographs of each family member, their names , relationship between the members, contact and address details of that family. We make separate folders for each family and store all the collected data by giving the folder name as the family name.

The next step is conversion of RGB image and then resize it to(224,224) in dimension. Then using the augmentation methods like zooming, flipping, rotation, shear and shifting the number of images is increased.
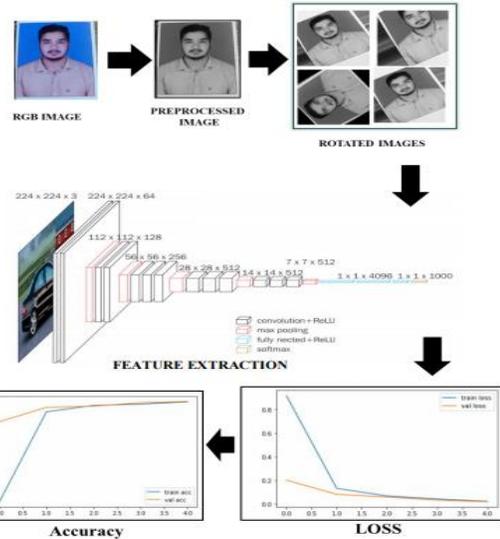
Fig 10: Steps of Implementation

Each image in the dataset undergoes 18 rotations by rotating 10 degrees each time randomly as shown in the   fig 4.8.

After augmentation 80% of the data collected is used as training data and the model is trained. The model extracts the features in different convolution layers during the training. There are 13 convolution layers each with the filter size (2,2) and five maxpooling layers each with size (2,2). Each neuron in the convolution layer has ReLU activation function attached to it. The features in the input image is summarized by the convolution layer and it maps the features. The ReLU layer is a activation function $F(x) = max(0,x)$, this function changes all the negative activations to 0.

The pooling layer operates on each feature map separately in order to create new set of same numbers of pooled feature map, in the proposed method we apply maxpooling it calculates maximum value for each batch of feature map. After this there is a dense layer followed by softmax classifier in the output layer it classifies and predicts the family. The loss and accuracy variation during training and validation is represented in the given two graphs. The loss decreases and the accuracy increases as we go on training the model. Once the training is complete after 5 epochs the model will be saved as a h5 file.

V. RESULT

The project outcome is displayed in the figure Outcome displays the family information with family name as primary name, along with their family members name, Mobile Number and the address and also the photographs of the family members are displayed as shown in the figure Using tensorflow the model is trained and verified for 21 family dataset using three different algorithms which are

i.   CNN algorithm with VGG19 architecture
ii. CNN algorithm with Resnet50 architecture
iii. CNN algorithm with VGG16 architecture

Fig 10: Displaying Family Name and Family Details
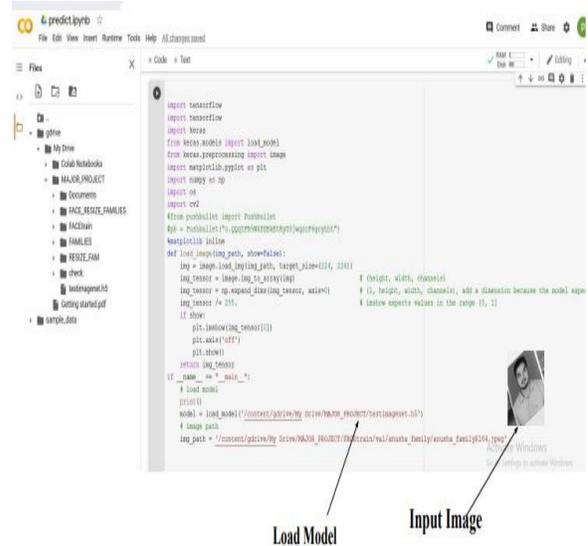
The model is prepared in Google Colab, which is an online stage given by Google for Preparing of complex AI models. Figure 5.2 shows the preparation of a model in colab stage.
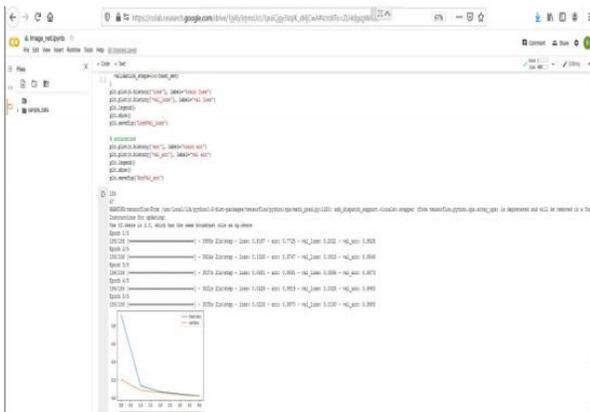
Fig 11: Google Colab platform

Loading of respective model and inputting the loaded image along with the inputed image implements is shown in fig 5.3

Fig 12: Prediction Slide
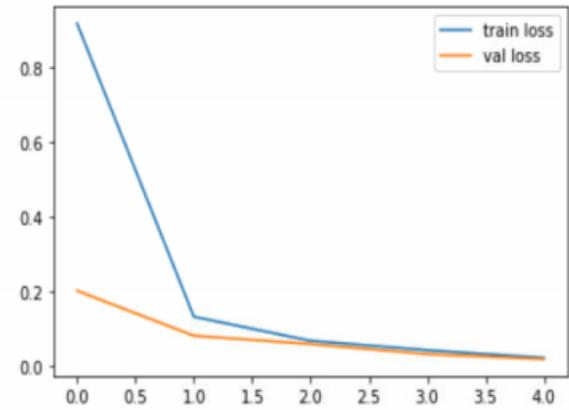
Fig 13: Train loss vs validation loss of VGG16 architecture
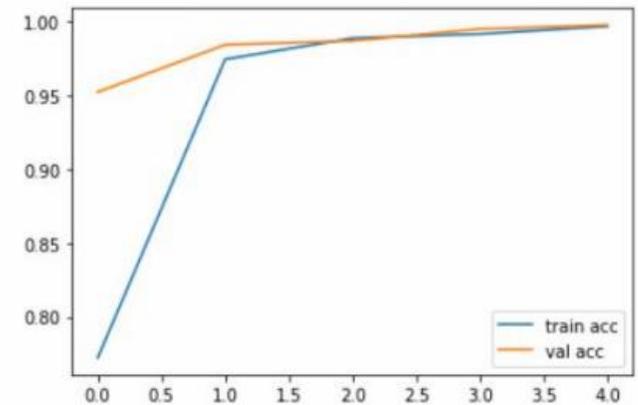
Fig 14: Train accuracy vs validation accuracy of VGG16 architecture

VI. CONCLUSIONS

The proposed framework using the individuals face and family subtleties verifies an individuals character. This framework utilizes CNN calculation with VGG16 engineering which is more precise than some other calculations for mark and unique mark order. this work stands separated by actualizing the prepared model on equipment stage and increment in the exactness and reduction in the bogus expectation rate.

Comparing with the previous work done mentioned in literature review, this work stands apart by implementing the trained model on hardware platform and increase in the accuracy and decrease in the false prediction rate. And proposed system is effective solution for refugee crisis locally and nationally.

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]   K. Yan, S. Huang, Y. Song, W. Liu and N. Fan, "Face recognition based on convolution neural network," 2017 36th Chinese Control Conference (CCC), Dalian, 2017, pp. 4077-4081, doi: 10.23919/ChiCC.2017.8027997

[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.

[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.

[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. (2018). Deep Learning Models for Facial Expression Recognition. 1-6. 10.1109/DICTA.2018.8615843.

[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.

[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 128,2019,Pages 529-535,ISSN 0167-8655,https://doi.org/10.1016/j.patrec.2019.10.034.

[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]   A. S. Tolba, A.H. El-Baz, and A.A. El-Harby. Face Recognition: A Literature Review,

[11]   AnimeshPatcha, Jung-Min park, "An Overview of Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends", Bardley Department of Electrical And Computer Engineering. 2007.

[12]   Analyticsvidhya.com/blog/2017/12/fundamentals-of-deep-learning-introduction-to-lstm

[13]   Thara D K, Dr. B G Premasudha, "A Review On Computer Aided Diagnosis Of Epilepsy Using Machine Learning And Deep Learning", IJRAR - International Journal of Research and Analytical Reviews (IJRAR), E-ISSN 2348-1269, P- ISSN 2349-5138, Volume.6, Issue 2, Page No pp.25-33, May 2019, Available at : http://www.ijrar.org/IJRAR19K1171.pdf

[14]   towardsdatascience.com/deep-inside-autoencoders-7e41f319999f

[15]   Strengths and Weaknesses of Deep Learning Models for Face Recognition Against Image Degradations KlemenGrm, VitomirStruc , AnaisArtiges, Matthieu Caron, Hazim Kemal Ekenel.

[16]   2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia) An Ensemble of VGG Networks for Video-Based Facial Expression Recognition.

[17]   Family Member Identification from Photo Collections 2015 IEEE Winter Conference on Applications of Computer Vision.

[18]   Families in the Wild (FIW): Large-Scale Kinship Image Database and Benchmarks Joseph P. Robinson, Ming Shao, Yue Wu, Yun Fu.

[19]   Thara D.K., PremaSudha B.G, Fan Xiong,"Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques",Pattern Recognition Letters,Volume 128,2019,Pages 544-550,ISSN 0167-8655,https://doi.org/10.1016/j.patrec.2019.10.029.

[20]   Face Recognition Based on Convolution Neural Network. Kewen Yan, Shaohui Huang, Yaoxian Song1, Wei Liu1, Neng Fan.