International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

Intelligent Face Recognition System for Enhancing the Security System Using Artificial Intelligence

Sruthi Thanugundala, Gundu Yashwin
Assistant Professor, Department of CSE, Princeton Institute of Engineering & Technology for Women, Hyderabad.
Department of CSE (AI&ML), Sri Indu College of Engineering and Technology, Hyderabad.

Abstract: AI (artificial intelligence) is being used in more and more areas, such as face recognition systems, because people want more improved security measures. This essay goes into great detail about an Intelligent Face Recognition System that is meant to make security better. Using AI to provide reliable and effective face recognition is what the suggested system does, which helps make security systems better overall. The paper talks about the tools behind such a system, its problems, and the things it could be used for. Deep learning-based methods have been getting more and more attention lately because they learn in different ways than standard machine learning methods. They have recently become very popular. This study shows a new way to recognise faces in a free-form world using deep residual ResNet50, a common type of convolution neural network architecture. In a free setting, the network is changed and fine-tuned for different face recognition situations.

Keywords: Face Recognition, Convolution Neural Network, Unconstrained Environment

I.INTRODUCTION

This Automatic Face Recognition (AFR) is meant to make recognition better by lowering the False Acceptance Rate (FAR) and False Rejection Rate (FRR). The proposed framework was trained and tried on the difficult datasets face94, face95, face96, Grimace, and LFW, which are all available to the public. It works better than current state-of-the-art methods without the need for extra resources.

The market for home protection has grown very quickly in the last few years. As intelligence becomes a major trend in the industry, intelligent security has slowly become the way that security companies are changing and getting better, and it will play a bigger role in the security industry as a whole. China's security business had a market size of about 720 billion yuan in 2019. It is thought that clever security will make a market worth around 100 billion yuan in 2020. This market will be very important in the security field. The Internet of Things can be used in many different ways in smart towns, public safety, and some specific businesses. One of the most important and necessary of these is fingerprint technology. The study of face recognition technology using big data platforms is very important. Big data and AI are both fields that are developing quickly. The big data environment not only makes it possible to develop face recognition systems in more depth, but it also lets people share feature databases across a wider range of fields, which helps to create more face feature databases.

II. RELATED WORK

In the last few decades, face recognition that uses computers has been very successful. When it comes to uses in limited environments, the face recognition method works well [1]. AFR can be used in many ways that are important to the economy, such as to find lost babies, stop passport fraud, and keep unwanted people out of hotels and gambling [4]. The hardest thing about the AFR system is that it has to deal with a lot of uncontrolled situations, like changes in posture like head movement, lighting, and facial emotions. Face recognition problems like these are now being faced by police and other intelligence agencies, as well as users on Google, Facebook, and Twitter [5]. In the last ten years, different ideas for how to solve the problems listed above have been put forward [6]. Besides the issues we already talked about, the methods we have now are strong enough for feature extraction and face normalisation. Eigen faces [1], Fisher faces [2], and their extensions [3] are common ways to get global features. They take into account changes that happen because of small poses and simple lighting changes. It's not as good with these methods when there are big changes in lights and positions. These methods also don't work well when things change, like when blockage and expression happen.

Other descriptors, like Har [7], Gabor [8], and LBP [9], have come a long way over time and have shown that they can handle local distortion in face recognition. Using traditional machine learning methods like feature extraction and feature selection can be time-consuming, mostly because they need to do a lot of image processing, like gradient math, to get useful information from pictures.

Recently, depth models like Convolutional Neural Networks (CNN) have shown that they can pull out advanced features from a library set of images that were taken in an unrestricted environment. This has led to the creation of boundaries that can identify faces in uncontrolled settings. Krizhevsky et al. showed how useful the CNN model could be by showing that it could do very well on the ImageNet dataset for more difficult visual classification tasks. Being able to get big data sets, using different training model methods, and stopping overfitting with techniques like Dropout [14] are what make things possible.

Researchers have come up with a number of deep CNN designs since then to improve the accuracy of face recognition in big data sets [13]. On the ImageNet dataset, the deep standard ConvNet model did a lot better. Very deep ConvNets, like VGG, GoogLeNet , and ResNet, get the best results by adding more depth to the weight layer. In the deep CNN design, the network's depth is slowly grown so that more convolution layers can be added.

III. REVIEW WORK

A lot of work has also been done on systems for recognising faces in 3D. The algorithm created by Vezzetti et al. [12] automatically detects six basic facial expressions: anger, disgust, fear, joy, sadness, and surprise. It does this by splitting the picture of the face into 79 regions and comparing the feature values in each region with a cutoff. An method for recognising faces is looked into by Marcolin et al. [13]. It uses geometric variables (like mean, Gaussian, path, shape index, curvature, and coecients of basic form) to get 3D facial traits and analyse them. They used a method that blends 2D and 3D facial traits to look at and recognise facial expressions in a study [14]. People who are suspected of committing crimes often hide their faces with makeup, which makes it harder to recognise them. In order to solve this issue, a database and an experiment are used to suggest a way to find the problem and fix it.

CNNs don't have neurons that are fully connected to each other like other neural networks do. Instead, neurons in each layer are only connected to sub-regions. CNN cut down on the number of links, share weights, down sampling, and factors because of this. For the picture that it is given, each layer of the CNN makes a feature map, also called an activation. The network's first layer picks up basic details like the image's lines and spots. For a deeper convolutional layer, the features that were taken are moved and sub-sampled. The fully linked layer, which is the last one, is in charge of network thinking. As a result, the network gives you the chance that an incoming picture belongs to each of the possible classes. Pre-trained CNNs can be used as feature extractors without having to go through the time and work of training them.

Krizhevsky et al. [1] won the ILSVRC (ImageNet Mass Visual Recognition challenge) competition in 2012 because they did well on the ImageNet classification standard.

The big deep neural network has 8 layers, and they trained it. For top1, the mistake rate was 37.5%, and for top5, it was 17.0%.

A more detailed network design was shown by Zeiler and Fergus [3] in 2013 and won the ILSVRC competition that same year. At each level of the model, they show how the features, filters, and weights look. On the ImageNet collection, the feature information leads to better classification, with a 14.8% mistake rate.

Simonyan and Zisserman test how well the ConvNet design works by adding more weight layers to the network [16]. They said that the best results came from making ConvNet deeper by adding 19 layers, which made a CNN model known as VGGNet.

In 2015, Szegedy et al. suggested a deep CNN called GoogLeNet. It had 22 convolutional layers and beat the ILSVRC-2014 with a 6.7% error rate. The model is different from the other designs because it adds a new layer structure in the form of a starting module. In contrast to standard designs, the first module creates a space where different tasks, like pooling or convolution, can be done at the same time.

Deep learning was used by Sun et al. to learn advanced facial features, which they called DeepID (Deep identification features) for face detection. The DeepID function is based on the structure of feature extraction in deep neural networks and adds up layers of features from different sizes.

The DeepID-obtained discriminant features were very small, and they made face verification 97.45% more accurate in the LFW field dataset. Sun et al. [4] said that LFW could verify faces with 99.15% accuracy. They show that using monitoring data for face recognition and checking depth features has a big effect on the system (DeepID2). The DeepID3 LFW data set, which is part of another ConvNets architecture, has an improved recognition rate of 99.53%. This is thanks to the stacked convolution and initial layer design from.

It was suggested by Schroff et al. in 2015 that ConvNet could be used to directly learn embedded Euclidean space for face verification. On the LFW dataset and the YouTube Faces DB, a method called FaceNet got a good score of 99.63%.

Another idea from Zhou et al. is the Megvii Face Recognition System (MFRS), a face recognition system that works well with a 10-layer deep network design for guided learning. In the LFW benchmark, they also got 99.5% of the competition ranking.

Microsoft Research Asia released ResNet at the end of 2015. It is a very deep CNN design with 152 layers that sets a new standard in picture classification, localization, and detection. With an error rate of 3.6%, the ResNet design won the 2015 ILSVRC and set new records for the number of levels. Because the other networks have come up with so many new ideas, ResNet has become the best CNN design.

In 2016, Masi et al. [3] suggested a gesture-aware method that uses CNN to recognise faces in pictures with big changes in pose and says that the IJB-A and PIPA data sets work better.

Shi and Jain recently used another deep CNN to improve the LFW dataset performance of face recognition systems. The method uses techniques for spatial change to find distinguishing face features. This group says that their suggested method works better on harder data sets than the most advanced face recognition methods.

In 2018, Elitsa Popova et al. suggested using an artificial neural network to improve a pairwise optimisation method. The writers say that the suggested method needs fewer ad hoc settings to make multiclass convolution networks run more quickly.

IV. METHODLOGY

Components of the Intelligent Face Recognition System

a. Image Acquisition:

The system begins with the acquisition of facial images through various sources such as CCTV cameras, smartphones, or dedicated surveillance devices. High-quality images are crucial for accurate face recognition.

b. Preprocessing:

Preprocessing techniques, including image normalization, alignment, and noise reduction, are applied to enhance the quality of acquired facial images. This ensures that the subsequent recognition process is robust to variations in lighting, pose, and facial expressions.

c. Feature Extraction:

AI algorithms extract distinctive facial features from the preprocessed images. Convolutional Neural Networks (CNNs) and deep learning techniques are commonly employed for this purpose. These features form the basis for accurate identification and authentication.

d. Face Matching:

The extracted features are matched against a database of stored facial templates. Machine learning algorithms, such as Support Vector Machines (SVM) or Neural Networks, are utilized to determine the similarity between the captured image and the stored templates.

e. Decision Making:

Based on the matching results, the system makes a decision regarding the identity of the individual. Advanced decision-making algorithms, often incorporating probabilistic models, contribute to the system's accuracy.

Human Face Recognition Method

Because the camera angle and the way the face is posed can change, a 3D picture of a face can't hold all of its information. This means that some features of the face are missing, which affects how well it can be recognised. With the face turned to the side, the side contour is the shape of the front end at its tallest point. Compared to other characteristic curves [24], it has a lot of facial traits and can show the most important parts of the face. Figure 6 shows that using the face contour for face recognition is more reliable and stable because the side outlines of different people are not all the same. Because the root of the nose is in the middle of both eyes, the position of the side outline is based on where the two inner eye corners are. The side outline is taken out based on where the nose's tip and root are located.

Figure 1 shows the suggested architecture's schematic structure. It has a convolution layer, a rectified linear function (ReLU) layer, a max pooling layer, and a normalisation layer.

A design with 27 layers has been put forward, with 24 layers being convolution layers and 3 levels being FC (fully connected layers). The picture that you send is 227x227x3, with 3 being the colour planes. The depth of the convolution layers in the first layer is set to 32 by default. After each max-pooling layer, it grows by two, and so on, until it reaches 512. The first layer's size of In the first layer, the kernel is 11x11 and has a stride of 4. It is then run through a ReLU function with a max sharing size of 3x3 and a stride of 2. A ReLU layer sets 0 to any number that is not negative. The kernel size is set to 5x5 with speed 1 and pad 2 in the second layer. After that, the feature maps that were made are put through a corrected linear function with max pooling that is 3x3 and uses step 2. Its kernel size is 3x3, and its speed and padding are both 1. The 3, 4, and 5 levels all have this. The result feature maps are run through the same corrected linear function with max pooling once more. In a normal setup, the top layers are fully linked. Three FC levels are used in the blueprint for the building.

Which has three layers: the first two have 4096 channels each, and the third has 6144 channels. As the last layer, the softmax algorithm is used.

We train the model with the chosen big number of photos. To sort pictures into groups, the cross entropy loss function is used. To change the weights and make the Loss function as small as possible, a gradient descent method is used.

The suggested CNN gives off a softmax forecast function with a chance of N different identities, that is,

$$y_i = \frac{exp(y_i^{'})}{\sum_{j=1}^{n} exp(y_j^{'})}$$

where, $y_j^{'} = \sum_i x_i . w_{i,j} + b_j$ linearly combines the features *xi* and *yj* is its output. The CNN is learned with minimizing *–logyt*, with the *t*-th target class.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 03 | March - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

V.RESULTS and DISCUSSION

We used five publicly available face datasets: LFW, Face94, Face95, Face96, and Grimace. These have 13,233 images of people (5749 people), 3,060 images of people (153 people), 1440 images of people (72 people), and 3,040 images of people (152 people). In Table I, a full summary of the information is given.

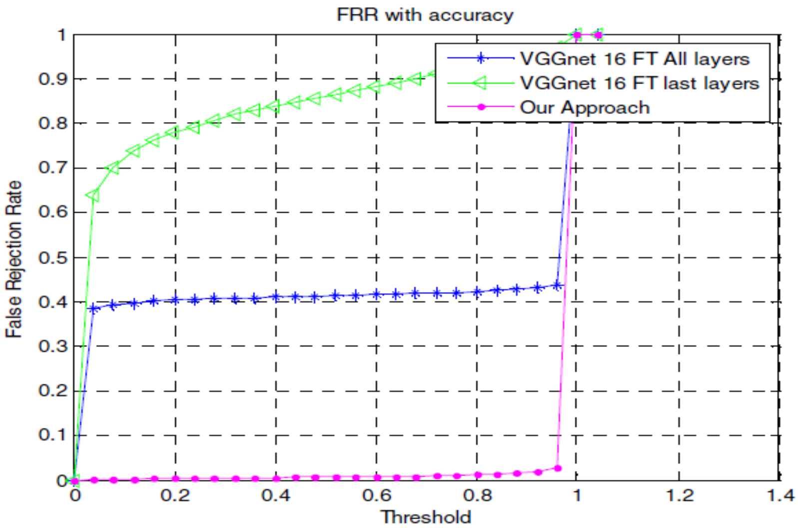

Fig. 1. False Rejection Rate for Our Approach, VGGNet16 last layer and VGGNet16 all layers

In this study, the experiment is done on 6,144 people using a total of 21,133 pictures from the five collections listed above. Out of all the pictures, 14,493 are used for training, which is 70% of the total. The other 30%, or 6,340 images, are used for testing. As you can see in Figure 2, our method has a very low false rejection rate compared to the VGGNet16 design when fine tuning all layers or just the last layers. This is true for both situations. Figure 1, on the other hand, makes it clear that the wrong acceptance rate is also very low compared to VGGNet16, especially when all the layers are fine-tuned. Here are the specifics of FRR and FAR for our method, as well as VGGNet16 for the top layer and all the levels below it.

There is a big difference between our method and VGGNet16's last layer (0.7798) and all of its layers (0.4036) when the threshold number is set to 0.2, as shown in Table II. In the same way, the FRR keeps going down compared to the VGGNet16 design at cutoff values of 0.4, 0.6, and 0.8. But for FAR, the value on VGGNet16's last layers is about the same as ours, but it's better than VGGNet16's entire set of layers and keeps doing better than VGGNet16's entire set of layers.

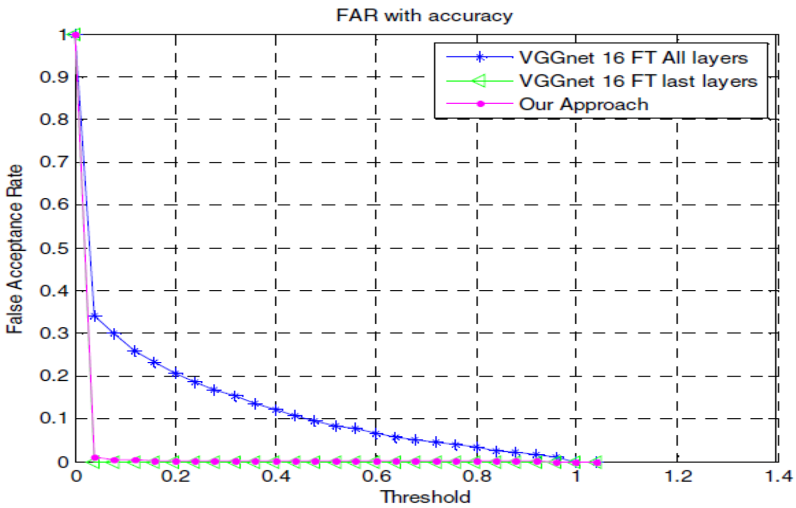

Fig. 2. False Acceptance Rate for Our Approach, VGGNet16 last layer and VGGNet16 all layers.

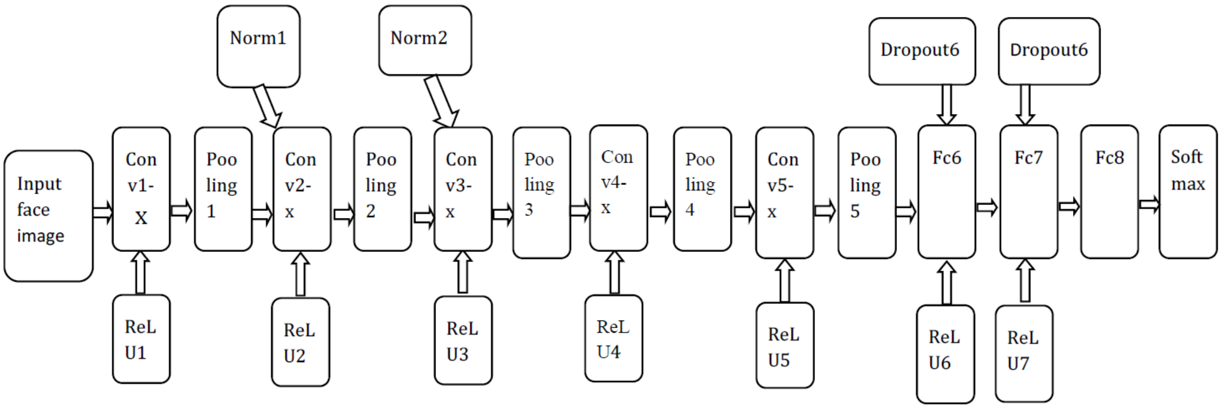

Figure 3. Framework of proposed CNN based architecture for face recognition.

It looks like combining an Intelligent Face Recognition System with Artificial Intelligence could make protection much better. These systems can quickly and correctly identify people by using complex algorithms and deep learning methods.

S.No.	Dataset	Individuals	Total Images	Description
1	LFW	5749	13233	Complex faces under pose, illumination and expression
2	Face94	153	3060	Less variation of head tilt and turn, plane background of green colour, considerable expression variation
3	Face95	72	1440	Considerable lighting and shadow effect due to artificial lighting, Less variation of head tilt, turn; plane background of red colour, considerable expression variation
4	Face96	152	3040	Varying background and lighting condition, very complex in terms of facial expression and large scale variation, large head scale variation
5	Grimace	20	400	large head scale variation due to translation of face position, considerable expression variation

Table I. Database Description For LFW, FACE94, FACE95 AND FACE96.

This helps make a variety of settings safer and more secure. To make sure that this technology is used in a responsible and useful way, problems like private issues, differences in the environment, and moral concerns must be dealt with. As AI keeps growing, more research and development will be done on Intelligent Face Recognition Systems to make them even better. This will make a big difference in the progress of security systems in many areas.

Threshold	FRR (False Rejection Rate)			FAR (False Acceptance Rate)		
	VGG Net16 (Last Layer)	VGGNet16 (All layers)	Our Approach	VGG Net16 (Last Layer)	VGGNet16 (All layers)	Our Approach
0.2	0.7798	0.4036	**0.0032**	0.001	0.206	**0.00237**
0.4	0.8382	0.4105	**0.0056**	0.0007	0.1209	**0.00107**
0.6	0.8808	0.4154	**0.00068**	0.0004	0.06701	**0.00043**
0.8	0.9278	0.423	**0.12**	0.000	0.03275	**0.00043**

Table II. Performance Of FRR AND FRR on Different Threshold Values.

The recognition rate is 91.5% when stacked matching is only used for rank recognition. The rate of recognition is also 76.5% when point-distance matching is only used for rank recognition. The stacked elastic matching includes both local and global features of the curve. However, the recognition rate is not great because the method only uses the two-dimensional coordinate information of the sampling places and throws away their one-dimensional information.

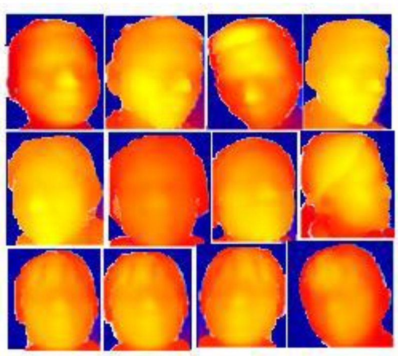

Figure 4. 3D face reconstructed in various poses.

The point-distance matching also looks at how far away the sampling places are from the nose-tip point. This method may have a lower recognition rate, but it does a good job of keeping the three-dimensional spatial information, which makes up for the problems with stacked elastic matching. We get a higher recognition rate when we use both point-distance matching and stacked elastic matching together in this study. Figure 15 shows the cumulative match characteristic (CMC) plots for all three tests.

VI.CONCLUSION

Face recognition systems have become a lot more accurate in the last few years thanks to deep learning. There are already a lot of architectures built on convolution neural networks that are better at recognizing objects. It is clear that by making a few small changes to the design that is already in place, the outcome can be greatly enhanced for other uses. Because of this, this study shows a changed version of the ResNet50 architecture for a face recognition system. The framework is trained and tested on five face datasets that are open to the public: LFW, face94, face95, face96, and Grimace. With a lower False Acceptance Rate and False Rejection Rate, this work shows that the recognition system is more accurate than the norm. When we compare our network to VGGNet16 (a normal CNN-based design), it's clear that our changed network consistently does better.

REFERENCES

[1] M. Turk and A. Pentland,"Eigenfaces for recognition". Journal of Congnitive Neuroscience, 3(1):71-86, 1991.

[2] P.Belhumer, J. Hespanha, and D. Kriegman. "Eigenfaces vs.Fisherfaces: Recognition using class specific linear projection". TPAMI, 19(7):711-720, 1997.

[3] X. Wang and X. Tang. "A unified framework for subspace face recognition". TPAMI, 26(9):1222-1228, 2004.

[4] A. Agrawal, Y.N.Singh, "An efficient approach for face recognition in uncontrolled environment [J]. Multimedia Tools and Applications 76 (8):1-10, 2017.

[5] Ravindra Changala, Block Chain and Machine Learning Models to Evaluate Faults in the Smart Manufacturing System, International Journal of Scientific Research in Science and Technology, Volume 10, Issue 5, ISSN: 2395-6011, Page Number 247-255, September-October-2023.

[6] Ravindra Changala, AIML and Remote Sensing System Developing the Marketing Strategy of Organic Food by Choosing Healthy Food, International Journal of Scientific Research in Engineering and Management (IJSREM), Volume 07 Issue 09, ISSN: 2582-3930, September 2023.

[7] P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J.O'Toole, D. S. Bolme, J. Dunlop, Y. M. Lui, H. Sahibzada, and S. Weimer, "An introduction to the good, the bad, & the ugly face recognition challenge problem," in 2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG). IEEE, pp. 346–353, 2011.

[8] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. "Deepface:Closing the gap to human-level performance in face verification", IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 1701–1708, 2014.

[9] P. Viola and M. Jones. "Rapid Object detection using a boosted cascade of simple features". CVPR, 2001.

[10] J. Daugman. "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by twodimensional visual cortical filters". Optical Society of America, Journal, A: Optics and Image Science, 2:1160-1169, 1985.

[11] Ravindra Changala, A Dominant Feature Selection Method for Deep Learning Based Traffic Classification Using a Genetic Algorithm, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, ISSN : 2456-3307, Volume 8, Issue 6, November-December-2022, Page Number : 173-181.

[12] Ravindra Changala, A Novel Approach for Network Traffic and Attacks Analysis Using Big Data in Cloud Environment, International Journal of Innovative Research in Computer and Communication Engineering: 2320-9798, Volume 10, Issue 11, November 2022.

[13] Ravindra Changala, Analysis and Prediction of Water Quality Data using Machine Learning Approaches and Exploratory Data Analysis, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, ISSN : 2456-3307, Volume 8, Issue 6, November-December-2022, Page Number : 188-193.

[14] T.Ahonen, A. Hadid, and M. Pietikainen. "Face description with local binary patterns: Application to face recognition".TPAMI, 28(12):2037-2041, 2006.

[15] D. Chen, X. Cao, and F. W. J. Sun. "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification". CVPR, 2013.

[16] N. Dalal, B. Triggs "Histograms of oriented gradients for human detection". International Conference on Computer Vision & Pattern Recognition, Vol. 2, INRIA Rhone-Alpes, ZIRST-655, av. de l'Europe, Montbonnot-38334, pp. 886–893, 2005.

[17] K. O'Shea, and R. Nash. "An Introduction to Convolutional Neural Networks", arXiv:1511.08458v2, 2015.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "Imagenet classification with deep convolutional neural networks". In Proc. NIPS, 2012.

[19] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever and R. R. Salakhutdinov "Improving neural networks by preventing co-adaptation of feature detectors" Department of Computer Science, University of Toronto, 2012.

[20] Ravindra Changala, "Statistical Models in Data Mining: A Bayesian Classification" in International Journal of Recent Trends in Engineering & Research (IJRTER), volume 3, issue 1, pp.290-293. in 2017.

[21] Ravindra Changala, "Secured Activity Based Authentication System" in " in Journal of innovations in computer science and engineering (JICSE), Volume 6, Issue 1,Pages 1-4, September 2016.ISSN: 2455-3506.

[22] Ravindra Changala, "Automated Health Care Management System Using Big Data Technology", at Journal of Network Communications and Emerging Technologies (JNCET), Volume 6, Issue 4, April (2016), 2016, pp.37-40,ISSN: 2395-5317, ©EverScience Publications.

[23] A. G. Howard. "Some Improvements on Deep Convolutional Neural Network Based Image Classification", 2013.

[24] K. Simonyan and A. Zisserman. "Very deep convolutional networks for large-scale image recognition". In ICLR, 2015.

[25] C. Szegedy,W. Liu, Y. Jia, P. Sermanet, S. Reed, D.Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. "Going deeper with convolutions". CVPR, 2015.

[26] K. He, X.Zhang, S.Ren, and J.Sun. "Deep Residual Learning for Image Recognition", In CVPR, 2016.