International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-10, August 2019

Enhanced Deep Learning with featured transfer learning in Identifying Disguised Faces

Yedavalli Venkata RaghavaRao, Kuthadi Venu Madhav, Selvaraj Rajalakshmi

Abstract: Paper The objective of face recognition is, given an image of a human face identify the class to which the face belongs to. Face classification is one of the useful task and can be used as a base for many real-time applications like authentication, tracking, fraud detection etc. Given a photo of a person, we humans can easily identify who the person is without any effort. But manual systems are biased and involves lot of effort and expensive. Automatic face recognition has been an important research topic due to its importance in real-time applications. The recent advance in GPU has taken many applications like image classification, hand written digit recognition and object recognition to the next level. According to the literature Deep CNN (Convolution neural network) features can effectively represent the image. In this paper we propose to use deep CNN based features for face recognition task. In this work we also investigate the effectiveness of different Deep CNN models for the task of face recognition. Initially facial features are extracted from pretrained CNN model like VGG16, VGG19, ResNet50 and Inception V3. Then a deep Neural network is used for the classification task. To show the effectiveness of the proposed model, ORL dataset is used for our experimental studies. Based on the experimental results we claim that deep CNN based features give better performance than existing hand crafted features. We also observe that the among all the pretrained CNN models we used, ResNet scores highest performance.

***Index Terms*: pre-training, Deep CNN features, CNN, DNN, transfer learning.**

I. INTRODUCTION

The objective of face recognition is identifying the label of a given class. When a group of person's images are available in the dataset, recognizing the class label of a person is called as face recognition. A set of Face recognition is one of the useful tasks and can be used as a base for many real-time applications like authentication, fraud detection, Criminal identification, Human computer interface (HCI) etc [2]. In the traditional face recognition methods are: Data collection, pre-processing, feature extraction, training the model and evaluate the model with test data. After collecting the data and pre-processing the data the most important step is the feature extraction. The performance of a classification algorithm is decided by the features that represent the given data. If these features are more representative and discriminative then the performance the further steps would be better.

Revised Manuscript Received on August 01, 2019

Yedavalli Venkata RaghavaRao[1], Professor, Department of Computer Science and Engineering University AP, India.

Kuthadi Venu Madhav, Department of CS & IS, Faculty of Science, BIUST, Botswana

Selvaraj Rajalakshmi, Department of CS & IS, Faculty of Science, BIUST, Botswana

In the traditional face recognition systems hand crafted features like gist [20], Hog[20], LBP[13], are used. These features can represent the entire image as a global feature vector. As a further improvement a set of features representation is introduced. In these representations instead of using a single static feature vector a set of features are used. Block features and sift[14] are the examples for local feature representation methods. In the traditional face recognition methods any of these handcrafted features or a combination of these is used. Once the features are extracted then they are passed to the classifier like SVM [4] or any other classifier. The downside of these methods is that they are task specific, time consuming and extracting for large set is not feasible. In addition, they require domain knowledge before processing and their generalization capability is limited. With the introduction of Convolutional Neural Networks (CNNs), machine learning field flourished [15]. Most of the success in the computer vision and NLP is only with the introduction of Deep learning. The recent advance in GPU has taken many applications like face recognition, hand written digit recognition and object recognition to the next level. Deep neural networks based models give outstanding performance compared to the other shallow models like SVM, MLP etc;. With the introduction of deep CNNs, it has been possible to learn more robust, discriminative and representative features [19]. The performance of the classifier is better when these features are used for training as the generalization ability of these models is better compared to other models. In the last few years there has been tremendous progress in the area of facial image classification. In facial image classification, the objective is to identify the person to whom the given face belongs to. A major breakthrough in this face classification task is possible with the introduction of deep neural network (DNN) based models. The reason for the success of deep learning models is the effective feature representation [15]. Representation of the images is a key factor for the success of face classification, object classification, and many other computer vision tasks. The neuron output of fully-connected layers of pre-trained CNN models are excellent representation of images and can better represent the most discriminative features that are useful for tasks like [18] image classification, scene classification etc;. The neuron outputs of convolution layers of pre-trained CNN models can better represent the regional features of the image and are useful for detecting objects in images and to draw bounding boxes over the object regions [19]. In this paper a novel simple deep face recognition (FR) model is introduced that makes use of deep convolution features.

In this work we try to stress the importance of deep convolution features and how they can help to improve the accuracy of the face recognition algorithm. Especially how the deep neural features can affect the task is studied in detail. Features from fully connected layers of the pre-trained model are obtained and then fed as input to the deep neural classifier. We propose to use the deep convolutional features to better represent the given facial image instead of using the traditional hand crafted features. The downside of the deep learning models is that they require large datasets to obtain better performance. To make use of deep learning models without the requirement of large datasets is to use pre-trained deep CNN models. For feature extraction different pre-trained Convolutional Neural Networks models like VGG16, VGG19, Resnet50, InceptionV3 are used and the concept of Deep Neural Networks model is used for classification. Based on the experimental results we observed that the proposed method is simple and gives better performance compared to the other methods. To show the performance of the proposed model, ORL dataset is used for the experimental studies. Based on the experimental results we claim that these unsupervised features better represent the images compared to the hand crafted features. In addition we also observe that Resnet features give better performance among the all pre-trained models used. The reason could it is deeper compared to the other models and hence could be able to better represent the images.

II. Background work

The objective of face recognition is to identify the class to which a human face belongs to. Several studies have been conducted in this area as there are large number of applications which directly or indirectly uses face recognition. Face recognition can be used as a part of many interesting and useful applications like authentication, criminal detection, HCI applications, and many more similar applications.

A genetic algorithm-based face recognition method is proposed in [1]. The authors proved that the proposed method has better generalization ability than PCA and LDA. In [5] a multimodal face recognition algorithm has been proposed. In[3] different distance measures are proposed for face recognition task. A two-dimensional PCA is introduced in [6]. In [7] a support vector-based classification is used for face recognition task. A multimodal face classification method is proposed in [8] that uses a two-dimensional Fisher linear discriminant. Discussed methods outperforms the existing models in addition it is computationally efficient.

A two-dimensional PCA based face recognition is proposed in [9]. In [10] wavelet features are extracted and then a KNN based classifier is used for classification. An ICA (independent component analysis) based gabor features are used in [11]. An orthogonal laplacian face method is used for face recognition in [12]. In [13] elongated LBP features are extracted in addition to a new feature AMDGM. ORL and FERET datasets are used for experimental results. Sift feature based facial recognition algorithm is introduced in [14].

III. Proposed work

In this paper a simple face recognition method is introduced that makes use of deep CNN based features. We call this method as deep facial recognition (deep FR). Similar to all the traditional face recognition methods, steps involved are: Data collection, preprocessing, feature extraction, training the model and evaluate the model with test data. To extract features, from the faces instead of using the hand-crafted features we use the deep CNN based features. Initially the face region is extracted from the given face images using viola jones face detection algorithm. Then from the cropped face regions, are fed as input to the deep CNN model after removing the last fully connected layer. In our work we have used 4 different pretrained models: VGG16, VG19, InceptionV3 and ResNet50. After extracting features deep neural network models are used for classification. Our experimental studies show that the deep feature-based FR method is simple and gives better accuracy compared to the other exiting methods that uses hand crafted features. Following is the brief description of the pretrained models used in our work.

VGG-16, a pretrained convolution model, is known for its performance improvement in various applications [15] like image classification, object detection etc;. It is trained on ImageNet dataset. VGG16 consists of a total of 13 layers with 5 convolution layers interleaved with pooling layers. Features are extracted from the FC2 layer of VGG16, which gives 2048-dimensional feature vector for each image. Static feature representation is used where all the images are represented with equal size.

VGG-19, pushed the number of layers from 16 to 19 which can improve the discriminative power of images. Google Net contains 22 layers. In this model a novel architecture called as inception is introduced. Inception V3 is an extension to this GoogLeNet. ResNet50 is deeper than the above models and contains 50 layers. This model uses residual layers and skip connection.

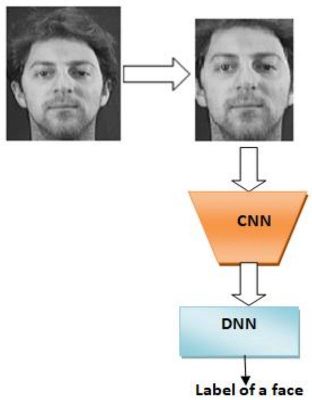


Fig1: Block diagram of proposed face recognition model

IV. EXPERIMENTAL ANALYSIS

The objective of this work is to identify the face label of a human face given an image of that human. Due to its applications in the real time world, this face recognition has become so popular in the recent past. In this paper we have proposed an algorithm for facial emotion recognition that makes use of deep neural based features. To prove the efficiency of our proposed method we use ORL dataset, the bench mark dataset for facial recognition, for experimental studies. Summary of the Dataset (ORL): To investigate the performance of the proposed method, benchmark dataset is used for experimental studies. The ORL bench mark dataset is used for our experimental studies as it is the standard dataset used extensively for the facial recognition task. This dataset contains a total of 400 faces of 40 different people. Each class contains 10 faces with different poses. The images are taken at different times with different poses and under varying lighting conditions and varying expressions. Figure2 shows few sample images from the dataset. Out of the total data, 80% is used for training and 20% is used for testing.

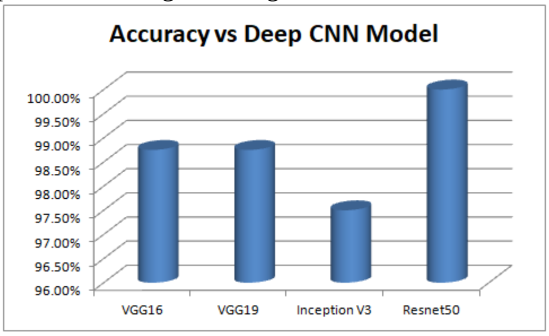
Figure2: Example faces from ORL dataset

To understand the performance of the proposed model, ORL dataset is used for our experimental studies. All the images of the dataset are of gray scale. The input image is given as input to the viola jones algorithm to get the face region part. Then these cropped images are fed as input to the pre-trained models to extract the most representative features from faces. We use VGG16, VGG19, InceptionV3 and Resnet models for our experiments. In all the cases the features are extracted from the last fully connected layers, to get a static representation of an image. These feature vectors are classified using DNN to get the label of face. Accuracy is used as the performance measure to understand the efficiency of the proposed method. After Feature extraction DNN is used with 3 hidden layers for classification. In all the hidden layers ReLu is used as the activating function. Softmax is used as the activation function at the output layer. 200 hidden neurons are used at each hidden layer. Cross entropy loss is used and adam optimization is used. The model is trained in batches with a batch size of 10.

Pretrained model	**#Features**	**Accuracy**
VGG16	4096	98.75%
VGG19	4096	98.75%
Inception V3	2048	97.5%
Resnet50	2048	100%

Table1: Accuracy on ORL datasets of pre-trained models for feature extraction

Table1 shows the performance of the proposed model for face recognition with deep CNN features. The accuracy obtained for 4 different methods are provided. With VGG16 and VGG19 models 4096 features are extracted and an accuracy of 98.75% is achieved. With Inception V3 there is a drop-in performance. Resnet gives 100% accuracy on the dataset even with less number of features used for representation of given image.



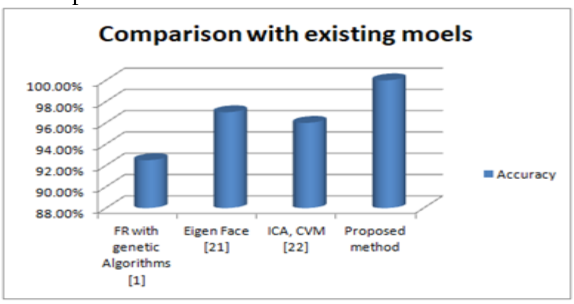
Graph1: Contrasting pretrained feature extraction models

Graph1 shows that Resnet gives 100% accuracy on the dataset when compared to the other models. The reason could be that even with less number of features used for representation of given image, the representations are more discriminative than others. In this subsection the deep feature based facial recognition method is compared with the existing methods. To understand the performance of the proposed deep feature-based algorithm we compare the model performance with the existing methods.

Model	**Accuracy**
FR using genetic Algorithms [1]	92.50%
Eigen Face [21]	97%
ICA, CVM [22]	96%
Proposed method (Deep FR)	**100%**

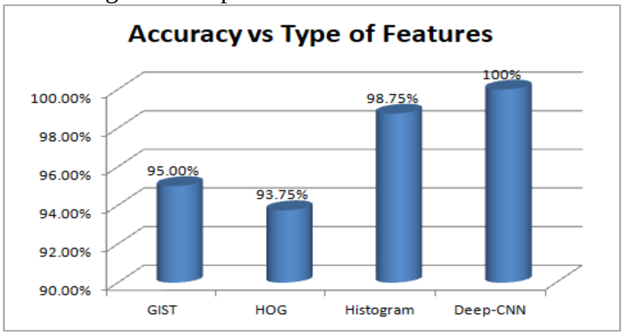
Table2. Comparison of the deep feature-based method with the existing models.

Table2 shows that the proposed FR method is better compared to the existing methods that uses hand crafted features. Based on the results we observe that the deep CNN features can better represent the facial features. The methods that uses these discriminative features has high generalization ability compared to the other methods. Graph2 shows that the proposed method is better compared to the existing methods. Based on the results we claim that our proposed method that uses deep features is far better than the other methods.



Graph2: comparison of the proposed with existing models

Table3 shows that the proposed method is better compared to the methods that uses hand crafted images. In this experiment we use different had craft features for face recognition task. Gist, HoG, Histogram based features are extracted and then SVM based classification is applied on them. Their performance is compared to the proposed Deep-CNN based Face recognition outperforms the other methods



Graph3: Comparison of Deep CNN based vs hand craft feature based models

Graph3 shows that the proposed Deep FR method is better compared to the methods that uses hand crafted images.

V. CONCLUSIONS AND PROSPECTIVE FUTURE

The proposed deep face recognition method gives better performance than the existing methods. The deep neural based features can represent the face better than the hand-crafted features. That is more discriminative features are extracted and hence the generalization ability is good. Especially the performance of ResNet based model is far better when compared to the other models. The deep features can represent the face features better than the hand-crafted features. In future we are planning to apply this on large datasets to test the scalability of the proposed model.

REFERENCES

1. Sukhija, Pratibha, Sunny Behal, and Pritpal Sing, "Face recognition system using genetic algorithm", Procedia Computer Science 85 (2016), 410-417.
2. Zhao, Wenyi, et al, "Face recognition: A literature survey", ACM computing surveys (CSUR) 35.4 (2003): 399-458.
3. Perlibakas, Vytautas, "Distance measures for PCA-based face recognition", Pattern recognition letters 25.6 (2004): 711-724.
4. N. Cristiana, T. Shawe, "An Introduction to Support Vector Machine, Cambridge University Press", 2000.
5. Pantic, Maja, and Leon JM Rothkrantz. "Toward an affect-sensitive multimodal human-computer interaction" Proceedings of the IEEE 91.9 (2003): 1370-1390.
6. Yang, Jian, et al. "Two-dimensional PCA: a new approach to appearance-based face representation and recognition." IEEE transactions on pattern analysis and machine intelligence(2004).
7. Guo, Guodong, Stan Z. Li, and Kap Luk Chan. "Support vector machines for face recognition." Image and Vision computing19.9-10 (2001): 631-638.
8. Xiong, Huilin, M. N. S. Swamy, and M. Omair Ahmad. "Two-dimensional FLD for face recognition." Pattern Recognition38.7 (2005): 1121-1124.
9. Zhang, Daoqiang, and Zhi-Hua Zhou. "(2D) 2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition." Neurocomputing 69.1-3 (2005): 224-231.
10. Chien, Jen-Tzung, and Chia-Chen Wu. "Discriminant waveletfaces and nearest feature classifiers for face recognition." IEEE Transactions on Pattern Analysis & Machine Intelligence 12 (2002): 1644-1649.
11. Liu, Chengjun, and Harry Wechsler. "Independent component analysis of Gabor features for face recognition." IEEE transactions on Neural Networks 14.4 (2003): 919-928.
12.] Cai, Deng, et al. "Orthogonal laplacianfaces for face recognition." IEEE transactions on image processing 15.11 (2006): 3608-3614.
13. Liao, Shu, and Albert CS Chung. "Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude." Asian conference on computer vision. Springer, Berlin, Heidelberg, 2007.
14. Geng, Cong, and Xudong Jiang. "Face recognition using sift features." 2009 16th IEEE international conference on image processing (ICIP). IEEE, 2009.
15. Johnson, Justin, Andrej Karpathy, and Li Fei-Fei. "Densecap: Fully convolutional localization networks for dense captioning." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
16. Zhao, Guoying, and Matti Pietikäinen. "Boosted multi-resolution spatiotemporal descriptors for facial expression recognition." Pattern recognition letters 30.12 (2009): 1117-1127.
17. Zhang, Ligang, Dian Tjondronegoro, and Vinod Chandran. "Random Gabor based templates for facial expression recognition in images with facial occlusion." Neurocomputing145 (2014): 451-464.
18. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
19. Johnson, Justin, Andrej Karpathy, and Li Fei-Fei. "Densecap: Fully convolutional localization networks for dense captioning." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
20. Douze, Matthijs, et al. "Evaluation of gist descriptors for web-scale image search." Proceedings of the ACM International Conference on Image and Video Retrieval. ACM, 2009.
21. Gupta, Sheifali, et al. "A new optimized approach to face recognition using eigenfaces." Global Journal of Computer Science and Technology 10.1 (2010)

AUTHORS PROFILE

Y. Venkata Raghav Rao has B. Tech in computer science and engineering from Nagarjuna University, M.Tech degree in CSE from JNTU and Ph.D. in computer science from JNTU, Hyderabad. He is working as professor in University. Rao research work was on distributed hyper ledger, IOT based fog computing, big data analytics. Rao has published more than 65 number of publications in various reputed journals and international conference proceedings and successfully supervised master's students and supervising PhD students.

Venu Madhav Kuthadi obtained his bachelor's degree in computer science and engineering from Nagarjuna University, India in 1998. He obtained his master's degree in computer science and engineering from JNT University, India in 2001.He completed his PhD in computer science in 1998 and obtained doctorate in engineering from University of Johannesburg in 2018. Currently he is working as a senior lecturer in the Department of CS & IS, BIUST, Botswana. Kuthadi's research work was on the network security to generate a security pattern to safeguard data. Kuthadi got more than 50 publications in peer reviewed journals and international conference proceedings and successfully supervised masters and Ph.D. students.

Rajalakshmi Selvaraj is working as a senior lecturer in the Department of CS & IS, BIUST, Botswana. Selvaraj has earned her doctorate in the area of network security using honeypot.
. Selvaraj also got Ph.D. in computer science. Her earlier work has included use of association rule in data mining. Dr Selvaraj collaborated her projects with research scholars globally journals, conference proceedings and book chapters, supervising good number of master and PhD students,