International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-07, Special Issue, MAY 2021

Transfer Learning for Recognizing Face in Disguise

[1]Prof. Vishal R. Shinde, [2]Miss. Sayali Raut, [3]Mr. Sahil Tikam

[1]Asst. Professor,[2,3]UG Student,[1,2,3]Computer Engg. Dept. Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharshatra, India.

[1]mailme.vishalshinde@gmail.com,[2]sayaliraut412@gmail.com,[3]sahilofficial101@gmail.com

Abstract- Face recognition is a method in Machine Learning to recognize objects in the image or video. People have a memory to understand other people and recognize some objects like creatures, plants, living things, and non-living things. Machine Learning is the method in Computer Vision that can be used, so computers can understand one person's face to another person contained in the image or video. In this paper, the author introduces testing some popular Convolutional Neural Network Model Architecture to see which one is better to recognize the person face dataset in disguised. The author uses the "Recognizing Disguised Faces" dataset to distinguish 75 classes of faces and then tries to train and test how accurate it can be recognized by the machine, where it will be useful to anyone who needs to search and develop an Architecture of Deep Learning. This paper is awaited to contribute to the field of Machine Learning like an algorithm that is applied to solve the problem in image classification. The trial results show significant improvement using transfer learning in VGG Models. Then gather that ImageNet weight is best used for face-recognizing using VGG Models.

Keywords - face recognition, transfer learning, deep learning, machine learning

I. INTRODUCTION

In the area of Machine Learning, there are many areas and ways of recognizing something, whether what is detected is a non-living or living thing (human, creature, plant). For biometric identification, the face of a person is used in recognizing someone else. People can identify someone else from one another because people have the memory and brain to process their thinking. But a machine cannot do that itself, hence begins a field that makes machine thinking, that is Machine Learning which is discovered by Arthur Samuel [1].

In the past decade, many methods to recognizing a person's faces which are Eigenfaces [2] and Principal Component Analysis [3], to Convolutional Neural Networks [4] which following that, the ability to recognize face became higher and higher. Transfer learning is a method applied in machine learning where the first training task produces a model, then does another test using the model of the first training task. Transfer learning is opposed to traditional machine learning because it entails using a pre-trained model as a springboard to start a subsequent task [5]. With the many benefits of using CNN, such as transfer learning, for example, CNN has done broadly used in various fields of research. Which are image classification [6], pedestrian detection [7], object detection [8], video analysis [9], food detection [10], and face recognition[4],[11].

In this paper, analyze some successful Pre- Trained CNN Model Architecture provided by Keras which is an open-source neural network library written in Python [12]. The architecture used is VGG16, VGG19, ResNet50, ResNet152 v2, InceptionV3 and Inception-ResNet V2. Later, break it into two parts: applying the vector to train the classifier model, and appraising the efficiency and cost function of the classifier model. From this research, expected to understand the most suitable Pre-Trained Architecture model with the highest level of exactitude, also the lowest cost function in the optimal hyperparameter situation. The author uses the "Recognizing Disguised Faces" dataset [13], which is a data set of 75 pictures of a person's face using a disguised tool like a bandana, masker, fake beard, glasses, etc. Every person in dataset mostly get 7-8 image as disguised and end 2 is his/her real face.

II. AIMS AND OBJECTIVE

a) Aim

Testing some famous Convolutional Neural Network (CNN) Model Architecture to understand which one is ampler to recognize the person face dataset in disguised. The author uses the "Recognizing Disguised Faces" dataset to identify 75 classes of faces and then tries to train and test how perfect it can be recognized by the machine,

 © 2021, IJREAM All Rights Reserved.

where it will bebeneficial to anyone who wants to explore and develop an Architecture of DeepLearning.

b) Objective

The objective of this project is to see the mostsuitable Pre-Trained Architecture model with the immense level of exactitude, and the cheapest cost function in the optimal hyperparameter situation.

III. LITERATURE SURVEY

Paper 1: Face Recognition with Convolutional Neural Network and Transfer Learning.

An age - invariant face recognition method using discriminative model with deep feature training is proposed. In this work, AlexNet is used as the transfer learning CNN model to learn high level deep features. These featuresare then encoded using a code book into a code word with higher dimension for image representation. The encoding framework ensures similar code word for the same person's face images photographed during different time scale. Linear regression-based classifier is used for face recognition and the method is tested on three datasets including FGNET which are publicly available. A face recognition system which incorporates theConvolutional neural network, auto encoder and denoising is proposed which is called Deep Stacked Denoising Sparse Autoencoders (DS-DSA).

Paper 2: Face Recognition Based on Two-stage CNN Combined with Transfer Learning.

In this paper, a method for face recognition under an unrestricted environment. The first stage aims at searching face windows and their bounding box regression vector. The second merge facial candidates which are covered highly. Then, the network finishes the task of face detection and alignment whose outputs are face regions and the position of the five facial landmarks. Experiments prove that our method of face detection and alignment is better than other methods. Secondly, faces are sent to the convolutional neural network for training. With very limited training data, retraining the entire network seems impossible.

Paper 3: Face Key Point Location Methodbased on Parallel Convolutional Neural Network.

Based on convolutional neural network and face detection algorithm, this paper proposes a training sample expansion strategy, and a parallel convolutional network face detection algorithm for face features, occlusion and illumination detection, combined with Relu activation function and Dropout random regularization strategy network, but also improves the generalization ability. The results show that the method is considerably improved in accuracy and authenticity.

IV. COMPARTIVE STUDY

Table 1: Comparative Analysis

SR NO.	PAPER TITLE	AUTHORNAME	METHOD	ADVANTAGE	DISADVANTAGE
1.	Face Recognition with ConvolutionalNeural Network.	R.Meena Prakash, N.Thenmoez hi, M.Gayathri	AlexNet is used as the transfer learning CNN modelto learn highlevel deep features	Handling multi- dimensional and multi-variety data	It has High dimensional feature vector.
2.	Face Recognition Based on Two-stage CNN Linkedwith TransferLearning	Anqi Zhou, Jianxin Chen,Jie Ding, Zhaolai Pan	A method for face recognition under an unrestricted environment.	Exceeds the traditional algorithms and general CNN insmall dataset.	Low detection rate.
3.	Face Key Point Location Method basedon Convolutional Neural Network	Zhou Pu, KaiWang, Kai Yan	Convolutional neural network and face detection algorithm	It can achieverobust and accurate estimation of key points.	Convolution cost ishigh.

V. EXISTING SYSTEM

In Existing System, disguise facedetection uses the PCA technique. The PCA when resembled with the quality PCA has an augmented recognition speed for face images with substantial modifications in countenance. within the PCA, the face pictures are broken into small sub-images and therefore the PCA is implemented to every one of those sub- images. Because many of the restricted countenances of an individual don't change even when the countenance vary, expect the stated technique to be ready to manage these changes. The performance of the quality PCA and modular PCA is set following the situations of varying appearance and pose applying regular face databases.

VI. PROBLEM STATEMENT

Create a project using transfer learning solving various problems like Face Recognition and Image Classification,

 © 2021, IJREAM All Rights Reserved.

applying CNN, like transfer learning, for instance, CNN has been broadly applied in various fields of research which is image classification, pedestrian detection, object detection, video analysis, food detection, and face recognition.

VII. PROPOSED SYSTEM

In the proposed system, we compare some popular Pre-Trained CNN Model Architecture provided by Keras which is an open-source neural network library written in Python. The architecture we used is VGG16, VGG19, ResNet50, and InceptionV3. Then, we divide into two parts: using the vector to train the classifier model, and evaluating the accuracy and cost function of the classifier model from this research, we expected to see the best Pre-Trained Architecture model with the highest level of accuracy, and the lowest cost function in the optimal hyperparameter state. In transfer learning, the information of an already trained machine learning model is implemented to a separate but relatedproblem.

VIII. ALGORITHM

Step 1: Start

Step 2: Loading a sample image

from tensorflow.python.keras.applicationsimport vgg16

from tensorflow.python.keras.applications.vgg16import preprocess_input

from tensorflow.python.keras.preprocessing.imag e import ImageDataGenerator, load_img from tensorflow.python.keras.callbacksimport ModelCheckpoint

from tensorflow.python.keras import layers,models, Model, optimizers

Step 3: Image Augmentation train_datagen <- ImageDataGenerator(

rescale <- 1./255, rotation_range <- 10,

zoom_range <- 0.1,

width_shift_range <- 0.1,

height_shift_range <- 0.1, horizontal_flip <- False, brightness_range <- (0.9,1.1),fill_mode <- 'nearest'

)

Step 4: training/validation accuracy and lossdef load_train_data(n):

lst <- []

for i in range (1, n+1):

filename <- str(i)+'.jpg'

x <- mpimg.imread('face reco/train/'+filename)

x <- cv2.resize(x, (160,160))lst.append(x)

df <- np.array(lst)

y_train <- pd.read_csv('face reco/train/y_train.csv',header <- None) y_train <- y_train.values

return df,y_train

Step 5: Making the predictiondef predict (path, filename):

x <- mpimg.imread(path+'/'+filename)x <- cv2.resize(x, (160,160))

x <- x.reshape(1,160,160,3)

x <- embedder.embeddings(x)

if(max(model.predict_proba(x)[0])<0.25):#0.25 is the threshold

print ("Face not found!")return

Step 6: End

IX. MATHEMATICAL MODEL

1. Artificial Neural Network (ANN)

An artificial Neural Network may be a processing system that has several performance characteristics almost like our biological brain neural networks. McCulloch & Pitts first designed ANN in 1943. ANN has been incurred as a mathematical model of generalization of neurobiology or human cognition, then supported this assumption:Neurons are a component where information science occurs. All attachment link has an associated weight that flows the transmitted signal.

- Neurons are an element where information processing occurs.
- Each connection link has an associated weight that streams the transmitted signal.
- The connection between neurons is called the connection link that passesa signal.
- To determine each neuron output, it implements an activation function that is usually non-linear to the input of its network.

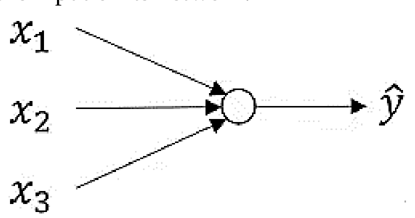

Fig.1: Single Neuron network

Let's see the simple neural network in Fig.1, let There are many elements of the processing units in ANN which are commonly called neurons, units, cells, or nodes. Every neuron is connected via a communication link and associated with the load. Weight represents the information that will be used to solving problems, and among the widely used neural networks implementation is for the pattern classification challenge output neurons be

 © 2021, IJREAM All Rights Reserved.

ŷ, it then receives input from activation function neuron a. And respectively a get the input from three other activation x1, x2, and x3. Which symbolize X1, X2, and X3 for their neuron's name. Moreover, the weight associated from X1, X2, and X3 to activation function are w1, w2, and w3. So, the calculation for output can be denoted by (1).

$$\hat{y} = a = w1x1 + w1x1 + w1x1 \quad (1)$$

After that, calculate the loss function network above. The loss function is a measure of the difference between the prediction of ŷ, and the true value (ground truth), in other words, it is an error calculation for one stage of training. This function can be seen (2).

$$L|\left(\hat{y}^{(i)}, y^{i}\right) = \frac{1}{2}\left(\hat{y}^{(i)}, y^{i}\right)^{2} \qquad (2)$$

In this research, use categorical cross-entropy as a loss function, because we want to classifyeach person by his/her face. This function will compare the distribution of predicted face, by true and false which is set to 1 for true and 0 if false. The true class of a person'sface represents as a one-hot encoded vector, which is we get the lower loss if the model output vector is closer to the true class. The loss function is as follows:

$$L\left(X_i, Y_i\right) = -\sum_{j=1}^{C} y_{ij} * \log(p_{ij}) \qquad (3)$$

$$y_{ij} = \begin{cases} 1, & if\ i_{th} \text{ element is in class } j \\ 0, & if\ i_{th} \text{ element is not in class } j \end{cases}$$

Class symbolizes by C, where Xi is the input vector for one-hot encoded target vector Yi, and pij is a probability that Ith element in class j.

2. Convolutional Neural Network

Convolutional Neural Networks is one of the special cases of the Artificial Neural Network which is currently considered the best technique to solve object recognition and digit detection problems [4]. In a Neural Network as deep as CNN many models are being developed until now, but in this paper, just focus on 3 different model architectures.

2.1 AlexNet

AlexNet Created in 2012, this architecture is that the first deep network which will classify some objects with significant accuracy withinthe ImageNet dataset, compared to traditional methodologies that were before AlexNet. This network includes convolutional layers supported by three completely combined layers [15].

2.2 VGG16

Created in 2013, this architecture comes from the VGG group, Oxford. VGG was made to improve the AlexNet architecture by replacing large kernel filters with some 3x3kernel filters. With a given receptive field, small-sized kernels that are stacked are better than large-sized kernels, because several non-linear layers increase the depth of the network which makes it possible to learn more complex features [12].

2.3 GoogLeNet / Inception

Created in 2014, while VGG achieved phenomenal accuracy in the ImageNet dataset, but its use requires high computation, even though it uses a GPU (Graphics Processing Unit). This has become inefficient due to the large width of the convolutional layer used. GoogLeNet builds on the idea that most activations in deep networks are not needed (zero value) or excessive because of the correlation between them [16]. GoogLeNet designed a module called the Inception module which numbered roughly like a thin CNN with solid construction. Because only a small fraction of the neurons is effective as mentioned previously, the width/number of convolutional filters of the kernel size is small. This module also uses convolution of various sizes to capture details at various scales (5x5, 3x3, 1x1).

2.4. ResNet

Following what has been discussed so far, namely, to improve accuracy in the network must increase the intensity of the layer, as long as it can keep over-fitting. However, increasing the deep network does not work by directly adding layers. Deep networks are difficult to practice because of the problem ofdisappearance gradients, where gradients are re-propagated to the previous layer, repeatedrepetition can make the gradient very small. As a result, as the network grows, the performance becomes saturated or evenbegins to degrade quickly.

X. SYSTEM ARCHITECTURE

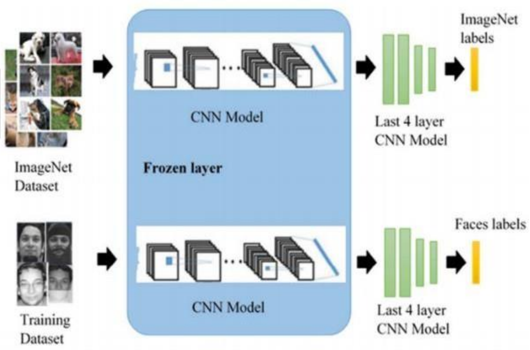

Fig.2: System Architecture for RecognizingDisguise Faces

Use 2 setups for training, which is Freezing all layers in CNN Model (Setup 1) and Unfreezing the last 4 layers (Setup 2). Freezing all layers mean that use the weight from the previous training to CNN Model.

XI. ADVANATGES

1) Helps solve advanced real-world issueswith many constraints.

2) It mechanically detects the vital optionswith none

© 2021, IJREAM All Rights Reserved.

ISSN : 2454-9150 Vol-07, Special Issue, MAY 2021

human superintendence.

3) It mechanically detects the vital optionswith none human superintendence.

XII. DESIGN DETAILS

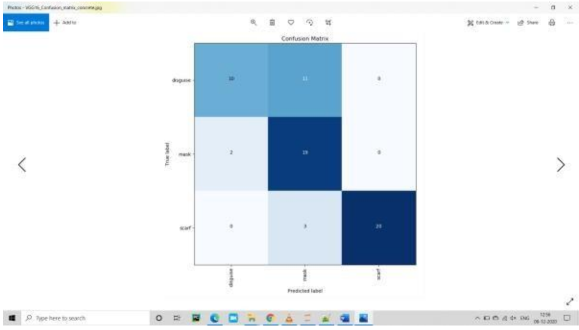

Fig.3: VGG16 Confusion matrix page

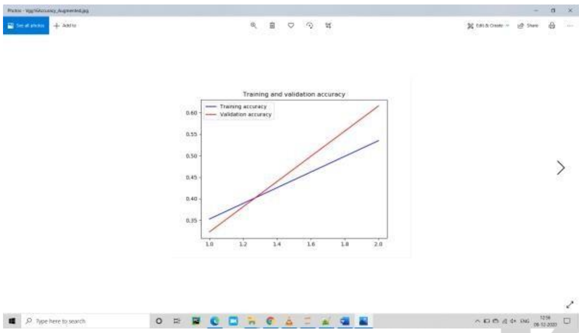

Fig.4: VGG16 Accuracy

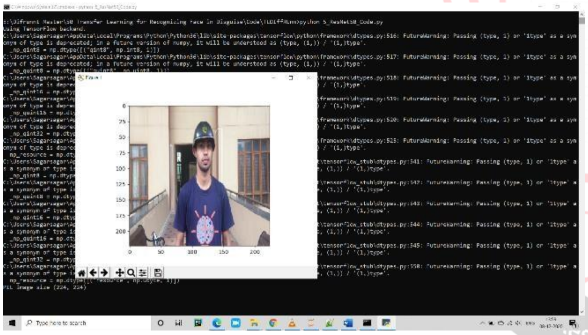

Fig.5: ResNet50 Loa

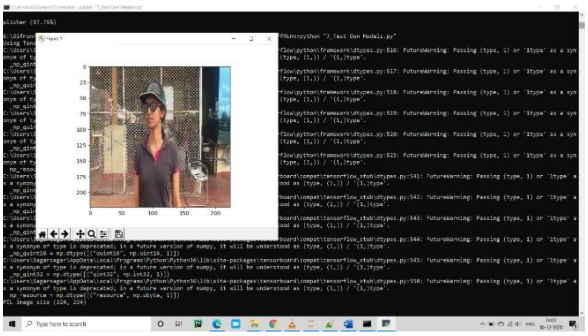

Fig.6: Testing Own Model

XIII. CONCLUSION

Thus, we have tried to implement the paper “Fauzan Nusyura, I Ketut Eddy Purnama, Reza Fuad Rachmadi”, “*Transfer Learning for Recognizing Face in Disguise*”, IEEE 2020 and according to the implementation the conclusion is as follows in training, the result shows that the VGG model has a balanced accuracy of training and validation, and on the other side, the ResNet152 v2 Model has better efficiency than VGG in the train set. But in the test result shows, that the VGG model is the highest performance than other CNN Models. ImageNet weight can be used for Transfer Learning to Recognize face using VGG Model.

REFERENCE

[1] Samuel, "Some Studies in Machine Learning Using the Game of Checkers," IBM Journal of Research and Development, vol. 3, no. 3, p. 210–229, 1959.

[2] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, p. pp. 586–591, 1991.

[3] J. Yang, D. Zhang, A. F. Frangi and J. Yang, "Two dimensional PCA: a new approach to appearance-based face representation and recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, p. 131– 137, 2004.

[4] Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," CVPR, p. 1701–1708, 2014.

[5] J. West, D. Ventura and S. Warnick, "A Theoretical Foundation for Inductive Transfer," Spring Research Presentation, 2007.

[6] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama and a.T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in ACM MM, 2014.

[7] D. Tomé, F. Monti, L. Baroffio, L. Bondi, M. Tagliasacchi and S. Tubaro, "Deep Convolutional Neural Networks for pedestrian detection," Signal Processing: Image Communication, vol. 47, pp. 482-489, 2016.

[8] Z. Deng, H. Sun, S. Zhou, J. Zhao and H. Z. Lin Lei, "Multi-scale object detection in remote sensing with convolutional neural networks," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 145, no. Part A, pp. 3-22, 2018.

[9] M. Perez, S. Avila, D. Moreira, D. Moraes, V. Testoni, E. Valle, S. Goldenstein and A. Rocha, "Video pornography detection through deep learning techniques and motion information," Neurocomputing, vol. 230, pp. 279-293, 2017.

[10] R. D. Yogaswara and A. D. Wibawa, "Comparison of Supervised Learning Image Classification for Food and Non-Food Objects," in CENIM, Surabaya, 2018.

[11] Z. Yang and R. Nevatia, "A multi-scale cascade fully convolutional network face detector," in International Conference on Pattern Recognition, Mexico, 2016.

[12] F. Chollet, Keras, GitHub: GitHub repository, 2015.

[13] T. Dhamecha, A. Nigam, R. Singh, and M. Vatsa "Disguise detection and face recognition in visible and thermal spectrums," in IEEE International Conference on Biometrics, pp. 1–8, 2013.

[14] Q. Le, J. Ngiam, Z. Chen, D. Chia, and P. Koh, "Tiled Convolutional Neural Networks," in Neural Information Processing Systems Foundation, 2010.

[15] Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional," 2012.

[16] Szegedy, Y. Jia, and S. Reed, "Going deeper with convolutions," in 2015 IEEE Conference on Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015.

 © 2021, IJREAM All Rights Reserved.