FACE ASSESSMENT LEARNED FROM EXISTING IMAGES IN ORDER TO CLASSIFY THE GENDER OF THE IMAGES BASED ON IMPROVED FACE RECOGNITION

N. Srivani, Dr Prasadu Peddi

Research Scholar, ShriJagdishprasadJhabarmalTibrewala University, Rajasthan.

Assistant Professor, Dept of CSE, ShriJagdishprasadJhabarmalTibrewala University.

vani.medipally@gmail.com

ABSTRACT:

Face is among the most vital biometric traits. When we study the face, we can gather details such as age and gender, ethnicity, expression, identity, etc. A system for gender classification uses the faces of people from an image to identify what nature of the gender (male/female) that is the individual. A gender-based approach that is successful can improve the performance of numerous other applications, including facial recognition and a sophisticated human-computer interface. This paper provides the fundamental processing steps required for gender classification based on face images. In this paper, a variety of techniques employed in different stages in gender categorization, i.e., features extraction as well as classification are presented and contrasted.

Keywords: Biometrics, feature, extraction,classifier.

1. INTRODUCTION

Face Recognition (FR) is among the most popular applications of understanding and analyzing images that have attracted considerable attention, particularly in past few years. In image processing, FR is among the most vital biometric traits and is increasingly used for various applications. It now used in multiple areas like surveillance. Over the past few years, gender-based identification on facial images has become a hugely important issue and has become a well-known field of study. In this section, we'll examine different methods for gender classification and the use of various facial features like eyes, nose, mouth, etc., to determine gender by using machine learning methods.

Face detection that is successful within an image using an image of a frontal view of a human face. Unfortunately, when you have static images there is a vast range of possible positions of a face within the image. Faces could be large or small and can be placed in any direction in the middle of the image, from left up to right corner of the image.

The majority of face detection systems employ an approach based on examples to determine if faces are present in the display at the moment in time. While it's quite easy to find faces, how does you find a reliable image that represents non-faces.Thus, systems for face detection using the principle of example-based learning require millions of "face" as well as 'non-face images to train effectively.

II Related Work

To speed up the learning process and speed up learning, the normalization layers within the network built in LRN (local response normalization) (LRN)are changed to normalization layers for batch training [30]. A batch normalization layer is an integral component of the model and is responsible for normalization of each mini batch. This allows the network to be trained at greater rate of learning that otherwise is difficult and unstable when using regular LRN layers. The use of batch normalization was proven to provide the same high-quality performance but using 14 times fewer training repetitions. It is a great way to improve the performance of your custom CNN. The overview of the architecture of the proposed custom CNN is shown in Figure1.Subsequently, the layers, size of kernels, number of kernels, and strides for each layer used in custom CNN architecture, all in all the custom CNN contains seven convolution layers as well as 6 batch normalization layers that are in between each layer of convolution. Between each convolution layers, a maximum layer of pooling is included to limit the size of the input is used to create subsequent layers. The activation function used in the system uses a rectified linear unit (ReLu). Many dropout layers are integrated into the ReLu to prevent overfitting. To show the fact that this specific CNN is more effective in its parameters than other CNN that are utilized to create the customized CNN layout are calculated. This article provides in-depth information on ways to estimate the amount of the output features of each layer, as well as methods to estimate the quantity of parameters related to convolution layers and fully connected layers within the CNN.

III Methodologies

The recognition of gender in faces is a major issue in the machine vision domain. The first gender recognition technology within computer vision mostly built on machine learning from a neural network. One of them was a neural network with two layers. SEXNET developed by Gollomb [1]. Its face-to-face input dimensions were 30x30 and the accuracy of classification was 91.9 percent tested using 90 images of faces that were 45-man faces and 45-woman images, respectively. Another neural network developed by Edelman [2] was trained using 3 different faces, that is the top half and bottom half of face images, and the entire face image. The linear neural cells employed within the system, and three different results of classification were evaluated.

Face detection that is successful within an image using an image of a frontal view of a human face. Unfortunately, when you have static images there is a vast range of possible positions of a face within the image. Faces could be large or small and can be placed in any direction in the middle of the image, from left up to right corner of the image.

The majority of face detection systems employ an approach based on examples to determine if faces are present in the display at the moment in time. While it's quite easy to find faces, how does you find a reliable image that represents non-faces (Rowley and co. 1996)? Thus systems for face detection using the principle of example-based learning require millions of "face" as well as 'non-face images to train effectively.

IV PROPOSED WORK

Data preprocessing.
The data we gathered was received from participants' self-scores before an examination. We scrutinized our data to identify three kinds of incorrect trials. These included there was An incorrect response was typed, and they were able to score themselves as 'correct as well, and the third) There was no answer. However, the participant was able to score themselves as correct. In all of these cases, if more than that 50 percent of trials revealed If this pattern is seen in any participant, the entire trial was taken off for further investigation (2.71 percent of participants). If the trial had a score below 50% during these tests the trial was eliminated (3 per cent of trials) instead of participants. We wanted to evaluate the effectiveness of participants' FR, we only looked at faces that they had been exposed to and who they knew. We eliminated faces from which participants said they were not familiar with
A t-test independent of the study revealed that males with fame (M = 5.51, SD = 0.99) were significantly higher in famous score (t (67) = 2.66, P = 0.010 (d = 0.66) and then popular women (M 4.89, SD = 4.89 + SD 0.82). Additionally, we report that fame score and the age of the celebrities were significantly related (r = 0.33) this suggests that celebrities who are older are more popular.

We also observed that the famous males were substantially older than females with well-known names. This suggests that subjects may have had exposure to males who were more famous than females. In addition to the typical accuracy scores for statistics, our study also attempted to assess the accuracy of our analysis after taking into account fame to consider the bias resulting from either or both or both of these factors.

Calculation of facial embeddings

It is the Convolution Network used to downsize images and transform an (224*224*3) size image (which is used as an input to the network) in to the (128*1) Vector of the the same size. The same design employed in the Facenet model proposed by those who wrote the paper has utilized here.[4]. In short, the network is comprised of an 3D convolution filter, then 2 Max-Pooling and Batch Normalization layers. Then, there are Ten Inception module. Each module is comprised of 5*5 convolution filters and 1*1,3*3. The results of all the filters is mixed and then transferred to the layers below. The resultant image, after being passed through the Inception Modules is 7x7x1024 pixels. The 1024 channels that make up the picture are subject the average of pooling. The result is a vector that has 1024 dimensions. The 1024 output units pass through an extensive network that is fully connected and has 128 output units, which have an activation layer of output ReLU. The output of all nodes is combined into an 128-dimensional vector. These layers are more complex and assist network in understanding deeper representations of the image. To speed up back propagation residual connections are made between layers. Transfer Learning is applied as well as the weights that were learned from Facenet model in the initial training are being used in this research. In a nutshell, an image is 224*224*3 transferred into network, resulting in an 128-dimensional vector at the other end that contains an embedding for facial features of the image input. This network has 24 layers deep and was created with the efficiency of computation in mind of. Original authors selected 128-dimensions for their facial embeddings since they were able to achieve greater precision, and the identical dimensions will be employed in this paper.

One of the greatest challenges when it comes to categorizing gender with facial photos is the impact on the posture of the person, their illumination, and background noise. While neural networks are utilized for self-representation, they are also subject to specific physical conditions that pictures with soft input. By pre-processing, this technique will assure that all images input are consistent. This increases the effectiveness of Machine Learning Techniques.

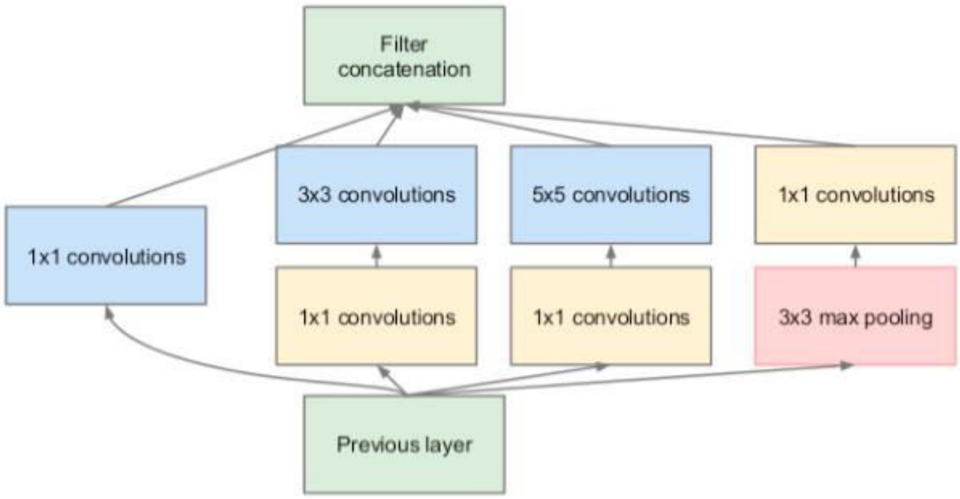

Fig 1. Inception Modules

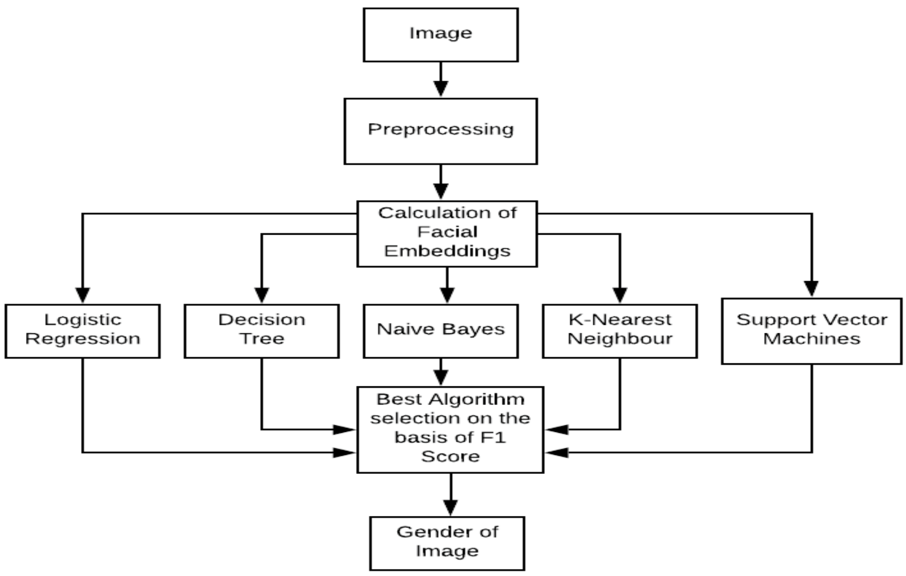

Fig. 2. Architecture Diagram of Proposed Model

Table1 Several existing gender recognition methods and their accuracy

Feature	Dataset	No. of Sample	Accuracy
Neural Network	Personal	90	91.90%
Boosted LBP	LFW	7443	94.81%
Haar Features	Web Images	3500	79.00%
NABP	LFW	13000	92.74%
Fisher Vector encoding	LFW	13233	92.50%
Color	FERET	1762	98.00%
MSLBP	FERET	1762	95.19%

Proposed Algorithm CNN with ensemble module: -

Algorithm CNN Algorithm

Require: image I_{in}

Ensure: image I_{out}

1: if I_{in} is colored then

2: I_{yuv}<= Convert I_{in} from RGB to YUV

3: R <= Y channel of Iyuv

4: else

5: R<= I_{in}

6: end if

7: RH_{in}<= Calculate histogram of R

8: Divide RH_{in} into R_{HL} and R_{HU} using Harmonic Mean

9: Apply HE separately on R_{HL} and R_{HU}

10: R_{out}<= $R_{HL} * R_{HU}$

11: if I_{in} is colored then

12: R_{out}<= Combine R_{out} with U and V channels

13: I_{out}<= Convert R_{out} from YUV to RGB

14: end if

15: I_{out}<= R_{out}

Dataset

We will make use of a small portion of the Celebrity face dataset to conduct our research. Celebrity Face is a massive set of data on the face attributes of celebrities. It comprises more than 100k celebrity faces that each have 40 face attribute binary annotations including gender as one of the 40. We've chosen an unintentional subset of 200 faces that have 100 female and male faces are each used for the purposes of this study. Let's look at the distribution of testing and training splits. The first step is to execute our program on a training-set comprising 160 images. The test set of forty images. Label the image in dataset using label=0. This is for “female" and label=1 is for male. The same process is followed for the LFW face datasets as well.

Deep CNN (D-CNN) has been used in this area also. It is a method of estimating articulated poses and body configuration parsing. facial parsing, face recognition as well as identification of things,, of paths, the detection of disease with the help of images of leaves on plants and facial recognition using the human facial facial key-point detection . Speech recognition.

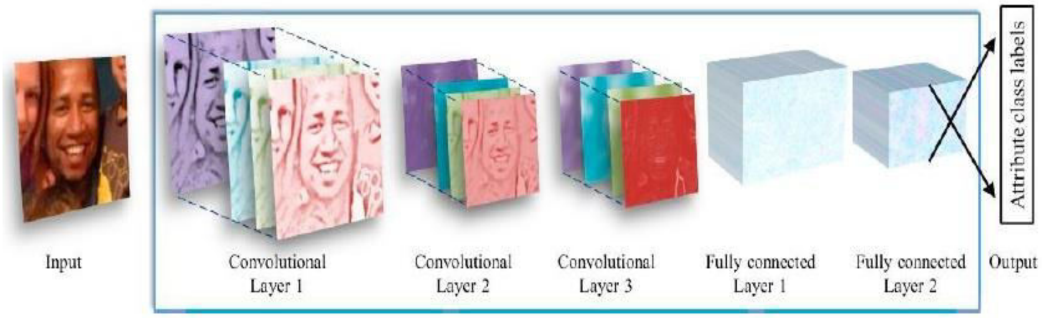

Figure 3: How CNN work with image for Gender Recognition

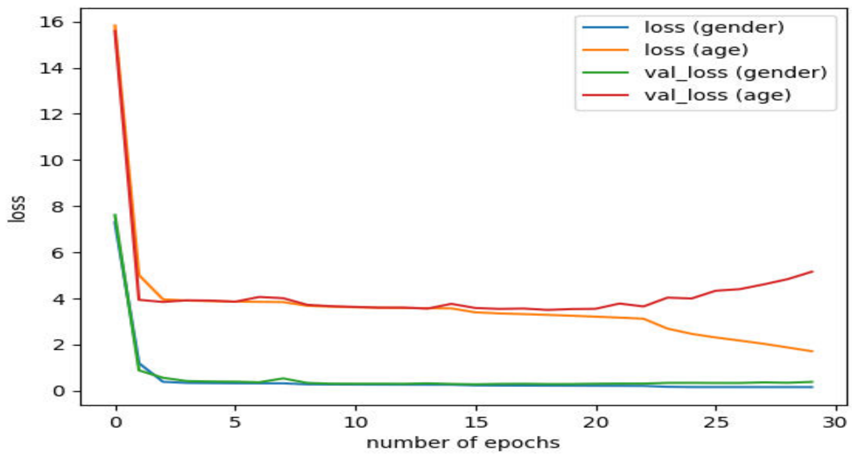

Figure: Loss of the error rate from the neural network model

Gathering a large and identified image from the dataset used for training, estimate gender based on the collection of social images or images that do not need access to personal information of people visible in the pictures such as their birth date, and traditional methods that involve gathering other information of an individual and the basis of these, we can determine gender of manually labeled sets of data for gender recognition. This is why we use DCNN that is directly dependent on an image and helps in the estimation of gender with accuracy. Over fitting is a minor issue that typically occurs when deep learning or machine learning-based methods are applied with a small number of face images in our database.

Figure4 : Network Architecture for Gender classification

Images are then scaled again to the size of 256 x 256 images, and then it is used to crop the 227 x 227 before it is transmitted onto the network. These three layers are described as

The following fully connected layers are described in the following manner:

The first procedure is to construct the FC layer that will receive the output from the third convolutional layer which contains 500 neurons and is superseded by an activation function known as Rectified Linear Unit(Relu) and the dropout layer.

Following that, you need to design another FC layer that will receive information in 512 dimension from the prior FC layer. The same process that was followed for the initial layer.

The next step is to construct an unconnected, fully connected layer, which is then mapped to the last classes to identify gender.

The result of the previously connected layer is redirected towards the function of soft-max, and the function assigns probabilities for each label in order to facilitate gender recognition. The anticipation is created using the most likely label to be at risk over the remainder of the test image used in gender recognition.

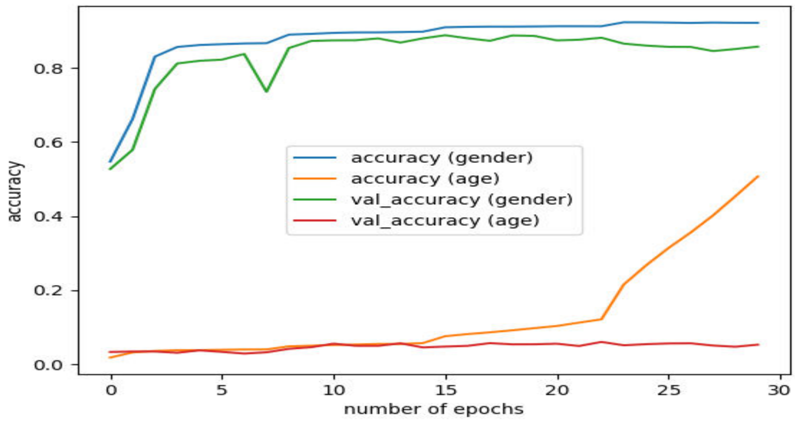

Figure5 : Accuracy of the model trained on the data set and on the testing data set

Figure 6: Model testing on the Validating data

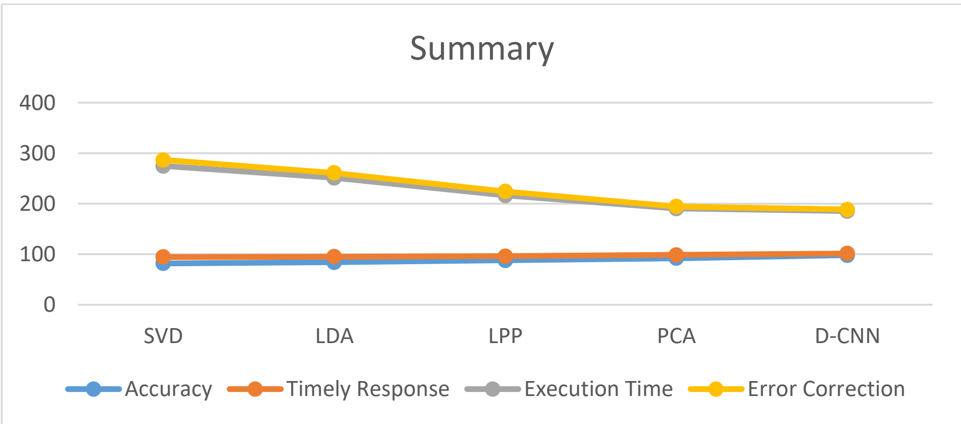

Figure 7: Summary of the algorithm usage and their performance

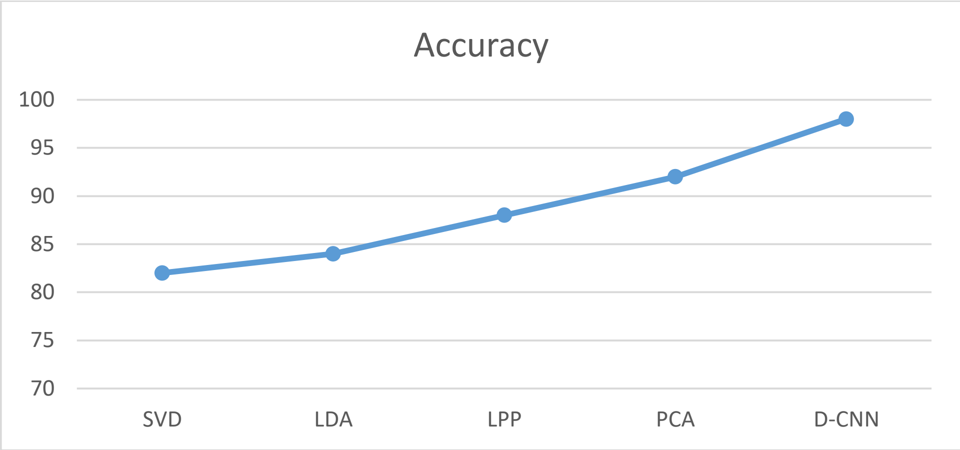

Figure 8: Accuracy of the algorithms when compared to existing models to proposed model

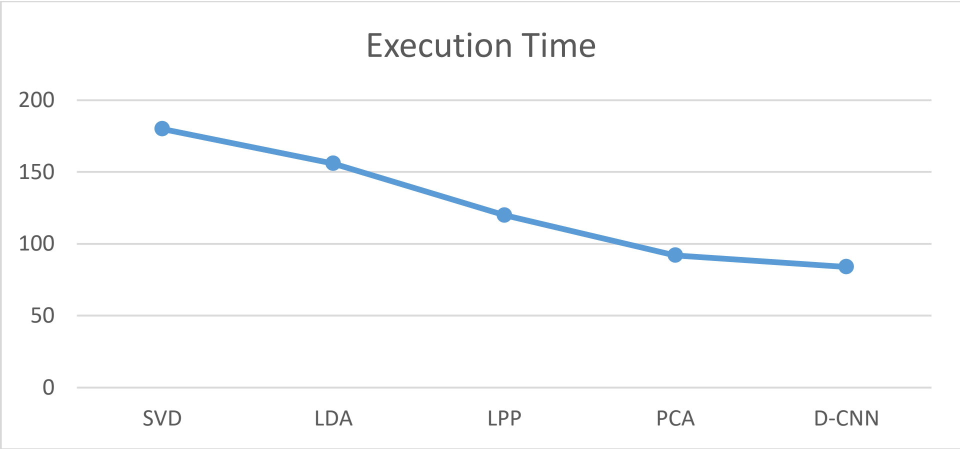

Figure 9: Execution time and performance of the algorithms when compared to existing models to proposed model

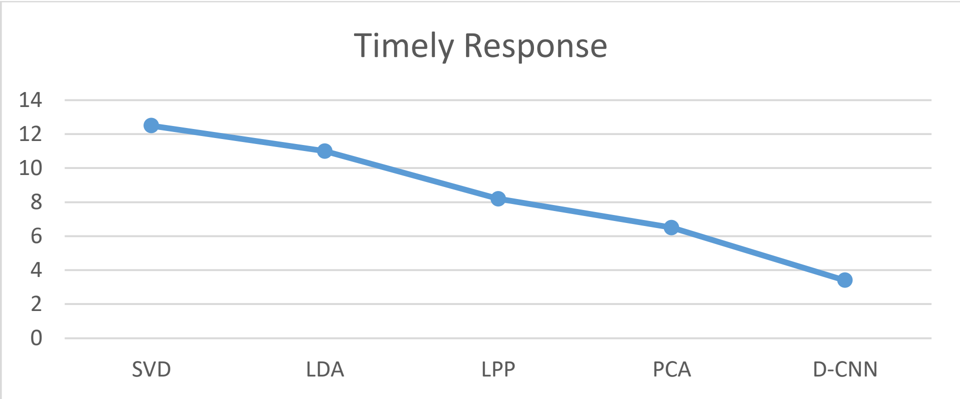

Figure 10: Timely Response of the algorithms when compared to existing models to proposed model

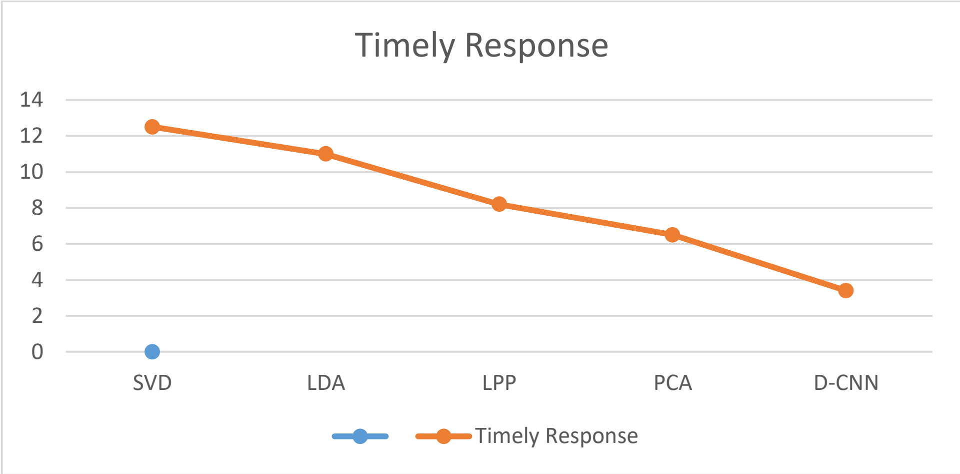

Figure: 11 Timely Functionality of the algorithms when compared to existing models to proposed model

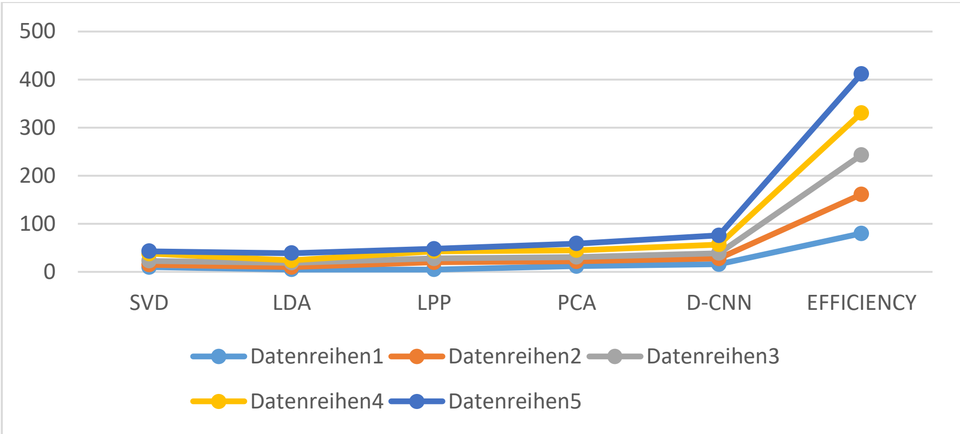

Figure 12: Summary of the algorithm usage and their performance

VCONCLUSION

This paper describes how we developed and developed a self-organized map algorithm. This algorithm allows for better understanding of the information gathered by the standard SOM neurons. This method can not only help to understand the characteristics of SOM clusters but can also depict the entirety of SOM neurons and their similarities or differences to the initial data set. We built a neighbourhood graph with the help of SOM neurons to illustrate potential self-organized routes to travel within the large, densely filled image space. The graph visualization technique provides specific information regarding the size and characteristics of the clusters that represent the data that is being investigated. The proposed algorithm may be an effective tool for SOM analysis. It provides a clear explanation of the topologically constrained manifolds modelled by SOM and reveals some perceptual characteristics that are common to well-framed facial image analysis, such as facial expressions and ethnicity.

REFERENCES

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face Description with Local Binary Patterns: Application to Face Recognition. IEEE Trans on PAMI, 28:2037–2041, 2006. 1, 3, 5

[2] P. Belhumeur, D. Jacobs, D. Kriegman, and N. Kumar.Localizing parts of faces using a consensus of exemplars. In CVPR, pages 545–552. IEEE, 2011. 2

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Trans on PAMI, 1997. 1, 2, 3, 5

[4] T. Berg and P. N. Belhumeur.Tom-vs-Pete Classifiers and Identity-Preserving Alignment for Face Verification.In British Machine Vision Conference, 2012. 1, 7

[5] Y. Boureau, F. Bach, Y. LeCun, and J. Ponce. Learning midlevel features for recognition. In Computer Vision and Pattern Recognition, pages 2559–2566. IEEE, 2010. 2, 3

[6] Y. Boureau, N. Le Roux, F. Bach, J. Ponce, and Y. LeCun. Ask the locals: multi-way local pooling for image recognition. In ICCV, pages 2651–2658, 2011. 3

[7] X. Cao, Y. Wei, F. Wen, and J. Sun. Face alignment by explicit shape regression. In Computer Vision and Pattern Recognition, pages 2887 –2894, June 2012. 1, 2

[8] Z. Cao, Q. Yin, X. Tang, and J. Sun. Face recognition with learning-based descriptor.In Computer Vision and Pattern Recognition, pages 2707–2714, 2010. 1, 3, 5

[9] C. Chan, J. Kittler, and K. Messer.Multi-scale local binary pattern histograms for face recognition. Advances in biometrics, pages 809–818, 2007. 2

[10] D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun. Bayesian face revisited: A joint formulation. In European Conference on Computer Vision, pages 566–579, 2012. 1, 2, 3, 5

[11] L. Clemmensen, T. Hastie, D. Witten, and B. Ersboll.Sparse discriminant analysis.Technometrics, 2011.

[12] A. Coates, H. Lee, and A. Ng. An analysis of singlelayer networks in unsupervised feature learning. Ann Arbor, 1001:48109, 2010.

[13] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In Computer Vision and Pattern Recognition, volume 1, pages 886–893, 2005.

[14] M. Douze, H. Jegou, H. Sandhawalia, L. Amsaleg, and C. Schmid.Evaluation of gist descriptors for web-scale image search.In International Conference on Image and Video Retrieval, page 19, 2009. 2

[15] J. Friedman. Greedy function approximation: a gradient boosting machine. Ann. Statist, 2001.

[16] Prasadupeddi(2020), MINING POSTS AND COMMENTS FROM ONLINE SOCIAL NETWORKS, Turkish Journal of Computer and Mathematics Education, Vol 11, No 3, pp: 1018-1030.

[17] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker.Multi-pie.In International Conference on Automatic Face and Gesture Recognition, 2008.

[18] M. Guillaumin, T. Mensink, J. J. Verbeek, and C. Schmid.Automatic face naming with caption-based supervision. In CVPR, pages 1–8, 2008.

[19] Prasadu Peddi (2019), Data Pull out and facts unearthing in biological Databases, International Journal of Techno-Engineering, Vol. 11, issue 1, pp: 25-32.

[20] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. The Journal of Machine Learning Research, 3:1157–1182, 2003.

[21] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang. Face recognition using laplacianfaces. IEEE Transactioins on Pattern Analysis and Machine Intelligence, 27(3):328–340, 2005. 2

[22] C. Huang, S. Zhu, and K. Yu. Large scale strongly supervised ensemble metric learning, with applications to face verification and retrieval. In NEC Technical Report TR115, 2011.

[23] Yogesh Hole et al 2019 J. Phys.: Conf. Ser. 1362 012121