




A Proposed Framework for Identity Verification in Passport Management Using Model Scaling and Semantic Similarity



Rishita Khurana¹, Madhulika Bhatia¹, Manika Battan¹, Nripendra Narayan Das^{2*}, Meenakshi Memoria³

¹ Department of Computer Science and Engineering, Amity University, Noida 201303, India

² Department of Information Technology, Manipal University Jaipur, Rajasthan 303007, India

³ Department of Computer Science and Engineering, Uttaranchal University, Uttarakhand 248007, India

Corresponding Author Email: nripendradas@gmail.com

<https://doi.org/10.18280/isi.280211>

ABSTRACT

Received: 19 September 2022

Accepted: 30 November 2022

Keywords:

efficient net, face detection, image similarity, neural networks, passport management

The dependence on images for face detection and identity verification was at odds with broad assessment which showed that coordinating with sets of non-similar images was profoundly inclined to make mistakes. The current passport management lacks security of data and requires more manpower and time, manual calculations, and verification of documents as well as the photographs. Additionally, the computerized passport examination requires accurate processing in order to carry out crucial tasks like identifying passport forgeries, looking for wanted criminals, or finding those who are ineligible for immigration, among other things because the utilization of fake passports addresses a huge danger to the security of the nation. Therefore, the purpose of this paper is to enhance the identity verification process in passport management. The paper comprises implementation of various technologies and technology driven models for serving its purpose by using the latest and efficient techniques that includes neural networks, face detection, image similarity, ROI (region of interest) and model scaling. The dataset utilized in the paper was self-prepared by the authors and consisted of 5 images each for 10 different individuals. It had been trained from scratch on different models based on CNN architecture. Out of all the three models used for training, the model that resulted in highest accuracy was further considered for research. EfficientNetB0 resulted in 82% accuracy being the highest out of all the three models. The threshold for the model was thereby calculated using z-score which compared the similarity scores and classified the images as similar or non-similar. Therefore, it can be concluded that the techniques utilized in the paper are efficient enough to determine the similarity of an image with its corresponding image.

1. INTRODUCTION

The face is a significant component of the human body and must be detected for several purposes, including security and forensic analysis. The challenges of various facial expressions, position fluctuations, occlusion, ageing, and resolution, whether in the frame of a stationary item or video sequencing images, call for the use of correct approaches for face identification and recognition. Face is the fundamental point of convergence of thought in public activity having a critical part in passing on character and sentiments. Different appearances which are learnt all through our life can be perceived and faces can be recognized at the very first moment by following quite a while of partition. This ability is very vigorous not withstanding huge varieties in visual improvement because of evolving condition, maturing and interruptions like facial hair, glasses, or changes in hair. Computers that distinguish and perceive countenances could be applied to a wide assortment of undertakings including criminal ID, security framework, face similarity, labeling purposes, and immigration system. Lamentably, fostering a model of face detection and face similarity and acknowledgment is very troublesome since countenances are tangled, complicated and multidimensional. Face

identification and verification is utilized in various places nowadays, especially the sites enabling pictures such as Instagram, Snapchat and Facebook.

The naturally labeling highlight adds a different feature for dividing pictures between individuals who are in the image and then gives the idea to others about the individual who is in the picture. In this undertaking, a beautiful yet basic model is considered and carried out which identifies the facial features of images and compares and computes the similarity of facial features of one image with another image. The model also contributes to the effortless renewal process in passport management and security framework for identification in passport management systems. The capacity to deliver a safe, dependable type of distinguishing proof on demand is underestimated by numerous residents, particularly those living in nations with cutting edge economies.

This ability gives various improvement advantages to people, from getting to government and business administrations to building up their right of home and work in an area. Moreover, cross country utilization of solid methods for recognizable proof can assist with combatting wrongdoing and unlawful migration. Several efforts have been taken by a few countries to introduce identity verification systems that would prompt a wide scope of improvement in the passport

and immigration management [1].

The major point, which is accepted and reached, was to foster an approach for face verification that is quick, effortless, and exact with moderately basic and straightforward calculations, predictions, and procedures.

Therefore, the proposed system, first detects the faces in the photograph, extracts the faces and eliminates the background. In the second step, the system computes the vectors of the photograph/image provided, compares it to the vectors of the corresponding image and measures the image similarity for the identity verification process.

Section 2 in this paper comprises literature review and related works carried out. Section 3 describes the methodology adopted in this paper for implementation of the model and briefly introduces the algorithms and metrics used. Section 4 shows the results obtained. Section 5 discusses the conclusion of the research. Section 6 shows and describes the future scope of this research.

2. LITERATURE REVIEW

In the field of face identification and verification, various works have been done so far. Many research papers and journals have been taken into consideration to explore the work done until today and the techniques utilized in each work are displayed under this segment in Table 1.

This section reviews the research done previously which is somehow related to the problem discovered by the authors. The goal was to understand and make sense of the challenging problems and to pinpoint the abilities needed for research. According to a review of the literature, the main discrepancies between authors' descriptions of processes are differences in language, methodology, and hierarchy of these processes. Table 1 is created that identifies and depicts the key phases of research and the comparison of methodologies adopted. The literature was found to contain implicit knowledge about the skills needed to engage in research, explore the key findings of the work done and highlighting the limitation so as the methodology adopted in this paper by the authors is different and more efficient.

Table 1. Literature review

Serial no.	Title	Datasets used	Methodology	Limitation
1.	Parameterization of a stochastic model for human face identification [2]	ORL.	HMM parameterizations are analyzed.	HMMs have been utilized with a bit less success around face identification.
2.	Face Recognition: A Convolutional Neural Network Approach [3]	400 images of 40 individuals	Hybrid neural network solution is presented which compares favorably with other methods.	Generalization and speed up handwritten digit recognition can be improved by selecting appropriate Convolutional network architecture.
3.	Learning a similarity metric discriminatively, with application to face verification [4]	Purdue/AR face	A method for training a similarity metric from data is presented which maps the input patterns into a target space which is further applied in face verification	Similarity metrics used in this research can be furthermore implemented to build invariant kernel functions.
4.	Face Verification Across Age Progression [5]	Retrieved from the passports of 465 individuals	Pre-processing methods for minimizing variations that are seen gradually with age are presented and overcome.	There is a lack in datasets of age progressed face images which leads to less research in this area.
5.	Face verification and identification using Facial Trait Code [6]	3575 images from 840 people under different illumination conditions are collected.	Facial Trait Code (FTC) to encode human facial images is presented.	The FTC cannot be used to handle faces under larger out-of- plane rotations.
6.	Large Scale Online Learning of Image Similarity Through Ranking [7]	Large, web scale datasets	An Online Algorithm for Scalable Image Similarity learning that learns a bilinear similarity measure over sparse representations is implemented.	The model presented is class independent whereas class dependent models can prove to be more efficient and are likely to increase precision.
7.	Cascade shallow CNN structure for face verification and identification [8]	Labeled Face in the Wild dataset and the AT&T dataset of faces	A cascade classifier based on shallow Convolutional neural networks.	The classifier is more efficient for a greater number of samples in the dataset.
8.	Triplet probabilistic embedding for face verification and clustering [9]	IJB-A and LFW datasets	A deep CNN-based approach is coupled with an embedding step to address the face verification problem.	The clustering algorithm does not prove efficient for handling large scale scenarios.
9.	Template adaptation for face verification and identification [10]	IJB-A	Comparison of template adaptation to Convolutional networks with metric learning, 2D and 3D alignment.	Various face recognition tasks can be evaluated with new template-based face datasets.
10.	Detecting morphed passport photos: A training and individual differences approach [11]	Set of 49 identity pairs used by Robertson et al.	Model is presented which quantifies the threat posed by a morphed passport photos fraud.	This research can be further deployed to evaluate whether SRs could be an effective morph fake antidote.

3. RELATED WORK

3.1 Face detection

A generally well-known subject called Face detection has an enormous scope of uses. Advanced smartphones and laptops accompany in-fabricated face identification features, which can validate the identity of the customer. There are various applications that can catch, recognize, and measure a face feature progressively. The rundown is not restricted to these versatile applications, as this technology likewise has a wide scope of utilizations in passport management and immigration system to meet the security purposes. In any case, the beginning of its Success stories traces all the way back to 2001, when Viola and Jones proposed the first since forever Face Detection Framework for detecting faces.

Haar cascade algorithm: Viola and Jones proposed the algorithm that utilizes edge or line location highlights in their exploration paper “Rapid Object Detection using a Boosted Cascade of Simple Features” issued in the year 2001 [12]. The algorithm is trained using a ton of positive and negative pictures consisting of faces and no faces, respectively [13].

3.2 Region of interest

This technique of retrieving the needful from an image was proposed by T. Chen, L.-H. Chen, K.-K. Ma in 1999. The researchers implemented this strategy on about 500 images.

This technique provides high accuracy in aftereffects of picture recovery. Pivoted and scaled forms of pictures can additionally be recognized by utilizing the proposed strategy.

3.3 Neural networks

CNN (Convolutional Neural Network) embedded with multiple layers have as of late been displayed to accomplish incredible outcomes on some undeniable level errands, for example, image similarity, object detection and more as of late additionally semantic similarity and segmentation. Thus, Convolutional networks are prepared to give great limited pixel-wise highlights for the subsequent advance being generally a more worldwide graphical model [14].

Alexander G. Schwing and Raquel Urtasun exhibit their technique on the semantic picture division errand and show empowering results on the difficult PASCAL VOC 2012 dataset. Anastasios Doulamis, Nikolaos Doulamis, Klimis Ntalianis, and Stefanos Kollias proposed a solo video object (VO) segmentation and following calculation dependent on a versatile neural-network design. Bharath Hariharan, Pablo Arbel'aez, Ross Girshick, and Jitendra Malik distinguish all occurrences of a class in a picture and, for each case, mark the pixels that have a place with it. They call this undertaking Simultaneous Detection and Segmentation (SDS). In contrast to old style jumping box identification, SDS requires division and not simply a case. In contrast to old style semantic division, we require singular article cases [15].

A group of models, called EfficientNets has several baselines B1-B7 while B0 being the main baseline of the network developed by AutoML MNAS from which the other baselines have been developed. These models are based on CNN architecture, which accomplish much better exactness and proficiency than previous ConvNets such as MobileNet and ResNets [16].

3.4 Image similarity

The architecture applied for retrieval of image comprises of the following three steps:

- Transforming features
- Representing features
- Similarity function

There is a vast history of evaluating similarity functions in various fields which includes face or objects detection, face verification, and similarity of faces. Apart from this, many similarity functions have also been introduced for the evaluation of image features retrieved from an image [17].

4. METHODOLOGY

4.1 Dataset

The dataset utilized in this research paper is a collection of images of different individuals and is created manually by the authors. The images are collected from family and friends. The images taken are preprocessed and augmented so as to make it relevant for the work. The dataset consisted of three folders: train, test and validation. Each of these folders consisted of 10 sub-folders corresponding to each individual. Initially, the dataset consisted of about 50 raw images, 5 images each for 10 different individuals. As this research-based project focuses primarily on passport identification and immigration system, the images are also collected on the same criteria i.e., according to the 5-year gap from the year 2017 to 2021. The year gap remains the same for all individuals. Initially, the mean values and shape of the images were estimated as shown in Figure 1 and Figure 2 below. The significance behind taking the mean values is to elaborate that which image in the dataset corresponds to the other image displayed. It extracts the useful information from the image.

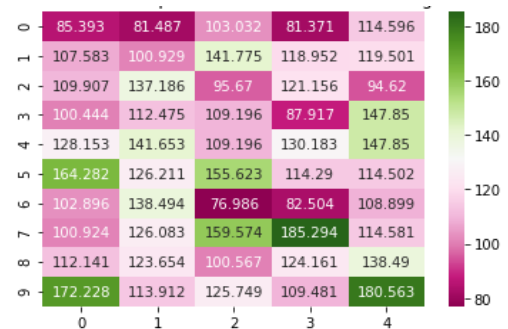


Figure 1. Heatmap representation of mean values of images

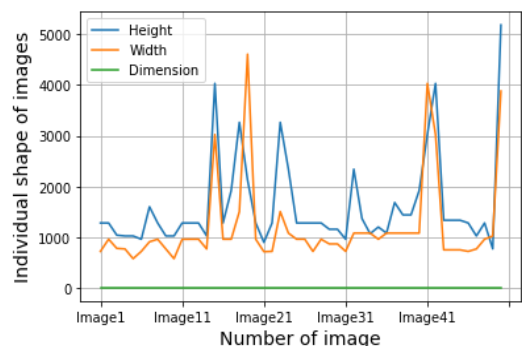


Figure 2. Line graph representation of shape of images

4.2 Data preprocessing

Reading images: The images are read using OpenCV by specifying and storing the path to our image dataset into a variable.

Resizing images: All images in the dataset are captured by a camera and are to be fed to our face detection model but they vary in size, therefore, we have established a base size for all images.

Converting images into grayscale: The images are converted into a grayscale format because the face detection algorithm works more efficiently, the reason being that differentiating such images from any other sort of color image requires less information to be provided for each pixel.

Cropping images: The images are cropped by passing parameters for x-axis, y-axis, width, and depth. As the images are captured in different angles for different individuals, the parameters are passed according to our requirements.

4.3 Model architecture

The architecture of our model is described below and depicted in Figure 3:

- The images are taken as an input and fed into the face detection model which detects the faces in the images and gives the count of the number of faces present in the image.
- After the process of face detection is complete, an efficient technique ROI i.e., Region of Interest is implemented which is used to extract the faces that are detected by the face detection model so that the target area gets separated, and the results obtained are more accurate.
- The faces that are extracted are thereafter combined into another modified dataset.
- The modified dataset or the original dataset is not sufficient to obtain good results. Therefore, the size of the dataset is increased using the process of data augmentation which makes it enough for the process of training. Initially, there were 50 images in the dataset which were later increased to about 1400 images.
- The augmented dataset is furthermore trained using three models based on neural network architecture which are: EfficientNetB0, MobileNet, DenseNet121.
- The training results are compared for all three models and the efficient model out of three is chosen. The outputs of the dataset trained on the most efficient model are then fed into the similarity model which gives the similarity scores between the images.
- The similarity between two images is computed using two different metrics - Cosine similarity and Pearson similarity.

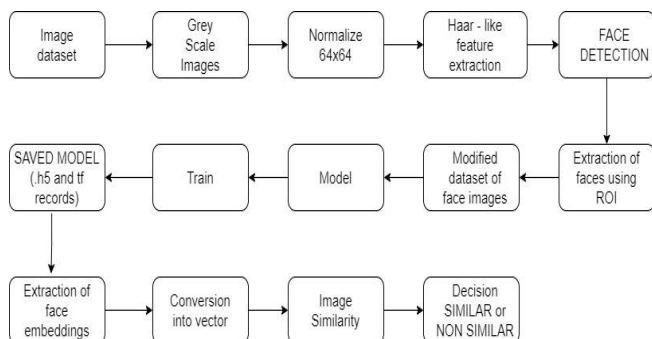


Figure 3. Model architecture

4.4 Face detection

Face detection is a biometric programming application adjusted to distinguish people through detecting faces and to perceive the identity of individuals. This methodology can be executed basically in fields like passport management and immigration system for security. The technique used in this research for the face detection process is done with the help of Viola Jones algorithm. Viola Jones algorithm works with a solid classifier [18]. The Viola-Jones algorithm is made from three basic ideas that make it conceivable to foster an ongoing face identifier: Haar-like features, Image identifier, Adaboost preparing and Cascading classifier. By applying these features, the framework can decide the presence of a human face [19]. Feature can be described as a single value procured by deducting the amount of the pixels underneath the white square shape from the amount of the pixels underneath the dark square shape [20]. For the computation of each feature, it is important that the amount of the pixels which are under the white and dark square shapes should be discovered [21].

In this research, the pre trained Haar cascade classifier is downloaded and moved into the current working directory. After that, an instance of the classifier has been created and the faces in the images are detected. Thereafter, a bounding box is printed around each detected face using OpenCV as shown in Figure 4 [22].

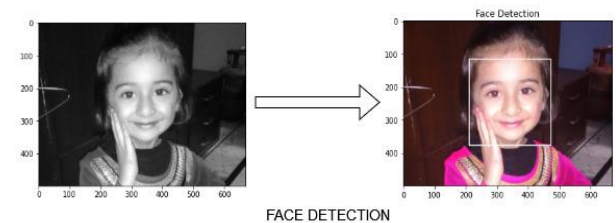


Figure 4. Face detection using Viola-Jones

4.5 Region of interest

Region of interest (ROI) can at times be pretty much as little as one pixel. In the configuration of ROI, the read-out window is diminished to the space pertinent for examination. ROI is an efficient technique to extract the detected faces from an image. After the face detection process in this model, the detected faces have been extracted using this technique. The mouse drag technique is utilized to select the target area and get ROI. As a result of this program, coordinates of the target region and the target image are obtained that is primarily a face in this project. Figure5 shows one batch of extracted images.

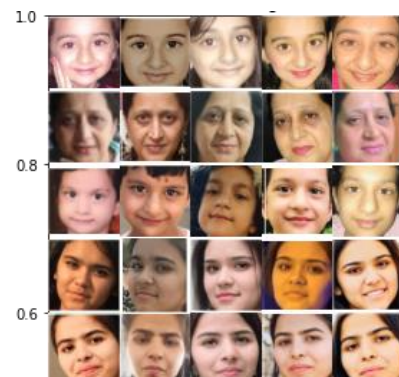


Figure 5. Extracted images using ROI

4.6 Data augmentation

Image data augmentation is a technique that can be used to extend the size of a dataset by making amendments in the dataset falsely. Training extensive learning of neural organization models on more data can bring about more capable models, and the expansion strategies can make assortments of the images that can work on the limit of the fit models to summarize what they have sorted out in some way to new images [23]. In this model, the Keras deep learning library is utilized which gives the ability to fit models utilizing picture information expansion by means of the ImageDataGenerator class. Many different techniques and methodologies are provided along with the pixel scaling methods.

4.7 Model training

Neural networks require an underlying arrangement of information, called a preparation or training dataset, to go about as a gauge for additional application and usage. The training dataset should be precisely named or labeled before the model can measure and gain from it. To prepare a model, a training dataset is highly required. The processes of training a dataset on a model begins with creating a train and test split of the dataset. The process involved in model training are shown in Figure 6. This method not only quickly evaluates the performance of an algorithm but also helps in making the model more robust according to our problem.

The dataset used in this project has been divided into various sub-folders: train, test, validation using a piece of code which takes the path of an input and output dataset and splits the input dataset into the specified ratio such as .6 for testing, .2 for training and .2 for validation using splitfolders. Thereafter, it is trained on three models which are based on neural network architecture:

1. MobileNet
2. DenseNet121
3. EfficientNetB0

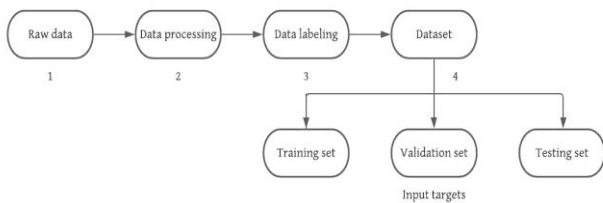


Figure 6. Processes involved in model training

The dataset is trained by passing various parameters required by the models and with 200 epochs. The number of epochs is a hyper parameter which indicates and controls the number of times an algorithm visits the dataset. The history of epochs of the model - EfficientNetB0 is shown below in Table 2 and Figure 7 respectively. In Table 2, the epoch number, loss, and accuracy corresponding to each epoch is mentioned for 6 epochs commencing from 191 to 196. The computation complexity of CNN can be determined by taking the sum of complexity of Convolutional layers. Model summary shows all the necessary information about the number of parameters and layers in a model. EfficientNetB0 proves to be the most efficient model for the dataset used as it gives the highest test

accuracy of about 82% and 100% accuracy for training. Therefore, EfficientNetB0 has been implemented in this project to compute the similarity between images. The .h5 file is converted into .pb file to be used as the input for the similarity model.

Table 2. Few epochs for training the dataset

Epoch	Loss	Accuracy
191/200	0.04	0.996
192/200	5.22	1.00
193/200	0.01	0.996
194/200	3.09	1.00
195/200	0.0014	1.00
196/200	1.56	1.00

```

block7a_se_reshape (Reshape) (None, 1, 1, 1152) 0 block7a_se_squeeze[0][0]
block7a_se_reduce (Conv2D) (None, 1, 1, 48) 55344 block7a_se_reshape[0][0]
block7a_se_expand (Conv2D) (None, 1, 1, 1152) 56448 block7a_se_reduce[0][0]
block7a_se_excite (Multiply) (None, None, None, 1) 0 block7a_activation[0][0]
block7a_se_expand[0][0]
block7a_project_conv (Conv2D) (None, None, None, 3) 368640 block7a_se_excite[0][0]
block7a_project_bn (BatchNormal (None, None, None, 3) 1280 block7a_project_conv[0][0]
top_conv (Conv2D) (None, None, None, 1) 409600 block7a_project_bn[0][0]
top_bn (BatchNormalization) (None, None, None, 1) 5120 top_conv[0][0]
top_activation (Activation) (None, None, None, 1) 0 top_bn[0][0]
global_average_pooling2d (Globa (None, 1280) 0 top_activation[0][0]
dense (Dense) (None, 1024) 1311744 global_average_pooling2d[0][0]
dense_1 (Dense) (None, 10) 10250 dense[0][0]
=====
Total params: 5,371,565
Trainable params: 5,329,542
Non-trainable params: 42,023
  
```

Figure 7. Model summary of EfficientNetB0

4.8 Semantic similarity

The measure of similarity between two images concludes image similarity. As such, it measures the level of comparability between embedded features of two pictures [24]. For building an image similarity model, the meaningful features need to be extracted into a vectorized form. Mapping images to vectors is required which can further be obtained using Convolutional neural network (CNN), and the vectors obtained can be used as embeddings [25].

The model proposed in this paper aims at predicting similarity between images using the criteria described above and further evaluates it to be similar or non-similar. This is implemented using a state-of-art model - EfficientNetB0 [15]. This model is based on CNN architecture which comprises of three basic layers along with two important parameters such as dropout layer and activation function as described below:

1. Convolution layer: This is the primary layer which is utilized to extricate the different features from the dataset. The scalar value obtained from dot product is taken between the channel and the pieces of the inputted image as for the size of the channel (MxM) by passing the channel over the inputted image.

2. Pooling layer: The essential point of pooling layer is to diminish the size of the convolved feature map to lessen the computational expenses. This is done by diminishing the associations among layers and autonomously works on each element map.

3. Fully Connected layer: This layer comprises the loads and predispositions alongside the neurons. The image taken as an

input is straightened in this. The straightened vector then, at that point, goes through scarcely any more FC layers where the numerical capacities activities ordinarily occur. The classification process begins at this stage.

4. Dropout layer: When every one of the features is associated with the FC layer, overfitting in the training dataset can be caused. Overfitting happens when a specific model functions admirably on the training dataset causing an adverse consequence in the performance of the model when utilized on another dataset.

To defeat this issue, a dropout layer is used wherein a couple of neurons are dropped from the neural network during the process of training bringing about diminished size of the model.

5. Activation function: Activation function is the main parameter of the CNN model. They are utilized to learn any sort of persistent and complex connection between factors of the network. Activation functions like the ReLU, Softmax, tanH and the sigmoid function have different utilization. Sigmoid and soft max functions are favored for a binary-classification problem whereas in the case of multi-class classification, softmax is utilized [26].

EfficientNet-B0 is the base network based on the inverted bottleneck residual blocks of MobileNetV2 which is trained on ImageNet dataset.

Step (a): The total number of layers in EfficientNet-B0 is 237 and all these layers can be made from 5 modules [26]. The modules of the layers are depicted in the Figure 8.

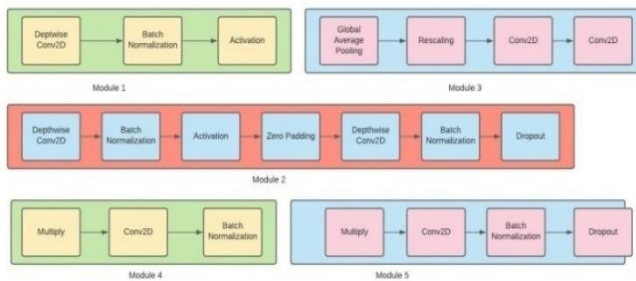


Figure 8. Modules of layers [24]

Step (b): The modules shown above are further combined and the complete architecture of EfficientNetB0 is depicted in the Figure 9 [27].

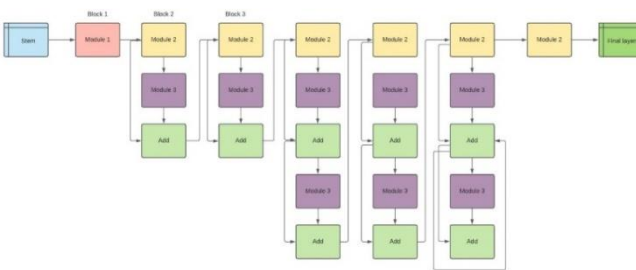


Figure 9. Architecture of EfficientNetB0 [24]

4.9 Similarity metrics

Cosine similarity: This metric is used to measure the similarity between the images regardless of their size. The cosine of the angle between two vectors obtained from images projected in a multi- dimensional space is measured [28]. In

this context, the two vectors which are being referenced are obtained using the feature extractor class which outputs the features of two images in vectorized form [29].

Pearson similarity: It is defined as the covariance between two vectors divided by their standard deviations. Its value can range from -1 to +1 [30].

5. RESULTS

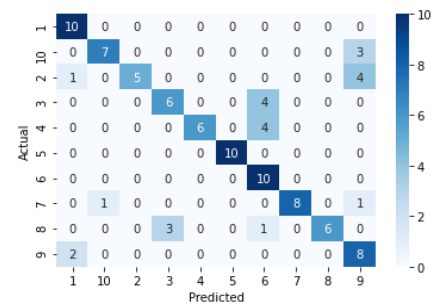
5.1 Model performance on the dataset

The results of accuracy obtained by training the dataset on three different state-of-art models are shown below in Figure 10, Figure 11, and Figure 12. From Figure 13, it can be proved that EfficientNetB0 proves to be the appropriate and best model for the dataset utilized in this research as it has achieved the test accuracy of 82%. The model is further implemented in the process of feature extraction and for measuring the similarity.

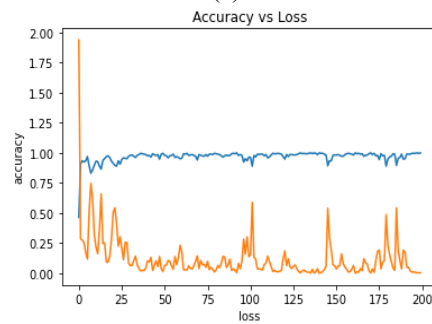
MobileNet: The test accuracy of MobileNet has come to be 76 %. The plots are shown below in Figure 10.

EfficientNetB0: The test accuracy of EfficientNetB0 has come to be 82 %. The plots are shown below in Figure 11.

DenseNet121: The test accuracy of DenseNet121 has come to be 75 %. The plots are shown below in Figure 12.

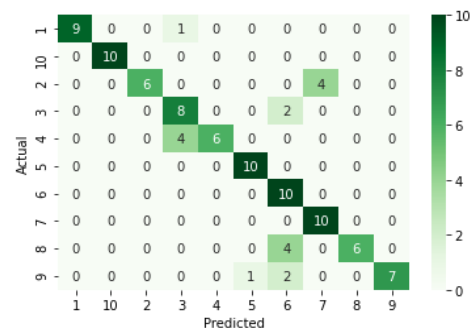


(a)



(b)

Figure 10. Accuracy of model MobileNet on the dataset



(a)

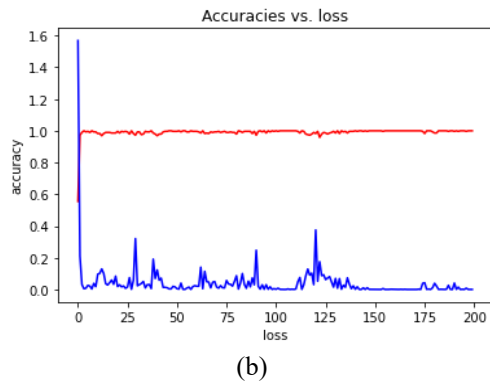


Figure 11. Accuracy of model EfficientNetB0 on the dataset

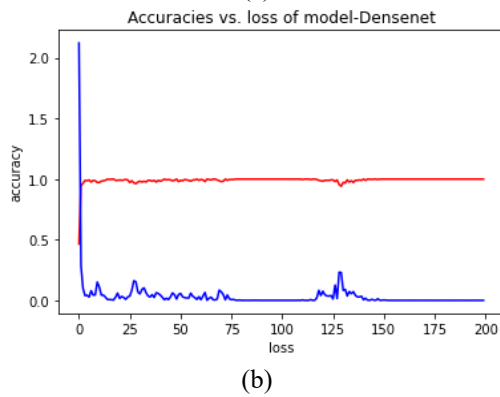
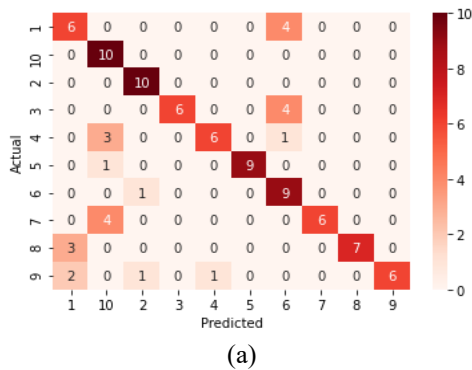


Figure 12. Accuracy of model DenseNet121 on the dataset

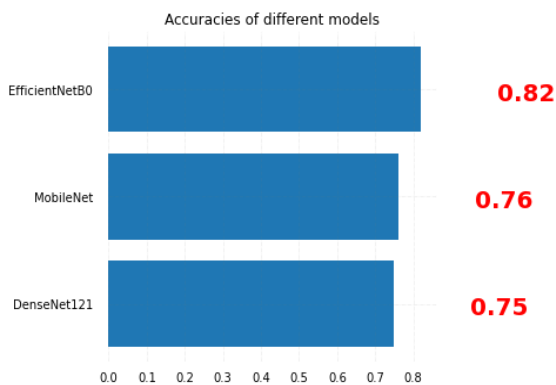


Figure 13. Accuracy comparison of different models

5.2 Similarity trends

Figure 14 shows the similarity trends for ten different individuals according to the differences in years ranging from 2017-2021. The similarities have been computed for the years

2017-2018 (1-year gap), 2017-2019 (2-year gap), 2017-2020 (3-year gap), 2017-2021 (4-year gap). This can be concluded that every individual has a unique similarity trend and there is no specific pattern for increasing or decreasing range of similarity corresponding to the gaps taken between years.

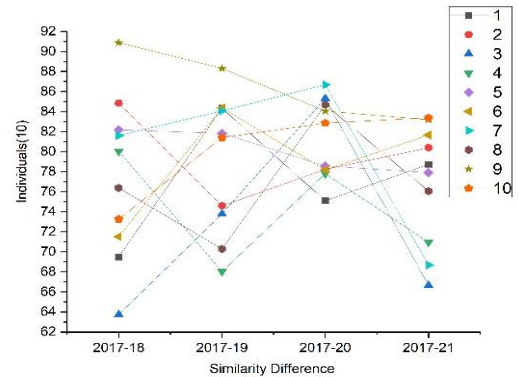
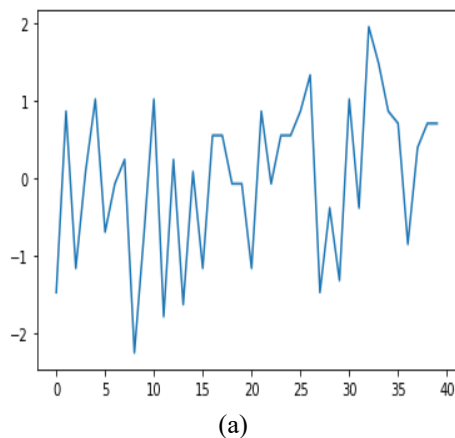


Figure 14. Row-wise graph of similarity trends

5.3 Calculation of threshold using Z-score

For the determination of threshold, Z-score or standard score is estimated using all the values obtained as similarities. The threshold for this model has come to be 6.41 which states the criteria for evaluating the image as similar or non-similar.

In the Figure 15 given below, the graph for z-scores obtained is shown and the outliers are detected in the dataset used which proves that there is no outlier i.e., all images in the dataset are similar corresponding to their test images.



```
print(sd)
6.41463171195354

threshold = 6.41# create empty list to store outliers
outliers = []# detect outlier
for i in data:
    z = (i-mean)/sd # calculate z-score
    if abs(z) > threshold: # identify outliers
        outliers.append(i) # add to the empty list# print outliers
print("The detected outliers are: ", outliers)
```

The detected outliers are: []

(b)

Figure 15. Results of Z-score

5.4 Semantic similarity

Few outputs for image similarity are shown below in Figure 16:

Similarity: 69.49 %	Similarity: 78.73 %
Actual: Similar	Actual: Similar
Prediction: Similar	Prediction: Similar
2017 2018	2017 2021



Figure 16. Outputs of image similarity

6. CONCLUSION

In this research, the degree to which identity of a person can be verified through images provided by matching the images from the previous records saved in the database of passport administration for the renewal process has been evaluated.

Moreover, it has also been evaluated whether the model verifies the identity of the individual present in the current image with the old image. The study also revealed that the similarity trends are different for different people i.e., there is no regular criteria for grouping the range of similarity scores.

Findings suggest that images with similarity above the threshold i.e., 64% are estimated as similar and the images with similarity below the threshold are estimated as non-similar. Further research is required to assess the problem of verifying the identical twins which could be deployed as an effective approach for verifying identity of both related and unrelated humans.

7. FUTURE SCOPE

The model implemented in this research gives satisfying results for unrelated humans but fails at detecting and computing similarity of related or identical twins when implemented on sample images of twins. The outputs that the model gave are shown below in Figure 17.

Biology has found Iris to be the most unique feature of the human face [31]. Therefore, it can prove to be an excellent feature to be used for distinguishing identical twins in the process of biometric identification and digital passport management. Hence, this drawback can be addressed using detection of iris by integrating feature extraction and Principal Component Analysis (PCA) or by using the watermarking-based approach [32-35].

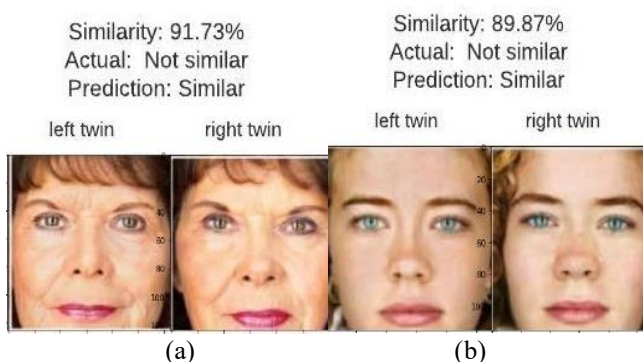


Figure 17. Future work

ACKNOWLEDGMENT

The authors thank Mrs. Monika Khurana, Mr. Dinesh Khurana, Ms. Ruhani Khurana, Ms. Suhani Khurana, Ms. Diksha Singra, Ms. Vandana, Ms. Sonam Ohri and Mr. Astitva Narang for contributing their images for this research.

REFERENCES

- [1] Hjelmås, E., Low, B.K. (2001). Face detection: A survey. *Computer Vision and Image Understanding*, 83(3): 236-274. <http://dx.doi.org/10.1006/cviu.2001.0921>
- [2] Samaria, F.S., Harter, A.C. (1994). Parameterisation of a stochastic model for human face identification. In *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision*, pp. 138-142. <https://doi.org/10.1109/ACV.1994.341300>
- [3] Lawrence, S., Giles, C.L., Tsoi, A.C., Back, A.D. (1997). Face recognition: A convolutional neural-network approach. *IEEE Transactions on Neural Networks*, 8(1): 98-113. <http://dx.doi.org/10.1109/72.554195>
- [4] Chopra, S., Hadsell, R., LeCun, Y. (2005). Learning a similarity metric discriminatively, with application to face verification. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 539-546. <http://dx.doi.org/10.1109/CVPR.2005.202>
- [5] Ramanathan, N., Chellappa, R. (2006). Face verification across age progression. *IEEE Transactions on Image Processing*, 15(11): 3349-3361. <http://dx.doi.org/10.1109/CVPR.2005.153>
- [6] Lee, P.H., Hsu, G.S., Hung, Y.P. (2009). Face verification and identification using facial trait code. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1613-1620. <https://doi.org/10.1109/CVPR.2009.5206830>
- [7] Chechik, G., Sharma, V., Shalit, U., Bengio, S. (2010). Large scale online learning of image similarity through ranking. *Journal of Machine Learning Research*, 11(3). http://dx.doi.org/10.1007/978-3-642-02172-5_2
- [8] Leng, B., Liu, Y., Yu, K., Xu, S., Yuan, Z., Qin, J. (2016). Cascade shallow CNN structure for face verification and identification. *Neurocomputing*, 215: 232-240. <http://dx.doi.org/10.1016/j.neucom.2015.08.134>
- [9] Sankaranarayanan, S., Alavi, A., Castillo, C.D., Chellappa, R. (2016). Triplet probabilistic embedding for face verification and clustering. In *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pp. 1-8. <http://dx.doi.org/10.1109/BTAS.2016.7791205>
- [10] Crosswhite, N., Byrne, J., Stauffer, C., Parkhi, O., Cao, Q., Zisserman, A. (2018). Template adaptation for face verification and identification. *Image and Vision Computing*, 79: 35-48. <http://dx.doi.org/10.1016/j.imavis.2018.09.002>
- [11] Robertson, D.J., Mungall, A., Watson, D.G., Wade, K. A., Nightingale, S.J., Butler, S. (2018). Detecting morphed passport photos: A training and individual differences approach. *Cognitive Research: Principles and Implications*, 3(1): 1-11. <http://dx.doi.org/10.1186/s41235-018-0113-8>
- [12] Savaş, B.K., İlkin, S., Becerikli, Y. (2016). The realization of face detection and fullness detection in

- medium by using Haar Cascade Classifiers. In 2016 24th Signal Processing and Communication Application Conference (SIU), pp. 2217-2220. <https://doi.org/10.1109/SIU.2016.7496215>
- [13] Hossen, A.M.A., Ogla, R.A.A., Ali, M.M. (2017). Face detection by using OpenCV's Viola-Jones algorithm based on coding eyes. *Iraqi Journal of Science*, 735-745. <http://dx.doi.org/10.1093/oxfordhb/9780199349791.013.16>
- [14] Li, Y., Hao, Z., Lei, H. (2016). Survey of convolutional neural network. *Journal of Computer Applications*, 36(9): 2508-2515.
- [15] Albawi, S., Mohammed, T. A., Al-Zawi, S. (2017). Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET), pp. 1-6. <http://dx.doi.org/10.1109/ICEngTechnol.2017.8308186>
- [16] Abedalla, A., Abdullah, M., Al-Ayyoub, M., Benkhelifa, E. (2021). Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures. *PeerJ Computer Science*, 7: e607. <http://dx.doi.org/10.7717/peerj-cs.607>
- [17] Aljanabi, M.A., Hussain, Z.M., Lu, S.F. (2018). An entropy-histogram approach for image similarity and face recognition. *Mathematical Problems in Engineering*, 2018. <http://dx.doi.org/10.1155/2018/9801308>
- [18] Dabhi, M.K., Pancholi, B.K. (2016). Face detection system based on Viola-Jones algorithm. *International Journal of Science and Research (IJSR)*, 5(4): 62-64. <http://dx.doi.org/10.21275/23197064>
- [19] Khan, M., Chakraborty, S., Astya, R., Khepra, S. (2019). Face detection and recognition using OpenCV. In 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), pp. 116-119. <http://dx.doi.org/10.1109/ICCCIS48478.2019.8974493>
- [20] Gupta, V., Mittal, M., Mittal, V., Saxena, N.K. (2021). A critical review of feature extraction techniques for ECG signal analysis. *Journal of The Institution of Engineers (India): Series B*, 102(5): 1049-1060. <http://dx.doi.org/10.1007/s40031-021-00606-5>
- [21] Singh, A., Herunde, H., Furtado, F. (2020). Modified Haar-cascade model for face detection issues. *International Journal of Research in Industrial Engineering*, 9(2): 143-171. <https://doi.org/10.22105/riej.2020.226857.1129>
- [22] Van Dyk, D.A., Meng, X.L. (2001). The art of data augmentation. *Journal of Computational and Graphical Statistics*, 10(1): 1-50. <http://dx.doi.org/10.1198/10618600152418584>
- [23] Guo, T., Dong, J., Li, H., Gao, Y. (2017). Simple convolutional neural network on image classification. In 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), pp. 721-724. <http://dx.doi.org/10.1109/ICBDA.2017.8078730>
- [24] Elavarasi, S.A., Akilandeswari, J., Menaga, K. (2014). A survey on semantic similarity measure. *International Journal of Research in Advent Technology*, 2(3): 389-398. <http://dx.doi.org/10.32622/ijrat>
- [25] Yildirim, M., Cinar, A. (2020). Classification of Alzheimer's disease MRI images with CNN based hybrid method. *Ingénierie des Systèmes d'Information*, 25(4): 413-418. <https://doi.org/10.18280/isi.250402>
- [26] Tan, M., Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning, pp. 6105-6114. <http://dx.doi.org/10.48550/arXiv.2209.11604>
- [27] Bloch, L., Boketta, A., Keibel, C., Mense, E., Michailutschenko, A., Pelka, O., Friedrich, C.M. (2020). Combination of image and location information for snake species identification using object detection and efficientnets. In CLEF (Working Notes).
- [28] Nguyen, H.V., Bai, L. (2010). Cosine similarity metric learning for face verification. In Asian Conference on Computer Vision, pp. 709-720. http://dx.doi.org/10.1007/978-3-319-16817-3_4
- [29] Wilford, R.E., Raja, V., Hershey, M., Anderson, M. (2022). The treachery of images: Objects, pictures, words and the role of affordances in similarity judgements. In Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44). <https://doi.org/10.31234/osf.io/s4p9c>
- [30] Starovoytov, V.V., Eldarova, E.E., Iskakov, K.T. (2020). Comparative analysis of the SSIM index and the pearson coefficient as a criterion for image similarity. *Eurasian Journal of Mathematical and Computer Applications*, 8(1): 76-90. <http://dx.doi.org/10.32523/2306-6172-2020-8-1-76-90>
- [31] Neha, K., Gupta, R., Mahajan, S. (2010). Iris recognition system. *International Journal of Advanced Computer Science and Applications*, 1(1): 34-40. <http://dx.doi.org/10.14569/IJACSA.2010.010106>
- [32] Rana, H.K., Azam, M.S., Akhtar, M.R. (2017). Iris recognition system using PCA based on DWT. *SM Journal of Biometrics & Biostatistics*, 2(3): 1015.
- [33] Gupta, V., Mittal, M., Mittal, V., Chaturvedi, Y. (2022). Detection of R-peaks using fractional Fourier transform and principal component analysis. *Journal of Ambient Intelligence and Humanized Computing*, 13(2): 961-972. <http://dx.doi.org/10.1016/j.jksuci.2022.05.009>
- [34] Pandey, M., Bhatia, M., Bansal, A. (2016). IRIS based human identification: Analogizing and exploiting PSNR and MSE techniques using MATLAB. In 2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH), pp. 231-235. <http://dx.doi.org/10.1109/ICICCS.2016.7542309>
- [35] HimaBindu, G., Anuradha, C., Chandra Murty, P.S.R. (2019). Feature extraction techniques in associate with opposition based whale optimization algorithm. *Ingénierie des Systèmes d'Information*, 24(4): 403-410. <https://doi.org/10.18280/isi.240407>