International Journal of Safety and Security Engineering
Vol. 14, No. 1, February, 2024, pp. 185-190
Journal homepage: http://iieta.org/journals/ijsse

Biometric Identification for a Secured Environment Using AI-Based Facial Recognition

Preetha Shivanna*, Sheela Samudrala Venkatesiah

Vivesvaraya Technological University, Department of Information Science & Engineering, B.M.S. College of Engineering, Basavanagudi, Bengaluru 560019, India

Corresponding Author Email: preetha.ise@bmsce.ac.in

https://doi.org/10.18280/ijsse.140118

Received: 11 July 2023
Revised: 16 January 2024
Accepted: 22 January 2024
Available online: 29 February 2024

Keywords:
artificial intelligence, facial, signature, biometric authentication, Siamese neural network, machine learning

ABSTRACT

Smart, intelligent and automated environments solves real-world issues by modeling artificial intelligence based applications. AI such as functional, analytical, textual and visual AI are applied to enhance capabilities and intelligence of an application. Information security is significant in the field of information technology. AI–powered biometric solutions have initiated to flourish in securing individual's identification to combat unauthorized access. Security deals with authentication; biometric user authentication systems should adopt reliable schemes to determine or confirm individual's identity who seek services. Such schemes should render service to only legitimate users to access secured environments. Face Detection and Recognition is emerging as ideal solution to facilitate secure verification and authentication in secured systems. In this paper, Siamese Neural Network method which employ unique structure is adopted to naturally rank input similarities of Biometric traits. The network is tuned to benefit from dominant discriminative features for generalizing predictions in authority of all new classes from unfamiliar distributions. Proposed approach for face recognition is compared with few current state-of-the-art approaches. The approach achieved an accuracy of 96.5%. Novelty of the research lies in combining Siamese network and Biometric authentication.

1. INTRODUCTION

Present biometric applications reveal several relations to artificial intelligence (AI). AI and IoT innovations have increased the inter-connectivity among devices and people. Digital platforms should provide a safe environment for people without compromising their data, this is where AI and biometric come in. Humans reveal a robust proficiency to recognize and acquire innovative patterns. In precise, presentation with stimuli is observed when individuals are able to comprehend new concepts swiftly and then identify discrepancies on such concepts in impending percept [1]. Effective usage of Machine learning (ML) benefits in attaining state-of-the-art performance in many diverse applications such as caption generation, speech & image recognition spam detection and web search. Though, such algorithms fail to predict data when forced and provided with less supervised information. Computer Vision uses One-shot learning, a machine learning based object classification algorithm to evaluate difference along with similarity among two images. One-shot learning aims to explicate the model to establish its own conventions on similarities based on marginal number of visuals.

Swift progress in the field of deep learning, artificial intelligence, machine learning and neural networks are convenient in analysing the evolution of AI and biometric technologies. AI-based biometric authentication provides more accurate, proficient identification and verification of individuals. AI benefits to identify fraudulent actions by analyzing human behavioral patterns. Convergence of biometric and AI has led novel innovations to enable fast, user-friendly and secure authentication protocols. IoT technologies are creating new and powerful models by combining software and hardware developments. Rising security risks due to introduction of Internet of Things (IoT) environments can be addressed by new type of AI-powered biometric identification facilities to secure and improve overall security of individuals and world businesses.

Simplification of unfamiliar classifications without providing wide retraining either expensive or impossible owed to inadequate data or in an online prediction setting is the key objective. One predominant motivating task is classification under constraint is to only perceive a single sample of each probable class prior to determine prediction of a test instance. This is the primary focused model in the studies of Lake et al. [1] and Li et al. [2] called One-shot Learning. Zero-shot 4 learning is distinguished as it is unable to peek at any samples from the target classes [3].

Organization of the paper is as follows: Section 2 presents approach adopted for experimentation. Section 3 discusses image verification using deep Siamese networks. Section 4 explains proposed architecture, AI Model, Secure classification and experimentation. Section 5 analyses results, concludes exploration of the proposed scheme and discusses

performance, challenges and scope of research.

2. APPROACH

The Siamese network scheme involves two identical subnetworks, to process one input each. Firstly, the inputs are processed through a convolutional neural network which extracts substantial features from the given images.

Representation of biometric images are learnt over a supervised metric–based technique along with Siamese neural networks, later reuse those network's attributes for one-shot learning deprived of any retraining. Siamese in the convolutional neural networks are employed experiment as they are

- Paired samples from source data can be trained easily by adopting standard optimization techniques
- Providing a modest method independent to domain- specific knowledge rather misusing deep learning procedures.
- Proficiency in learning general image features is useful in making predictions of unknown class distributions, though examples are minimal from new distributions.

Unlike conventional CNNs, a Siamese Network studies a similarity function which are trained on numerous amounts of data to predict multiple classes. This function permits us to distinguish between classes consuming marginal data, rendering them mainly effective for one-shot classification. Advancement of a one-shot image classification model aims to study neural network that discriminates class- identity of image pairs, typical verification task for recognizing an image.

3. RELATED WORK

3.1 Image verification using deep Siamese networks

In early 1990s Siamese networks were presented to resolve verification of signatures as an image matching problem [4]. A Siamese neural network comprised of identical networks that receives distinctive inputs merged with energy function at the top. Some metric is used by the function on every side to compute maximum-level feature representation as shown in Figure 1 [5]. A modest two hidden layer in Siamese network is used for binary classification along with logistic prediction p. Network structure is recreated through the topmost and bottom most sections to arrange twin networks, which shares weight matrices at every layer.

These Identical networks tie the parameters among them. Since same function is computed by each network, Weight binding assures that two images that are exceedingly similar may not surely be mapped by their corresponding networks to vastly dissimilar places in feature space. Also, network is symmetric and topmost conjoining layer shall calculate the same metric when two distinctive images are presented to identical networks assuming the similar two images are presented to opposite networks. (This explains how Siamese networks are implemented). Siamese networks are effective with fewer labeled examples. They excel in scenarios where labeled data is limited because they learn the similarity or dissimilarity between pairs of samples rather than relying on a large dataset for direct classification.

Siamese networks are excellent for metric learning tasks. They learn a similarity metric between pairs of input data, enabling them to determine how similar or dissimilar two faces are.

Siamese networks can learn from just one or a few examples of a new person, making them suitable for scenarios where new faces need to be added to a system without retraining on the entire dataset.

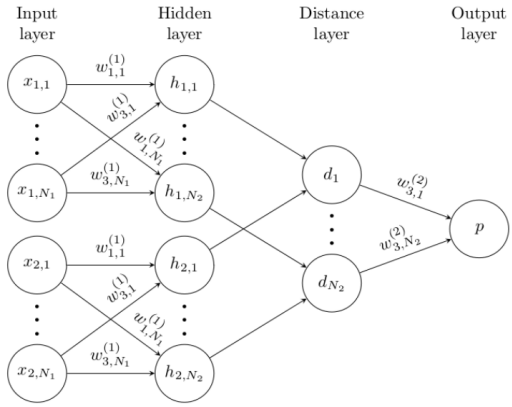

Figure 1. Siamese network with two hidden layers

A distinctive energy function containing dual terms was used to reduce energy loss in Bromley et al. [4] though metrics are fixed in Facebook's DeepFace approach proposed by Taigman et al. [6]. Convolutional neural networks own significant achievements and outcomes in various large- scale applications of computer vision, predominantly in image recognition tasks in the studies of Bengio [7], Krizhevsky et al. [8], Simonyan and Zisserman [9], and Srivastava 10]. Multiple convolution layers are chosen by best performing models. When compared to fully-connected layers and top-level energy function. However convolutional layers are computational expensive when compared to standard nonlinearities. Local connectivity in the model can predominantly reduce number of factors which basically offer some arrangement of built-in regularization.

Several such factors mark convolutional networks interesting. Also, these network's convolution operations use direct filtering interpretation, which maps each feature against input features to recognize groupings of pixels as patterns. Hence, the output of every convolution layer are mapped to essential spatial features in the original input space and suggest robustness for meek transforms. Currently a very swift CUDA libraries are accessible for designing large convolution networks without an intolerable amount of training time [8, 9, 11]. Datasets from Kaggle are considered for comparison study with the proposed method. Detailing the structure of Siamese networks and specifications of the learning algorithm adopted in the experiments are explained in detail in the following section [12-15].

4. PROPOSED ARCHITECTURE

Preprocessing Steps: Proposed model architecture for a novel AI-based biometric identification system is depicted in Figure 2. The images are collated from a University of Massachusetts database for labelled face images (https://vis-www.cs.umass.edu/lfw/). The images are passed through a

mean filter with a size 3 kernel.

The architecture combines AI, security parameters and authentication protocols like WEP or WPA in a biometric security scenario. Individual's unique facial images of people & faces are collated as an initial step and filtered. AI model should correctly identify the images and recognize the faces. Identified images are passed through WEP authentication protocol, which stores signature pairs mapped with face images. A 2-way WEP authentication is done to check signatures & correctly identified image match to determine if authentication is passed or failed. Proposed model adopts Siamese classifier based on CNN. Tensor flow and Jira's framework are employed for training and implementing the model. Model is trained considering the necessary parameters. Siamese model comprises of two CNN input layers followed by sequential layers, distance layers followed by fully connected layer.

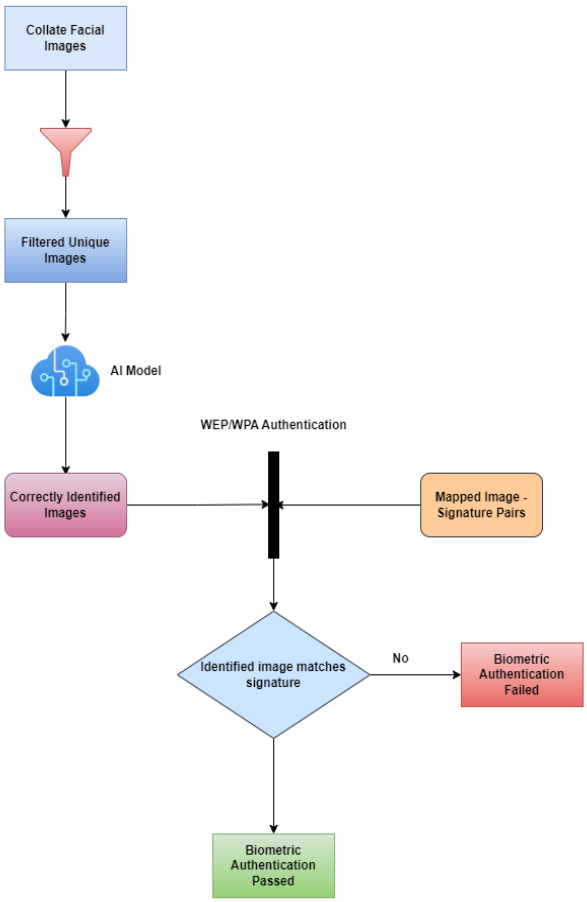

Figure 2. Model architecture and workflow diagram

A convolutional neural network is a kind of neural network that practices an arrangement of layers to extract features and process images. CNN has four layers.

- Convolutional layer: Feature maps are produced by applying filters to input images.
- Pooling layer: Important information is retained as this layer and the size of feature maps are reduced.
- ReLU layer: Applies a non-linear activation function to the feature maps and introduces non-linearity to the network.
- Fully connected layer: Performs classification by connecting all neurons from previous layer to output layer.
- Sequential layer: It is a default structure model with neural nets to stack up multiple layers from one's output goes into another's ahead.

4.1 Experimentation

Initially faces of famous personalities is collected. Face dataset contains 1690 unique faces. Dataset is also part of a Kaggle Face Recognition challenge. As a first step, all images are pre-processed by applying a Gaussian filter for smoothening the images. AI model however, accepts an input of 128×128, therefore the images are cropped to a dimension of 128×128×3. Images are provided to the network with RGB information intact, as converting the images to greyscale would lead to inferior detection capabilities. RGB images carry more information (color) which provides another dimension of potential features that could help correctly identify faces. For example, people of different color with similar facial features can be identified correctly based on the color. Loss of color information would lead to loss in recognition ability of the network. However, this would result in fewer trainable parameters, making the network run faster. However, speed gained is not worth the compromise that arises network performance. In this instance, classes of 1690 unique images that make up the dataset's classes are unevenly distributed; some have just two examples of a given personality, while others have eight. As a result, future AI model will have to be built and trained using non- IID (Independent or Identical) data, which will be difficult to do. AI model performs at its best to guarantee the problem of class imbalance has been taken into consideration. Siamese networks are used to address the non-IID aspect of the data. Class imbalance here was not addressed. The network performed despite the class imbalance.

4.2 AI model

Network used for face recognition is a Siamese Network. The summary of the network is shown in Figure 3.

```
Model: "model"
_________________________________________________________________________________________________
Layer (type)                 Output Shape              Param #   Connected to
=================================================================================================
Input1 (InputLayer)          [(None, 128, 128, 3)] 0
_________________________________________________________________________________________________
Input2 (InputLayer)          [(None, 128, 128, 3)] 0
_________________________________________________________________________________________________
sequential (Sequential)      (None, 2048)              25057832  Input1[0][0]
                                                                 Input2[0][0]
_________________________________________________________________________________________________
Distance (Lambda)            (None, 2048)              0         sequential[0][0]
                                                                 sequential[1][0]
_________________________________________________________________________________________________
Prediction (Dense)           (None, 1)                 2049      Distance[0][0]
=================================================================================================
Total params: 25,059,881
Trainable params: 8,947,201
Non-trainable params: 16,112,680
_________________________________________________________________________________________________
```

Figure 3. Siamese network model summary

Images are cropped to a size of 128×128×3 to meet specific conditions of the AI model prior being fed into it. Two of these input layers are present in the model, and a following sequential layer condenses the input into a one- dimensional vector with a 2048-by-2048-pixel size. Class's sole output is placed in last layer, which is a dense, completely linked layer. 25,059,881 parameters make up the whole model, although only 8,947,201 of those parameters are trainable. Architecture of the network in shown in Figure 4.

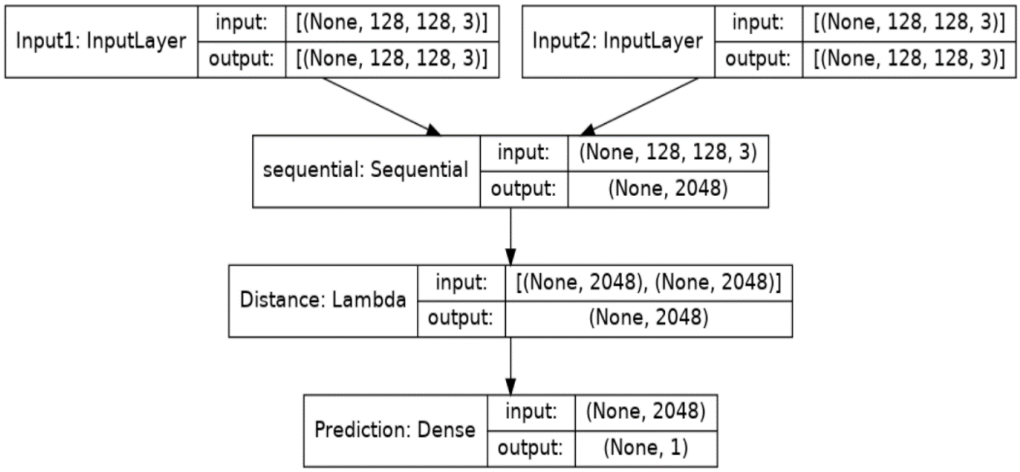

Figure 4. Siamese network model architecture

4.3 Siamese network

Siamese networks are used to perform identification, recognition and verification tasks for signature verification and face recognition. These networks consist of two subnetworks that mirror each other. Two inputs are needed for two subnetworks. Two pairs of images are needed to check if they belong to same class or different class of same person. Random samples are considered to train the network for similarity. Binary code entropy is the loss function used to train Siamese networks.

Contrasting a conventional CNN, Siamese Network calculates the distance between any two given images. Euclidean distance is calculated between images if faces instead of facial key points. Network learn the parameters if the images have same label i.e. smaller distances between two images are produced based on weights and preferences, larger distance if they fall under different labels.

Design of the model is purposefully kept minimal due to the time frame in which it will run. This will help to retain a high degree of accuracy while making quick predictions. However, a more sophisticated design can be adopted to boost performance if the system's speed is never a priority. Adam Optimizer with a learning rate of 0.01 was used with Binary Cross Entropy Loss.

4.4 Secure classification

Trained AI model can identify and recognize a user's face with accuracy. Recognized face is then sent to a simulated secure environment, which may be hosted on a cloud-based platform or secured by WEP/WPA authentication. A rigorous Python class is presently used to illustrate and replicate this environment. Python environment is set to include functions that accepts an input set of images and labels and then constructs image pairs. The function trains Siamese network with two separate arrays named "image" and "label". Montages of images help to visually validate multiple images at once. Two parameters, images of dataset and class labels associated with images are passed into method to compute total number of unique class label for the dataset. Unique function finds all unique class labels from dataset and corresponds to unique digits class labels. Finally builds list of indexes of each data points in a super compact and efficient manner. Array is allocated for side-by-side visualization, horizontally stack two images and then adding the image pair to output array. Image resolution is increased using RGB channel for better clarity of images. Network learns to differentiate between every pair of face images and determines if it is the same person in both photos and if the signature is a forgery or not. However, this may be adjusted to any secure environment seen in real world. Each category of recognized faces has its own hash code. When AI model correctly recognizes a face, secure environment receives the associated hash code for that class.

Secured signatures of people are kept in a secure environment. An authentication test is executed to ensure that hash code supplied with an identifiable face matches the hash code of related signature prior to secure environment transfer. A message informing the user that authentication was successful is sent if a match exists. In contrast, a notice informing the user that authentication has failed is sent if hash codes do not match. Authenticity of the authentication procedure fully depends on AI model's performance as it determines the hash code that is generated and matched with hash code of signature in the secure environment.

5. RESULTS AND DISCUSSION

5.1 Results

Siamese Network is trained utilizing a PC with an 8-core CPU, 32 GB RAM, and an 8 GB RTX 3070 GPU for completely 4 hours, 35 minutes. Training was limited to a shorter duration due to availability of resources. Training process is carried out with a batch size of 1024 across 5 epochs. To guarantee that future data addition may be included while retaining an appropriate training duration, decision to restrict the process of training to 5 epochs is considered. The face dataset contains 1680 unique faces. 90 unique faces were used for testing. A final accuracy of 96.32% on test set was achieved, appreciating the reserve of 5% of data during training. Variations in accuracy seen across the training epochs are shown in Figure 5.

Accuracy of the model shows a considerable rise between the first and second epochs. After fifth epoch, accuracy changes very minimally. This discovery leads to conclusion that, up to past fifth epoch, there was no justifiable trade-off between accuracy and training time. A decrease in the number

of training epochs to 5 was done in order to shorten training periods without sacrificing accuracy. Development of loss function across 5 epochs and 55 batches is presented in Figure 6.

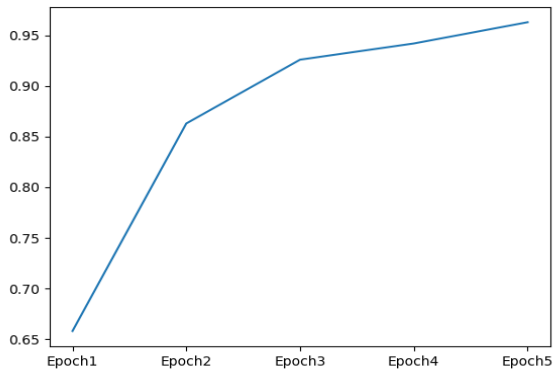

Figure 5. Test Accuracy increases as training progresses

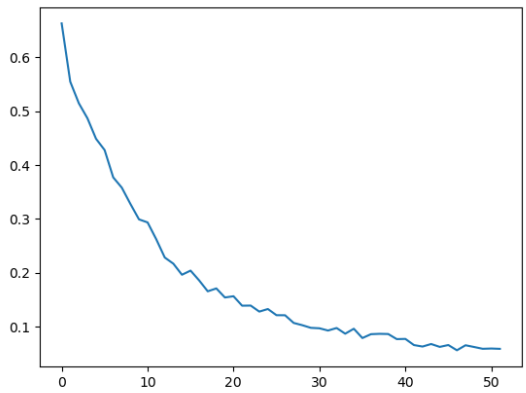

Figure 6. Loss curve for training. Y-axis is the loss metric and the X-axis is batch number

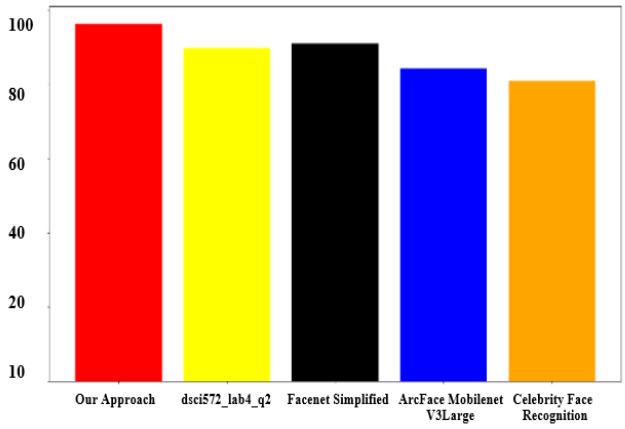

Figure 7. Accuracy comparison of our approach with other state-of-the-art approaches

Number of batches shows horizontal axis of loss curve, which is 55 batches spread across 5 epochs with an average of 11 batches each epoch. Progressive reduction in loss throughout various batches and epochs is observed in Figure 6.

Comparison of face recognition system to four state-of-the-art methodologies [1-4] to assess its efficacy is performed. Figure 7's bar graph shows accuracy performances of each method. One of the bars in this graph represents the proposed model accuracy, making it easy to compare it to other methods. With a score of 96.3%, our method had the best accuracy of all the approaches that were evaluated. On the same dataset, the second-best performing method [2], had an accuracy of 91%.

5.2 Discussion

In comparison to more conventional methods, suggested Siamese networks method for face recognition provides a number of advantages. Siamese Network design is very helpful in applications like facial recognition that call for the comparison of similarities between two inputs. Because of its capacity to learn intricate patterns and characteristics from images, the network is resistant to changes in lighting, position, and facial expression.

Network's capacity to recognize objects is also preserved by using RGB photos rather than monochrome ones. In face recognition applications, loss of color information in grayscale images might cause network's capacity to recognize objects to decline, which is undesirable. Facial images can be collected from several postures and angles with good lighting conditions to enable predictor model to facilitate successful predictions.

The problem of class imbalance was taken into consideration when the AI model was designed, preventing the model from favoring classes with greater data. This is crucial in situations where there are no equal amount of images for each class since it might result in subpar performance and inaccurate person identification. Model can be trained to consider diverse data and equal number of data images for enhancing the performance of the model.

Effectiveness of AI model, which may be constrained in some circumstances, was a need for the authentication procedure. Performance of the AI model may be impacted, for instance, when individual's face is not clearly apparent or when image quality is low. However, using a secure environment for authentication aids in guarding against unapproved system access.

Finally, the propose method for face recognition using Siamese networks demonstrated good results in terms of precision and speed. It is a strong and trustworthy method for face recognition applications since it uses RGB images, address class imbalance, and conducts authentication in a secure setting. Research scope is enabled to enhance the model's performance in various settings and investigate its scalability in practical applications.

6. CONCLUSION

Proposed Model is a promising method for facial recognition applications since it accurately identifies and verifies people. The network is trained to take advantage from main discriminative features for simplifying calculations in authority of unfamiliar distributions from all new classes.

Siamese networks is used in the model to provide an innovative approach for face identification. On a dataset comprising 1690 photos of 10 distinct people, proposed method accomplished exceptional accuracy of 96.3%. When compared to other cutting-edge methods, the model's performance showed a notable improvement in accuracy. On

a PC with 32 GB RAM, an 8-core CPU, and an 8 GB RTX 3070 GPU, network was efficiently trained.

Siamese Design of networks was chosen because it can compare and detect similar patterns across two different inputs. Network was built to accept images with all of the RGB information present to retain strong identification capabilities. Avoidance of AI model getting biased regarding classes pertaining more data, imbalance in the number of images per class was addressed during the design phase. With a batch size of 1024, model is trained for 5 epochs in 4 hours, 35 minutes.

Hash code of recognized image is compared with hash code of the signature saved in secure environment, after the identified face was delivered to the simulated secure environment. Efficiency of the AI model, which calculated the hash code communicated to secure environment, was entirely responsible for the procedure of authentication.

As for future directions, mechanism of network optimization can be further explored to combine Siamese network with several compatible techniques to enable real-life scenarios and applications. Edge intelligence is one among the great applications.

REFERENCES

[1] Lake, B., Salakhutdinov, R., Gross, J., Tenenbaum, J. (2011). One shot learning of simple visual concepts. In Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33): 2568-2573.

[2] Li, F.F., Fergus, R., Perona, P. (2006). One-shot learning of object categories. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(4): 594-611. https://doi.org/10.1109/TPAMI.2006.79

[3] Palatucci, M., Pomerleau, D., Hinton, G.E., Mitchell, T.M. (2009). Zero-shot learning with semantic output codes. Advances in Neural Information Processing Systems, 22.

[4] Bromley, J., Guyon, I., LeCun, Y., Säckinger, E., Shah, R. (1993). Signature verification using a "Siamese" time delay neural network. Advances in Neural Information Processing Systems, 6.

[5] Koch, G., Zemel, R., Salakhutdinov, R. (2015). Siamese neural networks for one-shot image recognition. ICML Deep Learning Workshop, 2(1).

[6] Taigman, Y., Yang, M., Ranzato, M.A., Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA. pp. 1701-1708. https://doi.org/10.1109/CVPR.2014.220

[7] Bengio, Y. (2009). Learning deep architectures for AI. Foundations and trends® in Machine Learning, 2(1): 1-127. http://doi.org/10.1561/2200000006

[8] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6): 84-90. https://doi.org/10.1145/3065386

[9] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

[10] Srivastava, N. (2013). Improving neural networks with dropout. University of Toronto, 182(566): 7. https://api.semanticscholar.org/CorpusID:17084851

[11] Mnih, V. (2009). Cudamat: A CUDA-based matrix class for python. Department of Computer Science, University of Toronto.

[12] https://www.kaggle.com/code/shlrley/dsci572-lab4-q2.

[13] https://www.kaggle.com/code/rushikeshhiray/facenet-simplified.

[14] https://www.kaggle.com/code/whysetiawan27/arcface-mobilenetv3large.

[15] https://www.kaggle.com/code/ravehgillmore/celebrity-face-recognition.