



廣西科技師範學院

Guangxi Science & Technology Normal University

# 2022 届本科毕业论文（设计）

## Class of 2022 Undergraduate Thesis (Design)

题 目 ( Thesis Title)	Design and Implementation of Face Recognition System Based on Convolutional Neural Network (CNN)
学 院 名 称 (College Name)	College of Mathematics & Computer Science
专业名称(Major)	Computer Science and Technology
学号(Student ID )	180503104
姓名(Name)	Touhidur Rahman
指导教师姓名（职称） (Supervisor's Name (Title))	Wu Lingmei

教务处制

2022 年 3 月



## Table of Contents

1 Introduction.....	1
1.1 Research Background and research meaning.....	1
1.2 Approach.....	2
2 Literature Review.....	3
2.1 Research Background Study.....	3
3 System Design .....	5
3.1 System Functional Structure Design.....	5
3.2 Data Collection.....	7
3.3 Data Loading And Preprocessing.....	7
4 Model Building And Training.....	8
4.1 Model Layers.....	8
4.2 Model Details.....	9
4.3 Model Training.....	12
5 Experimental Result.....	12
5.1 Test Solutions.....	12
5.2 Positive Result.....	13
5.3 Negative Result.....	14
5.4 Evaluation Metrics.....	15
6 Acknowledgment .....	17
7 Conclusion .....	17
8 References.....	17
9 Author Profile .....	19
10 Declaration.....	19



## **Design and Implementation of Face Recognition System Based on Convolutional Neural Network (CNN)**

**Major:** Computer Science and Technology

**Name:** Touhidur Rahman

**Abstract:** Face recognition is one of the crucial corridors of active exploration areas and its major applications are in biometrical security, computer vision, health care, and marketing among other recognition technique face recognition have attained the most popular research area because of its high scaling and recognition rate and uniqueness. Nowadays deep learning methods are the most popular and Convolutional Neural Network is one of them and is widely used among many people. Basically in this research paper, we explore a particular system of face recognition that is the Siamese neural networks which serve a unique structure to naturally rank parallels between inputs. A convolutional neural network was designed and constructed, and the model was trained on 3 types of data, the positive and anchor data are collected from the webcam using OpenCV, and the negative data set is from labeled faces in the wild data set from the sanctioned website. The accuracy rate attained on the test set was around 97. Eventually, the trained model was applied to recognize faces using the method described in this paper and OpenCV. The system model is simple, occupies a small amount of memory, and can fast and accurately detect and recognize human faces.

**Keywords:** Face recognition, Convolutional Neural Network, Classification, Siamese Neural Networks



# 1 Introduction

## 1.1 Research Background and research meaning

Face recognition is one of the key parts of active research areas and also refers to a biometric feature that has an extensive range of applications such as identity management and security. It belongs to the category of computer vision and is the field of artificial intelligence such as deep learning and pattern recognition [1] in the new era, The face recognition system is currently employed for numerous applications with the introduction of deep convolutional neural networks, the accuracy of face recognition has been greatly improved and is widely used in access control security, attendance, candidate identification, character recognition, face payment,[1] healthcare technology, emotion recognition facilitates automated diagnosis of diseases such as Down syndrome. And other fields of diagnosis of diseases such as Down syndrome. It is also being devoted by the advertising media to understanding the customer's facial expressions and possible emotions. Face recognition poses certain challenges due to conditions such as poor lighting, differences in poses and background, occlusion, etc [2].

There are two orders for the face recognition task. The first one is the face verification task, it's a one-to-one matching problem. For illustration, when you unleash your phone using your face you use face verification, another illustration is in some airfields, you should pass through a system that scans your passport and your face to corroborate you're the correct person. The alternate bone is a face recognition task, to find answers to the question of who's this person. It's a one-to-numerous matching problem. Since CNN grounded ways have been used in the performance of some grueling tasks like face verification, face discovery is largely bettered. Another fashion to break the below tasks is called one-shot literacy. This system of learning representations from an illustration. In the Siamese neural networks (Siamese NNs), we calculate the encodings of the taken input image, also, with the same network without doing any updates on network parameters we take an image as an input of a different person and calculate its encoding. After these computations, we can check if there's a similarity between the two images[3].

The structure of the paper is as follows, discuss related works in section 2, and section 3 system design datasets. About model building and model training in section 4. Section 5 discusses the results of the experiment and then the conclusion of the work.

## 1.2 Approach

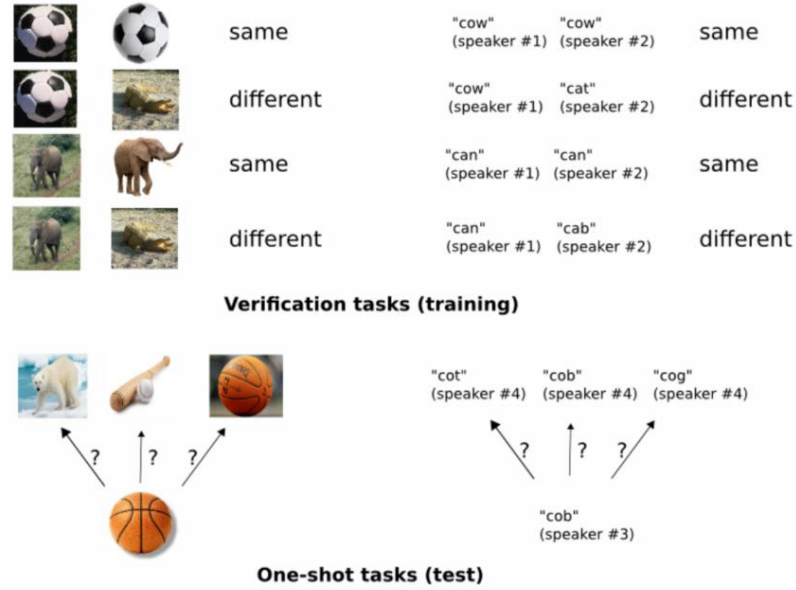


Figure 1: Siamese neural network basic structure.

Figure 1: In the paper, our general strategies are: 1) Train a model to differentiate between a collection of the same or different pairs. 2) Generalize to evaluate new categories based on learned characteristic mappings for verification.

Typically, we can learn image representations via a supervised metric-predicated approach with Siamese neural networks, and also apply that network's features for one-shot learning without any retraining. In our trials, we restrict our concentration to character recognition, although the primary approach can be replicated for nearly any modality ( Figure 1). For this sphere, we employ large Siamese CNN which a) are suitable for learning general image features useful for forming predictions about unknown class distributions indeed when truly numerous samples from these new distributions are available; b) are easily trained using standard



optimization ways on couples tried from the source data, and c) give a competitive way that doesn't calculate upon sphere-specific knowledge by rather exploiting deep knowledge ways. If we want to develop a model for a one-shot image recognition type, first we need to learn a neural network that can make the differentiation between the class identity of image pair, which is the standard verification job for image recognition system. We postulate that networks that do well at verification should generalize to a one-shot type. The verification model learns to pinpoint input couples according to the probability that they belong to identical classes or nonidentical classes. This model can also be used to estimate new images, precisely one per individual class, in a pairwise manner against the test image. The pairing with the topmost score corresponding to the verification network is also awarded the topmost probability for the one-shot task. Still, also they ought to be sufficient for other rudiments, handed that the model has been exposed to a variety of rudiments to encourage disunion amongst the learned features. If the features can be learned by the verification model are sufficient to confirm or deny the identity of characters from one set of rudiments [4].

## **2 Literature Review**

### **2.1 Research Background Study**

Proposes an age-invariant facial recognition system based on a discriminative model with deep feature training [5]. AlexNet is utilized as a transfer learning CNN model to learn high-level deep features in this study. These characteristics are then encoded into a code word with a higher dimension for visual representation using a codebook. The encoding system assures that the same person's face photos captured at different times have comparable code words. For face recognition, a linear regression-based classifier is utilized, and the algorithm is tested on three datasets, including the publicly accessible FGNET. Deep Stacked Denoising Sparse Autoencoders is a suggested face recognition system that comprises a Convolutional Neural

---

Network, an autoencoder, and denoising (DS-DSA). The classification technique employs Multiclass Support Vector Machine (SVM) and Softmax classifiers. ORL, Yale, and Pubfig are among the four publicly available datasets on which the approach is tested. The use of a deep local descriptor learning framework for cross-modality face recognition is proposed, in which both compact local information and discriminant features are directly learned from raw facial patches. A service robot-oriented face recognition system was proposed using the own dataset collected from the webcam and used a convolutional neural network to train on the dataset through many convolutional layers [6].

The proposed research has shown the implementation state of art face recognition methods and compared them. The advantages and disadvantages of Convolutional and Siamese neural networks are explored for the face recognition task. The novel face recognition system accuracy is checked on the widely used Labelled Faces in the Wild (LFW) dataset[3]. Authors of [7] have used the Weighted PCA-EFMNet deep-learning feature extraction method to solve problems related to changes in expression, position, illumination, and occlusion. Authors of [8] have proposed a part-based learning method for face verification, in which feature representation is extracted by a convolutional fusion network (CFN).

Star element analysis( PCA) on the network affair in confluence with an ensemble of Support Vector Machines (SVM) is used for the face verification task. In other words, the multi-stage fashion is used to align faces to a general 3D shape model [9]. The authors have trained over four thousand individuals to perform face recognition. The work explored the Siamese network and optimized the L1 distance between two face features. The exploration's stylish result on the LFW dataset was 97.35. The experimenter suggested a simple and cheap cipher network [10]. The study used an ensemble of twenty-five networks, each network operating on a different face patch. The final performance was 99.47 on the LFW dataset. In that work, the proposed system doesn't bear unequivocal 2D/ 3D alignment. Using a combination of bracket

---

and verification loss, the network is trained. The verification loss is analogous to the triplet loss that we bandy in the coming section and emplace in our face recognition system.

There are some of the authors have thought of other modalities. Lake et al. also have some truly recent work that uses a generative Hierarchical Hidden Markov model for speech heathens combined with a Bayesian conclusion procedure to fete individual words spoken by different speakers [11]. Maas and Kemp have some of the only published work using Bayesian networks to prognosticate attributes for Ellis Island passenger data by modifying the tentative distributions at each knot to be burdened by a soft index  $\lambda_i$  [12]. This is the so-called attention parameter" which measures whether the child is a near deterministic function of its parents. For certain observances, this allows some attributes to be linked as vital for inferring the idle variables of the graph. Wu and Dennis declamation of one-shot knowledge in the area of way planning algorithms for robotic actuation [13]. Actuation percepts arbitrated during knowledge can't inevitably be rephrased straightway to new jobs because environmental restrictions may have been modified. The authors present a mapping from the learned action template to a revised template by integrating the distance between these two point spaces parameterized by the coordinates of the Cartesian distortional place, implicitly defining an energy function that can be minimized.

## **3 System Design**

### **3.1 System Functional Structure Design**

The one specifically intriguing task is a bracket under the constraint that we may only observe a single illustration of each possible class before prognosticating a test case. This is called one-shot literacy, and this paper uses the Siamese convolutional Neural Network for One-shot face recognition. In our paper on Siamese convolutional neural network for one-shot face recognition, we are using three-part of images the anchor image, the positive image, and the negative image, and training the CNN model with these three types of images to learn the

---

pattern of the faces, then recognize the face using the trained CNN using one-short method model.

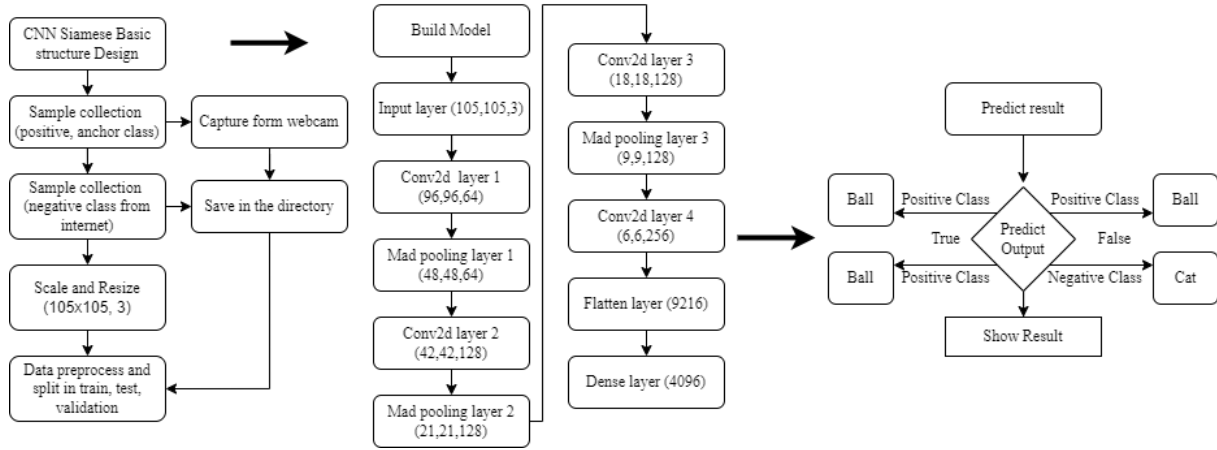


Figure 2: Siamese neural network basic structure design.

Figure 2: In the basic structure of the siamese convolutional neural network shows the basic flow chart of the operational steps from data sample collecting to data set building with preprocessing then model building steps with conventional and max-pooling layers, flatten and dense layer after that predicting result step.

The emergence of the convolutional neural network answered some failings of neural networks well, like the computational burden, the over-fitting of operation results, and the lack of original characteristics. Through its original open field, participating weights, and the time sphere or spatial sphere samples, the relegation, scaling, and deformation invariance of the results keep maintained. Siamese convolutional networks are two-inflow networks which mean two inputs. It can be used in object discovery, visual shadowing, and other tasks. It directly makes use of the characteristics of the complication operation, so the result can be expressed in some way[14]. In [15] the authors explore the training of deep networks to make these simple and abstract representations using labeled data, which are in the form of double analogous markers for training images to make dyads. [16] Shows us the successful operation of siamese

convolutional networks in visual shadowing. Tracking is also an important field of computer vision and the problem of arbitrary object Shadowing is solved by this system.

### **3.2 Data Collection**

This model uses the Lenovo Ideapad 5 laptop webcam to collect the data of 2 different classes in 2 particular folders labeled as positive and Anchors using the OpenCV library which is installed in the Miniconda environment, with the help of the webcam it takes 600 images of each class by pressing different keys in the keyboard, saves those images in positive and anchor directories while taking the images it also uses the library called UUID and names the images using some random naming system in a particular range. While taking images for positive and anchor classes it also saves the images using the frame of 250 x 250 which means it can reshape the pre-defined image size, by pressing "A" on the keyboard, it will take anchor image, by pressing "P" it will take the positive image and by pressing "Q" the image taking program will quit. For the negative class, it uses the data set from Labeled Faces in the Wild data set from the sanctioned website a database of face photos designed for studying the problem of unconstrained face recognition. The data set contains further than images of faces collected from the web. Each face has been labeled with the name of the person pictured. 1680 of the people pictured have two or further different prints in the data set. The only constraint on these faces is that they were detected by the Viola-Jones face sensor.

### **3.3 Data Loading And Preprocessing**

To load the image and preprocess we took 1000 images of each class through Tensorflow, Regarding the use of data sets, we need to take care of some important parts such as:

- 1) We used Data Augmentation in the image to enhance and increase the quality of the image and image samples.
- 2) To preprocess the image we need to scale and resize the images to 105 x 105 shapes so the image can fit in the CNN model, first, we need to read each image using TensorFlow IO and then decode the image into jpeg format after that preprocessing step use image

resize the image to be 105 x 105 x 3 using image function from TensorFlow framework then finally resize the images and preprocess it too. Create the labeled data set using TensorFlow we created leveled data set for positive and negative images because the original data can not be fitted in the training part so we need to convert them to a form that the computer can understand.

- 3) Then we created the data loader pipeline in the pipeline we preprocessed the data and make twin data using the buffer size of 10,000. we need to use the zip function from TF data and the dataset and labeled those images with the respective labels in the directory.
- 4) Then we need to split the data into two different partitions training and testing partition for the training partition. First, we need to build a data loader pipeline referencing the labeled dataset that we have created, to build the data loader pipeline we will map the data through the process twin function then we will cash the data and finally shuffle with the buffer size of 10000 to maintain the equality of the data.
- 5) In the training part, we are taking the total length of our dataset and then taking 70% of the data, and for the training partition then we are creating a batch of 16 and prefetch of 8 using the training data. All the training data will be used for the training model and the data will fit in the model using batch operations.
- 6) In the training part, we are taking the total length of our dataset and skipping the first 70% and taking the rest of the 30% of data, and for the testing partition then we are creating a batch of 8 and prefetch of 8 using the training data. All the testing data will be used for the testing model and the data will fit in the model using batch operations.

## 4 Model Building And Training

### 4.1 Model Layers

Our standard model is a siamese convolutional neural network with  $L$  layers each with  $N_1$  units, where  $h_{1,l}$  represents the retired vector in the layer of  $l$  for the first twin, and  $h_{2,l}$  denotes the equal for the alternate twin. We use simply remedied direct (ReLU) units in the first

---

$L - 2$  layers and sigmoidal units in the remaining layers. The model consists of a sequence of convolutional layers, each of which uses a single channel with pollutants of contrasting sizes and a fixed stride of 1. The number of convolutional pollutants is specified as a multiple of 16 to optimize performance. The network applies a ReLU activation function to the affair point charts, voluntarily followed by maximum-pooling with sludge size and stride of 2.

$$a_{1,m}^{(k)} = \max\text{-pool}(\max(0, W_{l-1,l}^{(k)} * h_{1,(l-1)} + b_l), 2) \quad (1)$$

$$a_{2,m}^{(k)} = \max\text{-pool}(\max(0, W_{l-1,l}^{(k)} * h_{2,(l-1)} + b_l), 2) \quad (2)$$

Where  $W_{l-1,l}$  is the 3-dimensional tensor representing the feature maps for layer  $l$  and we have taken  $*$  to be the valid convolutional operation corresponding to returning.

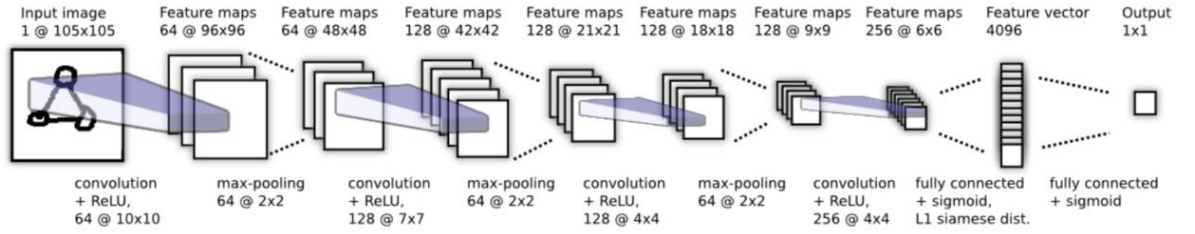


Figure 3: Different layers of convolutional neural network.

Figure 3: Better convolutional architecture is selected for the verification task. Siamese twin is not delineated but joins incontinently after the 4096 unit it fully-connected layer wh1 and the L1 component-wise distance between vectors is computed.

## 4.2 Model Details

Our standard model is a siamese convolutional neural network with 4 conventions, 3 max polling, 1 flattened, and 1 dense layer with Relu and Sigmoid activation.

Layer (type)	Output Shape	Param #
=====		
input_image (InputLayer)	[(None, 105, 105, 3)]	0
conv2d_4 (Conv2D)	(None, 96, 96, 64)	19264
max_pooling2d_3 (MaxPooling 2D)	(None, 48, 48, 64)	0
conv2d_5 (Conv2D)	(None, 42, 42, 128)	401536
max_pooling2d_4 (MaxPooling 2D)	(None, 21, 21, 128)	0
conv2d_6 (Conv2D)	(None, 18, 18, 128)	262272
max_pooling2d_5 (MaxPooling 2D)	(None, 9, 9, 128)	0
conv2d_7 (Conv2D)	(None, 6, 6, 256)	524544
flatten_1 (Flatten)	(None, 9216)	0
dense_1 (Dense)	(None, 4096)	37752832
=====		
Total params: 38,960,448		
Trainable params: 38,960,448		
Non-trainable params: 0		
=====		

Figure 4: The convolutional neural network model.

As shown in Figure 4, the constructed convolutional neural the network has the following segments:

**First block:** The first input convolution layer will take input of 105x105 shape with the ReLu of (64, 10x10) size. Then the max-pooling of (64, 2x2) size will take the input from the first layer and the feature map will have (64, 96x96) output.



**Second block:** Then second convolution layer with the ReLu of (128, 7x7) shape will take input from the first convolution layer. And the max-pooling will be (64, 2x2) size.

**Third block:** Then third convolution layer with the ReLu of (128, 4x4) shape will take input from the second convolution layer. And the max-pooling will be (64, 2x2) size.

**Fourth and Final block:** Then the final convolution layer with the ReLu of (256, 4x4) shape will take input from the third convolution layer and will go through a flattened layer then will go through a dense layer and will have a sigmoid activation. And at the will have Total params: 38,960,448, Trainable params: 38,960,448, Non-trainable params: 0. After adding the L1 Distance Layer the status will be Total params: 38,964,545, Trainable params: 38,964,545, Non-trainable params: 0.

Layer (type)	Output Shape	Param #	Connected to
input_img (InputLayer)	[(None, 105, 105, 3 ) ]	0	[ ]
validation_img (InputLayer)	[(None, 105, 105, 3 ) ]	0	[ ]
embedding (Functional)	(None, 4096)	38960448	['input_img[0][0]', 'validation_img[0][0]']
distance (L1Dist)	(None, 4096)	0	['embedding[2][0]', 'embedding[3][0]']
dense_3 (Dense)	(None, 1)	4097	['distance[0][0]']
=====			
Total params: 38,964,545			
Trainable params: 38,964,545			
Non-trainable params: 0			

Figure 5: Final summary of Siamese CNN Model.

The siamese network model in the final stage has the Input\_img layer (Input layer) with the shape of (None, 105,105,3) and 0 param. The Validation\_img layer (Input layer) with the shape of (None, 105,105,3) and 0 param. The Embedding layer (Functional) with the shape of (None, 4096) and 38960448 param and connected to input\_img[0][0] and validation\_img[0][0] .

The Distance layer (L1Dist) with the shape of (None, 4096) and 0 param. and connected to embedding[2][0] and embedding[3][0]. The Dense\_3 layer (Dense) with the shape of (None, 1) and 4097 param. and connected to distance [0][0]. Then total params will be 38,964,545, trainable params will be 38,964,545 and non-trainable params will be 0.

### 4.3 Model Training

The siamese one-short CNN model is trained with 1000 samples on each class of the dataset over 50 EPOCHS the achieve a good prediction rate.

Topic	Start status	End status
Epoch	1	50
Loss	0.2126203179359436	9.872399459709413e-06
Recall	0.7819225192070007	1.0
Precision	0.9909090995788574	1.0
Time	68s 349ms/step	64s 365ms/step

Table 1: Model training status during the training period.

As we can see from table 1: The binary cross-entropy loss function on the first EPOCHS, the status was: estimated time 68s 349ms/step, binary cross-entropy loss was small but at the last epoch, the binary cross-entropy loss becomes high a little. The recall of the model during the training period was low but at the last epoch become 1.0 which is good, as well as the precision, which was low in the first epoch but become 1.0 at epoch 50.

## 5 Experimental Result

### 5.1 Test Solutions

The software environment of this system is a window the system, the programming language used is Python, and Tensorflow is used as a framework. With those assets, the model

---

is having 1.0 precision and 1.0 Recall and the prediction rate is 90% on average when the light environment is good. From the actual results. The system can achieve practical results.

## 5.2 Positive Result:

From the result, we can see the test input and validation input index are the same and both results are shown in figure 6.

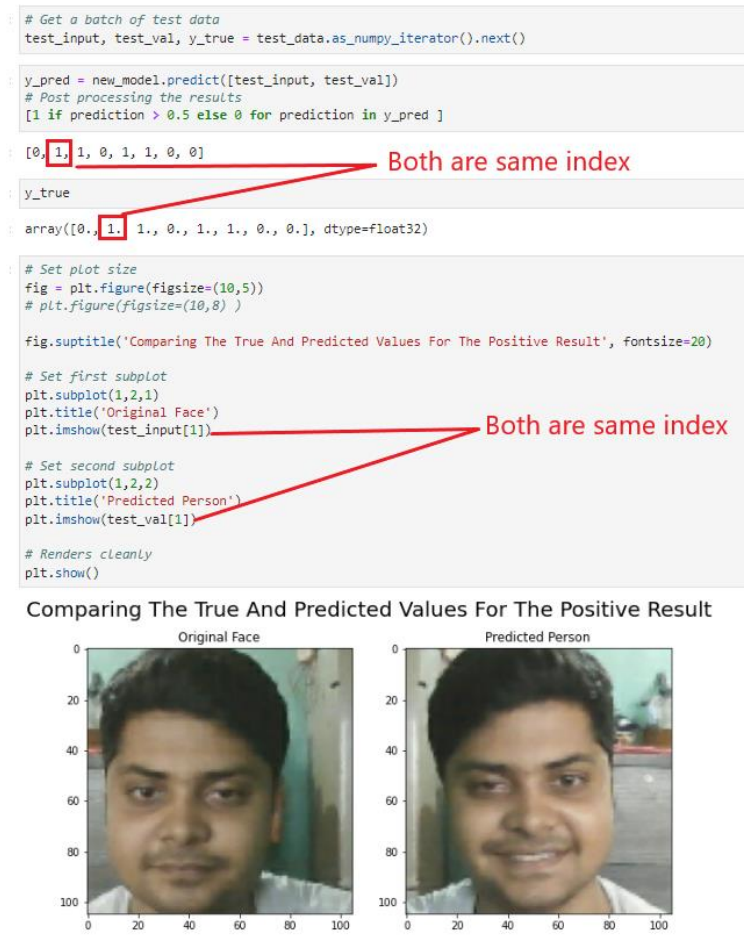


Figure 6: Actual result of CNN model showing a positive result.

In figure 6: As we can see the test input index and test validation index are the same and the predicted person is also the same, which means the model is working fine and can recognize

the face of the person using a siamese network model. The model is comparing the test result and the validation result to show the result here and recognize the real person. If the test result and test validation result do not match then it is regarded as the nonmatching result.

### 5.3 Negative Result:

From the result, we can see the test input and validation input index are the same and both results are shown in figure 7.

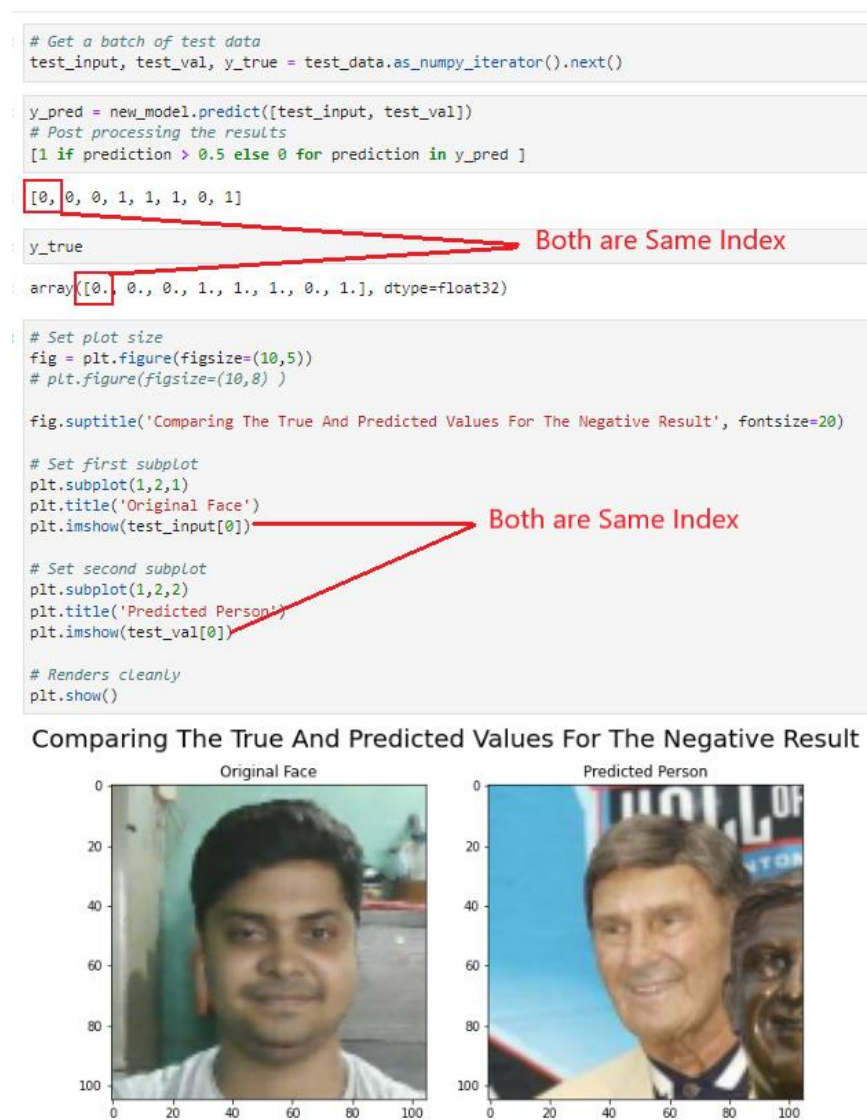


Figure 7: Actual result of CNN model showing a negative result.

In figure 7: As we can see the test input index and test validation index are the same and the predicted person is not the same, so the model is showing a correct negative result. which means the model is working fine and can recognize and can identify the negative result too using a siamese network model. The model is comparing the test result and the validation result to show the result here.

#### 5.4 Evaluation Metrics:

**Confusion matrix:** Bellow, we can see the confusion matrix and get a proper idea of true or false positive and negative prediction rates.

Confusion Matrix of Siamese Face Recognition System Based on CNN

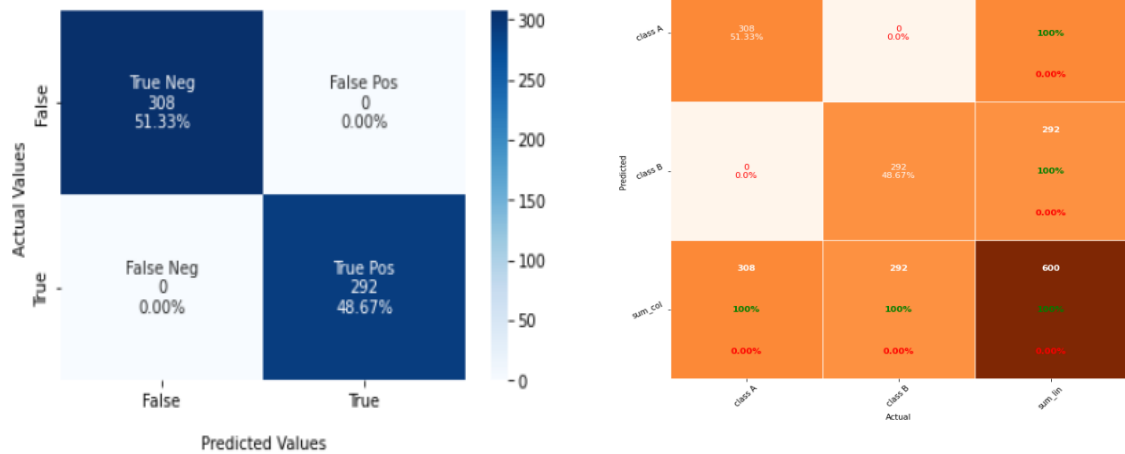


Figure 8: The confusion matrix of the Siamese CNN Model based on the test data sample.

Figure 8: The confusion Matrix shows that there are 308 true negative predictions with 51.33 % and a true positive prediction of 292 and 48.67 % based on testing data samples. respectively the is 0 false positive with 0.00% and a false negative of 0 with 0.00% which is also based on the same sample test data.

**Precision and Recall Curve:** When using classification models in machine learning, two metrics we often use to assess the quality of the model are precision and recall.

**Precision:** Correct positive predictions relative to total positive predictions.  $\text{Precision} = \text{True Positives} / (\text{True Positives} + \text{False Positives})$ .

**Recall:** Correct positive predictions relative to total actual positives.  $\text{Recall} = \text{True Positives} / (\text{True Positives} + \text{False Negatives})$

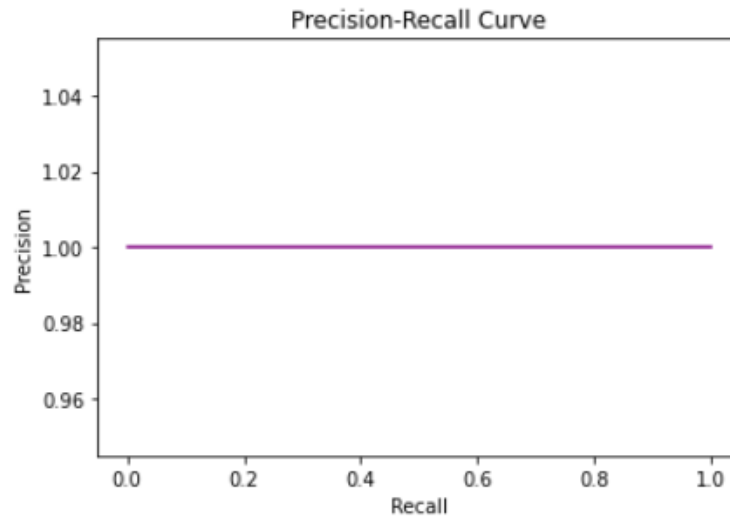


Figure 9: The confusion matrix of the Siamese CNN Model based on the test data sample.

Figure 9: shows the precision and recall of the siamese CNN model for face recognition. To visualize the precision and recall for a certain model, we can create a precision-recall curve. This curve shows the tradeoff between precision and recalls for different thresholds.

## 6 Acknowledgment

This thesis is granted by Guangxi Science and Technology Normal University I am thankful to get a wonderful opportunity to research on this particular topic, and I am really thankful to Professor Wu Ling Mei. It helped me a lot during this thesis and I am also thankful to those people who help me to complete this thesis paper by giving good references and other materials.

## 7 Conclusion

The face recognition system completed four steps face collection and preprocessing, model establishment, model training, and real-time recognition. Based on the self-built data set and collected dataset from the internet, a convolutional neural network was built and trained. After continuously adjusting parameters, the accuracy rate on the test set was 97.63%. The actual test was performed during the testing period, and the average recognition rate was 90% under good light conditions. The advantages of the face recognition system are that the model is simple, and the memory capacity is small. It is also suitable for practical application scenarios. It can also achieve better recognition results for face recognition within a certain angle. However, the robustness of light needs to be strengthened, and it should be improved in subsequent work.

## 8 References

- [1] M. Nakada, H. Wang, and D. Terzopoulos, "AcFR: Active Face Recognition Using Convolutional Neural Networks."- July 2017
- [2] Francis Xavier Engineering College and Institute of Electrical and Electronics Engineers, *Proceedings of the 2nd International Conference on Smart Systems and Inventive Technology (ICSSIT 2019) : 27-29, November 2019.*

- [3] W. Cui, W. Zhan, J. Yu, C. Sun, and Y. Zhang, "Face recognition via convolutional neural networks and Siamese neural networks," in *Proceedings - 2019 International Conference on Intelligent Computing, Automation, and Systems, ICICAS 2019*, Dec. 2019, pp. 746–750. DOI: 10.1109/ICICAS48597.2019.00161.
- [4] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese Neural Networks for One-shot Image Recognition." Vol 2. Year 2015
- [5] M. S. Shakeel and K. M. Lam, "Deep-feature encoding-based discriminative model for age-invariant face recognition," *Pattern Recognition*, vol. 93, pp. 442–457, Sep. 2019, DOI: 10.1016/j.patcog.2019.04.028.
- [6] Chinese Association of Automation, M. IEEE Systems, and Institute of Electrical and Electronics Engineers, *Proceedings, 2019 Chinese Automation Congress (CAC2019) : Nov. 22-24, 2019, Hangzhou, China.*
- [7] *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*. IEEE, 2018.
- [8] C. Xiong, L. Liu, X. Zhao, S. Yan, and T. K. Kim, "Convolutional fusion network for face verification in the wild," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 3, pp. 517–528, Mar. 2016, DOI: 10.1109/TCSVT.2015.2406191.
- [9] E., et al van der Spoel, "Siamese Neural Networks for One-Shot Image Recognition," vol. 7, pp. 956–963, 2015.
- [10] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep Learning Face Representation by Joint Identification-Verification." year 2014, <https://doi.org/10.48550/arXiv.1406.4773>
- [11] B. Lake, C.-Y. Lee, J. Glass, B. M. Lake, J. R. Glass, and J. B. Tenenbaum, "UC Merced Proceedings of the Annual Meeting of the Cognitive Science Society Title One-shot learning of generative speech concepts Publication Date One-shot learning of generative speech concepts," no. 36, p. 36. Year 2014
- [12] A. L. Maas and C. Kemp, "One-Shot Learning with Bayesian Networks." Year 2019
- [13] Ieee and Ieee, *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops.*



- [14] Q. Li, Institute of Electrical and Electronics Engineers, and IEEE Engineering in Medicine and Biology Society, *CISP-BMEI 2017 : proceedings, 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics : 14-16 October 2017, Shanghai, China*.
- [15] C. Liu, “Probabilistic Siamese Network for Learning Representations,” 2013.
- [16] G. Hua and H. Jégou, Eds., *Computer Vision – ECCV 2016 Workshops*, vol. 9914. Cham: Springer International Publishing, 2016. DOI: 10.1007/978-3-319-48881-3.

## 9 Author Profile

Name: Touhidur Rahman

College of Mathematics & Computer Science

Major: Computer Science and Technology

Class of 2018

## 10 Declaration

I solemnly declare that this thesis (design) is completed independently under the guidance of my supervisor during my study at the Guangxi Institute of Science and Technology. Its content is true and reliable, if there is plagiarism, plagiarism, I am willing to assume full responsibility. At the same time, I fully understand and am willing to abide by the regulations of Guangxi Institute of Science and Technology Teacher's College regarding the preservation and use of the thesis (design), which include.

1. The college has the right to keep and submit the original and copies of the thesis (design) to the relevant departments.

2. The college may adopt photocopying, microfilming, or other reproduction methods to preserve the thesis (design).
3. The college may present and exchange the thesis (design) for academic exchange.
4. The college may allow the thesis (design) to be consulted or borrowed.
5. The college may publish all or part of the thesis (design) following the provisions of the copyright law (confidential thesis (design) is subject to this provision after declassification).

The scientific research results of this thesis belong to the Guangxi Institute of Science and Technology unless otherwise constrained by scientific research contracts or other legal documents.

Hereby declare!

Signature of declarant: Touhidur Rahman

2022 year   month 04   day 16