**Department of Electrical and Computer Engineering**

**North South University**



**CSE498R**

**Directed Research**

**Automated Restaurant Food Review System using CNN**

**Niazi Mahrab ID-1711248042**

**Abdullah Ibne Ali ID: 1711200042**

**Israt Jahan Mim ID: 1711392042**

**Faculty Advisor:**

**Dr. Riasat Khan**

**Assistant Professor**

**ECE Department**

**Fall, 2020**

**LETTER OF TRANSMITAL**

21 February, 2021

To

Dr. Mohammad Rezaul Bari

Chairman,

Department of Electrical and Computer Engineering

North South University, Dhaka

Subject: **Submission of Directed Research Report on “**Automated Restaurant Food Review System using CNN**”**

Dear Sir,

With due respect, we would like to submit our **Directed Research Report** on **“**Automated Restaurant Food Review System using CNN**”** as a part of our BSc program. This project was very much valuable to us as it helped us gain experience from practical field and apply in real life. We tried to the maximum competence to meet all the dimensions required from this report.

We will be highly obliged if you kindly receive this report and provide your valuable judgment. It would be our immense pleasure if you find this report useful and informative to have an apparent perspective on the issue.

Sincerely Yours,

Niazi Mahrab

.........................................................

ECE Department

North South University, Bangladesh

Abdullah Ibne Ali

........................................................

ECE Department

North South University, Bangladesh

Israt Jahan Mim

.........................................................

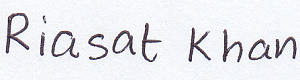
ECE Department

North South University, Bangladesh

**APPROVAL**

Niazi Mahrab (ID # 1711248042), Abdullah Ibne Ali (ID # 1711200042) and Israt Jahan Mim (ID # 1711392042) from Electrical and Computer Engineering Department of North South University, have worked on the Directed Research “Automated Restaurant Food Review System using CNN” under the supervision of Dr. Riasat Khan partial fulfillment of the requirement for the degree of Bachelors of Science in Engineering and has been accepted as satisfactory.

**Supervisor’s Signature**

…………………………………….

**Dr. Riasat Khan**

**Assistant Professor**

Department of Electrical and Computer Engineering

North South University

Dhaka, Bangladesh.

**Chairman’s Signature**

…………………………………….

**Dr. Mohammad Rezaul Bari**

**Associate Professor**

Department of Electrical and Computer Engineering

North South University

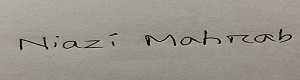
Dhaka, Bangladesh.

**DECLARATION**

This is to certify that this Project is our original work. No part of this work has been submitted elsewhere partially or fully for the award of any other degree or diploma. Any material reproduced in this project has been properly acknowledged.

Students’ names & Signatures

**1.** Niazi Mahrab

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

**2.** Abdullah Ibne Ali



\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

**3.** Israt Jahan Mim



\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

**ACKNOWLEDGEMENTS**

By kindness of the Almighty we have successfully completed our directed research entitled “Automated Restaurant Food Review System using CNN”

Our deep gratitude goes first to my faculty advisor Dr. Riasat Khan, who expertly guided us in our senior design project throughout the whole CSE498R. His guidance helped us in all type of research, writings and completing the project.

Our sincere thanks also go to North South University, Dhaka, Bangladesh for giving us such a platform where we can have an industrial level experience as a part of our academics.

We would also like to thank my friends Saad Ahmed Salim for helping us in this project.

Last but not the least, we would like to thank our family as their inspiration and guidance kept us focused and motivated.

Contents

[**Abstract ----** 8](#_Toc64970471)

[*I.* ***Keywords****— deep learning, convolutional neural network (CNN), face algorithm features, binary classification* 8](#_Toc64970472)

[**II.** **Introduction** 8](#_Toc64970473)

[**III.** **Related Work** 9](#_Toc64970474)

[**IV.** **Proposed System Methodology** 11](#_Toc64970475)

[A. **Dataset** 11](#_Toc64970476)

[B. **Data pre-processing and augmentation** 11](#_Toc64970478)

[C. **Modeling** 12](#_Toc64970479)

[D. **Image Cases:** 13](#_Toc64970480)

[**V.** **Performance evaluation** 15](#_Toc64970481)

[**VI.** **Conclusions** 19](#_Toc64970482)

[**VII.** **References** 20](#_Toc64970483)

**Automated Restaurant Food Review System using CN****N**

# **Abstract ----** A large amount of money is added every year to the economy through the restaurant business in a country. Nowadays, the restaurant business in Bangladesh has become very popular because of the increasing number of customers and profitability. Different people prefer various types of food in the restaurant; moreover, they order food without knowing the quality and the taste of the food. Because there are a few reviewed systems for customers, they are mostly App-based. As a result, the customer does not have any appropriate knowledge about the restaurant and the food. We tried to apply such deep learning techniques for the food review system by recognizing facial expressions with the help of Convolutional Neural Network (CNN) and the FER (2013) dataset which is an open-source dataset. The experiment results show that CNN performs significantly better and we got high accuracy rate by using face detection algorithm.

# ***Keywords****— deep learning, convolutional neural network (CNN), face algorithm features, binary classification*

# **Introduction**

Bangladesh is a developing country with a large population. The people of working sectors play a tremendous role in this development. A big part of these working people works far from their home. Besides, they have no other option without a restaurant. So, they eat at the restaurant to save time and for getting a good quality of food. Not only working people but also many people go to the restaurant for several reasons. Above all, the restaurant becomes an important and busy place day by day. People eat food in a different restaurant without knowing the quality. Some countable restaurants started a food review system; those systems are mostly app-based. We have decided on a food review system using MI by observing this issue. We basically apply deep learning for the food review system.

# **Related Work**

In this section, we have discussed about some papers that are related to our work. [1] facial expression recognition system has been used through deep learning. For our project, we will use only Convolutional Neural Network (CNN). We will follow the article instructions like Haar-like based face detection algorithm, OpenCV library, Caffe for developing real-time application operated on Windows system. This conference paper has been chosen because we will use FER-2013 dataset as it is an open source dataset with 35887 images. And in our project, we will use happy and sad images for our detection.

[2] Imane Lasri et al., made an intelligent classroom management system using Convolutional neural networks (CNN) that extracts facial emotions of students. Our system is consisted with three phases like the Imane Lasri. But we select two types of emotions instead of seven. The system includes face detection, normalization and emotion recognition. Like Imane Lasri, we used the FER-2013 database which contains 35887 images to train CNN architecture and OpenCV library which are used capture live frames from web camera. They achieved 70% accuracy rate at the 106 epochs and we achieved 81% accuracy rate at the 95 epochs. Their main idea was to use this live emotion of students to inform the teacher to improve his/her material.

[3] Shima Alizadeh, Azar Fazel made a model for a facial expression recognition task. They developed CNNs with variable depths so that it evaluates the performance of the models. They used a dataset of 37,000 gray scale images from Kaggle website which are 48x48 pixel well-structured images of faces. Our dataset images are also same pixel, so it is easier for us to implement CNN and to train with 4 convolutional layer and we’ve trained one test and one deep network We got 55% validation accuracy and 58% train accuracy in our first epoch at first. But after several times, with integrations, it gets higher. The detection of facial expression is processed as a binary classification and represent human emotions. We have used the face detection algorithm, OpenCV library. We get an 81% accuracy rate using 18968 images.

[4] by Mao Xu et. al, author wanted to show an efficient facial expression recognition based on transfer features from deep convolution network. They take 2062 imbalanced samples depending on four facial expression databases like CK+, JAFFE, KDEF, Pain expressions form PICS. They achieve 80.49% recognition rate with the seven-class SVM classifier on the self-built facial expression database. They process the database by using Viola-Jones face detection algorithm and artificial selection. To show transfer features from deep convolution network they construct the validation sets of different occluded ratio from 0 to 0.5 and then test their model on the validation sets.

# **Proposed System Methodology**

## **Dataset**

## We are using FER (2013) dataset [5]. This is an open source. This is merged by the emotions of Happiness and Surprise under interested, Anger, Disgust and Fear under Disappointed and lastly the emotions of neutral and indifferent under the neutral folder. We make images into binary classification by taking only interested and disappointed. There are two folders. They are train and test images. There are total 24000 images in our dataset. There are 18968 training images and 6016 testing images. We separate images as our need. Each interested and disappointed class contains 9484 training images and 3008 test images. The dataset comes in CSV format with three columns emotion, pixels and usage. The pixel column contains the pixel values of the images in 1D. So, we converted the 1D list of pixels into 2D to get the width and height of the images which is 100 by 100 according to the dataset.

## **Data pre-processing and augmentation**

Pre-processing is the input image data to convert it into meaningful floating-point tensors for feeding into Convolutional Neural Networks. It is very important to get a good accuracy. Tensors are used to store data, they can be assumed as multidimensional arrays. It is very challenging to know how to best prepare image data when training a convolutional neural network. Because at first images are not ready to train. They are in different shape, different size, different angle. This involves both scaling the pixel values and use of augmentation techniques during both the training and evaluation of the model. Our images are 64\*64 pixels. We reshaped it by 100\*100 pixels. The system also requires a dimension. A tensor representing a 100 X 100 image having 3 channels will have its dimensions (100, 100, 3). We have also used rescale=1./255,horizontal\_flip=.2,zoom\_range=.5, rotation\_range=0.4. These help to get a good validation accuracy.

## **Modeling**

Computer vision: It is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do [6]. Image processing applies mathematical functions to images. The outcome of image processing may or may not provide detailed data. Image processing performs operations such as sharpening, smoothing, stretching, and contrasting on an image. Computers perceive images as 2D signals that consist of rows and columns of pixels.

Convolutional Neural Network: A Convolutional Neural Network (CNN) is a deep artificial neural network that can identify visual patterns from input image with minimal pre-processing compared to other image classification algorithms. That means the network learns the filters that in traditional algorithms were hand engineered [7]. Here is an image of our hyperparameter we have used in our model:

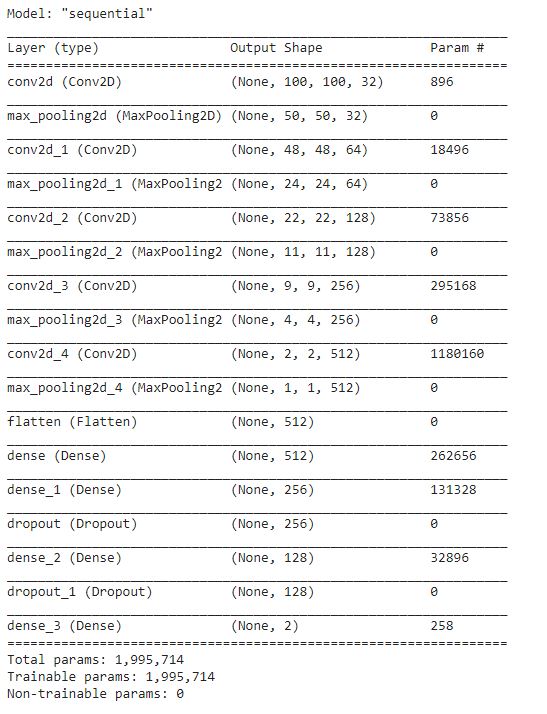


Figure 1: Hyperparameters

Max Pooling: Its nothing but selecting maximum value from the matrix of specified size. This method is helpful to extract features with high importance or which are high-lighted in the image. [8]

Rectified Linear Unit: The rectified linear unit (ReLU) in a neural network is used to introduce the non-linearity. The ReLU simply works like a filter, ignores the negative signals and pass the positive signals.

## **Image Cases:**

We developed a facial expression detection using a Convolutional Neural network (CNN) architecture. In this part we are going to describe the methods we proposed in our system.

We used binary Classification for this system where we can detect a Happy or Disappointed face. First, the system detects the face from input then it gets cropped and normalize the face. Then it goes to the CNN input to predict it’s a happy or a disappointed face. Here is a structure for our proposed system.

relu function (rectified linear: f(x)=max (0, a) = max (0, summation of wixi+b))

The relu simply works like a filter, ignores the negative signals and pass the positive signals. [9]

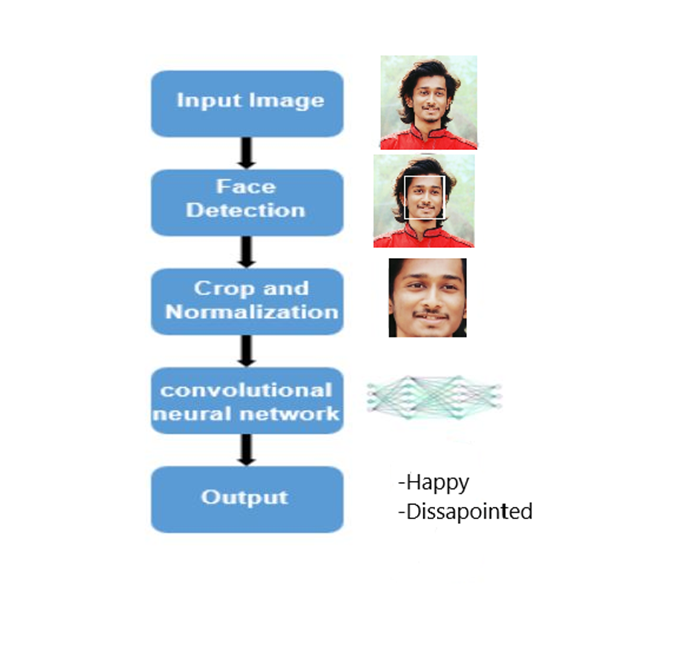
sigmoid function (Sigmoid activation function, sigmoid(x) = 1 / (1 + exp(-x)). It ranges from 0 to 1.

Sigmoid is equivalent to a 2-element Softmax, where the second element is assumed to be zero. We use this function

to predict the probability as an output. (As the sigmoid function always returns a value between 0 and 1.)

In this scenario, our number of epochs to train the dataset is 20 and total number of steps before declaring one epoch to finish and start the next epoch is 95. Validation data will override itself and the number of steps to draw before stopping when performing validation, steps need 31.

It reads all the images from the folder specified and we also defined the arguments precisely to arbitrary transformations. Then we've selected our path directory and the transformation images are returned in the form of generator like augmented data.



# **Performance evaluation**

We got 55% validation accuracy and 58% train accuracy in our first epoch which is the lowest. Finally, we are able to get 81% validation accuracy and 97% train accuracy from our model. Though there are many pre-trained models are available, but we tried to make binary classification.

We took datasets from Kaggle named Facial-Expression-Classification-Dataset and make it 2 classes. Then we tried to give input photo of ours. We’ve seen that every 10 photos, we got 8 corrected facial expression review. 2 was wrong because of less accuracy.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Epochs** | | **Accuracy** | | **Loss Function** | | **Processing Time(s)** |
| 1 | | 0.5554 | | 0.6712 | | 358 |
| 2 | | 0.5994 | | 0.6615 | | 330 |
| 3 | | 0.6644 | | 0.6247 | | 330 |
| 4 | | 0.6965 | | 0.5686 | | 326 |
| 5 | | 0.7252 | | 0.5327 | | 331 |
| 6 | | 0.6902 | | 0.6012 | | 328 |
| 7 | | 0.7768 | | 0.4600 | | 331 |
| 8 | | 0.7914 | | 0.4533 | | 329 |
| 9 | | 0.7621 | | 0.4752 | | 380 |
| 10 | | 0.8130 | | 0.4266 | | 364 |
| 11 | | 0.7970 | | 0.4512 | | 350 |
| 12 | | 0.8135 | | 0.4768 | | 334 |
| 13 | | 0.7955 | | 0.4954 | | 383 |
| 14 | | 0.8182 | | 0.5197 | | 357 |
| 15 | | 0.8105 | | 0.6602 | | 358 |
| 16 | | 0.8055 | | 0.6539 | | 328 |
| 17 | | 0.7984 | | 0.6892 | | 330 |
| 18 | 0.8138 | | 0.7637 | | 332 | |
| 19 | 0.8113 | | 0.8902 | | 330 | |
| 20 | 0.8032 | | 0.9301 | | 351 | |

Table 1: Showing accuracy with the help of epoch

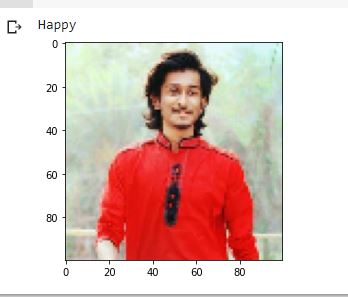
The accuracy can be increased by augmentation and increasing epochs. By increasing epochs and adding augmentation we can increase our accuracy a little bit. From final output we can make an

Figure 2: Output result

average review on a restaurant food. We could also use other features such as Histogram of Oriented Gradients (HOG) for feature extractions. These could help in getting better accuracy and improve the review prediction.

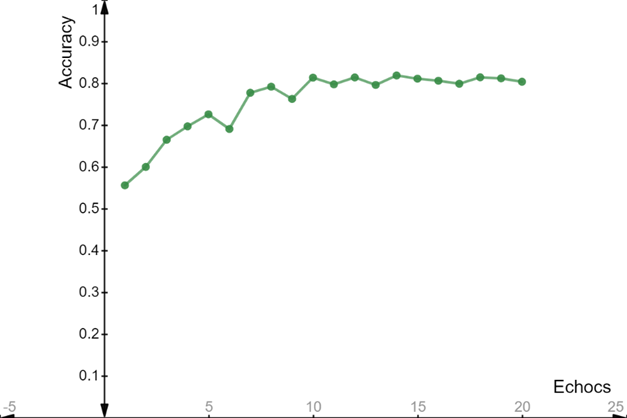
We can get a graph from this relation between epoch and accuracy: 

Figure 3: Epochs vs. accuracy graph

This graph shows us, accuracy is increasing in every epoch.

Here is percentage of happy and disappointed:

|  |  |  |
| --- | --- | --- |
| Step | Happy | Disappointed |
| 1 | 67.04 | 68.56 |
| 2 | 70.55 | 71.45 |
| 3 | 72.66 | 73.67 |
| 4 | 74.44 | 75.44 |
| 5 | 80.33 | 79.78 |

Table 2: happy and disappointed percentage

**Confusion Matrix:**

The confusion matrix shows the ways in which your classification model is confused when it makes predictions. We got 8 correct prediction from 10 predicts.

We know, classification accuracy = correct predictions / total predictions \* 100

= 8/10 \*100 = 80%

error rate = (1 - (correct predictions / total predictions)) \* 100

= (1- (8/10)) \* 100 = 20%

Here there is a graph between loss and iteration.

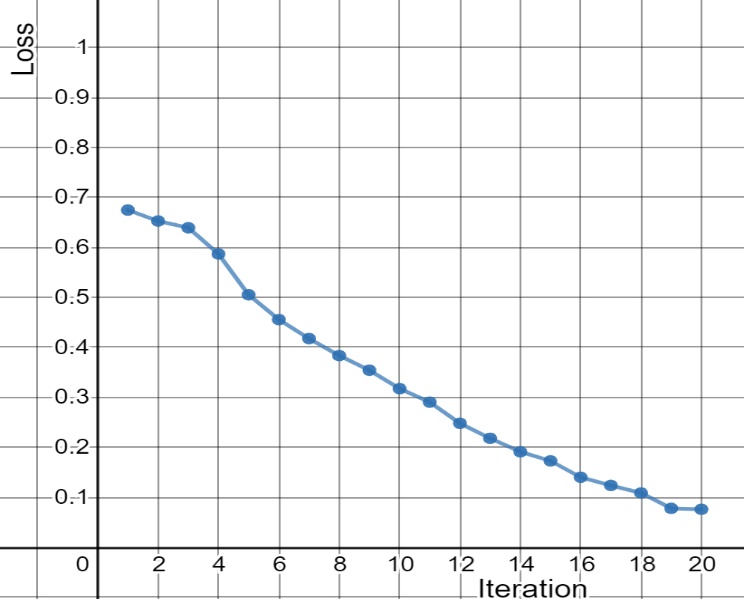


Figure 4: Loss vs. iteration graph

# **Conclusions**

Our main object in the project to get a review from a customer by AI. So, accuracy was not our target. If we can make such kind of project then customers will get an accurate review. They don’t have to be suffered by wrong reviews from different food bank reviews. In future we have plan to make it real time project by using IOT. Then we can use it with direct security camera and able to see real time reviews.

# **References**

[1] H. Jung et al., "Development of deep learning-based facial expression recognition system," 2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), Mokpo, 2015, pp. 1-4, doi: 10.1109/FCV.2015.7103729.

[2] Imane Lasri, Anouar Riad Solh, Mourad El Belkacemi, “Facial Emotion Recognition of Students using Convolutional Neural Network”.

[3] Shima Alizadeh, Azar Fazel, Convolutional Neural Networks for Facial Expression Recognition, Stanford University, 2017.

[4] Mao Xu, Wei Cheng, Qian Zhao, Li Ma and Fang Xu, "Facial expression recognition based on transfer learning from deep convolutional networks," 2015 11th International Conference on Natural Computation (ICNC), Zhangjiajie, 2015, pp. 702-708, doi: 10.1109/ICNC.2015.7378076

[5] https://www.kaggle.com/nightfury007/fercustomdataset-3classes

[6] Shima Alizadeh, Azar Fazel ‘Convolutional Neural Networks for Facial Expression Recognition’

[7] Imane Lasri, Anouar Riad Solh, Mourad El Belkacemi ‘Facial Emotion Recognition of Students using Convolutional Neural Network’

[8,9] https://arxiv.org/pdf/1804.02051Natural Computation (ICNC), Zhangjiajie, 2015, pp. 702-708, doi: 10.1109/ICNC.2015.7378076