Symposium on
Machine Learning and Metaheuristics Algorithms, and Applications
SOMMA'19
IIITM-Kerala
December 18 - 21, 2019 |Trivandrum, Kerala, India

AMRITA VISHWA VIDYAPEETHAM | School of Engineering

Facial Emotion Recognition Using Shallow CNN

Sachin Saj T K, G Seshu Babu, Vamsi Reddy, Gopika, Sowmya V, Soman K P
Amrita School of Engineering,
Coimbatore, Tamil Nadu, India

Agenda

- Objective
- Introduction
- Flow Chart
- Dataset - JAFFE
- Proposed CNN Architecture
- Results
  - Confusion Matrix
  - Classification Report
  - Comparison of existing approaches
- Conclusion
- Reference

Objective

To Propose a model which can do Facial Emotion Recognition with reduced computational complexity as well as gives better performance.

Introduction

- Emotion recognition is a technique that allows reading the emotions on a human face using advanced image processing.
- Emotions are divided into six types: Happiness, Anger, Sadness, Fear, Disgust, and Surprise.

Applications

- During healthcare, determining patients feeling and comfort level about the treatment
- Security Systems
- Interactive Computer Simulations/designs
- Driver Fatigue Monitoring
- Psychology and Computer Vision

Flow Chart - For our Proposed Approach

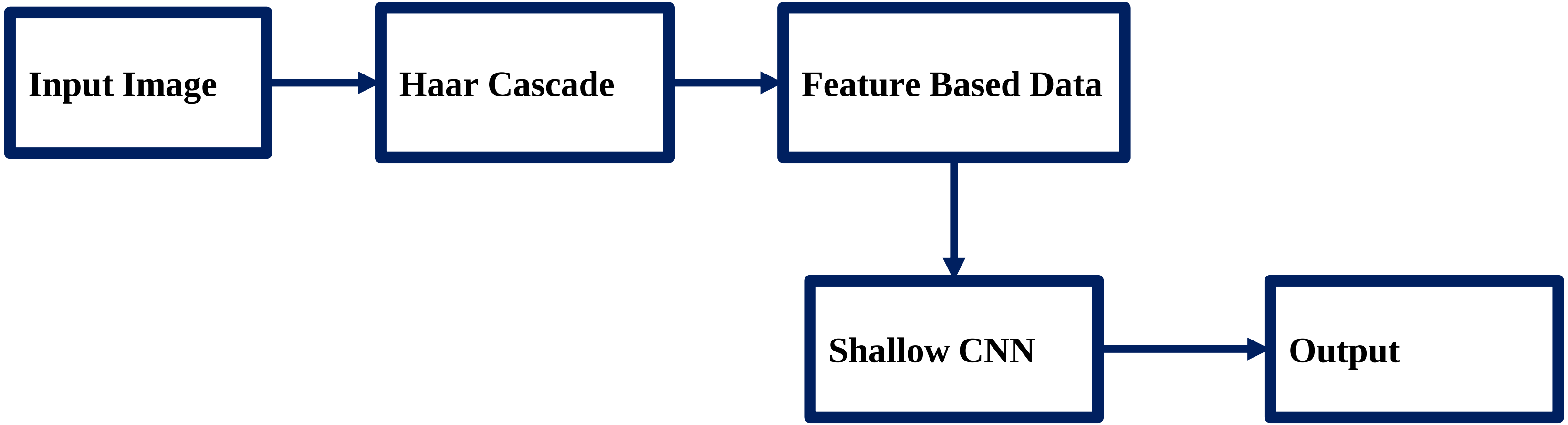

Figure 1 : *Flowchart for Proposed Approach*

Dataset Description

- The dataset which is used for **Facial Emotion Recognition** is **JAFFE dataset.**
- It consisted of **213 images** of **10 different Japanese female** posing for 7 different emotions: **Happiness, Sadness, Surprise, Anger, Disgust, Neutral and Fear**
- Data Split **70:30**
- **Feature Based** Dataset[2]

Figure 2 : *Different Emotions in JAFFE Dataset*

Proposed Work **(Initial Phase)**

Data	CNN	Epochs	Test Accuracy
Raw Input	1 Layer	1000	22.27%
Feature Based Data	1 Layer	1000	62.39%
Augmented Data(Raw)	1 Layer	1000	50.32%

Table 1: *Test Accuracy For different Dataset input*

Proposed **Shallow CNN** Architecture

- **Input Shape:** 100x100
- **1st Convolution Layer:** Kernel Size: 3x3, No of Filter: 128
- **2nd Convolution Layer:** Kernel Size: 3x3, No of Filters: 64
- **Max Pooling:** 2x2
- **3rd Convolution Layer:** Kernel Size: 3x3, No of Filter: 64
- **Max Pooling:** 2x2
- **Fully Connected Layer**: 512
- **Output:** 7 Classes

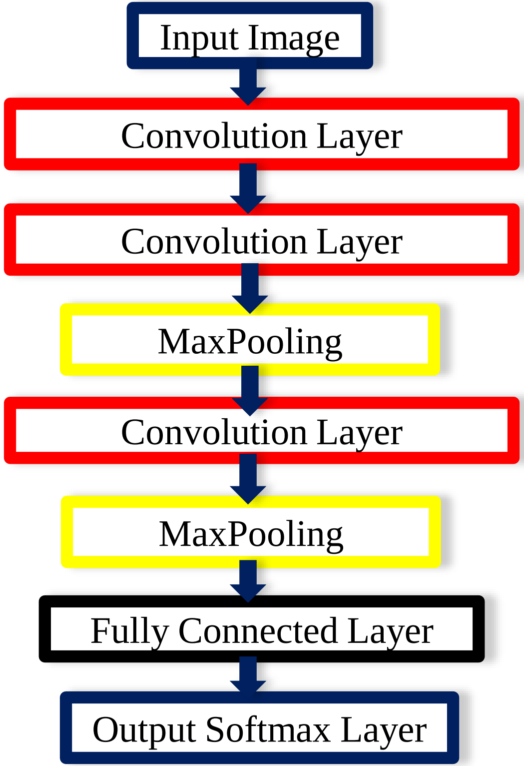

Figure 3: *Proposed CNN Architecture*

Results

Hyper- parameter Tuning

Layer	Epoch	Batchsize	Test accuracy
2	1000	20	75.09%
3	1000	20	85.5%
4	1000	20	71.52%
5	1000	20	70.09%

Table 2: *Hyper Parameter Tuning*

Results

Hyper- Parameter tuning (Continued)

No of Layers	No of Filters (Each Layer)	Batch Size	Epochs	Test Accuracy(%)
3	**128:64:64**	**20**	**3000**	**91.93**
3	128:64:64	20	2000	90.30
3	64:32:32	20	2000	87.09
3	256:128:128	20	2000	88.70
3	512:256:256	20	2000	88.70
4	128:128:64:64	20	2000	87.09

Table 3: *Hyper Parameter Tuning (Continued)*

Results: Comparison

Confusion Matrix

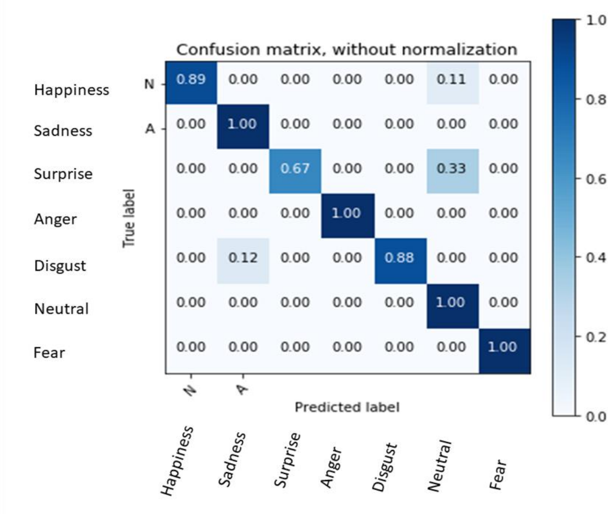

__Figure 4__: Proposed Architecture Confusion Matrix

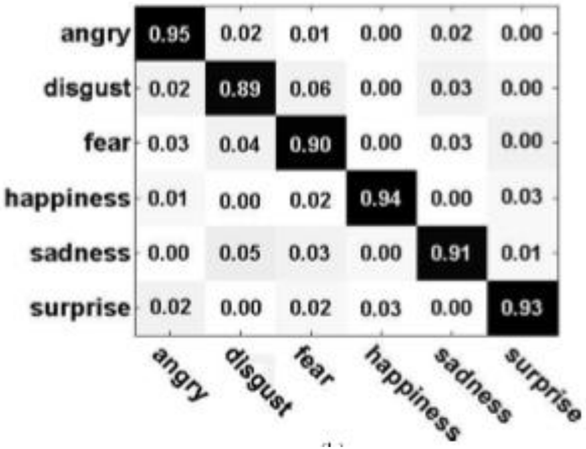

***Figure 5:** Benchmark Paper Architecture Confusion Matrix*

Classification Report

Class	Precision	Recall	F1-Score
0	0.89	1.00	0.94
1	1.00	0.90	0.95
2	0.67	1.00	0.80
3	1.00	1.00	1.00
4	0.88	1.00	0.93
5	1.00	0.69	0.82
6	1.00	1.00	1.00

Table 4: *Classification Report*

Comparison with existing approaches

Existing Approaches	Testing Accuracy(%)
Aly *et al* [3]	87.32
Rivera *et al* [4]	88.75
Zhang *et al* [5]	91.48
Yang *et al* [1]	92.21
Prudhvi *et al* [2]	86.38
Proposed Architecture	**91.93**

Table 5 : Comparison of Our proposed approach with Existing Approaches

Approaches	Learnable Parameters
Yang *et al* [1]	18,556,203
Prudhvi *et al* [2]	79,195
Proposed Architecture	**163,687**

Table 6: No of Learnable Parameter compared with proposed architecture and existing architecture

Conclusion

- Our proposed architecture was able to achieved state of art accuracy of 91.93% which is comparable with benchmark accuracy of 92.21 % and better than base paper 86.38 %.
- Our proposed architecture was able to reduce the computational complexity compared to benchmark paper by reducing the number of convolution layer and learnable parameter still gives comparable performance with the existing approaches for Facial Emotion Recognition.

Reference

- Biao Yank, Jinmeng, Rongrong Ni, YuYu Zhang “**Facial Expression Recognition using Weighted Mixture Deep Neural Network based on Double-Channel Facial Images**”, IEEE Access, 2018.
- Prudhvi Raj, “**Facial Emotion Detection using Convolutional Neural Networks and Representation Autoencoders Units**”, IEEE,2017
- Sherin Aly, A. Lynn Abbott, Marwan Torki “ **A Multi-modal Feature Fusion Framework for Kinect-Based Facial Expression Recognition using Dual Kernel Discriminant Analysis**”,IEEE, 2014
- Adin Ramierz Rivera, “**Local Directional Number Pattern for Face Analysis:Face and Expression Recognition**”, IEEE, 2015
- Wel Zhang,Youmei Zhang, Lin Ma “ **Multimodal learning for facial expression recognition**”, Elsevier,2015.

Thank You