FACIAL EXPRESSION RECOGNITION BY CONVOLUTIONAL NEURAL NETWORKS

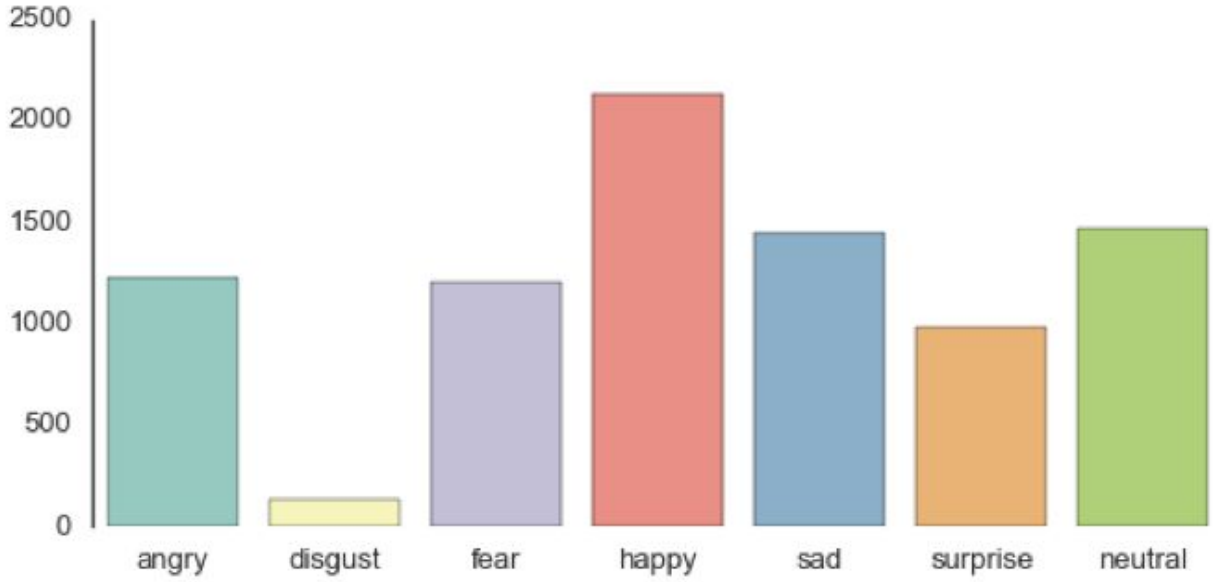
BY
KARANSINGH PUROHIT
linkedin.com/in/KaranPurohit
karansingh.p@somaiya.edu

1.Motivation

Human facial expressions can be easily classified into 7 basic emotions: happy, sad, surprise, fear, anger, disgust, and neutral. Our facial emotions are expressed through activation of specific sets of facial muscles. These sometimes subtle, yet complex, signals in an expression often contain an abundant amount of information about our state of mind. Through facial emotion recognition, we are able to measure the effects that content and services have on the audience/users through an easy and low-cost procedure. For example, retailers may use these metrics to evaluate customer interest. Healthcare providers can provide better service by using additional information about patients' emotional state during treatment. Entertainment producers can monitor audience engagement in events to consistently create desired content.

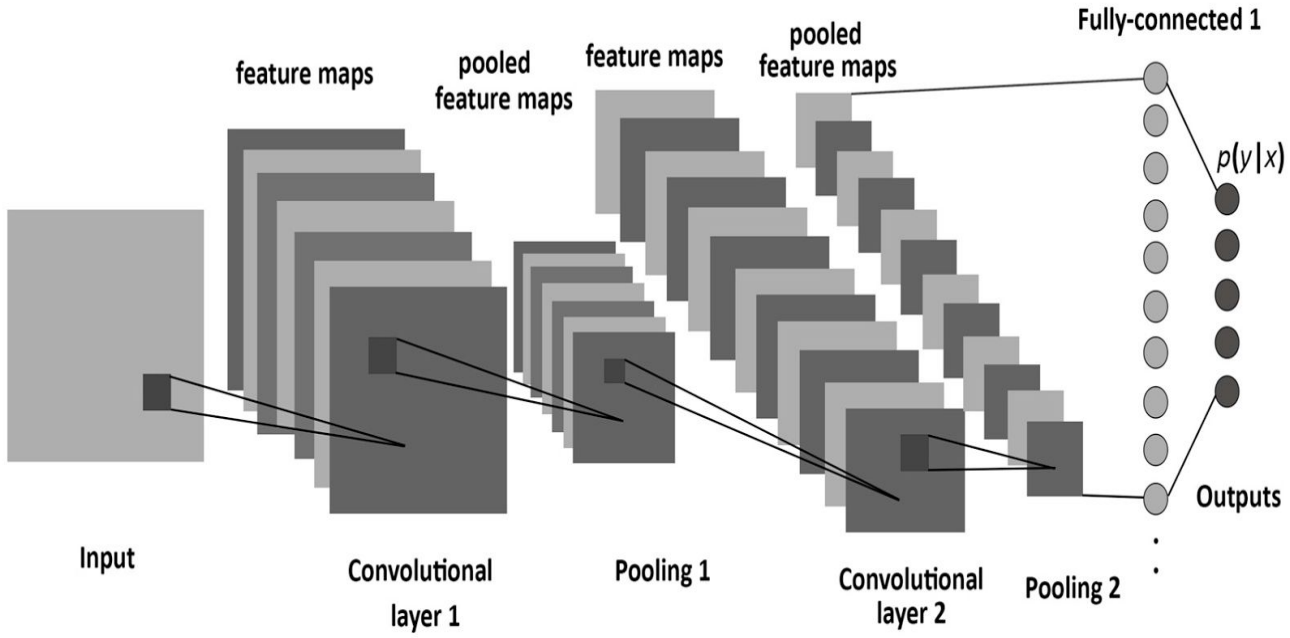
2. The Database

The dataset I used for training the model is from a Kaggle Facial Expression Recognition Challenge a few years back (FER2013). It comprises a total of 35887 pre-cropped, 48-by-48-pixel grayscale images of faces each labeled with one of the 7 emotion classes: anger, disgust, fear, happiness, sadness, surprise, and neutral.



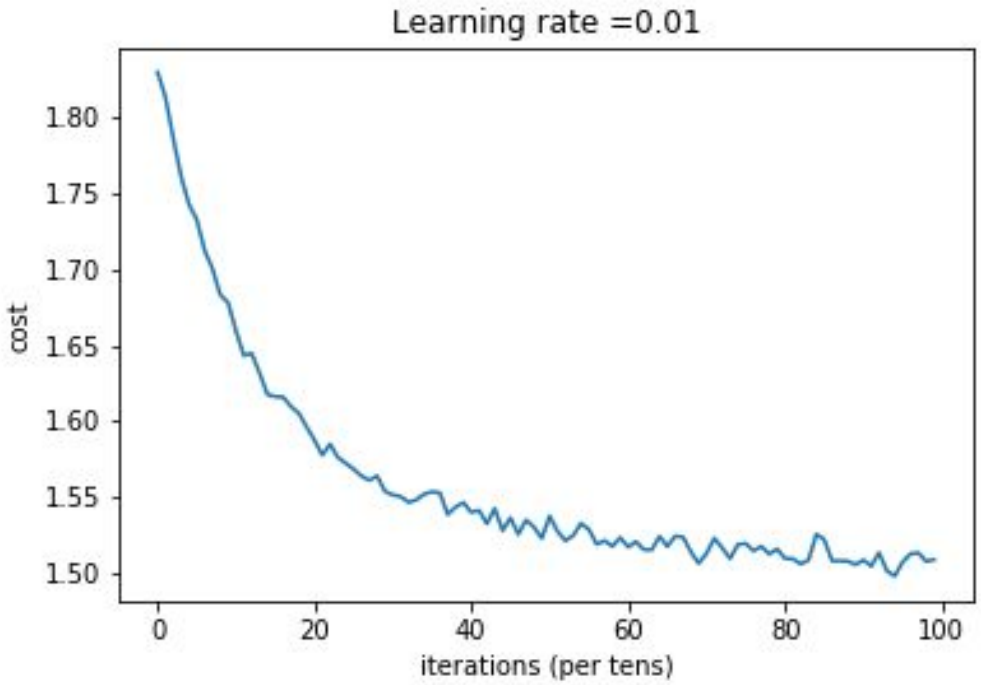
3. The Model

Deep learning is a popular technique used in computer vision. I chose convolutional neural network (CNN) layers as building blocks to create my model architecture. CNNs are known to imitate how the human brain works when analyzing visuals.A typical architecture of a convolutional neural network will contain an input layer, some convolutional layers, some dense layers and an output layer.



Before CNN I used simple neural network which gave very poor accuracy of 26.8%
In my convolution network there are 2 convolutional layers as well as 2 max pooling layers as shown in figure above.

	Activation shape	Activation size
Input	(48,48,1)	2304
CONV1(f=4,s=1)	(48,48,8)	18,432
POOL1	(6,6,8)	288
CONV2(f=2,s=1)	(6,6,16)	576
POOL2	(2,2,16)	64
Dense layer	(64,1)	64
Softmax	(7,1)	7



Train Accuracy: 45.95%
Test Accuracy: 42%

4.Future Work

Accuracy of network can further be increase by a adding more convolutional layers and dense layers. Due to lack of computing power I only used 6000 images. After utilising whole dataset (36000 images) state of the art accuracy can be achieved.

5.Conclusion

For image classification Convolutional neural networks are very effective. Simple neural networks performs very poor in this case. Images which are not properly alligned are not easy to recognise.

6. References

1. *"Dataset: Facial Emotion Recognition (FER2013)"* ICML 2013 Workshop in Challenges in Representation Learning, June 21 in Atlanta, GA.
2. *"Andrej Karpathy's Convolutional Neural Networks (CNNs / ConvNets)"* Convolutional Neural Networks for Visual Recognition (CS231n), Stanford University.
3. Coursera deeplearning specialisation by Andrew Ng
4. Duncan, D., Shine, G., English, C., 2016. *"Report: Facial Emotion Recognition in Real-time"* Convolutional Neural Networks for Visual Recognition (CS231n), Stanford University.