Mood Detection and Management using Facial Expression

Fatema Hasta, Maniraj Mohareer, Sriram Mentay and Sunil Madikanti
University of Missouri-Kansas City, Missouri, USA
{fahvdx, mm72d, sjmd3n, smgv6}@umkc.edu

Abstract— The core idea of our project is to classify a person's mood using his facial expression and provide certain services based on the mood. The features include suggesting and playing a song as per the mood, suggesting movies and management tips to improve the mood. We have designed a web application that takes the input in the form of an image. This image can be fed to our model that detects the emotion and classifies the image either into happy, sad, anger, disgust or surprise. The moods of the person is tracked every day and an analysis chart could be given at the end of particular month. We trained our model using Shallow learning techniqies (Spark), Deep learning techniques (Softmax, CNN and Inception) and Clarifai Api and we noted down the accuracy values for each of them. Inception model gives us the best value for our design.

I. INTRODUCTION

Let's start with the basic features of our model. Firstly, they can analyze the mood of the person with the facial expression. The model can detect five different moods of the person. The classes can be divided based on the emotions happiness, sadness, anger, surprise and disgust. Based on the facial expression the model predicts the emotion that would be useful further in the application. Secondly, it gives some suggestions based on the current emotion like songs, movies etc. On the user's profile in the website, he/she can have access to a lot of information that is sorted in detail. They include: Suggesting movies and playing them based on the current emotion of the user, redirecting to music and suggest and play songs from playlists based on a particular mood, Emotion Management tips. Thirdly, it provides a monthly analysis of the emotions to the user using charts and Graphs.

Some of the system features were:

- Image as an Input:
  - We can upload the image as the input to the application.
  - On analyzing the image, the emotions are predicted and the conclusion is omitted out.
  - Most precise conclusion
  -
- Output:
  - An emotion is given as an output along with music, movie suggestions.
  - A graphical chart that gives the user an idea of how to manage his/her mood.

II. RELATED WORK

We did a lot of ground work and what we concluded that,

- There is no project of this kind so far. Our application gives an all-in-one package with music and movie suggestions which helps the user to graphically view his monthly emotion chart and manage it.
- According to our study on various similar projects, they offer similar things differently but not altogether.
- We want to make this completely functional and achieve all the objectives and make this one-of-a-kind.

III. APPROACH

- Data Sources:
  - We have taken images of faces from http://archive.ics.uci.edu/ml/machine-learning-databases/faces-mld/
  - We have classified the dataset into 5 classes: 'angry', 'happy', 'disgust', 'sad' and 'surprise'
- Analytic Algorithms:
  - For analysis of our model we have used:
    - Spark (Shallow Learning)
      - K-Means clustering and Random Forest classification.
    - Convolutional Neural Network (CNN)
    - Softmax Classification
    - Inception
    - Clarifai Api
  - Training and testing of data was done on algorithm to get the best one that can precisely classify our training data and give us a concise and accurate output.
- Analytical tools:
  - Apache Spark
  - Tensorboard
- IDE:
  - Intellij
  - Pycharm
- Expected Input/Outputs:
  - Input: Input image of facial expression of a person.
  - Output: The expression of the person and suggestions for Movies and Music.

- Evaluation and Validation:
    - Clarifai API
        - Accuracy: 86.2%
    - Shallow Learning (Spark)
        - Accuracy: 63.6%
    - Deep Learning (CNN)
        - Accuracy: 72.4%
    - Deep Learning (Inception)
        - Accuracy: 92.7%
    - Deep Learning (Softmax)
        - Accuracy: 69.3%

IV. IMPLEMENTATION

- System Specification:
    - MacOS Sierra+/ Windows 7 + Ubuntu 2.3 GHz Intel Core i5 and above Minimum 8 GB 2133 MHz LPDDR3 and above

- Software Architecture

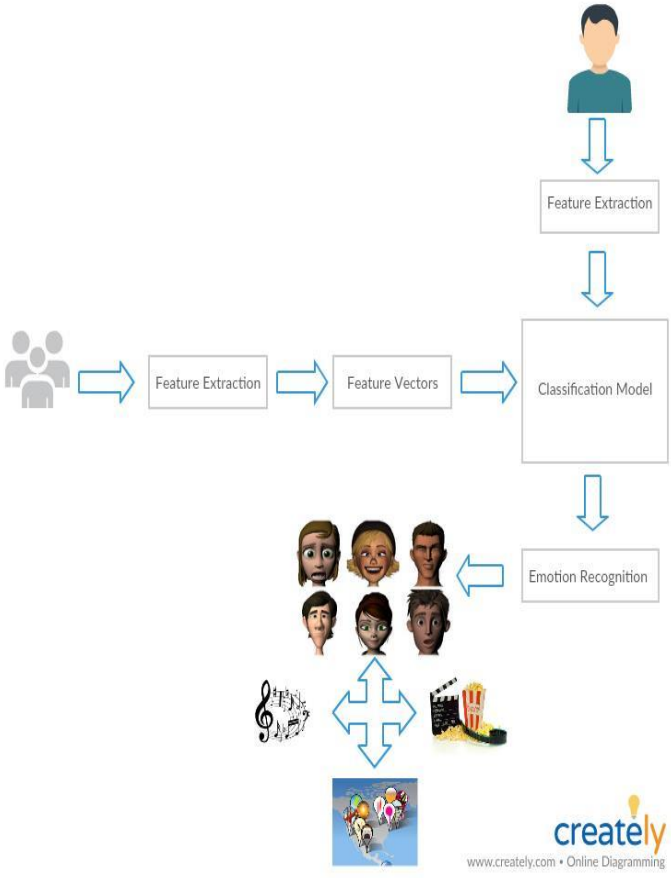

Fig. 1. The entire software architecture of our system with all components

- Sequence Diagram

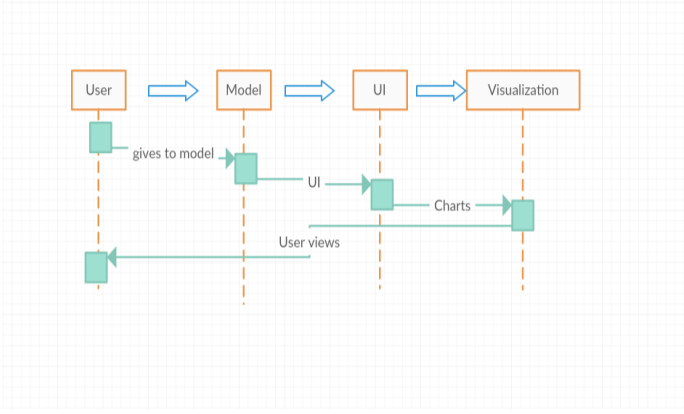

Fig. 2. The sequence diagram of our system

- Activity Diagram

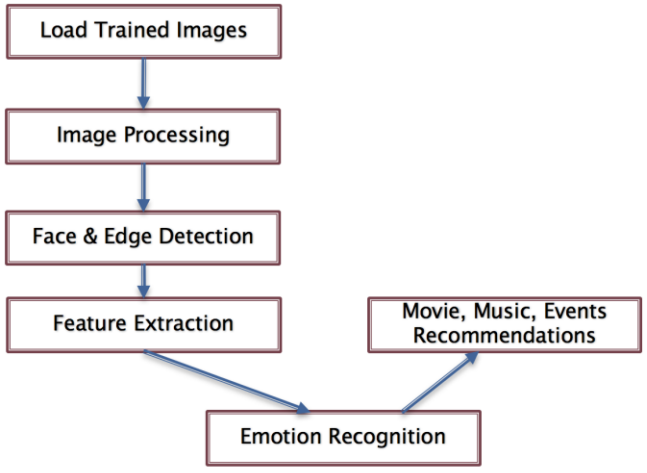

Fig. 3. The activity diagram of our system

- Feature Specification
    - Image as an Input:
        - We can upload the image as the input to the application.
        - On analyzing the image, the emotions are predicted and the conclusion is omitted out.
        - Most precise conclusion
    - Output:
        - An emotion is given as an output along with music, movie suggestions.
        - A graphical chart that gives the user an idea of how to manage his/her mood

- Existing Applications/Services Used:
    - Kairos – Offers a wide variety of image recognition solutions through their API. Their API endpoints include identifying gender, age, emotional depth, facial recognition in both photo and video, and more.
    - Trueface.ai – One flaw with some facial recognition APIs is that they are unable to differentiate between a face and a picture of a face. TrueFace.ai solves that problem with their ability to do spoof detection through their API.
    - Amazon Recognition – This facial recognition API is fully integrated into the Amazon Web Service ecosystem. Using this API will make it really easy to build applications that make use of other AWS products.
    - Face Recognition and Face Detection by Lambda Labs – With over 1,000 calls per month in the free pricing tier, and only $0.0024 per extra API call, this API is a really affordable option for developers wanting to use a facial recognition API.
    - EmoVu by Eyeris – This API was created by Eyeris and it is a deep learning-based emotion recognition API. EmoVu allows for great emotion recognition results by identifying facial micro-expressions in real-time.
    - Microsoft Face API – One cool feature that I found while doing research on the Microsoft Face API, is that the API has the ability to do "similar face search." When this API endpoint is given a collection of faces, and a new face as a query, the API will return a collection of similar faces from the collection.
    - Animetrics Face Recognition – Using advanced 2D-to-3D algorithms, this API will convert a 2D image into a 3D model. The 3D model will then be used for facial recognition purposes.
    - Face++ – This API also has an offline SDK for iOS & Android for you to use. The offline SDK does not provide face recognition, but it can perform face detection, comparing, tracking and landmarks, all while the phone does not have cell service.
    - Google Cloud Vision – By being integrated into the Google Cloud Platform, this API will be a breeze for you to integrate into applications that are already using other Google Cloud Platform products and services.

- Implementation Details using:
    - Shallow Learning using Random Forest Classification: We have used Random Forest classification to predict our images facial expression with an accuracy value of 63.6%
    - Clarifai API: We are using Clarifai API to predict the facial expressions of the images in our data set using their training model. This gives us a good accuracy value of 86.2%.
    - Softmax Classification: We used the Softmax Function to classify our testing images with an accuracy value of 69.3%
    - Convolutional Neural Network (CNN): Trained Model using CNN to get an accuracy value of 72.4%
    - Inception model: Compared to all the models, the inception model gives the best results with an accuracy value of 92.7% and helps us classify all types of images correctly.

- Project Management:
    - Work Completed:
        - Application using Clarifai API
        - Shallow learning using Random Forest approach
        - Softmax Classification
        - Convolutional Neural Network (CNN)
        - Inception model
        - Web Application using Spotify API
    - Zenhub Issues

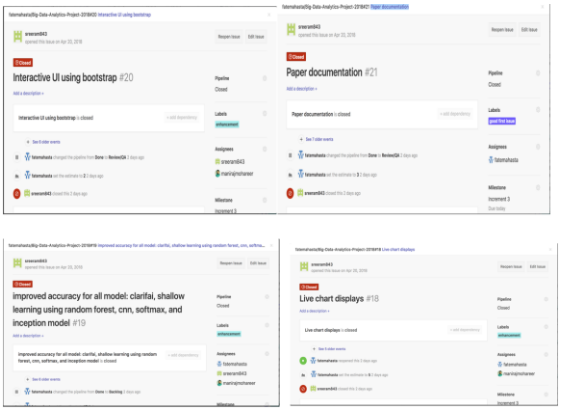

Fig. 4. Zenhub Issues

- Zenhub Board

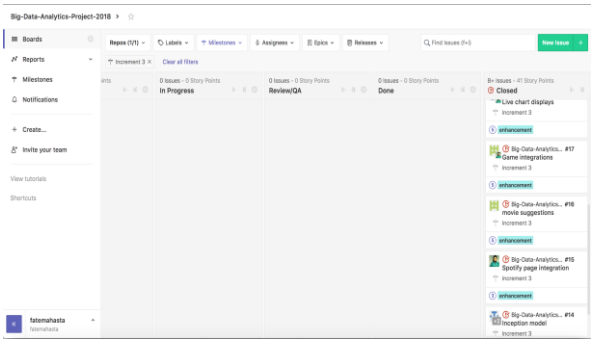

Fig. 5. Zenhub Board

- Zenhub Milestones

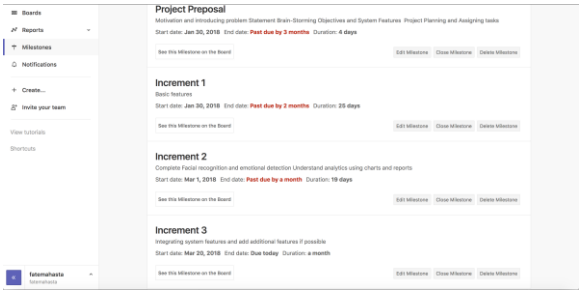

Fig. 6. Zenhub Milestones

- Burndown Chart

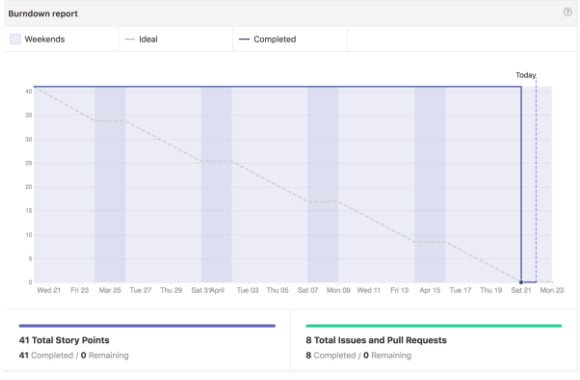

Fig. 7. Burndown Chart

- Time Taken: Each of us worked for around 12-15 hours each.
- Contributors: All of us worked towards this aim equally. No task is achieved by only one. But, we split the work equally overall.

V. RESULTS

- Data Set: The data set used is 5 classes of different emotional expressions. The classes are images of 5 different categories as follows: "angry", "disgust", "happy", "sad", "surprise". Each class contains about 300 images.
- Evaluation Results and Tensorboard
  - Clarifai Model

```
1.0 : (0,67.09511568123393);
(1,32.904884318766065);
(2,0.0)
(0.0,2.0)
(0.0,0.0)
(0.0,1.0)
Accuracy:0.8623065318738433
Confusion Matrix:
1.0  0.0  0.0
1.0  0.0  0.0
1.0  0.0  0.0

Process finished with exit code 0
```

Fig. 8. 86.2% Accuracy of Clarifai model

  - Shallow Learning using Random Forest Classification

```
 |==================== Confusion matrix ==========================
7.0  0.0   1.0  0.0  0.0  1.0  1.0  0.0  0.0  0.0
0.0  11.0  1.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
0.0  0.0   5.0  2.0  1.0  1.0  2.0  0.0  0.0  2.0
1.0  0.0   0.0  9.0  1.0  0.0  0.0  0.0  0.0  0.0
0.0  0.0   0.0  0.0  6.0  0.0  1.0  0.0  0.0  0.0
0.0  1.0   1.0  3.0  1.0  8.0  1.0  1.0  0.0  2.0
0.0  0.0   0.0  0.0  0.0  0.0  3.0  0.0  0.0  0.0
0.0  0.0   0.0  0.0  1.0  0.0  2.0  7.0  0.0  0.0
2.0  3.0   5.0  3.0  2.0  0.0  0.0  0.0  6.0  1.0
0.0  1.0   0.0  0.0  0.0  0.0  0.0  0.0  0.0  4.0
0.6365945945945946
numTrees 4 featureSubsetStrategy all impurity entropy maxDepth 6
```

Fig. 9. 63.6% Accuracy of Shallow Learning model

  - Softmax Classification

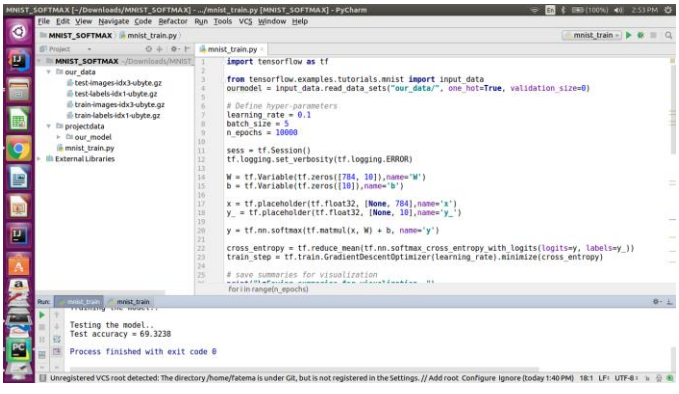

Fig. 10. 69.3% Accuracy of Softmax Classification

- Convolutional Neural Network (CNN)

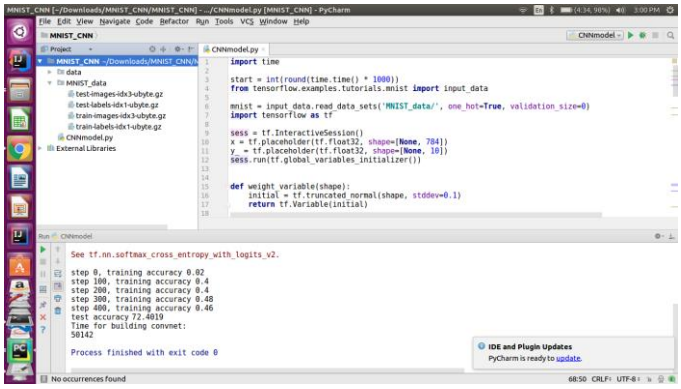

Fig. 11. 72.4% Accuracy of CNN model

- Inception model

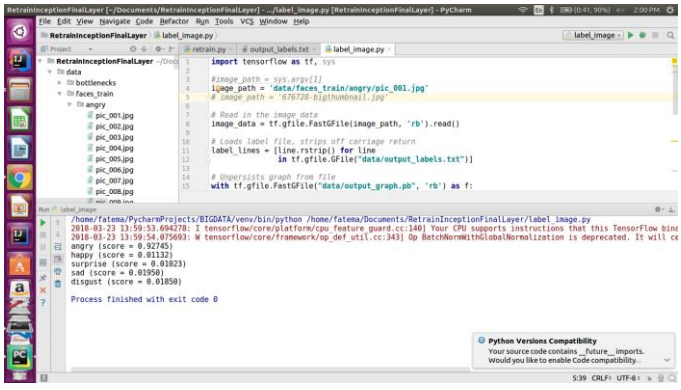

Fig. 12. 92.7% Accuracy of CNN model

- Web Application

Fig. 13. Dashboard

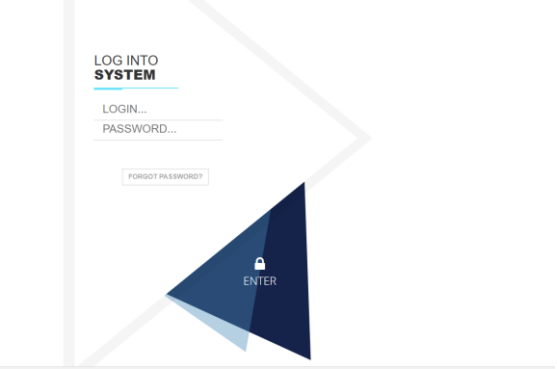

Fig. 14. Login Page

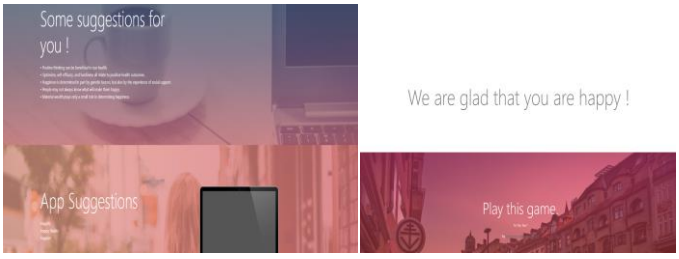

Fig. 15. Happiness emotion suggestions

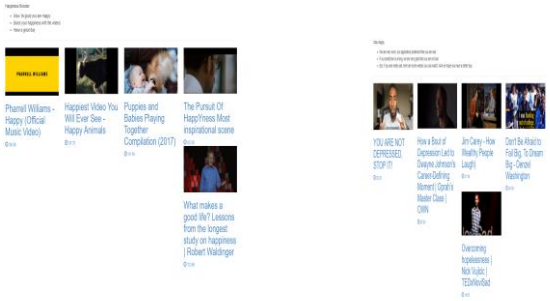

Fig. 16. Video suggestions based on emotion

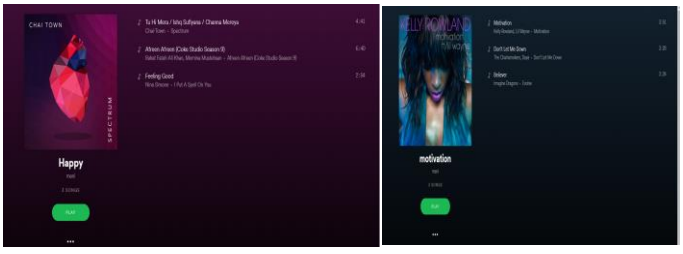

Fig. 17. Music suggestions based on emotion

VI. COMPARISON TABLE

Accuracy Comparison Table

Model	Accuracy	Image Used Happy	Image Used Sad	Image Used Angry
Shallow Learning	63.6%			
Clarifai	86.2%			
CNN	72.4%			
Softmax	69.3%			
Inception	92.7%			

VII. CONCLUSION

In this paper, we have described the experiences on developing a facial expression recognition web application. Our dataset was trained on multiple networks and we found different results for different networks. After our comparisons on different network, we found that the Inception model gives us the best and accurate result.

As future work, we plan to perform real time robust facial expression recognition on a smartphone.

VIII. REFERENCES

[1] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, pp. 1-127, 2009.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *NIPS*, 2012.

[3] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *AISTATS*, 2010.

[4] D. Cireşan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," *CVPR*, 2012.

[5] https://code.google.com/p/cuda-convnet

[6] T. Ojala and M. Pietikäinen, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *TPAMI*, vol. 24, no. 7, pp. 971-987, 2002.

[7] D. Erhan, P.-A. Manzagol, Y. Bengio, S. Bengio, and P. Vincent, "The difficulty of training deep architectures and the effect of unsupervised pre-training," *AISTATS*, 2009.