JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

BEYOND LIKES: SA MULTI MODEL MUSIC RECOMMENDATION SYSTEM USING FACIAL RECOGNITION AND HAND GESTURES

Mrs. K. Uma
Professor department of AIML

E. Robin
B. Tech Student
AIML
Siddhartha Institute of Technology and sciences

A. Shashank kumar
B. Tech Student
AIML
Siddhartha Institute of Technology and science

C. Bharath Raj
B. Tech Student
AIML
Siddhartha Institute of Technology and sciences

M.Satya Prakash
B. Tech Student
AIML
Siddhartha Institute of Technology and science

J. Apur Patel
B. Tech Student
AIML
Siddhartha Institute of Technology and sciences

ABSTRACT

Now-a-days people are crazy about music, people used to listen music in different aspects of life. Most of the people listens to music to cherish their mood and to keep themselves motivated. Good Music connects us with the environment and boosts our mood. Every individual wants to listen music according to his/her state of emotion at that time. Listening music according to his/her state of emotion will really gives peace to mind and soul. This can be possible only with the facial expressions of the individuals. Facial expressions denote an individual state of emotion at that time. In Proposed System, Person's facial expressions can be captured by the camera and by applying Machine Learning techniques it analyzes person's expression can suggests a playlist. One such technique is MobileNet model with Keras. Abstract

Keywords: Face Recognition, Image Processing, Computer Vision, Emotion Detection.

INTRODUCTION

- Human emotions are classified as fear, disgust, rage, surprise, sad, glad, or neutral. This emotional umbrella includes a wide range of different emotions, such as cheerful (a modification scorn (a modification of disgust). These are mild feelings.
- Facial muscle contortions are quite tiny, and recognizing them can be challenging because even minor variations result in distinct expressions. Furthermore, because emotions are largely context dependent, expressions of different or even the same persons for the same feeling may differ. Also, because emotions are very context dependent, expressions of various or even the same people may change for the same feeling.
- While the focus may be on the areas of the face that convey the greatest emotion, such as around the mouth and eyes, how these motions are extracted and classified remains an important topic. These tasks have been completed successfully using neural networks and machine learning. Machine learning techniques have been quite useful in pattern identification and classification, and can thus be used to detect mood.
- With the evolution of technology in digital music, the creation of a personalised music recommendation system that offers music to consumers is crucial. Making recommendations based on the large amount of data available on the internet is a significant issue. Companies such as Amazon and eBay provide customized suggestions to individuals based on their tastes and history, whilst Spotify and Pandora provide relevant recommendations using Machine Learning and Deep Learning approaches. Personalised music recommendation can be approached in two ways. Content-based filtering is one way, which evaluates the content of music that people have previously liked and recommends music with relevant content.

- The model is able to generate recommendations according to the user's existing interests, which is the biggest downside of this strategy. In brief , the model's potential to build on existing user interests is limited. The collaborative filtering strategy, on the other hand, suggests music that was loved by a peer group with same preferences. Both recommendation systems depends on the user's listening behavior and preferences. The main disadvantage of this strategy is the popularity bias problem, which occurs when popular (i.e., often rated) items receive a lot of attention while less popular items are under-represented in the suggestions. In general, a hybrid method is adopted, which combines both content and collaborative procedures to maximise accuracy and mitigate the limitations of both types. The goal of this project is to develop a music recommendation system that can recognise a user's face, detect their facial emotion, and then offer a playlist according to the detected emotion.
-

SCOPE AND MOTIVATION

The scope and motivation for music recommendations through facial recognition and hand gestures are quite broad and exciting. Here are some key points:

Scope:

The scope and motivation for music recommendations through facial recognition and hand gestures are quite broad and exciting. Here are some key points:

1. Personalized Music Experience: This technology can analyze a user's facial expressions in real-time and recommend music that best matches their current emotional state[1]. It aims to personalize music recommendations based on the listener's facial expression[1].

2. Applications Beyond Music: This technology can also have practical applications in other fields such as healthcare, education, and entertainment[1].

3.Emotion Recognition: The system uses facial expression analysis to understand the listener's emotional state and recommend music that aligns with that emotion[1].

Motivation:

1. Emotion-Based Recommendations: The relationship between music and human emotions has long been recognized, with music having the ability to evoke various emotional states. A facial emotion-based music recommendation system is an innovative solution that aims to personalize music recommendations based on the listener's facial expression.

2. Efficiency: Manually classifying songs based on a user's emotional preference is a laborious task. This system removes the laborious and time-consuming chore of manually sorting music into distinct lists.

3. Revolutionizing Interaction with Music: The increasing availability of high-quality facial recognition technology and the ever-growing music streaming services make this technology a promising solution for personalized music recommendations. This system has the potential to revolutionize the way we interact with music, making it more personalized and emotionally engaging.

Literature Survey

2. METHODOLOGY

a. Proposed System

- This system comprises 1 major module.
- The user needs to complete the registration first to access the system. If the user forgets their password, the system will send an OTP via email and they can reset the password. They can sign in using their credentials. They profile can be managed and the password can be changed if they want. The user can choose an emoji and the system will play the song according to the chosen emoji. The system can also capture the user’s facial emotions through the camera and play a song according to the detected emotion.
- The face emotions detection will be done using Mobilenet and Keras or either OpenCV. For both cases, data is trained and the model is created. The facial emotions are detected using Google Vision and Machine Learning algorithms. And the songs are uploaded to the Firebase storage and Firebase real-time database will be used. The song links will be accessed from there.
- For this project, XML is used on the front end and the backend is developed through Firebase. The programming language is Java. The IDE used is Android Studio.

The scope of the music recommendation system based on users’ facial emotion is to provide personalized music recommendations to users according to their current emotional state as detected through facial recognition technology. The goal of this system is to enhance the user's music listening experience by playing music that corresponds to their present emotional state. The technology can help users regulate their moods and improve their overall well-being by recommending music that matches their feelings. Additionally, the system can also be used to introduce users to new music that they may not have otherwise discovered.

b. Design System Architecture

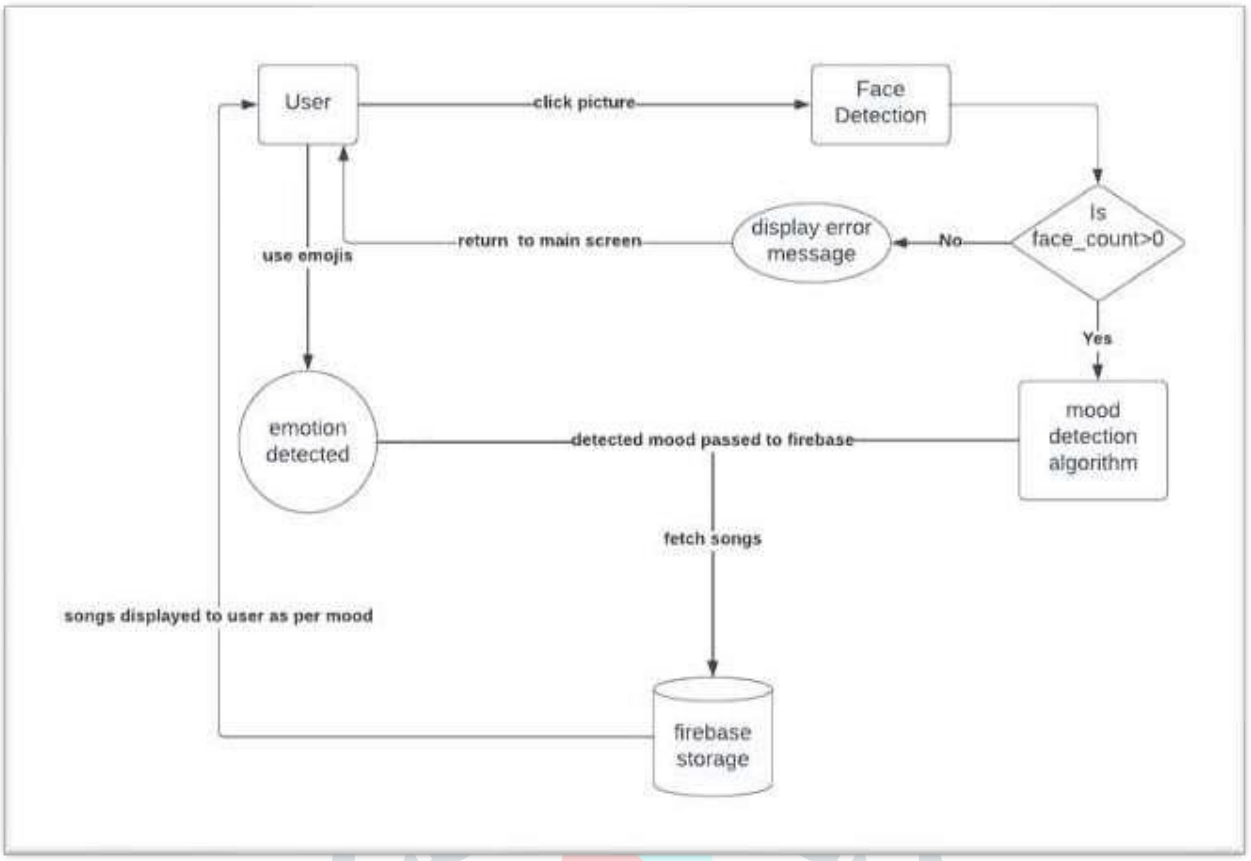


Fig 1: System Architecture

Data Flow Diagram

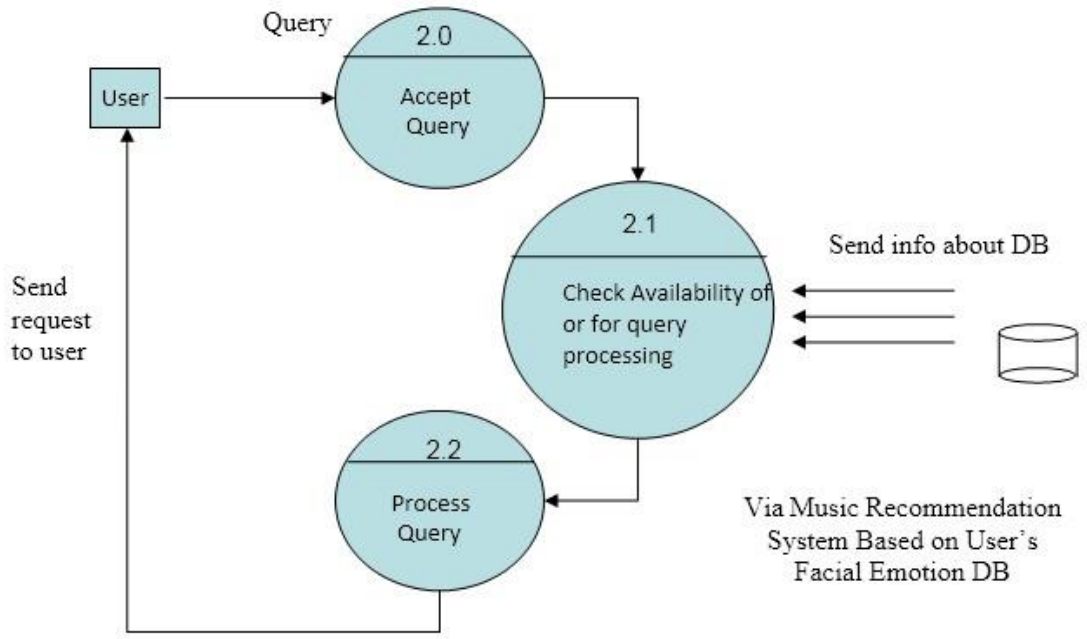


Fig 2: Data Flow Diagram

Use Case Diagram

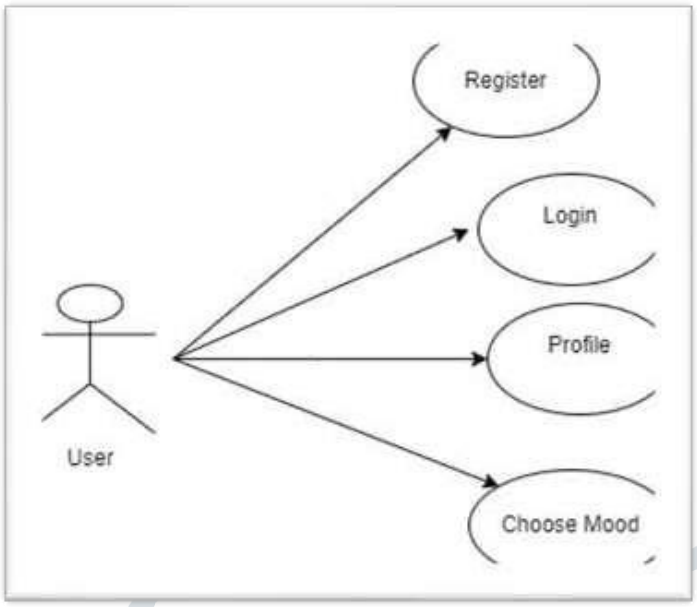


Fig 3: Use Case Diagram

c. Modules and their Description

The system is made up of two major modules and their sub-modules, which are as follows:

❖ **User:**

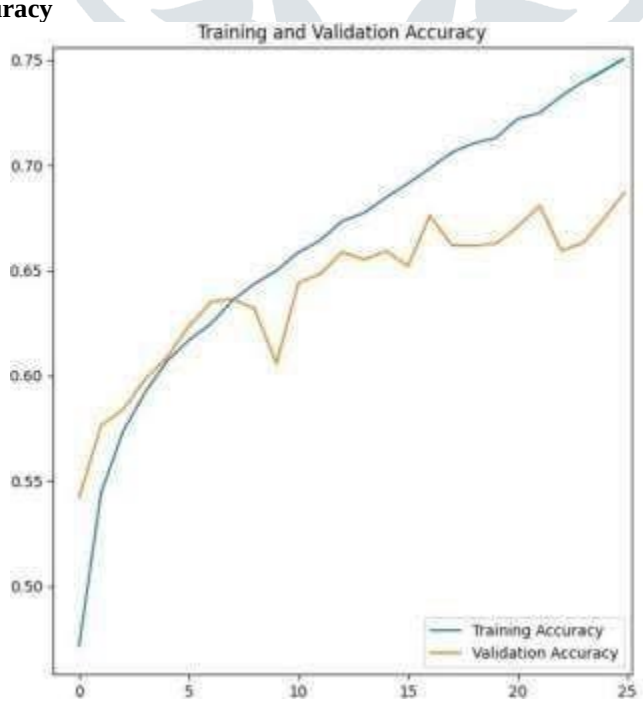
In the user module, the user needs to complete the registration using the credentials. If the user forgets their password, an OTP will be sent through email to reset the password. The user can log in using their credentials. The user also has an option to manage their profile. The user can change their password if they want.

❖ **Choose Mood:**

❖

After the user successfully logins, the user has option to choose either emoji option or face capture. If the user chooses emoji option, songs will be played based on that emoji. If the user choose face capture option, the system will identify the emotion of the user captured by the camera. Accordingly, the system plays songs based on their detected emotion.

Training and Validation Accuracy

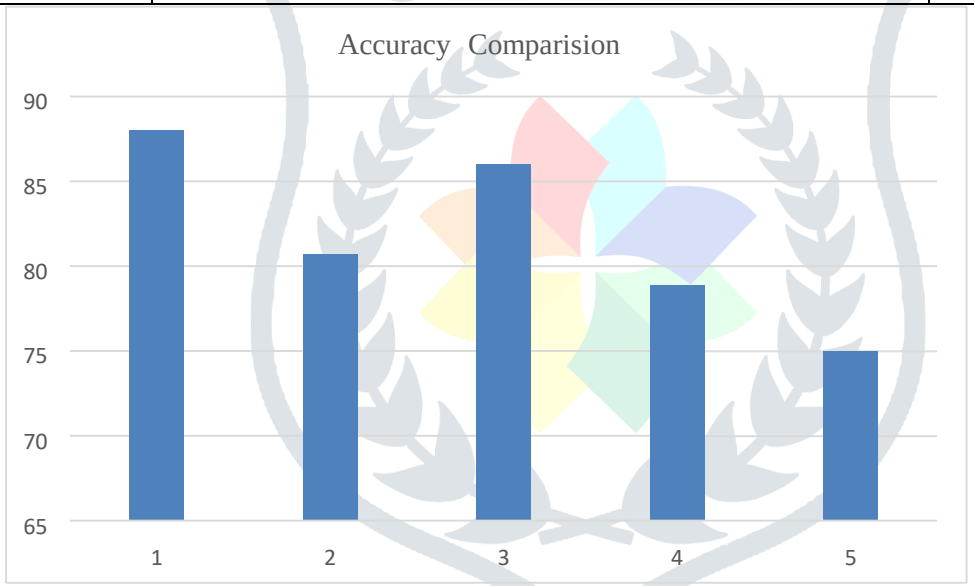


The graph depicts our model's accuracy, with the X-axis indicating the number of iterations and the Y-axis indicating the accuracy. Our model obtained roughly 75% accuracy, as seen in the picture. It understands emotions in the manner in which it has been trained because it is fully computer-based.

Accuracy Comparison

The table below compares the accuracy of various existing systems and tools to our proposed system.

SI.No	Existing Systems and Tools	Accuracy (%)
1	Keras-based face recognition through deep learning	88
2	Music suggestion using a hybrid method	80.7
3	Viola Jones' framework for object detection	86
4	Music.AI	78.84
5	Music Recommendation System based on Facial Emotion	75



The above figure represents graphical representation of the above shown accuracy comparison table. Here X-axis represents Existing systems and tools and Y-axis represents accuracy percentages.

ER Diagram:

- **Purpose**: To describe the structure of the software system, including classes, their attributes, methods, and relationships.
- **Components**: Classes, attributes, methods, associations, and inheritance relationships.
- **Usage**: Class diagrams provide an overview of the system's object-oriented design, representing entities like users, chat data, analysis components, and more.

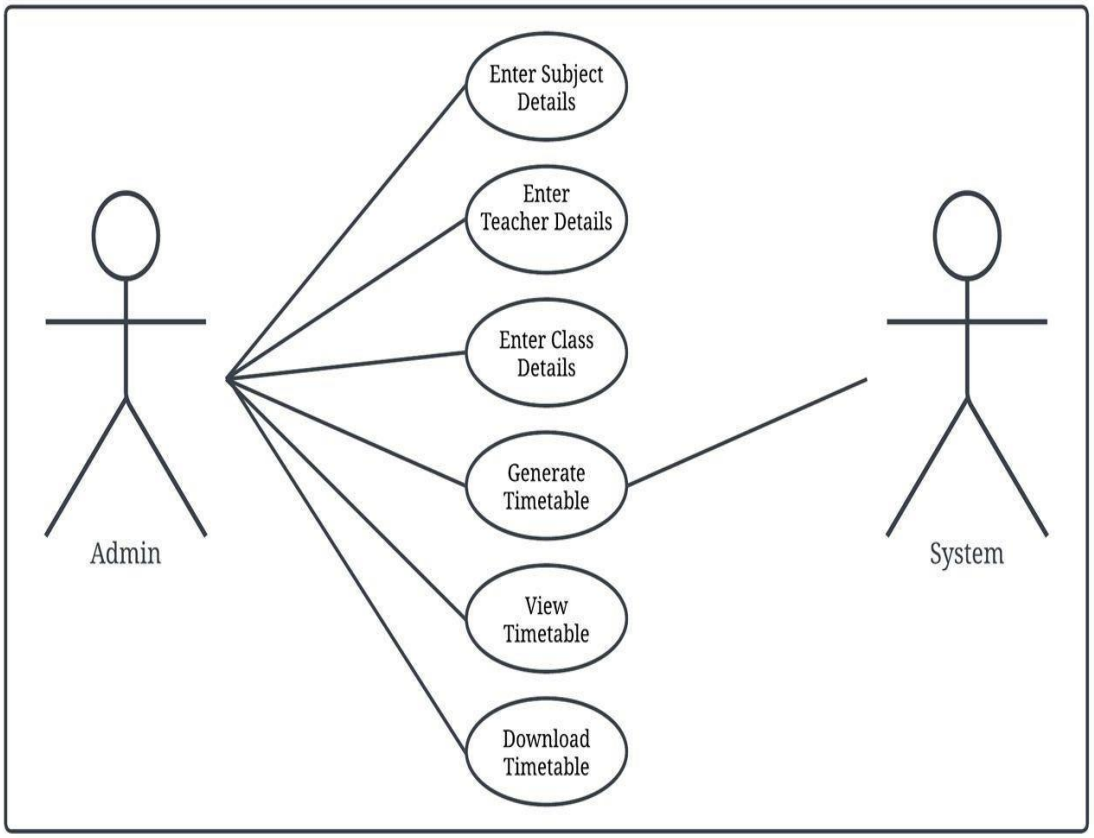


Fig 6 ER diagram

CONCLUSION & FUTURE ENHANCEMENTS

CONCLUSION

We conclude that our interface will work according to user wish and satisfies all user need for music according it the user emotion. It saves time of user. The application can be improved by modifying and adding few functionalities. Current application uses Affective SDK that has a lot of limitations, creating custom emotion recognition system that can be merged into the current application improves functionality and performance of the system. Making the application run without needing an internet connection. Including other emotions.

FUTURE ENHANCEMENTS

In the future, we would like to focus on improving the recognition rate of our system. Also, we would like to develop a mood-enhancing music player in the future that starts with the user's current emotion (which may be sad) and then plays music of positive emotions, thereby eventually giving the user a joyful feeling. The Emotion Based System will be of great advantage to users looking for music based on their mood and emotional behavior. It will help reduce the searching time for music thereby reducing unnecessary time and hence increasing the overall accuracy and efficiency of the system. The system will not only reduce physical stress but will also act as a boon for the music therapy systems and may also assist the music therapist to treat the patient. In future it can also be used to detect the sleepy mood of the driver, driving the car and many more uses. This technology has the potential to enhance the way we interact with music and make it a more personalized and impactful experience.The system can select music that aligns with user’s emotions, creating a more immersive and emotionally resonant music experience.

REFERENCES

JOURNALS

[1] Hafeez Kabani, Sharik Khan, Omar Khan, Shabana Tadvi" Emotion Based Music Player" International Journal of Engineering Research and General Science Volume 3, Issue 1, January-February, 2015.

[2] Shlok Gikla, Husain Zafar, Chintan Soni, Kshitija Waghurdekar" SMART MUSIC INTEGRATING AND MUSIC MOOD RECOMMENDATION"2017 International Conference on Wireless Networking(WiSppNET). Communications, Signal Processing and

[3] Srushti Sawant, Shraddha Patil, Shubhangi Biradar," EMOTION BASED MUSIC SYSTEM", International Journal of Innovations & Advancement in Computer Science, IJIACS ISSN 2347-8616 volume 7, Issue 3 March 2018.

[4] Ekman, P., & Rosenberg, E. L. (2005). What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press.

[5] Dabbish, L., Kraut, R., Fussell, S., & Kiesler, S. (2005). Understanding email use: Predicting action on a message. Proceedings of the SIGCHI conference on Human factors in computing systems, 691-700.

[6] Logan, R. K. (2000). The extended mind: The emergence of language, the human mind, and culture. University of Toronto

Press.
[7] S.L.Happy and A. Routray,"Automatic facial expression recognition using features of salient facial patches,"in IEEE Transactions on Affective Computing, vol. 6, no. 1, pp. 1 12, 1 Jan.-March 2015.
[8] Rahul ravi, S.V Yadhukrishna, Rajalakshmi, Prithviraj, "A Face Expression Recognition Using CNN 2020 IEEE. 35
[9] S.Deebika, K.A.Indira, Jesline, "A Machine Learning Based Music Player by Detecting Emotions",2019 IEEE.

BOOKS

1. “This Is your Brain on Music” by Daniel J.Levitin and “Musicophilia” by Oliver Sacks.
2. “Music and the Mind” by Tom Cochrane.
3. “The Emotional Life of Your Brain” by Richard J. Davidson and Sharon Begley.
4. Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. Emotion, 3(1), 3-22.
5. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. Springer.
6. Shiffman, D. (2012). The Nature of Code: Simulating Natural Systems with Processing. Self-published.

WEB LINKS

1. https://www.tensorflow.org/hub/tutorials/tf2_image_retrainig
2. https://en.wikipedia.org/wiki/Convolutional_neural_network
3.https://opencv.org