WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Emotion Based Intelligent Music Recommendation System Using Faster R CNN Method

[1]Mrs. Padmapriya, [2]Vishaly K, [3]Ilamathi T, [4]Vijayalakshmi R

[1,2,3,4]Sri Manakula Vinayagar Engineering College, Madagadipet, Pondicherry, India

Abstract

Human emotions are dynamic and keep on changing over time, hence observing human emotions and storing values for classification is an important factor. To provide an interface between the music system and also provide very good entertainment for the users this system has been proposed. Humans often use nonverbal cues such as hand gestures, facial expressions, and tone of voice to express feelings in interpersonal communications. The face of the human is an important organ of an individual's body and it plays an important role in the extraction of an individual's behavior and emotional state. Facial expression provides the current mental state of a person. It is very time consuming and difficult to create and manag e large playlists and select songs from these playlists. Thus, we create the music player that itself chooses a song according to the current mood of the user. Here, in this system, we propose to merge both the things FER system and the music recommendation system for the convenience and ease of use of the user who can get songs played according to the mood and can get user-generated playlists. In this system, the main focus will be on the real-time working of the system. The system here will capture real-time images for emotion prediction. These real-time working of the system can lead us to simplicity and effortlessness in life. The predicted emotion will be the input to the system after that the music will be classified according to the predicted emotion to play the songs related to your mood.

Keywords: Facial Expression Recognizer (FER); Music recommendation system; Deep Learning(DL)

1. Introduction

The human inclination is a unique one and continues to change instantly, consequently noticing human inclination and putting away qualities for characterization is a significant element. To give a connection point between the music framework and furthermore give a generally excellent diversion to the clients, to implement the ideas of machine learning it is programmed to search and interpret the data and thus create a playlist based on the rules or the parameters provided, we provide a new age platform for music lovers and to bridge gap between growing technologies and music techniques.

Deep Learning is new field in AI where we train huge, profound Brain Organization models for various undertakings from picture handling to discourse and language displaying. Deep Learning has demonstrated to beat every one of the benchmarks on various AI assignments over shallow brain organizations or shallow SVM models. Already there were endeavors to prepare profound brain network models for various discourse, picture and regular language handling assignments however those endeavors have fizzled for absence of enormous handling abilities and less accessibility of GPUs. Additionally there are issues like the evaporating inclination issues. As of late the term Profound Advancing again gotten some momentum in mid-2000's Geoffrey Hinton and Salakhutdinov demonstrated the way that a many-layered feed-forward brain organization could be successfully pre-prepared each layer in turn treating each layer at a time as a solo Rbn's. Since its resurgence Profound Learning has been the cutting edge in many trains mostly for picture acknowledgment/PC vision and Programmed discourse acknowledgment.

Machine Learning vs. Traditional Programming

Traditional programming contrasts fundamentally from AI. In Traditional programming, a developer code every one of the guidelines in meeting with a specialist in the business for which programming is being created. Each standard depends on a sensible establishment; the machine will execute a result understanding the intelligent proclamation. At the point when the framework develops intricate, more guidelines should be composed. It can immediately become unreasonable to keep up with.

Traditional programming contrasts essentially from AI. In Traditional programming, a developer code every one of the principles in counsel with a specialist in the business for which programming is being created. Each standard depends on a legitimate establishment; the machine will execute a result understanding the intelligent explanation. At the point when the framework develops complicated, more principles should be composed. It can immediately become impractical to keep up with.

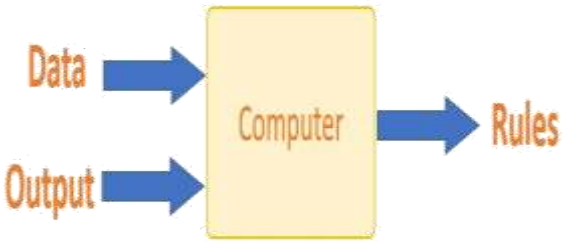

Fig. 1.1 Traditional Programing

AI should conquer this issue. The machine figures out how the info and result information are connected and it composes a standard. The developers don't have to compose new principles each time there is new information. The calculations adjust because of new information and encounters to further develop viability over the long run.

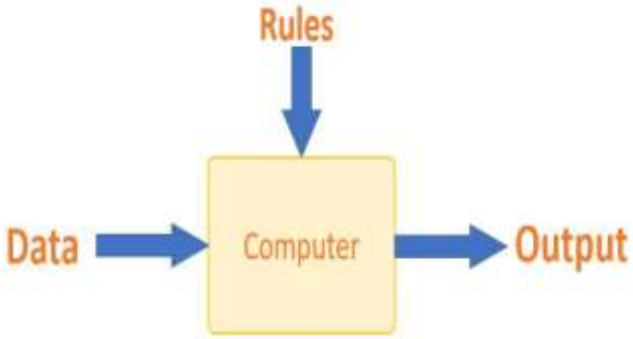

Fig. 1.2 Machine Learning

The primary resurgence for Profound Learning is a direct result of the accompanying reasons:

2. Learning Portrayals:

For AI assignments, the contributions to the frameworks are mostly handmade elements removed from the first picture/text. Programmed learning of the elements is one of the main parts of Profound Designs.

2.1. Circulated Portrayals:

Numerous NLP models takes in count the count of words or sack of word approach for characterization or recovery, however that will sum up the new test information and the portrayals are exceptionally meager as well. As of late the word2vec approach utilizing a shallow brain organization and skip n gram has shown generally excellent outcomes in figuring out semantic connection between words. Deep learning models can be taken care of as information this vector portrayal of words for various NLP undertakings. Different Deep Learning Calculations that are cutting edge for various AI errands are as per the following:

a) Recurrent Neural Network

b) Faster RCNN (Faster Regional Convolutional Neural Network)

c) Auto Encoders

d) Deep Belief Networks and so on

2.2 How in all actuality does Machine Learning Function?

Machine Learning is the cerebrum where all the learning happens. The manner in which the machine learns is like the person. People gain for a fact. The more we know, the more effectively we can anticipate. By similarity, when we face what is happening, the probability of achievement is lower than the known circumstance. Machines are prepared something very similar. To make an exact expectation, the machine sees a model. At the point when we give the machine a comparable model, it can resolve the result. Notwithstanding, similar to a human, on the off chance that its feed a formerly inconspicuous model, the machine experiences issues to predict.

The center target of ML is the learning and derivation. Most importantly, the machine learns through the disclosure of examples. This revelation is made thanks to the information. One critical piece of the information researcher is to select cautiously which information to give to the machine. You can consider a component vector a subset of information that is utilized to handle an issue.

The machine utilizes a few extravagant calculations to work on the truth and change this revelation into a model. Hence, the learning stage is utilized to portray the information and sum up it into a model.

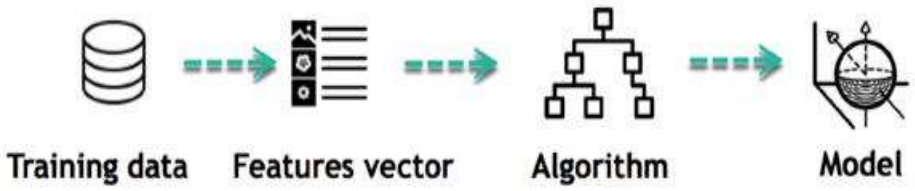

Fig 1.3 Learning Phase

The machine utilizes a few extravagant calculations to work on the truth and change this revelation into a model. Hence, the learning stage is utilized to portray the information and sum up it into a model. For example, the machine is attempting to comprehend the connection between the pay of an individual and the probability to go to an extravagant eatery. It turns out the machine tracks down a positive connection among wage and going to a top of the line café: This is the model

2.2.1 Inferring

Exactly when the model is gathered, it is plausible to test serious areas of strength for how is on never-seen-before data. The new data are changed into a features vector, go through the model and give an assumption. This is all the great piece of ML. There is convincing explanation need to revive the rules or train again the model. You can use the model as of late ready to make acceptance on new data.

The existence of ML programs is clear and can be summed up in the accompanying focuses:

1. Define an inquiry
2. Collect information
3. Visualize information
4. Train calculation
5. Test the Calculation
6. Collect input
7. Refine the calculation
8. Loop 4-7 until the outcomes are fulfilling
9. Use the model to make an expectation

When the calculation significantly improves at reaching the right determinations, it applies that information to new arrangements of information.

3. Machine learning Algorithms

ML can be gathered into two expansive learning errands: Regulated and Unaided. There are numerous different calculations

3.1. Supervised learning

It's algorithm utilizes preparing information and criticism from people to get familiar with the relationship of given contributions to a given result. For example, a professional can involve promoting cost and weather conditions gauge as information to foresee the deals of jars.You can utilize regulated realizing when the result information is known. This calculation will anticipate new information.

There are two classes of regulated learning:

- Classification task
- Regression task

3.1.2. Classification

Envision you need to anticipate the orientation of a client for a business. You will begin gathering information on the level, weight, work, pay, buying container, and so forth from your client data set. You know the orientation of every one of your client, it must be male or female. The goal of the classifier will be to dole out a likelihood of being a male or a female (i.e., the name) in light of the data (i.e., highlights you have gathered). At the point when the model figured out how to perceive male or female, you can utilize new information to make a forecast. For example, you just got new data from an obscure client, and you want to find out whether it is a male or female. Assuming the classifier predicts male = 70%, it implies the calculation makes certain at 70% that this client is a male, and 30% it is a female.

The label can be of at least two classes. The above AI model has just two classes, yet on the off chance that a classifier needs to foresee object, it has many classes (e.g., glass, table, shoes, and so forth each item addresses a class)

3.1.3. Regression

At the point when the result is a persistent worth, the undertaking is a relapse. For example, a monetary investigator might have to gauge the worth of a stock in light of a scope of element like value, past stock exhibitions, macroeconomics file. The framework will be prepared to gauge the cost of the stocks with the most reduced conceivable mistake.

4. Conclusion

This was the project of system design about “Music Recommendation System Using Facial Expression Recognition” based on python. Development of this system takes a lot of efforts as different technologies is needed to be integrated with the software. This system can give a lot of satisfaction to the music lovers and the users. However, not every task is said to be perfect in this development field even more improvement may be possible in this system, but the main motto of this project is to play songs according to the emotion of the person and it is somehow satisfied as of now. I have learned so many things and gained a lot of knowledge about development field. I hope this will prove fruitful to us.

S.No	AUTHOR	PROPOSED STATEMENT	PROS	CONS	DATASET	METRICES
1	Ankita Mahadik 2021	In our proposed system, a mood-based music player is created which performs real time mood detection and suggests songs as per detected mood.	Used Realtime Facial Images	Less Accuracy	Face Expression Dataset	75%
2	S. L. Happy 2020	Facial landmark detection stage to remove noise by applying Gaussian Filter or mask	active facial patches are extracted	The training database consist less image	Facial landmark	65%
3	Sohom Sen 2020	The input images are selected from the training set. After this Landmark detection & Local representation will be done. By using LBP (Local Binary Pattern) algorithm, Local regions LBP features and Local Regions NCM (Normalized Centre Movement) features are extracted.	Used Realtime Facial Images	The system was trained with 4 types of facial expression.	CK+ dataset	84%
4	G.Priyadharshini, 2T.Gowtham 2021	Receives a user heart rate from a smart band or a face image from a mobile camera , it analyses what the user emotion is.	High Accuracy	Sad mood detection is the lowest accuracy with 40%.	JAFEE	90%

5	Sushmita G	They proposed a system in which they used PCA(Principal component approach) for feature extraction. To classify and recognize the expression	songs belonging to Multiple category	Accuracy Low	Database with 7 expressions of 4 individual's persons	65%
6	Anuja Arora 2021	They submitted a program in which the DEAM data set was used to classify the emotions. It has more than 2800 songs with 4 emotions annotated	Used Realtime Facial Images	4 Emotions only Implemented	DEAM data set	75%
7	Nutan Deshmukh 2020	Focused on creating a system that fetches the emotion of the user using a camera and then automates the result using the emotion detection algorithm.	Faster Process	Average calculated estimation	Own Facial Dataset	60%
8	Chang Liuet 2020	Described a system that makes use of Brain-Computer Interfaces, also called as BCI.	Used EEG Signals	EEG Signals Process was Complex	EEG Dataset	90%
9	Parul Tambe 2022	Proposed an idea which automated the interactions between the users and music player, which learned all the preferences emotions and activities of a user and gave song selection as a result.	High Accuracy	Lowest accuracy with 50%.	JAFEE	50%
10	Cyril Laurier 2020	Real-time music mood sync tool, Mood Cloud classifies music emotions into 5 types SVM library to analyze the emotion dataset.	No visuvalization supported	Accuracy Low	JAFEE	70%
11	Bhimani Perry Miyani Bansari 2021	It extracts the facial features of the user from the captured image. According to the emotion, the music will be played from the predefined directories.	Used realtime Facial Expression	Needs future improvement with new technologies	Face Expression Dataset	80%
12	Diana Kayser Hanke Eglrmanh 2021	Audience facial expressions detected by automated face analysis software reflect emotions in music	Used Realtime facial images	Less Accuracy	CK+ Dataset	60%

S.NO	**AUTHOR**	**PROPOSED STATEMENT**	**PROS**	**CONS**	**DATASET**	**METRICES**
13	Ebenezer Owusu Jacqueline Asor Kusni 2021	On Facial Expression Recognition Benchmarks	Used Realtime facial images	Less Accuracy	Benchmark Dataset	75%
14	Santosh ra,cjpmder Sowrabh.M 2020	Smart Music Player Based on Facial Expression.	High Accuracy	Sad mood detection is Low	JAFEE	70%
15	Cuiqing Huang Qiang Zhang 2021	Research on music emotion Recognition model of deepa learnng based om musical stage effect.	Used Musical Stage effect	Four emotion only implemented	MNIST	83%
16	S Metilda Folrence M Uma 2020	Emotiojnal Detection and music recommendation	Used Real time facial images	Lowest Accuracy	Benchmark Dataset	78%

		system based on user facial expression				
17	Jialie Shen Karen Rafferty 2021	It identifies four key research directions: content descriptor generation, personalization, playlist optimization and performance evaluation.	Music suggested can improve related people's well-being.	Doesn't involve advanced techniques	CK+dataset	60%
18	Anand Raju 2021	It convert audio data into image and to visualize the data, plots the amplitude versus time graph. Thus the data becomes more reluctant at low frequency.	Categorizes only audio signals	Precision is sufficient for only 10 epochs and needs to be trained every time	SVN (or) KNN classifier	67%
19	Ziyang Yu Mengda Zhao 2020	It uses micro-expression recognition technology to identify a model that recognizes facial micro-expressions and recommends music according to corresponding mood.	Emotion based Classification	Less accuracy	FER 2013	70%

References

1. Lyons MJ, Budynek J, Akamatsu S. Automatic classification of single facial images. IEEE Transactions on Pattern Analysis and Machine Intelligence 1999;21:1357±62.
2. Moe Moe Htay, Zin Mar Win. Survey on Emotion Recognition Using Facial Expression. International Journal of Computer (IJC) (2019) Volume 33, No 1, pp 1-10.
3. https://www.slideshare.net/Nehapevekar/genetic-algorithm-based-music- recommender-system.
4. H. I. James, J. J. A. Arnold, J. M. M. Ruban, M. Tamilarasan, and R. Saranya, "EMOTION BASED MUSIC RECOMMENDATION SYSTEM," vol. 06, no. 03, p. 6, 2019.
5. V. R. Ghule, A. B. Benke, S. S. Jadhav, and S. A. Joshi, "Emotion Based Music Player Using Facial Recognition," vol. 5, no. 2, p. 7, 2007.
6. S. Gilda, H. Zafar, C. Soni, and K. Waghurdekar, "Smart music player integrating facial emotion recognition and music mood recommendation," in 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Mar. 2017, pp. 154–158, doi: 10.1109/WiSPNET.2017.8299738.
7. Gokul Krishnan K, Parthasarathy M, Sasidhar D and Venitha E, "Emotion detection and music recommendation system using machine learning" in 2018 International Journal of Pure and Applied Mathematics, 2018, vol. 119, pp. 1487-1498.
8. "12_Emotion.pdf." Accessed: Jun. 09, 2020. [Online]. Available: http://ijircce.com/upload/2019/february/12_Emotion.pdf.
9. https://opencvpythontutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/_py_face_det ection.html (accessed May 07, 2020). "Face Detection using Haar Cascades — OpenCV- Python Tutorials 1 documentation."

10. "python - AttributeError: 'module' object has no attribute 'createFisherFaceRecognizer,'" Stack Overflow. https://stackoverflow.com/questions/42018103/attributeerror-module-object-has-no- attribute-createfisherfacerecognizer (accessed May 07, 2020).
11. https://www.w3resource.com/pandas/dataframe/dataframe-dropna.php (accessed May 07, 2020). "Pandas DataFrame: dropna() function - w3resource."
12. "argparse — Parser for command-line options, arguments and sub-commands — Python 3.8.3rc1 documentation." https://docs.python.org/3/library/argparse.html (accessed May 07, 2020).
13. "OpenCV: cv::VideoCapture Class Reference." https://docs.opencv.org/3.4/d8/dfe/classcv_1_1VideoCapture.html (accessed May 07, 2020).
14. https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.read_excel.html (accessed May 07, 2020). "pandas.read_excel — pandas1.0.3 documentation."
15. https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_api.html (accessed May 07, 2020). "FaceRecognizer — OpenCV 2.4.13.7 documentation."
16. A. Mahmood, S. Hussain, K. Iqbal, and W. S. Elkilani, "Recognition of Facial Expressions under Varying Conditions Using Dual-Feature Fusion," Mathematical Problems in Engineering, 2019. [Online]. Available: https://new.hindawi.com/journals/mpe/2019/9185481/. [Accessed: 08-Jan-2020].
17. "Recognition of Facial Expressions under Varying Conditions Using Dual-Feature Fusion." [Online]. Available: https://www.hindawi.com/journals/mpe/2019/9185481/#B3. [Accessed: 30-Dec-2019].
18. A. Fathallah, L. Abdi, and A. Douik, "Facial Expression Recognition via Deep Learning," in 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, 2017, pp. 745–750, doi: 10.1109/AICCSA.2017.124.
19. H. Bahuleyan, "Music Genre Classification using Machine Learning Techniques," arXiv:1804.01149 [cs, eess], Apr. 2018.
20. A. Elbir, H. Bilal Çam, M. Emre Iyican, B. Öztürk, and N. Aydin, "Music Genre Classification and Recommendation by Using Machine Learning Techniques," in 2018 Innovations in Intelligent Systems and Applications Conference (ASYU), Oct. 2018, pp. 1–5, doi: 10.1109/ASYU.2018.8554016.
21. "Musical genre classification of audio signals - IEEE Journals & Magazine." Accessed: Jan. 18, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/1021072.
22. H. Bahuleyan, "Music Genre Classification using Machine Learning Techniques," ArXiv180401149 Cs Eess, Apr. 2018, Accessed: Jan. 18, 2020. [Online]. Available: http://arxiv.org/abs/1804.01149.
23. S. Kumar, "Haar cascade Face Identification," Medium, Jun. 01, 2020. https://medium.com/analytics-vidhya/haar-cascade-face-identification-aa4b8bc79478 (accessed Jul. 04, 2020).