IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 4, April 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104227 |

Video and Music Recommendation Based on Facial Expressions

Neetu M D[1], Rashmi B R[2], Sahana N Jain[3], Sinchana U Ghatge[4], Dr. Naveen Kumar K R[5], Prof. Arjun H[6]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[1]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[2]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[3]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[4]

Associate Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[5]

Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[6]

ABSTRACT: Recommendation of video and music using facial expression play a vital role for Millennials, Gen Z, Gen Alpha in helping uplift their mood as they spend 80% of their time on the OTT platforms. Our project mainly focuses on facial expressions, which plays a crucial role in identifying the emotion of a person. Our project is an facial expressions and recommendation system, which uses machine learning algorithms to detect the status of a person's facial expressions and recommend movies and music based on it.This is an facial expressions system with music and video recommendations based on the emotions. The model to be built uses the transfer learning approach for which the MobileNet model will be used. Based on the classification of 7 emotions, different templates will be created for both video as well as music recommendations.Since music therapy has been studied for use in managing numerous medical conditions, mood enhancement and anxiety/stress relief, This movie and music recommendation feature based on the detected emotion of a person plays vital role in current and upcoming generations

KEYWORDS: Machine Learning, Transfer learning approach, Recommendation System, MobileNet

I. INTRODUCTION

Recommendation systems have been based on context and content, and now the technological challenge of making personalized recommendations based on the user's emotional state arises through physiological signals that are obtained from devices or sensors. Recommending music and videos based on a user's music preference is a way to improve user experience. One of the most studied ways of mood detection is by using facial expressions, which is still one of the challenging fields in pattern recognition and machine learning science. Deep Neural Networks (DNN) have been widely used in order to overcome the difficulties in facial expression classification. Biologically, Facial Expressions are derived from the relative position or motion of muscles that lie under the skin. According to certain controversial theories, these also convey the emotional state of the individual at a given time. They are indeed controversial because one can easily fake their expressions. But in the world where communication is one of the most important acts, facial expression is the means of non-verbal communication. Recommender system, as the name says, simply means a system

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104227 |

that could be used to recommend items to a user on the basis of some information or criterion like past feedback of user, or other user pattern.

This project is aimed at not using the user's past feedback or other patterns, it will rather use the user's facial expression to recommend him entities like videos or songs. Hence creating a recommendation system that will require less user data and should still be able to work nicely as user requirements might not be related to his past but with the present that is signified by his/her expressions.

II. LITERATURE SURVEY

In order to get required knowledge about various concepts related to the present application, existing literature was studied. Some of the important conclusions were made through those are listed below.

Bakhtiyariet al.[1] discussed a fuzzy model for multi-level human emotions recognition by computer systems through keyboard keystrokes, mouse and touch-screen interactions. This model can also be used to detect the other possible emotions at the time of recognition. Accuracy measurements of human emotions by the fuzzy model are discussed through two methods; the first is accuracy analysis and the second is false positive rate analysis. This fuzzy model detects more emotions, but on the other hand, for some of the emotions, a lower accuracy was obtained with the comparison with the non-fuzzy human emotions detection methods. This system was trained and tested by Support Vector Machine (SVM) to recognize the users' emotions. Overall, this model represents a closer similarity between human brain detection of emotions and computer systems.

Mala Saraswat et al.[2]analyzed the top–N recommendation model based on emotion analysis of the reviews of the item. This paper compared emotion based method with rating based cosine item to item similarity model. The results showed that emotion based item similarity provided more accurate recommendation than rating based cosine item similarity.It used 42 fuzzy rules for classifying each movie as liked or disliked based on its emotional content. Results conclude that fuzzy emotion features do not increase the prediction accuracy as compared to discrete emotion features.

Seanglidetet al.[3] proposed an application which has 3 main steps:

1.Face feature detection
2.State interpretation
3.Emotion prediction

OpenCV (Open Source Computer Vision) is a well-known library of programming functions that can be used to detect and recognize faces. Once the face is identified, the instance of the active shape model is created in order to do the fitting by calling image alignment method.Next the parameters were chosen based on two factors. First factor was that it should clearly represent the changes of 6 emotions: Anger, Disgust, Fear, Happy, Sad, and Surprise. Second factor is that it shall focus on permanent features such as face, eyes, nose and lips. Classification was done using Support Vector Machines (SVMs) which is a supervised learning model with an associated learning algorithm that analyzes data and recognizes patterns for classification and regression analysis.In addition, a music player function is included where users can listen to the music which changes according to the predicted mood.

Ai Thanh Ho et al.[4] proposed a system which is a hybrid of two techniques: collaborative filtering and content-based filtering.In order to do so, it analyzed the color sequence by using the following logic: if at least two of three chosen colors indicated the same emotion and then this emotion becomes the current emotional state of the user. User profiles were used which contained information about movie ratings and the emotion related to each movie. This information was stored in the User Profiles Database.

Rishabh Ahuja et al.[5] proposed a movie recommender system built using the K-Means Clustering and K-Nearest Neighbor algorithms. The data was taken from movielens data set. The system was implemented in the python

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104227 |

programming language. It was seen that after implementing the system in the python programming language the RMSE value of the proposed technique was better than the existing technique. It was also seen that the RMSE value of the proposed system achieved the same value as the existing technique but with less no of clusters.

Prudhvi Raj Dachapallyet al.[6] proposed two independent methods for this very task. The first one using autoencoders, though the intuition was right and the idea was fairly original, it did not generate the best results. One of the main reasons is since they concatenated all the pixels vertically to feed into the autoencoder, which led to some structural integrity of the image. In the future,they want to replace the normal hidden units with convolution filters in the encoder part, and deconvolution units for the decoder. The second method was using a convolutional neural network, and with three convolutional layers, three pooling layers and two fully – connected layers, which gave a good accuracy on the JAFFE test set as well as on the LFW test set. Results show that with more fine-tuning and depth, our CNN model can outperform the state-of-the-art methods for emotion recognition by creating an 8-layer CNN with three convolutional layers, three pooling layers, and two fully connected layers.

FerdosFessahayeet al.[7] focused on an approach to improving music recommendation systems, although the proposed solution could be applied to many different platforms and domains, including Youtube (videos), Netflix (movies), Amazon (shopping), etc.Current systems lack adequate efficiency once more variables are introduced. Their algorithm, Tunes Recommendation System (T-RECSYS), uses a hybrid of content-based and collaborative filtering as input to a deep learning classification model to produce an accurate recommendation system with real-time prediction. They applied Their approach to data obtained from the Spotify Recsys Challenge, attaining precision scores as high as 88% at a balanced discrimination threshold. The algorithm utilizes an input vector carrying information useful for both content-based and collaborative filtering. This input is fed into a deep neural network, and through a training process, the network learns to recognize patterns in a user's listening history, ultimately recommending songs it has confidence the user will enjoy.

Shabab Bazrafkanet al.[8] shown that the classification accuracy is significantly lower when the network is trained with one database and tested with a different database.Deep Network has been proposed which uses the DNN to extract the most representative features.in [20] a DNN with five layers and 65K neurons has been designed to classify the expression into five categories (Neutral, happy, sad, angry and surprised).Three databases has been used in the research. Radboud Faces Database (RaFD), Cohn-Kanade AU-Coded Facial Expression Database Version 2 (CK+) and The Japanese Female Facial Expression (JAFFE) Database.The goal of this research is to investigate the amount of error that occurs from training a DNN network for a database and test it on other databases.

Keunwoo Choi et al.[9] introduced auralisation of CNNs, which is an extension to CNNs visualization. This is done by inverse-transformation of a deconvolved spectrogram to obtain a 9 time-domain audio signal. Listening to the audio signal enables researchers to understand the mechanism of CNNs when they are applied to spectrograms. One effective way to understand and explain the CNNs was introduced in, where the features in deeper levels are visualized by a method called deconvolution. By deconvolving and un-pooling layers, it enables people to see which part of the input image is focused on by each filter. Visualization of CNNs, which showed how high level features (postures/objects) are combined from low-level features (lines/curves). Visualization of CNNs helps not only to understand the process inside the black box model, but also to decide hyper-parameters of the networks.

Luz Santamaria-Granados et al.[10] applied the deep learning approach using a deep convolutional neural network on a dataset of physiological signals (electrocardiogram and galvanic skin response), in this case, the AMIGOS dataset. The detection of emotions is done by correlating these physiological signals with the data of arousal and valence of this dataset, to classify the affective state of a person. In addition, an application for emotion recognition based on classic machine learning algorithms is proposed to extract the features of physiological signals in the domain of time, frequency, and nonlinear. This application uses a convolutional neural network for the automatic feature extraction of the physiological signals, and through fully connected network layers, the emotion prediction is made. The

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104227 |

experimental results on the AMIGOS dataset show that the method proposed in this paper achieves a better precision of the classification of the emotional states, in comparison with the originally obtained by the authors of this dataset.

R Raja Subramanian et al.[11] proposed a CNN model to detect different facial emotions in real time. The work done by the authors made the model work in a better way such that it can recognize various emotions in different illumination and pose. The model uses Keras which is a python library that helps in deep learning models, tensor flow and OpenCV to detect emotions in face at different situations. The authors initially built a model and trained it such that it was capable of detecting and classifying facial emotion in real-time. The model gave good accuracy but took too much time to train the data.

Pranshu Diwan et al.[12] an FER model that uses Support Vector Machines(SVM) and Convolutional Neural Networks (CNN).The paper covers all the necessary datasets and algorithms that are required to build an efficient facial emotion recognition model.Optimization of the parameters was explained using the following tuning parameters: regularization, gamma and the margin parameters. The important algorithms that were used to train a model like this were Support Vector Machines, Linear Discriminant Analysis, Basic CNN,Hidden Markov models and Deep learning.

AmoghN.Parab et al.[13] proposed a system that has a newfangled band and a web backend. It comprises an Internet of Things (IoT) based system representing the band and backend server along with prediction models representing the web backend. The authors used the amalgamation of concepts like Electrodermal Activity, Pulse Rate Sensing, Emotion and Sentiment Analysis, an individual's stress as well as emotional state, under different scenarios, can be identified. The proposed design also deals with critical areas involving judging the mental ability of an individual.

RamalathaMarimuthuet al.[14] proposed a model where emotions are classified using deep belief networks (DBN), which is trained by moth Search Optimization Algorithm (MSA) along with the standard gradient descent algorithm (SGD). The performance of the proposed method of emotion recognition is analyzed using evaluation measures, such as False Rejection Rate (FRR),accuracy, and False Acceptance Rate (FAR).The proposed MSA-DBN method outperformed the existing models with a maximal accuracy, FAR, FRR values of 98.5%,0.63%, and 0.77 %, respectively.

M. Kalpana Chowdhury et al.[15] dealt with emotion recognition by using transfer learning approaches. In this work pre-trained networks of Resnet50, vgg19, Inception V3, and Mobile Net are used. The fully connected layers of the pre-trained ConvNets are eliminated, and we add our fully connected layers that are suitable for the number of instructions in our task. Finally, the newly added layers are only trainable to update the weights. The experiment was conducted by using the CK + database and achieved an average accuracy of 96% for emotion detection problems

2.1 Literature review summary

Emotion detection using conventional approaches having the drawback of mutual optimization of feature extraction and classification. To overcome this problem, researchers are showing more attention toward deep learning techniques.The paper compared emotion based method with rating based cosine item to item similarity model. It used 42 fuzzy rules for classifying each movie as liked or disliked based on its emotional content.The design deals with critical areas involving judging the mental ability of an individual to handle pressure and perform under situations where they are not the most comfortable. An algorithm Tunes Recommendation System (T-RECSYS) which uses a hybrid of content-based and collaborative filtering as input to a deep learning classification model to produce an accurate recommendation system with real-time prediction.It was seen that after implementing the system in the python programming language the RMSE value of the proposed technique was better than the existing technique such as Collaborative filtering.A fuzzy model for multi-level human emotion recognition. Accuracy was measured by 2 methods- accuracy analysis and false positive rate analysis. This system was trained and tested by the Support Vector Machine (SVM).

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104227 |

III. COMPONENTS USED

1.Camera: The camera is the primary input device that captures the user's facial expressions. A high-quality camera with good resolution is needed to accurately capture and analyze the user's facial features.

2.Facial recognition algorithm: A facial recognition algorithm is used to detect and recognize facial expressions. The algorithm should be able to accurately recognize a wide range of facial expressions and emotions.

3.Machine learning algorithms: Machine learning algorithms are used to analyze the data captured by the facial recognition algorithm and provide recommendations based on the user's emotional state. These algorithms can be trained to adapt to the user's preferences and provide more personalized recommendations over time.

4.Recommendation engine: The recommendation engine is responsible for analyzing the user's emotional state and providing relevant video and music recommendations. The engine can use a variety of techniques such as collaborative filtering, content-based filtering, and hybrid filtering to provide recommendations.

5.Music and video database: The music and video database contains a large collection of video and music tracks that can be recommended to the user. The database should be regularly updated with new content to provide fresh recommendations.

6.User interface: The user interface provides the user with a way to interact with the system. The interface should be intuitive and user-friendly, allowing the user to easily access recommended content and provide feedback.

7.Cloud infrastructure: Cloud infrastructure can be used to host and scale the system. This allows for easy deployment and maintenance of the system.

IV. METHODOLOGY

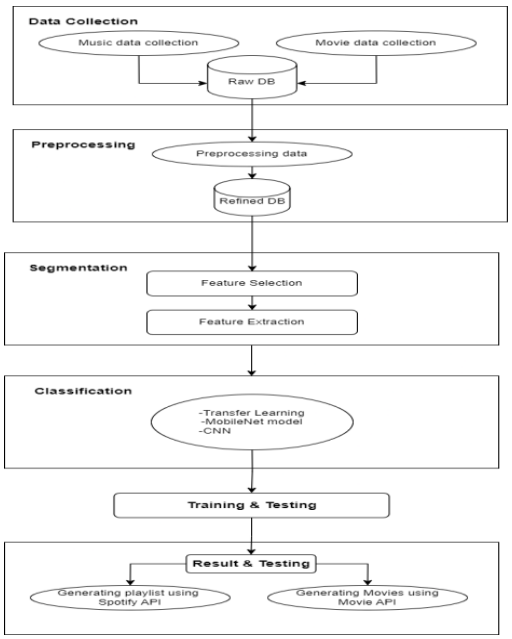

Fig 4.1 : Emotion Detection

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104227 |

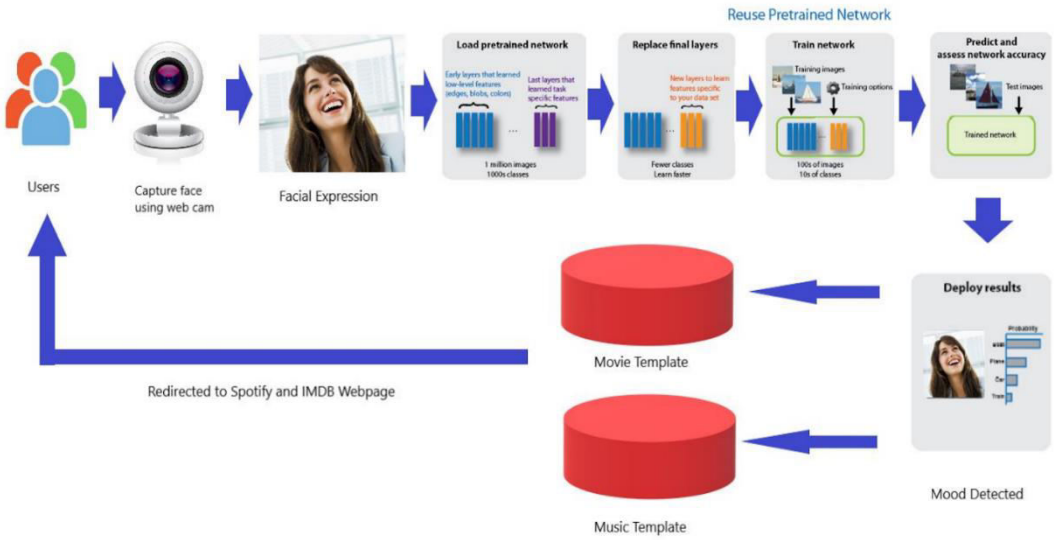

Fig 4.2 : Training the model

For the project we will implement the transfer learning algorithm for detecting the mood of the user. It works by transferring as much knowledge as possible from an existing model to a new model designed for a similar task. For example, transferring the more general aspects of a model which make up the main processes for completing a task. This could be the process behind how objects or images are being identified or categorized. Extra layers of more specific knowledge can then be added to the new model, allowing it to perform its task in new environments. The deep learning architecture is nothing but a set of weights. A set of continuous epochs would provide good accuracy to our deep learning model. Transfer learning can be used to achieve good accuracy with few epochs and modifying only those weights which are related to our problem. Several deep learning architectures like the RestNet101,VGG19, AlexNet, GoogleNet and MobileNet have overcome previous architectures to make good predictions with very high accuracy predicting for nearly classifying thousand classes.The main benefits of using transfer learning include:

- Removing the need for a large set of labeled training data for every new model.
- Improving the efficiency of machine learning development and deployment for multiple models.
- A more generalized approach to machine problem solving, leveraging different algorithms to solve new challenges.
- Models can be trained within simulations instead of real-world environments.

Mobilenet is a model which does the same convolution as done by CNN to filter images but in a different way than those done by the previous CNN. It uses the idea of Depth convolution and point convolution which is different from the normal convolution as done by normal CNNs. This increases the efficiency of CNN to predict images and hence they can be able to compete in the mobile systems as well. Since these ways of convolution reduce the comparison and recognition time a lot, it provides a better response in a very short time and hence we are using them as our image recognition model.

The FER2013 dataset consisted of 48x48 px grayscale images which were re-sampled using bilinear interpolation and converted to 224x224px for higher resolution and getting the same weights. As we know,The number of neurons inside the fully connected layer is always equal to the number of classes, which in this case is 7, So the thousand layered pre-trained model (which in this case is MobileNet) is converted into a fully connected layer consisting of 7 neurons.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104227 |

IV. CONCLUSION

In this research paper, the recommendation system basically combines two different recommender systems i.e. one for videos and other one for music using the human emotion conveyed through expression using face detection and classification algorithm so as to get the emotion. The emotions that the system can detect were happy, sad, angry, neutral, fear or surprised. After determining the user's emotion, the proposed system provided the user with a playlist that contains music and video matches that detected the mood. Here, we provide an overview of how music and videos can affect the user's mood and how to choose the right content to improve the user's moods.

REFERENCES

1. Bakhtiyari, Kaveh, and Hafizah Husain. "Fuzzy model on human emotions recognition."arXiv preprint arXiv:1407.1474 ,2014.
2. Saraswat, Mala, Shampa Chakraverty, and Atreya Kala. "Analyzingemotion based movie recommendation system using fuzzy emotion features." International Journal of Information Technology 12, no.2: 467-472,2020
3. Seanglidet, Yean, Bu Sung Lee, and Chai Kiat Yeo. "Mood prediction from facial video with music "therapy" on a smartphone." In 2016 wireless telecommunications symposium (wts), pp. 1- 5. IEEE, 2016.
4. Ho, Ai Thanh, Ilusca LL Menezes, and YousraTagmouti. "E-mrs: Emotion-based movie recommender system." In Proceedings of IADIS e-Commerce Conference. USA: University of Washington Both-ell, pp. 1-8,2006.
5. Ahuja, Rishabh, Arun Solanki, and Anand Nayyar. "Movie recommender system using k-means clustering and k-nearest neighbor." In 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 263-268. IEEE, 2019.
6. Dachepally, Prudhvi Raj. "Facial emotion detection using convolutional neural networks and representational autoencoder units." arXiv preprint arXiv:1706.01509,2017
7. Fessahaye, F., Perez, L., Zhan, T., Zhang, R., Fossier, C., Markarian, R. & Oh, P.T-recsys: A novel music recommendation system using deep learning. In 2019 IEEE international conference on consumer electronics (ICCE) (pp. 1-6). IEEE,2019
8. Bazrafkan, Shabab, Tudor Nedelcu, Pawel Filipczuk, and Peter Corcoran. "Deep learning for facial expression recognition: A step closer to a smartphone that knows your moods." In 2017 IEEE International Conference on Consumer Electronics (ICCE), pp. 217-220. IEEE, 2017.
9. Choi, Keunwoo, et al. "Auralisation of deep convolutional neural networks: Listening to learned features." Proceedings of the 16th International Society for Music Information Retrieval Conference, ISMIR,2015.
10. Santamaria-Granados, L., Munoz-Organero, M., Ramirez-Gonzalez, G., Abdulhay, E., &Arunkumar, N. J. I. A.Using deep convolutional neural networks for emotion detection on a physiological signals dataset (AMIGOS). IEEE Access, 7, 57-67,2018
11. Subramanian, R. R., Niharika, C. S., Rani, D. U., Pavani, P., &Syamala, K. P. L.Design and Evaluation of a Deep Learning Algorithm for Emotion Recognition. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS),2021
12. Balasubramanian, B., Diwan, P., Nadar, R., & Bhatia, A.Analysis of Facial Emotion Recognition. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI),2019
13. Parab, A. N., Savla, D. V., Gala, J. P., &Kekre, K. Y. Stress and Emotion Analysis using IoT and Deep Learning. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA),2020
14. Haridas, A. V., Marimuthu, R., & Chakraborty, B. (2020). Emotion Recognition System for Specially Needed People with Optimized Deep Learning Algorithm.Fourth International Conference on Inventive Systems and Control (ICISC),2020
15. Chowdary, M. Kalpana, Tu N. Nguyen, and D. Jude Hemanth. "Deep learning-based facial emotion recognition for human–computer interaction applications." Neural Computing and Applications : 1-18, 2020

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details