e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704005 |

Human Emotion Recognition and Notification Using AI

Dr.S.Govindaraju[1], Arulkumaran A.S[2]

[1]Associate Professor, PG& Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu India

[2]UG Student, PG& Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamil Nadu India

ABSTRACT: Emotion detection is the process of using technology to identify and interpret human emotions based on various signals, such as facial expressions, voice intonation, and body language. Emotion detection is a growing field in artificial intelligence and is used in various applications, including customer service, mental health diagnosis, and entertainment. There are various techniques used for emotion detection, including machine learning algorithms, neural networks, and natural language processing. These techniques analyze patterns in human behavior to recognize emotional states such as happiness, anger, fear, and sadness. The proposed system will identify the emotion of the person using facial expression analysis and generates the notification to the user device such as mobiles. There are several algorithms used for emotion detection using face expression, and the choice of algorithm depends on the specific requirements and application. CNN is a type of neural network that is commonly used for image processing and analysis. CNN is particularly effective for facial expression recognition as it can extract features from the image that are relevant to the task. CNNs have achieved state-of-the-art performance in many facial expression recognition benchmarks.

Keywords: CNN, Emotion detection, Machine Learning, expression analysis, Artificial intelligence.

I.INTRODUCTION

Emotion detection using face expression is a popular approach in artificial intelligence. It involves the use of computer vision algorithms to analyze and interpret facial expressions to detect emotions such as happiness, sadness, anger, surprise, and fear. Facial expression analysis works by capturing images or videos of a person's face and analyzing the movements and changes in the facial muscles and features. This information is then used to identify the emotions being expressed. There are different techniques used for facial expression analysis, including the use of feature extraction and machine learning algorithms. Feature extraction involves identifying key features in the face such as the position of the eyebrows, mouth, and eyes, which are used to determine facial expressions. Machine learning algorithms, on the other hand, use statistical models to learn patterns in the data and classify emotions based on the facial features detected. These algorithms can be trained on large datasets of labeled images, allowing them to learn to recognize emotions accurately.

Facial expression analysis has many practical applications, including in marketing, customer service, mental health diagnosis, and entertainment. For example, companies can use facial expression analysis to assess consumer emotions towards a product or service, while therapists can use it to help diagnose and treat patients with mental health disorders such as depression or anxiety. However, there are also concerns about the ethical implications of using facial expression analysis, particularly with regards to privacy and data protection. Additionally, there is the risk of bias in data analysis, particularly when training machine learning algorithms on biased datasets. There are various algorithms used for emotion detection using face expression, including:

Support Vector Machines (SVMs): SVMs are a popular classification algorithm used for emotion recognition. They work by finding the optimal hyperplane that separates the data into different classes. Convolutional Neural Networks (CNNs): CNNs are a type of deep learning algorithm that have shown promising results in facial expression analysis. They work by learning hierarchical representations of the input data, allowing them to detect complex patterns in the images. Random Forest: Random Forest is an ensemble learning algorithm that combines multiple decision trees to classify emotions based on facial expressions.
K-Nearest Neighbors (KNN): KNN is a simple machine learning algorithm used for classification tasks. It works by finding the k closest data points in the training set and assigning the most common class label to the new data point. Deep Belief Networks (DBNs): DBNs are a type of deep learning algorithm that have shown promising results in facial expression analysis. They work by training multiple layers of artificial neural networks, allowing them to learn

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704005 |

hierarchical representations of the input data. The choice of algorithm depends on various factors, such as the size and complexity of the dataset, the computational resources available, and the specific application requirements. Researchers and developers often compare the performance of different algorithms on benchmark datasets to choose the most suitable one for their application.

II.LITERATURE REVIEW

According to researchers, Facial expressions form a major part of human communication, 55% of the entire human face-to-face communication spectrum [2],[6]. This is to imply that the inability to read facial expressions means you are missing out on more than half of the total meaning of what someone wants to communicate. Aside from facial expressions, humans express emotions through other means like voice intonation, which accounts for about 38% of all conversations, and just 7% are the actual choice of words used to press emotions. [2],[6],[7]. There are a couple of models on emotions. These include, among others, Plutchik's Wheel of Emotions, Izard, Panksepp & Watt, Levenson, and Ekman [8]. which accounts for about 38% of all conversations, and just 7% are the actual choice of words used to press emotions. [2],[6],[7]. There are a couple of models on emotions.

FER combines three inherent tasks: detecting face area, extracting and representing data of interest, and recognizingthe expression. Machine Learning models of FER have twophases; a training phase and a testing phase. Two approaches have been adopted for data representation, viz., the feature-based geometric and the appearance-based approaches. During the implementation of the feature-based geometric approach, image processingtechniques to extract vital facial points (i.e., corners of the lip, middle of the eye, the ends of the eyebrows, and the tip of the nose). The resulting coordinates are used to construct a facial geometry made of the extracted characteristics vectors. The appearance-based approach examines video frame by frame and generates an attribute vector using an image filter. Recent studies have added on more facial emotion, contempt, which makes interest seven (7)[11]. Facial Emotion Recognition is a typical classification problem that can be solved using several classification methods such as k-Nearest Neighbours (KNN), Decision Tree (DT), Learning Vector Quantization (LVQ), and multilayer Feed- forward Neural Network (MFFNN) [12].

In recent years ANNs such as Convolution Neural Network (CNN) and Deep CNN have been used for image classification with very high accuracies. ANNs are systemsinspired by the brain's biological neural networks. They are made up of a collection of artificial neurons which are interconnected. Each neuron can send signals to another connected neuron which can also process the signal received and transmit it after that. Neurons are arranged inlayers, and each layer may perform a different type of transformation on the receiving input signal. Some ANNs are capable of Backpropagation which allows the data to flow backward in the network to adjust the network's effectiveness. Deep Learning (Deep Neural Network) is based on ANN.

A 2-dimensional array is created, which is then passed to the flatten layer to be converted to a single-dimensional vector and then passed to a two-layered network used to classify the emotions. For a probabilistic output, softmax is used as the activation function for the output layer. The model was trained and tested for 11 epochs with a learning rate of 0.01. The model had an accuracy of 78.04%. The dataset used was manually created using a 48 MP camera and contained a total of 2550 images, each having a pixel size of 1920x2560. The dataset was split into 2040 training images, 255 validation images, and 255 testing images. The emotions tested for were angry, happy, neutral, sad, and surprised. An experiment FER system was developed, which is divided into three processes: preprocessing with the ViolaJones algorithm, feature extraction with local fisher discriminant analysis (LFDA) for dimensionality reduction and k-nearest neighbors (KNN), and classification with a feed-forward artificial neural network (ANN) [6]. The dataset used for this experiment is JAFFE, but the researchers focused on four (4) emotions. Happy, Neutral, Surprised, and Sad out of the seven emotions present in the dataset. The two classifiers were compared in terms of performance. The 1NN algorithm performed better with sad and neutral emotions, while the ANN algorithm performed well with happy and surprise emotions, but for the average performance, the ANN algorithm outperforms the 1NN algorithm with an average performance of 66.66% as against 54.16%. Jaiswal et al.

[14] proposed a deep learning architecture of two different CNN networks for facial emotion detection. The proposed model uses Keras and has an input shape of 48*48*1 and two models with the same kernel size for feature extraction. The submodels are flattened into vectors and then concatenated into one long vector before transmitting to a fully connected softmax layer for classification. The performance of their architecture was evaluated using two datasets, FER2013 and JAFFE, and the accuracies realized were 70.14% and 98.65%, respectively, for the two datasets. The choice of datasets by the researchers was to make the model more robust in terms of diversity. Lasri et al. [15] also proposed a CNN architecture to recognize students' facial emotions. Their architecture consists of 4 convolution layers,

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704005 |

each with maximum pooling layers and two fully connected layers. The proposed make use of softmax to predict facial emotions. The model was trained and tested using the FER2013 dataset. Their model also scored 70% on accuracy. Naik and Mehta [16] proposed a CNN model that is used for FER in a hand-over-face situation called Hand-over-Face Gesture-based Facial Emotion Recognition Method (HFG_FERM). Hand-over- face is typically considered in other facial emotion recognition as an instance of occlusion, and as such, images with hand-over-face are exempted from the experiments. The proposed research provides extensive coding schemas with additional hand signs that help identify unexplored emotions such as confidence, making a decision, and scared, ashamed, angry, and ok signs along with basic emotions.

The authors validated their model with images from the Cam3d corpus, FER2013 dataset, and public domains summing up to a total of 18 emotion categories. The performance of their model was compared with two other models, Multimodal Fusion Approach (MMFA) and Emotion Recognition through Facial Gestures (ERTFG). From their experiments, their model HFG_FERM outperforms MMFA and ERTFG on different levels. The researchers in [17] proposed different architectures of Convolution Neural Networks (CNNs) of two models to classify seven facial emotions. The first CNN model had four convolutional layers, four max pooling, one dropout, and two fully connected layers. The second model used the same model as the first but with data augmentation. The dataset used in this experiment is iCV MEFED (Multi- Emotion Facial Expression Dataset). Their choice of the dataset was relatively new and had compound/mixed emotion, e.g., angrily surprised, tears of joy. The dataset contains 5,750 images, each of size 5184x3456 pixels. It was shown that the model performed better with images that are not distorted than with augmented images. It was also revealed that emotions such as sadness and contempt are under-predicted compared to the other emotions. Bouzakraoui et al. [18] proposed a model that automatically detects facial expressions displayed by clients when they react to a product or service. First, the researchers extracted geometric features from the customers' faces and then used adapted SVM to predict the customers' satisfaction. The image is converted to geometric primitives such as points and curves during the geometric feature extraction by measuring relative distances between distinct features like the eyes, eyebrows, nose, mouth, and chin. A vector of 19 values that represents the customer's facial expression from these distances is generated. The dataset used for this experiment is the JAFFE dataset. The researchers reclassify the emotion in the dataset into three classes; satisfied, not-satisfied and neutral. Jain et al. [19] proposed a model based on a single Deep Convolutional Neural Networks (DNNs). The proposed model consists of six convolution layers, two deep residual blocks, and two fully connected layers, each with a ReLU as activation function and dropout. These features of the model help the model to learn subtle features which are related to certain emotions. Two residual blocks contain four convolution layers with varying sizes, two short connections, and one skip connection. For the classification of emotion, softmax is used. Their proposed model was trained and tested on two datasets, CK+ and JAFFE, and it was found that the combination of fully connected networks and residual block improved the overall performance of the proposed model. Researchers in [20] proposed a facial expression recognition based on the Valence-Arousal dimensional emotion model. The proposed model uses a valence dimension prediction which has nine levels. The proposed model implores the use of CNN to predict a result equal to the weighted fusion of valence value and its corresponding probability. The CNN architecture consists of 4 convolution layers, each with ReLU as activation function, three maximum pooling layers placed after convolution layers 2, 3, and 4 respectively, and Two (2) fully connected layers. The model uses softmax for the classification

III.METHODOLOGY OF PROPOSED SURVEY

3.1 Emotion detection using face expression

Data Collection: The first step is to collect a large and diverse dataset of facial expressions. The dataset should contain images or videos of people expressing different emotions such as happiness, sadness, anger, surprise, fear, and disgust. The dataset should also be balanced, meaning that there should be an equal number of samples for each emotion. The dataset can be collected using various methods such as crowdsourcing, online platforms, or in-person data collection.

Preprocessing: The collected data is preprocessed to remove noise and irrelevant features. This involves resizing, cropping, and normalization of the images. The images are also converted to grayscale to reduce the complexity of the data and improve the performance of the model.

Feature Extraction: Features are extracted from the preprocessed images to represent the key facial components and expressions. The goal of feature extraction is to identify the relevant facial components that are most informative for predicting emotions accurately. Popular features used for facial expression analysis include local binary patterns, facial landmark points, and geometric features.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704005 |

Training the Model: The extracted features are used to train a machine learning or deep learning model. The choice of algorithm depends on the size and complexity of the dataset, as well as the specific application requirements. For example, support vector machines (SVMs) are a popular classification algorithm used for emotion recognition. Convolutional neural networks (CNNs) are another popular approach that has shown promising results in facial expression analysis. The model is trained using labeled data, where each image or video is labeled with the corresponding emotion.

Testing and Evaluation: Once the model is trained, it is tested on a separate set of data to evaluate its performance in predicting emotions accurately. The test set should be representative of the population and should contain samples that are not included in the training set. Metrics such as accuracy, precision, recall, and F1-score are used to evaluate the performance of the model. The model can be fine-tuned based on the evaluation results to improve its performance.

Deployment: The final step is to deploy the trained model in the target application. This may involve integrating the model into a larger system or developing a standalone application. The model should be tested thoroughly in the target environment to ensure that it performs accurately and reliably.

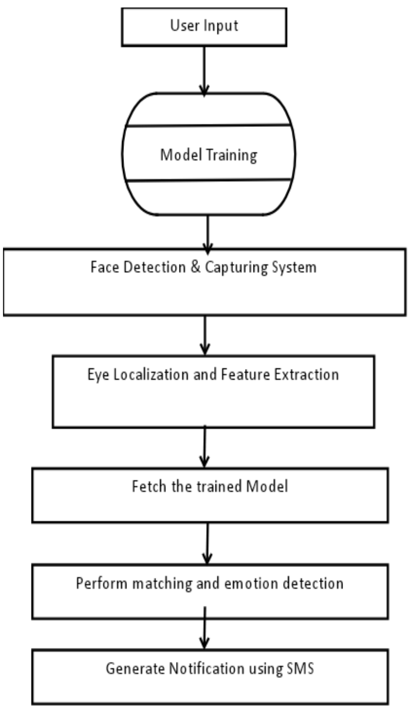

Figure-1 System Architecture

An emotion detection system offers several advantages, including:

- Accurate emotion detection: An emotion detection system can accurately detect emotions in a variety of situations and contexts, including facial expressions, vocal tone, and text analysis. This can provide valuable insights into the emotions of individuals and groups, and help to inform decision-making processes.
- Real-time feedback: An emotion detection system can provide real-time feedback on the emotions of individuals, enabling immediate responses and interventions in situations where emotions are a critical factor.
- Enhanced customer experience: An emotion detection system can be used in customer service contexts to identify the emotions of customers and respond appropriately, improving the overall customer experience and satisfaction.
- Increased efficiency: An emotion detection system can automate the process of emotion detection, reducing the need for manual analysis and saving time and resources.
- Objective analysis: An emotion detection system can provide objective analysis of emotions, reducing the potential for bias or subjective interpretation of emotional cues.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704005 |

- Data-driven insights: An emotion detection system can generate insights from large volumes of data, helping to identify patterns and trends in emotional responses and behaviors.
- Improved decision-making: An emotion detection system can provide valuable information for decision-making processes in a variety of contexts, such as marketing, healthcare, and education.

Overall, an emotion detection system can provide valuable insights into emotions and help to inform decision-making processes, leading to improved outcomes and experiences for individuals and organizations.

IV.RESULTS AND DISCUSSION

Convolutional Neural Networks (CNNs) are a popular choice for emotion detection systems, especially when using facial expressions as input. Here are the steps involved in using CNNs for an emotion detection system: Data collection and preparation: Collect a large dataset of labeled facial expressions and preprocess it for training the CNN. The dataset should include a variety of facial expressions, such as happiness, sadness, anger, fear, and surprise. Data augmentation: To increase the size of the dataset and improve the robustness of the CNN, data augmentation techniques can be used. Common techniques include image rotation, scaling, cropping, and flipping. Building the CNN: Design the architecture of the CNN by selecting the number of convolutional layers, pooling layers, and fully connected layers. The architecture of the CNN can vary depending on the complexity of the emotion detection problem and the size of the dataset. Training the CNN: Train the CNN on the preprocessed dataset using Backpropagation and stochastic gradient descent. The CNN should be trained until it reaches a satisfactory level of accuracy on the validation dataset. Model improvement: If the performance of the CNN is not satisfactory, adjust the architecture of the network or the training parameters and retrain the model. Deployment: Deploy the trained CNN to the emotion detection system for real-time detection of emotions in facial expressions.

CNNs are effective for emotion detection because they are able to automatically extract relevant features from the input images, reducing the need for manual feature engineering. However, training CNNs can be computationally expensive, especially when using large datasets and the performance of the model may be affected by factors such as lighting conditions and facial expression variations. Once a CNN model has been trained for emotion detection, there are several ways to improve its performance: Fine-tuning: Fine-tuning is a process where a pre-trained model is further trained on a smaller dataset specific to the task at hand. For example, a pre-trained CNN model trained on a large dataset of images can be fine-tuned on a smaller dataset of facial expressions to improve its performance on emotion detection. Hyper parameter tuning: The performance of a CNN model can be affected by its hyper parameters, such as the number of convolutional layers, the learning rate, and the dropout rate. Tuning these hyper parameters can improve the performance of the model. Transfer learning: Transfer learning is a technique where a pre-trained model is used as a starting point for a new task. For example, a pre-trained CNN model trained on a large dataset of images can be used as a feature extractor for emotion detection by removing the last few layers of the network and replacing them with new layers specific to the emotion detection task.

Ensemble learning: Ensemble learning is a technique where multiple models are trained on the same dataset and their predictions are combined to improve performance. For example, multiple CNN models with different architectures or hyper parameters can be trained and their predictions combined using averaging or voting.

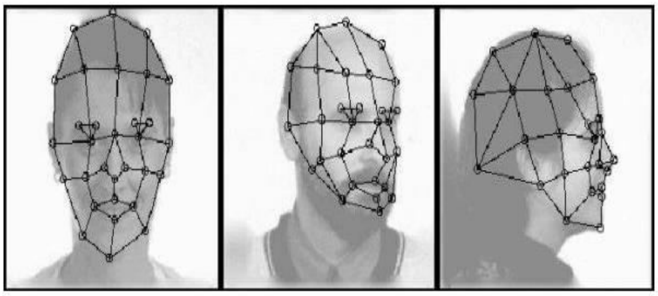

Figure-2 Face landmarks

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704005 |

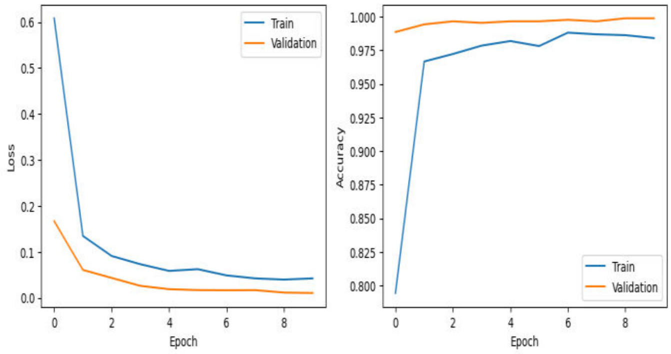

Figure-3 Validation Accuracy

REFERENCES

[1] Jaiswal, S. and Nandi, G.C., 2020. Robust real-time emotion detection system using CNN architecture. Neural Computing and Applications, 32(15), pp.11253-11262.
[2] Dachapally, P.R., 2017. Facial emotion detection using convolutional neural networks and representational autoencoder units. arXiv preprint arXiv:1706.01509.
[3] Jaiswal, A., Raju, A.K. and Deb, S., 2020, June. Facial emotion detection using deep learning. In 2020 international conference for emerging technology (INCET) (pp. 1-5). IEEE.
[4] Ali, M.F., Khatun, M. and Turzo, N.A., 2020. Facial emotion detection using neural network. the international journal of scientific and engineering research.
[5] Santamaria-Granados, L., Munoz-Organero, M., Ramirez-Gonzalez, G., Abdulhay, E. and Arunkumar, N.J.I.A., 2018. Using deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS). IEEE Access, 7, pp.57-67.
[6] Abdullah, M., Hadzikadicy, M. and Shaikhz, S., 2018, December. SEDAT: sentiment and emotion detection in Arabic text using CNN-LSTM deep learning. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 835-840). IEEE.
[7] Acheampong, F.A., Wenyu, C. and Nunoo-Mensah, H., 2020. Text-based emotion detection: Advances, challenges, and opportunities. Engineering Reports, 2(7), p.e12189.
[8] Kanjo, E., Younis, E.M. and Ang, C.S., 2019. Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. Information Fusion, 49, pp.46-56.
[9] Bertero, D. and Fung, P., 2017, March. A first look into a convolutional neural network for speech emotion detection. In 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 5115-5119). IEEE.
[10] Polignano, M., Basile, P., de Gemmis, M. and Semeraro, G., 2019, June. A comparison of word-embeddings in emotion detection from text using bilstm, cnn and self-attention. In Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization (pp. 63-68).
2. Tibbits, S., van der Harten, A., & Baer, S. (2011). Rhino *Python* Primer (3rd ed.).

[11] Downey, A. B. (2015). *"*Think Python: How to Think Like a Computer Scientist (2edition)". Sebastopol, CA: O'Reilly Media.
[12] Greg Wilson. "Data crunching: solve everyday problems using Java, Python and more. The pragmatic programmers", Pragmatic Bookshelf, Raleigh.
[13] Guido van Rossum and Fred L. Drake, Jr. "The Python Tutorial — An Introduction to Python". Network Theory Ltd., Bristol,
[14] Michael Dawson. "Python programming for the absolute beginner". Premier Press Inc., Boston, MA, USA, 2003.
[15] Harvey M. Deitel, Paul Deitel, Jonathan Liperi, and Ben Wiedermann"Python How To Program". P T R Prentice-Hall, Englewood Cliffs.
[16] Brad Dayley. "Python phrasebook: essential code and commands. Developer's library". SAMS Publishing, Indianapolis,
[17] Liza Daly. "Next-generation web frameworks in Python". O'Reilly & Associates, Inc., 103a Morris Street, Sebastopol.
[18] Mike Dawson. "Python programming for the absolute beginner". Thomson Course Technology, Boston,
[19] Peter Norton, Alex Samuel, David Aitel, Eric Foster-Johnson, Leonard Richardson, Jason Diamond, Aleatha Parker, Michael Roberts, "Begining Python", 2005.

BIOGRAPHY

Dr S Govindaraju MCA MPhil PhD he pursed Master of Computer Applications @ Gobi Arts and Science College from Bharathiar University, Coimbatore in the year 2005 and completed MPhil in Computer Science from Bharathiar University in the year 2011 and he completed PhD in Bharathiar University, Coimbatore in the year 2019 and currently working as an Associate Professor PG and Research Department of Computer Science Sri Ramakrishna College of Arts and Science (Formerly SNR Sons College), Bharathiar University, Coimbatore. He has published more than fourteen research papers in reputed international journals including Thomson Reuters (SCOPUS) and conferences and it's also available in online. His main research work focuses on Image Retrieval using Medical Images. He has seventeen years of Teaching experience and twelve years of Research experience.