International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2023* ***Revised:*** *12 December 2023* ***Accepted:*** *30 January 2024*

Emoticon Generation, Expression Recognition, and Gender Classification Using Deep Learning in Real-Time

***[1]Bineet Kumar Joshi, [2]Bansidhar Joshi**
[1]Himalayan School of Science & Technology Swami Rama Himalayan University Dehradun, India
[2]Jaypee Institute of Information Technology Noida, India

***Abstract*-**Images play an increasingly important role in identifying a person's gender and emotional state in today's digital environment, but there are still methodological hurdles to overcome. Image processing utilizing deep learning algorithms is the way to go. Our study's overarching goal is to find ways to bridge the communication gap through the use of emoticons based on the emotions conveyed in photographs and snapshots. We have utilized the Keras framework to implement a deep learning algorithm called a Convolutional Neural Network (CNN) and evaluated it using Tensor Flow to predict gender. The goal is to create a new dataset of pictures free of noise and then utilize those images as inputs to a convolutional neural network (CNN). The algorithm's result is supposed to be more trustworthy gender identification based on increased accuracy. We have implemented an LSTM-RNN (Long short-term memory recurrent neural network) for emotion identification and facial expression detection. Feature selection is the most crucial step since it will ultimately aid in emoticon generation**.**

***Keywords*:** Deep Learning; Emotional Recognition; Face Attributes; CNN**.**

1. INTRODUCTION

The ever-increasing communication gap between individuals is no longer an issue because to advances in population and technology. In today's technological age, where snapping selfies is all the rage. Even if you use a filter that makes you look younger, your features will reveal your true age. Do you think it's fair to suggest that men and women look completely different? Well, there is, but the question is how to convey it to a machine. Inaccurate and unreliable outcomes are produced by the image processing method. With the use of cloud computing, users of social media and messaging apps may now more easily convey their emotions through the usage of emoticons [1].

Photos always tell the truth about the subjects in them. But modern technology allows us to determine such details about an individual as their gender, age, emotional state, physical attributes, and more. It's possible to learn the truth about someone without sacrificing their anonymity by extracting attributes from an image and analyzing the data for a pattern [2].

Although images contain a wealth of information, unlike words, that information cannot be parsed with ease. The purpose is not to transcribe visual information into text but rather to cherry-pick and extract facts that will be useful in the future. The value of data extraction from images lies in the subsequent information generation.

The state of the art in image processing still doesn't give a workable solution, but this may be overcome by combining image processing methods with machine learning algorithms. Machine learning algorithms can be used to the data extracted from a picture using image processing techniques to get the desired results. Collectively, these data will shed light on widespread phenomena, such as gender ratios in a nation.

In the past, image processing was commonly employed to analyze and collect data from images, but it seldom produced useful insights. When it comes to automating various operations that call for facial recognition, findings from image processing alone are not reliable. To combat the growing temporal complexity of our efforts to enhance image processing algorithms for better outcomes, we have turned to a different method: machine learning. Machine learning, when combined with image processing, yields superior outcomes.

Because of the complexity of the task at hand, we turned to deep learning, an improved kind of machine learning. Our goal is to improve precision by combining machine learning with image processing. We expect our algorithm to produce more reliable results that may be put to practical use in the business world. In the United States, anti-theft facial recognition systems are commonly used in retail establishments including electronics stores and clothing boutiques. However, there have been instances where the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2023* ***Revised:*** *12 December 2023* ***Accepted:*** *30 January 2024*

matched pair was incorrect. We hope to prevent such occurrences by developing an algorithm with great accuracy. When compared to other methods, our Algorithm is a significant improvement that yields usable findings in just a few microseconds [3-5].

2. METHOD AND EXPERIMENTAL DESIGN

The ever-increasing communication gap between individuals is no longer an issue because to advances in population and technology. In today's technological age, where snapping selfies is all the rage. Even if you use a filter that makes you look younger, your features will reveal your true age. Do you think it's fair to suggest that men and women look completely different? Well, there is, but the question is how to convey it to a machine. Inaccurate and unreliable outcomes are produced by the image processing method. With the use of cloud computing, users of social media and messaging apps may now more easily convey their emotions through the usage of emoticons.

Photos always tell the truth about the subjects in them. But modern technology allows us to determine such details about an individual as their gender, age, emotional state, physical attributes, and more. It's possible to learn the truth about someone without sacrificing their anonymity by extracting attributes from an image and analyzing the data for a pattern [1-2]. Although images contain a wealth of information, unlike words, that information cannot be parsed with ease. The purpose is not to transcribe visual information into text but rather to cherry-pick and extract facts that will be useful in the future. The value of data extraction from images lies in the subsequent information generation.

The state of the art in image processing still doesn't give a workable solution, but this may be overcome by combining image processing methods with machine learning algorithms. Machine learning algorithms can be used to the data extracted from a picture using image processing techniques to get the desired results. Collectively, these data will shed light on widespread phenomena, such as gender ratios in a nation.

In the past, image processing was commonly employed to analyze and collect data from images, but it seldom produced useful insights. When it comes to automating various operations that call for facial recognition, findings from image processing alone are not reliable. To combat the growing temporal complexity of our efforts to enhance image processing algorithms for better outcomes, we have turned to a different method: machine learning. Machine learning, when combined with image processing, yields superior outcomes. Because of the complexity of the task at hand, we turned to deep learning, an improved kind of machine learning.

Our goal is to improve precision by combining machine learning with image processing. We expect our algorithm to produce more reliable results that may be put to practical use in the business world. In the United States, antitheft facial recognition systems are commonly used in retail establishments including electronics stores and clothing boutiques. However, there have been instances where the matched pair was incorrect. We hope to prevent such occurrences by developing an algorithm with great accuracy. When compared to other methods, our Algorithm is a significant improvement that yields usable findings in just a few microseconds [6-8]

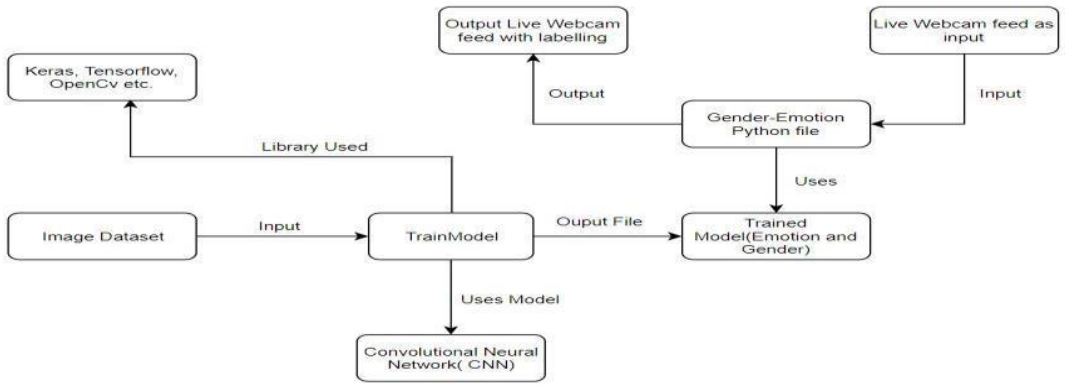

Figure 1: Flow Activity Diagram of the methodology used [8]

As can be seen in Fig. 1, the libraries and picture dataset are fed into the CNN model to form the trained-model file, which is then utilized in the emotion-gender Python file in conjunction with the live webcam feed to output the gender and emoticon.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2023* ***Revised:*** *12 December 2023* ***Accepted:*** *30 January 2024*

3. SIGNIFICANCE OF WORK

In the modern era, it is extremely difficult to accurately predict a person's gender and emotions. The use of filters in images via various applications makes it even more challenging. In this endeavor, we employed deep learning techniques such as CNN to achieve superior results. We have implemented these techniques on both still images and live webcams in order to classify gender and detect emotions in real time. The distinctiveness resides in the emoticon generation in addition to emotion detection and classification. Since a person's emotions can alter every second, live webcams are used for the same purpose. Because a person's facial expression may change, the output displays the real-time emotion with the appropriate emoticon based on the individual's face, as well as a complete graph of each emotion.

The ability to accurately detect a person's emotions enables its practical application. Without precision, the work would not be trustworthy enough to extract information from. This endeavor will function as a rational judge to collect a person's moment-by-moment emotions throughout the day. Such data will enable virtual assistants to make recommendations regarding the user's medication, sleep cycle, and dietary patterns, among others.

4. RELATED WORKS

Face detection, face recognition, gender recognition, and age classification are the primary phases of the gender estimation procedure for humans. The first step is essential, as all subsequent steps are rendered meaningless without it, and a Facial image helps to determine age, gender, ethnicity, scarring, and marks. AdaBoost classifier can be employed for face detection. Face recognition will necessitate feature extraction, which will pave the way for gender recognition based on facial characteristics such as the nose, eyes, jawline, cheeks, and ear [3]. Numerous systems utilize face images of individuals as digital files because they can reveal a variety of personal characteristics, including age, gender, ethnicity, and identification. In multiple vision systems, the use of Support Vector Machine (SVM) for gender estimation in conjunction with the widely used Local Binary Pattern (LBP) feature has been observed to potentially perpetuate victimization. The Local Binary Pattern (LBP) technique was created using a construction and multi-scale block concept, which permits the extraction of both micro- and macro-features from an image [4]. DCT (Discrete Cosine Transform) is utilized to represent the appearance caused by light and facial expressions, which is classified using SVM for gender using Binary classification between male and female [5,8].

Table 1: Depiction of an integrated summary of all previous work

Sr. No.	Paper Name	Paper Summary
1	Age Estimation from Face Images: Challenging Problem for Audience Measurement Systems [3]	The paper talks about age estimation from images which have a human face in it. For face detection, the AdaBoost classifier is used. Authors talk about how this technology will benefit, as this can be used in surveillance feeds.
2	FDDB-360: Face Detection in 360degree Fisheye Images [4]	This paper presents a study conducted by researchers on an age and gender estimate system that incorporates ethnic variations in face photos through the use of a Convolutional Neural Network (CNN) and Support Vector Machine (SVM). The facial picture was inputted into one of the ethnically-specific support vector machines (SVMs) based on the ethnicity prediction made by the convolutional neural network (CNN). This SVM then predicts the final age and gender.
3	Gender and Age Detection using face images [5]	In this research paper author discuss about finding age and gender using Artificial Neural Networks from a face image. Author used Discrete Cosine Transform to represent appearance caused by light and facial expressions, classified using Support Vector Machin for gender using Binary classification between male and female.

4	Age Estimation of Face Images Based on CNN and Divide-and-Rule Strategy[6]	Researchers have mostly directed their attention towards the extraction of facial geometry ratio features for the purpose of extracting human face traits. The utilization of a six-layer Convolutional Neural Network (CNN) is employed for the purpose of extracting facial age features, while the estimation of age is accomplished through the application of manifold learning techniques.
5	Face Image-Based Age and Gender Estimation with Consideration of Ethnic Difference [7]	This paper presents a study conducted by researchers on an age and gender estimate system that incorporates ethnic variations in face photos through the utilization of a Convolutional Neural Network and Support Vector. A proposed classification system for age and gender has been put out by researchers, which incorporates considerations for ethnic variation.
6	HyperFace: A Deep Multi-Task Learning	The authors propose an algorithm named HyperFace which uses CNN for simultaneous face detection and gender recognition. For
7	Deep Learning analysis of Mobile Physiological, Environmental and location sensor data for Emotion Detection.[9]	The authors discussed the importance of emotion detection in human-robot interaction and the type of modeling required to enable this in various applications. The authors also discussed the scope to use low-cost wearable sensors such as fitbits to gather data which will connect actions to emotions.

Face age analysis is utilized and frequently involved in face recognition applications. Hindawi, a mathematical approach, is utilized. In order to extract options for external body parts, the face pure mathematics magnitude relation feature extraction takes precedence. For the purpose of extracting facial features, a six-layer Convolutional Neural Network (CNN) is utilized. To estimate the gender, numerous supervised learning techniques are implemented. Experimental findings indicate that the intended strategy is more discriminative and robust than the ancient physicist, LBP (Local Binary Pattern), and BIF (Best Individual Features) approaches. [6] For gender estimation, standard performance metrics such as Mean Absolute Error, Cumulative Score, and Probability Density of Error can be used.

CNN is used by Hyperface for concurrent face detection and gender recognition. The algorithm detects a visage, localizes landmarks, estimates the pose, and identifies the gender. The two models upon which Hyperface is based are Alex Net and Reset. As post-processing techniques, the authors employ IRP (Iteartive Region Proposal) and L-MNS (Landmark-based Non-Maximum Supression). For the localization of Landmark, the author utilized 21-point markups for a visage. Due to entire pose variations, some landmarks may be obscured. For face selection, a selective search yielding a high detection score was used. For face detection, a feature map displaying activations alongside various units was utilized. One potential improvement to the HyperFace method is the use of a high recall rapid face detector to generate regional recommendations, thereby enhancing the algorithm's computational efficiency. This would enable the algorithm to be applied to live surveillance feeds and provide Face detection and gender recognition [8]. This study investigates the applied mathematics strategies utilized in the manual extraction of physiological signals in comparison to the automated extraction methods that are more closely associated with emotional states. Establishing a correlation between physiological signals and the arousal and valence information contained in the dataset enables the classification of an individual's affective state when identifying emotions. The efficacy of emotion prediction depends on the extraction and correlation of physiological signals and galvanic skin response. By pre-processing the peaks in the Electrocardiogram (ECG) and Galvanic Skin Response (GSR) data, [10] the identification of morphological features valuable for state prediction was facilitated.

Deep Alignment Network (DAN) is used to improve facial features extraction results. CNN is used, with two convolutional, two subsampling and one fully connected layer. The primary advantage of DAN is that it can process the entire face image rather than in segments, which would result in a performance decrease. DAN is also able to manage images with significant variation and produce superior results. DAN is the most accurate model compared to CNN(2), CNN(3), Inception-V3, and EmotionNet 2. The AffectNet database is used to store facial expressions and the locations of 68 facial landmarks [11].

Advantages of Human-Computer Interaction devices, Surveillance Cameras for Age Restriction in Traffic, and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2023* ***Revised:*** *12 December 2023* ***Accepted:*** *30 January 2024*

Prohibited Products for Certain Age Limit [5]. This technology will be advantageous because it can be used in surveillance feeds to filter minors from bars. This does not pose a security risk because it does not identify a person. There is no detection and identification of an individual, only a general notion of the person's age irrespective of his uniqueness [3]. The possibility of using low-cost wearable sensors, such as Fitbits, to collect data that connects actions to emotions, thereby creating a more accurate emotion based on facial expressions, surroundings, and actions [9]. Table I depicts an integrated summary of all previous work.

5. TESTING AND EXPERIMENTAL RESULTS

We tested our algorithm using multiple inputs, including individual images, group images, and a live webcam feed to analyze our results for gender and emotion detection. In individual images, the algorithm can accurately predict a person's gender with a high degree of accuracy. In photographs with numerous faces, the algorithm can only discern a face up to a certain threshold, at which point it accurately predicts the subject's gender. The issue occurs in images where the number of features degrades pixel quality. In the case of emoticon generation, singular face detection is performing well. Multiple emotions can exist on a single detected visage, such as happiness and surprise. Therefore, the algorithm is able to identify each emotion. As depicted in Fig. 2, the algorithm worked well to detect gender when a single or multiple features were present in an image.

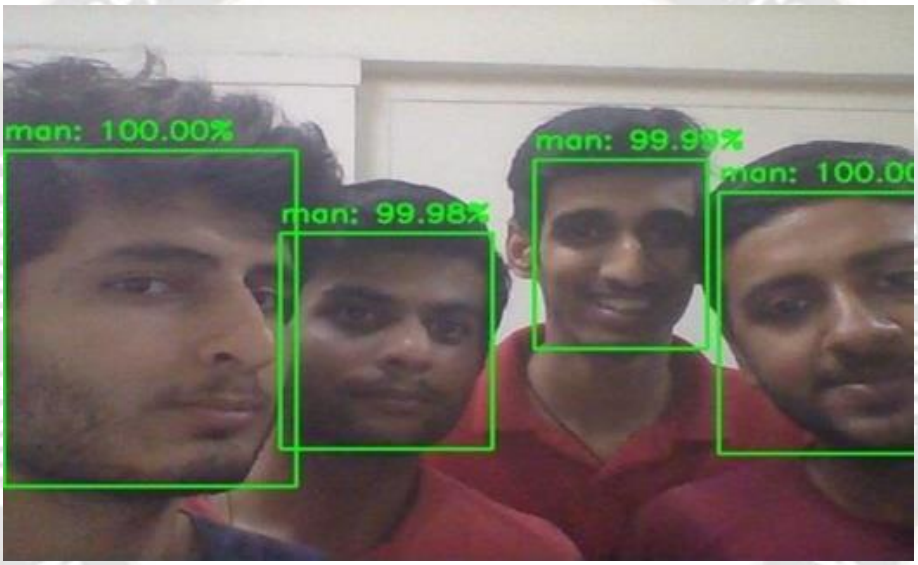

Figure 2: Gender detection via Webcam for multiple faces

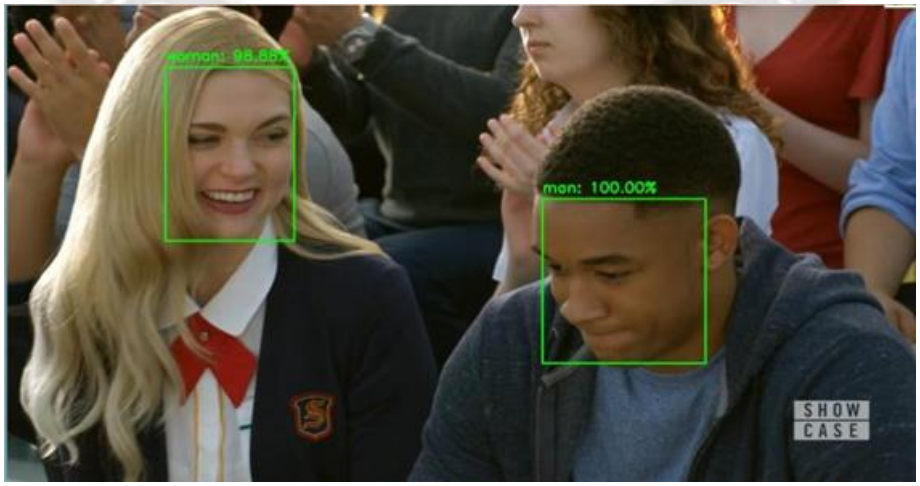

Figure 3: Gender detection for determining different gender

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: 25 November 2023 ***Revised:*** 12 December 2023 ***Accepted:*** 30 January 2024

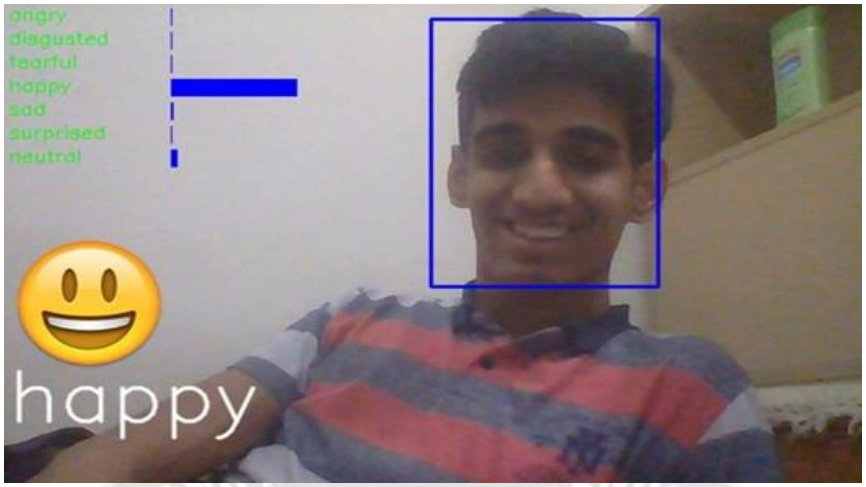

Figure 4: Emotion Detection with Emoticon- Happy

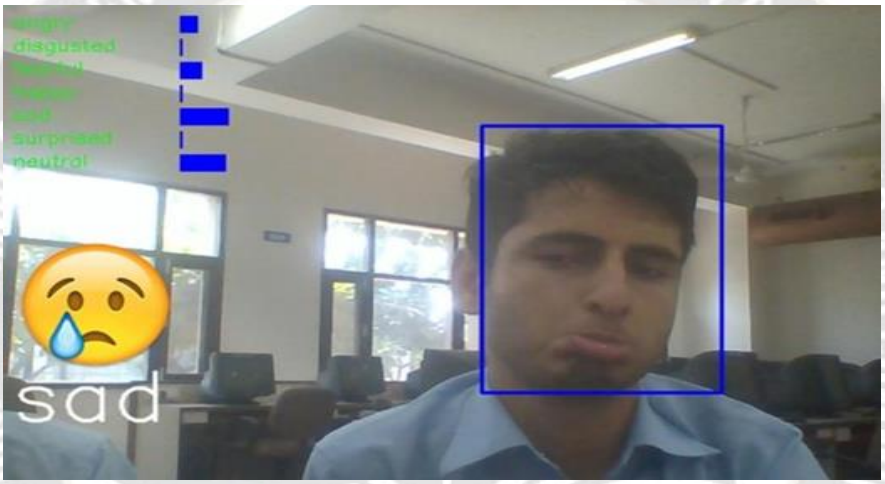

Figure 5: Emotion Detection with Emoticon- Sad

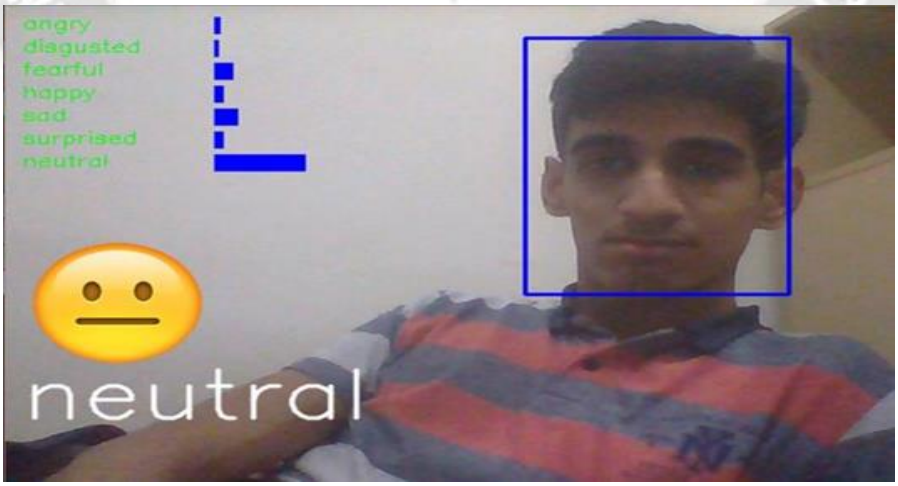

Figure 6: Emotion Detection with Emoticon- Neutral

A Neutral face implies markup points which couldn't be categorized into an emotion. Such emotion will also have characteristics of other emotions as well.

6. CONCLUSION

The algorithm can accurately determine the gender of a person based on an image or a live webcam transmission. As only two categories, namely man and woman, are to be used

to classify each case, it is reasonable to anticipate high accuracy. In cases where pixels are not very distinct or the orientation of the face is not linear, the algorithm may incorrectly identify females as males and vice versa. Such cases are rare in case of an image, but as we shift to webcam feed, the number of cases increases by a substantial quantity. A limitation of the above solution is that the image resolution must be high, i.e., the number of pixels must be high, so that the distinguishing characteristic between the two genders is distinct. Either the image should be of high quality or the face-to-image ratio should be high (in the case of group photos, this ratio would be low).

The purpose of this research is to determine a person's gender, and the algorithm achieves a very high level of accuracy in a variety of circumstances.

7. FUTURE SCOPE

Taking into account the live webcam feed, the algorithm can provide rapid real-time results at each millisecond of time, thereby allowing this application to be used for day-today problems and reducing human manual labor. The analysis of an image or a surveillance broadcast offers an excellent opportunity to supplant manual work with automated calls. There are numerous locations where only women are permitted, such as the first coach of the Delhi Metro. An analysis of the live surveillance feed from inside the train will assist in identifying anomalies and finding the perpetrator.

REFERENCES

[1] Dhawan H, Joshi BK. A Road to Cloud Computing. Lap Lambert Academic Publ. 2012.

[2] Joshi B, Joshi B, Mishra A, Arya V, Gupta AK, Peraković D. A comparative study of privacy-preserving homomorphic encryption techniques in cloud computing. International Journal of Cloud Applications and Computing (IJCAC). 2022 Jan 1;12(1):1-1.

[3] Khryashchev V, Ganin A, Stepanova O, Lebedev A. Age estimation from face images: challenging problem for audience measurement systems. InProceedings of 16th Conference of Open Innovations Association FRUCT 2014 Oct 27 (pp. 31-37). IEEE.

[4] Fu J, Alvar SR, Bajic I, Vaughan R. FDDB-360: Face detection in 360-degree fisheye images. In2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) 2019 Mar 28 (pp. 15-19). IEEE.

[5] Joshi B, Joshi BK, Kumar S. A Novel Deep Learning Based Gender Classification, Expression Recognition and Emoticon Generation. In2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT) 2024 Feb 9 (Vol. 5, pp. 29-33). IEEE.

[6] Liao H, Yan Y, Dai W, Fan P. Age estimation of face images based on CNN and divide-and-rule strategy. Mathematical Problems in Engineering. 2018 Jun 5;2018.

[7] Shin M, Seo JH, Kwon DS. Face image-based age and gender estimation with consideration of ethnic difference. In2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) 2017 Aug 28 (pp. 567-572). IEEE.

[8] Ranjan R, Patel VM, Chellappa R. Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. IEEE transactions on pattern analysis and machine intelligence. 2017 Dec 8;41(1):121-35.

[9] Kanjo E, Younis EM, Ang CS. Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. Information Fusion. 2019 Sep 1;49:46-56.

[10] L.Granados, M.Organero, G.Gonzalez, E.Abdulhay and N.Arunkumar, "Using Deep Convolutional Neural Network for Emotion Detection on a Physiological Signals Dataset (AMIGOS)", IEEE Access, 2019.

[11] Kopalidis T, Solachidis V, Vretos N, Daras P. Advances in Facial Expression Recognition: A Survey of Methods, Benchmarks, Models, and Datasets. Information. 2024 Feb 28;15(3):135.