VOL. 12 NO. 9,
MAY, 2026
E-ISSN 3026-958X
P-ISSN 3027-1169

African Scholar Publications & Research International
www.africanscholarpub.com

Journal of Science Innovation & Technology Research (JSITR)

Patient Satisfaction Feedback Analysis: Facial Emotion Recognition Using Convolutional Neural Network

Nne Rena Saturday[1]; Hachikaru Ngozi Okwu[1]; & Vivian Onyinyechi Anthony[2]

[1]Computer Science Department, Rivers State University. [2]Computer Science Department, University of Port Harcourt.

Corresponding Author: saturday.rena@ust.edu.ng

DOI: https://doi.org/10.70382/ajsitr.v12i9.082

Abstract

Telehealth has become very crucial for remote healthcare services, yet measuring patient satisfaction poses a snag. This study introduces a feedback-enabled telehealth system that adopts Convolutional Neural Networks (CNNs) for real-time facial emotion recognition. Using the CK+ dataset, a MobileNet-based CNN was trained to categorized patient emotions during remote consultations. The model achieved 65.78% accuracy, with performance affected by imbalanced class distributions across facial expressions. These results demonstrate the feasibility of CNN-based emotion recognition as a tool for having real-time access to patient satisfaction in telehealth. Future work will be in the direction of improving the accuracy of the result through balanced datasets and applying privacy-preserving techniques such as federated learning. This research contributes to advancing patient-centric telehealth systems by integrating AI-driven feedback mechanisms into remote healthcare services.

Keywords: Telehealth, CNN, e-Health, Artificial Intelligence, Patient's Satisfaction, Patient Feedback.

Introduction

Tele-health, as a part of e-health, describes the wide list of digital information and communication technologies and systems that allow the delivery of health-related

services (Andrikopoulou, 2023). The fast growth of tele-health, especially during the coronavirus disease 2019 (COVID-19) pandemic, combined with variable regulations and guidelines built increased potential for liability and legal issues (Cascella, 2018). Medical experts can leverage on technology to enhance health care services delivery via tele-health. Tele-health comprises of three distinct types namely: synchronous, asynchronous, and remote (tele) monitoring or remote patient monitoring (RPM); Which has brought about a rapid shift to tele-health models in many settings in both inpatient and outpatient scenarios, so as to prevent and ameliorate spread of deadly disease, patients and providers had to quickly adapt to tele-health models.

However, a notable limitations and flaws, specifically regulatory and technological does exist. The inconsistencies in accessing and use of tele-health were already discovered but was evident during the COVID-19 pandemic (Ortega et al., 2020). This led to the emerging and fast-growing field of digital determinants of health, which is made up of the following factors - digital literacy, access to AI/technology, and community infrastructure (Al-Shorbaji, 2021). AI technology has helped patients to encounter AI-based programs that detect emotions, which could help clinicians offer better mental and behavioral health services. Other applications of AI technology include: Clinicians, such as neurologists and physical therapists, using AI to measure patient locomotion to determine neurological and physiological impairments so as to track treatment success.

Affective computing is one of the AI technologies used in recognition and analysis of human emotion (Hossain and Mohammad, 2019). The body intuitions can send vital information regarding the state of the human body which is termed human emotions. Correct identification of human emotions can help medical experts to obtain health information of a patient in order to provide a more appropriate treatment. The robustness of affective computing can extend its applicability in patient support, mental health monitoring, customized treatment (Yildirim-Celik et al., 2022, Liu et al, 2024). In its basic definition, it is based on enabling computers to recognize, interpret and simulate human emotions (Shingjerji et al, 2026).

Facial expressions of emotions can be a reflex action or a voluntary movement that occur when one or more of the 43 facial muscles are stimulated. Facial expressions are very vital for non-verbal communication and exposes a huge amount of emotional and cognitive information. tele-health systems leverage on facial emotion recognition as it aids the medical practitioner in mental treatment, disease diagnosis and monitoring, patient communication, clinical research, customer service and others. Due to the advancement in technology, recognition of patients' emotions using deep learning algorithms has gained researchers' attention.

There has been no form of assessment that provides feedback to medical professionals in order to ascertain the quality of service rendered for every session with a patient. Thus, this can limit the performance or improvement of these professionals and thereby leave patients dissatisfied. This paper focuses on the analysis of the performance of healthcare delivery as it relates to patients' satisfaction; and bringing quality healthcare services.

Customarily, tele-health principles for special fields are pivoted on patient self-empowerment in order to enhance patients' health and avoid worst case scenarios.

Related Literature

Patients' satisfaction and wellbeing is a strong factor in determining the quality and efficiency of healthcare service delivery. Feedback from patients is important for the measurement of quality assurance of healthcare delivery. A few research is reviewed in this paper to ascertain the use of CNN in patient feedback analysis as a very basic groundwork for subsequent expansion of this research.

Lexicon-Based and Machine Learning Approaches to Patient Feedback

Khaleghparast et al. (2023) developed a hybrid system blending lexicon-based sentiment analysis with machine learning to classify feedback on patient satisfaction. Their model correctly identified positive and negative sentiments and correlated them with specific hospital wards and staff. The study expressed the importance of preprocessing techniques such as tokenization and stop-word removal, and showed that blending rule-based and statistical methods can produce interpretable and actionable insights in healthcare environments. Furthermore, their study provides a basis for integrating sentiment analysis with hospital service evaluation; however, CNNs was not used directly.

Deep Neural Networks with Graph Convolutional Learning

Zhang and Liu (2023) proposed a DNN-based model using Graph Convolutional Neural Networks (GCNNs) for prediction of patient' satisfaction using structured hospital data. Their approach highlighted problems like data sparsity and cost in feedback collection. GCNNs were used to capture relational patterns including service features and patient responses, which improved accuracy of the prediction and scalability. This further showed the potential of CNN variants (GCNNs) for modeling structured healthcare data and improving the prediction of patient' satisfaction.

CNN and LSTM for Sentiment Classification in Healthcare

Altun-Dogan and Dogan (2024) presented a deep learning framework blending CNN and LSTM architectures for the analysis of patient's feedback. Their model had an accuracy of over 85% in sentiment classification, which shows that CNNs are efficient in extracting local features from text, while LSTMs capture long-term dependencies. It also highlighted the importance of patient centric design and interpretable AI in healthcare applications.

Their study validates textual sentiment analysis using CNN technique in healthcare, with major performance metrics and real-world implications.

Methodology

The deep learning model adopted in this paper is the Convolutional Neural Network (CNN) which is appropriate for recognizing images (like pictures) and discovering patterns. It has the capacity to scan through piles of images and extract the patterns for further classification. CNN has different classes; amongst which MobileNet is listed. The diagrams in figures 1 and 2 shows the block diagram of the architecture and the detailed architecture of the system.

The MobileNet CNN Model is a light-weight computer vision model used for training classifiers and is based on depth-wise separable convolutions. The operation on the image is given by:

$$S(i,j) = (I * K)(i,j) = \sum_{m} \sum_{n} I(i+m, j+n) \, . K(m,n)$$

Where:

I = the input image

S(i, j) = the output feature map

K = kernel or filter

i,j = coordinates of the feature map

m,n = coordinates of the kernel

* = convolution of the kernel

The depth-wise separable convolution reduces computational cost when compared to standard convolution and is given by D_K x D_K x M x D_F x D_F

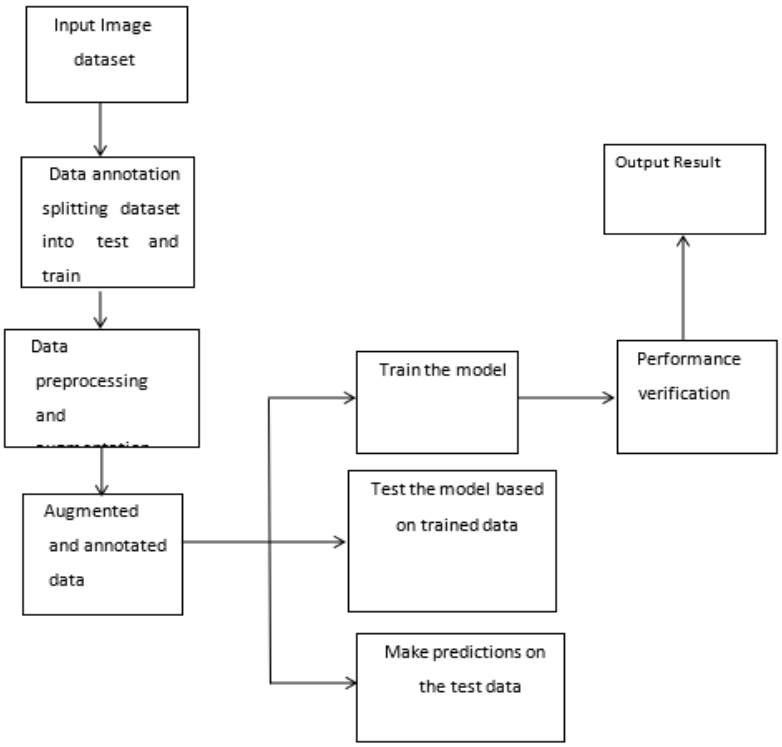


Figure 1: Block Diagram of the System Architecture

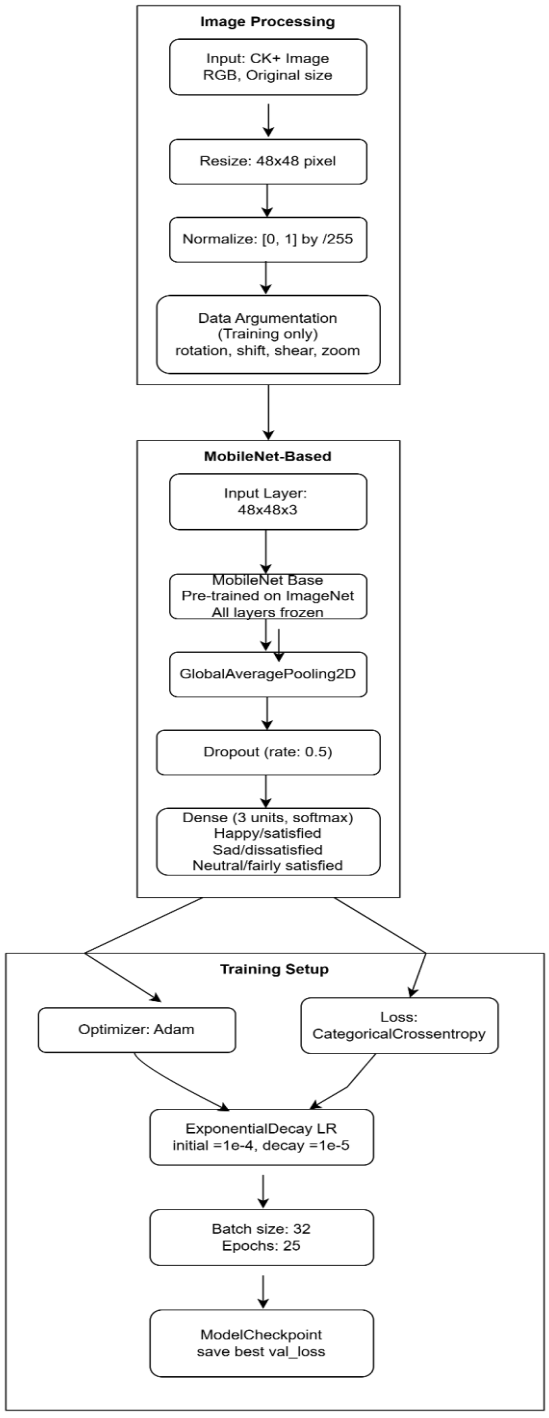


Figure 2: System Architecture

Results and Discussion

The model results and implementation are captured in this section. The model was first trained and evaluated against the public test generator that was already split (into test and train) in the

CK+ dataset website. The CK – dataset contains 6587 gray-scale images of heads (faces) that have been tagged with various expression. A subset of the CK+ dataset consisting of 3 emotion classes – happy, sad, neutral; which were relabeled for our satisfaction-based analysis.
The heatmap shows the performance evaluation of the model using confusion matrix in figure 3. It captured the correct and incorrect predictions of each class. The class performance evaluation is thus as follows:

Recall (Probability of detection): $\frac{TP}{TP + FN}$

For Happy; R = $733/733 + 98 = 733/831 = 0.6928$

For Neutral; R = $283/283 + 250 = 283/533 = 0.5309$

For Sad R = $182/182 + 366 = 182/548 = 0.3321$

Precision:

P = $\frac{TP}{TP+FP}$

For Happy; R = $733/733 + 421 = 733/1154 = 0.6352$

For Neutral; R = $283/283 + 202 = 283/485 = 0.5835$

For Sad; R = $182 / 273 = 0.6666$

F1 Score:

F1 = 2 x $\frac{P \times R}{P + R}$

For Happy; F1 = $2 \times {}^{0.6357 \times 0.8928}/_{0.6357 + 0.8928} = 0.7426$

For Neutral; F1 = $2 \times {}^{0.5835 \times 0.5309}/_{0.5835 + 0.5309} = 0.5559$

For Sad; F1 = $2 \times {}^{0.5667 \times 0.3321}/_{0.5667 + 0.3321} = 0.4433$

The ‘Happy’ class has a low value for precision and a high value for recall which indicates that there are more false alerts on the prediction of this class, the ‘Sad’ class has a low recall and a moderate precision value which indicates that this class was mostly predicted as ‘Happy’

or ‘Neutral’; the neutral class was moderately predicted in its recall and precision values. These values are tabularized in table 1.

Also, from the heatmap of the confusion matrix in figure 3, the correct predictions include 733 + 283 + 182 = 1198 and the total predictions equal 1912. Therefore, the overall accuracy equals 1198 / 1912 = 0.6267 which translates to 62.7%. Observe that we have some misclassifications that is calculated as 1912 – 1198 = 714 which means that the error rate will be 100 - 62.7 = 37.3

The model has a significant bias on predicting the ‘Happy’ class. It is observed from figure 3 that in 227 cases ‘Sad is misclassified as ‘Happy’, and in 194 cases ‘Neutral is classified as ‘Happy’.

Accuracy refers to the ratio between correct predictions and the total number of predictions in the training data; which means a higher value for accuracy indicates a positive performance of the model. This is shown in figure 4 where the value of accuracy of the model increases as the epoch increases. The loss value over the training data after each epoch is shown in figure 5; indicating the reduction of the loss value as the epoch increases. The graphs in figure 3 and 4 shows these values on the train and validation datasets performed for 25 epochs.

The model was tested with live data showing how it accurately detects a satisfied, unsatisfied and a neutral emotion via facial emotion recognition as shown in figures 6, 7, and 8.

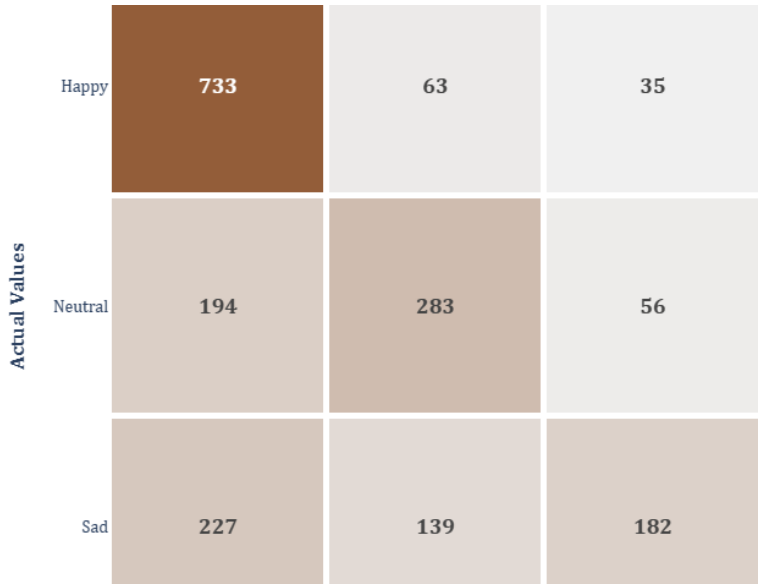


Figure 3: The Heatmap of the Confusion Matrix

Table 1: Precision, Recall and F1 Score

	Precision	Recall	F1-Score
Happy	0.6352	0.6928	0.7426
Neutral	0.5835	0.5309	0.5559
Sad	0.6666	0.3321	0.4433

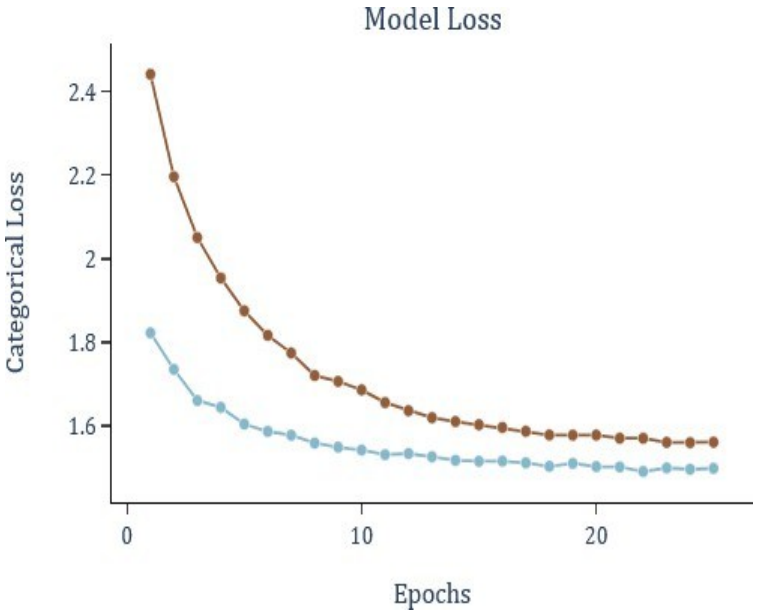


Figure 4: Graph showing Model Loss

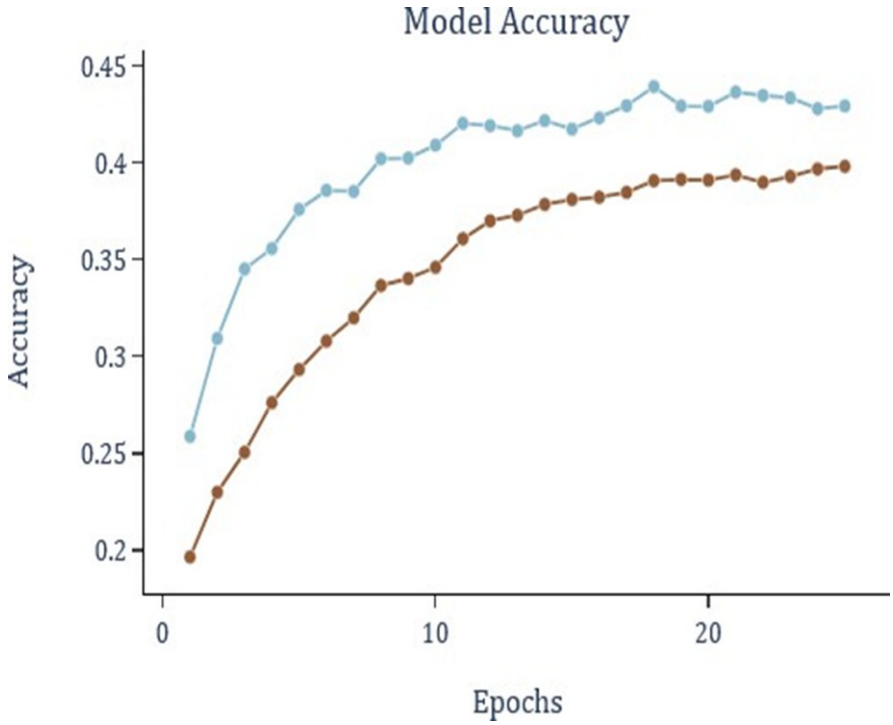


Figure 5: Graph Showing Model Accuracy

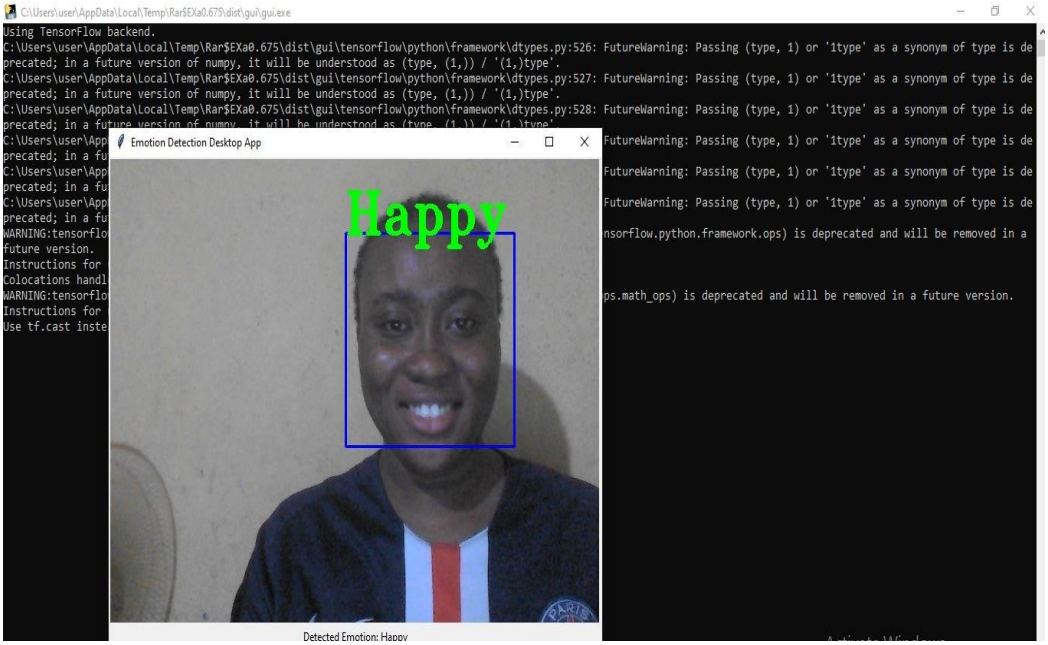


Figure 6: Detection of Happy or Satisfied Emotion

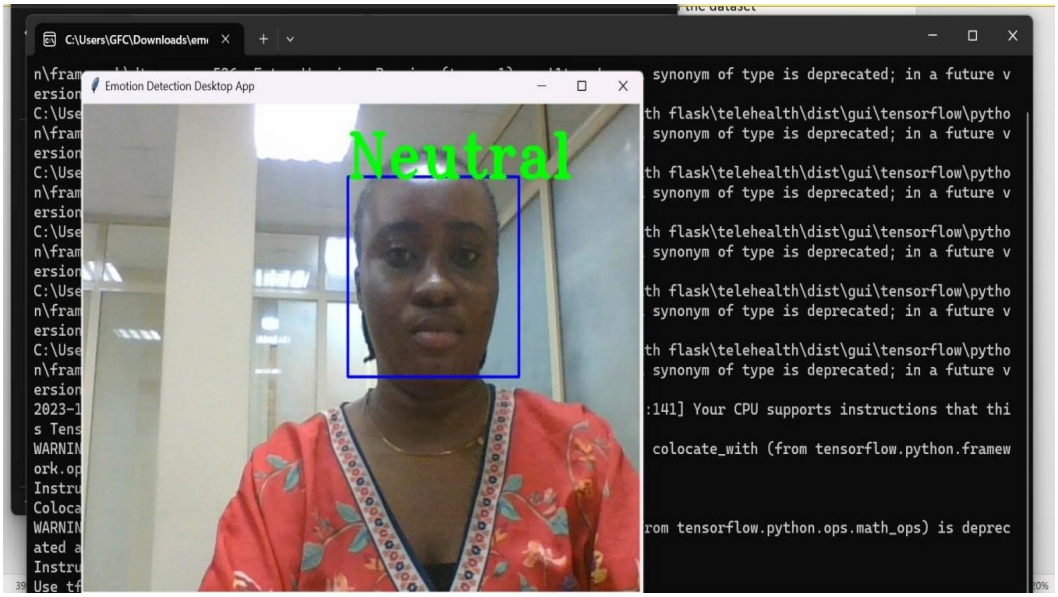


Figure 7: Detection of Neutral Emotion

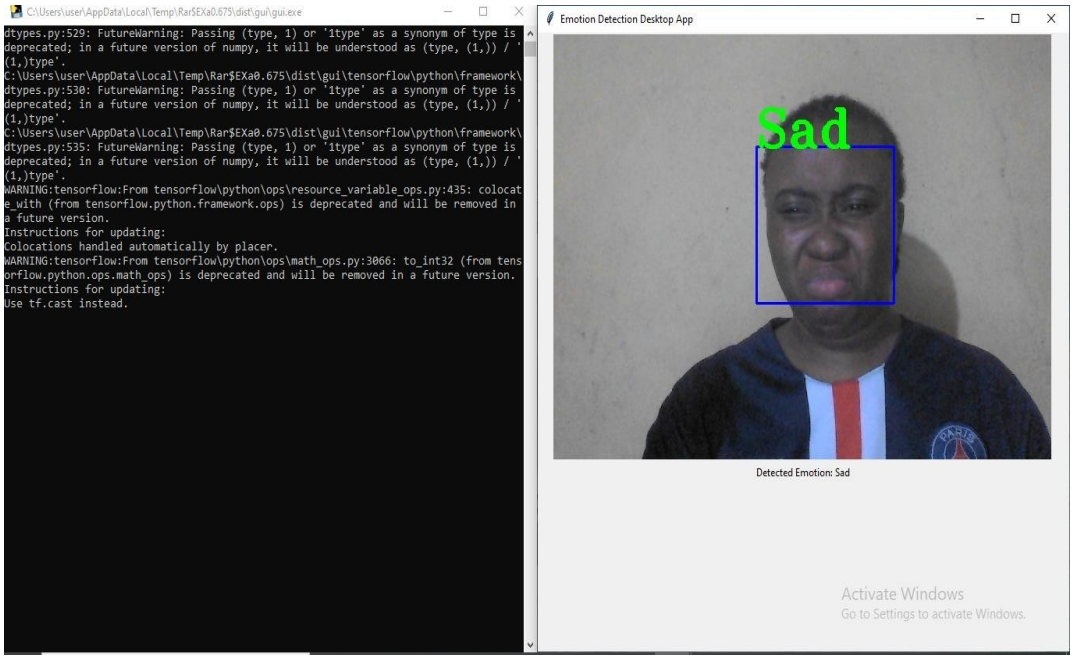


Figure 8: Detection of Sad or Dissatisfied Emotion

Conclusion and Future Work

In this paper we have built and implemented a remote patient' feedback for a supposedly telehealth system by using Convolutional Neural Network model. The results show expressly the alignment with Altun-Dogan and Dogan (2024) that emphasized the importance of patient-centric design in AI-assisted healthcare system. Also, it has provided a springboard for the implementation of AI systems in healthcare with unstructured dataset thereby strengthening the effectiveness of CNN models. The model achieved a 65.78% accuracy and a descending model loss value approximately 1.4 as the epoch increases. The distribution of samples available in the CK+ dataset is observed to be lopsided. The "sad" emotion has the least number of data sample, while the "happy" has the largest number of data samples. This translates to mean that the model would be more likely to predict a „happy" facial expressions correctly than detecting or predicting „sad" facial expressions. The higher the number of a specific class of data samples in the database that the model will be trained on, the more biased the model will be in its prediction.

Therefore, a more balanced and evenly distributed class of dataset should be used in future research and experiment in order to achieve a better accuracy value. Also, the adoption of federated learning for privacy-preserving implementation.

References

Al-Shorbaji, N. (2021). Improving Healthcare Access through Digital Health: The Use of Information and Communication Technologies. *Healthcare Access*, *10*.

Altun-Dogan, M. and Dogan, A., 2024. Patient-centered healthcare: A framework for analyzing patient feedback using CNN and LSTM. In *Artificial Intelligence Applications in Healthcare*. Springer, pp.145–162.

Andrikopoulou, E. (2023). The rise of AI in telehealth. In *Emerging Practices in Telehealth* pp.183-207. Academic Press.

Cascella L. (2018). Virtual risk: an overview of tele-health from a risk management perspective, 5(2), 128-137.

Khaleghparast, S., Gharib, M., & Gharib, M. (2023). Development of a patients' satisfaction analysis system using machine learning and lexicon-based methods. *BMC Health Services Research, 23*(1), 1–12. https://doi.org/10.1186/s12913-023-09260-7

Liu, Y., et al. (2024). Affective Computing for Healthcare: Recent Trends, Applications, Challenges and Beyond. [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2402.13589

Shingjergji, K., Iren D., Urlings C., Klemke, R. (2026). Affective computing in online higher education: A systematic literature review, Computers and Education: Artificial Intelligence, Volume 10, 100499, ISSN 2666-920X, https://doi.org/10.1016/j.caeai.2025.100499.

Ortega, G., Rodriguez, J. A., Maurer, L. R., Witt, E. E., Perez, N., Reich, A., & Bates, D. W. (2020). Telemedicine, COVID-19, and disparities: policy implications. *Health policy and Technology*, *9*(3), 368-371.

Zhang, Y. and Liu, H., 2023. DNN-based hospital service satisfaction using GCNNs learning. *International Journal of Advanced Computer Science and Applications*, 14(6), pp.112–120.