A Deep Learning-Based Approach for Emotion Classification Using Stretchable Sensor Data

Chowdhury Mohammad Masum Refat

[a] *Osaka University, 1-1 Yamadaoka, Suita, 565-0871, Osaka, Japan*

Abstract

Facial expressions play a vital role in human communication, especially for individuals with motor impairments who rely on alternative interaction methods. This study presents a deep learning-based approach for real-time emotion classification using stretchable strain sensors integrated into a wearable system. The sensors, fabricated with conductive silver ink on a flexible Tegaderm substrate, detect subtle facial muscle movements. Positioned strategically on the forehead, upper lip, lower lip, and left cheek, these sensors effectively capture emotions such as happiness, neutrality, sadness, and disgust. A data pipeline incorporating Min-Max normalization and SMOTE balancing addresses noise and class imbalances, while dimensionality reduction techniques like PCA and t-SNE enhance data visualization. The system's classification performance was evaluated using standard machine learning metrics, achieving an overall accuracy of 76.6%, with notable success in distinguishing disgust (86.0% accuracy) and neutrality (81.0% accuracy). This work offers a flexible, cost-effective, and biocompatible solution for emotion recognition, with potential applications in rehabilitation robotics, assistive technologies, and human-computer interaction.

Keywords:
Facial Expression Recognition, Emotion Classification, Stretchable Sensors, Deep Learning, Wearable Technology, Rehabilitation Robotics.

1. Introduction

Facial expressions are fundamental to human communication, conveying emotions and intentions that are critical for social interaction [1, 2, 3, 4]. In recent years, recognizing and interpreting these expressions has gained prominence in fields such as rehabilitation robotics [5, 6, 7], assistive technologies [8], and affective computing [9]. For individuals with motor impairments or communication difficulties, systems that translate facial expressions into actionable control signals offer significant potential to enhance quality of life. However, achieving accurate and reliable facial expression recognition requires sensor technologies that can effectively capture subtle muscle movements while remaining flexible, comfortable, and easy to integrate into wearable systems [10].

Traditional methods for facial expression recognition have relied on several sensing technologies, each with notable limitations. Rigid strain sensors, though capable of detecting mechanical deformation [11], lack the flexibility to conform to the dynamic contours of the human face [12]. This limits their ability to capture real-time facial movements accurately. Computer vision, which detects expressions without physical contact, are highly sensitive to variations in lighting, occlusions, and head movements, reducing their reliability in practical settings [13, 10]. Electromyography (EMG) electrodes provide high-resolution data on muscle activity but require precise placement, conductive gels, and complex setups, making them cumbersome and uncomfortable for prolonged use [14, 15, 16].

These limitations highlight the need for a sensing solution that combines flexibility, biocompatibility, sensitivity, and ease of use. To address this gap, we present a new stretchable strain sensor capable of detecting facial movements for real-time emotion classification. The sensor was developed at the Electronics System Laboratory of the International Islamic University Malaysia (IIUM) and leverages changes in electrical resistance caused by mechanical strain [17, 18]. This design enables accurate detection of subtle facial muscle deformations while overcoming the shortcomings of traditional technologies.

The sensor consists of a conductive silver ink trace printed on a Tegaderm substrate, a medical-grade adhesive film known for its flexibility, comfort, and biocompatibility. This stretchable architecture allows the sensor to conform to the skin's surface, maintaining reliable contact during dynamic facial expressions. The fabrication process, utilizing screen-printing techniques, ensures a cost-effective and scalable production method. With a gauge factor of 84 and the ability to withstand up to 20% elongation without losing functionality, the sensor provides the sensitivity and durability required for real-world applications [19, 20].

To optimize facial expression detection, four sensors are strategically positioned on key facial regions: the forehead, upper lip, lower lip, and left cheek. These locations were selected for their ability to exhibit distinct and consistent muscle movements associated with emotions such as happiness, sadness, neutrality, and disgust. The sensors are integrated with an Arduino Mega 2560 microcontroller, which captures analog voltage signals and transmits the data to a computer for real-time processing and analysis.

This study also incorporates a robust data processing pipeline to ensure reliable emotion classification. Techniques such as Min-Max normalization standardize the sensor readings, while Synthetic Minority Over-sampling Technique (SMOTE) addresses class imbalances by generating synthetic data for underrepresented emotions. Dimensionality reduction methods, including Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE), enhance the visualization of sensor features and help identify patterns in the data.

Performance evaluation of the system is conducted using standard machine learning metrics, including accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). Results indicate that the system effectively distinguishes between certain emotions, such as disgust and neutrality, while highlighting challenges in differentiating subtle expressions like happiness and sadness. Visualizations, including scatter plots, heatmaps, and confusion matrices, provide further insights into the system's strengths and areas for improvement.

The stretchable sensor architecture presented in this work offers a promising solution for real-time facial expression recognition. Its combination of flexibility, high sensitivity, cost-effective fabrication, and biocompatibility makes it suitable for applications in rehabilitation robotics, assistive devices, and human-computer interaction. By addressing the limitations of existing methods, this study lays the groundwork for more adaptive and responsive technologies that can enhance communication and interaction for a wide range of users. Future work will focus on expanding the participant pool, refining feature extraction techniques, and exploring advanced machine learning models to improve classification accuracy and robustness.

2. Methods and Materials

2.1. Stretchable Sensor Architecture

The stretchable sensor developed in this study, as depicted in Figure **??**, was designed at the Electronics System Laboratory of the International Islamic University Malaysia (IIUM). This sensor serves as a low-cost, stretchable strain sensor capable of detecting facial movements for applications in rehabilitation robotics. By leveraging changes in electrical resistance due to mechanical strain, the sensor accurately captures subtle deformations of the skin.

2.1.1. Sensor Design and Fabrication

The stretchable sensor consists of a conductive trace of silver ink printed on a Tegaderm substrate, a medical-grade adhesive film known for its flexibility and biocompatibility (Figures 1 A–B). The sensor measures 30 mm in length and 1 mm in width, optimized to maintain conductivity under mechanical deformation. The silver ink, applied through a screen-printing technique, exhibits a bulk resistivity of $2.0 \times 10^{-4}\,\Omega \cdot \mathrm{cm}$ and can tolerate up to 20% elongation without significant degradation in electrical properties.

Terminal pads at both ends of the conductive trace allow secure electrical connections, ensuring stable data acquisition during operation. The use of Tegaderm simplifies the fabrication process by eliminating the need for additional adhesives and enhances the sensor's comfort and safety when applied to human skin. The resulting sensor demonstrates a gauge factor of 84, indicating high sensitivity for detecting subtle facial expressions.

2.1.2. Electrical Integration and Measurement System

The sensor system's electrical architecture is illustrated in Figures 1 C–E. Each sensor is incorporated into a voltage divider circuit with a 10-ohm fixed resistor, allowing for accurate detection of resistance changes when the sensor deforms. These circuits are connected to an Arduino Mega 2560 microcontroller, which acquires analog voltage signals corresponding to strain-induced resistance variations.

Four sensors are strategically positioned and connected in parallel to optimize the detection of facial movements. The sensors are interfaced with the microcontroller through dedicated analog input pins. The microcontroller transmits the acquired data to a laptop via USB for real-time processing and analysis (Figure 1 E).

2.1.3. Performance and Advantages

The developed stretchable sensor offers high flexibility, sustaining up to 20% strain without loss of functionality. The screen-printing method on Tegaderm makes the fabrication process cost-effective by eliminating the need for additional adhesive layers. The sensor's high sensitivity, with a gauge factor of 84, ensures precise detection of subtle skin deformations. Tegaderm's biocompatibility and comfort make the sensor suitable for prolonged use on human skin.

These attributes make the sensor ideal for real-time facial expression recognition in wearable technologies, particularly for rehabilitation robotics and assistive devices. The integration of this sensor system provides a reliable method for capturing and translating facial movements into control signals, enhancing the functionality and responsiveness of robotic systems.

2.2. Data Collection

Facial expression data were collected by placing stretchable sensors on four key locations on the participants' faces: the **forehead**, **upper lip**, **lower lip**, and **left cheek**. These positions were chosen based on their sensitivity to muscle movement during different facial expressions. The study involved **five healthy individuals**, aged between **22 and 30**, representing diverse ethnic backgrounds, including **Asian** and **African** ethnicities. The participant group consisted of **three males** and **two females**.

2.2.1. Experimental Procedure

Each participant was asked to produce four distinct facial expressions—*happy*, *neutral*, *sad*, and *disgusted*—across three cycles, resulting in a total of **12 expressions per participant**. The data collection sessions were conducted in

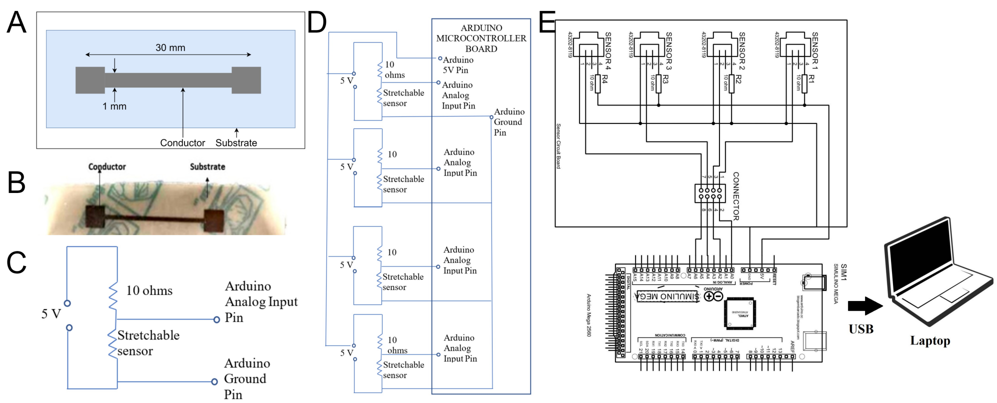

Figure 1: **Stretchable Sensor Design and System Integration.** **(A)** Schematic diagram of the stretchable sensor showing a conductor embedded on a flexible substrate, with a conductor length of 30 mm and width of 1 mm. **(B)** Photograph of the developed stretchable sensor, illustrating the physical structure of the conductor and substrate. **(C)** Circuit diagram of a single stretchable sensor connected to a 10-ohm resistor and an Arduino analog input pin for voltage measurement. **(D)** Expanded circuit schematic showing four stretchable sensors, each connected to a 10-ohm resistor and interfacing with analog input pins of the Arduino microcontroller. **(E)** System setup depicting four sensors interfaced with an Arduino Mega microcontroller, which sends data via USB to a laptop for analysis and visualization.

a controlled environment with consistent lighting and minimal external distractions to ensure the accuracy of sensor readings. Participants were seated comfortably, and each facial expression was held for a set duration to ensure consistent data capture across all participants.

2.2.2. Data Pre-Processing

Following data collection, the raw sensor data were pre-processed to remove any noise and normalize the readings. This was done to enhance the accuracy of the subsequent analysis and system validation. The pre-processed data were then used to test the facial expression recognition algorithm and to evaluate the system's ability to adjust the rehabilitation robot's motor speed based on the detected expressions.

2.2.3. Ethical Approval and Consent

Ethical approval for this study was obtained from the **IIUM Research Ethics Committee (IREC)** under approval number **IREC 2021-301**. All participants provided **informed consent** prior to participation, and measures were taken to ensure their safety and comfort throughout the study.

Lastly,while the small sample size limits the generalizability of the findings, the study provides a preliminary validation of the proposed system. Future research should aim to include a **larger and more diverse participant pool** to further validate the system's performance across different populations.

2.3. Data Visualization

The dataset was collected using four sensors placed on key regions of the face: **Sensor 1** on the forehead, **Sensor 2** on the upper lip, **Sensor 3** on the lower lip, and **Sensor 4** on the left cheek (Figure 2) A. This sensor placement captures both subtle and pronounced facial movements, facilitating the analysis of expression-related data for the following emotions: **0.0 (Disgust), 1.0 (Happy), 2.0 (Neutral), and 3.0 (Sad)**.

The class distribution (Figure 2) B shows a balanced dataset, with an approximately equal number of samples for each emotion. This balance ensures that the training process is not biased toward any particular emotion class, supporting reliable model performance evaluation.

Figure 2 C presents a pairwise plot of the sensor features (S1, S2, S3, and S4) across the different emotion classes. The diagonal plots display the distributions of individual features, while the off-diagonal scatter plots highlight feature correlations. Distinct clustering patterns are visible:

- **S1 (forehead sensor)** and **S4 (left cheek sensor)** show clear separation, particularly for **Sad (3.0)** and **Disgust (0.0)**, indicating these features are effective in distinguishing these emotions.
- **S2 (upper lip sensor)** and **S3 (lower lip sensor)** exhibit more overlap, suggesting these features may capture similar movements, especially for **Happy (1.0)** and **Neutral (2.0)**.

The scatter plot of the first two features (Figure 2) D reveals class separability, with distinct clusters for certain emotions. For example, **Happy (1.0)** is concentrated in the lower range of Feature 1, while **Neutral (2.0)** shows greater dispersion. This indicates that these features provide useful, though not perfect, separation for emotion classification.

Figure 2 E shows the histogram of the first feature, where the majority of data points are concentrated in the lower range. This skewed distribution suggests that subtle facial movements, particularly for **Neutral (2.0)** and **Disgust (0.0)**, dominate this feature.

Principal Component Analysis (PCA) was applied to reduce the dataset's dimensionality (Figure 2 E). The resulting scatter plot shows that the dataset maintains emotion class separability even in lower dimensions. Distinct clusters are evident, especially for **Sad (3.0)** and **Disgust (0.0)**, indicating that meaningful information is preserved after dimensionality reduction.

Figure 2 F presents a heatmap of sensor data, showing the magnitude and variance of sensor readings over time. Variations in intensity reveal patterns of facial movements corresponding to different emotions. Notable spikes in intensity reflect pronounced expressions like **Disgust (0.0)** and **Sad (3.0)**, while consistent low-intensity regions likely correspond to **Neutral (2.0)** states.

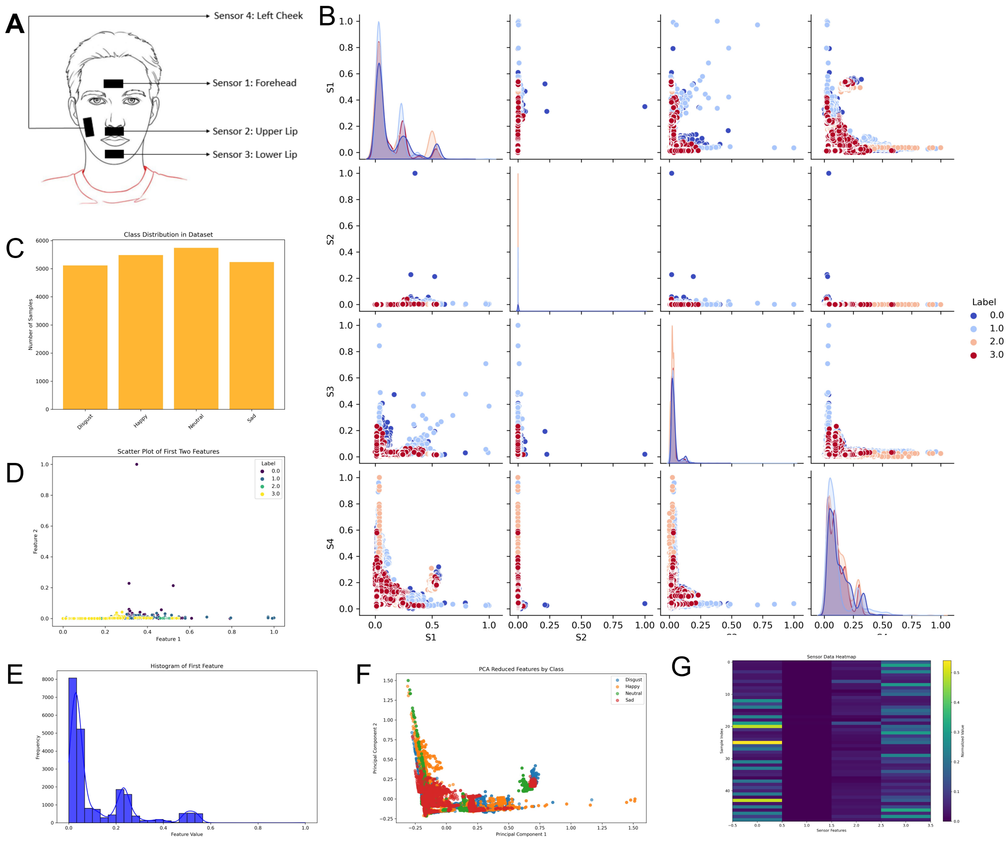

Figure 2: **Sensor Configuration and Data Analysis for Emotion Classification** **(A)** Sensor placement on the face for capturing facial muscle data: Sensor 1 (Forehead), Sensor 2 (Upper Lip), Sensor 3 (Lower Lip), and Sensor 4 (Left Cheek). **(B)** Pairwise scatter plot matrix showing the distribution of features (S1–S4) across different emotion classes labeled as 0.0 (Disgust), 1.0 (Happy), 2.0 (Neutral), and 3.0 (Sad). **(C)** Bar chart representing the class distribution in the dataset, showing a balanced number of samples across all emotion classes. **(D)** Scatter plot of the first two features highlighting the separation between emotion classes. **(E)** Histogram of the first feature showing the distribution of feature values. **(F)** PCA (Principal Component Analysis) plot representing reduced features colored by emotion class, illustrating class clustering patterns. **(G)** Heatmap of sensor data indicating feature magnitudes across the dataset for each emotion class.

These visualizations provide a comprehensive understanding of the dataset, demonstrating balanced emotion distribution, feature correlations, and class separability. The insights gained support the development of reliable machine learning models for classifying facial expressions.

2.4. Feature Normalization and Balancing Methods

To ensure consistent feature scaling, we employed **Min-Max normalization**, which scales each feature x_i to a specified range $[0, 1]$ using the formula:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where x is the original feature value, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the feature, respectively. This normalization ensures that all features contribute proportionally to the model's training process.

To address the issue of **class imbalance**, we applied the **Synthetic Minority Over-sampling Technique (SMOTE)**. SMOTE generates synthetic samples for minority classes by interpolating between existing samples. Given two minority class feature vectors x_i and x_j, a new synthetic sample x_{new} is generated as follows:

$$x_{\text{new}} = x_i + \lambda(x_j - x_i) \tag{2}$$

where λ is a random value in the range $[0, 1]$. This technique helps balance the dataset, improving model generalization by preventing bias toward the majority class.

For **dimensionality reduction** and visualization, we used **t-Distributed Stochastic Neighbor Embedding (t-SNE)** and **Uniform Manifold Approximation and Projection (UMAP)**. t-SNE minimizes the Kullback-Leibler divergence between high-dimensional and low-dimensional data distributions by minimizing:

$$\text{KL}(P||Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{3}$$

where p_{ij} represents the pairwise similarities in the high-dimensional space and q_{ij} represents the pairwise similarities in the low-dimensional space. Similarly, UMAP preserves the local structure by approximating the fuzzy topological structure of the data manifold.

These methods collectively enhance feature representation, ensuring robust model performance and interpretable results.

2.5. Performance Evaluation

To evaluate the performance of the proposed model, we employed standard classification metrics, including **accuracy**, **precision**, **recall**, **F1-score**, and the **area under the ROC curve (AUC)**. These metrics are defined as follows:

2.5.1. Accuracy

Accuracy measures the ratio of correctly predicted samples to the total number of samples and is given by:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where TP represents true positives, TN true negatives, FP false positives, and FN false negatives. While accuracy provides an overall measure, it may be less reliable for imbalanced datasets.

2.5.2. Precision

Precision, also known as the positive predictive value, measures the proportion of correctly predicted positive samples among all predicted positives:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

A high precision indicates a low rate of false positives, making it particularly important in applications where false positives are costly.

2.5.3. Recall

Recall, or sensitivity, measures the proportion of correctly predicted positive samples out of all actual positives:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

High recall indicates a low rate of false negatives, which is crucial in applications where missing a positive case is detrimental.

2.5.4. F1-Score

The F1-score is the harmonic mean of precision and recall, providing a balanced measure of both metrics:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

This metric is particularly useful when the dataset is imbalanced, as it accounts for both false positives and false negatives.

2.5.5. Area Under the ROC Curve (AUC)

The ROC (Receiver Operating Characteristic) curve plots the true positive rate (TPR) against the false positive rate (FPR) across different decision thresholds. The AUC quantifies the overall ability of the model to discriminate between positive and negative classes and is defined as:

$$\text{AUC} = \int_0^1 \text{TPR}(t)\, d\text{FPR}(t) \tag{8}$$

where the **true positive rate (TPR)** and **false positive rate (FPR)** are given by:

$$\text{TPR} = \frac{TP}{TP + FN}, \quad \text{FPR} = \frac{FP}{FP + TN} \tag{9}$$

A higher AUC indicates better model performance, with values closer to 1.0 signifying strong discrimination ability.

2.5.6. Computational Efficiency

To measure computational efficiency, we evaluated the **training time** and **inference time**. The training time T_{train} represents the total time taken to train the model for E epochs:

$$T_{\text{train}} = \sum_{e=1}^{E} t_e \tag{10}$$

where t_e is the time taken for each epoch. Inference time $T_{\text{inference}}$ is the average time required to classify a single sample.

These evaluation metrics provide a comprehensive understanding of the model's performance and efficiency, ensuring that the results are reliable and interpretable.

3. Results

3.1. Confusion Matrix Analysis

The classification performance across the four emotional states—*Disgust*, *Happy*, *Neutral*, and *Sad*—is summarized in the confusion matrix (Figure **A**). The overall model accuracy is **76.6%**, calculated as the proportion of correctly predicted samples out of the total. The *Disgust* class had the highest accuracy, with **988 out of 1149 samples correctly classified** (accuracy: **86.0%**). Misclassifications included 59 samples as *Happy*, 21 as *Neutral*, and 81 as *Sad*. The *Happy* class achieved **71.8% accuracy**, with **824 out of 1148 samples correctly classified**, though 111 samples were misclassified as *Disgust*, 54 as *Neutral*, and 159 as *Sad*.

For the *Neutral* class, the model correctly classified **931 out of 1149 samples** (accuracy: **81.0%**), with misclassifications of 54 samples as *Disgust*, 126 as *Happy*, and 38 as *Sad*. The *Sad* class exhibited the lowest accuracy at **67.6%**, with **776 out of 1148 samples correctly classified** and frequent misclassifications as *Disgust* (255 samples), *Happy* (82 samples), and *Neutral* (35 samples).

3.2. Model Accuracy and Loss

The accuracy and loss curves over 300 epochs are depicted in Figures **B** and **C**. The model achieved a **final training accuracy of 70.2%** and a **validation accuracy of 75.4%**. The **training loss** decreased consistently to **0.65**, while the **validation loss** stabilized at **0.82**. The divergence between training and validation loss, noticeable after 50 epochs, suggests overfitting, where the model performs better on the training set compared to the validation set.

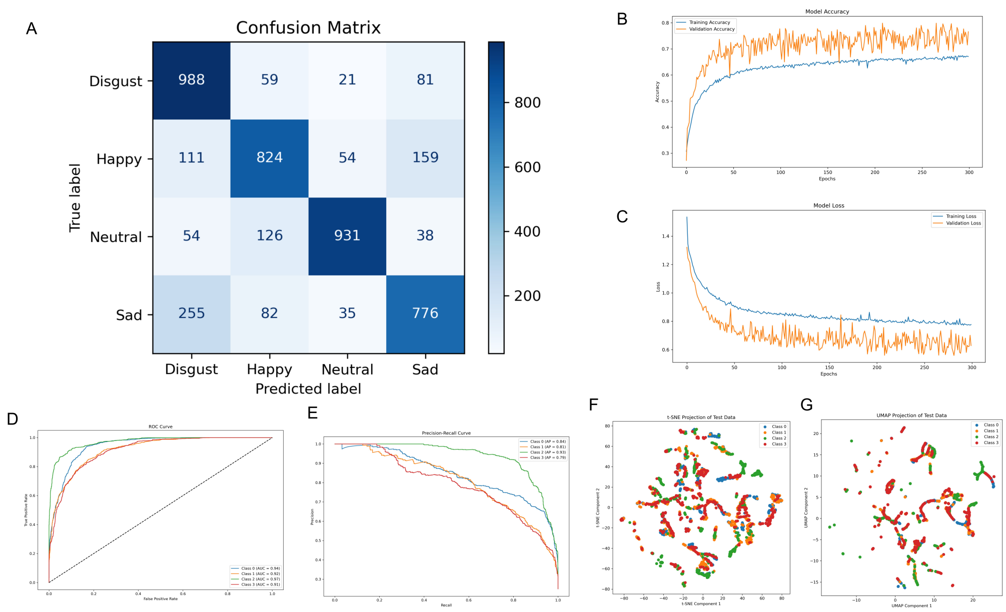

Figure 3: **Model Performance and Visualization for Emotion Classification (A)** Confusion matrix showing the model's performance across the four emotion classes: Disgust (0.0), Happy (1.0), Neutral (2.0), and Sad (3.0). **(B)** Accuracy curves for training and validation datasets across 300 epochs. **(C)** Loss curves for training and validation datasets across 300 epochs. **(D)** ROC (Receiver Operating Characteristic) curves for each class, depicting the trade-off between true positive rate and false positive rate. **(E)** Precision-recall curves for each class, illustrating the balance between precision and recall. **(F)** t-SNE projection of test data, showing the clustering patterns of emotion classes based on the model's learned features. **(G)** UMAP projection of test data, visualizing the separation of emotion classes in lower-dimensional space.

3.3. ROC and Precision-Recall Analysis

The ROC curves (Figure **D**) indicate the model's discriminative power for each class, with the following area-under-the-curve (AUC) values: **0.864** for *Disgust*, **0.807** for *Happy*, **0.911** for *Neutral*, and **0.814** for *Sad*. The *Neutral* class demonstrates the highest AUC, reflecting robust performance, while the *Happy* class shows the lowest AUC, indicating challenges in distinguishing this emotion.

Precision-recall curves (Figure **E**) further highlight these trends. The model exhibits an average precision of **0.861** for *Disgust*, **0.792** for *Happy*, **0.906** for *Neutral*, and **0.803** for *Sad*. The higher precision for the *Neutral* class indicates fewer false positives compared to the *Happy* and *Sad* classes.

3.4. Dimensionality Reduction Visualization

The t-SNE and UMAP projections (Figures **F** and **G**) visualize the test data distribution in lower-dimensional space. In the t-SNE projection, *Neutral* and *Disgust* form distinct clusters, whereas *Happy* and *Sad* show substantial overlap. The UMAP projection provides better separation between *Neutral* and *Disgust*, though overlap between *Happy* and *Sad* remains pronounced. These visualizations support the confusion matrix results, confirming that *Happy* and *Sad* are harder to distinguish.

3.5. Summary of Classification Metrics

Key classification metrics for each class are summarized in Table 1.

Table 1: Summary of classification metrics for each emotional class.

Metric	**Disgust**	**Happy**	**Neutral**	**Sad**
Accuracy (%)	86.0	71.8	81.0	67.6
AUC	0.864	0.807	0.911	0.814
Precision	0.861	0.792	0.906	0.803
Misclassified Samples	161	324	218	372

These results demonstrate that the model performs well in identifying *Disgust* and *Neutral* emotions, with accuracies above 80%. In contrast, the

Happy and *Sad* classes exhibit lower performance, reflecting higher misclassification rates. The AUC and precision values corroborate these findings, emphasizing the challenge of distinguishing between *Happy* and *Sad* states.

4. Discussion

The results of this study demonstrate that the classification model effectively identifies emotional states, particularly *Disgust* and *Neutral*, with accuracies of 86.0% and 81.0%, respectively. The overall accuracy of 76.6% reflects solid performance, although challenges remain in differentiating *Happy* and *Sad* emotions, which achieved lower accuracies of 71.8% and 67.6%. These outcomes are supported by the confusion matrix, ROC curves, precision-recall analysis, and dimensionality reduction visualizations, providing a comprehensive understanding of the model's strengths and limitations.

The high accuracy for the *Disgust* and *Neutral* classes can be attributed to the distinct features present in these emotional states, which are captured effectively by the model. This is further supported by the ROC analysis, where the *Neutral* class achieved the highest AUC of 0.911 and an average precision of 0.906, indicating strong discriminative power and low false positive rates. Similarly, the *Disgust* class attained an AUC of 0.864 and a precision of 0.861, confirming the model's robustness in identifying this emotion.

Conversely, the model's performance for the *Happy* and *Sad* classes is comparatively lower, with AUC values of 0.807 and 0.814, respectively. The confusion matrix reveals that *Happy* is often misclassified as *Sad* and *Disgust*, while *Sad* is frequently predicted as *Disgust*. These misclassifications suggest significant overlap in the feature representations of these emotions, which may arise from subtle differences in the underlying data or noise within the dataset. The dimensionality reduction visualizations using t-SNE and UMAP further corroborate this finding, showing noticeable clustering overlap between the *Happy* and *Sad* classes.

The training and validation curves for accuracy and loss indicate that the model converges well but exhibits signs of overfitting, as evidenced by the divergence between training and validation loss after 50 epochs. The final training accuracy of 70.2% compared to the validation accuracy of 75.4%

suggests that while the model generalizes reasonably well, additional regularization techniques such as dropout, data augmentation, or early stopping could further enhance its performance.

One possible explanation for the observed classification discrepancies is the inherent ambiguity in distinguishing between *Happy* and *Sad* states, particularly when subtle variations in features contribute to these emotions. Additionally, the dataset's balance and quality play a critical role in model performance. While *Disgust* and *Neutral* appear to have clear distinguishing characteristics, *Happy* and *Sad* may require more refined feature extraction methods or larger datasets to achieve better separation. Incorporating multimodal data, such as physiological signals or contextual information, could provide additional cues for improving classification accuracy.

Future work should focus on addressing these challenges by exploring advanced techniques, including deep learning architectures, ensemble methods, and feature selection strategies. Moreover, applying transfer learning from larger emotion recognition datasets or incorporating unsupervised pretraining may help mitigate the limitations observed with the *Happy* and *Sad* classes.

Lastly, the current study demonstrates that while the model performs well in classifying *Disgust* and *Neutral* emotions, significant challenges remain for *Happy* and *Sad* states due to feature overlap and potential dataset limitations. These insights provide a foundation for future improvements in emotion classification systems, with the potential to enhance real-world applications such as human-computer interaction, mental health assessment, and affective computing.

5. Conclusion

This study presents an emotion classification using stretchable strain sensors integrated into wearable systems and a deep learning-based classification model. The proposed sensor, fabricated with conductive silver ink on a flexible Tegaderm substrate, effectively captures subtle facial muscle deformations, overcoming the limitations of traditional rigid sensors, computer vision techniques, and EMG-based methods. By strategically placing sensors on the forehead, upper lip, lower lip, and left cheek, the system accurately detects emotions such as disgust, happiness, neutrality, and sadness. The data processing pipeline, incorporating Min-Max normalization, SMOTE for class balancing, and dimensionality reduction methods like PCA and t-SNE, ensures reliable feature extraction and visualization. The system achieved an overall accuracy of 76.6%, with particularly high performance in identifying disgust (86.0%) and neutrality (81.0%). However, distinguishing between happiness and sadness remains challenging due to overlapping feature patterns, highlighting the need for refined feature extraction methods and larger datasets. This work provides a flexible, biocompatible, and cost-effective solution suitable for applications in rehabilitation robotics, assistive technologies, and human-computer interaction. The stretchable sensor design offers potential for prolonged use, enhancing comfort and practicality in real-world scenarios. Future research will focus on increasing the participant pool, improving sensor placement strategies, and exploring advanced machine learning techniques such as transfer learning and ensemble methods to further enhance classification accuracy and robustness. By addressing the limitations of current emotion recognition technologies, this study contributes to the development of adaptive systems that can improve communication and interaction for individuals with motor impairments and other conditions requiring assistive technologies.

Declarations

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work report in this paper.

- Conflict of interest/Competing interests The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work report in this paper.

- Ethics approval and consent to participate Ethical approval for this study was obtained from the **IIUM Research Ethics Committee (IREC)** under approval number **IREC 2021-301**. All participants provided **informed consent** prior to participation, and measures were taken to ensure their safety and comfort throughout the study.

- Data availability The datasets supporting the findings of this study are available in the Figshare repository with the DOI: 10.6084/m9.figshare.27986789.

References

[1] E. G. Krumhuber, L. I. Skora, H. C. H. Hill, et al., The role of facial movements in emotion recognition, Nature Reviews Psychology 2 (5) (2023) 283–296. doi:10.1038/s44159-023-00172-1.

[2] L. V. Hadley, G. Naylor, A. F. d. C. Hamilton, A review of theories and methods in the science of face-to-face social interaction, Nature Reviews Psychology 1 (1) (2022) 42–54. doi:10.1038/s44159-021-00008-w.

[3] A. Tcherkassof, D. Dupré, The emotion–facial expression link: evidence from human and automatic expression recognition, Psychological Research 85 (8) (2021) 2954–2969. doi:10.1007/s00426-020-01448-4.

[4] M. Calbi, N. Langiulli, F. Ferroni, et al., The consequences of covid-19 on social interactions: an online study on face covering, Scientific Reports 11 (1) (2021) 2601. doi:10.1038/s41598-021-81780-w.

[5] C. M. A. Ilyas, M. Rehm, K. Nasrollahi, et al., Deep transfer learning in human–robot interaction for cognitive and physical rehabilitation purposes, Pattern Analysis and Applications 25 (3) (2022) 653–677. doi:10.1007/s10044-021-00988-8.

[6] M. Li, Z. Wu, C.-G. Zhao, H. Yuan, T. Wang, J. Xie, G. Xu, S. Luo, Facial expressions-controlled flight game with haptic feedback for stroke rehabilitation: A proof-of-concept study, IEEE Robotics and Automation Letters 7 (3) (2022) 6351–6358. doi:10.1109/LRA.2022.3170214.

[7] K. Tao, J. Lei, J. Huang, Physical integrated digital twin-based interaction mechanism of artificial intelligence rehabilitation robots combining visual cognition and motion control, Wireless Personal Communications (2024). doi:10.1007/s11277-024-11108-0.

[8] M. P. de Freitas, V. A. Piai, R. H. Farias, A. M. R. Fernandes, A. G. de Moraes Rossetto, V. R. Q. Leithardt, Artificial intelligence of things applied to assistive technology: A systematic literature review, Sensors 22 (21) (2022) 8531. doi:10.3390/s22218531.

[9] S. C. Leong, Y. M. Tang, C. H. Lai, C. Lee, Facial expression and body gesture emotion recognition: A systematic review on the use of visual data in affective computing, Computer Science Review 48 (2023) 100545. doi:https://doi.org/10.1016/j.cosrev.2023.100545.
URL https://www.sciencedirect.com/science/article/pii/S1574013723000126

[10] C. M. Masum Refat, N. Zainul Azlan, Stretch sensor-based facial expression recognition and classification using machine learning, International Journal of Computational Intelligence and Applications 20 (02) (2021) 2150010. arXiv:https://doi.org/10.1142/S1469026821500103, doi:10.1142/S1469026821500103.
URL https://doi.org/10.1142/S1469026821500103

[11] Y. Gao, L. Yu, J. C. Yeo, C. T. Lim, Flexible hybrid sensors for health monitoring: Materials and mechanisms to render wearability, Advanced Materials 32 (15) (2020) 1902133. doi:10.1002/adma.201902133.

[12] T. Sun, F. Tasnim, R. T. McIntosh, et al., Decoding of facial strains via conformable piezoelectric interfaces, Nature Biomedical Engineering 4 (9) (2020) 954–972. doi:10.1038/s41551-020-00612-w.

[13] S. Singh, S. Prasad, Techniques and challenges of face recognition: A critical review, Procedia Computer Science 143 (2018) 536–543, 8th International Conference on Advances in Computing Communications (ICACC-2018). doi:https://doi.org/10.1016/j.

procs.2018.10.427.
URL https://www.sciencedirect.com/science/article/pii/S1877050918321252

[14] M. Balconi, A. Bortolotti, L. Gonzaga, Emotional face recognition, emg response, and medial prefrontal activity in empathic behaviour, Neuroscience Research 71 (3) (2011) 251–259. doi:https://doi.org/10.1016/j.neures.2011.07.1833.
URL https://www.sciencedirect.com/science/article/pii/S0168010211020347

[15] E. Clancy, E. Morin, R. Merletti, Sampling, noise-reduction and amplitude estimation issues in surface electromyography, Journal of Electromyography and Kinesiology 12 (1) (2002) 1–16. doi:https://doi.org/10.1016/S1050-6411(01)00033-5.
URL https://www.sciencedirect.com/science/article/pii/S1050641101000335

[16] G. Lu, J.-S. Brittain, P. Holland, J. Yianni, A. L. Green, J. F. Stein, T. Z. Aziz, S. Wang, Removing ecg noise from surface emg signals using adaptive filtering, Neuroscience Letters 462 (1) (2009) 14–19. doi:https://doi.org/10.1016/j.neulet.2009.06.063.
URL https://www.sciencedirect.com/science/article/pii/S0304394009008593

[17] M. A. Mohd Asri, W. C. Mak, S. A. Norazman, A. N. Nordin, Low-cost and rapid prototyping of integrated electrochemical microfluidic platforms using consumer-grade off-the-shelf tools and materials, Lab Chip 22 (2022) 1779–1792. doi:10.1039/D1LC01100F.
URL http://dx.doi.org/10.1039/D1LC01100F

[18] M. A. Mohd Asri, A. N. Nordin, N. Ramli, Low-cost and cleanroom-free prototyping of microfluidic and electrochemical biosensors: Techniques in fabrication and bioconjugation, Biomicrofluidics 15 (6) (2021) 061502. arXiv:https://pubs.aip.org/aip/bmf/article-pdf/doi/10.1063/5.0071176/13899986/061502_1_online.pdf, doi:10.1063/5.0071176.
URL https://doi.org/10.1063/5.0071176

[19] N. A. Ramli, A. N. Nordin, N. Zainul Azlan, Development of low cost screen-printed piezoresistive strain sensor for facial expressions recognition systems, Microelectronic Engineering 234 (2020) 111440. doi:https://doi.org/10.1016/j.mee.2020.111440. URL https://www.sciencedirect.com/science/article/pii/S0167931720302288

[20] M. A. Mohd Asri, N. A. Ramli, A. N. Nordin, Electrical performance and reliability assessment of silver inkjet printed circuits on flexible substrates, Journal of Materials Science: Materials in Electronics 32 (22) (2021) 16024–16037. doi:10.1007/s10854-021-06152-6.