SCIENTIFIC Publishing Limited

Journal of Intelligent Communication
http://ojs.ukscip.com/index.php/jic

Article

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

Chowdhury Mohammad Masum Refat

Department of Mechanical Engineering, Osaka University, 1-1 Yamadoka, Suita 565-0871, Japan; masumrefat3@gmail.com

Received: 13 January 2025; **Revised**: 21 January 2025; **Accepted**: 5 March 2025; **Published**: 10 March 2025

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 [1–4]. In recent years, recognizing and interpreting these expressions has gained prominence in fields such as rehabilitation robotics [5–7], 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 [10,13]. 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 [14–16]. 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 [21]. 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 [22,23].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. Materials and Methods

2.1. Architecture of Stretchable Sensor

The stretchable sensor developed and fabricated in this study, shown in **Figure 1**, was created at the Electronics System Laboratory of the International Islamic University Malaysia (IIUM). This cost-effective strain sensor is designed to detect facial movements and is intended for use in rehabilitation robotics. It functions by measuring variations in electrical resistance caused by mechanical strain, allowing it to precisely track subtle skin deformations [24].

2.1.1. Custom Sensor Design and Fabrication Methods

The stretchable sensor features a conductive silver ink trace printed onto a Tegaderm substrate, a flexible and bio- compatible medical-grade adhesive film (**Figures 1 A–B**). With dimensions of 30 mm in length and 1 mm in width, the sensor is designed to retain conductivity during mechanical deformation. The silver ink, applied via screen printing, has a bulk resistivity of $2.0 \times 10^{-4}\ \Omega \cdot$ cm and maintains stable electrical properties even under 20% elongation.

The conductive trace includes terminal pads at both ends, providing reliable electrical connections for stable data acquisition during operation. Tegaderm serves as both the substrate and adhesive, streamlining the

fabrication process and enhancing user comfort and safety when applied to human skin. The sensor achieves a gauge factor of 84, reflecting its high sensitivity in detecting subtle facial expressions [24].

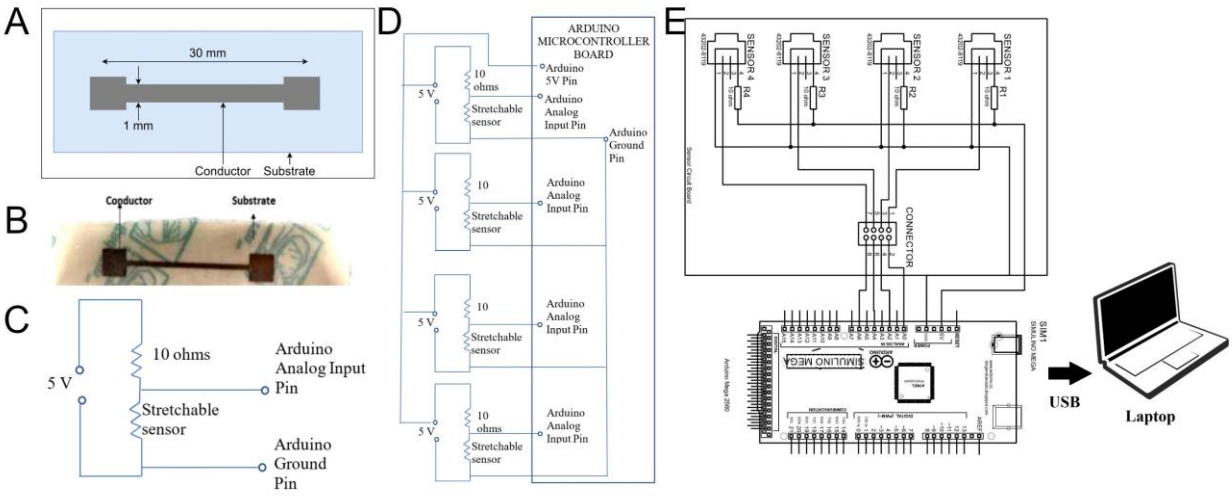

Figure 1. Stretchable sensor design and system integration. (**A**) Schematic diagram of the stretchable sensor showing a conductor embedded on a flexible substrate, with a conductor length of 30 mm and width of 1 mm. (**B**) Photograph of the developed stretchable sensor, illustrating the physical structure 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.

2.1.2. Electrical Circuit and Data Acquisition System

The electrical architecture of the sensor system is depicted in **Figures 1 C–E**. Each sensor is integrated into a voltage divider circuit alongside a 10-ohm fixed resistor, enabling precise detection of resistance changes caused by sensor deformation. These circuits interface with an Arduino Mega 2560 microcontroller, which captures analog voltage signals reflecting strain-induced variations in resistance [24]. Four sensors are strategically placed and connected in parallel to enhance the detection of facial movements. Each sensor is interfaced with the microcontroller through individual analog input pins. The microcontroller sends the collected data to a laptop via USB for real-time processing and analysis (**Figure 1 E**).

2.1.3. Performance and Advantages

The stretchable sensor exhibits excellent flexibility, maintaining functionality even under 20% strain. The use of a screen-printing technique on Tegaderm streamlines the fabrication process, reducing costs by eliminating additional adhesive layers. With a high gauge factor of 84, the sensor delivers exceptional sensitivity for accurately detecting subtle skin deformations. Additionally, Tegaderm's biocompatibility and comfort make the sensor ideal for extended use on human skin. These features make the sensor highly suitable for real-time facial expression recognition in wearable technologies, especially in rehabilitation robotics and assistive devices. By integrating this sensor system, facial movements can be reliably captured and translated into control signals, significantly improving the functionality and responsiveness of robotic systems [24].

2.2. Data Collection

Facial expression data were collected by placing stretchable sensors at four key locations on the participants' faces: the forehead, upper lip, lower lip, and left cheek. These positions were selected for their responsiveness to muscle activity during various facial expressions. The study included five healthy individuals,

aged 22 to 30, representing a range of ethnic backgrounds, including Asian and African ethnicities. The participant group comprised three males and two females [24].

2.2.1. Experimental Procedure

Each participant was instructed to perform four distinct facial expressions—*happy, neutral, sad*, and *disgusted*— repeated across three cycles, yielding a total of 12 expressions per participant. Data collection sessions were conducted in a controlled setting with uniform lighting and minimal external distractions to maintain the accuracy of the sensor readings. Participants were seated comfortably, and each facial expression was sustained for a specified duration to ensure consistent data capture across all individuals [24].

2.2.2. Data Pre-Processing

After data collection, the raw sensor data underwent pre-processing to eliminate noise and normalize the readings. This step was essential to improve the accuracy of subsequent analysis and system validation. The pre-processed data were then utilized to test the facial expression recognition algorithm and assess the system's capability to adjust the rehabilitation robot's motor speed in response to the detected expressions [24].

2.2.3. Ethical Approval and Consent

Ethical approval for this study was granted by the IIUM Research Ethics Committee (IREC) under approval number IREC 2021-301. All participants provided informed consent before participating, and appropriate measures were implemented 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 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 most 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.

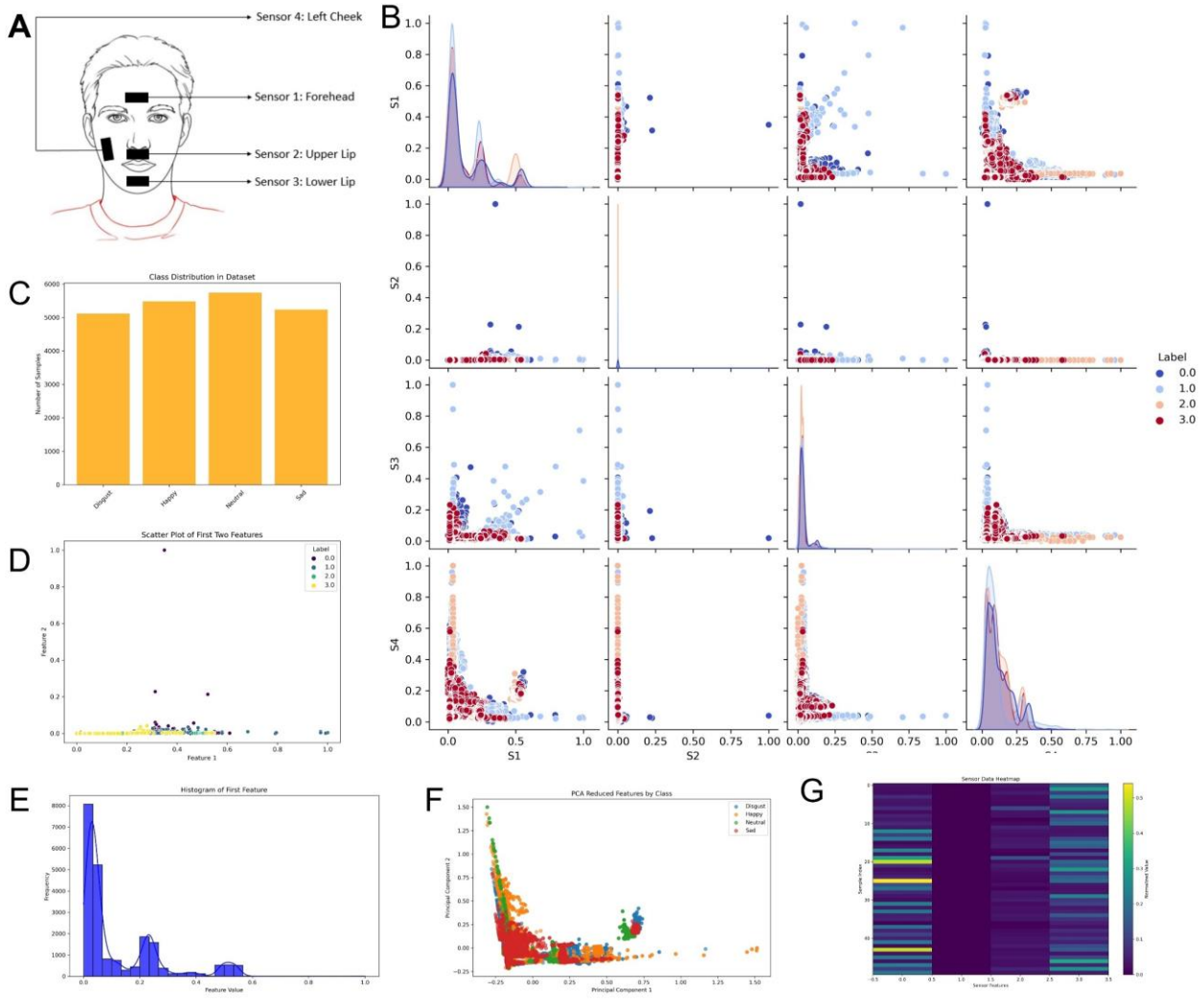

Figure 2. Sensor configuration and data analysis for emotion classification. (**A**) Sensor placement on the face for capturing facial muscle data: Sensor 1 (Forehead), Sensor 2 (Upper Lip), Sensor 3 (Lower Lip), and Sensor 4 (Left Cheek). (**B**) Pairwise scatter plot matrix showing the distribution 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.

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.

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 ∞i to a specified range [0,1] using the formula:

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

where ∞ is the original feature value, ∞min and ∞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 ∞_i and ∞_j, a new synthetic sample ∞_{new} is generated as follows:

$$\infty_{new} = \infty_i + \lambda(\infty_j - \infty_i) \tag{2}$$

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:

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

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:

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

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 means of precision and recall, providing a balanced measure of both metrics:

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

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:

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

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:

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

where t_e is the time taken for each epoch. Inference time $T_{\text{interence}}$ 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 3A**). 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).

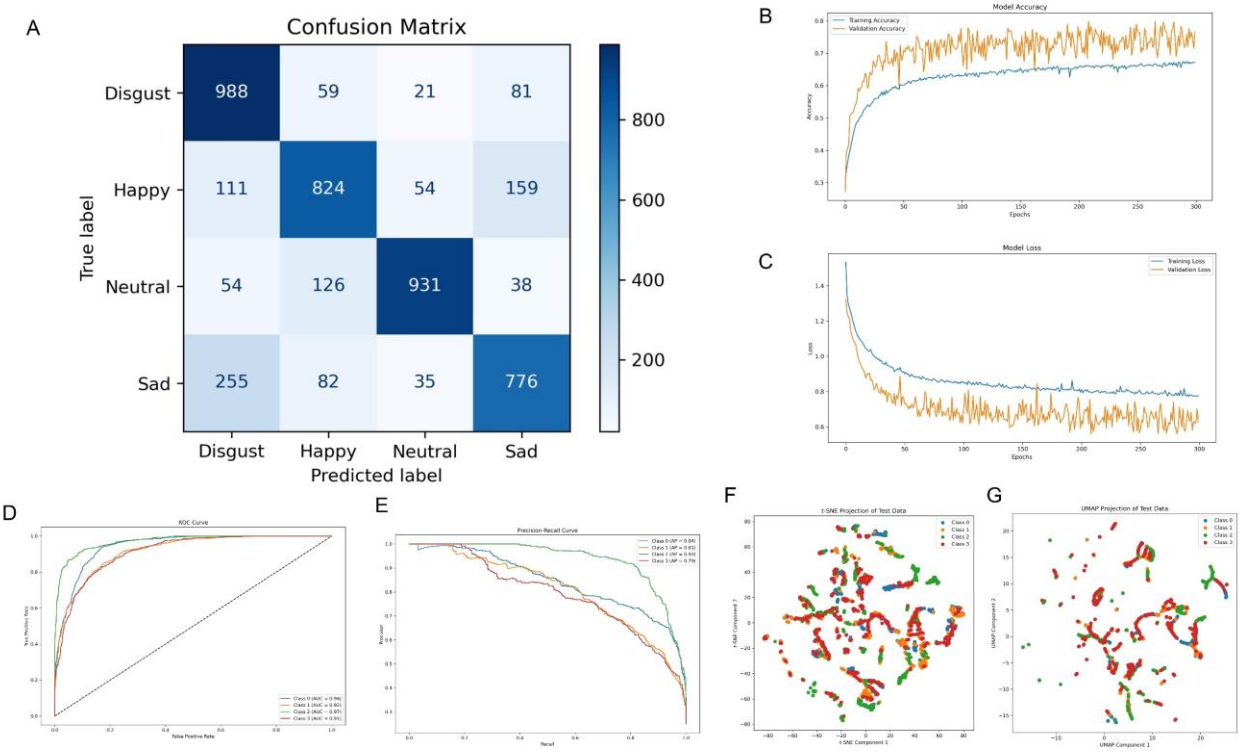

Figure 3. Model performance and visualization for emotion classification. (**A**) Confusion matrix showing the model's performance across the four emotion classes: *Disgust* (0.0), *Happy* (1.0), *Neutral* (2.0), and *Sad* (3.0). (**B**) Accuracy curves for training and validation datasets across 300 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.

3.2. Model Accuracy and Loss

The accuracy and loss curves over 300 epochs are depicted in **Figure 3B**,**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.

3.3. ROC and Precision-Recall Analysis

The ROC curves (**Figure 3D**) 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 (**Figure 3F**,**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 acceptable.

However, to mitigate overfitting and improve model generalization, several strategies can be considered. Regularization techniques, such as adding dropout layers, can reduce reliance on specific neurons, preventing the model from overfitting to training data. Additionally, data augmentation, such as introducing minor temporal shifts or controlled noise to the sensor signals, can increase variability and improve robustness. Early stopping could be applied to terminate training when validation loss stops improving, ensuring that the model does not continue learning noise. Another approach is to optimize model complexity by fine-tuning the number of hidden layers or reducing the number of parameters to prevent unnecessary memorization. Finally, cross-validation (e.g., k-fold cross-validation) can be employed

to assess model reliability across different data splits, ensuring that performance is not overly dependent on a specific subset of the dataset. While expanding the dataset remains the most direct solution, these techniques can significantly improve the model's ability to generalize even with the current sample size.

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 pre-training 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* due to feature overlap and potential dataset limitations. To address this, future improvements could include optimizing sensor placement (e.g., incorporating nasolabial fold sensors), refining feature extraction methods, or integrating advanced classification models that enhance class-specific feature discrimination. While dataset expansion is ideal, our findings indicate that targeted sensor realignment and additional feature engineering could significantly improve the model's ability to differentiate these closely related emotions. 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.

Future Work

This study presents a novel approach for emotion classification using stretchable strain sensors and deep learning techniques. While the proposed system demonstrates promising results, several areas warrant further investigation to enhance its effectiveness and applicability.

• Dataset Expansion and Diversity: One of the main limitations of this study is the small participant pool. Future work should focus on increasing the number of participants and incorporating diverse demographic groups to improve generalizability and ensure robust model performance across different populations.

• Sensor Placement Optimization: The current study utilizes four sensor placements, primarily on the forehead, upper lip, lower lip, and left cheek. Additional sensors on other facial regions, such as the nasolabial folds or jawline, could provide more comprehensive data and improve the classification of emotions with subtle muscle movements, such as *Happy* and *Sad*.

• Multimodal Integration: While stretchable sensors effectively capture muscle activity, integrating other modalities—such as computer vision-based facial analysis, electromyography (EMG), or physiological signals (e.g., heart rate variability, electrodermal activity)—could enhance emotion recognition by providing complementary data sources.

• Improving Model Generalization: The model exhibited signs of overfitting, particularly in distinguishing between *Happy* and *Sad* emotions. Future work could explore regularization techniques, transfer learning, or ensemble learning approaches to improve generalization. Additionally, investigating the use of attention mechanisms in deep learning architectures may enhance feature extraction and classification accuracy.

• Real-Time Implementation and Practical Applications: The current system operates in a controlled experimental setting. Future studies should focus on developing a real-time, wearable prototype that can be integrated into assistive technologies, rehabilitation robotics, and human-computer interaction applications. Evaluating the system in real-world environments will provide valuable insights into its practicality and user acceptance.

• Cross-Cultural and Context-Based Emotion Analysis: Emotion expression varies across cultural backgrounds and situational contexts. Future research should investigate how different cultural groups express emotions through facial muscle movements and adapt the classification model accordingly.

By addressing these challenges, future developments will refine the proposed system, making it more robust, adaptable, and applicable to a wide range of real-world scenarios.

5. Conclusions

This study introduces a deep learning-based approach for real-time emotion classification using stretchable strain sensors, offering a flexible, biocompatible, and cost-effective solution for facial expression recognition. The sensor design, which utilizes conductive silver ink printed on a Tegaderm substrate, enables precise detection of subtle facial muscle movements while maintaining comfort and adaptability for long-term use. By strategically placing sensors on key facial regions, the system effectively classifies emotions such as disgust, happiness, neutrality, and sadness, demonstrating its potential in rehabilitation robotics, assistive technologies, and human-computer interaction. To ensure robust feature extraction and classification, the system incorporates a comprehensive data pipeline, including Min-Max normalization for standardization, SMOTE for class balancing, and dimensionality reduction techniques like PCA and t-SNE for enhanced feature visualization. Performance evaluation yielded an overall accuracy of 76.6%, with notably strong recognition of disgust (86.0%) and neutrality (81.0%). However, the distinction between happiness and sadness remains challenging, likely due to overlapping feature patterns, underscoring the need for further sensor optimization and data expansion. Compared to conventional facial expression recognition methods, such as rigid strain sensors, computer vision-based approaches, and electromyography (EMG), the proposed system offers enhanced adaptability, real-time usability, and practical integration into wearable devices. Its ability to capture dynamic facial movements without requiring extensive calibration or con- trolled lighting conditions make it a promising alternative for assistive communication and emotion-aware technologies. Future work will focus on expanding the dataset to improve model generalizability, optimizing sensor placement to capture a broader range of facial expressions, and exploring advanced machine learning techniques, such as transfer learning and ensemble models, to enhance classification accuracy. Additionally, integrating multimodal approaches, incorporating physiological signals and computer vision techniques—could further refine emotion recognition capabilities. Real-world validation in healthcare, rehabilitation, and human-computer interaction settings will be crucial in assessing the system's usability and impact. By addressing these challenges, this study lays the foundation for next-generation, adaptive emotion recognition systems that can enhance human-computer interaction, accessibility, and assistive communication. The findings contribute to the broader field of affective computing and intelligent wearable technologies, paving the way for more inclusive, responsive, and human-centered applications.

Funding

This work received no external funding.

Institutional Review Board Statement

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 Statement

The datasets supporting the findings of this study are available in the Fig share repository with the DOI: 10.6084/m9.figshare.27986789.

Acknowledgments

I would like to express my sincere gratitude to my master's supervisor, Dr. Norsinnira Bt. Zainul Azlan, for her valuable guidance and support during the data collection phase of my master's research. I also extend my appreciation to my co-supervisor, Dr. Anis Nurashikin Nordin, for her insightful input and encouragement throughout the research process. Their mentorship laid the foundation for the dataset used in this current work.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Krumhuber, E.G.; Skora, L.I.; Hill, H.C.H.; et al. The role of facial movements in emotion recognition. *Nat. Rev. Psychol.* **2023**, *2*, 283–296.
2. Hadley, L.V.; Naylor, G.; Hamilton. A review of theories and methods in the science of face-to-face social interaction. *Nat. Rev. Psychol.* **2022**, *1*, 42–54.
3. Tcherkassof, A.; Dupré, D. The emotion–facial expression link: evidence from human and automatic expression recognition. *Psychol. Res.* **2021**, *85*, 2954–2969.
4. Calbi, M.; Langiulli, N.; Ferroni, F.; et al. The consequences of Covid-19 on social interactions: an online study on face covering. *Sci. Rep.* **2021**, *11*, 2601.
5. Ilyas, C.M.A.; Rehm, M.; Nasrollahi, K.; et al. Deep transfer learning in human–robot interaction for cognitive and physical rehabilitation purposes. *Pattern Anal. Appl.* **2022**, *25*, 653–67.
6. Li, M.; Wu, Z.; Zhao, C.G.; et al. Facial expressions-controlled flight game with haptic feedback for stroke rehabilitation: A proof-of-concept study. *IEEE Robot. Autom. Lett.* **2022**, *7*, 6351–6358.
7. Tao,K.; Lei, J.; Huang, J. Physical integrated digital twin-based interaction mechanism of artificial intelligence rehabilitation robots combining visual cognition and motion control. *Wirel. Pers. Commun.* **2024**.
8. de Freitas, M.P.; Piai, V.A.; Farias, R.H.; et al. Artificial intelligence of things applied to assistive technology: A systematic literature review. *Sensors* **2022**, *22*, 8531.
9. Leong, S.C.; Tang, Y.M.; Lai,C.H.; et al. Facial expression and body gesture emotion recognition: A systematic review on the use of visual data in affective computing. *Comput. Sci. Rev.* **2023**, *48*, 100545.
10. Refat, C.M.M.; Azlan, N.Z. Stretch sensor-based facial expression recognition and classification using machine learning. *Int. J. Comput. Intell. Appl.* **2021**, *20*, 2150010.
11. Gao,Y.; Yu, L.; Yeo,J.C.; et al. Flexible hybrid sensors for health monitoring: Materials and mechanisms to render wearability. *Adv. Mater.* **2020**, *32*, 1902133.
12. Sun,T.; Tasnim, F.; McIntosh, R.T.; et al. Decoding of facial strains via conformable piezo- electric interfaces. *Nat. Biomed. Eng.* **2020**, *4*, 954–972.
13. Singh, S.; Prasad, S.V.A.V. Techniques and challenges of face recognition: A critical review. In Proceedings of the 8th International Conference on Advances in Computing Communications (ICACC-2018). *Procedia Comput. Sci.* **2018**, *143*, 536–543.
14. Balconi, M.; Bortolotti, A.; Gonzaga, L. Emotional face recognition, emg response, and medial prefrontal activity in empathic behaviour. *Neurosci. Res.* **2011**, *71*, 251–259.
15. Clancy, E.A.; Morin, E.L.; Merletti, R. Sampling, noise-reduction and amplitude estimation issues in surface electromyography. *J Electromyogr Kinesiol* **2022**, *12(1)*, 1–16.
16. Lu, G.; Brittain, J.S.; Holland, P. et al. Removing ecg noise from surface emg signals using adaptive filtering. *Neurosci Lett* **2009**, *462(1)*, 14–19.
17. Mohd Asri, M.A.; Mak, W.C.; Norazman, S.A. et al. Low-cost and rapid prototyping of integrated electrochemical microfluidic platforms using consumer-grade off-the-shelf tools and materials. *Lab Chip* **2022**, *22*, 1779–1792.
18. Mohd Asri, M.A.; Nordin, A.N.; Ramli, N. Low-cost and cleanroom-free prototyping of microfluidic and electrochemical biosensors: Techniques in fabrication and bioconjugation. *Biomicrofluidics* **2021**, *15*, 061502.
19. Ramli, N.A.; Nordin, A.N.; Azlan, N.Z. Development of low cost screen- printed piezoresistive strain sensor for facial expressions recognition systems. Microelectron En*g* **2020**, *234*, 111440.
20. Mohd Asri, M.A.; Ramli, N.A.; Nordin, A.N. Electrical performance and reliability assessment of silver inkjet printed circuits on flexible substrates. *J Mater Sci Mater Electron* **2021**, *32(22)*, 16024–16037.
21. Chawla, N.V.; Bowyer, K.W.; Hall, L.O. et al. Smote: synthetic minority over-sampling technique. *J Artif Intell Res* **2002**, *16(1)*, 321–357.
22. Sehgal, S.; Singh, H.; Agarwal, M. et al. Data analysis using principal component analysis. In Proceedings of the 2014 International Conference on Medical Imaging, m-Health and Emerging Commu- nication Systems (MedCom), pp. 45–48.
23. Rogovschi, N.; Kitazono, J.; Grozavu, N. t-distributed stochastic neighbor embedding spectral clustering. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), pp. 1628–1632.
24. Nordin, A.N.; Refat, C.M.M.; Azlan, N.Z. Facial expression-driven rehabilitation robotics: A machine learning approach with stretchable sensors. *Res Sq* **2024**, PREPRINT (Version 1).