Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

E-Taste Signal Encoding and Classification via Multi- Attention Residual Networks Using sEMG Streams in Virtual Sensory Systems

Zhendong Song
Shenzhen Polytechnic University
Asif Ullah
asifkh@szpu.edu.cn

Shenzhen Polytechnic University
Majad Mansoor
University of Namur
You Wang
Zhejiang University

Research Article

Keywords: Electronic taste recognition, sEMG signal, MACB-ResNet model, residual learning, brain-computer interface (BCI)

Posted Date: July 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10132620/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

E-Taste Signal Encoding and Classification via Multi-Attention Residual Networks Using sEMG Streams in Virtual Sensory Systems

Zhendong Song[1,2], Asif Ullah[1,*], Majad Mansoor[3], and You Wang[4]

[1] Institute of Intelligent Manufacturing, Shenzhen Polytechnic University, Shenzhen 518055, China

[2] Guangdong Provincial Key Laboratory of Intelligent Morphing Mechanisms and Adaptive Robotics, Harbin Institute of Technology (Shenzhen), Shenzhen, China 518055

[3]Faculty of Medicine, University of Namur, Belgium

[4] State Key Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310027, China

* Correspondence: asifkh@szpu.edu.cn

Abstract

The purpose of this study was to improve the precision of electronic taste (E-Taste) categorization through the use of surface electromyography (sEMG) data. A multi-attention channel block and residual learning approach, called MACB-ResNet, was developed to emphasize the most important features for distinguishing among the several E-Taste classes. Silver electrodes were used to capture sEMG data from the facial muscles stimulated by various E-Taste experience levels (none, salty, sprite, bitter, mint, sour) while delivering these stimuli to the tip of the tongue. Data were detrended and filtered, then converted to spectrograms. All data obtained by the MACB-ResNet classifier were compared with three other standard classifiers: KNN, Random Forest (RF), and DenseNet-201. Model performance was evaluated using multiple means, including accuracy, F1 Score, area under the curve (AUC), confusion matrix analysis, principal component analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE) visualizations, and an ablation study. The results of this study demonstrate that the MACB-ResNet classifier generated the highest classification accuracy (86.32%) compared to KNN (65.11%), RF (78.35%), and DenseNet-201 (81.93%). In addition to high accuracy scores, the class-level analyses demonstrated robust classification of each taste type, with minimal misclassifications among similar feature classes (i.e., confusion between bitter and sour). PCA and t-SNE visualizations supported MACB-ResNet's ability to separate classes, and ablation studies indicated that both residual connections and multi-attention channel blocks significantly affected model performance. The MACB-ResNet model's ability to improve the accuracy of sEMG-based E-Taste classification will be useful for many future applications in personal nutrition, digital health, food technology, neuromarketing, and brain-computer interface systems.

Keywords: Electronic taste recognition; sEMG signal; MACB-ResNet model; residual learning; brain-computer interface (BCI)

1. Introduction

Taste perception is the foundation for an organism to identify and discern various substances in its environment, a function upon which human physiology heavily depends [1]. Nevertheless, anorexia, diabetes, cancer, and the COVID-19 virus are among the many medical issues that can lead to a loss of taste. [2-5]. Evaluating patients' gustatory perception may facilitate the early diagnosis of various disorders. Additionally, findings on virtual taste in remote multimodal interfaces have gained prominence as virtual reality has evolved. [6, 7]. Cruz et al. were among the pioneers in the field of thermal stimulation of the tongue. They observed that the tongue elicits taste through temperature variation. [8], as subsequently demonstrated by Lawless et al., who investigated metallic-based flavors [9].

Recently, an approach incorporating electrical and thermal stimulation of taste receptors has been employed, enabling the perception of fundamental tastes and flavors. A typical electrical stimulation system consists of a Peltier element for heating or cooling, silver electrodes for conducting electricity, and an additional temperature sensor for temperature feedback [10]. Research has shown that silver electrodes are the best choice for electrical stimulation because they do not produce a metallic taste or otherwise affect taste. Heat- and electrically induced taste sensations are typically very dissimilar. Consequently, both types of stimuli help achieve a fuller taste. Given the examples above, it is worthwhile investigating the measurement of natural and electronic taste sensations for virtual taste, disease detection, and other taste-related brain-computer interfaces (BCIs) [11-13].

Taste sensations depend on tongue receptors, which respond to diverse stimuli and convey information to the cerebral cortex [14, 15]. The vagus, glossopharyngeal, and facial nerves transport the electrical signals generated by taste receptor cells to the brainstem [16]. The facial nerve has two branches: the superficial petrosal and the chorda tympani. The chorda tympani nerve innervates the anterior tongue and transmits information, whereas the superficial petrosal nerve activates the soft palate [17]. These impulses are transmitted from the brainstem's pontomedullary junction to the nucleus solitarius. The glossopharyngeal nerve transmits somatosensory, tactile, and gustatory information from the posterior tongue to the brainstem via the jugular foramen, which synapses at the nucleus solitarius. Impulses from the vagus nerve communicate with the epiglottic area of the body and travel through the medulla and synapse to the nucleus solitarius. After processing taste input, the brain sends motor signals via the facial nerve (VII) to innervate the facial muscles and produce appropriate facial expressions.

Surface electromyography is a non-invasive technique that can compute human facial activity [18]. When muscles contract, a series of electrical impulses called sEMG signals is recorded across the skin. Sensory information from food tasting adversely affects the activation patterns of the suprahyoid muscles during pharyngeal swallowing. This approach may assess taste responses in newborns and those with impaired senses [19]. Examples of sEMG applications include cosmetics, clinical diagnostics, sports engineering, and synthetic electromyography management [20]. The non-invasive sEMG approach, which detects salivary gland activation and facial activity in individuals, has recently gained recognition for assessing muscular fatigue [21-24]. Sensory neurons and muscle fibers generate electrical impulses that the sEMG records. Significantly, sEMG devices are used in speech, gesture recognition, and prosthesis control [25-28]. Numerous studies have examined sEMG attributes during eating and swallowing. According to the author in [29], sEMG is affected by food temperature and carbonation during swallowing.

In contrast, the author of [30] measured and recorded sEMG signals from subjects as they swallowed 18 common food gels, with three textures and six flavors. Furthermore, Manda et al. recorded electromyographic activity of the posterior tongue while subjects chewed soft, semi-solid, and solid foods, and compared free and effortful chewing. While masticating a solid meal, the results demonstrated a considerable increase in chewing cycles [31]. The results showed that chewing cycles increased significantly during free and rigorous mastication of a solid meal. Conversely, no significant difference in mastication was observed between free and harsh semi-solid and soft foods. Nonetheless, neither swallowing nor chewing can be classified as muscle activity in response to taste stimuli. Taste perception occurs only when the taste material is still in the oral cavity or on the tongue.

Previous studies predominantly employed manually designed statistical methods to interpret sEMG signals. However, a more adaptable and valuable classification model is required for taste-related BCI applications, comparable to emotion identification and motor imagery (MI) recognition. Traditional feature engineering and machine learning techniques have shown efficacy [32, 33]. Recent advancements in deep learning have led some researchers to replace hand-extracted features and classifiers with convolutional neural networks (CNNs) and long short-term memory (LSTM) networks for validating sEMG signals [34, 35]. For sEMG-based estimation of simultaneous multi-finger forces, Zanghieri et al. proposed an incremental online training approach using a small Temporal Convolutional Network suitable for embedded training on a device [36]. Ma et al.

implemented a fully connected neural network (FNN) on an unlabeled sEMG dataset derived from muscle force measurements to identify muscle-tendon parameters [37].

Numerous research projects have used CNNs or LSTM architectures in their study of biosignal analysis. Examples include fatigue detection and action recognition [38-40]. Human interaction recognition is another type of biosignal analysis that has used filter-based methods, such as extended Kalman filtering [41]. He et al. employed four deep learning models for gesture recognition, including temporal TCN, CNN, GRU, and Transformer [42]. Sakorn et al. employed a ResNeXt-based model to evaluate sEMG of human knees for abnormality detection, achieving 96.73% accuracy and 93.77% F1 score [43]. The studies indicate that multiscale convolution effectively learns feature representations from sEMG. However, these models' performance is inadequate for direct application to sEMG-based detection of taste sensation.

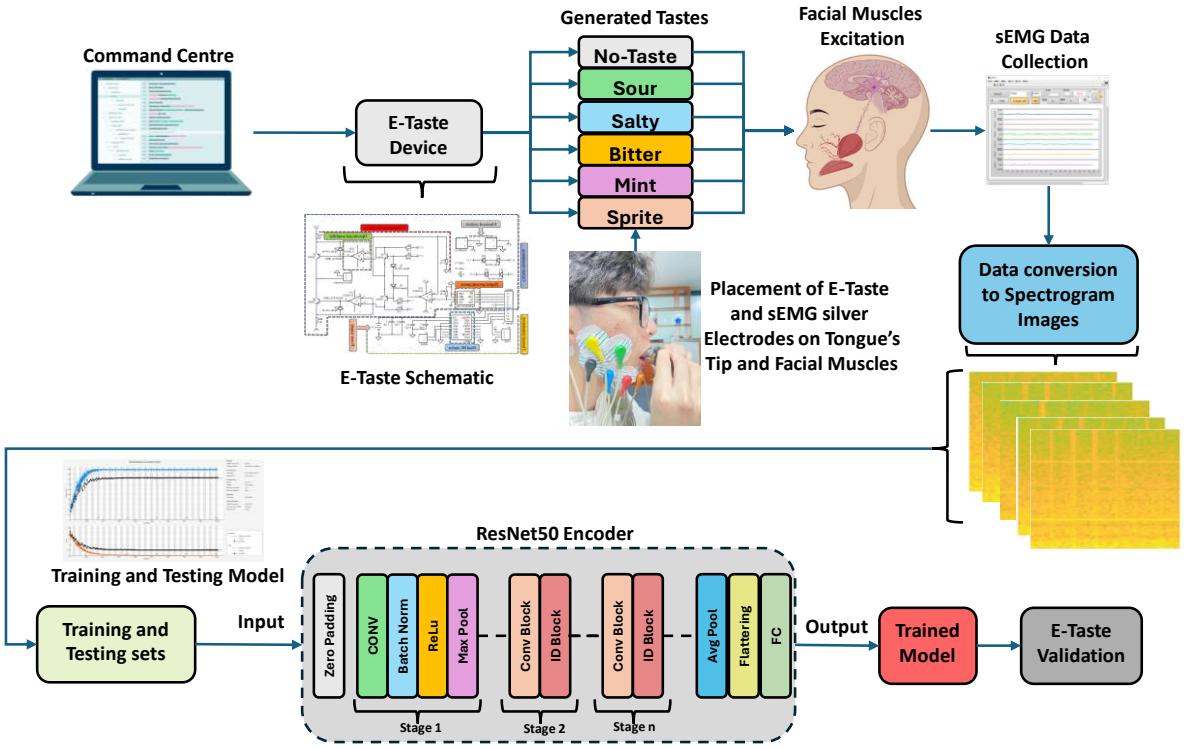


Figure 1. Schematic Representation of the E-Taste System Framework for Multi-Taste Classification using sEMG Data and ResNet50 Model

In this work, we introduce MACB-ResNet, a deep learning architecture that employs multiscale attention convolutional blocks (MACB) within the ResNet framework to recognize electronic taste sensations (ET-Sensations), as shown in **Figure 1**. To the best of our knowledge, this is the first study to investigate six-class E-Taste recognition using a residual dual-attention CNN framework. Unlike conventional sEMG classification tasks, E-Taste recognition involves subtle and highly overlapping muscle activation patterns across

multiple taste categories. To address this challenge, the proposed MACB-ResNet combines multiscale convolutional feature extraction with joint channel-wise and spatial attention mechanisms to enhance the representation of discriminative and weak sEMG features. Residual connections further facilitate stable training and efficient gradient propagation. Extensive experimental results, including class-level evaluation, visualization techniques (PCA and t-SNE), and quantitative ablation studies, demonstrate the effectiveness of the proposed framework for accurate E-Taste classification and highlight its potential applications in brain–computer interfaces (BCI), food technology, and neuromarketing.

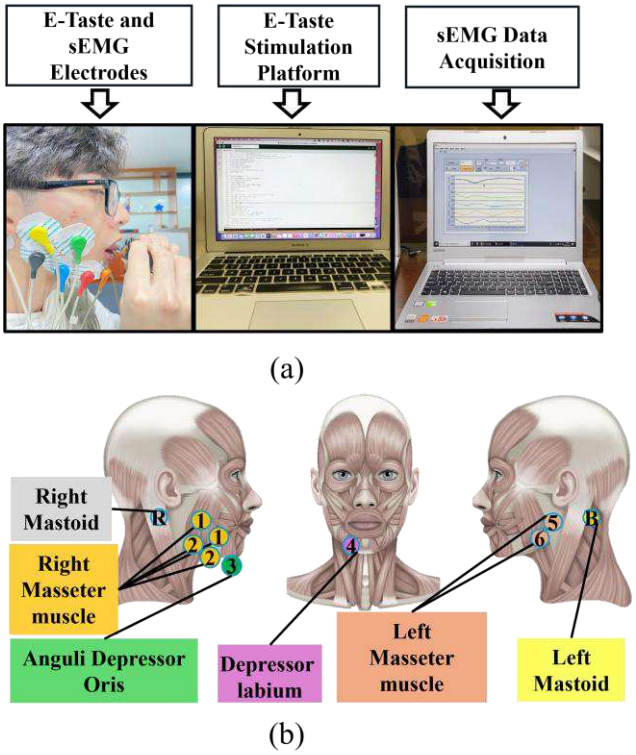


Figure 2. (a) Data Collection Setup for E-Taste System using sEMG Signals **(b)** Placement of sEMG Silver Electrodes on Facial Muscles for E-Taste Data Collection

2. Data Acquisition

2.1. Data Recording Platform

2.1.1. Configuration

The experimental environment, electronic taste stimulation device, and data acquisition equipment have been customized to ensure the quality of the gathered data. We set up a comfortable and soundproof environment for the participants. Due to the unique nature of E-Taste perception, participants were asked to abstain from eating or drinking (except water)

for two hours before the experiment and from spicy food for twenty-four hours. Taste was stimulated using an E-Taste device, while sEMG data were acquired using a multi-channel sEMG acquisition system at a sampling frequency of 1 kHz. The sEMG acquisition system used 10 Ag/AgCl electrodes configured to obtain six recording channels, along with one reference electrode (R) and one bias electrode (B). The configuration included two bipolar channels (Channels 1 and 2) and four monopolar channels (Channels 3–6). Electrodes were placed on the masseter, depressor anguli oris, and depressor labii inferioris muscles. This configuration enabled the acquisition of six synchronized sEMG signals per trial for subsequent processing [44]. The system, therefore, provided six-channel sEMG recordings from a total of ten electrodes. **Figure 2 (a)** depicts the data-collection setup, while the sEMG electrode placement is illustrated in **Figure 2 (b)**.

2.1.2. Electronic Taste (E-Taste) Stimulation

An E-Taste device was designed to generate various taste sensations, including sour, salty, sprite bitter, and mint. Moreover, the no taste (Physiological state) was included as a reference for other taste sensations. The current and temperature ranges for taste stimulation are shown in **Table 1**.

Table 1 E-Taste-Based Stimulation Ranges for Different Tastes and Flavors

Tastes and Flavors	Range of Stimulation
No Taste (no)	Nil
Salty (st)	40-70 μA
Sprite (sp)	60-180μA and 30-20 ºC
Bitter (bt)	60-140 μA
Mint (mt)	Below 25 ºC
Sour (sr)	60-180 μA

2.1.3. Participants

Eight participants (average age 23 ± 3 years, five males and three females) voluntarily participated in the experiments. Each participant passed the taste function preliminary and was in good physical and mental health, with no reported neurological or gustatory disorders. The participants in the study do not suffer from any illnesses or other symptoms of dysgeusia. Zhejiang University's Ethical Board approved this study for Human Safety. All methods were performed in accordance with the relevant guidelines and regulations, as approved by the ZJU review board. Informed consent was obtained from all participants and/or their legal guardians before the study. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

2.1.4. Dataset

Eight participants contributed to this study, and 90 experiments were conducted, each comprising 4 sessions. One session contained six trials. Overall, 360 sessions were conducted, incorporating 15,120 samples, with each taste sensation comprising 2,520 samples.

2.2. Data Recording Procedures

sEMG electrodes were placed on the participant's face, and he was instructed to place the tip of the tongue between the silver electrodes of the E-Taste device. The stimulation lasted twenty seconds to generate various taste sensations, while the sEMG signals from facial muscles were recorded for eight seconds. Specifically, sEMG acquisition was initiated at the 19th second of stimulation and continued for 8 seconds. The initial 1 second of sEMG recording overlapped with the 20th second of stimulation and was therefore discarded, resulting in an effective 7-second sEMG segment used for analysis. This step was implemented to avoid transient artifacts and noise introduced by electrical and thermal stimulation, ensuring that the analyzed sEMG data reflect stable muscle activity. Every taste sensation was followed by a one-minute interval during which the participant was instructed to rinse his mouth or drink water. To minimize confounding effects from the rinsing procedure, a 60-second rest period was enforced after rinsing and before the next taste stimulation. This interval allowed salivary conditions and facial muscle activity to return to baseline levels. Pilot testing confirmed that facial sEMG signals returned to baseline within 30–40 seconds post-rinsing; therefore, the 60-second rest period was sufficient to eliminate carryover effects from the previous stimulation. The participant was asked to take a 5-minute break after experiencing all 6 taste sensations before beginning the next session. The steps mentioned above were repeated for each participant [13].

2.3. Data Preprocessing

The acquired sEMG data were segmented into 7-second intervals. The initial 1 second, which overlapped with electrical and thermal stimulation, was removed to avoid transient artifacts. The remaining signal was then divided into 1-second subsegments for further analysis. Following the data structure, the dataset underwent four preprocessing steps: data detrending, high-pass filtering, adaptive notch removal, and removal of aberrant amplitude signals. By removing the polynomial trend components, the DC and base-signal deviations

are reduced, bringing the signal's fluctuation range back to an acceptable level (approximately 2000 μV for this acquisition system). After detrending, a high-pass Butterworth filter with a cutoff frequency of 10 Hz was applied to the dataset to eliminate baseline drift caused by motion artifacts. Afterward, an adaptive notch filter was applied to the sEMG waveform sample to remove harmonics and fluctuations caused by power-frequency interference. Signals with abnormal amplitudes were removed after base drifting, and interference from power frequencies was removed. A criterion was established to eliminate the abnormal signals. If the sample signal's fluctuation range is between -2000 μV and 2000 μV, the sEMG signal is considered standard. In contrast, the signals beyond this threshold were eliminated. An example of raw and filtered data is shown in **Figure 3**.

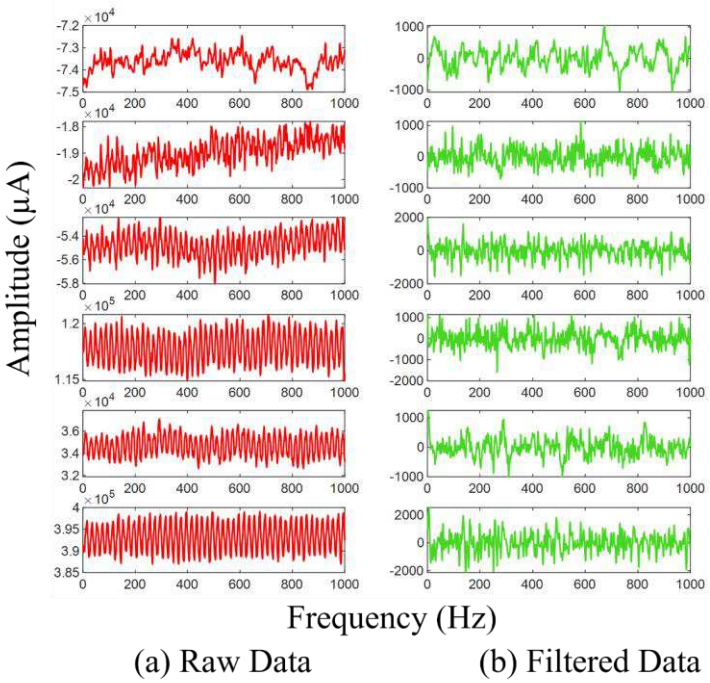


Figure 3. Comparison of Raw and Filtered sEMG Data for an E-Taste System after Preprocessing

2.4. Spectrogram Dataset

We applied the Short-Time Fourier Transform (STFT) to the filtered sEMG data to obtain spectrogram representations that capture variations in frequency content over time. STFT enables the extraction of transient time–frequency characteristics of muscle activation that are not easily separable in the time domain and are critical for distinguishing subtle taste-induced sEMG patterns**.** All collected channels were filtered, and segments from each channel were standardized to a common duration of 7 seconds. To ensure uniformity, each

segment was divided into 256-sample windows with 128-sample overlap, and a 512-point Fast Fourier Transform (FFT) was applied at a sampling rate of 1000 Hz. These parameters yielded high-resolution time–frequency representations for each taste label.

The raw sEMG signals comprised 6 channels × 7 seconds × 1000 Hz = 42,000 samples per trial. After preprocessing and 1-second segmentation with 50% overlap, STFT was applied to each segment (1000 samples), producing spectrograms of size 129 × 63 per channel. The spectrograms from all six channels were stacked to form a 6 × 129 × 63 tensor, which served as the input to the MACB-ResNet model.

The resulting spectrograms were organized by their respective taste labels and used as inputs to both machine-learning and deep-learning classification models. An example spectrogram for sour taste is shown in **Figure 4**.

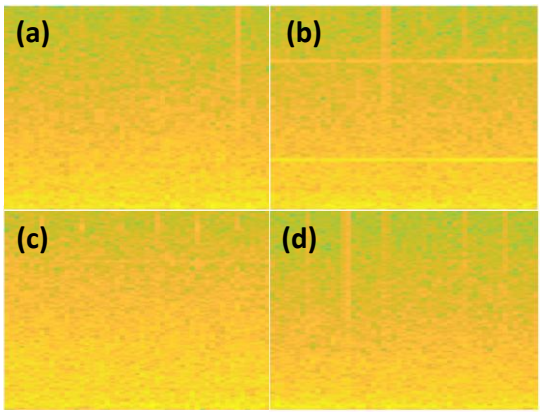


Figure 4. Spectrogram Samples for Sour Taste in E-Taste System

3. Methodology

3.1. Structure of MACB-ResNet

This part describes an effective and adaptable neural network design, MACB-ResNet, that blends the strengths of deep feature extraction in ResNets with MACB's ability to fine-tune features to identify E-Taste sensations from sEMG signals (see **Figure 5**). MACB's attention mechanism allows the model to weight input features when determining which provide the most information about the signal; once the features are refined, they are transformed into probabilistic classifications in the fully connected (FC) and softmax layers. This method can accurately classify very similar E-Taste sensations influenced by minute changes in sEMG signals.

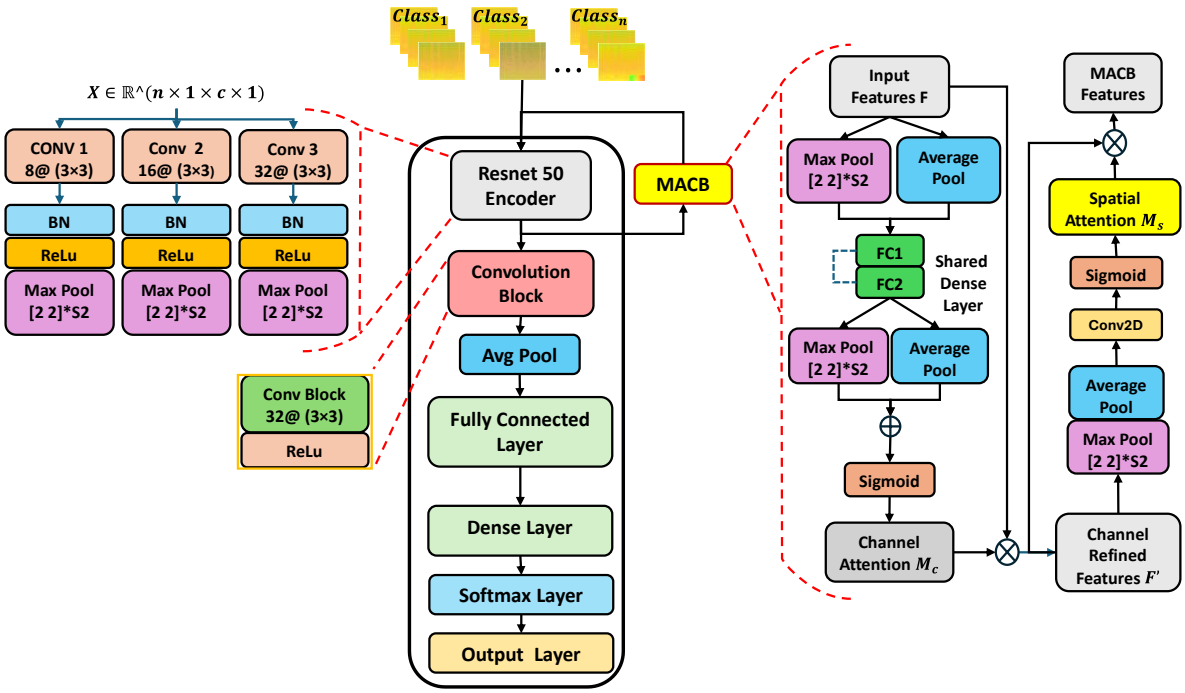


Figure 5. A Detailed Deep Learning Framework for E-Taste Classification Using ResNet50 Encoder with Multi-Attention Convolutional Block (MACB) Module

ResNet50 was selected because the STFT transforms sEMG signals into two-dimensional time–frequency representations that capture spatial and spectral locality. Convolutional neural networks are well-suited for extracting such localized features. CNN-based architectures have also demonstrated strong performance in spectrogram-based sEMG analysis compared to traditional machine learning methods, supporting the suitability of the selected design. Residual connections further enable stable training of deep architectures by improving gradient flow. In addition, the proposed Multi-Attention Channel Block (MACB) enhances feature representation by jointly modeling channel-wise and spatial dependencies, allowing the network to emphasize informative sEMG patterns associated with different taste sensations while suppressing irrelevant features.

3.2. Feature Extraction Module

This encoder architecture combines ResNet with the Multi-Attention Channel Block to extract deep features, with a primary focus on channel-based and spatially distinctive features.

3.2.1. ResNet-50 Encoder

Residual networks are designed with residual blocks that provide shortcuts through which gradients can bypass a few layers, allowing for the extraction of deep, numerous

complex features from sEMG. Temporal and spatial features are vital components of any sEMG variable, as they enable us to evaluate very small differences over time in the activation of taste-related muscle groups. Within each residual block, a sequence of operations consisting of a convolution operation, a batch normalization operation, and a ReLU operation will be operated on input X, and then added back onto the original input X, as shown below:

$$X_{out} = f(X) + X \quad (1)$$

f(X) represents the transformation of the input via a convolutional procedure. Because of its architecture, the network can learn from deeper layers more efficiently, which helps prevent gradient vanishing.

3.2.2. Multi-Attention Channel Block (MACB)

For each sEMG channel, this module used an attention mechanism to assign a weight that allows the model to emphasize the channel carrying the most valuable information for taste classification. Attention scores were generated by calculating both Average Pooling and Max Pooling before passing them through two Fully Connected Layers and then being combined with a Sigmoid Function to create final attention scores as seen in the following Equation,

$$Mc = \sigma\left(FC2\left(ReLU\left(FC1\left(AvgPool(F) + MaxPool(F)\right)\right)\right)\right) \quad (2)$$

M_C stands for channel attention map, which emphasizes significant channels by applying a scaling factor.

Spatial attention focused on the most significant spatial regions within the feature maps. It can be achieved by combining average and max-pooled features along the channels, applying a 2D convolution, and passing the result through a sigmoid function, stated as,

$$Ms = \sigma\left(Conv2D([AvgPool(Mc) \oplus MaxPool(Mc)])\right) \quad (3)$$

The output M_s serves as a spatial mask, allowing the model to focus on key spatial features associated with distinct taste sensations.

The output channel-refined features from the MACB were calculated as follows:

$$F' = Mc \odot Ms \quad (4)$$

where $\odot$ denotes element-wise multiplication. This results in channel- and spatially optimized features for taste-sensation classification.

3.2.3. Convolutional Layers for Deep Feature Extraction

After MACB-ResNet, the network included a series of convolutional blocks to further refine the extracted features. Each Convolutional Block consisted of:

(a) Convolution Layer: Each convolution layer applied a filter to capture spatial relationships within the sEMG signals. For example, in the first convolutional block:

$$Y_{i,j} = \sum_{k=0}^{m} \sum_{l=0}^{n} X_{i+j+k+l} . K_{k,l} \quad (5)$$

where K is the convolutional kernel. This process extracted localized patterns relevant to taste sensation from the data.

(b) Batch Normalization: BN stabilized training by normalizing each layer's inputs. It adjusted and scaled the layer activations during training. The final BN can be defined as:

$$y_i = \gamma \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (6)$$

Where μ is the mean of the batch, σ^2 is the variance, $\hat{x}_i$ represents batch normalization by subtracting the mean and dividing it by the standard deviation (SD), and γ and β are learnable parameters after the transformation of the normalized input $\hat{x}_i$. The mathematical equations of these parameters can be represented as

$$\mu = \frac{1}{m} \sum_{i=1}^{m} x_i \quad (7)$$

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^{m} (x_i - \mu)^2 \quad (8)$$

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \quad (9)$$

$$y_i = \gamma \hat{x}_i + \beta \quad (10)$$

(c) ReLU Activation: Rectified linear unit (ReLU) applied a non-linear transformation, allowing the network to capture non-linear patterns associated with taste sensations. This function can be defined as:

$$f(x) = max\,(0, x) \quad (11)$$

For an input x, if $x > 0$, then $f(x) = x$, and if $x \leq 0$, then $f(x) = 0$.

(d) Max Pooling: MP reduced the spatial size of the feature maps, preserving only the most significant features, thus reducing computational complexity. It can be defined as:

$$y_{i,j} = max\,(X_{m,n}),\ (m, n) \in windows_{i,j} \quad (12)$$

Where $y_{i,j}$ is the output value after max pooling at position (i, j), $X_{m,n}$ are the values within the $k \times k$ pooling window corresponding to the position (i, j) in the output feature map, and $windows_{i,j}$ represents the set of indices (m, n) covered by the pooling window centered at (i, j).

3.2.4. Classification Model

The refined features extracted from the convolutional layers were then passed through a classification model consisting of fully connected layers and a softmax layer to deliver the final taste prediction.

(a) Fully Connected Layer (Dense Layer): The fully connected layer aggregated information from the preceding layers. It transformed the multidimensional feature map into a score for each taste class. Mathematically, this transformation is:

$$y = W_z + b \qquad (13)$$

W and b are weights and biases, and z represents the feature map. This layer generated output vectors representing the scores for the six taste categories.

(b) Softmax Layer: Softmax converted scores into probabilities, with each probability representing the likelihood of a given taste class. The softmax function is described as:

$$p_i = \frac{e^{y_i}}{\sum_{j=1}^{6} e^{y_i}} \qquad (14)$$

where y_i is the score for class i. The sum of probabilities for all classes is 1, making it suitable for multi-class classification.

(c) Output Layer: The output layer took the probabilities from the softmax layer and assigned the input sEMG signal to the taste class with the highest probability. This final classification step allowed the system to predict the specific taste sensation.

Table 2 Hyperparameter Values of Training/Testing

Parameters	Value
Epochs	30
Iterations	21750
Learning Rate	0.0001
Batch size	8
Validation Frequency	30Hz
Dropout Rate	0.1
Train/Test Ratio	9/1
Model Size	27,910,760

4. Experimental Results

4.1. Model Training and Evaluation Metrics

4.1.1. Training Configuration and Hyperparameters

Training was performed by the model for 30 epochs (i.e., 21,750 iterations). The learning rate for each training iteration was set to a constant of 0.0001, and the batch size for each iteration was kept at 8. Every 30 iterations of the training process, validation would occur to assess how well the model performed on new data. The training dataset was divided into a 9-to-1 split relative to the test dataset.

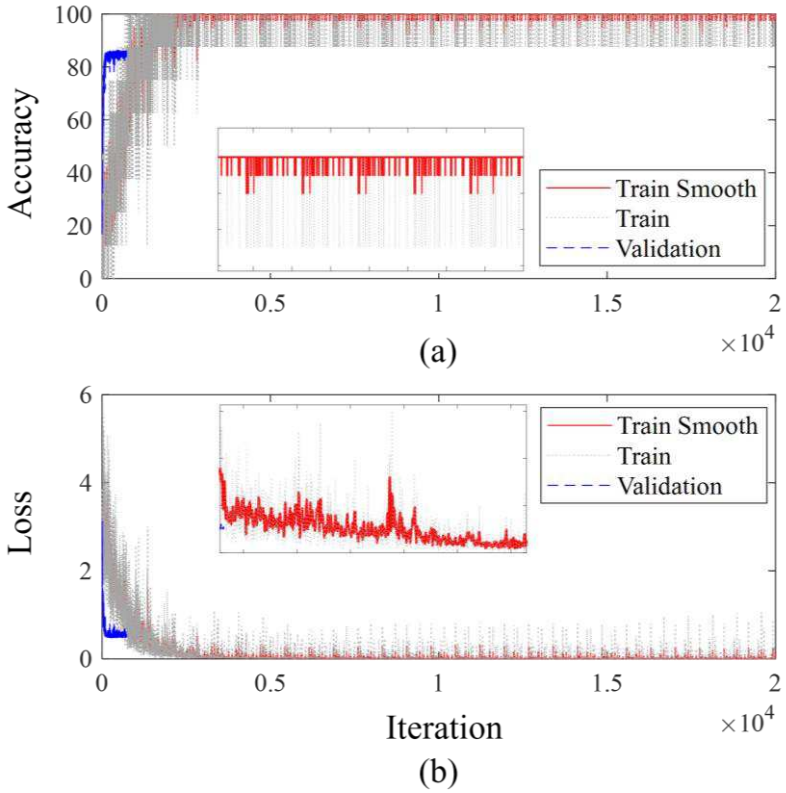


Figure 6. Training and Validation Performance of the E-Taste Classification Model (a) Accuracy and (b) Loss

4.1.2. Training and Validation Performance

Figure 6(a) illustrates the plot of training and validation accuracies, and **Figure 6(b)** displays the plot of loss during training and validation. The training accuracy increased rapidly during an early learning phase (between 100 and 200 iterations), and stabilized at around 100%. The validation accuracy also stabilized at approximately 86.32%, suggesting it can generalize to unseen data.

During the first iteration of the model's training, the loss was high; the metric then decreased and stabilized, indicating successful learning of the training dataset's features and attained a steady state. The similarity between the training and validation loss trends and their stabilization processes supports the model's performance on unseen data.

There was a very small difference between the training and validation accuracies, and both losses have reached steady states; these results indicate that convergence has occurred for this model, with no noticeable signs of decreasing model quality.

4.2. Performance of MACB-ResNET

4.2.1. Class-Level Performance Metrics

The classification achieved 86.32% accuracy, demonstrating that the model performed effectively. **Table 3** shows the class-specific metrics, including accuracy, F1 score, TPR, TNR, FPR, and FNR, which were used to evaluate the model. The model effectively discriminates between taste categories and shows high TPR and TNR, with relatively low FPR and FNR across all classes. The F1 scores also support this conclusion: it gives balanced performance across all classes, but the highest F1 score is for the "no" class (0.92), indicating that its precision and recall are very good. The "sp", "mt", and "st" classes also have competitive F1 scores, indicating the model's reliable classification performance.

On the other hand, the "bt" and "sr" classes have relatively low F1 scores (0.84 and 0.82, respectively), indicating that it may be harder to classify between these two classes due to greater similarity in their feature traits. The higher FNR for those two classes supports the observation that further improvement could be made to the classification of the "bt" and "sr" classes.

Table 3 Performance Metrics of E-Taste Classification Model across Different Taste Classes

Class	Accuracy	F1 Score	TPR	TNR	FPR	FNR
no	0.97	0.92	0.92523	0.97944	0.020561	0.074766
st	0.96	0.89	0.8981	0.9757	0.024299	0.14019
sp	0.95	0.90	0.86916	0.96636	0.033645	0.13084
bt	0.94	0.84	0.84112	0.95888	0.041121	0.15888
mt	0.93	0.87	0.81308	0.95327	0.046729	0.18692
sr	0.96	0.82	0.80374	0.98879	0.011215	0.19626

4.2.2. Data Representation and Visualization

Figure 7 demonstrates dimensionality reduction using two-dimensional (2D) and three-dimensional (3D) Principal Component Analysis (PCA). With PCA, the original feature space was transformed into a smaller, more manageable set of principal components that

retain as much variance as possible in each reduced dimension, helping visualize the separability of the different classes.

The clusters formed by the different classes on the 2D plot (see **Figure 7(a)**) showed some overlap, although the "no" cluster appeared relatively compact and contained a more consistent set of features that could be more easily classified. In contrast, the "st", "sp", and "bt" classes exhibited substantial overlap, which likely contributed to misclassifying samples among these classes. The overlap among the clusters formed by these three classes indicates that their feature representations are similar, making them difficult to classify.

In the 3D projection, as shown in **Figure 7(b),** the separation between the classes was improved compared to the 2D representation. Additionally, the 3D model allowed for better separation of the grouped classes, which clustered closely together. This indicates that the dimensions of the feature space provide more information that helps in class separation. However, there is still some overlap between classes, which could be due to their similar feature representations, and led to the misclassification of some samples from these classes.

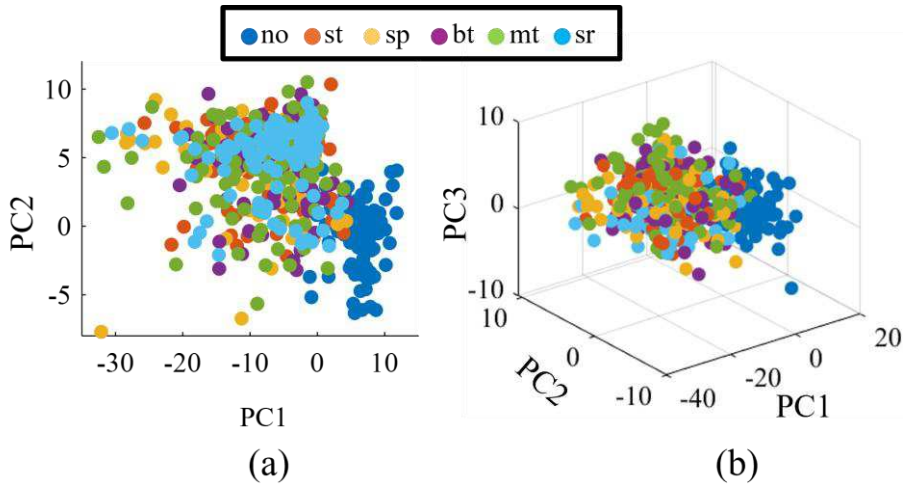


Figure 7. PCA Visualization of (a) 2D (b) 3D PCA Intra-Class Analysis for E-Taste Classification

4.2.3. T-SNE Visualization

Figure 8 shows the t-distributed stochastic neighbor embedding (t-SNE) and indicates the separation between the classes when projecting high-dimensional feature representations into a lower-dimensional space (specifically, a two-dimensional projection). In most cases, class clustering is well-defined, and remarkably good separation was observed between the “no” and “sr” classes. The features learned by the model are highly discriminative for accurate classification, as demonstrated by the compactness of clustering between classes. However, the classes “sp,” “mt,” and “bt” showed partial overlap, suggesting similar

underlying feature distributions and indicating that additional engineering or advanced techniques may be required to facilitate their separation in future work.

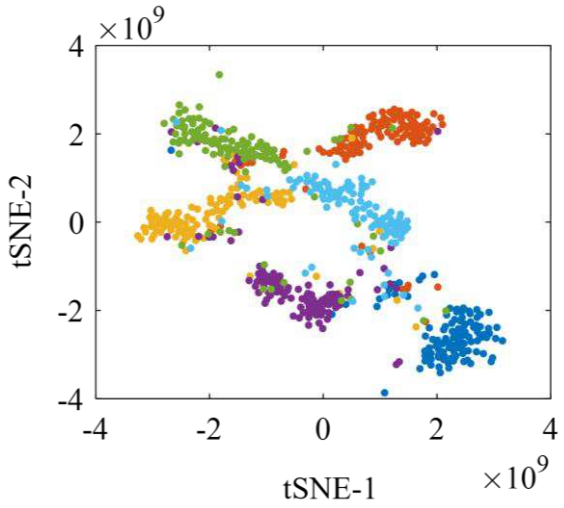


Figure 8. t-SNE Visualization of E-Taste Classification Model's Feature Embeddings

4.2.4. Class Distribution Analysis

Figure 9 shows the prediction's distribution across the 6 classes. Overall, the distribution was quite even, with "bt" and "mt" having the most predictions (112 each), followed by "sp" (111), "no" (110), "st" (105), and "sr" (92).

Overall, the distributions were fairly uniform across all taste categories, indicating that the model showed little bias towards any particular category. Minor differences were observed in the class distributions, but they were not statistically significant and suggest consistent predictive performance across all taste categories.

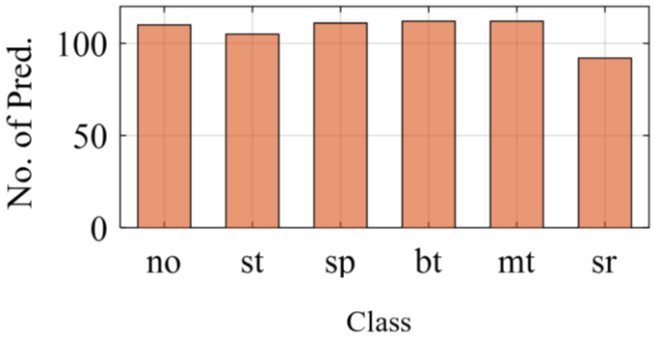


Figure 9. Class Distribution of Predictions in the E-Taste Classification Model

4.2.5. Confusion Matrix Analysis

To ascertain the efficacy of a machine learning model, it is important to evaluate its performance. One standard approach to assessing a model's predictive accuracy on actual data is the confusion matrix (shown in **Figure 10**). The diagonal values in this matrix

represent the true positives, which count how many times the model correctly identified the class. In the observed results, the classification model performs well overall, with high true-positive counts across all categories. Remarkably, the classes of no and sp also have rather high true-positive rates, which means that the precision and recall of these classifications are high.

On the contrary, the off-diagonal cells show the incorrectly predicted instances by the classification model. In particular, there is significant confusion among instances in the "sp", "bt", and "sr" categories, suggesting that the input data distributions for these three classes are similar. For example, there were many instances of the "bt" category that were incorrectly predicted by the model as either "sp" or "sr", and instances of both "sp" and "bt" were predicted as "sr". As a result, the classification model has a very difficult time identifying closely related tastes, which is a significant area for improvement.

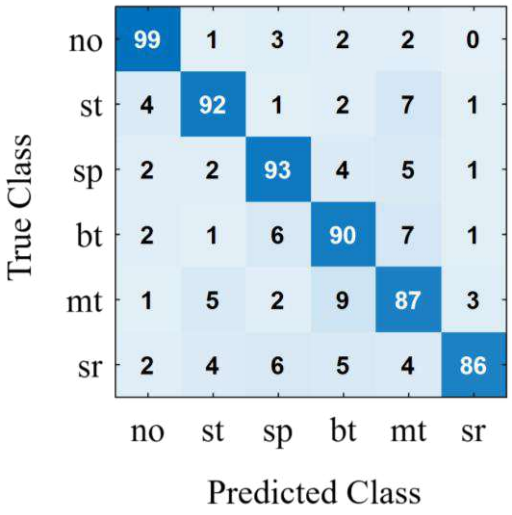


Figure 10. Confusion Matrix for E-Taste Classification Model

4.2.6. Precision-Recall Curve

The Precision-Recall curve shown in **Figure 11** illustrates the classes' respective performance at threshold settings. The Precision-Recall curves for all classes show that the model achieves adequate precision and recall across a range of thresholds, indicating that it can classify many TP instances with relatively few FP instances. Precision–recall performance is particularly important when FP and FN have a significant impact on system performance. Therefore, the class bt and mt models appeared to have slightly higher degrees of difficulty in terms of FP as compared to TP, as demonstrated by the abrupt decline of Precision scores as Recall increases; thus, finding optimized threshold settings for the two

classes will increase overall model performance, especially when Recall is of utmost importance.

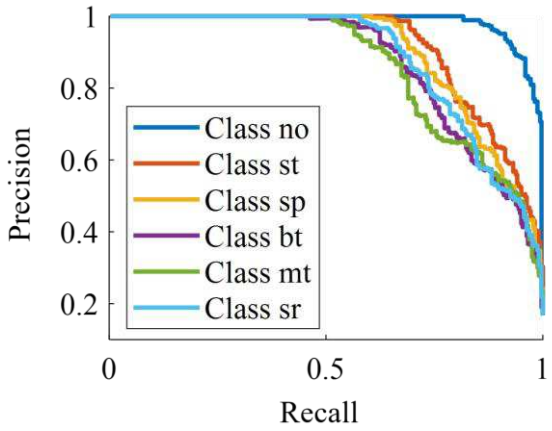


Figure 11. Precision-Recall Curve for Each Class

4.2.7. ROC Curve

The ROC curve in **Figure 12** shows the model's ability to separate the classes across different thresholds. The AUC values show that the majority of the model's classes are performing well because they possess high AUC values for the model overall (in fact, the "No" class has an AUC of 1; all other classes are above .95). A high AUC reflects that the model has very good TPR and low FPR and can therefore serve as a robust classifier. However, the relative difference in AUC values also indicates a difference in the degree of class separation. For instance, the AUC values for the “bt” and “mt” classes are approximately 0.95. Although slightly lower than that of the “no” class, these values still indicate strong classification performance. Overall, although the model is generally good at separating classes, there are opportunities to improve class separation in some instances.

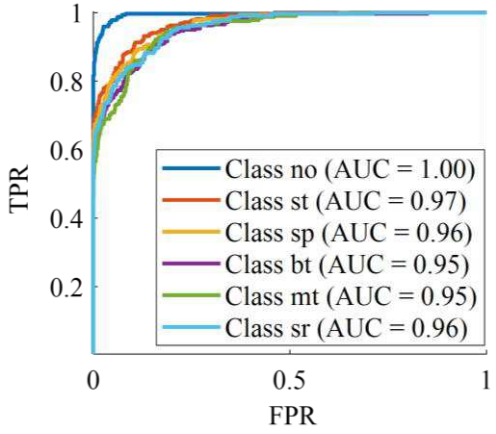


Figure 12. ROC Curve and AUC Scores for Each Class in the E-Taste Classification Model

4.2.8. Misclassified Samples

Figure 13 illustrates the misclassification patterns of the proposed model. The findings showed that most errors occurred between classes with similar feature characteristics. For example, some instances of the “bt” class were sometimes misclassified as the “mt” or “sp” class, indicating that their underlying feature representations are similar.

These results indicate that improved feature discrimination is required, either by adopting more effective feature extraction algorithms or by adding more labeled training data. Additional work on misclassified samples can provide insight into how to better define class boundaries and the model's overall performance.

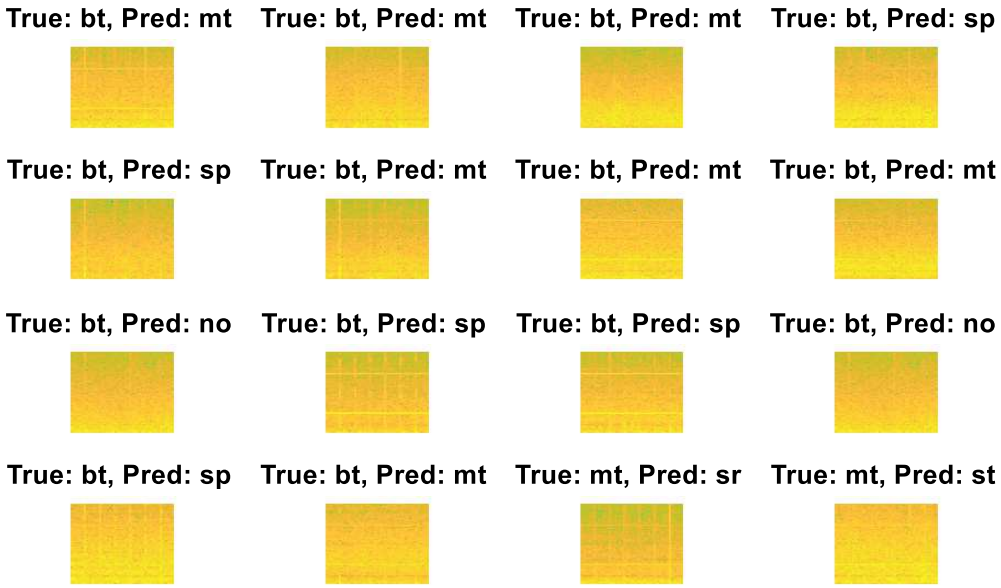


Figure 13. Misclassified Samples in the E-Taste Classification Model

4.2.9. Comparison with Other Methods

This section presents a comparative analysis of RF, KNN, DenseNet-201, and the proposed MACB-ResNet models on the sEMG-based E-Taste dataset, using a 90%–10% training–testing split. The classification results and average accuracies are shown in **Figure 14**.

Although KNN achieved an average accuracy of 65.11%, demonstrating moderate performance on the “bt” and “no” classes, it performed relatively poorly on other classes, such as “sr” and “st”. These findings suggest that KNN has limited capability in capturing discriminative features across all taste categories.

Compared to KNN, the random forest (RF) classifier showed a notable improvement, achieving an average accuracy of 78.35%, with particularly strong results for the “no” class (94.13%). However, its performance decreased for “bt” and “mt,” suggesting limitations in modeling subtle variations in taste-related features.

DenseNet-201 further improved performance, achieving an average accuracy of 81.93% and offering more consistent classification across categories, particularly for “sp” and “bt,” where accuracies exceeded 80%.

In contrast, the proposed MACB-ResNet achieved the highest performance, with an average accuracy of 86.32% and over 93% accuracy in five out of six classes. This improvement highlights the effectiveness of the proposed architecture in extracting discriminative features from sEMG signals and accurately classifying complex taste patterns.

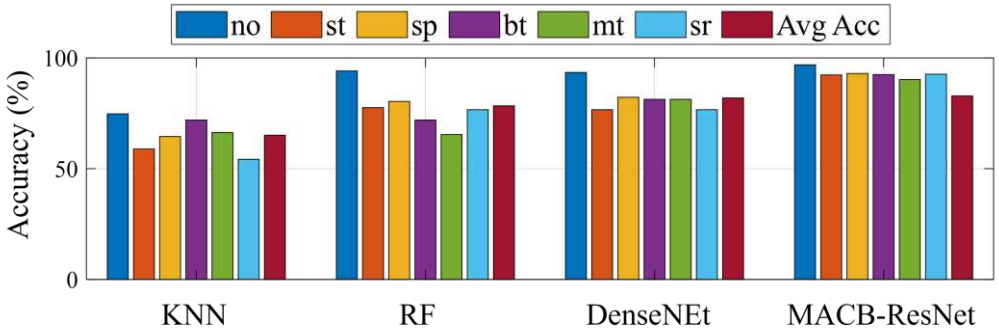


Figure 14. Comparison of Classification Accuracy Across Six Taste Categories Using Multiple Classifiers.

4.3. Ablation Study Overview

A series of studies was performed to evaluate the efficacy of each component of the proposed structure in proper taste detection. To accomplish the aim, we systematically suppressed or manipulated the inherent properties of blocks to identify the contribution of each block to taste identification using sEMG data. Quantitative results for all ablated components are summarized in **Table 4** (Section 4.3.4).

4.3.1. Impact of Residual Blocks

The ResNet architecture features Residual Blocks that help extract features from the input sEMG recordings and enable the model to learn common temporal and spatial patterns. Performance dropped significantly for classes with similar taste features (e.g., "mt" and "sr") after removing the residual connection from the model. This implies that residual connections allow the network to maintain higher gradient flow, leading to more complete feature extraction and reduced degradation, which is commonly experienced with deep networks.

526
527

528 4.3.2. Role of Multi-Attention Channel Block (MACB)

529 Recognizing taste in the MACB module was enhanced by increasing feature
530 discrimination across channels and space. The channel attention mechanism assigned
531 adaptive weights to channels and focused on those that contained discriminative information
532 for a particular taste class. As a result, improved performance was observed for the “st,” “sr,”
533 and “no” classes. The accuracy was significantly reduced upon removing the channel
534 attention component, confirming its importance in capturing channel-specific characteristics.
535 Likewise, the spatial attention mechanism was used to improve classification,
536 particularly for the “bt” and “sp” classes. Focusing on informative regions within the feature
537 maps enabled more effective extraction of localized sEMG patterns. Removing spatial
538 attention resulted in impaired performance across several classes, demonstrating its
539 significant contribution to refining spatial feature representations.

540 4.3.3. Contribution of Convolutional Layers

541 The MACB was followed by additional convolutional layers to enhance features for final
542 classification. The convolutional layers identified the localized patterns of taste-related
543 features in the channel and spatially optimized sEMG signals. The resulting lower
544 convolutional layers made the model less accurate on complex taste categories like "mt" and
545 "sp" that cannot be perfectly classified with fewer convolutional layers. By combining
546 improved feature extraction with attentional mechanisms, the MACB-ResNet achieves
547 superior performance and is therefore the best model for sEMG-based taste classification
548 among the various models evaluated.

549 4.3.4. Quantitative Ablation Results

550 The quantitative overview of the ablation study is presented in **Table 4**, which reports
551 the validation accuracy, macro F1 score, and accuracy drop for each ablation component
552 compared to the full MACB-ResNet Model. The quantitative findings support the conclusion
553 that the use of residual connections and multi-attention channels synergistically enhances E-
554 taste classification performance. The MACB-ResNet outperformed all ablated versions and
555 showed the largest performance decline when all parts of the MACB module were removed.
556 Thus demonstrating the importance of using both channel and spatial attention for taste-
557 related sEMG feature extraction.
558

559 **Table 4** Quantitative ablation study results relative to the full MACB-ResNet baseline
560 across different ablated components.

Ablated Component	Validation Accuracy (%)	Macro F1 Score	Δ Accuracy
None (Full MACB-ResNet)	86.32	0.87	—
− Residual Connections	79.14	0.78	−7.18%
− Channel Attention	81.23	0.80	−5.09%
− Spatial Attention	82.45	0.81	−3.87%
− Full MACB Module	76.58	0.75	−9.74%

561 **Table 5** Comparison of Training Configuration and Its Impact on Model Performance

Training Configuration	Setup 1	Setup 2	Setup 3
Toral Epochs	35	50	30
Iteration per Epochs	141	161	725
Maximum Iteration	4935	8050	21750
Training/Testing	7/3	8/2	9/1
Validation Accuracy (%)	76.92	82.02	86.32
Training Time	22m, 22s	23m, 10s	47m, 0s

562 ***4.4. MACB-ResNet's Comparison Study***

563 To determine the impact of Model Performance across different Training and Testing
564 configurations, multiple experiments were conducted using various methodologies. The
565 results of the various Training configurations and their effect on Model Performance are
566 provided in **Table 5**.

567 In setup 1, training and validation accuracies converged over time, with a maximum
568 validation accuracy of 76.92%, indicating the model was underfit. The training loss was
569 substantial, indicating insufficient optimization, whereas the validation loss improved
570 generalization.

571 In setup 2, training accuracy reached nearly 100% while validation accuracy was about
572 82.02%, indicating overfitting. The training loss continued to decrease consistently. The
573 validation loss showed minimal improvement, suggesting the likelihood of overfitting.
574 Compared to Setup 1, Setup 2 yielded fewer overall iterations but slightly greater overfitting,
575 which improved validation performance.

576 Validation accuracy for setup 3 stabilized at approximately 86.32% while training
577 accuracy rapidly approached 100%. The training loss decreased consistently throughout this
578 setup, indicating effective optimization. The dissimilarity between the training and validation
579 losses in setup 3 indicated potential overfitting. In this setup, training performance improved

with more epochs and iterations, which contributed to overfitting and thus limited generalization.

Table 6 Computational cost and hardware requirements of the proposed MACB-ResNet model.

Computational-cost metric	Actual / estimated value
Trainable parameters	27.91 M
Parameter memory, FP32	≈106.48 MB
Input window size	6 × 129 × 63 spectrogram
MACs/window	≈765.0 M
FLOPs/window	≈1.53 GFLOPs
Training epochs/iterations	30 epochs / 21,750 iterations
Total training time	47 min 0 s
Average training time/epoch	≈94.0 s
Average training time/iteration	≈0.130 s
Estimated inference latency/window	≈5.4 ms
Hardware/software	MATLAB R2023a; Intel Core i7-12700F CPU, 12 cores, 32 GB RAM; no GPU

4.5. Computational Cost and Hardware Requirements

For BCI and food-technology applications, computational efficiency is an important consideration because the model should provide accurate classification while remaining practical for future real-time deployment. The proposed MACB-ResNet contains 27.91 million trainable parameters, corresponding to approximately 106.48 MB of parameter memory when stored in 32-bit floating-point format (see **Table 6**). For each input spectrogram window of size 6 × 129 × 63, the model requires approximately 765.0 million multiply–accumulate operations (MACs), equivalent to approximately 1.53 GFLOPs per forward pass.

The main experiment was trained for 30 epochs, corresponding to 21,750 iterations, and required 47 min 0 s in MATLAB R2023a on a desktop system with a 12th Gen Intel Core i7-12700F CPU, 32 GB of RAM, and no GPU acceleration. This corresponds to an average training time of approximately 94.0 s per epoch and 0.130 s per iteration. Based on the reported CPU training time and batch size, the estimated inference latency is approximately 5.4 ms per spectrogram window. Although the FLOPs and latency values were estimated from the reported architecture and training configuration rather than measured using a dedicated profiler, these results indicate that MACB-ResNet provides a feasible

computational profile for sEMG-based E-Taste classification. Future work will include direct CPU/GPU benchmarking and model compression techniques, such as pruning and 8-bit quantization, to further reduce computational complexity and support real-time deployment.

5. Discussion, Limitations, and Future Work

5.1. Discussion

The proposed MACB-ResNet architecture improves upon existing methods by incorporating residual learning, multi-attention channel blocks, and convolutional feature extraction. It outperformed traditional machine learning algorithms (KNN and Random Forest) and deep learning models such as DenseNet-201 in classifying sEMG-based taste signals, especially for complex, closely related groups. Experimentally, the MACB-ResNet architecture, which includes residual blocks and channel-spatial attention mechanisms, enables the model to prioritize discriminative features, leading to higher accuracy and greater robustness in taste classification than current methods.

Experimental results showed that the proposed MACB-ResNet outperformed the compared models across various taste classes, with significant improvements on classes that required fine discrimination. This is because it views residual learning as coupled with channel-spatial attention, which enhances the ability to extract discriminative features from sEMG signals. On the other hand, DenseNet-201 does not have task-specific attention mechanisms as compared to traditional classifiers like KNN and Random Forest.

E-Taste is an emerging application of sEMG-based classification that enables non-invasive measurement of taste in the areas of food science, health care, and assistive technologies. The results highlight the potential to have advanced deep learning models to improve the reliability and accuracy of sEMG-based systems.

5.2. Limitations and Future Work

Despite the strong performance of MACB-ResNet, the current dataset size remains moderate for deep-learning applications. The size and diversity of the datasets may also influence model performance, as sEMG signals can be affected by electrode position, muscle fatigue, and individual physiological factors. Further research will enhance the data by incorporating additional participants, repeated recording procedures, and cross-subject validation to make the data more robust and generalizable.

Similarly, the current participant sample included eight young adults (23 ± 3 years; five male, three female). The sEMG responses may depend on age-related variations in taste sensitivity, hormonal factors, and differences in muscle activation patterns. Thus, the applicability of the proposed model to other populations (e.g., older adults, children, or people with gustatory impairments) remains limited and requires further confirmation. Further work will involve stratified data collection by age and sex.

Another limitation is that STFT-based spectrograms provide only a fixed time-frequency representation and may not capture the time-varying nature of taste-induced muscle activity. Minor dynamic effects, such as the gradual change in the strength of the taste impression, might thus be underrepresented. Temporal models, including BiLSTM, TCN, or transformer-based networks, should be included in future studies to analyze raw sEMG sequences and infer long-range dependencies.

Furthermore, the evaluation of the ablation study was limited to accuracy and F1-score. To provide a more comprehensive assessment of each module's contribution, additional metrics such as parameter efficiency, inference speed, and memory usage should be employed.

The experimental setup also experienced practical limitations, such as the safe and comfortable positioning of the E-Taste device electrodes on the tongue. The new designs will aim to enhance the devices' ergonomics, safety certification, and user comfort to enable reliable data collection.

Lastly, although this study focused on sEMG-based E-Taste classification, applying the proposed MACB-ResNet framework to other biosignal tasks or multimodal sensory systems remains an open research direction. Moreover, the evaluation of transformer-based architectures was not considered due to dataset size and computational constraints; their integration will be explored in future work.

6. Conclusion

The proposed study used the MACB-ResNet model to classify electronic taste sensations from sEMG signals, achieving a validation accuracy of 86.23%. The combination of multi-attention-channel blocks and residual connections enhances feature extraction and facilitates effective discrimination between closely related taste categories. The proposed framework demonstrates improved performance compared with traditional classifiers (KNN and Random Forest) and deep learning models such as DenseNet-201, as it can capture complex patterns in sEMG features.

The MACB module uses channel and spatial attention to concentrate on the discriminative feature representations, with residual connections to support stable training and efficient gradient propagation. The PCA and t-SNE visualizations also indicate that the model can generalize across taste classes.

Overall, MACB-ResNet is a solid framework for sEMG-based E-Taste classification and has potential applications in virtual reality, neuromarketing, food technology, and brain-computer interface (BCI) systems, and can serve as a basis for future developments in non-invasive assessment of taste.

Data Availability: The data will be made available upon request from the corresponding author.

Acknowledgments: This work was supported by the Guangdong Provincial Department of Education Characteristic Innovation Project (Grant: 2022KTSCX307), Research Projects of Department of Education of Guangdong Province（Grant: 2025ZDZX3069）, Postdoctoral Foundation (Grant: 6024331025K), and Research Projects of Shenzhen Polytechnic University (Grant: 6023310034K, 6023310026K).

References

[1] J. Ye *et al.*, "A novel biomimetic electrochemical taste-biosensor based on conformational changes of the taste receptor," vol. 249, p. 116001, 2024.

[2] D. Rosati *et al.*, "Taste and Smell Alterations (TSAs) in Cancer Patients," vol. 12, no. 6, p. 130, 2024.

[3] S.-Y. Tan *et al.*, "Alterations in sweet taste function in adults with diabetes mellitus: A systematic review and potential implications," vol. 63, no. 16, pp. 2613-2625, 2023.

[4] C. J. J. B. P. C. N. McAdams and Neuroimaging, "Expectations Matter: Outcomes in Anorexia Nervosa," vol. 9, no. 3, pp. 249-250, 2024.

[5] J. R. Lechien *et al.*, "COVID-19: post-vaccine smell and taste disorders: report of 6 cases," vol. 103, no. 2, pp. NP104-NP107, 2024.

[6] K. Karunanayaka, N. Johari, S. Hariri, H. Camelia, K. S. Bielawski, and A. D. Cheok, "New thermal taste actuation technology for future multisensory virtual reality and

693 internet," *IEEE transactions on visualization and computer graphics,* vol. 24, no. 4,
694 pp. 1496-1505, 2018.
695 [7] E. J. T. S. Kerruish and Society, "Arranging sensations: smell and taste in augmented
696 and virtual reality," vol. 14, no. 1, pp. 31-45, 2019.
697 [8] A. Cruz and B. G. Green, "Thermal stimulation of taste," *Nature,* vol. 403, no. 6772,
698 pp. 889-892, 2000.
699 [9] H. T. Lawless, D. A. Stevens, K. W. Chapman, and A. Kurtz, "Metallic taste from
700 electrical and chemical stimulation," *Chemical senses,* vol. 30, no. 3, pp. 185-194,
701 2005.
702 [10] A. Ullah *et al.*, "E-Taste: Taste sensations and flavors based on tongue's electrical
703 and thermal stimulation," *Sensors,* vol. 22, no. 13, p. 4976, 2022.
704 [11] R. Anbarasan, D. Gomez Carmona, and R. J. F. E. R. Mahendran, "Human taste-
705 perception: brain computer interface (BCI) and its application as an engineering tool
706 for taste-driven sensory studies," vol. 14, no. 3, pp. 408-434, 2022.
707 [12] H. Wang *et al.*, "Unsupervised cross-user adaptation in taste sensation recognition
708 based on surface electromyography," vol. 71, pp. 1-11, 2022.
709 [13] A. Ullah *et al.*, "Surface Electromyography-Based Recognition of Electronic Taste
710 Sensations," vol. 14, no. 8, p. 396, 2024.
711 [14] J. Chandrashekar, M. A. Hoon, N. J. Ryba, and C. S. J. N. Zuker, "The receptors and
712 cells for mammalian taste," vol. 444, no. 7117, pp. 288-294, 2006.
713 [15] D. A. Yarmolinsky, C. S. Zuker, and N. J. J. C. Ryba, "Common sense about taste:
714 from mammals to insects," vol. 139, no. 2, pp. 234-244, 2009.
715 [16] J. R. Gibbons and N. M. Sadiq, "Neuroanatomy, neural taste pathway," 2019.
716 [17] A. Komiyama, T. Kobayakawa, H. Toda, N. Gotow, M. Ikeda, and S. Saito, "A high-
717 concentration NaCl solution does not stimulate the human trigeminal nerve at the tip
718 of the tongue," *Acta Otolaryngol,* vol. 127, no. 7, pp. 754-759, 2007.
719 [18] A. Ullah *et al.*, "Gender Differences in Taste Sensations Based on Frequency Analysis
720 of Surface Electromyography," vol. 130, no. 3, pp. 938-957, 2023.

[19] Y. Miyaoka, I. Ashida, D. Inagaki, and S. Y. KAWAKAMI, "Differentiation of activity patterns in the suprahyoid muscles during swallowing of foods with five taste qualities," *Journal of Sensory Studies,* vol. 20, no. 6, pp. 473-483, 2005.

[20] R. Merletti and P. J. Parker, *Electromyography: physiology, engineering, and non-invasive applications*. John Wiley & Sons, 2004.

[21] M. Cifrek, V. Medved, S. Tonković, and S. Ostojić, "Surface EMG based muscle fatigue evaluation in biomechanics," *Clinical biomechanics,* vol. 24, no. 4, pp. 327-340, 2009.

[22] C. Dawes *et al.*, "The functions of human saliva: A review sponsored by the World Workshop on Oral Medicine VI," *Archives of oral biology,* vol. 60, no. 6, pp. 863-874, 2015.

[23] A. Del Vecchio, F. Negro, F. Felici, and D. Farina, "Associations between motor unit action potential parameters and surface EMG features," *Journal of Applied Physiology,* vol. 123, no. 4, pp. 835-843, 2017.

[24] K. Mukai and K. Watanabe, "Intentional facial expression variation per taste preference for beverages," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2024, vol. 46.

[25] P. Kaczmarek, T. Mańkowski, and J. Tomczyński, "putEMG—a surface electromyography hand gesture recognition dataset," *Sensors,* vol. 19, no. 16, p. 3548, 2019.

[26] N. Nasri, S. Orts-Escolano, and M. Cazorla, "An semg-controlled 3d game for rehabilitation therapies: Real-time time hand gesture recognition using deep learning techniques," *Sensors,* vol. 20, no. 22, p. 6451, 2020.

[27] Y. Wang *et al.*, "Silent speech decoding using spectrogram features based on neuromuscular activities," *Brain sciences,* vol. 10, no. 7, p. 442, 2020.

[28] N. Nasri, S. Orts-Escolano, F. Gomez-Donoso, and M. Cazorla, "Inferring static hand poses from a low-cost non-intrusive sEMG sensor," *Sensors,* vol. 19, no. 2, p. 371, 2019.

[29] Y. Miura, Y. Morita, H. Koizumi, and T. Shingai, "Effects of taste solutions, carbonation, and cold stimulus on the power frequency content of swallowing submental surface electromyography," *Chemical senses,* vol. 34, no. 4, pp. 325-331, 2009.

[30] W. Sato, K. Minemoto, A. Ikegami, M. Nakauma, T. Funami, and T. Fushiki, "Facial EMG correlates of subjective hedonic responses during food consumption," *Nutrients,* vol. 12, no. 4, p. 1174, 2020.

[31] Y. Manda, N. Kodama, N. Maeda, and S. Minagi, "Effect of food properties and chewing condition on the electromyographic activity of the posterior tongue," *Journal of oral rehabilitation,* vol. 46, no. 6, pp. 511-517, 2019.

[32] Y. Wang *et al.*, "Qualitative recognition of primary taste sensation based on surface electromyography," vol. 21, no. 15, p. 4994, 2021.

[33] R. Ueda *et al.*, "Characterization of palatability and ease of deglutition of the five basic tastes by partial least squares regression analysis using myoelectric potential parameters of the submental muscle," vol. 28, no. 1, pp. 75-86, 2022.

[34] F. Rahimi, M. A. Badamchizadeh, S. Ghaemi, and A. D. Vecchio, "Simultaneous Estimation of Digit Tip Forces and Hand Postures in a Simulated Real-Life Condition With High-Density Electromyography and Deep Learning," *IEEE Journal of Biomedical and Health Informatics,* vol. 28, no. 10, pp. 5708-5717, 2024, doi: 10.1109/JBHI.2024.3350239.

[35] R. Gupta, V. Kavanal, and A. Joshi, "Deep Learning Models for Recognition of Hand Gestures in Basketball Sports," in *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU)*, 1-2 March 2024 2024, pp. 1-6, doi: 10.1109/IC-CGU58078.2024.10530848.

[36] M. Zanghieri, P. M. Rapa, M. Orlandi, E. Donati, L. Benini, and S. Benatti, "sEMG-Driven Hand Dynamics Estimation With Incremental Online Learning on a Parallel Ultra-Low-Power Microcontroller," *IEEE Transactions on Biomedical Circuits and Systems,* vol. 18, no. 4, pp. 810-820, 2024, doi: 10.1109/TBCAS.2024.3415392.

[37] S. Ma, J. Zhang, C. Shi, P. Di, I. D. Robertson, and Z. Q. Zhang, "Physics-Informed Deep Learning for Muscle Force Prediction With Unlabeled sEMG Signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 32, pp. 1246-1256, 2024, doi: 10.1109/TNSRE.2024.3375320.

[38] T. Wang, J. Li, H.-N. Wu, C. Li, H. Snoussi, and Y. Wu, "ResLNet: deep residual LSTM network with longer input for action recognition," *Frontiers of Computer Science,* vol. 16, no. 6, p. 166334, 2022.

[39] Z. Li, J. Cai, Q. Chen, L. Chen, M. Qing, and S. X. Yang, "An LSTM Network With Neural Plasticity for Driver Fatigue Recognition on Real Roads," *IEEE Transactions on Industrial Electronics,* 2025.

[40] L. Chen, S. Wen, Y. Tang, Y. Ma, H. Wang, and M. W. Geda, "Brain power mapping with optimized deep learning for EEG-based pilot fatigue detection," *Biomedical Signal Processing and Control,* vol. 114, p. 109284, 2026.

[41] T. F. N. Bukht *et al.*, "Robust Human Interaction Recognition Using Extended Kalman Filter," *Computers, Materials & Continua,* vol. 81, no. 2, 2024.

[42] J. He, X. Niu, P. Zhao, C. Lin, and N. Jiang, "From Forearm to Wrist: Deep Learning for Surface Electromyography-Based Gesture Recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 32, pp. 102-111, 2024, doi: 10.1109/TNSRE.2023.3341220.

[43] S. Mekruksavanich, W. Phaphan, and A. Jitpattanakul, "Deep Learning Networks for Human Knee Abnormality Detection Based on Surface EMG Signals," in *2024 47th International Conference on Telecommunications and Signal Processing (TSP)*, 10-12 July 2024 2024, pp. 160-163, doi: 10.1109/TSP63128.2024.10605943.

[44] A. Ullah, Z. Song, W. Riaz, and Y. Wang, "GTMH-TasteNet: Advanced Deep Learning for sEMG-Based Taste Sensation Recognition," *Tsinghua Science and Technology,* 2025, doi: 10.26599/TST.2025.9010143.