

Fog Computing-based Real-Time Emotion Recognition using Physiological Signals

Omur Fatmanur Erzurumluoglu¹, Kadir Ozlem¹, Asli Tuncay Atalay², Ozgur Atalay², Gökhan Ince¹

¹Faculty of Computer and Informatics Engineering, Istanbul Technical University, Istanbul, Turkey

²Faculty of Textile Technologies and Design, Istanbul Technical University, Istanbul, Turkey
{erzurumluoglu18, kadir.ozlem, aatalay, atalayoz, gokhan.ince}@itu.edu.tr

Abstract—Emotion recognition based on physiological signals has become a crucial area of research in affective computing and human-computer interaction, with applications in smart homes, workplaces, educational institutions, healthcare, and entertainment. In this study, a real-time emotion recognition system utilizing fog computing architecture was developed by considering the challenges of latency, total response time, resource usage, and security in IoT environments. The random forest machine learning model was trained with time-based statistical features by using the DREAMER dataset. Even though the model achieved an accuracy of 84.21% with 104 features, to meet real-time performance requirements, the system was optimized to calculate 24 features, maintaining a commendable accuracy of 79.70%. Extensive experiments demonstrated the superior performance of fog computing compared to edge and cloud computing in terms of latency, queuing delay, jitter, and most importantly total response time. The results highlight the proposed system's ability to support multiple users simultaneously.

Index Terms—Emotion recognition, fog computing, machine learning, signal processing, electrocardiogram, electroencephalogram

I. INTRODUCTION

Emotion recognition is a prominent research subject in the fields of affective computing and human-computer interaction. To address the emotion recognition problem, the physiological signals which consist of ElectroEncephaloGram (EEG), ElectroCardioGram (ECG), ElectroMyoGram (EMG), respiration, temperature, and Galvanic Skin Response (GSR) signals, etc. are considered more reliable than facial [1], gesture [2], and speech [3] signals because they are less susceptible to subjective control by individuals. Additionally, using speech signal is often ineffective in noisy environments, and language barriers further limit its practicality. Also, camera-based methods raise significant security concerns, particularly in systems that involve continuous indoor monitoring, where there is a high risk of privacy violations. Furthermore, video data demands more storage and processing resources compared to physiological signals. Therefore, physiological data not only provides a more objective and consistent basis for emotion recognition but is also more secure and efficient to handle.

This study was partially funded by the Scientific Research Project Unit of Istanbul Technical University (Grant No: MDK-2021-43137), by Scientific and Technological Research Council of Turkey (TUBITAK) (Grant No: 221N310), and by the European Research Council (ERC) under the European Union's Horizon Research and Innovation Programme (Grant No: 101042402).

There are several uses for emotion recognition to monitor multi-user emotional states simultaneously in real time. For smart homes, emotion recognition is integrated to provide automation for family members by modifying lighting, music, and other environmental elements according to the overall emotional states of the households [4], [5]. For workplace wellness, employers obtain the ability to track the health of many employees, recognizing signs of stress and tiredness; thus, they can take timely actions for a healthier work environment [6]. In educational settings, emotion recognition provides educators with insights into student engagement and understanding for groups of students [7], [8]. That helps educators adjust their teaching strategies to suit the demands of each class and improves the quality of the learning process overall. In a clinical setting, emotion recognition allows healthcare professionals to monitor the emotional states of multiple patients simultaneously, receiving immediate alerts and providing timely interventions [9]–[11]. In the entertainment industry, especially in multi-player gaming and virtual reality environments, emotion recognition enhances the experience by adapting environmental settings to the emotional states of multiple users in real-time while ensuring low latency and high performance [12], [13].

For those use cases, real-time emotion recognition systems should be implemented on the Internet of Things (IoT) computing architecture. This architecture needs to acquire physiological signals from sensors and process them using prediction models to determine the corresponding emotion. To continuously gather and process these signals, the system requires significant computational power and energy resources. While sensor-to-cloud architectures are preferable due to their scalability and high-performance capabilities, they may not be appropriate due to high bandwidth requirements, security concerns, and potential latency issues [14]. Edge computing, on the other hand, faces several challenges, including limited programmability, scalability concerns for multi-user systems, and difficulties in service management, even though it offers low bandwidth and latency [15]. Given these requirements, fog computing architecture is a suitable candidate to be used for real-time emotion recognition from physiological signals due to its reduced bandwidth requirements, security measures, computation capabilities, effortless integration with existing infrastructure, flexible placement options, and energy efficiency [16].

Consequently, there are several clear benefits that fog computing offers for real-time emotion recognition from physiological signals. Introducing a hierarchical processing structure ensures optimum resource utilization and performance, and efficiently supports numerous users by allowing data to be processed at different locations between the edge and the cloud [17]. Fog computing greatly lowers latency by processing data closer to the source; this is important for applications such as virtual reality, interactive gaming, and mental health monitoring that require immediate responses. By eliminating the need to send sensitive physiological data to the cloud, it improves security and privacy while reducing the risk of data breaches and maintaining compliance with privacy standards. Moreover, it is capable of scaling and accommodating several users simultaneously, while guaranteeing reliability and consistent performance. Additionally, fog computing effectively disperses computational workloads among several nodes, preventing any device from becoming a bottleneck. It reduces the strain on edge devices by distributing the processing load, preserving their battery life, and providing longer run times for wearable or mobile devices used for emotion recognition.

In this study, a real-time emotion recognition system was implemented utilizing fog computing. Initially, an emotion recognition model was trained using features derived from physiological signals through offline experiments with an open-source dataset. Subsequently, to enable real-time emotion prediction from physiological signals, the pre-trained machine learning model was embedded within a fog computing architecture. This architecture also performed normalization and feature extraction operations before prediction. Furthermore, the architecture was evaluated in real-time scenarios using signals from an open-source dataset, assessing its resource utilization, time performance, and network bandwidth usage.

The main contribution of this study lies not only in the development of a high-performing model using an open-source dataset but also in the implementation of an architecture capable of using the model in real-time scenarios for various use cases. Alongside achieving optimal accuracy in model development, considerations were given to real-time usability, time performance, bandwidth requirements, and resource usage.

II. RELATED WORKS

A. Existing Emotion Models

Emotions are classified in various ways among which dimensional emotion space model and discrete emotion model have found widespread use [18]. For the dimensional emotion space model, valence and arousal dimensions are mostly used. The valence dimension indicates the polarity of emotion as either negative or positive, and the arousal dimension shows the intensity of an emotion, so this model distinguishes between states of activation and inactivation. This bipolar emotion model approach is indicated by Russell [19].

The discrete emotion model follows the approach of classifying emotions into discrete emotion labels. Ekman defined the emotion model with six basic emotion states: happy, sad,

anger, fear, surprise, and disgust [20]. Plutchik suggested eight basic bipolar emotions: joy versus sorrow, anger versus fear, acceptance versus disgust, and surprise versus expectancy [21]. Izard defined ten fundamental emotions: interest, joy, surprise, anguish, fear, shyness, guilt, anger, disgust, and contempt [22].

B. Discrete Emotion Recognition using Machine Learning

In the literature, some studies perform 4 discrete emotion recognition using machine learning models. A study by Disanayake et al. [23] that used features obtained from ECG signal achieved an accuracy of 70.09% using the Extra Tree Classifier (ETC). A study by Valenza et al. [24] utilizes Support Vector Machine (SVM) model and extracts features from the instantaneous spectrum and bispectrum of ECG signals. It achieved an accuracy of 79.29%. A study by Zheng et al. [25] applied a feature extraction method into different frequency bands of the EEG signals and achieved an accuracy of 70.33% using SVM model. A study by Kim et al. [26] utilized preprocessing using wavelet-independent component analysis and feature extraction using wavelet packet transform operations with SVM model. It achieved 77.23% accuracy on the DREAMER dataset.

C. Emotion Recognition using IoT Systems

IoT-based real-time emotion recognition systems utilize data such as facial expressions [4], speech [3], and physiological signals [7]. For computational capability, cloud, edge, and fog computing architectures are employed [27].

Chen et al. [3] introduced the I-Fabric system that integrates wearable technology with edge and cloud computing to monitor physical/psychological health and emotion. To determine the emotional state of a user, SVM, decision tree, attention-based bidirectional Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN) models were used. The system was validated through user testing.

Izountar et al. [12] developed a virtual reality-based exercise game system that personalizes rehabilitation by adapting the difficulty of exercises based on real-time emotion recognition. The system utilizes CNN model built on the DeepFace framework by using facial data that is captured through a web camera integrated into the VR setup. The platform processes facial data using edge computing.

Yang et al. [4] proposed an AI-enabled emotion-aware system that integrates a personal robot, smart clothing, and edge cloud computing to enhance mental health monitoring. The system proposed CNN for facial expressions and LSTM networks for physiological signals collected from smart clothing, and entity/relationship recognition. However, the study implemented emotion recognition using only facial expressions with Fer2013 dataset. The study reported an accuracy of 99.9% in classifying sad and happy emotions.

Dai et al. [7] developed a cloud-edge computing-based IoT system for smart classrooms, which integrates various sensor data to recognize students' emotional states. The system collects and processes environmental, identity, and physiological data, including GSR, ECG, and EEG.

TABLE I
COMPARISON OF EXISTING STUDIES WITH PROPOSED STUDY

Author	Model	IoT	Fog Comp.	Signals	Classes	Dataset	Evaluation Metrics
Meng et al. [11]	FFNN	Yes	No	EEG	2	Original dataset	Accuracy, Hit rate
Izountar et al. [12]	CNN	Yes	No	Facial	3	Original dataset	Accuracy, Response time, User experience and satisfaction
Chen et al. [3]	SVM, Decision tree, CNN and LSTM	Yes	No	ECG, Blood pressure, Temperature, Heart rate, Sitting posture, and Speech	N/A	Original dataset	Delay, User experience and satisfaction
Yang et al. [4]	CNN and LSTM	Yes	No	Facial	2	Fer2013	Accuracy, Hit rate
Dai et al. [7]	N/A	Yes	No	Environmental data, Facial, GSR, ECG, EEG	N/A	Original dataset	Surveys with Likert-scale questions
Liu et al. [13]	SLDA and SVM	Yes	No	EEG	2	Original dataset	Accuracy
Ejaz and Shah [8]	CNN	Yes	Yes	Facial	2	N/A	Accuracy
Proposed Study	Random Forest Model	Yes	Yes	EEG and ECG	4	DREAMER	Accuracy, Latency, Queueing delay, Execution time, Response time, Jitter, Network bandwidth, Resource usage

Liu et al. [13] developed a real-time system for recognizing discrete emotions induced by movie clips using EEG signals. The system employs a three-level classification scheme, combining Sparse Linear Discriminant Analysis (SLDA) for feature selection with SVM for classification. Utilizing the original dataset, the system achieved an average accuracy of 92.26% for the arousal dimension and 86.63% for the valence dimension. By using the valence-arousal coordinate space, this study also classified discrete emotions: amusement, joy, tenderness, anger, disgust, fear, and sadness. The study used edge computing for real-time EEG data processing and cloud computing for storing and analyzing the data.

Meng et al. [11] conducted a study for an emotion-aware monitoring system based on EEG signals. The system is designed to predict binary emotions, comfortable and uncomfortable, with a classifier of Feedforward Neural Network (FFNN). The evaluation, which involved an original dataset with 60 participants, demonstrated an accuracy of 85.7%. Cloud and edge computing were used to process the data. The study mentioned that fog computing could be used as an alternative to edge computing.

Ejaz and Shah [8] introduced a real-time system for analyzing student engagement in smart classrooms using fog computing and IoT devices. The proposed system processes facial images to predict binary engagement levels of students with an accuracy of 93% with a fine-tuned CNN model. Their work highlights the potential of integrating fog computing to enhance classrooms to obtain continuous engagement in learning.

Table I includes a comparison between the proposed and existing works in the literature, where the development of fog-based real-time emotion recognition from physiological signals emerges as an open research area.

III. EMOTION RECOGNITION SYSTEM

The pipeline of the proposed real-time emotion recognition system is seen in Fig. 1. Operation blocks shown in the pipeline will be explained in this section. In this work, to ensure a fair comparison, it was decided to use offline and public dataset instead of real-time sensor data.

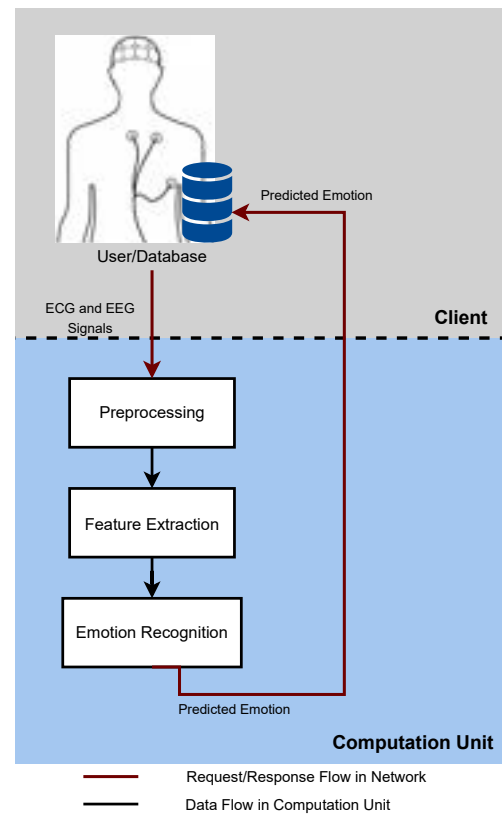


Fig. 1. Real-Time Emotion Recognition Pipeline.

A. Dataset

The DREAMER dataset [28] includes 14 channels of EEG and 2 channels of ECG signal data which were gathered using audio and video stimuli. The ECG signals were recorded at 256 Hz using a SHIMMER wireless sensor. The Emotiv EPOC wireless EEG headset was utilized to capture EEG signals. The EEG channels contain AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. Before applying stimuli, baseline signals were collected from the participants. As recommended by Gabert-Quillen et al. [29], 18 film clips were shown for durations ranging from 65 to 353 seconds to elicit various emotions from the 23 volunteers. These 18

clips, selected to evoke nine specific emotions—amusement, excitement, happiness, calm, anger, disgust, fear, sadness, and surprise—were chosen with 2 clips per emotion.

In our research, amusement, calmness, disgust, and surprise emotions were selected to focus on. This selection ensures that the study addresses a wide spectrum of emotions, from high-arousal (surprise) to low-arousal (calmness), and from positive (amusement) to negative (disgust). The emotion videos in those emotion categories were divided into two sets containing the equal number of videos showing four emotions. One set was used for model training, and the other set was used for model testing. This approach ensures that, unlike in many other studies, the same signal was not segmented and used for both training and testing the model [30]–[33].

B. Preprocessing and Feature Extraction

The ECG data was down-sampled from 256 Hz to 128 Hz, whereas the frequency of EEG signals is already 128 Hz. Each signal was segmented into 3-second length moving windows, and the window was shifted by 1 second. In order to eliminate the information that the windows belong to the same signal, Z-score normalization was applied to each window. After that, 36 statistical features were extracted. Features obtained across simultaneous windows from ECG and EEG signals were concatenated using feature-level fusion. Since 36 features were derived from 16 simultaneous windows, a feature vector with 576 features was obtained. Table II lists extracted features.

TABLE II
EXTRACTED FEATURES FROM DATA

Feature Name	Feature Name
Arithmetic Mean	Arithmetic Mean of Instant Exchange
Median	Quadratic Mean of Instant Exchange
Mean Absolute Deviation	Frequency Above Median
Mean Absolute Value	Frequency Below Median
Minimum Value	Frequency Between Median
Maximum Value	Frequency Above Mean
Range Value	Frequency Below Mean
Skewness	Frequency Between Mean
First Degree Difference	Maximum Consecutive Above Median
Second Degree Difference	Maximum Consecutive Below Median
Average Gradient	Maximum Consecutive Between Median
Minimum Reduction	Maximum Consecutive Above Mean
Maximum Reduction	Maximum Consecutive Below Mean
Minimum Increase	Maximum Consecutive Between Mean
Maximum Increase	Moment m ($m \in \{3, 4, 5, 6\}$)
Kurtosis	Percentile p ($p \in \{25, 75\}$)

In order to select features having higher contribution to classification, the Mean Decrease in Impurity (MDI) [34] feature selection method was used in the training dataset. After applying MDI, the contribution rates of the features were obtained and sorted in descending order.

C. Emotion Recognition

A Random Forest (RF) machine learning (ML) model was employed to predict emotions from ECG and EEG signals. This model is an ensemble method that performs classification using multiple decision trees [35]. A decision tree predicts the value of a target variable by inferring decision rules from data.

It divides the dataset into groups as feasibly as possible in terms of the variable to be predicted [36]. Different subsets of features are used to create a forest of decision trees randomly, enabling the random forest model to handle high-dimensional feature data.

The RF model was trained using features derived from the ECG and EEG signals in the DREAMER dataset. Log loss was utilized as a criterion, and maximum depth was not specified. Once the model is trained, the emotion recognition process begins with real-time signals from the user. These signals are preprocessed using the same steps applied in the training phase. The system provides immediate prediction outputs, making it applicable in real-time.

IV. FOG COMPUTING ARCHITECTURE

For the fog computing architecture of this study, the FogE-*Tex* framework designed for general IoT systems developed by the authors was utilized as described below [37].

A. Hardware Components

Cloud Device: In fog computing, the role of the cloud is to provide centralized management in our system. It was not chosen as the primary computation unit due to high bandwidth requirements, security concerns, and potential latency issues while it can perform complex computations.

Broker Node: The role of the broker is to manage communication between devices in the network. The broker coordinates the distribution of tasks in the fog nodes to optimize performance and reduce latency while ensuring efficient resource allocation.

Worker Node: The role of the worker is to manage data processing tasks at the edge of the network. The worker handles ECG and EEG signal transmission and performs emotion recognition. That eliminates the need for data to be transmitted to the cloud, which minimizes latency and improves response times.

Client: The role of the client is to generate and send data to the fog nodes for processing. Clients are typically end-user devices or sensors that collect data. For this study, the client device is responsible for sending ECG and EEG signals from the DREAMER dataset at a frequency of 32 Hz.

B. Network Setup

The formation of the fog computing architecture of our system begins with the interconnection of the components in the network and provides efficient management for intensive data transmission and execution of ML tasks. Initially, the broker node connects to the cloud, registers itself, and provides information including its IP address, available CPU and memory resources, and current load. This registration allows the cloud to recognize the broker as the main coordinator of fog nodes to manage and distribute tasks effectively. Once the broker node is connected and acknowledged by the cloud, the worker node connects to the cloud to obtain information about the broker. Then, the worker node establishes a connection with the broker node. Because worker nodes are the primary

executors of ML tasks and handle high-density data streams, this connection is crucial. The broker node keeps receiving task requests from client devices and distributes the workload to the worker nodes with the lowest load. Users connect with available worker nodes. This allows two-way communication between the user and the worker node. Consequently, the fog computing system is able to create and sustain a robust, effective network that utilizes localized processing while ensuring seamless integration with the cloud while performing intensive data transfer and completion of ML tasks.

C. Computing System

The computing system of fog architecture is responsible for executing the tasks requested by users. As shown in Fig. 1, initially, the computing system receives raw physiological signals from the user or database, then handles the processing of signals by applying preprocessing, and feature extraction operations. Then, extracted features are applied to the trained emotion recognition model to estimate the user's present emotion in real time.

Cloud, broker, and worker nodes within the architecture are able to perform computational tasks as a computation unit, so computation tasks can be carried out on cloud and broker nodes when needed instead of worker nodes.

V. EXPERIMENTS

A. Experimental Setup

The devices and their respective configurations used in the experiments are listed below.

Cloud Device: Digital Ocean VPS (2x Dedicated Premium Intel CPU, 8GB RAM, 50GB NVMe SSD, Cloud Location: Frankfurt/Germany)

Broker Node: Raspberry Pi 4 Model B (Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit SoC CPU @1.8GHz, 8GB LPDDR4-3200 SDRAM, 32GB Micro SDHC Class 10)

Worker Node: Raspberry Pi 4 Model B (Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit SoC CPU @1.8GHz, 8GB LPDDR4-3200 SDRAM, 32GB Micro SDHC Class 10)

Wi-Fi Access Point (AP): Infinity 401 Modem (ISP: NetMaster, Download speed: 100Mbps, Upload speed: 5Mbps)

Client Device: Asus ROG Strix G15 G513IC (AMD Ryzen 7 4800H CPU @2.90GHz, 16GB DDR4 RAM, 512GB SSD, and Windows 11 Pro 64-bit).

Fig. 2 shows the connection diagram of the devices used for the experimental setup. The client and the fog units reside within the same network. The broker node and worker node are connected to the Wi-Fi AP with ethernet, facilitating communication between the worker, broker, and Wi-Fi AP. The fog node accesses the internet through fiber internet via the Wi-Fi AP, and devices in the network reach the cloud node through internet connection. The client connects to the Wi-Fi AP using the IEEE 802.11ac wireless networking standard. To ensure performance comparisons were conducted in a consistent environment, tests were performed using a single

worker device, given the presence of one cloud node and one broker node in the system.

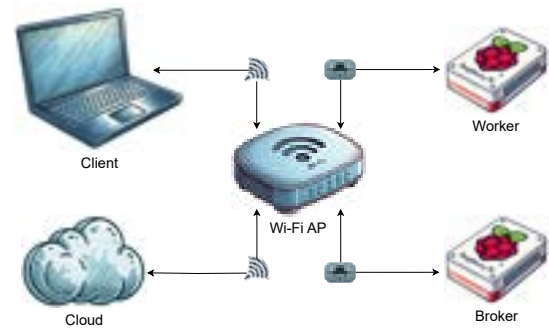


Fig. 2. Experimental setup device connection diagram (generated using the assistance of ChatGPT and DALL-E).

B. Evaluation Methods

1) Evaluation of Recognition Performance: After preprocessing and feature extraction operations were applied to the ECG and EEG signals in the DREAMER dataset, the data obtained using the stimuli from film clips were split evenly into training and test sets for all emotion categories. Instance numbers of data in the training and test sets corresponding to the emotions of participants are shown in Table III. It is noteworthy to mention that the number of instances in the dataset is dependent on the duration of the film clips. For example, while the training set includes physiological signals from a 131-second amusement film clip for all participants, the test set contains signals from a 199-second film clip from the same category of amusement.

TABLE III
INSTANCE NUMBERS CORRESPONDING TO EMOTIONS IN TWO SETS

	Amusement	Calmness	Disgust	Surprise
Training set	2838	8624	3938	4048
Test set	4334	4180	2068	5588

2) Evaluation of Transmission Performance: The fog computing architecture was built using a pre-trained model to test it in real-time. The performance of the architecture was evaluated through a system transmitting recorded physiological signals in the DREAMER dataset from the client. The mock user sent 32 requests per second, with 31 of them streaming ECG and EEG data to the computation unit and adding this data to a buffer. On the 32nd request, the computation unit processed the accumulated data, which represented the latest 1 second of ECG and EEG signals appended to the previous 2 seconds, resulting in a total of 3 seconds of data. This 3-second data was then normalized, and the relevant features were calculated. These feature vectors were merged, and the emotion was predicted using the learned model and returned to the client.

Only the emotion recognition requests were considered in the experiment results, while the data streaming requests were disregarded.

The worker node is primarily responsible for computational tasks for the clients; however, the client device, broker node,

and cloud node can also provide these services. Thus, the worker node, broker node, edge device, and cloud node are all classified as computation units, and experiments were conducted on four of them. A client connected arbitrarily to these devices to receive emotion recognition service. To assess the time characteristics, CPU load, memory usage, and bandwidth of the four computation units, amusement emotion signals from the first participant in the DREAMER dataset, spanning 199 seconds, were utilized. Emotion recognition was performed in real-time with 3-second windows for each request per second, resulting in a total of 197 emotion recognition requests and 6107 data stream requests.

To examine the impact of the number of features calculated on the system's response time and to determine the final system configuration based on accuracy and time efficiency evaluation metrics, the system was tested with varying numbers of features. After sorting the features in order of importance, the number of features used was selected from 4 to 576 in increments of 4.

C. Evaluation Criteria

The metrics used in the evaluation of the proposed system proposed system's recognition performance are accuracy, precision, recall, and f1 score whereas the transmission performance is measured using the metrics below.

Latency: It is the amount of time that elapses between a client sending data and the data arriving in the computation unit.

Queuing Delay: It is the amount of time that a request remains in the queue before execution.

Execution Time: It is the total duration for which the emotion recognition pipeline steps are sequentially applied to incoming requests.

Response Time: It is the total amount of time between sending the emotion recognition requests request and receiving the response.

Jitter: It represents the variation in total response time.

Resource Usage: It represents CPU and memory usage data for the respective computation units. Elevated resource utilization may result in heightened system latency and diminished service capacity for concurrent users.

Network Bandwidth Usage: It represents the total of incoming and outgoing data. The total network bandwidth usage for different computation units was considered.

VI. RESULTS AND DISCUSSION

A. Results of Emotion Recognition

The model was trained and tested as explained in Section V-B1. It achieved an 80.12% accuracy, 80.12% recall, 83.67% precision, and 80.27% F1 score as shown in Table IV. According to those evaluation metric results, the model shows balanced performance. These results were obtained on the client device detailed in Section V-A.

Table IV shows the comparison between existing studies that perform discrete emotion analysis using ML models and our study. Our study outperforms existing studies that utilize

ML models for discrete emotion recognition from ECG and EEG signals in terms of accuracy.

TABLE IV
PERFORMANCE METRICS OF THE MODEL

Paper	Method	Dataset	Signals	Classes	Accuracy
[23]	ETC	Original	ECG	4	70.09%
[24]	SVM	Original	ECG	4	79.29%
[25]	SVM	SEED-IV	EEG	4	70.33%
[26]	SVM	DREAMER	EEG	4	77.23%
Our study	RF	DREAMER	ECG, EEG	4	80.12%

The plot in Fig. 3 illustrates the relationship between response time and accuracy as the number of features increases processed in the worker which is the primary device for computational tasks. As the number of features increases, response time exhibits an upward trend because of the computing load of features. Accuracy, depicted on the secondary y-axis, peaks at 84.21% with 104 features. The stability that occurs around 82% suggests that the increase in the number of features does not significantly impact accuracy. However, a slight decline is observed after 200 features indicating potential overfitting or redundancy in the feature set.

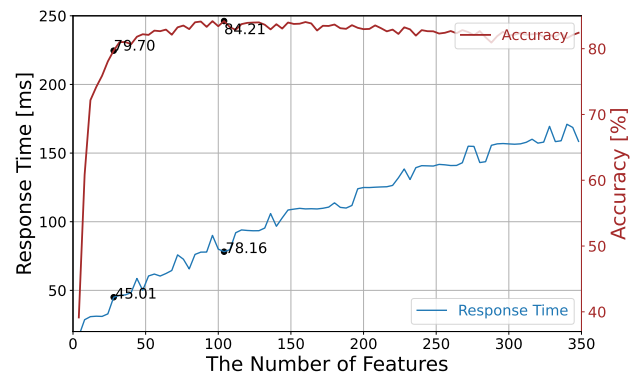


Fig. 3. Comparison of response time of worker node and accuracy results in terms of the number of features.

In this study, the main goal was not only to obtain a high-performance model but also to design a real-time emotion recognition system by taking into account all evaluation criteria. Therefore, instead of the model with 84% accuracy achieved with 104 number of features and an emotion response time of 78.16 ms, a model with 24 number of features achieving 80% accuracy and a response time of 45.01 ms, half the duration of the best model, was chosen to get optimum results for real-time emotion recognition system. Although there is a 5% loss in accuracy, the system operates approximately with half the response time. The evaluation criteria results are based on tests conducted on the system with 24 features.

B. Results on Transmission Performance

1) Results on Latency: Fig. 4 gives the mean latency for different computing devices. The data shows that the cloud node had the highest mean latency at 32.2 ms due to its multi-hop connection and longer data travel distances. In contrast, the edge device demonstrates the lowest mean

latency at 1.6 ms because the data source also performs as a computational unit. The worker and broker nodes show mean latencies of 4.7 and 4.8 ms respectively. These results show the potential of edge device and fog computing architecture to significantly reduce latency, making them more suitable for real-time emotion recognition applications that require swift and responsive data processing.

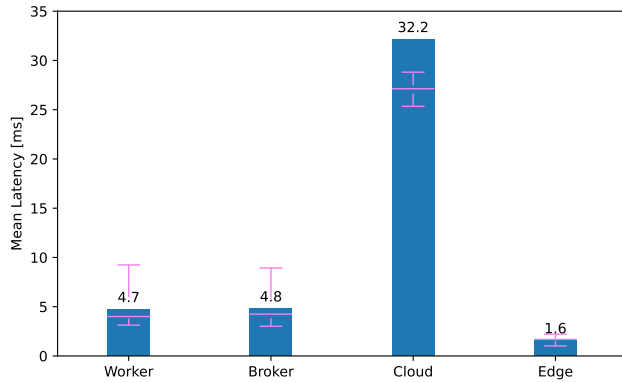


Fig. 4. Latency results.

2) *Results on Queuing Delay:* The mean queuing delay results for different computing devices are presented in Fig. 5. The data shows that the edge device has a significantly higher mean queuing delay of 13.7 ms, compared to the other devices, which all exhibit mean queuing delays below 1 ms. This difference shows that the client device demonstrates substantial queuing delays due to its inherent network characteristics or load-handling capacity. In contrast, worker, broker, and cloud nodes demonstrate comparable and minimal queuing delays, indicating more efficient queuing processes. Worker and broker nodes show slightly better performance compared to the cloud node.

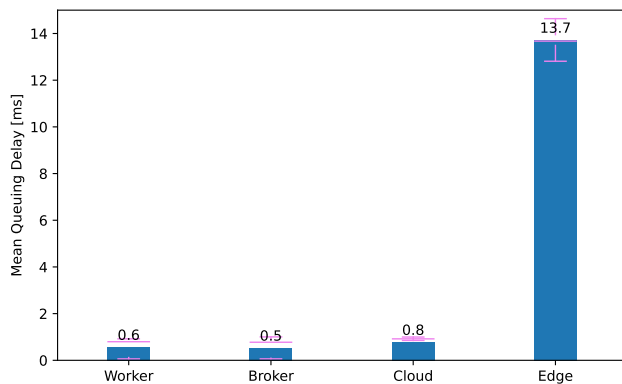


Fig. 5. Queuing delay results.

3) *Results on Execution Time:* The mean execution time results for various computing architectures are displayed in Fig. 6. As expected, the worker and broker nodes exhibit similar and the highest mean execution times of 34.7 ms and 35.3 ms, respectively because they have the same processor power. On the other hand, the cloud and edge nodes show lower mean execution times, with the cloud node at 8.9 ms

and the edge node at 5.7 ms as they have more powerful processors.

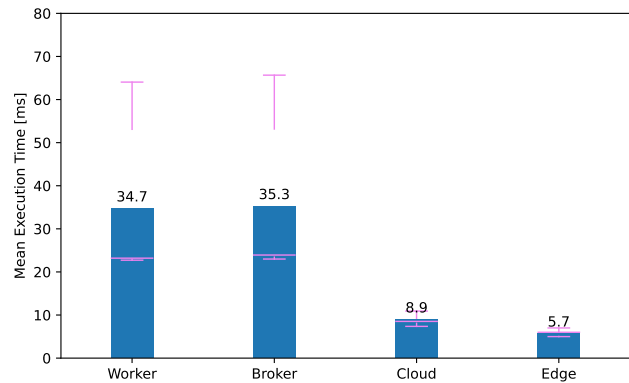


Fig. 6. Execution time results.

4) *Results on Response Time:* Fig. 7 provides the total response time in different devices. Worker device has a lower response time than the broker and cloud devices, indicating that workers perform their duty both adequately and effectively in the fog architecture. The cloud component has the highest response time at 73.3 ms, followed by the worker and broker devices at 43.3 ms and 43.9 ms, respectively. The edge device exhibits the lowest mean response time at 22.5 ms. Although the edge device has the lowest mean response time, it is not suitable for our real-time emotion recognition system which demands extensive and simultaneous multi-user interactions and integrations. Also, because the system needs continuous battery-powered operations, it can result in increased battery drain in edge devices. Furthermore, our system requires centralized data analysis to ensure consistency and compliance with regulatory standards, which is more effective on fog computing systems rather than on edge devices.

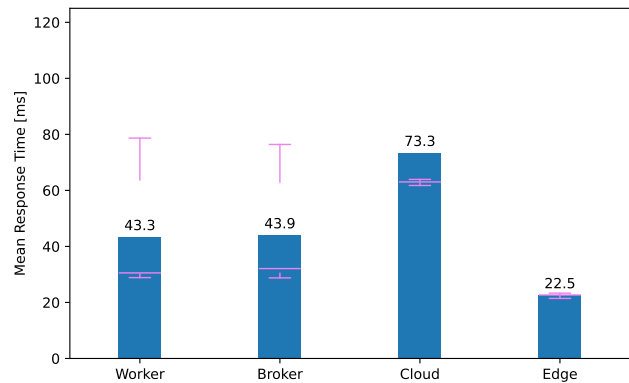


Fig. 7. Total response time results.

5) *Results on Jitter:* In Fig. 8, The results indicate that the cloud node demonstrates the highest jitter at 36.6 ms, suggesting a significant variability in latency. The worker and broker nodes show comparable jitter levels of 17.4 ms and 16.8 ms, respectively, indicating more stable and consistent network performance compared to the cloud node.

6) *Results on Network Bandwidth Usage:* The analysis of mean bandwidth usage across different computing layers re-

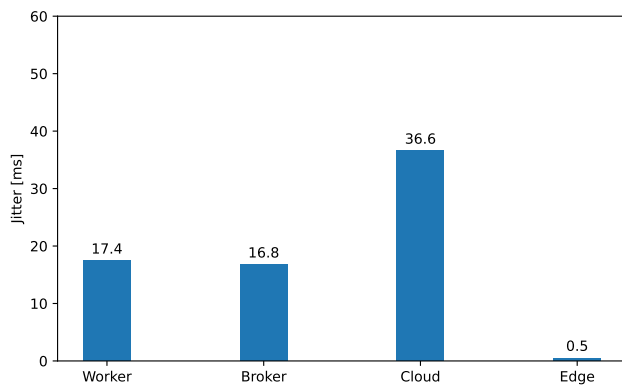


Fig. 8. Jitter results.

veals in Fig. 9. The worker, broker, and cloud layers exhibit mean bandwidths of 373.9 kbps, 392.9 kbps, and 384.2 kbps respectively, indicating relatively high and consistent network usage. In contrast, the edge layer demonstrates a dramatically lower mean bandwidth of 8.5 kbps.

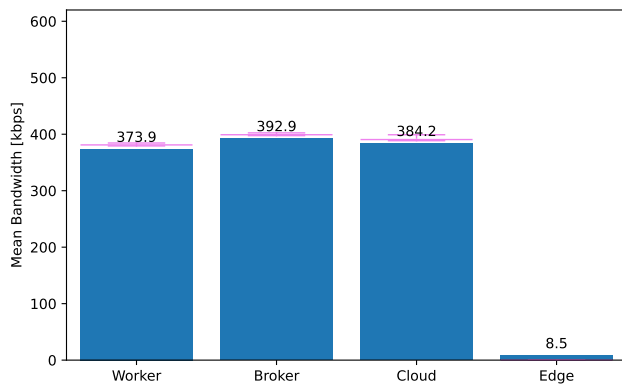


Fig. 9. Network bandwidth usage in different devices.

7) *Results on Resouce Usage:* Fig. 10 shows the CPU and memory usage results of computing service devices. The worker and broker nodes exhibit similar and balanced resource usage, with CPU usage at 6.7% and memory usage at 7.2% because they have the same hardware system. The CPU usage of cloud node is similar to broker and worker devices with 6.2% although it is relatively more powerful and is optimized for computation. And, cloud node shows a slightly higher memory usage at 12.2%. Moreover, the edge device stands out with minimal CPU usage at just 0.4%, but significantly higher memory usage at 41.9%.

C. Discussion

When a user connects to our real-time emotion recognition system, it is assigned to a worker node within the fog computing architecture, and the user's physiological signals begin to be streamed to the worker and are added to the worker's buffer. Data streaming occurs 32 times per second. Once the buffer contains 3 seconds of data, the emotion recognition process commences every second. While receiving 32 data streams per second from the user, the oldest data is deleted and new data is added to the buffer, so the buffer always contains 3 seconds

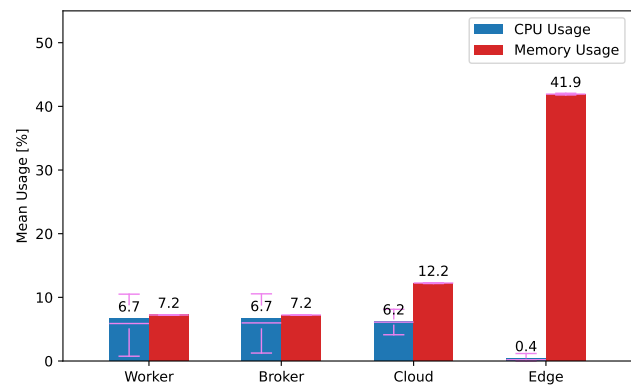


Fig. 10. CPU and memory usage in different devices.

of data. Every second, the system uses the 3-second data in the buffer to recognize the current emotion of the user and then sends the predicted emotion to the user. This procedure takes approximately 43 ms, including a latency of 4.7 ms and an execution time of 34.7 ms. In that case, a single worker device can serve approximately 20 users. Also, the system can be scaled to serve more users by adding additional worker devices.

In terms of latency, total response time, queuing delay, and jitter, the worker device has demonstrated successful performance among cloud and edge devices in our system. A study by Chen et al. [3] extracted 743-dimensional features from speech and fed it to an attention-based bidirectional LSTM and a CNN model for emotion recognition. The procedure took 265.8 ms. A study by Gia et al. [38] extracted heart rate, P wave, and T wave features from ECG signals. This feature extraction procedure took 96.3 ms. In terms of response time, our fog computing system outperforms existing studies that extract features from signals in real time. Therefore, these results prove that our fog computing architecture has the required qualities for a real-time emotion recognition system using physiological signals.

VII. CONCLUSION AND FUTURE WORK

In this study, we developed a real-time emotion recognition system based on physiological signals utilizing fog computing architecture. This system is especially suited for environments where the emotional data of users can be processed in places such as educational institutions, workplaces, smart homes, and healthcare facilities. Random forest machine learning model was used and to train the model time-based statistical features of physiological signals were calculated using the DREAMER dataset. After feature selection was applied, 84.21% accuracy was achieved with 104 features. However, considering real-time system requirements such as resource utilization and time performance, the system has been configured to calculate 24 features from physiological signals with 79.70% accuracy. The emotion recognition procedure commences every second, and takes approximately 43 ms, allowing a single worker to serve around 20 users.

The usage of the fog computing architecture for the system was tested in terms of latency, queuing delay, jitter, total

response time, and resource usage to prove its superiority over edge or cloud computing. The experimental results showed that a worker node can efficiently handle computational tasks with low latency, queueing delay, and most importantly total response time. In overall, The emotion recognition procedure commences every second, and takes approximately 43 ms, including a latency of 4.7 ms and an execution time of 34.7 ms. The system can support 20 users with a single worker device, and can be scaled to serve more users by adding additional worker devices.

There are several unexplored research directions related to this study, optimizing computational resources, expanding the dataset and emotion classes, conducting real-world trials to assess the practical usability of the proposed system, and implementing deep learning models within the fog computing framework.

ACKNOWLEDGMENT

Omur Fatmanur Erzurumluoglu would like to thank the DeepMind Scholarship Programme for their support. We are grateful to Digital Ocean for providing their Cloud Device.

REFERENCES

- [1] Neha Jain, Shishir Kumar, Amit Kumar, Pourya Shamsolmoali, and Masoumeh Zareapoor. Hybrid deep neural networks for face emotion recognition. *Pattern Recognition Letters*, 115:101–106, 2018.
- [2] Jingjie Yan, Guanming Lu, Xiaodong Bai, Haibo Li, Ning Sun, and Ruiyu Liang. A novel supervised bimodal emotion recognition approach based on facial expression and body gesture. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 101(11):2003–2006, 2018.
- [3] Min Chen, Yingying Jiang, Nadra Guizani, Jun Zhou, Guangming Tao, Jun Yin, and Kai Hwang. Living with i-fabric: Smart living powered by intelligent fabric and deep analytics. *IEEE Network*, 34(5):156–163, 2020.
- [4] Jun Yang, Rui Wang, Xin Guan, Mohammad Mehdi Hassan, Ahmad Almogren, and Ahmed Alsanad. Ai-enabled emotion-aware robot: The fusion of smart clothing, edge clouds and robotics. *Future Generation Computer Systems*, 102:701–709, 2020.
- [5] Yekta Said Can and Cem Ersoy. Smart affect monitoring with wearables in the wild: An unobtrusive mood-aware emotion recognition system. *IEEE Transactions on Affective Computing*, 14(4):2851–2863, 2022.
- [6] Zhongshan Chen, Xinning Feng, and Shengwei Zhang. Emotion detection and face recognition of drivers in autonomous vehicles in iot platform. *Image and Vision Computing*, 128:104569, 2022.
- [7] Zhicheng Dai, Qianqian Zhang, Liang Zhao, Xiaoliang Zhu, and Dongbo Zhou. Cloud-edge computing technology-based internet of things system for smart classroom environment. *International Journal of Emerging Technologies in Learning (iJET)*, 18(8):79–96, 2023.
- [8] Javeria Ejaz and Munam A Shah. Real-time analysis of student's behavioural engagement in digital smart classrooms using fog computing and iot devices. In *Competitive Advantage in the Digital Economy (CADE 2022)*, volume 2022, pages 49–54. IET, 2022.
- [9] Tao Zhang, Minjie Liu, Tian Yuan, and Najla Al-Nabhan. Emotion-aware and intelligent internet of medical things toward emotion recognition during covid-19 pandemic. *IEEE Internet of Things Journal*, 8(21):16002–16013, 2020.
- [10] Heng Yu and Zhiqing Zhou. Optimization of iot-based artificial intelligence assisted telemedicine health analysis system. *IEEE access*, 9:85034–85048, 2021.
- [11] Weizhi Meng, Yong Cai, Laurence T Yang, and Wei-Yang Chiu. Hybrid emotion-aware monitoring system based on brainwaves for internet of medical things. *IEEE Internet of Things Journal*, 8(21):16014–16022, 2021.
- [12] Youssa Izountar, Samir Benbelkacem, Samir Otmene, Abdallah Khababa, Mostefa Masmoudi, and Nadia Zenati. Vr-peer: A personalized ex-game platform based on emotion recognition. *Electronics*, 11(3):455, 2022.
- [13] Yong-Jin Liu, Mingjing Yu, Guozhen Zhao, Jinjing Song, Yan Ge, and Yuanchun Shi. Real-time movie-induced discrete emotion recognition from eeg signals. *IEEE Transactions on Affective Computing*, 9(4):550–562, 2017.
- [14] Madhulika Bhatia, Shubham Sharma, Surbhi Bhatia, and M Alojail. Fog computing mitigate limitations of cloud computing. *International Journal of Recent Technology and Engineering (IJRTE)*, SCOPUS, 2020.
- [15] Weisong Shi, Jie Cao, Quan Zhang, Youhui Li, and Lanyu Xu. Edge computing: Vision and challenges. *IEEE internet of things journal*, 3(5):637–646, 2016.
- [16] Punit Gupta, Dinesh Kumar Saini, Pradeep Rawat, and Kashif Zia. *Bio-Inspired Optimization in Fog and Edge Computing Environments: Principles, Algorithms, and Systems*. CRC Press, 2023.
- [17] Hadi Zahmatkesh and Fadi Al-Turjman. Fog computing for sustainable smart cities in the iot era: Caching techniques and enabling technologies-an overview. *Sustainable cities and society*, 59:102139, 2020.
- [18] Nida Manzoor Hakak, Mohsin Mohd, Mahira Kirmani, and Mudassir Mohd. Emotion analysis: A survey. In *2017 international conference on computer, communications and electronics (COMPTELIX)*, pages 397–402. IEEE, 2017.
- [19] James A Russell. Affective space is bipolar. *Journal of personality and social psychology*, 37(3):345, 1979.
- [20] Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.
- [21] Robert Plutchik. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, 89(4):344–350, 2001.
- [22] Carroll E Izard. *Human emotions*. Springer Science & Business Media, 2013.
- [23] Theekshana Dissanayake, Yasitha Rajapaksha, Roshan Ragel, and Isuru Nawinne. An ensemble learning approach for electrocardiogram sensor based human emotion recognition. *Sensors*, 19(20):4495, 2019.
- [24] Gaetano Valenza, Luca Citi, Antonio Lanatà, Enzo Pasquale Scilingo, and Riccardo Barbieri. Revealing real-time emotional responses: a personalized assessment based on heartbeats dynamics. *Scientific reports*, 4(1):4998, 2014.
- [25] Wei-Long Zheng, Wei Liu, Yifei Lu, Bao-Liang Lu, and Andrzej Cichocki. Emotionmeter: A multimodal framework for recognizing human emotions. *IEEE transactions on cybernetics*, 49(3):1110–1122, 2018.
- [26] Sun-Hee Kim, Hyung-Jeong Yang, Ngoc Anh Thi Nguyen, Sunil Kumar Prabhakar, and Seong-Wan Lee. Wedea: A new eeg-based framework for emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, 26(1):264–275, 2021.
- [27] Xu Xu, Chong Fu, David Camacho, Jong Hyuk Park, and Junxin Chen. Internet of things for emotion care: Advances, applications, and challenges. *Cognitive Computation*, pages 1–21, 2024.
- [28] Stamos Katsigiannis and Naeem Ramzan. Dreamer: A database for emotion recognition through eeg and eeg signals from wireless low-cost off-the-shelf devices. *IEEE journal of biomedical and health informatics*, 22(1):98–107, 2017.
- [29] Crystal A Gabert-Quillen, Ellen E Bartolini, Benjamin T Abravanel, and Charles A Sanislow. Ratings for emotion film clips. *Behavior research methods*, 47:773–787, 2015.
- [30] Yuxuan Zhao, Xinyan Cao, Jinlong Lin, Dunshan Yu, and Xixin Cao. Multimodal affective states recognition based on multiscale cnns and biologically inspired decision fusion model. *IEEE Transactions on Affective Computing*, 14(2):1391–1403, 2021.
- [31] Jiaxin Ma, Hao Tang, Wei-Long Zheng, and Bao-Liang Lu. Emotion recognition using multimodal residual lstm network. In *Proceedings of the 27th ACM international conference on multimedia*, pages 176–183, 2019.
- [32] Pritam Sarkar and Ali Etemad. Self-supervised eeg representation learning for emotion recognition. *IEEE Transactions on Affective Computing*, 13(3):1541–1554, 2020.
- [33] Haiping Huang, Zhenchao Hu, Wenming Wang, and Min Wu. Multimodal emotion recognition based on ensemble convolutional neural network. *IEEE Access*, 8:3265–3271, 2019.
- [34] Leo Breiman, Jerome Friedman, R.A. Olshen, and Charles J. Stone. *Classification and regression trees*. Chapman and Hall/CRC, 1984.
- [35] Tin Kam Ho et al. Proceedings of 3rd international conference on document analysis and recognition. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 1995.
- [36] J. Ross Quinlan. Induction of decision trees. *Machine learning*, 1:81–106, 1986.
- [37] Kadir Ozlem, Asli Tuncay Atalay, Ozgur Atalay, and Gökhan Ince. Fogtex: Fog computing framework for electronic textile applications. *IEEE Internet of Things Journal*, 2024.
- [38] Tuan Nguyen Gia, Mingzhe Jiang, Amir-Mohammad Rahmani, Tomi Westerlund, Pasi Liljeberg, and Hannu Tenhunen. Fog computing in healthcare internet of things: A case study on eeg feature extraction. In *2015 IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing*, pages 356–363. IEEE, 2015.