IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 September 2023, accepted 11 October 2023, date of publication 20 October 2023, date of current version 25 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3326259

RESEARCH ARTICLE

HeMoFi4Q: Morse Communication Based on Wi-Fi and Head Motion for Quadriplegia With Environmental Robustness

MARWA R. M. BASTWESY[1,3], (Graduate Student Member, IEEE), HYUCKJIN CHOI[2], (Member, IEEE), AND YUTAKA ARAKAWA[2], (Member, IEEE)

[1]Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka 819-0395, Japan
[2]Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka 819-0395, Japan
[3]CCE, Tanta University, Tanta 31527, Egypt

Corresponding author: Marwa R. M. Bastwesy (bastwesy.marwa.377@s.kyushu-u.ac.jp)

This research is supported by the Grant-in-Aid for Scientific Research (KAKENHI No.JP19H05665) of Japan Society for the Promotion of Science (JSPS) and the Otsuka Toshimi Scholarship Foundation.

ABSTRACT Quadriplegics face a communication obstacle as their physical abilities are restricted, leaving them unable to speak or use their limbs, with only their upper neck being mobile. So, we propose a recognition system and a new communication language utilizing Morse code and head movements, to break this barrier. We aim to overcome the limitations of camera-based and wearable-sensor methods, including occlusion, privacy concerns, and user inconvenience. The goal is to passively detect quadriplegics' head movements and map them to their corresponding character. The dataset including all 26 alphabet letters, was gathered in various settings, including single-user and multi-human environments, with multiple locations for each setting. For evaluation, 2% samples are randomly selected from the unseen environment to be used with the seen environment as a training dataset. Based on the results, our system demonstrates practical feasibility for real-world implementation, with accuracy rates of 94% and 80% achieved in single-user and multi-human environments, respectively.

INDEX TERMS Wireless sensing, Wi-Fi CSI, quadriplegia, deep learning, location diversity.

I. INTRODUCTION

According to World Health Organization (WHO)[1] reports that between 250,000 to 500,000 people suffer from Spinal Cord Injury (SCI) [1] due to avoidable incidents such as accidents, falls, and violent crimes. SCI disrupts body systems and causes significant changes and limitations in many areas of life, resulting in disability. Quadriplegia, which affects the legs, arms, and trunk, is a common result of SCI. Patients suffering from C5 damage lose speech function, making it impossible to use known sign language for communication. To give them the spirit to complete their life and improve their emotional status, we propose a system that translates head nodding into characters to improve communication and emotional well-being for patients and their relatives, aiding in understanding patients' needs and pain sites. The capabilities of sign language recognition techniques have been investigated in many devices, based on computer vision such as infrared and depth image sensors [2], [3], wearable sensors [4], [5], and radio frequency signals [6], [7], [8]. However, camera-based systems require good lighting conditions and raise privacy concerns and bad performance in non-line-of-sight (non-LOS) scenarios. Furthermore, users of wearable sensors may find it uncomfortable to wear them constantly for them to function. In contrast, wireless sensing mechanism based on Radio Frequency (RF) signals has been taken into consideration in the context of interior environments owing to its vast range of coverage, abundance of availability, noninvasive nature, user privacy, and identity protection. Recently, RF signals have shown a rise in passive sensing techniques. The utilization of radar [9], RFID [10], and Wi-Fi signals [11], [12] is important in various applications, including indoor

The associate editor coordinating the review of this manuscript and approving it for publication was Renato Ferrero.

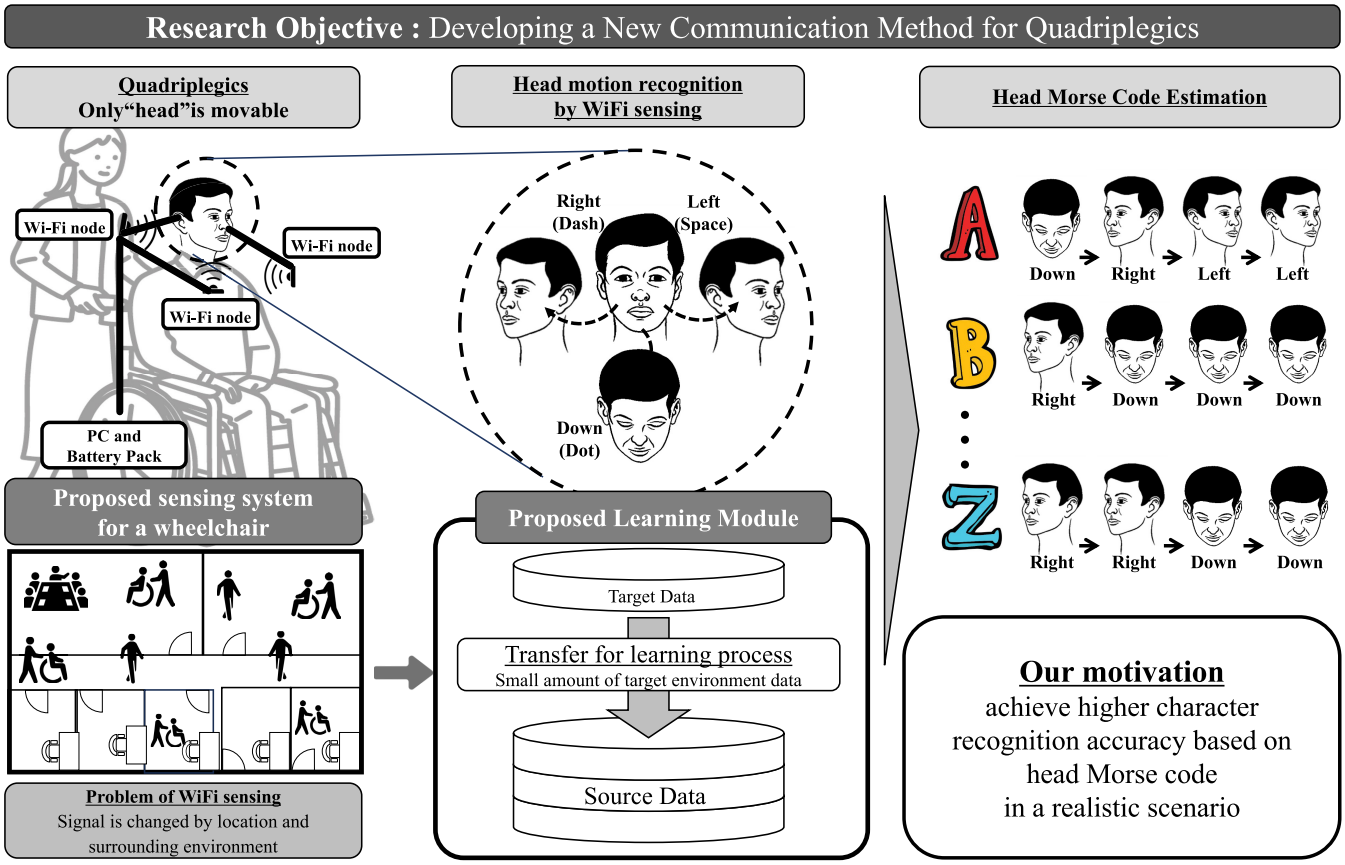

FIGURE 1. Conceptual diagram of the proposed morse code based on the head-motion framework. The system consists of three modules: data collection of CSI reading of the 26 English alphabets, the noise removal techniques and learning module, and the classification module.

localization [13], [14], [15], human behavior recognition (HAR) [16], and identification [17], and healthcare, where they play a key role in tracking and identifying objects and people. In [18], the authors leverage the radar signals from an ultra-wideband (UWB) sensor along with Wi-Fi channel state information (CSI) measurements through universal software radio peripheral (USRP) devices to read lips under the mask. The lip reading with mask system achieves recognition accuracy above 80% based on deep learning (DL) models. Meanwhile, RFID and radar-based systems are contactless sensing techniques and can track movement and detect human presence without capturing any visual data, they require special hardware with expensive installation costs. While, Wi-Fi sensing systems make use of the infrastructure and communication channels already in place, such as Wi-Fi routers. Moreover, it can detect and track macro and micro motions in both line of sight (LOS) and non-LOS.

To overcome the above limitations of existing sensing systems, the proposed system shown in Fig. 1 detects the head motions based on the Wi-Fi signals. Nowadays, Wi-Fi signals are considered a promising sensing technique because it is ubiquitous and readily available, as they are commonly used in homes, offices, and public spaces. This means that the proposed system could potentially be implemented without the need for additional sensors to be attached to the target object or special hardware to be installed. Furthermore, wireless signals can track the patients at any location and achieve better performance thanks to deep learning techniques. The head gestures act as scatters that result in variations in the channel frequency response (CFR) which also refers to CSI. Unique signatures associated with head motion are provided to the DL algorithm to classify into their appropriate character by capturing, monitoring, and analyzing these variations. Herein, we have leveraged Wi-Fi CSI waveforms to track the head motions in a passive manner and extract the gesture signatures of each character. The data was gathered using a real wheelchair and ESP32 microcontroller. The previous HAR CSI systems tackle the domain independent problem using few-shot learning in a single-user environment which is different from the real-time scenarios. Inspired by the few shot algorithms, we have tackled the location robustness issue in a multi-human context environment by adding a few samples from the unseen environment to the seen environment. To achieve the domain independent, the impact of amplitude and phase features is studied to improve the recognition accuracy with the smallest samples from the target source. The classification findings emphasize the effectiveness of diagonal links configuration

and smoothing the time-domain CSI amplitude variations gives the best in different domain independent scenarios.

Our main contributions are as follows:

- We propose a new communication method that uses Morse code generated by a head motion. Also, we propose to apply Wi-Fi sensing for detecting a head motion without a camera and wearables.
- We examine the effects of location diversity in a multi-context environment where patients are surrounded by others, making it more practical and closely aligned with real-world scenarios.
- Inspired by few shot learning algorithms, we merged randomly small samples from the target environment in the learning phase to improve the system performance to make the classifier able to extract the signature of each alphabet.
- We conducted the practical experiments in two different environments with different locations, and also gave diverse analytic comparisons of system performance by studying the impact of different base signals, different link configurations, and state-of-art classifiers.

In this paper, we follow a specific structure. Firstly, we review the existing literature on the topic in Section II. Then, in Section III, we provide a detailed explanation of the theory and properties of Wi-Fi CSI and introduce our proposed system. The evaluation results for our system are presented in Section IV, followed by a discussion of these results in Section V. Finally, we conclude this paper in Section VI.

II. RELATED WORK

Numerous studies have already been conducted that utilize Wi-Fi CSI to identify and detect human activities. These studies typically involve either manually extracting features from the data and using machine learning models or utilizing deep learning methods for automatic feature extraction.

A. HEAD MOTION SYSTEMS BASED ON WI-FI CSI

In our previous work [19], we presented a passive system called Wi-Nod, which utilizes Wi-Fi CSI to detect head nodding gestures based on time-frequency features to detect three basic symbols dot, dash, and space. These symbols are the base blocks for HeMoFi4Q system to build alphabet characters based on the combination of these symbols. Wi-Nod system checked the user and session diversity robustness in a fixed location with people surrounded to the target subject. The results indicate that the proposed system achieves over 95% recognition accuracy, demonstrating its feasibility for real-life deployment.

In [20], This system provides driver's face localization based on the variance of CSI amplitude and phase, with the aim of detecting distraction and fatigue activity. It consists of three modules: CSI preprocessing, feature extraction, and classification. The noise removal techniques employed are the Butterworth filter for amplitude and linear transformation for phase. The feature set includes mean, standard deviation, median absolute deviation, maximum peaks, 25th percentile, and 75th percentile, which are fed to the classification module. The system utilizes the KSVM classifier, which combines the advantages of SVM and KNN and achieves a recognition accuracy of over 91%.

The WiHead system [21] is a system that utilizes wireless signals to measure human head orientation in various directions, including yaw, roll, pitch, and their combinations, for obtaining feedback in online courses. It uses 56 subcarriers at 2.4 GHz with the Atheros CSI extractor tool to retrieve phase and amplitude, which are then filtered and calibrated to eliminate noise and unpredictability. The filtered data is then sent to a PCA method for dimensional reduction. Additionally, WiHead developed a CNN model that achieved recognition accuracy of 90% for three different head motion angles: pitch, roll, and yaw.

In [22], the authors utilize the CSI waveforms from ESP32 microcontroller to track the head motions. They aim to estimate the student's engagement in online courses based on their head movements. CSI amplitude is extracted and preprocessed via Hampel filter, discrete wavelet transforms and smoothed by Savitzky-Golay, After that, the filtered amplitude is fed to the XGBoost (XGB) model for the classification task and achieved 98% recognition accuracy.

B. ENVIRONMENTAL INDEPENDENT SYSTEMS BASED ON WI-FI CSI

Some recent studies have developed methods for recognizing human activities in different environments using Wi-Fi signals.

Authors in [23] present ReWis, which is a single-user environment human activity recognition system. They investigate the impact of using multiple receivers and multiple antennas per receiver to improve the system's accuracy. ReWis reduces the dimensionality of the time diversity based on singular value decomposition (SVD) and captures the correlation across subcarriers by estimating the Pearson correlation coefficients. During the training phase, the CSI waveforms of the source environment are fed to CNN model for representative feature extraction besides five samples of the four activities from the unseen/target environment to reduce the gap between the seen and unseen environment. The ProtoNet model aims to explore the similarity between the two different domains to solve the domain independent issue.

Another HAR work was introduced in [24] where the conjugate multiplication (CM) is utilized to enhance the CSI waveforms quality. CM aims to remove the randomness of the phase by capturing the best CSI quality and setting it as a reference. In other words, the reference CSI vector contains the maximum ratio of the mean of CSI amplitude and standard deviation over the subcarriers. Then, CM is calculated between all transmitter-receiver pairs and the reference vector. The authors apply principal component analysis for dimensionality reduction. Furthermore, to capture the activity

TABLE 1. Summary of related Wi-Fi CSI work.

Reference	Application	Base Signal	Preprocessing	Classifier	Location Independent	Multi-user Environment
[19]	Head motion detection	Amp	Short Time Fourier Transform(STFT)	Inception model	No	Yes
[20]	Driver's face localization	Amp+Phase	Butterworth filter+ Linear Transformation	SVM + KNN	No	No
[21]	Head motion detection	Amp+ Phase	STFT+ PCA+ Linear Transformation	CNN	No	No
[22]	Head motion detection	Amp	Savizky-Golay	XGBoost	No	No
[23]	HAR	Amp	SVD	CNN + ProtoNet	Yes	No
[24]	HAR	Amp+ Phase	CM + PCA +FFT	CNN-LSTM + MatNet	Yes	No
[25]	HAR	Amp+ Phase	Phase Sanitiziation	Inception model	Yes	No

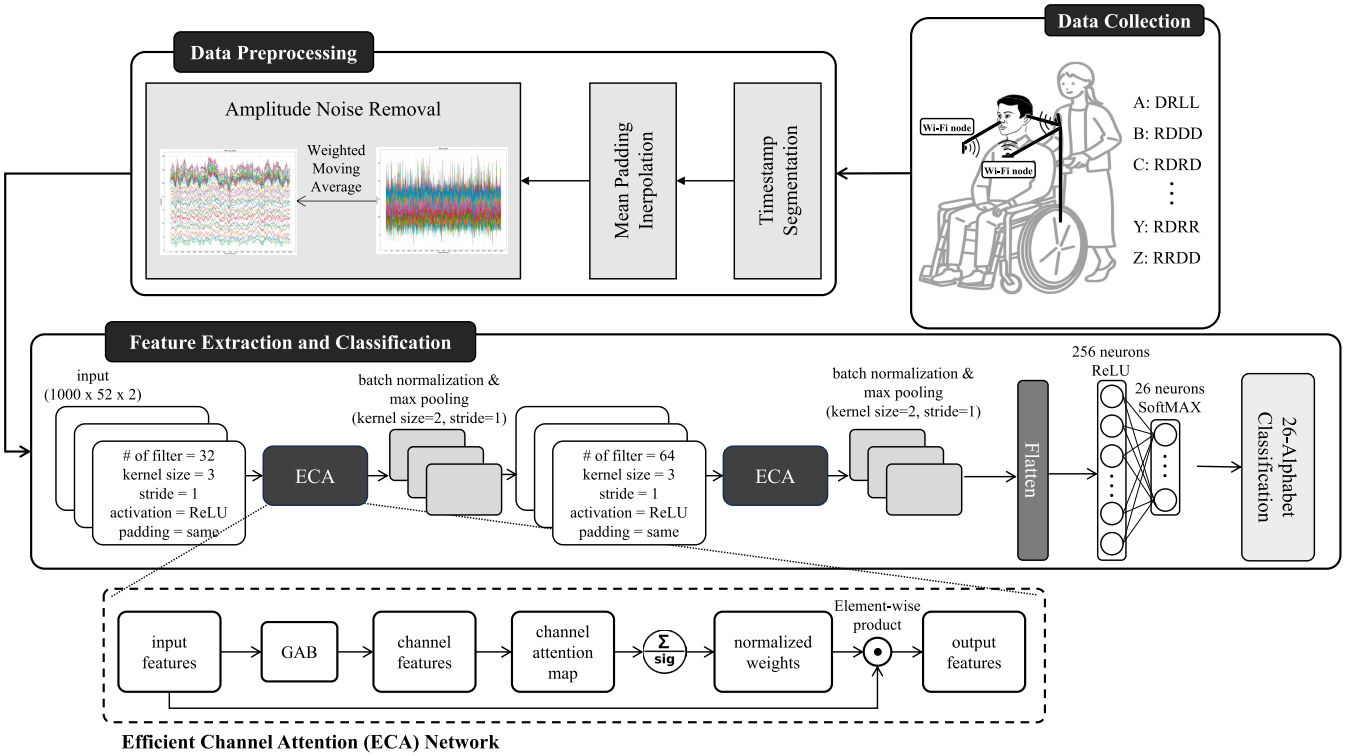

FIGURE 2. Overview architecture of HeMoFi4Q system.

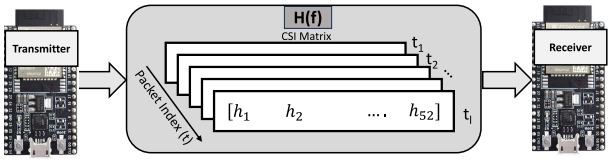

FIGURE 3. CSI vector for each t packet in ESP32 system.

related features, the Fast Fourier Transform (FFT) is applied to get the spectrogram of the filtered CSI. The spectrogram is fed to the CNN-LSTM network for the embedding task. Finally, the MatNet network aims to maximize the cosine similarity between the source environments to extract the representative features. The system achieves over 74% average recognition accuracy.

Francesca et al. proposed SHARP [25] which is a single independent environment HAR based on Wei-Fi CSI. The authors present a new phase sanitization method based on the strongest path to remove the phase offset from the raw CSI signals. After that, they extracted the Doppler trace from the normalized amplitude and filtered phase. The inception model is used to perform the classification task. SHARP data collection is done in three different environments using the Nexmon tool. The average accuracy in the worst scenario is about 95%.

The key features of current Wi-Fi CSI-based systems are presented in Table 1. However, based on our knowledge, there is currently no Wi-Fi CSI system that is capable of handling multi-human context environments with location diversity robustness.

III. PROPOSED SYSTEM

This paper presents a novel head motion system utilizing three modules: data collection, data preprocessing, and feature extraction and classification module as shown in Fig. 2. For the first time, a comprehensive Wi-Fi Channel

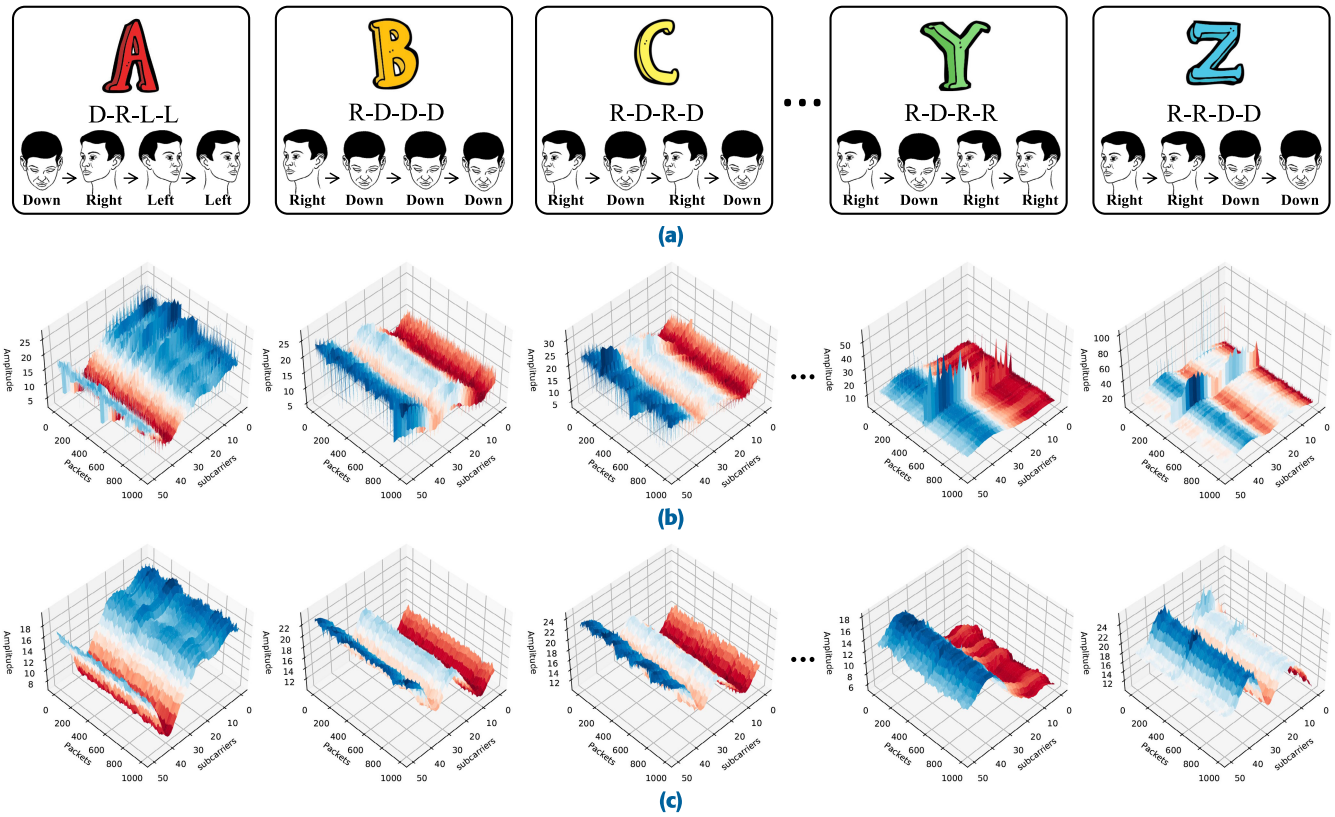

FIGURE 4. First three and last two alphabets motions corresponding to the raw and filtered CSI amplitude. a. Visual representation of A, B, C, . . . , Y, and Z characters. b. Raw CSI amplitude for each character across the first link. c. Filtered CSI amplitude after applying weighted moving average.

State Information (CSI) dataset has been gathered from various locations in two distinct settings using a cost-effective, low-power ESP32 microcontroller. This device is viewed as a promising tool for CSI sensing in the Internet of Things (IoT) realm due to its standalone capabilities. Then, the CSI readings are fed to the data preprocessing phase for signal segmentation and noise removal. After that, the filtered amplitudes are input to ECA to extract representative features for the classification task.

A. ESP32 WI-FI UNIT

The currently employed Wi-Fi CSI-based systems for HAR, head detection, and gesture recognition rely on open-source CSI tools developed by Halperin et al. [26] and Atheros CSI Tool [27]. Although these tools are widely used, they have certain drawbacks, such as limited device support and hardware compatibility, complex setup, limited data processing capabilities, and lack of official support and updates. The former can extract CSI across 30 subcarries from Intel 5300 wireless network card interface (NIC). However, it relies on hardware support that may not be available on all Wi-Fi chipsets. The latter provides a more accurate and details analysis. Furthermore, it supports both the 802.11n and 802.11ac Wi-Fi standards, which allows it to capture and analyze CSI data from the latest Wi-Fi technologies. However, it has some limitations such as hardware compatibility, limited device support, complex setup, limited data processing capabilities, and lack of official support and updates. The challenges of Wi-Fi CSI sensing systems include utilizing a small, low power and memory consumption, cost-efficient, and compatible CSI tool; handling the noisy CSI values in multi-human context environments, and extracting the variations of the most informative base signal related to the head motions. ESP32 unit is a microcontroller a single system on a chip, which means that it integrates multiple components, such as the processing unit, memory, and communication interfaces, onto a single chip. It is a low power consumption and cost-effectiveness.

The proposed system works in three different pairs of ESP32 microcontrollers attached to a wheelchair surrounding the target's head. ESP32 module utilizes orthogonal frequency division modulation (OFDM) to transmit wireless signals, which consists of 64 narrow band subcarriers, 12 of them are null subcarriers and the rest 52 are data subcarriers.

The CSI data collected by the ESP32 node is represented as a complex matrix with dimensions of $l \times m$, where l represents the number of packets and m represents the number of subcarriers as shown in Fig. 3.

$$Y_i(f) = H_i(f) \times X_i(f) + N \tag{1}$$

where i represents the subcarrier index, $H_i(f) \in \mathbb{C}^{N_{Rx} \times N_{Tx}}$, $Y_i(f) \in \mathbb{R}^{N_{Rx}}$, and $X_i(f) \in \mathbb{R}^{N_{Tx}}$ denote the CSI complex matrix, received and transmitted signals across i^{th}subcarrier, respectively where N_{Rx} and N_{Tx} are the number of receiving and transmitting antennas. Since the ESP32 device used in this context has a single antenna, both N_{Rx} and N_{Tx} is equal to 1, and the form of the CSI matrix for each packet is $H(f) = [h_1\ h_2\ \ldots\ h_{52}]$. **N** is a noise vector.

Typically, the **H** matrix records the the amplitude attenuation, $\|H_i(f)\|$, and phase response, $\angle H_i(f)$, of each subcarrier frequency.

$$H_i(f) = \|H_i(f)\| \times e^{j\angle H_i(f)} \tag{2}$$

B. DATA PREPROCESSING PHASE

The process was conducted in three stages, which involved 1) dividing the data into segments based on timestamps, 2) adding mean values to mitigate the effects of packet loss in Wi-Fi CSI measurements, and 3) eliminating any noise in the signal amplitude. More specific information regarding the segmentation, padding, and filtering procedures can be found in the following descriptions.

1) TIMESTAMP SEGMENTATION

The goal of signal segmentation is to divide the CSI measurements of each link based on their timestamps. This process is necessary to combine the signals from each link in order to create a distinct pattern for each user's head movement. By doing so, the character can be mapped to its corresponding signature. In other words, signal segmentation enables the system to identify and distinguish between different head movements made by each user, which are then translated into their corresponding characters. This process is essential for accurate and reliable communication through the system.

2) MEAN PADDING

Mean padding is a widely adopted approach to address the issue of packet loss in Wi-Fi CSI measurements while maintaining the distribution of remaining packets. Rather than simply discarding the missing packets, mean padding involves estimating their values based on the surrounding data points in the time series. By using the average value of the neighboring data points to fill in the gaps, mean padding can help to preserve the distribution of the packets and maintain the continuity of the overall time series.

3) AMPLITUDE NOISE REMOVAL

This section outlines the noise removal technique used to smooth the CSI amplitude. It is worth mentioning that we investigated three different features to evaluate the system's performance. These features are including the filtered amplitude, time-frequency feature, and the combination of the filtered amplitude and the calibrated phase. However, the variations of the filtered time-domain CSI amplitudes outperform other features in the recognition accuracy. they reveal distinct signatures for different alphabets.

Since it is not reliable to use raw amplitude directly as a feature selection due to the signal interference and environmental changes noise. Therefore, we adopt a weighted moving average filter (WMA) [28] to remove the outliers and smooth the CSI amplitude waveforms as shown in Eq.3

$$\hat{A}_{t,i} = \frac{m \times A_{t,i} + (m-1) \times A_{t\text{-}1,i} + \ldots + 1 \times A_{t\text{-}m\text{-}1,i}}{m + (m-1) + \ldots + 1} \tag{3}$$

where $\hat{A}_{t,i}$ and $\mathbf{A}_{t,i}$ are the filtered and raw amplitude corresponding to the subcarrier i at time t, and m that is empirically set to 30 in this paper. The illustration in Fig. 4a showcases the visual head motions of A, B, C, Y, and Z alphabets, while corresponding raw and filtered CSI amplitudes are shown in Fig. 4b and Fig. 4c, respectively. Further details of the experiment setup are described in the following sections.

C. FEATURE EXTRACTION AND CLASSIFIER PHASE

Nowadays, deep convolutional neural networks (DCNN) outperform traditional machine learning techniques (ML) in Natural Language Processing (NLP), Computer Vision (CV), sensing techniques, etc. The reason is that the performance of a ML-based system depends on the quality of extracted features. The ML depends on hand-crafted features extracted from the data which can be inefficient for the classification task. On the other hand, the DL is a black box that is able to automatically capture more discriminative features from the input. The main issue with the current DCNN is that more layers added to it will improve performance, but will make the model more complex. Additionally, the current attention technique non-identically weights the extracted feature depending on how crucial it is to the classification task. Therefore, information is lost as a result of the unnecessary and ineffective dependencies across several channels.

To get over these restrictions, the ECA module [29] is presented to aggregate the local inter-channel information. During the learning process, ECA assigns the weights of each channel based on its significance in the channel attention map (CAM). Then, to emphasize the distinct patterns and suppress the uncorrelated data, the input features of each channel are multiplied by the corresponding CAM weights. The following is a description of how Fig. 2 depicts the ECA network's process. First of all, the global average pooling (GPA) layer aggregates the input features f_a. Then, the fast one-dimensional convolution layer $1d$ with k kernel size f_{1d}^{k} is applied to each input channel to generate the weights of CAM f_b. The k establishes how many neighbors take part in the attention prediction via each channel. In the proposed system, k value is equal to 3. After that, the sigmoid function normalizes the attention weights $f_b^{'}$. Finally, the significant input features are generated by multiplying the input features with the channel weights $f_b^{'}$. In ECA network, the adaptive

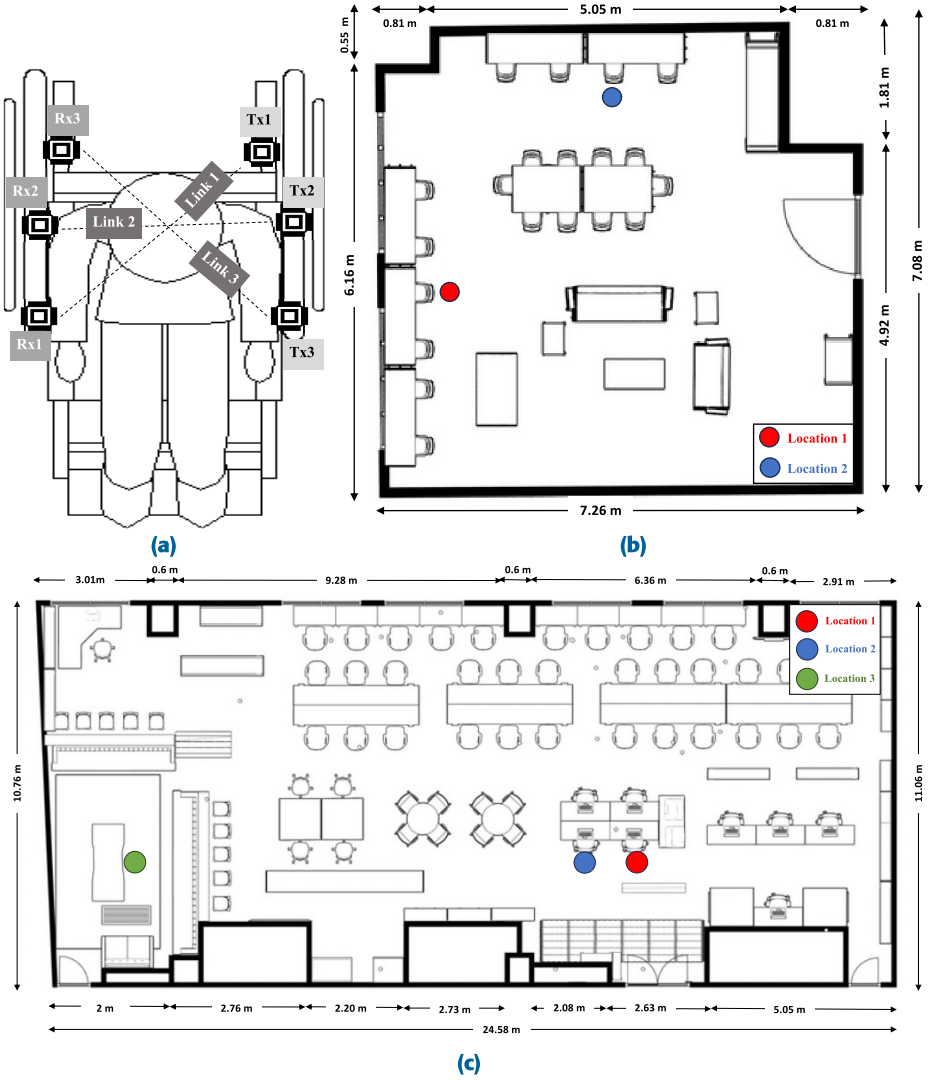

FIGURE 5. Data collection setup based on ESP32 microcontroller. a. Top view of wheelchair setup used in both environments. b. Single-user environment layout. c. Multi_human context environment layouts.

TABLE 2. HeMoFi4Q Code: D-down motion, R-right motion, L-left motion.

Char	Code	Char	Code	Char	Code	Char	Code	Char	Code
A	D-R-L-L	G	R-R-D-L	M	R-R-L-L	S	D-D-D-L	Y	R-D-R-R
B	R-D-D-D	H	D-D-D-D	N	R-D-L-L	T	R-L-L-L	Z	R-R-D-D
C	R-D-R-D	I	D-D-L-L	O	R-R-R-L	U	D-D-R-L		
D	R-D-D-L	J	D-R-R-R	P	D-R-R-D	V	D-D-D-R		
E	D-L-L-L	K	R-D-R-L	Q	R-R-D-R	W	D-R-R-L		
F	R-R-D-L	L	D-R-D-D	R	D-R-D-L	X	R-D-D-R		

selection of the kernel size is an exponential function that depends on the number of channels C.

This model is a convolutional neural network (CNN) with an ECA module. The input to the model is a tensor of shape (1000,52,2). The model has two convolutional layers, each followed by an ECA module, batch normalization, and max-pooling layers. The output of the second max-pooling layer is flattened and fed into a dense layer with 256 units,

TABLE 3. Training and testing size for HeMoFi4Q performance evaluation.

	Env1		Env2		Env1 → Env2		Env2 → Env1	
Training samples	Loc1 + 2% Loc2	2652	Loc1 + Loc2 + 2% Loc3	5252	All_locs_Env1 + 2% Loc3_Env2	5356	All_locs_Env2 +2% Env1	7904
Testing samples	98% Loc2	2548	98% Loc3	2548	98% Loc3_Env2	2548	98% Env1	5096

followed by a dropout layer with a dropout rate of 0.25. The final output layer contains 26 units representing the number of characters with softmax activation.

The ECA module is a feature recalibration mechanism that enhances the performance of CNNs by recalibrating the feature maps. It applies a convolutional operation with a small kernel size to squeeze the feature maps into a single channel and then applies a sigmoid activation function to obtain attention maps. The attention maps are multiplied with the original feature maps to generate the scaled feature maps, which are then passed to the next layer.

IV. EVALUATION

A. DATA ACQUISITION

In our experiments, two different environments are considered where ESP32 modules are used as both transmitter, Tx, and receiver, Rx, devices. This work utilized six ESP32 microcontrollers using 2.4 GHz frequency band and IEEE 802.11n protocol for the CSI data collection task as shown in Fig. 5. Three of them were used as transmitters connected to a mini-PC having Ubuntu 16.04 operating system and the remaining served as receivers. The distance between each transmitter and its corresponding receivers is 84.3 cm, 52.3 cm and 84.3 cm for link 1, link 2, and link 3, respectively.

For both environments, *Env1* and *Env2*, the subject moved his head in a way outlined in Table 2, with each symbol and the end motion taking two seconds for a total of 10 seconds per character. We gathered data for each character in separate files, with a duration of 10 minutes for each character. Each of the three Txs concurrently transmit Wi-Fi frames to the Rxs at 100 Hz. *Env1* represented a single environment with only the subject existed and the data collected in two different locations in it as shown in Fig. 5b. Whereas *Env2* was a multi-human context environment as shown in Fig. 5c. The dimension of each character is represented as $1000 \times 52 \times n$, where 1000 refers to the number of packets, 52 is the number of subcarriers, and n depends on the number of links used. We investigate the impact of various link configurations, which are discussed later in IV-C. For a single link configuration, n is set to 1. When utilizing two links, n is equal to 2, and when utilizing all available links, n is set to 3. The collected dataset is balanced, meaning that an equal number of samples were collected from all locations. We had a total of 5 locations, and for each location, we gathered 100 instances of each character. This results in a total of 2600 alphabet samples for each location. Altogether, across all locations, we accumulated 13,000 samples in total. For our experiments, we focused on 26 alphabet characters as the number of classes. This choice is based on the functionality of HeMoFi4Q, which extracts the signature for Morse code using head motion and maps it to the corresponding character.

B. EVALUATION METRICS OF CLASSIFICATION MODELS

We evaluate the HeMoFi4Q effectiveness on different environments based on CSI data. Firstly, the HeMoFi4Q is implemented to be applicable in the real world. Inspired by few shot learning algorithms, we fuse a small amount (x) from the unseen/test environment to the seen/train dataset to tackle the location diversity robustness problem as shown in Fig. 1. The performance of the proposed system is evaluated through accuracy and F1-score metrics. Accuracy is defined as the total number of correct predictions divided by total number of predictions. The F1-score represents the harmonic mean of two measures (precision and recall). It is a range of numerical values from 0 to 1, where the worst and best values are 0 and 1, respectively. The performance metrics are calculated by following equations from Eq.4 and Eq.7:

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

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

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

where True Positive (TP) refers to the model accurate identification for the positive class, whereas True Negative (TN) refers to the predicted and actual values is negative. False Negative (FN) represents the cases when the actual is positive and the model classified them as positive while False Positive (FP) represents the cases where the actual is negative and the predicted is positive.

C. RESULT ANALYSIS

The major objective of this paper is to introduce passive communication between the quadriplegics and others based on head motions detected by Wi-Fi signals and DL algorithms.

To evaluate the robustness of location diversity, the source dataset is combined with 2% of the target dataset for training the model, and the rest 98% is then used for the testing phase. Table 3 provides information about the sample sizes used in

TABLE 4. Location diversity comparative results of different base signals at Env1.

Link Conf	ECA			CNN			ResNet		
	Amp	DWT_Amp	Amp+Phase	Amp	DWT_Amp	Amp+Phase	Amp	DWT_Amp	Amp+Phase
Link1	79.3	77.7	37.3	84.5	73.7	33.4	81	74.8	34.5
Link2	86.8	81	62	88	81.7	69	86.7	80	65
Link3	82.7	77.2	61.4	88.4	72.7	60.2	76.2	74.7	61
Link1_2	92.4	93.8	78.5	93.3	90.2	72.3	88.3	**91.7**	73.4
Link1_3	**94.6**	90.2	79.3	92.8	88.9	74.7	88.4	87.2	72
Link2_3	93.3	94	79.7	**93.7**	92.5	75.6	87.2	90.6	74 .3

TABLE 5. Location diversity comparative results of different base signals at Env2.

Link Conf	ECA			CNN			ResNet		
	Amp	DWT_Amp	Amp+Phase	Amp	DWT_Amp	Amp+Phase	Amp	DWT_Amp	Amp+Phase
Link1	65.9	76	48.3	76.8	73.3	44	57.5	70.6	42.6
Link2	74.6	83.4	55.6	84	70.6	52	55.8	68.6	48.7
Link3	74.7	88.9	57.6	82.8	**87**	51.7	58.8	83.8	40.5
Link1_2	86.4	87.1	69.3	84.6	85.4	59.7	57.8	73.1	51.4
Link1_3	**89.3**	88.3	71	83.3	83	52	60.7	76.3	45.8
Link2_3	87.4	85.5	70.2	84.4	83.8	60.3	57.5	**85**	57.8

different experimental settings. In particular, we compared the performance of ECA learning model with two state-of-art classifiers namely, CNN and ResNet on different locations of the same environment as in Table 4 and Table 5 and cross domain environments which is the combination of one environment's locations and 2% amount of the second environment locations as a training dataset and testing on the rest 98% amount of the second environment as shown in Table 6. The parameters for the two baseline models are as follows.

- CNN model: This is a simple convolutional neural network (CNN) model for image classification. The model consists of two CNN blocks, each composed of several sequential layers. The input shape is (1000, 52, 2), representing a 2D image with 1000 packets, 52 subcarriers, and 2 links used. The first block starts with a convolution layer with 32 filters of size P5 $\times$ 5 and a stride of 1, applying zero-padding with a padding value of 1. This is followed by a batch normalization layer for input normalization, a ReLU layer for introducing non-linearity, and an average pooling layer with a window size of 3 $\times$ 3 and a stride of 3. A dropout layer is employed to prevent overfitting by randomly dropping out units during training. The second block consists of two fully connected layers. The first fully connected layer has 1000 neurons with a ReLU activation function and a dropout rate of 0.5. The second fully connected layer has 26 neurons, corresponding to the number of classes, and utilizes the softmax function for classification. The model is trained using the Stochastic Gradient Descent with Momentum (SGDM) optimization algorithm, with a learning rate of 0.02 and momentum of 0.9.
- ResNet model: This is a deep neural network architecture commonly used for image classification tasks. The input shape of the model is (1000, 52, 2), where 1000 refers to the number of packets, 52 represents the subcarriers, and 2 denotes the links used. To maintain the spatial dimensions, the input is initially padded with zeros using a (3, 3) padding size. The model consists of two stages. In the first stage, a convolutional layer with 64 filters, a kernel size of (7, 7), and a stride of (2, 2) is applied to extract features. This is followed by a batch normalization layer for activation normalization, a ReLU activation layer to introduce nonlinearity, and a max pooling layer with a pool size of (3, 3) and a stride of (2, 2) for downsampling. The second stage includes a convolutional block with three identity blocks. Each identity block comprises three convolutional layers with filter sizes of [16, 32, 64], a kernel size of 3, and a stride of 1. The first identity block has a different shape due to the change in filter sizes. An average pooling layer with a pool size of (2, 2) is applied to further downsampling the data. The output is then flattened into a 1D vector and fed into a fully connected layer with 26 neurons, which corresponds to the number of classes in the

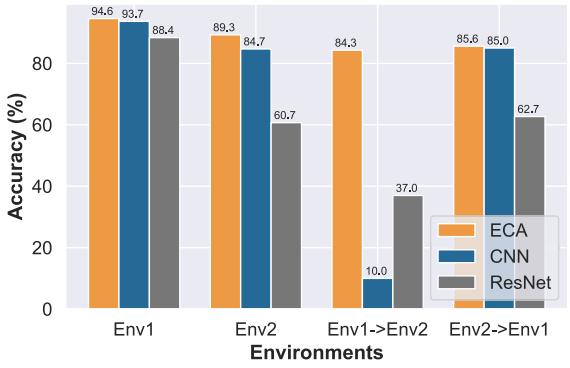

FIGURE 6. Overall accuracy of different learning models.

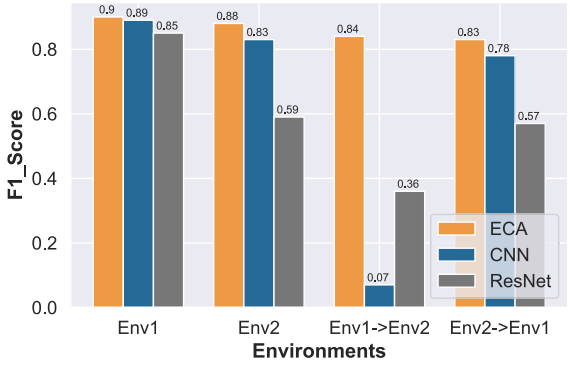

FIGURE 7. F1-score of different learning models.

TABLE 6. Cross domain results of different classifiers.

	Env1 →Env2			Env2→Env1		
	ECA	CNN	Resnet	ECA	CNN	Resnet
Accuracy	**84.3**	10	37	**85.6**	85	62.7
F1-score	0.84	0.07	0.36	0.83	0.78	0.57
Training_time(sec)	19	14	13	13	10	7
Testing_time (sec)	4	3	2	2	1	1

classification task. The softmax activation function is used to generate predicted probabilities for classification. The model is trained using the Adam optimizer with a learning rate of 0.001, beta_1 of 0.9, beta_2 of 0.999, and epsilon of 1e-08.

ECA classifier outperforms other algorithms as shown in Fig. 6 and Fig. 7 by achieving highest recognition accuracy and f1-score of location diversity evaluation for *Env1* and *Env2*, training on *Env1* with 2% of *Env2* and testing on 98% of *Env2* (*S1*), and training on *Env2* with 2% of *Env1* and testing on 98% of *Env1* (*S2*), respectively. CNN algorithm achieves the lowest performance when using the single-user environment for the learning phase because of overfitting. CNN and ResNet cannot capture the unique patterns for each character compared to ECA which uses an attention layer to highlight the signatures of the characters from the

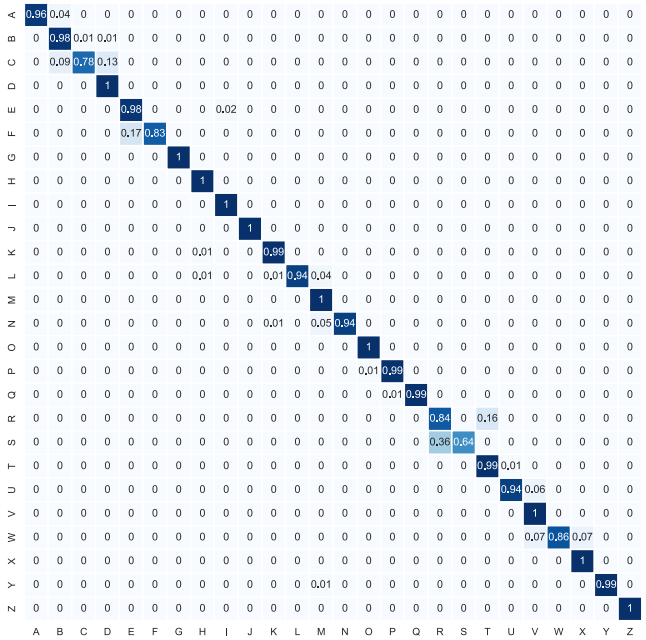

FIGURE 8. Confusion matrix of single user environment.

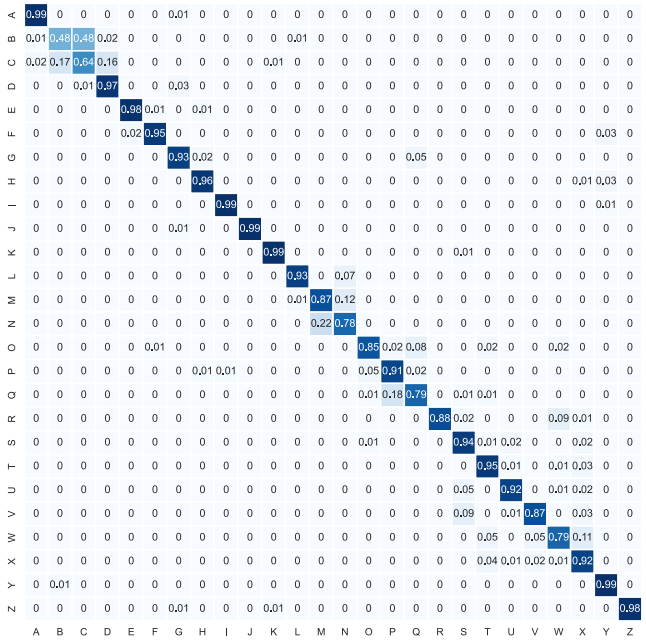

FIGURE 9. Confusion matrix of multi-human context environment.

single environment and the small amount of the multi-human sensing environment.

1) SINGLE USER ENVIRONMENT DATA

The system was trained in a single-user environment where only the participant existed, *Env1*, using the entire dataset collected from location 1 and an additional 2% of data from location 2, which equates to two samples for each character collected in the second location. This approach was taken to improve the system's performance and location

diversity robustness. Table 4 displays the evaluation outcomes of three distinct image classification algorithms, namely Efficient Channel Attention (ECA), Convolutional Neural Network (CNN), and Residual Network (ResNet). The results were obtained by assessing the accuracy of each classifier, and they are presented for comparison purposes.

It is observed from Table 4 that ECA slightly outperforms CNN on combined datasets of the diagonal links configuration (link1_3) for different base signals, while the ResNet classifier achieves the worst performance. Moreover, ECA achieves the highest recognition accuracy with 94% when using the variations of amplitude filtered by WMA filter from diagonal links datasets and transforming the filtered amplitude of the combined dataset between one of the diagonal links and the horizontal link to the wavelet domain by applying the discrete wavelet transform (DWT). The confusion matrix is introduced in Fig. 8. The confusion matrix shows that there are 13% and 9% misclassifications between *C* and *D* and *B*, respectively. Additionally, the model classified *R* character as *T* character with 16% rate and *S* as *R* character with 36% rate which is the highest misclassification rate. Finally, ECA classified *W* character with the same rate, 7%, as *V* and *X*.

2) MULTI-HUMAN CONTEXT ENVIRONMENT DATA

The classification accuracy of data collected from various locations in a multi-human context environment is presented in Table 5. Our study also investigated the optimal link configurations, different base signals, and three classifiers. The results demonstrated that the dataset of diagonal links outperformed other link configurations, achieving 89.3% accuracy based on the variations of filtered amplitude and using the ECA classifier. Furthermore, the ResNet classifier achieved its highest accuracy of 85% by utilizing the wavelet of the filtered amplitudes from one of the diagonal link data with the horizontal one. On the other hand, the combination of the variation of the filtered amplitude and calibrated phase was verified as the worst base signal that could be fed to the system because the randomness of the phase increased with the number of people in the sensing environment, leading to a degradation of the system's performance. Fig. 9 presents the confusion matrix. As it can be observed from this visual representation, the model classified the *B* character with about 48% accuracy as *C* character. Additionally, it misclassified the *C* character as 17% and 16% as *B* and *D* characters, respectively. Moreover, there is a misclassification rate between *M* and *N* with 12% and 22%, respectively.

3) CROSS DOMAIN RESULTS

This study examined the cross-domain or environment diversity robustness meaning that the model training on some locations of one environment merged with randomly selected 2% from another environment to investigate the effectiveness and robustness of the proposed system. There are two scenarios which considered as the worst scenarios. The first scenario (S1) is training on the single-user dataset with small samples from the multi-human one and testing on the multi-human environment (Env1→Env2). The second one (S2) uses a multi-user environment dataset for learning with small samples from the single-user environment and for the interference stage using the unseen samples of the single-user environment (Env2→Env1).

The classification results of different classifiers and confusion matrix are given in Table 6 and Fig. 10. Table 6 shows the overall performance of different classifiers. ResNet gives the lowest accuracy in both scenarios while ECA model gives the best performance among the other classifiers in terms of accuracy and F1-score metrics. However, ECA takes a little bit more time consumption in the learning and inference stages than others.

To evaluate the performance of *S1*, the training dataset consisted of the single-user environment (*Env1*) dataset merged with 2% of the third location dataset from the multi-human context environment (*Env2*), while the testing dataset comprised the remaining 98% of the data. The ECA algorithm outperformed the other algorithms, achieving 84.3% accuracy and 0.84 F1-Score with slightly more time consumption than CNN. On the other hand, ResNet achieved the shortest time consumption compared to CNN and ECA models but yielded an unacceptable classification accuracy as 37%. Interestingly, CNN achieved poor accuracy due to overfitting, where the model could not distinguish the signatures for each alphabet of the multi-human environment from the small target samples used in the learning phase. the model achieves 70% and higher for each character except the *C and T* characters, it achieves 53% and 55%, respectively. ECA wrongly classified the *C* character as *B* and *F* characters with accuracy of 18% and 24%, respectively. For *T* character, it misclassified it as *S* character with 18% accuracy.

The classification results for *S2* give better accuracy than S1 results because the classifiers are able to extract the most significant features for each character. ECA achieves the best accuracy in 85.6% which is slightly higher than CNN model performance and 0.83 F1-score. From the confusion matrix in Fig. 10, ECA achieves 72% and above for most of the characters. In particular, *W* character gives the worst accuracy 43% since the model wrongly classified it as *X* and *V* with 32% and 14% accuracy, respectively. Moreover, *B* character accuracy is slightly higher than *W* one 54.3% since there is a wrong classification as *C* and *A* 29% and 14%, respectively. The model classified *C* character as *B* with 33% accuracy. Furthermore, There is also misclassification between *S* as *U* and *R* with accuracy rates of 15% and 14%, respectively.

V. DISCUSSION

In this study, a Morse code based on Wi-Fi CSI head motion detection is presented using ESP32 microcontroller. we investigated the impact of different link configurations, base signals, and state-of-art deep learning classifiers on the location diversity performance as shown in Table 4 and Table 5. It is worth mentioning that these results by using

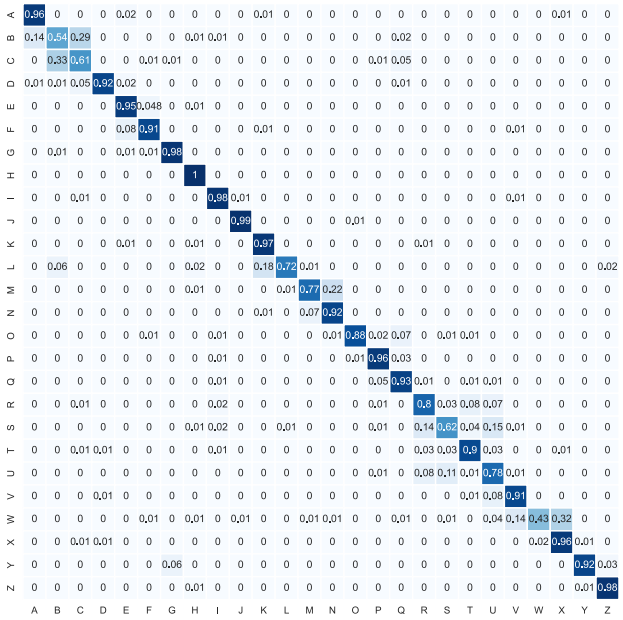

FIGURE 10. Confusion matrix of Env2 → Env1.

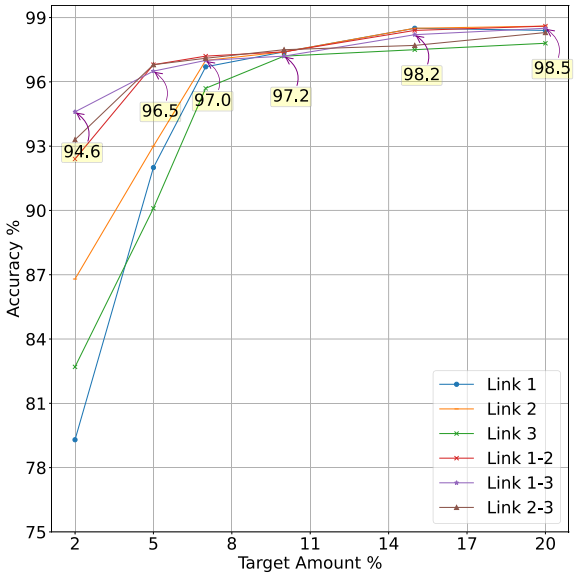

FIGURE 11. Classification accuracy using different amounts from the unseen location merged with the seen location in a single-user environment.

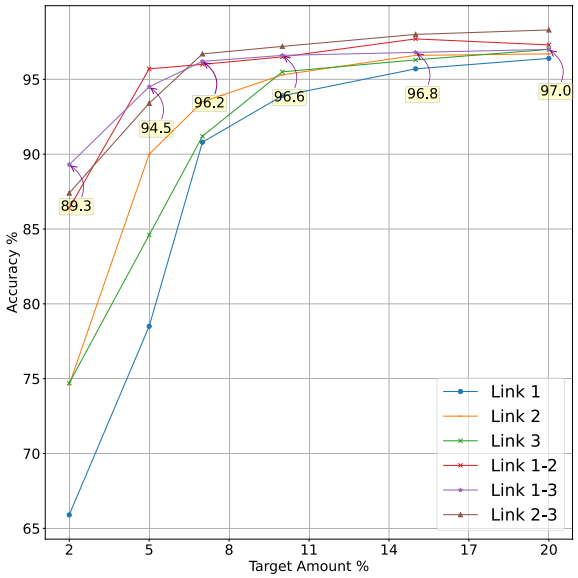

FIGURE 12. Classification accuracy for different amounts from the unseen location merged with the seen location in a multi-user environment.

2% amount From the target or unseen location. from these tables, it is obvious that the combination of the diagonal links' data outperforms other link configurations by extracting the variation of the CSI amplitude and filtering it using the weighted moving average algorithm outperforms other link configurations by achieving 94% and 89% recognition accuracy, for single and multi-human context environment, respectively.

A. IMPACT OF DIFFERENT TARGET AMOUNT

Our study aimed to explore the effects of altering the target amount incorporated into the training data on the enhancement of model accuracy during instances of limited data availability. Specifically, we conducted an experiment involving the testing of different target amount values within the training data and the subsequent evaluation of their impact on model performance in location diversity. Our findings contributed novel insights into the potential benefits of adjusting the target amount in the training data to improve the environmental robustness of the Wi-Fi CSI system.

Fig. 11 depicts the accuracy of the ECA algorithm, under various link configurations, for different amounts from the unseen location (Loc2) merged with the training dataset, which is the data gathered from Loc1, (Target Amount%) during the learning phase. According to the information presented in the figure, it is evident that the utilization of the diagonal links dataset results in the highest performance for the ECA algorithm. This configuration yielded the best results even when only two randomly selected samples were taken for each character collected in the target location. As shown in Fig. 12, the performance of each link alone yielded the worst recognition accuracies, ranging from 66% to 74.7% for the diagonal and horizontal links, respectively, due to the influence of the target's surrounded movements by individuals. However, when the diagonal links were combined, the system achieved the highest performance of 89.3%. This can be attributed to the ECA classifier, which utilizes a channel attention layer to highlight the most significant features from the filtered amplitude captured by these links.

B. IMPACT OF DIFFERENT LINK CONFIGURATION

Table 4 and Table 5 reveal a significant finding that the diagonal link configurations consistently achieve the highest

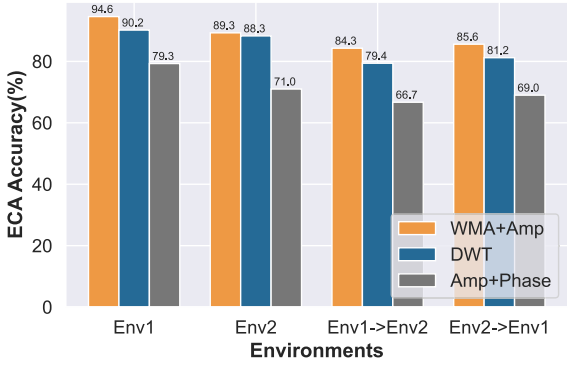

FIGURE 13. Accuracy of different base signals.

accuracy for both single-user and multi-human context environments. Moreover, the variations of the filtered amplitude were found to be a robust and reliable signature for each character, regardless of the location diversity. In addition, the channel attention layer of the ECA algorithm was identified as a crucial component that enhances the classification task. It is noteworthy that the overall accuracy of the system in a multi-human context environment is lower compared to a single-user environment. This is due to the fact that the presence of more individuals leads to increased interference and reflection, which ultimately degrades the system's performance. Overall, the study highlights the importance of an effective link configuration, signal processing technique, and utilizing advanced classification algorithms in achieving accurate and reliable character recognition based on passively tracking head motion via Wi-Fi CSI signals in various environments.

C. IMPACT OF DIFFERENT BASE SIGNALS

We investigated the impact of the variations of different base signals on the classification performance as shown in Fig. 13. The impact of different base signals like, filtered amplitude using WMA filter, transforming the filtered amplitude to the wavelet domain by applying discrete wavelet transform technique (DWT_Amp), and combining both the filtered amplitude based on WMA and calibrated phase based on the linear transformation algorithm (Amp+Phase) on the performance are studied. As the results show in the previous tables, the combination of the CSI amplitude and phase variations degrades the performance as it achieves 79.7% and 70.2% for the first and second environments, respectively. The impact of the randomness of the Wi-Fi CSI is obviously shown in the multi-human environment because there are a lot of reflections and interference due to the existence of many dynamic subjects in the sensing environment.

VI. CONCLUSION

In this paper, we introduce HeMoFi4Q which is a passive head motion detection system by analyzing the Wi-Fi Channel State Information to enable non-verbal communication with quadriplegia patients through new sign language. To the best of our knowledge, it is the first attempt to build a new sign language for quadriplegia patients based on Morse code and head motions. The proposed HeMoFi4Q extracts the CSI amplitude variations from ESP32 microcontrollers and feeds the filtered amplitudes using the weighted moving average to the Efficient Channel Attention (classifier). The dataset was collected from two different environments with multiple locations in each environment. Inspired by few shot learning techniques, the small sample amount, randomly 2% samples, from the unseen environment were merged with the training dataset in the model learning interference to verify the system effectiveness and location diversity robustness. A variety of performance metrics of HeMoFi4Q including accuracy, F1-score, and confusion matrix outperforms all baseline models. the study underscores the crucial role of an effective link configuration, signal processing technique, and advanced classification algorithm in achieving accurate and reliable character recognition through passive tracking of head motion via Wi-Fi CSI signals across diverse environments. In the future, we are planning to build a word recognition system based on Wi-Fi CSI using NLP algorithms for word correction to make the system applicable in real world.

REFERENCES

[1] M. Karsy and G. Hawryluk, "Modern medical management of spinal cord injury," *Current Neurol. Neurosci. Rep.*, vol. 19, no. 9, pp. 1–7, Sep. 2019.
[2] M. G. Morshed, T. Sultana, A. Alam, and Y.-K. Lee, "Human action recognition: A taxonomy-based survey, updates, and opportunities," *Sensors*, vol. 23, no. 4, p. 2182, Feb. 2023.
[3] X. Chen, L. Su, J. Zhao, K. Qiu, N. Jiang, and G. Zhai, "Sign language gesture recognition and classification based on event camera with spiking neural networks," *Electronics*, vol. 12, no. 4, p. 786, Feb. 2023.
[4] S. Suh, V. F. Rey, and P. Lukowicz, "TASKED: Transformer-based adversarial learning for human activity recognition using wearable sensors via self-knowledge distillation," *Knowl.-Based Syst.*, vol. 260, Jan. 2023, Art. no. 110143.
[5] X. Wu, X. Luo, Z. Song, Y. Bai, B. Zhang, and G. Zhang, "Ultra-robust and sensitive flexible strain sensor for real-time and wearable sign language translation," *Adv. Funct. Mater.*, vol. 33, no. 36, Sep. 2023, Art. no. 2303504.
[6] X. Li, Y. He, and X. Jing, "A survey of deep learning-based human activity recognition in radar," *Remote Sens.*, vol. 11, no. 9, p. 1068, May 2019.
[7] F. Costa, S. Genovesi, M. Borgese, A. Michel, F. A. Dicandia, and G. Manara, "A review of RFID sensors, the new frontier of Internet of Things," *Sensors*, vol. 21, no. 9, p. 3138, Apr. 2021.
[8] R. Shahbazian and I. Trubitsyna, "Human sensing by using radio frequency signals: A survey on occupancy and activity detection," *IEEE Access*, vol. 11, pp. 40878–40904, 2023.
[9] A.-I. Şiean, C. Pamparău, A. Sluÿters, R.-D. Vatavu, and J. Vanderdonckt, "Flexible gesture input with radars: Systematic literature review and taxonomy of radar sensing integration in ambient intelligence environments," *J. Ambient Intell. Humanized Comput.*, vol. 14, pp. 1–15, Apr. 2023.
[10] L. Lasantha, N. C. Karmakar, and B. Ray, "Chipless RFID sensors for IoT sensing and potential applications in underground mining—A review," *IEEE Sensors J.*, vol. 23, no. 9, pp. 9033–9048, May 2023.
[11] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 1472–1480.
[12] C. Chen, G. Zhou, and Y. Lin, "Cross-domain WiFi sensing with channel state information: A survey," *ACM Comput. Surv.*, vol. 55, no. 11, pp. 1–37, Nov. 2023.
[13] M. Mercuri, P. J. Soh, P. Mehrjouseresht, F. Crupi, and D. Schreurs, "Biomedical radar system for real-time contactless fall detection and indoor localization," *IEEE J. Electromagn., RF Microw. Med. Biol.*, early access, May 26, 2023, doi: 10.1109/JERM.2023.3278473.

[14] E. L. Gomes, M. Fonseca, A. E. Lazzaretti, A. Munaretto, and C. Guerber, "Clustering and hierarchical classification for high-precision RFID indoor location systems," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5141–5149, Mar. 2022.
[15] B. Zhang, H. Sifaou, and G. Y. Li, "CSI-fingerprinting indoor localization via attention-augmented residual convolutional neural network," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5583–5597, Aug. 2023.
[16] C. Jobanputra, J. Bavishi, and N. Doshi, "Human activity recognition: A survey," *Proc. Comput. Sci.*, vol. 155, pp. 698–703, Jan. 2019.
[17] D. Avola, M. Cascio, L. Cinque, A. Fagioli, and C. Petrioli, "Person re-identification through Wi-Fi extracted radio biometric signatures," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 1145–1158, 2022.
[18] H. Hameed, M. Usman, A. Tahir, A. Hussain, H. Abbas, T. J. Cui, M. A. Imran, and Q. H. Abbasi, "Pushing the limits of remote RF sensing by reading lips under the face mask," *Nature Commun.*, vol. 13, no. 1, p. 5168, Sep. 2022.
[19] M. R. M. Bastwesy, K. Kai, H. Choi, S. Ishida, and Y. Arakawa, "Wi-Nod: Head nodding recognition by Wi-Fi CSI toward communicative support for quadriplegics," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2023, pp. 1–6.
[20] Z. U. A. Akhtar, H. F. Rasool, M. Asif, W. U. Khan, Z. U. A. Jaffri, and M. S. Ali, "Driver's face pose estimation using fine-grained Wi-Fi signals for next-generation Internet of Vehicles," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–18, May 2022.
[21] Y. Liu and S. Konomi, "WiHead: WiFi-based head-pose estimation," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2022, pp. 69–86.
[22] V. K. Singh, P. Kar, A. M. Sohini, M. Rangaiah, S. Chakraborty, and M. Maity, "Monitoring engagement in online classes through WiFi CSI," in *Proc. 15th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2023, pp. 462–465.
[23] N. Bahadori, J. Ashdown, and F. Restuccia, "ReWiS: Reliable Wi-Fi sensing through few-shot multi-antenna multi-receiver CSI learning," in *Proc. IEEE 23rd Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2022, pp. 50–59.
[24] Z. Shi, Q. Cheng, J. A. Zhang, and R. Yi Da Xu, "Environment-robust WiFi-based human activity recognition using enhanced CSI and deep learning," *IEEE Internet Things J.*, vol. 9, no. 24, pp. 24643–24654, Dec. 2022.
[25] F. Meneghello, D. Garlisi, N. D. Fabbro, I. Tinnirello, and M. Rossi, "SHARP: Environment and person independent activity recognition with commodity IEEE 802.11 access points," *IEEE Trans. Mobile Comput.*, vol. 22, no. 10, pp. 6160–6175, Oct. 2023.
[26] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 4, pp. 159–170, Aug. 2010.
[27] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity WiFi," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, Sep. 2015, pp. 53–64.
[28] H. Li, X. He, X. Chen, Y. Fang, and Q. Fang, "Wi-motion: A robust human activity recognition using WiFi signals," *IEEE Access*, vol. 7, pp. 153287–153299, 2019.
[29] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11531–11539.

MARWA R. M. BASTWESY (Graduate Student Member, IEEE) received the B.E. and M.E. degrees from Tanta University, Egypt, in 2016 and 2021, respectively. She is currently pursuing the Ph.D. degree with the Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan. Her research interest includes passive sensing techniques.

HYUCKJIN CHOI (Member, IEEE) received the B.E. degree from Tongmyong University, in 2016, the M.E. degree from Pusan National University, Busan, Republic of Korea, in 2018, and the Ph.D. degree from the Nara Institute of Science and Technology (NAIST), Nara, Japan, in 2022. Since October 2022, he has been an Assistant Professor with the Graduate School of Information Science and Electrical Engineering, Kyushu University. He is majorly studying the applications of wireless sensing, such as crowd estimation, localization, and human activity recognition techniques.

YUTAKA ARAKAWA (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from Keio University, Japan, in 2001, 2003, and 2006, respectively. He is currently a Professor with the Graduate School and the Faculty of Information Science and Electrical Engineering, Kyushu University. He is also a Visiting Professor with the Nara Institute of Science and Technology and Osaka University. His current research interests include human activity recognition, behavior change support systems, and location-based information systems. He is a member of ACM, IPSJ, and IEICE.