Identification of Air-Writing Tamil Alphabetical Vowel Characters

Rukshani Puvanendran
Dept. of ICT
University of Vavuniya
Vavuniya, Sri Lanka

Vijayanathan Senthooran
Dept. of ICT
University of Vavuniya
Vavuniya, Sri Lanka

Abstract*—In recent years, there has been a lot of focus on gesture recognition because of its potential as a means of communication for cutting-edge gadgets. As a special category of gesture recognition, air-writing is the practice of forming letters or words in the air using one's fingers or the movements of one's hands. The primary objective of this study is to propose a classification framework with feature extraction techniques to enhance the recognition of vowel characters in the Tamil language. The data collection and classification procedure involved a set of 12 distinct letters. A methodology has been developed to facilitate the analysis of various configurations for the purpose of evaluation. To get useful features from the 2-second time window data segments, this study uses a one-dimensional convolutional neural network (1D CNN). In our approach, we employ five machine-learning methods to conduct our evaluation. These methods include Naive Bayes, Random Forest, K-Nearest Neighbor, Support Vector Machine, and Decision Tree. The classification algorithms are considered to be superior based on the results obtained from our dataset in this experiment. The results of the tests show that the suggested K-nearest neighbors (KNN) algorithm works very well when used with a k-1 and 0.6:0.4 split ratio for training and testing. Specifically, the KNN model achieved an accuracy rate of 91.67%. The present study builds upon previous research by utilizing applications that have been employed in prior studies. However, a unique aspect of our system is the integration of cutting-edge technology, which utilizes collected sensor data to classify the characters. The examination of the window size has the potential to enhance accuracy and performance.

Keywords*—Air-writing; Tamil alphabetical vowel; convolutional neural network; feature extraction; machine learning

I. Introduction

The recent advancement in technology breaks the barriers to communication between users and computers. The communication between users and computers includes emotion and gesture recognition. Gestures offer a complementary mode of interaction for standard human-computer communication. New types of user interfaces made possible by gestures are especially well suited for portable and wearable computer systems. As a subset of gesture recognition, air-writing recognizes characters and numbers written in the air using the hand. Touchscreens and other technological gadgets have become fashionable in today's information-driven society, and a majority of individuals are comfortable using them [1]. As a result, air-writing recognition systems are gradually becoming more adaptable using wearable sensors and state-of-the-art techniques [2], [3], [4], [5].

Languages consist of alphabets that are combined to make meaningful words and are utilized for communication [6]. Therefore, these alphabets possess both symbolic and phonetic characteristics [6], [7]. The recognition of the air-writing characters of a particular language can be distinguished based on their symbolic structure [7], [8]. Most alphabets in any language consist of several complex motions, which means that sometimes a single sensor may not be enough to recognize these alphabets.

Usually, traditional handwriting styles of languages encompass two primary forms: cursive and print letters. Air-writing character recognition is done using information gathered from six-degree free-range hand motions. In air-writing each isolated letter is written in a virtual space. For user interfaces that do not allow users to type on a keyboard or write on a trackpad or touch screen, as well as for text input for controlling smart systems and many other applications, air-writing is extremely helpful [9], [10], [11]. Air-writing has been found to be a highly effective method for various applications, including but not limited to device control, entertainment, health care, and education.

The categorization of air-writing recognition approaches can be delineated into two distinct classes: Vision-Based Recognition (VBR) algorithms and Sensor-Based Recognition (SBR) algorithms [3], [12], [13]. The VBR algorithms are designed to execute gesture recognition tasks using image data acquired through a camera device. According to Amma's research, it has been observed that achieving accurate classification in image processing tasks can be feasible [4]. However, it is important to note that this process often demands substantial computational resources. Specifically, significant computational efforts are necessary for extracting relevant information from images during both the training and inference stages. This finding underscores the importance of considering the computational aspect when designing and implementing image processing algorithms [12]. The study conducted by Amma et al. [4], sheds light on the potential challenges associated with image-based classification tasks and highlights the need for efficient computational strategies in order to achieve optimal results. The techniques employed in SBR (Sensor-Based Recognition) are fundamentally rooted in the utilization of sensors. Accelerometers, gyroscopes, flex sensors, electromyography (EMG), Radio Frequency Identification (RFID), and the integration of these sensors have been widely employed in various applications [5], [14].

In the context, of hand gesture identification using SBR, accelerometers are widely used sensors in various applications because of their practicality, affordability, and durability [13], [15]. They have been extensively utilized in numerous prediction systems [16]. Wearable technology, such as smart phones, provides an optimal platform for the collection and monitoring of data [10], [17]. Furthermore, it has been reported that placing the sensor component on the wrist is the optimal placement position for a high degree of accuracy [5]. In some SBR-based studies, high classification accuracy for gesture recognition has been observed [2].

However, many of these techniques do not support the recognition of a sequence of gestures. Only isolated gestures can be recognized [19]. Furthermore, some SBR techniques are used for gesture recognition, which may incur high computation complexities [18]. The Recurrent Neural Networks (RNNs) and their variants, such as the Long Short-Term Memory (LSTM) algorithm, are effective for the recognition of gesture sequences with low computation costs [12], [19], [20], [21].

Many emerging algorithms have been proposed as the availability of temporal data has grown substantially in recent years [22]. To show time series as feature vectors, people often use dynamic time warping (DTW), simple statistics, complex mathematical methods, and other similar methods. For classification, they use algorithms that range from shallow learning to deep neural network models [23]. In addition, models like recurrent and convolutional neural networks (CNNs) incorporate feature engineering internally and automatically [24]. The One-Dimensional (1D) CNNs are effective for several applications, such as the classification of ECG signals [19], human activities, and internet traffic [16], [18], [25], [26], [27]. Researchers have found that using CNNs to classify time series is better than other methods in a number of important ways. This is because CNNs are very good at ignoring noise and can pull out very useful, deep features that are not affected by time [28]. In the context of CNN, one-dimensional convolutional neural networks (1D CNNs) excel at extracting useful features from smaller (fixed-length) subsets of a larger data set [29] and at analyzing audio signals, cyber security, NLP, and time series data from sensors [18], [30].

The majority of previous studies focused on the recognition of the English alphabet, where most of the alphabets are mostly drawn with lines that could subsequently be identified with the changes of the sensors [31], [32]. Some of the previous studies dealt with the designated devices for data collection and identification of the characters in second-language learning [33], [34]. Moreover, no prominent dataset is shared in publicly available sources.

Sensor-based air-writing has significant importance for the Tamil language alphabets. The recognition of handwritten words from a digital writing pad using sensor-based techniques can greatly benefit Tamil language learners and users. Jayanthi & Thenmalar (2023) proposed a method for recognizing handwritten words from a digital writing pad using the MMU-SNet algorithm [8]. The Tamil script is conventionally written in a horizontal manner, progressing from left to right. Its fundamental repertoire of characters encompasses 247 letters, including 12 vowels, 18 consonants, 12 vowels by 18 consonants (216), and 1 unique character [7]. The Tamil script is commonly recognized for its distinctive rounded shapes, leading to its colloquial designation as the "round alphabet" [6]. This approach can accurately capture the unique characteristics of Tamil alphabets, which have a smaller number of lines, angles, curves, and bends.

The study mainly improves the motion gesture recognition approach for six degrees of freedom (DOF) for air-written Tamil language characters. In the initial phase of the study, in spite of the motion movements, isolated air-writing characters are analyzed and identified. Motion sensing is done with sensors that are attached to the back of the palm. This led to a great deal of interest in the potential of wearable technologies for air-writing. A comprehensive list of features is extracted from the dataset of 3-axis signals from five different sensors with 1D CNN.

The present study focuses on the investigation of the recognition process pertaining to isolated characters that are written in a continuous single-stroke manner. This study employed one specific dataset, which was collected as part of the study. To sum up, this study presents the following contributions to the field of air-writing:

The authors have

- Collected the sensor data for 12 vowels of the Tamil alphabet.
- Developed a light feature extraction approach using 1D CNN from the collected signal data and populated the dataset for the classification.
- Evaluated the performance of the classification models with few configurations.

By leveraging this study, this method enables the recognition of Tamil alphabets written in the air, providing a valuable tool for Tamil language learners and users.

The structure of this article is as follows: We go over the relevant earlier work in the part after that. The data collection techniques for air-writing, the feature extraction process, and the methods for modeling are described in Section III. Section IV presents the experimental design results and discussions. Section V contains the conclusion of the study.

II. Related Works

Air-writing refers to the process of writing characters or words in free space using finger or hand movements [31]. It is a form of gesture recognition where gestures correspond to characters and digits written in the air [32]. Air-writing has been studied in various contexts, including its effectiveness in language learning and its potential applications in handwriting analysis and recognition.

In previous studies, certain research endeavors have employed a specifically engineered apparatus to carry out the process of data acquisition. In contemporary times, a plethora of wearable devices, including but not limited to smartphones and smartwatches, have been integrated with various sensors such as accelerometers, gyroscopes, magnetometers, and other similar components. Various methodologies have been put forth in scholarly works to address the task of air-writing recognition, wherein it is regarded as a spatial-temporal signal [14].

The wearable input system for handwriting recognition proposed by Amma et al. in [4] utilizes an accelerometer and a gyroscope to effectively capture and analyze each of the complex gestures involved in air-writing. Another study conducted by [5], a range of alterations to the Hidden Markov Models (HMMs) were implemented in order to facilitate the identification and interpretation of air-writing patterns that were generated through the utilization of the Leap Motion Controller.

Agrawal et al. [2] introduced the PhonePoint Pen system on the Nokia N95, which leverages the inherent accelerometers present in mobile phones for the purpose of recognizing handwritten English characters. The findings of this study indicate that the identification of English characters can be achieved with a mean accuracy rate of 91.9%.

Furthermore, Xia et al. [10] presented a MotionHacker system that utilizes a smartwatch application to capture and analyze the dynamics of hand movement, thereby demonstrating the potential for motion sensor-based handwriting.

The proposed system in [15] introduces a novel approach to handwritten character recognition. It leverages the power of 3D accelerometer signal processing in conjunction with Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) models. In the experimental setup, a participant carrying a MYO armband. The data in the dataset corresponds to handwritten English lowercase letters (a–z) and digits (0–9) that were written in a freestyle manner. The findings indicate that the proposed system exhibits superior performance compared to other existing systems, surpassing them by a margin of 0.53%.

Wang and Chuang in [13] proposed a novel digital pen that utilizes an accelerometer for the purpose of recognizing handwritten digits and gesture trajectories. Subsequently, the collected acceleration signals are subjected to a process of feature extraction, wherein a hybrid approach was employed to identify the most discriminant features. This involved the utilization of kernel-based class separability to determine the salient features, followed by the application of Linear Discriminant Analysis (LDA) to reduce the dimensionality of the dataset.

Two air-writing recognition methods were discussed in this study [1] dynamic time-warping and a Convolutional Neural Network. To assess the efficacy of the methodologies, a total of 15 instances of the English alphabet letters were obtained through airborne inscriptions and subsequently captured utilizing a smart-band device. These data sets were procured from a diverse pool of 55 individuals. The findings from the comprehensive evaluation yield an exceptional accuracy rate of 89.2%.

The study [35], introduces a novel framework for air-writing using a convolutional neural network (CNN) that is dependent on a generic video camera. The utilization of transfer learning with the recently obtained data leads to a notable enhancement in the recognition accuracy. The framework that was proposed demonstrated recognition rates of 97.7%, 95.4%, and 93.7% in person independent evaluations conducted on English, Bengali, and Devanagari numerals, respectively.

Additionally, a novel approach is proposed for character recognition of air-written text through the utilization of a 2D-CNN model for seven datasets [27]. In recent years, there has been significant progress in the field of handwritten character recognition, largely attributed to the utilization of deep convolutional neural networks (CNNs). These advanced neural network architectures have demonstrated remarkable accomplishments, surpassing established benchmarks by substantial margins. Currently, there is a lack of a standardized dataset in the field of air-writing research, which hinders the ability to assess the effectiveness of various methodologies. In recent times, a number of datasets have been made available for the purpose of recognizing air-writing characters.

The study [25], was performed using four different datasets (WISDM, HARDS, SBRHA, PAMAP2) with various characteristics, as previously used by similar recognition research. However, this includes the investigation to test the selected features with a variety of sensor placement locations, such as waist, wrist, chest, and ankle, sampling frequency, and the performance of the dataset. The purpose of [20] is to investigate the impact of window lengths with orientation invariant heuristic features on the performance of 1D-CNN-LSTM using data from 42 participants to recognize six human activities: sitting, lying down, walking, and running at three different speeds using information gathered from the Samsung Galaxy s7 smartphone's accelerometer sensor with the pre-installed Ethica application.

There are multiple benefits associated with the utilization of one-dimensional convolutional neural networks (1D CNNs) [28]. These advantages include notable performance on datasets with limited amounts of data, reduced computational complexity in comparison to two-dimensional CNNs and other deep learning architectures, expedited training processes, and a strong capability to extract pertinent features from sequential data and time series, such as signal data [26], [36].

Attributes capable of capturing patterns over a sliding window are formulated in the feature extraction phase. Time-domain features, frequency-domain features, and time-frequency-domain features are the three broad categories established by applying elementary statistical techniques to the features.

The impact of the one-dimensional convolutional neural network (1D CNN) technique on the field of air-writing recognition has been substantial. Numerous studies have been conducted to investigate the application of deep convolutional neural networks (CNNs) in the field of air-writing recognition, and these investigations have consistently proved the efficacy of such networks [20], [29].

Additionally, CNNs can be employed to analyze 1D signals, including electrocardiograms (ECG), electroencephalograms (EEG), and electromyograms (EMG), in healthcare applications. Lastly, CNNs can be utilized for machine fault detection purposes, among other applications [23]. The use of 1D CNN algorithms in sensor-based air-writing has made significant contributions to the field. These algorithms have been employed to recognize and interpret air-written characters and gestures accurately. The researchers investigated air-writing recognition using a 2D CNN model [32]. They extensively studied different interpolation techniques on publicly available air-writing datasets and developed a method to recognize air-written characters. Their findings highlighted the importance of choosing the proper interpolation technique for accurate recognition.

Few researchers have developed a simple yet effective air-writing recognition approach based on deep CNNs [9]. Their method utilized a 1D CNN architecture to recognize air-written characters. The results showed that the suggested method was very good at reading handwritten characters, showing that 1D CNNs could be useful in this area. The authors proposed an approach for air-writing recognition based on deep convolutional neural networks (CNNs).

Furthermore, [21] developed a wearable IMU-based human activity recognition algorithm for clinical balance assessment using a 1D-CNN and GRU ensemble model. Although this study focused on human activity recognition, the use of a 1D-CNN model demonstrates the potential of this algorithm in capturing and analyzing sensor data for various applications, including air-writing.

Furthermore, the use of 1D CNNs has been explored in other domains as well. applied a 1D CNN algorithm for the detection of water pH using visible near-infrared spectroscopy [28]. They interpreted the learning mechanism of the 1D CNN through visual feature maps generated by the convolutional layers. This demonstrates the versatility of 1D CNNs in various applications, including air-writing recognition.

In the field of handwriting analysis and recognition, air-writing has also been explored. Researchers have highlighted the importance of analyzing handwriting not only on paper but also in the air before the pen touches the paper. Significant differences have been observed between these two writing conditions [37]. Additionally, studies have focused on developing recognition systems for air-writing using deep learning and trajectory-based approaches. These systems utilize techniques such as deep neural networks and depth sensors to accurately recognize and track air-written characters [32], [11].

In a nutshell, the use of 1D CNN algorithms has made significant contributions to air-writing recognition. The studies mentioned above have demonstrated the effectiveness of deep CNNs in accurately recognizing air-written characters. The utilization of one-dimensional convolutional neural networks (1D CNNs) in the analysis and interpretation of intricate patterns in air-writing movements has facilitated progress and innovation in this particular domain.

Furthermore, the detection and tracking of fingertip movements in air-writing are crucial for accurate recognition. In [31], developed a fingertip detection and tracking algorithm specifically for air-writing recognition. This algorithm outperformed state-of-the-art approaches, achieving a mean precision of 73.1

Sensor-based air-writing offers several advantages for Tamil language learners and users. It provides a more interactive and immersive learning experience, allowing learners to practice writing Tamil alphabets without the need for physical writing materials. This can be particularly beneficial for learners who may have limited access to writing resources or prefer a more hands-on approach to learning. Additionally, sensor-based air-writing systems can provide real-time feedback and evaluation, helping learners improve their writing skills and accuracy.

In conclusion, sensor-based air-writing methods like reading handwritten words from a digital writing pad and finding and following fingertip movements are very important for the Tamil language alphabets. These techniques enable accurate recognition and analysis of Tamil alphabets written in the air, offering valuable tools for language learning, practice, and evaluation.

Our study differs from the existing studies in multiple aspects. Most of the prior research only used statistical equations to extract small subsets of features from the time domain, frequency domain, and time-frequency domain. Some of these studies also relied on publicly available datasets that were used for the purpose of the research. In contrast, this paper utilizes 1D-CNN to extract a complete set of features from five sensors' raw data. This investigation follows a multi-stage process and employs state-of-the-art techniques. In addition, unlike previous work, which relied on publicly available datasets, this study used data on Tamil alphabet recognition that was collected independently.

III. Methodology

A flowchart for the proposed air-writing recognition system is shown in Fig. 1. An approach is employed in our research for the purpose of 6-DOF character recognition.

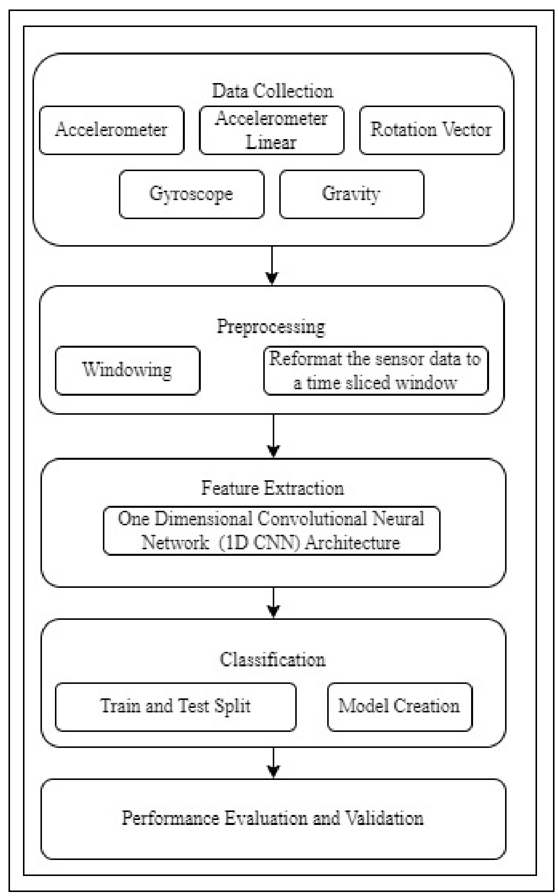

Fig. 1. The research flow.

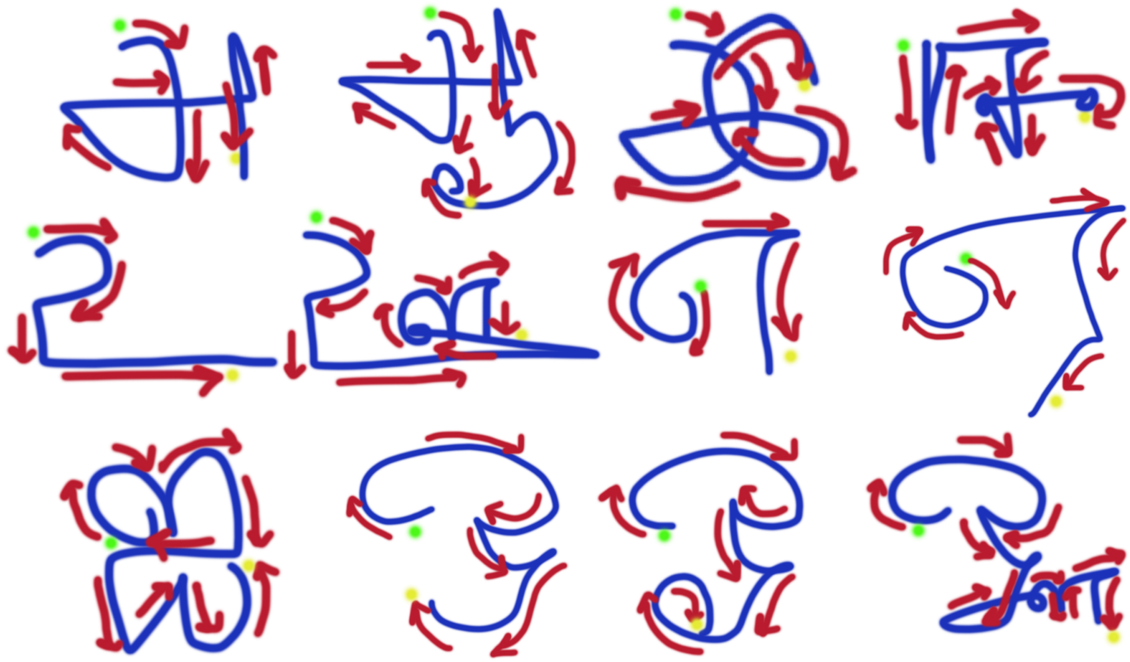

Fig. 2. The isolated characters of tamil vowel alphabets are written using a single continuous stroke, where the paths of the stroke are shown with arrows. The starting and ending points are denoted with green and yellow color dots, respectively.

A. Writing with Six Degree of Freedom (DOF)

Writing in the air differs significantly from writing by hand on paper or another surface. The writing is naturally reproduced in the air in uni-stroke without any pen-up or pen-down information, instead writing in a box in the imagined space without tactile input. Air-writing is tracked with a continuous stream of sensor data.To our best knowledge, we are the first to evaluate the rounded alphabet, i.e., Tamil language vowel recognition with axis-independent motions.

A total of five participants, all of whom were right-handed, were selected for the purpose of collecting air-writing data. Participants were instructed to maintain consistency in their writing style, including the use of a standardized writing scale, stroke order for each letter, and speed of writing. Each participant was given the opportunity to engage in a practice session before data collection commenced. Each instance of isolated motion characters was meticulously documented on different occasions by each participant.

The tracking system utilizes a smartphone device, which is monitored through an independent application. The application utilizes the input features to effectively acquire and record the sensor data associated with the task of writing. Subsequently, participants were provided with detailed instructions regarding the functionality and operation of the mobile application, as well as a comprehensive overview of the data collection procedure. Each participant carried an OPPO A12 smartphone attached to their right hand with the pre-installed application that can be recorded with five sensor data values with an interval of 10 ms. In order to mitigate the inclusion of extraneous movements, participants were instructed to execute the software with their non-dominant hand. Each letter was recorded separately. This was done in order to give the subjects the option to rest between recordings. and even change positions. The dataset pertaining to air-writing encompasses a comprehensive collection of 600 distinct samples for 12 vowels.

In order to enhance the machine's ability to recognize and expedite the user's writing process, the letters are subjected to a simplification process. Isolated characters are written using a single continuous stroke. The path of uni-stroke writing for different letters of Tamil vowel letters is illustrated in Fig. 2.

B. Preprocessing

Each individual sample, which corresponds to a single air-written letter, is regarded as a collection of distinct signals (time series). Based on the sensor data, it is possible to represent a writing motion as a spatio-temporal pattern. An analysis of the sensor values of the aforementioned sensors was performed in the spatio-temporal domain, wherein the relationship between sensor data is illustrated. The visual representation of the sensor data for the 12th vowel ('oow') is illustrated in Fig. 3. It serves as a tool for researchers to analyze and interpret patterns in the data collected.

C. Reformat Sensor Data into a Time-sliced Representation and Data Sampling

The variability in the duration of each signal's recording arises from the inherent characteristics of the writing and recording procedures. In accordance with the experimental design, a predetermined length, denoted as "l," has been established to ensure uniformity in the length of all signals across all samples. Achieving a suitable time window is a challenging task. Each window is considered non-overlapping as the characters are isolated. The raw data from each of the windows is considered an occurrence.

The sensor data from five sensors is accumulated into a single CSV format. We have tried truncating to make the signals consistent in length. This study selected an arbitrary window length as the intermediate time taken to write each character.

We have used the programming language Python to execute the required preprocessing of the data and extract the features, respectively. We have used packages named Pandas and Numpy for performing feature extraction.

D. The Motivation for using CNN

Significant advancements have been made in recent years in the domains of sensors, smart sensors, artificial intelligence, and their integration. In the field of machine learning, there is a current trend towards the utilization of Artificial Neural Networks (ANN), Deep Neural Networks (DNN) which are the prevailing techniques in the deep learning domain, and Convolutional Neural Networks (CNN). These concepts have gained popularity, indicating an increasing scientific interest and subsequent contributions from the scientific community. [19]

CNNs, also known as convolutional neural networks, are a prominent class of deep learning models that are widely used in various domains. These models are characterized by their architecture, which typically includes multiple layers of convolutions. These layers are designed to leverage a collection of adaptable multi-dimensional filters. These filters are systematically moved across the various axes of the input sample [35].

In various domains, including industrial, clinical, and environmental sectors, it is frequently observed that one-dimensional (1D) data is prevalent. Examples of such sensor data include electrocardiogram readings, temperature measurements, environment-control variables, motion data, and power consumption data. One potential approach for modeling this information is the utilization of one-dimensional convolutional neural networks (1D CNN) [28], [29], [36].

In the context of a one-dimensional convolutional neural network (1D CNN), the filters are applied by sliding them along a single axis of the input data. In the present scenario, it is observed that all dimensions of the filter's size, with the exception of one, are predetermined to align with the sizes of the fixed axes. In recent times, their utilization has extended to the realm of Temporal Sequence Recognition and Time Series Classification as well, as evidenced by the growing interest in the area. Temporal Sequence Classification refers to the task of accurately categorizing sequences of data that are captured by inertial measurement sensors. The objective is to assign these sequences to specific, pre-defined activities that occur over extended periods of time [26], [36]. This classification process takes place within a continuous stream of data, requiring robust and reliable algorithms to achieve accurate results.

In the present investigation, the decision was made to employ Convolutional Neural Networks (CNNs) due to two primary justifications. Traditional machine learning methods typically depend on the extraction of explicit features. In contrast, it is worth nothing that deep neural networks, particularly convolutional neural networks (CNNs), possess the ability to directly consume the raw input data without the need for explicit feature extraction procedures. Hence, it is plausible to hypothesize that a methodology that has demonstrated efficacy in the domain of temporal data analysis and computer vision holds promise for achieving favorable outcomes in our specific context.

Next, the set of preprocessed samples is introduced into the architecture, and a mini-batch learning process is implemented. The architectural design of our one-dimensional convolutional neural network (1D CNN) classifier draws inspiration from the model. The architectural design consists of convolutional layers that are strategically interspersed with pooling layers. In this study, it is observed that all of the convolutional layers in the experimental setup employ filters with a length of 3200 and 1600. Additionally, the Rectified Linear Units (ReLU) activation function is utilized across these layers.

Given the nature of our task, which involves classifying data into distinct vowels, we have opted to utilize categorical cross entropy as our chosen loss function. This particular loss function is well-suited for categorical classification tasks.

In the context of sensor technology, irrespective of the specific domain of application, it is frequently observed that data pertaining to the one-dimensional (1D) shape is prevalent [20], [29].

E. 1D CNN Architecture for Creating Feature Vector

The conventional architecture of a Convolutional Neural Network (CNN) consists of three primary layers, namely the convolutional layer, pooling layer, and fully connected layer. The layers in question employ various components and techniques, including convolutions, activation functions, pooling, dropout, batch normalization, and fully connected blocks, among others. These elements can be flexibly combined in numerous configurations.

A 1D CNN (Convolutional Neural Network) can be constructed based on the following considerations:

- The input data is formatted in one dimension. The data for this study was obtained from five sensors installed on a sliced window.
- The convolutional layers are tasked with performing feature extraction operations. The extraction process involves the application of convolution operations to the input data, with the resulting convolutions being passed as input to the subsequent layer. The convolutional operations are determined by a set of filters, a specified kernel size, padding, and stride. These operations result in the creation of a feature map, which is obtained by applying a ReLU activation function.
- Pooling layers are typically applied following a convolutional layer, serving the purpose of retaining the information produced by the feature maps. The techniques on these layers include max pooling.

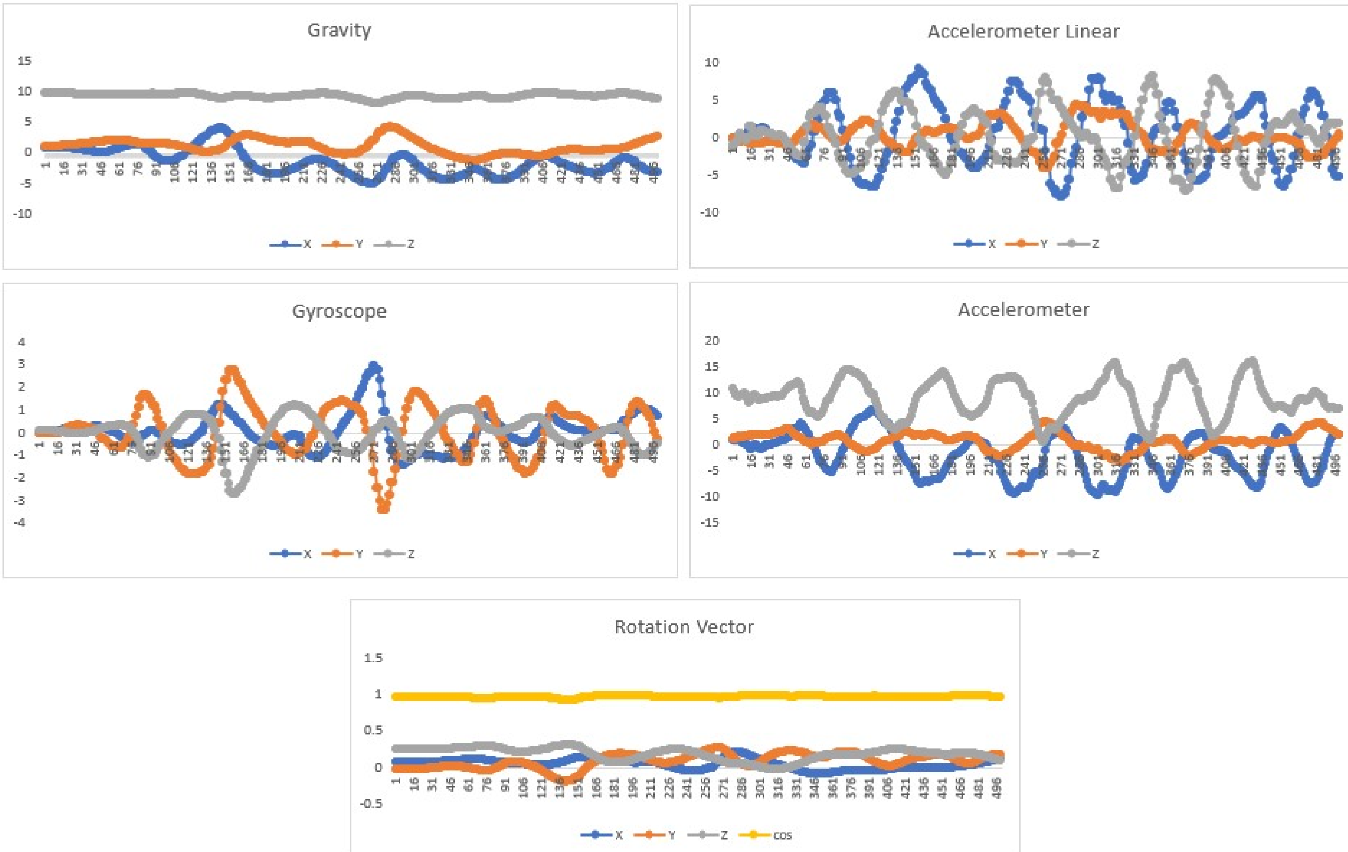

Fig. 3. The variation of the sensor data from five different sensors for the last letter of Tamil language vowel alphabet.

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d_15 (Conv1D)          (None, 3200, 32)          128

 conv1d_16 (Conv1D)          (None, 3200, 64)          6208

 conv1d_17 (Conv1D)          (None, 3200, 128)         24704

 max_pooling1d_5 (MaxPoolin  (None, 1600, 128)         0
 g1D)

 dropout_5 (Dropout)         (None, 1600, 128)         0

 flatten_5 (Flatten)         (None, 204800)            0

 dense_15 (Dense)            (None, 256)               52429056

 dense_16 (Dense)            (None, 512)               131584

 dense_17 (Dense)            (None, 128)               65664

=================================================================
Total params: 52657344 (200.87 MB)
Trainable params: 52657344 (200.87 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Fig. 4. The 1D CNN architecture.

The term “dropout” refers to an individual who leaves an educational institution before completing their program of study. The dropout technique is employed as a means of mitigating overfitting, and it can be implemented in both fully connected and convolutional layers. The technique involves randomly disabling certain connections within the neural network. This process ensures that individual neurons are able to be removed, thereby preventing them from exerting excessive influence on the model’s output. Alternatively, the technique enhances the generalizability of the model.

The application of the corresponding experiments, which have resulted in the 1D CNN architecture for the problem, is presented in Fig. 4. The input layer will consist of 300 inputs, which correspond to the sensor data for Tamil vowel characters.

F. Population of the Dataset

In order to facilitate analysis, it is necessary to extract features from the signal windows, as the raw data itself is not suitable for direct analysis. This process enables the identification and capture of patterns within the data. Previous studies in the field of recognition have utilized both time domain and feature domain features in their analysis. 1D CNN uses filters on each window to extract features automatically. 1D CNN maps these internally extracted features to different Tamil vowel alphabets. The extracted features are tabulated with their vowel labels using the label encoding technique.

However, when the 1D CNN feature vector is combined with the machine learning algorithms, the extracted features from the 1D CNN are used as inputs for the selected machine learning algorithms.

G. Split up the Data Set into Train and Test Sets

To facilitate the training process of our network, we initially partition the dataset into two distinct subsets: a training set and a testing set. The division is performed in such a way that the training set comprises 80% of the data, while the remaining 20% is allocated to the testing set.

H. Train the Dataset with Machine Learning Algorithms for Tamil Air-writing Alphabet Recognition

In this modeling and dataset training stage, we develop a model for classifying the features into the Tamil language vowel alphabet, we utilize various classifiers of machine learning. The use of Naïve Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT) and k-Nearest Neighbors (KNN) model for classification. These classification-based algorithms are based on performance metrics that include accuracy, precision, recall, and f1 score.

It is worth noting that our dataset is balanced, where each vowel is represented equally, ensuring a fair and unbiased evaluation of our model's performance. Therefore, we have selected accuracy as our performance measure, which will allow us to assess the effectiveness of our model in correctly classifying the data.

I. Validate the Performance of the Trained against the Test Data using Cross-validation and Confusion Matrix

In the conducted experiments, the relevant algorithm metrics that were evaluated included the following measures: The accuracy of the model is being evaluated with 80:20 split dataset. In the present scenario, the research pair has designated classification accuracy as the appropriate performance metric. The accuracy was assessed by measuring the performance in four different splits that were done on the original dataset. The datasets were designated as training, and testing sets, with a split of 80%, 20%, 70%, 30%, 60%, 40% and 50%, 50% from the original dataset, respectively. The objective was to develop a model that effectively generalizes learning from data. Consequently, a solution that exhibited consistent and homogeneous accuracy values was deemed favorable, rather than solely prioritizing the highest accuracy values within individual sub-datasets.

In summary, the experiments assessed the impact of the variant on the accuracy of the data sets with the parameter tuning of the selected model.

IV. Results and Discussions

A. Preprocessing of Sensor Data

The sensor data from five sensors is analyzed with regard to the variation in the data values. However, each of the letters lapses at different time intervals. We have chosen 2s as the preliminary window for the time-sliced segmentation, as shown in Fig. 5. As many of the vowel letters end up in the long lines or curve lines as shown in Fig. 2, the additional time can be ignored.

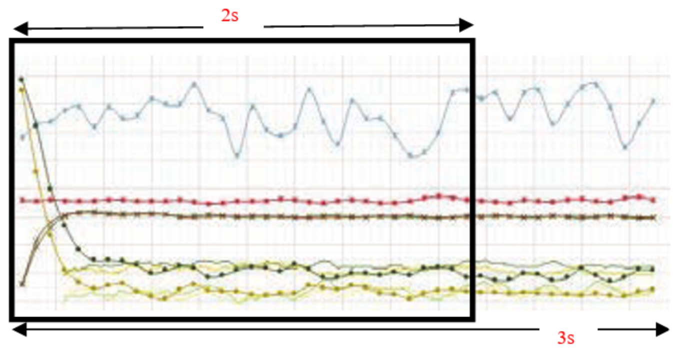

Fig. 5. The windowing process: the sensor data is sliced into the time window of 2s.

B. Feature Engineering Through 1D CNN

Each of the sliced windows of the sensor data was passed into the above-mentioned 1D CNN to create feature vectors. A total of 128 features were manipulated in the feature extraction stage, as illustrated in Fig. 4. Each of the feature rows in the feature matrix is labeled using a label encoder. The full feature set for the classification was populated in this manner.

C. Training and Testing of Dataset using Machine Learning Algorithms

As the initial step in this classification stage, the split training dataset is used for training, while the testing data is used for testing. The classification model was fitted with the training dataset. We have used five classification algorithms, such as Random Forest, Support Vector Machine, Decision Tree, Naive Bayes, and K-Nearest Neighbour. The analysis was first initiated with the 80:20 dataset split. It was found that we achieved the highest accuracy of 90.83% from the KNN algorithm. Fig. 6 shows the accuracy of the model in both testing and training. The accuracy for the training seemed to be 100% for RF, KNN and DT. However, in terms of testing accuracy, only KNN outperforms the others. We performed tests with numerous configurations. The first

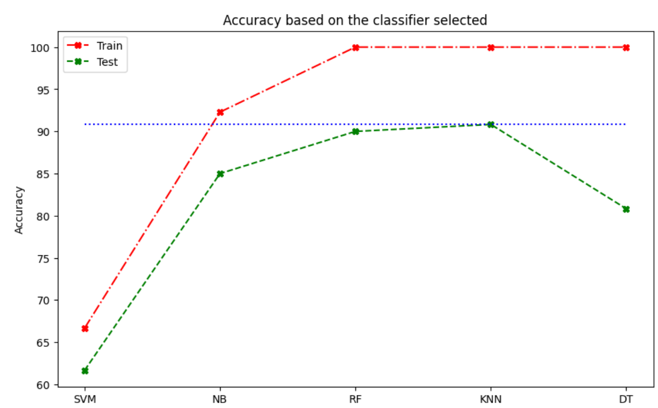

Fig. 6. The accuracy of the models (SVM,RF, NB, DT and KNN) during training and testing at 80:20 dataset split.

performance evaluation of the model is further extended using the accuracy for the different values of k. It was observed with the variation of the k value, ranging from 1 to 30. The accuracy of the model reduces with an increase in the k value. However, there are some fluctuations at certain places where a small rise prevails in the fall. Fig. 7 clearly illustrates the variation in accuracy with the k value. The accuracy for classifying the air-written characters with k = 1 and k = 2 depicts the maximum accuracy of 90.83%. This means that with k = 1 or k = 2, the classification is more accurate compared with the other k values.

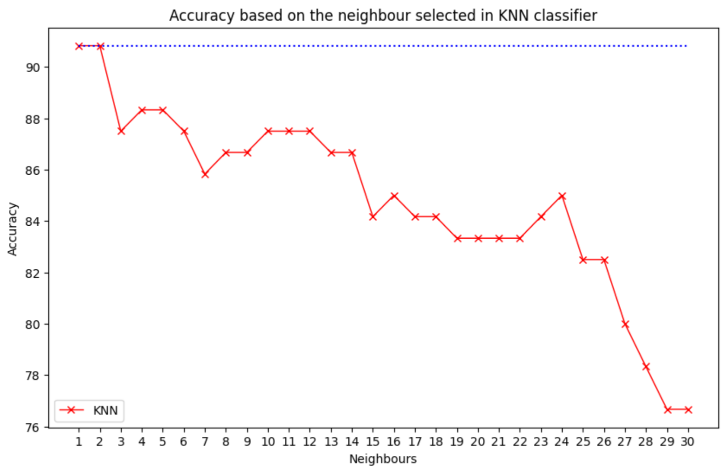

Fig. 7. The variation in accuracy with the change in k values.

Next configurations presented with the split percentage of 0.8, 0.7, 0.6 and 0.5 as training data and 0.2, 0.3, 0.4 and 0.5 as testing data, which had been taken into account for the evaluation of metrics. In this evaluation the k value is taken as 1, as it was obtained to be the highest accuracy. In Fig. 8, to classify the alphabets as per the samples used for training and testing, it has been observed that 0.6:0.4 achieves the maximum accuracy of 91.67% with the value k = 1, which shows the frequency of samples that are correctly classified to a specific air-written character within all the samples that are to be tested. Though the accuracy with the percentage of samples for 0.8, 0.7, and 0.6 increases, there is a sudden fall at the percentage of 0.5 of training data.

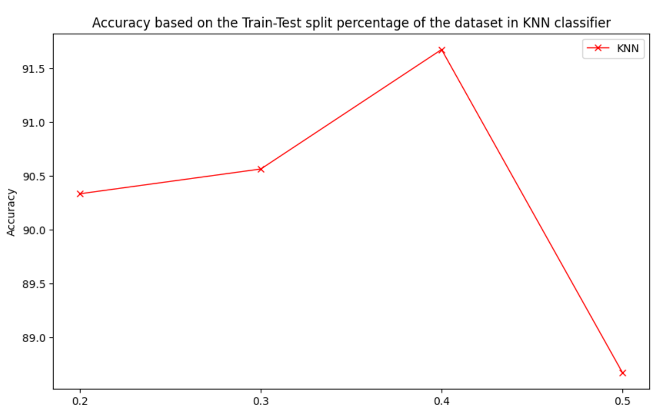

Fig. 8. The variation in accuracy with the change in the split percentage for KNN model, when k=1.

To provide the significance of the proposed model, its performance is being compared with multiple configurations. In this regard, the approaches are investigated that have been discussed previously.The confusion matrix of the model for the percentage samples of 0.6:0.4 and k=1 of KNN algorithm is shown in Fig. 9.

The accuracy, precision, recall, and F1-score of the classes of air-written characters are shown in Fig. 10. It was observed that the 3rd, 4th, 5th, 7th, and 10th characters can be easily distinguished from the others. The 6th, 11th, and 12th letters resembling the other letters show a lower accuracy rate. The evaluation has been performed with respect to the metrics used for the comparison.

The following are the findings of the evaluation:

- KNN outperforms the other classification algorithm.
- The value for k is chosen as 1 while varying from 1 to 30.
- The best test and train split is 0.6:0.4 for maximum accuracy.

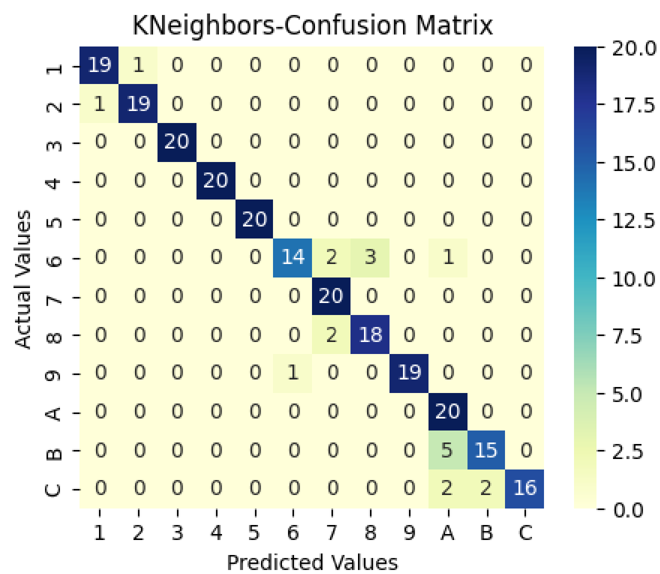

Fig. 9. The confusion matrix for KNN model when k = 1, with the 60:40 split percentage.

As this study has fully focused on the classification of the Tamil air-writing characters as of now, this is the first paper dealing with this objective.

D. Validation of the Performance

The validation of the model was tested with GridSearchCV with 10 folds for the above-identified configuration, and it was achieved with 93.33%. This study is summarized with the achievement of this accuracy in the initial phase of identifying the Tamil language air-written characters of vowel letters. We have achieved a considerable amount of achievement in the context of air-writing.

The process of identifying air-written Tamil characters poses numerous obstacles as a result of the distinct attributes

	precision	recall	f1-score	support
0	0.950000	0.950000	0.950000	20.000000
1	0.950000	0.950000	0.950000	20.000000
2	1.000000	1.000000	1.000000	20.000000
3	1.000000	1.000000	1.000000	20.000000
4	1.000000	1.000000	1.000000	20.000000
5	0.933333	0.700000	0.800000	20.000000
6	0.833333	1.000000	0.909091	20.000000
7	0.857143	0.900000	0.878049	20.000000
8	1.000000	0.950000	0.974359	20.000000
9	0.714286	1.000000	0.833333	20.000000
10	0.882353	0.750000	0.810811	20.000000
11	1.000000	0.800000	0.888889	20.000000
accuracy	0.916667	0.916667	0.916667	0.916667

Fig. 10. The performance evaluation report for KNN model when k = 1, with the 60:40 split percentage.

exhibited in handwritten Tamil, including variances in dimensions, styles, and angles of orientation. Moreover, the task of recognizing air-written characters presents intrinsic challenges due to the need to capture three-dimensional trajectories and employ suitable approaches for precise detection. This difficulty arises from the fact that many letters exhibit various features such as holes, loops, and curves. The study was conducted using a restricted dataset due to the unavailability of publicly accessible datasets. This work has the potential for further expansion through the comprehensive identification of all Tamil alphabetical characters. Additionally, enhancing the accuracy can be achieved by increasing the dataset and modifying the window size.

V. Conclusions

This paper focuses on developing a classification framework that employs a selection of features to assist us in recognizing Tamil vowel characters. A total of 12 different letters were used with the data collection and classification procedures. We have used a methodology for analyzing multiple configurations for evaluation. Initially, 1D CNN is used for feature extraction from the 2s of time window data segments. An evaluation is based on the results of five machine-learning methods: Naive Bayes, Random Forest, K-Nearest Neighbor, Support Vector Machine, and Decision Tree. The experimental results show that the proposed KNN achieves high accuracy with k = 1 and 0.6:0.4 train test split percentage with 91.67% accuracy. However, in validation using GridSearchCV this model has achieved an accuracy of 93.33% with 10 folds. Our findings were based on applications utilized in prior studies; however, as a novelty, the system provided employs collected sensor data to classify the characters by integrating the framework with cutting-edge technology. It is possible to increase the accuracy performance by examining the window size.

References

[1] Yanay, Tomer, and Erez Shmueli. ”Air-writing recognition using smart-bands.” Pervasive and Mobile Computing 66, 2020

[2] Agrawal, S., Constandache, I., Gaonkar, S., Roy Choudhury, R., Caves, K., & DeRuyter, F. Using mobile phones to write in air. Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services - MobiSys ’11. doi:10.1145/1999995.1999998, 2011.

[3] Amma, C., Georgi, M., & Schultz, T. Airwriting: Hands-Free Mobile Text Input by Spotting and Continuous Recognition of 3d-Space Handwriting with Inertial Sensors. 2012 16th International Symposium on Wearable Computers. doi:10.1109/iswc.2012.21, 2012.

[4] Amma, C., Georgi, M., & Schultz, T. Airwriting: a wearable handwriting recognition system. Personal and Ubiquitous Computing, 18(1), 191–203. doi:10.1007/s00779-013-0637-3, 2013.

[5] Chen, Mingyu, Ghassan AlRegib, and Biing-Hwang Juang. ”Air-writing recognition—Part I: Modeling and recognition of characters, words, and connecting motions.” IEEE Transactions on Human-Machine Systems 46, no. 3: 403-413, 2015.

[6] “World Languages”. [Online]. Available: https://www.mustgo.com/worldlanguages/tamil/ [Accessed: Oct. 3, 2023].

[7] “Tamil Script”. [Online]. Available:https://en.wiktionary.org/wiki /Appendix:Tamil_script [Accessed: Oct. 3, 2023].

[8] Jayanthi, V. and Thenmalar, S. Recognition of handwritten words from digital writing pad using mmu-snet. Intelligent Automation & Soft Computing, 36(3), 3551-3564. https://doi.org/10.32604/iasc.2023.036599, 2023.

[9] Hsieh, C., Lo, Y., Chen, J., & Tang, S. Air-writing recognition based on deep convolutional neural networks. Ieee Access, 9, 142827-142836. https://doi.org/10.1109/access.2021.3121093, 2021.

[10] Xia, Q., Hong, F., Feng, Y., & Guo, Z. MotionHacker: Motion sensor based eavesdropping on handwriting via smartwatch. IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). doi:10.1109/infcomw.2018.8406879, 2018.

[11] Alam, M., Alam, M., Abbass, M., Imtiaz, S., & Kim, N. Trajectory-based air-writing recognition using deep neural network and depth sensor. Sensors, 20(2), 376. https://doi.org/10.3390/s20020376, 2020.

[12] Chu, Yen-Cheng, Yun-Jie Jhang, Tsung-Ming Tai, and Wen-Jyi Hwang. ”Recognition of hand gesture sequences by accelerometers and gyroscopes.” Applied Sciences 10, no. 18: 6507, 2020.

[13] Wang, J.-S., & Chuang, F.-C. An Accelerometer-Based Digital Pen With a Trajectory Recognition Algorithm for Handwritten Digit and Gesture Recognition. IEEE Transactions on Industrial Electronics, 59(7), 2998–3007. doi:10.1109/tie.2011.2167895, 2012.

[14] Chen, Y., Yin, P., Peng, Z., Lin, Q., Zhao, Z., Fan, Q., … & Wei, Z. High-throughput recognition of tumor cells using label-free elemental characteristics based on interpretable deep learning. Analytical Chemistry, 94(7), 3158-3164. https://doi.org/10.1021/acs.analchem.1c04553, 2022.

[15] Lopez-Rodriguez, P.; Avina-Cervantes, J.G.; Contreras-Hernandez, J.L.; Correa, R.; Ruiz-Pinales, J. Handwriting Recognition Based on 3D Accelerometer Data by Deep Learning. Appl. Sci. 12, 6707. https://doi.org/10.3390/app12136707, 2022.

[16] Aakash, C., Kumar, P., Pilla, F., Skouloudis, A., Sabatino, S., Ratti, C., & Rickerby, D. End-user perspective of low-cost sensors for outdoor air pollution monitoring. The Science of the Total Environment, 607-608, 691-705. https://doi.org/10.1016/j.scitotenv.2017.06.266, 2017.

[17] Bastas, Grigoris, Kosmas Kritsis, and Vassilis Katsouros. ”Air-writing recognition using deep convolutional and recurrent neural network architectures.” 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 7-12. IEEE, 2020.

[18] “Human Activity Recognition”. [Online]. Available: https://towardsdatascience.com/human-activity-recognition-har-tutorial-with-keras-and-core-ml-part-1-8c05e365dfa0 [Accessed: Oct. 3, 2023].

[19] Rala Cordeiro, J.; Raimundo, A.; Postolache, O.; Sebastião, P. Neural Architecture Search for 1D CNNs—Different Approaches Tests and Measurements. Sensors 2021, 21, 7990. https://doi.org/10.3390/s21237990, 2021.

[20] Barua, A.; Fuller, D.; Musa, S.; Jiang, X. Exploring Orientation Invariant Heuristic Features with Variant Window Length of 1D-CNN-LSTM in Human ActivityRecognition. Biosensors 2022, 12, 549. https://doi.org/10.3390/bios12070549, 2022.

[21] Kim, Y., Kyunglim, J., Jeong, H., & Lee, S. Wearable imu-based human activity recognition algorithm for clinical balance assessment using 1d-cnn and gru ensemble model. Sensors, 21(22), 7628, 2021.

[22] “Time Series Classification with Deep Learning”. [Online]. Available:https://towardsdatascience.com/time-series-classification-with-deep-learning-d238f0147d6f [Accessed: Oct. 3, 2023].

[23] Khan, A., Sohail, A., Zahoora, U. et al. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53, 5455–5516, 2020. https://doi.org/10.1007/s10462-020-09825-6

[24] "Time Series Classification". [Online]. Available:https://towardsdatascience.com/how-to-use-convolutional-neural-networks-for-time-series-classification-56b1b0a07a57 [Accessed: Oct. 3, 2023].

[25] Bennasar, M.; Price, B.A.; Gooch, D.; Bandara, A.K.; Nuseibeh, B. Significant Features for Human Activity Recognition Using Tri-Axial Accelerometers. Sensors2022, 22, 7482. https://doi.org/10.3390/s22197482, 2022.

[26] Osman, Radwa Ahmed, Sherine Nagy Saleh, and Yasmine N. M. Saleh. "A Novel Interference Avoidance Based on a Distributed Deep Learning Model for 5G-Enabled IoT" Sensors 21, no. 19: 6555. https://doi.org/10.3390/s21196555, 2021

[27] Abir, F.A.; Siam, M.A.; Sayeed, A.; Hasan, M.A.M.; Shin, J. Deep Learning Based Air-Writing Recognition with the Choice of Proper Interpolation Technique. Sensors 2021, 21, 8407. https://doi.org/10.3390/s21248407, 2021.

[28] Li, D. and Li, L. Detection of water ph using visible near-infrared spectroscopy and one-dimensional convolutional neural network. Sensors, 22(15), 5809. https://doi.org/10.3390/s22155809, 2022.

[29] "Introduction to 1D convolutional neural networks in keras for time sequences". [Online]. Available:https://blog.goodaudience.com/introduction-to-1d-convolutional-neural-networks-in-keras-for-time-sequences-3a7ff801a2cf [Accessed: Oct. 3, 2023].

[30] Ismail Fawaz, Hassan, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. "Deep learning for time series classification: a review." Data mining and knowledge discovery 33, no. 4: 917-963, 2019.

[31] Mukherjee, Sohom, Sk Arif Ahmed, Debi Prosad Dogra, Samarjit Kar, and Partha Pratim Roy. "Fingertip detection and tracking for recognition of air-writing in videos." Expert Systems with Applications 136: 217-229, 2019.

[32] Abir, F., Siam, M., Sayeed, A., Hasan, M., & Shin, J. Deep learning based air-writing recognition with the choice of proper interpolation technique. Sensors, 21(24), 8407. https://doi.org/10.3390/s21248407, 2021.

[33] Mori, Y., Hasegawa, A., & Mori, J. The trends and developments of l2 japanese research in the 2010s. Language Teaching, 54(1), 90-127. https://doi.org/10.1017/s0261444820000336, 2020.

[34] Thomas, M. A role for "air writing" in second-language learners' acquisition of japanese in the age of the word processor. Journal of Japanese Linguistics, 30(1), 86-106. https://doi.org/10.1515/jjl-2014-0107, 2014.

[35] Roy, Prasun, Subhankar Ghosh, and Umapada Pal. "A CNN based framework for unistroke numeral recognition in air-writing." 16th international conference on frontiers in handwriting recognition (ICFHR), pp. 404-409. IEEE, 2018.

[36] Serkan Kiranyaz, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, Daniel J. Inman, 1D convolutional neural networks and applications: A survey, Mechanical Systems and Signal Processing, Volume 151, 107398, ISSN 0888-3270, https://doi.org/10.1016/j.ymssp.2020.107398, 2021.

[37] Diaz, M., Ferrer, M., Impedovo, D., Malik, M., Pirlo, G., & Plamondon, R. A perspective analysis of handwritten signature technology. Acm Computing Surveys, 51(6), 1-39. https://doi.org/10.1145/3274658, 2019.