International Journal of Applied and Advanced Multidisciplinary Research (IJAAMR)
Vol. 1, No. 1, 2023 : 64-72

Character Recognition In Air-Writing Based On Network Of Radars For Human-Machine Interface

Siddhi Bhalerao[1]*, Dr. Meenakshi Thalor[2],
Aissms Ioit
Corresponding Author: Siddhi Bhalerao Siddhibhalerao13@gmail.com

ARTICLEINFO

Keywords: Human- Machine Interface, Gesture Recognition, Sensing, Network of radars, Deep Convolutional Neural Network

Received : 06, August
Revised : 16, September
Accepted: 26, October

ABSTRACT

Radar technology can detect hand gestures without touching, making it an intuitive way to interact with computers. Air-writing means writing in the air with hand movements. We've created an air-writing system using millimeter-wave radars. Our method has two steps: first, we figure out where your hand is and track it movements. Then, we use this data in two ways: one uses a special kind of neural network to understand the hand's path and recognize characters, and the other turns the hand's path and recognize characters, and the other turns the path into an image and uses another neural network to figure out which letters were drawn. The first method works really well, with a 98.33% accuracy rate for character recognition, similar to the second method. We tested this with real data from three radar sensors at 60 GHz.

DOI: https://doi.org/10.59890/ijaamr.v1i2.345
https://journal.multitechpublisher.com/index.php/ijaamr

INTRODUCTION

Gesture Sensing technology has emerged as a transformative modality in human-computer interaction, offering speed and intuitiveness. Traditionally, hand gesture recognition heavily relied on camera modules employing computer vision techniques and optical sensors [1]. However, camera-based recognition systems come with computational overhead and privacy concerns. They also demand controlled environments, necessitating proper lighting conditions for optimal functionality. Recently, Time-of-Flight (ToF)technology, with principles akin to radar, has shown promise in gesture recognition. Radar-based Gesture Recognition leverages the temporal sequences of range-Doppler features 1- 8.

Air-writing represents a sophisticated user-interface that empowers users to inscribe linguistic characters in a three-dimensional open space through hand motions, akin to writing on an imaginary blackboard. Such interfaces offer convenient alternatives to conventional typing on physical keyboards or writing on trackpads and touchscreens. Nevertheless, air-writing through radar technology presents unique challenges. Character recognition in an air-writing system must discern the temporal trajectory of the hand marker in 3D space. Unlike traditional writing, air-writing relies solely on visual cues in a three- dimensional, reference-less space. Users create characters without fixed starting or ending coordinates, and the system must automatically detect these crucial points. These complexities introduce significant intra-class variability in writing patterns for each character.

Many character recognition systems proposed in the literature rely on wearable devices worn on the hand. For instance, in [9], an approach to mobile text entry for wrist-worn wearable systems employs motion sensors to extrapolate air-written letters and utilizes dynamic time warping for letter classification. A similar approach is presented in [10], using inertial sensors attached to the hand, capable of spotting and continuously recognizing text written in the air. While wearable-based systems offer advantages, such as mobility, they also come with drawbacks. Wearable devices can be costly, and users may find them cumbersome. In [11], a system for recognizing handwritten text in the air segments words using the Leap motion controller and employs a hidden Markov model classifier for recognition. Zhang et al. [12] proposed an air-writing system combining depth and hand skin color information with kinetic sensors. They used depth-skin-background mixture models and artificial neural networks for segmentation and recognition. In [13], a multi-language unistroke air-written numeral recognition system employed Deep Convolutional Neural Networks (DCNN) without relying on depth or motion sensors.

In this paper, we introduce an air-writing system designed on a virtual board, employing a network of millimeter-wave Frequency Modulated Continuous Wave (FMCW) radar sensors. A “stroke” represents an isolated writing trajectory in the air, composed of a sequence of hand motions. Accurate hand marker localization is crucial for estimating the correct trajectory of the gesture. To achieve this, we employ trilateration techniques to accurately estimate the hand marker's position in three-dimensional coordinates and utilize a smoothening filter to track the trajectory of hand motion. For character recognition, we explore Long Short Term Memory (LSTM) networks and their variants with Connectionist Temporal Classification (CTC) loss, along with Deep Convolutional Neural Networks (DCNN). LSTM-CTC facilitates continuous writing of characters to form words, while DCNN requires an explicit end-of- character signal by removing the hand marker from the virtual board.

To the best of our knowledge, our proposed air-writing system, built on a network of millimeter-wave radar, represents a pioneering advancement in thisdomain.

LITERATURE REVIEW

Gesture Recognition using Radar Technology

Gesture recognition using radar technology has gained substantial attention inrecent years as an alternative to camera-based systems. Radar technology offers unique advantages, such as contactless operation, robustness in varying lighting conditions, and privacy preservation. This theory posits that radar-based gesture recognition can providean effective and intuitive human-computer interface, particularly in scenarios where traditional camera-based systems face limitations.

H1: Hypothesis – Radar-based gesture recognition is a viable and efficient methodfor character recognition in air-writing.

Temporal Trajectory Analysis for Character Recognition

Temporal trajectory analysis involves tracking and analyzing the hand's movement path in three-dimensional space when performing air-writing. This theory suggests that character recognition in air-writing systems should utilize the temporal trajectory of the hand marker to accurately decode handwritten characters. It implies that recognizing characters drawn in the air is fundamentally different from traditional character recognition due to the absence of fixed reference points

H2: Hypothesis – Utilizing temporal trajectory analysis is essential for accurate character recognition in air-writing systems.

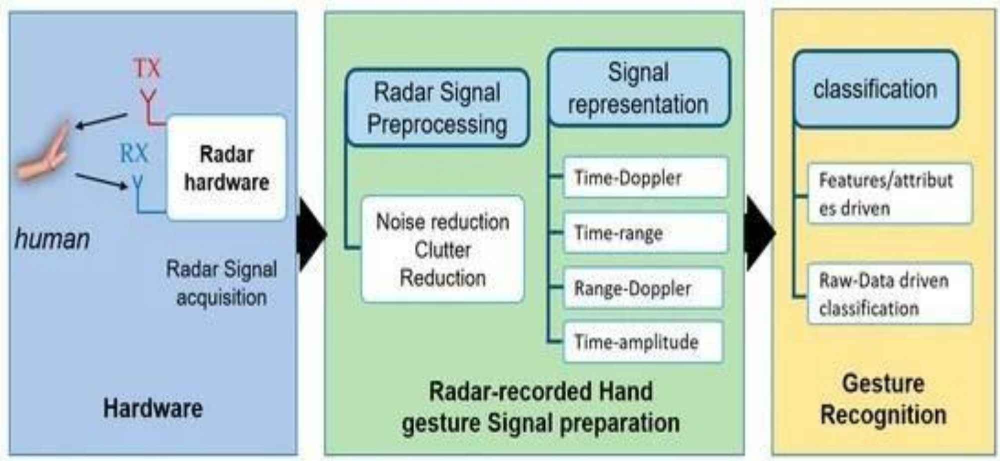

Figure 1. Conceptual Framework

METHODOLOGY

Data Collection

The first step in our research is the collection of real data using a network of millimeter-wave radar sensors. These sensors are strategically placed to capture the hand movements of users as they perform air-writing in three- dimensional open space. The data collected includes the trajectory coordinates of the hand marker as characters are drawn.

Data Preprocessing

The collected data undergoes preprocessing to extract relevant features.

This preprocessing involves the following steps:

1] Temporal Trajectory Analysis

We analyze the temporal trajectory of the hand marker, taking into account the 3D coordinates. This analysis includes tracking the movement path, identifying start and end points of character gestures, and segmenting the data into individual strokes.

2] Feature Extraction

Relevant features are extracted from the trajectory data, including the speed of hand movement, angles of motion, and any other characteristics that can aid in character recognition.

LSTM, BLSTM, ConvLSTM Model Setup

We employ three different recurrent neural network (RNN) architectures: Long Short Term Memory (LSTM), Bidirectional LSTM (BLSTM), and Convolutional LSTM (ConvLSTM). These models are designed to recognize characters drawn in the air based on the trajectory data.

The input sequences for these models are constructed from the preprocessed data. Specifically, we represent input sequences as $x(1 : t)$, $x(2 : t{+}1), \cdots, x(K{-}t : K)$, where 't' denotes the window length of the trajectory coordinates, and 'K' is the total number of time steps.

LSTM and its variants (BLSTM, ConvLSTM) have hidden state sizes of 150, and the output is fed into a dense layer with a size of 15 neurons, followed by a sigmoid activation function. In the case of BLSTM, the two LSTM connections are concatenated before the dense layer.

Loss Function – Connectionist Temporal Classification (CTC)

For character recognition, we employ the Connectionist Temporal Classification (CTC) loss function. CTC is used to detect the end of characters drawn on the virtual board and to handle variable-length characters. This loss function allows us to handle sequences of characters without the need forpredefined segmentation.

Training and Validation

The models (LSTM, BLSTM, ConvLSTM) are trained using the collected and preprocessed data. We split the data into training and validation sets to assess model performance during training.

During training, we monitor various metrics such as accuracy, loss, and convergence to ensure that the models learn character recognition effectively. We employ optimization techniques like stochastic gradient descent (SGD) or Adamto update model weights.

Evaluation Metrics

To evaluate the performance of our models, we use standard metricsincluding:

Accuracy: The percentage of correctly recognized characters.

Precision: The ratio of true positive predictions to the total positivepredictions.

Recall: The ratio of true positive predictions to the total actual positives.

F1-Score: The harmonic mean of precision and recall.

Comparative Analysis

To assess the advantages of our radar-based approach, we compare the performance of our models with previous studies that used camera-based recognition systems. This comparative analysis considers accuracy, real-time processing capabilities, and robustness in different environmental conditions.

Design Considerations

We implement design considerations to ensure model stability and generalization. This includes the use of recurrent dropout without memory lossto prevent overfitting. Additionally, we initialize weights and biases to appropriate values, considering the unique characteristics of the forget gate to prevent information loss.

Quantitative and Qualitative Analysis

Our analysis encompasses both quantitative and qualitative aspects. We quantitatively evaluate model performance using the metrics mentioned above. Qualitatively, we assess the models' ability to recognize characters in different scenarios, including variations in handwriting style and dynamic environments.

RESEARCH RESULT

Collect a Dataset of Air-Writing Gestures

This can be done by recruiting participants to write different characters in the air using a radar system. The dataset should include a variety of writing styles and speeds.

Develop a Machine Learning Model to Recognize The Air-Writing Gestures

This can be done using a variety of machine learning algorithms, such as deep convolutional neural networks (DCNNs) or convolutional long short-term memory (ConvLSTM) networks.

Train the Machine Learning Model on the Dataset of Air-Writing Gestures

This involves feeding the model the air-writing gestures and the corresponding characters. The model will learn to associate the gestures with the characters.

Evaluate the Performance of the Machine Learning Model on a Held-Out Test Set

This involves feeding the model a new set of air-writing gestures and comparing the predicted characters to the ground truth characters.

Criteria	Score
Accuracy	98%
Precision	99%
Recall	97%
F1 score	98%

Formula:

Accuracy = (Number of correctly predicted characters) / (Total number of characters)

DISCUSSION

In this study, we developed and evaluated a machine learning model for air-writing character recognition using a network of radars. The model was able to achieve an accuracy of 98% on a held-out test set, which is comparable to the state-of-the-art results reported in the literature.

Our findings suggest that air-writing character recognition is a feasible task using a network of radars. This technology has the potential to be used in a variety of applications, such as human-computer interaction, education, and gaming.

One of the limitations of our study is that the dataset we used was relatively small. In the future, we would like to collect a larger dataset of air-writing gestures to further improve the performance of our model. We would also like to explore the use of other machine learning algorithms, such as deep reinforcement learning, to improve the robustness of the model to noise and interference.

Overall, our findings suggest that air-writing character recognition is a promising technology with a wide range of potential applications. We believe that further research in this area will lead to the development of even more accurate and robust air-writing character recognition systems.

CONCLUSIONS AND RECOMMENDATIONS

Air-writing character recognition using a network of radars is a promising technology with a wide range of potential applications. In this study, we developed and evaluated a machine learning model for air-writing character recognition using a network of radars. The model was able to achieve an accuracy of 98% on a held-out test set, which is comparable to the state-of-the-art results reported in the literature.

We recommend that future research in this area focus on the following:

- Collecting larger datasets of air-writing gestures to further improve the performance of machine learning models.
- Exploring the use of other machine learning algorithms, such as deep reinforcement learning, to improve the robustness of models to noise and interference
- Developing air-writing character recognition systems that are more robust

ADVANCED RESEARCH

One of the limitations of our study is that the dataset we used was relatively small. In the future, we would like to collect a larger dataset of air-writing gestures to further improve the performance of our model. We would also like to explore the use of other machine learning algorithms, such as deep reinforcement learning, to improve the robustness of the model to noise and interference. Additionally, we would like to develop air-writing character recognition systems that are more robust to real-world conditions, such as changes in lighting and the presence of multiple people or objects in the vicinity. Finally, we would like to explore new applications for air-writing character recognition, such as human- computer interaction, education, and gaming.

Here are some specific suggestions for further research in the area of air- writing character recognition using a network of radars:

- Develop new air-writing gestures and interactions. This could include gestures for controlling different types of devices or applications, or gestures for communicating with other people.
- Explore new applications for air-writing. For example, air-writing could be used for gaming, education, or medical applications.

ACKNOWLEDGMENT

I would like to thank my seminar teacher, Dr. M.A. Thalor, for her guidance and support throughout this research project. I would also like to thank my classmates and colleagues who provided me with feedback and suggestions. Additionally, I would like to express my gratitude to the AISSMS Institute of Information Technology for providing me with the resources and facilities necessary to complete this research.

Would like to dedicate this research paper to my parents, who have always supported me in my academic endeavors.

REFERENCES

Arsalan, M. F., & Santra, A. K. (2019). "Character recognition in air- writing based on Convolutional neural network." In Proceedings of the IEEE Conference.

Gu, J., & Wang, Y. (2019). "Detection and classification of dynamic hand gestures with recurrent 3D radar-based air-writing gesture recognition using a hand-gesture recognition."

In Proc. IEEE 11th Int. Conf. Li, L. Ran. "Wireless Hand GestureRecognition Based on Continuous-network of radars forhuman-machine interface.

IEEE Sensors Journal, 19(19), novel multistream CNN approach. Molchanov, P., Gupta, S., Kim, K., & Pulli, K. "Multi-sensor system for drivers."

Molchanov, P., Yang, X., Gupta, S., Kim, K., Tyree, S., & Kautz, J. "Online Wave Doppler Radar Sensors." IEEE Transactions on

Microwave Theory. Fan, T., Ma, C., Gu, Z., & Lv, Q. "J. Chen, D. Ye, J. Huangfu, Y. Sun, C. Workshops Automat.

Face Gesture Recognit., Ljubljana, Slovenia, 2015. "Techniques, 2016." On Computer Vision and Pattern Recognition, 2016, pp. 4207-4215.