

Article

Real-Time Finger-Writing Character Recognition via ToF Sensors on Edge Deep Learning

Jiajin Zhang ¹, Guoying Peng ¹, Hongyu Yang ², Chao Tan ³, Yaqing Tan ^{1,*} and Hui Bai ^{4,*}¹ College of Big Data, Yunnan Agricultural University, Kunming 650201, China² College of Mechanical and Electrical Engineering, Yunnan Agricultural University, Kunming 650201, China³ Key Lab of Process Analysis and Control of Sichuan Universities, Yibin University, Yibin 644001, China⁴ College of Architectural Engineering, Yunnan Agricultural University, Kunming 650201, China

* Correspondence: yatsingt@ynau.edu.cn (Y.T.); zjjcl@ynau.edu.cn (H.B.)

Abstract: Human–computer interaction is demanded for natural and convenient approaches, in which finger-writing recognition has aroused more and more attention. In this paper, a device-free finger-writing character recognition system based on an array of time-of-flight (ToF) distance sensors is presented. The ToF sensors acquire distance values between sensors to a writing finger within a 9.5×15 cm square on a surface at specific time intervals and send distance data to a low-power microcontroller STM32F401, equipped with deep learning algorithms for real-time inference and recognition tasks. The proposed method enables one to distinguish 26 English lower-case letters by users writing with their fingers and does not require one to wear additional devices. All data used in this work were collected from 21 subjects (12 males and 9 females) to evaluate the proposed system in a real scenario. In this work, the performance of different deep learning algorithms, such as long short-term memory (LSTM), convolutional neural networks (CNNs) and bidirectional LSTM (BiLSTM), was evaluated. Thus, these algorithms provide high accuracy, where the best result is extracted from the LSTM, with 98.31% accuracy and 50 ms of maximum latency.

Keywords: finger-writing character recognition; time of flight; edge deep learning; distance sensor



Citation: Zhang, J.; Peng, G.; Yang, H.; Tan, C.; Tan, Y.; Bai, H. Real-Time Finger-Writing Character Recognition via ToF Sensors on Edge Deep Learning. *Electronics* **2023**, *12*, 685. <https://doi.org/10.3390/electronics12030685>

Academic Editor: Jaime Lloret

Received: 28 December 2022

Revised: 18 January 2023

Accepted: 23 January 2023

Published: 30 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Text input, transferring words in mind into digital information, occurs frequently in our day-to-day activities. Typing, the most used method for the majority of individuals in human–computer interaction, requires dedicated writing instruments, such as keyboards or touch screens, which are not always available in everyday life [1]. W. Chen et al. [2] presented an on-body typing system that allowed users to type on the backs of their hands on a T9-shaped keyboard. It is an advance in typing systems, but typing with one hand on a T9 keyboard is still awkward and slow. As a replacement for typing, speech/voice recognition is more friendly for those who struggle with keyboards and has been developed over many years [3,4]. However, it is sensitive to noise levels in the surroundings and could disclose sensitive information [5]. To fill this research gap, Silent Speech Interface (SSI) using deep neural network models and ultra-sound images to monitor a user’s unvoiced utterance and convert it into speech signals was proposed [6,7]. Unfortunately, real-time articulatory-to-acoustic mapping has not been accomplished by SSI.

Handwriting recognition is currently receiving more and more attention as a way to go beyond the constraints of typing and speech recognition. As a part of our body, hands and fingers are the earliest and most frequently used tool, and they are not an extra burden for us. In early human history, our ancestors painted on cave walls with their hands and fingers [8]. Various novel sensors have been applied in handwriting recognition to enable users to write freely and naturally like our predecessors did, as well as to accurately read what has been written. Inertial-based sensors are often used in wearable devices attached to hands or fingers to gather precise body motion data [9–11]. Optical sensors provided a

contact-free solution to track hand movement for recognition [12]. In addition, the use of acoustic sensors [13], radar sensors [14], ToF sensors [15], etc., in handwriting recognition has also been studied. All these sensing methods have their advantages and disadvantages, considering user convenience, computational complexity and noises in the environment.

The majority of existing handwriting recognition methods equipped with different sensors can be split into two groups based on whether they require direct contact with hardware [13]. Generally, digital pens [9,16,17], smart watches [18], smart bands [19] and other devices [10,20] using inertial-based sensors are attached to skin to acquire detailed body motion data for classification. M. Schrapel et al. [9] used a microphone and an inertial measurement unit (IMU) in a pen for handwritten digit recognition to record audio and motion data while writing. T. T. Alemayoh et al. [16] also operated with a smart pen equipped with an IMU sensor and three small force sensors for character recognition, achieving a validation accuracy of 99.05%. Z. Zhou et al. [18] collected accelerometer and gyroscope data from two smart watches, tested on the dataset of Hong Kong sign language, and revealed a significantly lower word error rate when compared to other existing machine learning or deep learning algorithms in the field. All of the research described above requires the user to hold or wear the devices, which might occasionally make the user feel constrained and unpleasant. Methods without device contact are more natural and relaxed approaches for those who utilize recognition systems. S.Z. Gurbuz et al. [21] used RF sensors for American Sign Language (ASL) recognition and classified 20 native ASL signs with 72.5% accuracy. D. S. Breland et al. [22] presented a robust hand-gesture recognition system based on high-resolution thermal imaging, which was light-independent and can be performed in low-light conditions. However, handwriting is distinct from ASL or hand gestures that are made up of a collection of specific movements that must be remembered before use. N. A. Khan et al. [12] applied optical sensors for air-writing recognition and reached an average accuracy of more than 90%, but the method light sensitive. B. Saez-Mingorance et al. [23] used one array of ultrasonic transceivers to track the pairwise distance of the hand with each transceiver for character classification. The best result had a latency of 71.01 milliseconds and an accuracy of 99.51%. These algorithms were operated on the local computers, limiting their application in a wide range.

An increasing number of deep learning algorithms have been studied for handwriting recognition. A shallow convolutional neural network (CNN) with a feature cube as its input for gesture detection and recognition achieved an encouraging classification result in real-time applications [20]. A multilevel decision (MLD) algorithm incorporating a lightweight support vector machine (SVM) algorithm was developed to operate computation for keystroke recognition and showed satisfactory performance [19]. Both algorithms mentioned above run on edge devices, while many other algorithms, e.g., long short-term memory recurrent neural network (LSTM) and BiLSTM, have achieved satisfactory performance in handwriting recognition, but they are rarely applied in edge-based tasks. BiLSTM is an improved version of LSTM, inheriting the strong sequential learning capabilities of LSTM while enabling the unit to collect knowledge from both its previous units and its future units simultaneously. BiLSTM has emerged as one of the most common methods for tackling a variety of sequential data problems, including speech recognition and handwriting recognition [18].

Edge devices for handwriting recognition have been developed because they are portable and convenient. H. Zhang et al. [10] presented a real-time wearable system for finger air-writing recognition based on an edge device, which was small enough to wear on the index finger. S. Boner et al. [24] developed an integrated small and low-power processor and gathered training data with the ToF sensors for gesture recognition. The system was significantly smaller than an adult's palm. T. Xing et al. [25] proposed a lightweight Wi-Fi gesture recognition system, which was designed and implemented for deployment on low-end edge devices. According to extensive experiments, the proposed system classified different activities quickly and accurately in 0.19 s, with an accuracy of up to 96.03%. This research on edge devices filled the gap for real-time and portable systems in text recognition,

but the primary problem is that they either employed wearable technology or focused on gesture recognition.

Edge devices deploying deep learning algorithms are also crucial in handwriting recognition because of their portability and convenience. After hand motion data are collected, they are sent to conventional workstations, Cloud or edge devices for analysis. Comparing with the other two, edge devices require lower energy consumption and latency [26,27]. Some works on the edge devices have used CNN for character classification tasks and achieved high accuracy [10,14,26]. While other machine learning or deep learning techniques, such as SVM, LSTM, BiLSTM, etc., have also been widely used in handwriting recognition [28–30], relatively little research on edge devices has been conducted. The application of different deep learning algorithms on low-cost and low-power edge devices is a promising research direction in handwriting recognition.

In this paper, to provide an efficient and low-cost text input solution targeted at small-screen devices (e.g., smartwatch) or people with poor eyesight, we propose a finger-writing character recognition system based on an array of ToF sensors. ToF sensors offer small size, accurate and low-cost options for three-dimensional (3D) imaging applications without body contact [31]. The ToF sensors acquire pair distance values between each sensor to a writing finger within an 9.5×15 cm zone on a flat surface at specific time intervals. Writing on a surface, such as a desk, is frequently favored because it provides users with touch-based feedback, enabling them to write more normatively. Collected distance data are sent to a low-power microcontroller STM32F401 equipped with deep learning algorithms for real-time inference and recognition tasks. The proposed method enables one to distinguish between 26 English lower-case letters by users writing with their fingers and does not require one to wear additional devices. Multiple classification algorithms have tested their performance in writing recognition, including LSTM, CNN and BiLSTM. These algorithms were chosen in light of their highly accurate outcomes for text-writing recognition obtained by other researchers [19,29,30,32].

Compared with previous work, the main contributions and innovations of this paper are summarized as follows:

- (1) We develop a finger-writing character recognition system based on TOF distance sensors, which is suitable for deployment on low-end devices for text input without the support of any additional high-end computing facility.
- (2) We employ simple and economically friendly sensors to decode the finger-writing pattern efficiently.
- (3) We design three deep learning algorithms for execution on low-end edge devices to perform fast recognition with high accuracy.
- (4) Meanwhile, our approach is an alternative text input solution to small-screen devices (e.g., smartwatch) or people with poor eyesight.

The rest of this paper is organized as follows: Section 1 introduces the current research status and related works about finger-writing recognition. Section 2 describes the architecture of the system. Section 3 presents the experiment for data acquisition. Section 4 describes the structure of deep learning algorithms and the specific parameters used for each algorithm. Section 5 summarizes the results and provides a detailed analysis of different algorithms used in the system. Section 6 focuses on the conclusions of this work.

2. System Description

The proposed system aims to realize a low-power and online finger-writing recognition system based on an edge device. Figure 1 illustrates the overall architecture of this system. Our system mainly includes three VL53L3CX ToF proximity sensors and an Arm Cortex[®]-M4F processor, which are all embedded in a driving module. Users draw letters with their index fingers in the handwriting area on a flat surface that is within the detection range of all VL53L3CX ToF sensors; the system realizes finger-writing character recognition online via edge computing. More specifically, data are collected by VL53L3CX sensors, which capture the distance signal data of the moving fingers and send them to the processor.

Then, deep learning models are applied to recognize finger writing by an Arm Cortex-M4F processor, and the recognition results are transmitted to the other devices (such as a laptop, a smart phone, etc.) to a display using a USB serial communication unit.

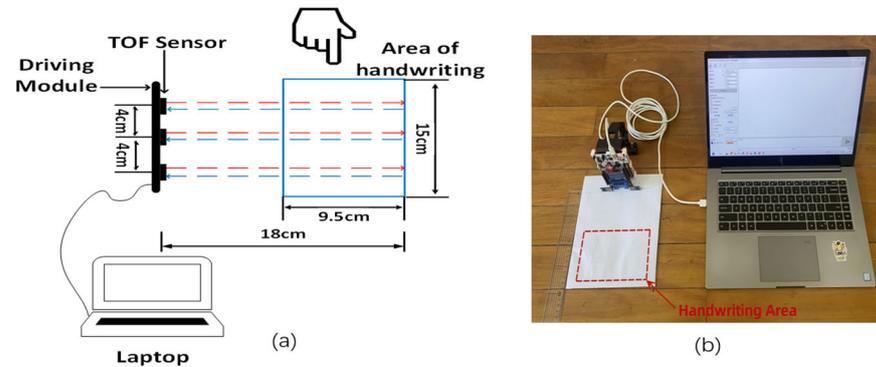


Figure 1. System overview. (a) Schematic of a distance-measuring ToF sensors. (b) Experimental measuring device.

2.1. Driving Module

The driving module, including the STM32F401RE Nucleo development board and the X-NUCLEO-53L3A2 expansion board, is shown in Figure 2. The STM32F401RE Nucleo is the Micro Controller Unit (MCU) in the X-NUCLEO-53L3A2, which enables USB communication with a laptop and is an expansion board that receives and processes the measurement data from a ToF sensor board as an input signal. X-NUCLEO-53L3A2 is a board with three built-in VL53L3CX sensors. It drives the VL53L3CX and transmits the measured data to the expansion board [33].

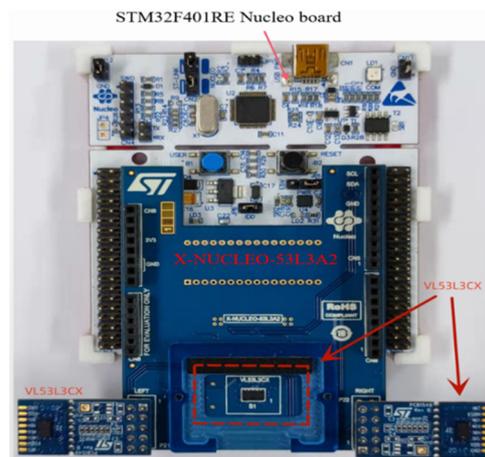


Figure 2. Driving module.

VL53L3CX ToF sensor:

ToF is an accurate and easy-to-understand technology used for distance measurement. It measures distances using the elapsed time that photons travel between two points, from the sensor emitter to a target and then back to the sensor receiver.

The following formula computes the distance:

$$d = \frac{1}{2}c\tau \quad (1)$$

where d is the measured distance, c is the speed of light and τ is the photon travel time.

In this study, three VL53L3CX ToF sensors embedded in the X-NUCLEO-53L3A2 expansion board are utilized to measure the distances between the static ToF sensors and

a writing finger within a 9.5×15 cm square on a flat surface. Continuous and dynamic distance data in finger-writing characters help neural network models extract more writing features and to enhance the classification capability of the models. The VL53L3CX, which integrates a single-photon avalanche diode (SPAD) array and physical infrared filters to achieve the best ranging performance in various ambient lighting conditions, combines the benefits of a high-performance proximity sensor with excellent short-distance linearity. The specifications for the sensor are shown in Table 1.

Table 1. Technical specifications of VL53L3CX sensor.

Feature	Detail
Package Size	Optical LGA12 $4.4 \times 2.4 \times 1$ mm
Operating voltage	2.6 to 3.5 V
Infrared emitter	940 nm
Interface	I ² C bus
Distance	1–300 cm
Field of view (FOV)	25°

2.2. Edge Computing Unit

The edge computing device in this work is an STM32F401RET6U MCU in a Nucleo-F401RE board. The MCU is based on the high-performance ARM Cortex M4 core equipped with a Floating-Point Unit (FPU), operating at frequencies up to 84 MHz. The board is supplied with 512 Kbytes of flash memory and 96 Kbytes of SRAM. The MCU offers a 12-bit Analog to Digital Converter (ADC), six general-purpose 16-bit timers and two general-purpose 32-bit timers. Moreover, thanks to its comprehensive set of power-saving modes, the MCU allows one to design custom low-power applications that can perfectly fit the requirements for TinyML.

In addition, VL53L3CX API is run on STM32F401RET6U MCU, enabling control of the operation of the ToF sensor and the display of the measured data in this study. An artificial intelligence package from STMicroelectronics, called X-Cube-AI, is used to load and benchmark the previously trained deep learning models on an STM32F401RET6U created in a PC server.

Briefly, the STM32F401RET6U, as an edge device, is in charge of distance sensor data collection, pre-processing, inference and recognition tasks on devices with edge deep learning algorithms.

3. Data Collection

3.1. Data Acquisition Module

Data acquisition begins when a finger moves in an area of 9.5×15 cm that is 8.5 cm from the ToF sensors for more than 15 ms. Three ToF sensors obtained distance data between themselves to the writing finger, with a time interval of 30 ms in sequence and repeated 20 times. The distance values more than 18 cm were removed in order to guarantee that the model was trained on reliable data. After trimming the raw data, each dataset has a shape of 20×3 , as in (2), where the row represents data gathered by three ToF sensors for once and the column is the width of the shifting window. These datasets can be directly fed into LSTM and BiLSTM, while being reshaped as 60×1 as the input of CNN.

Based on experimental results, writing on simple characters, e.g., “a”, “c” and “h”, which can be completed without leaving the surface, takes 0.7 s, on average, while the relatively

complex characters, such as “f”, “k” and “t”, take 1.2 s, on average, ensuring that the whole trajectory of finger motion is obtained within the setting time (1.8 s, 30 ms × 3 × 20).

$$d = \begin{cases} d_{1,1} & d_{1,2} & d_{1,3} \\ d_{2,1} & d_{2,2} & d_{2,3} \\ & \vdots & \\ d_{20,1} & d_{20,2} & d_{20,3} \end{cases} \quad (2)$$

3.2. Data Acquisition System Experiment

As the subjects in this study, 21 students from our institute were recruited to collect data for model training. Of the 21 students, both sexes were represented fairly equally, with a slight male majority of 12 (57.1%) students. Regarding the participants' dominant hand, there were 18 (85.7%) right-handed students and 3 left-handed students (14.3%) among the 21 students.

Without specific instructions on how to write characters, participants were asked to draw characters in their natural writing style within a 9.5 × 15 cm zone on a flat surface using their index finger as vertically as possible to minimize errors caused by different hand shapes (Figure 3). The writing speed for these characters was chosen freely by the subjects. When the subject finished writing a letter, the laptop displayed a list of 60 number integers, which is the motion data collected by three ToF sensors for the letter. At the same time, the subjects can choose to delete the recording or start again with the character. The subject then started writing the next letter. To avoid recording unconscious movements, such as movements that are not associated with the writing itself, each subject was asked to offer 15 sets of all 26 characters. Figure 4 shows examples of the data collected by ToF sensors for finger-written characters “i”, “m”, “p” and “x”. When comparing the three sensors of the four letters, there are clear differences to distinguish the characters, even by the naked eye.

In total, 8190 datasets were prepared. Out of these, 60% of the datasets were used for classification model training, 20% were used for validation during the training and the other 20% were left for testing the trained model.

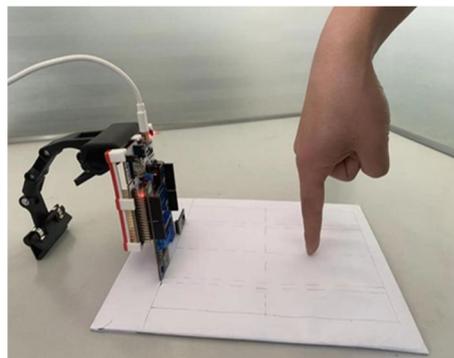


Figure 3. Drawing letters vertically on the flat surface.

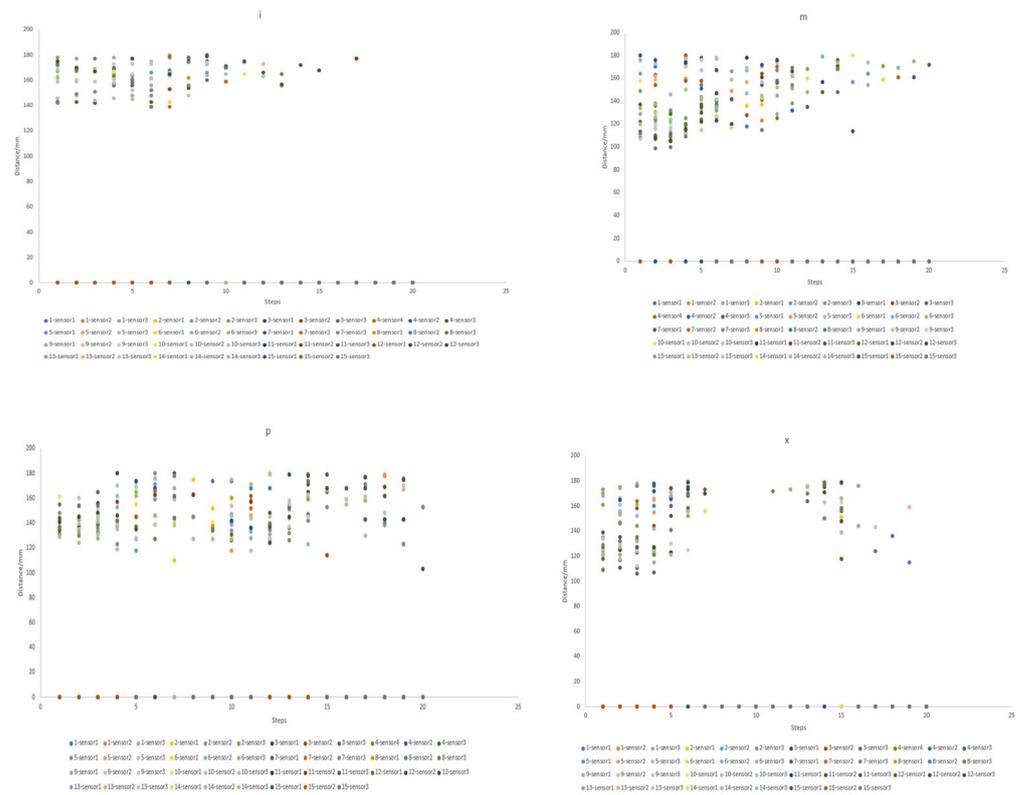


Figure 4. The signals of “i”, “m”, “p” and “x” from three ToF sensors.

4. Structure of the Deep Learning Algorithms

Unlike traditional machine learning, deep learning can automatically discover and extract the features from the raw sensor data without extracting hand-crafted features and is widely used in activity recognition. In this section, three deep learning models, including LSTM, CNN and BiLSTM, were investigated for the classification of 26 English lower-case letters by finger writing.

4.1. Long Short-Term Memory (LSTM)

Due to their capacity to transcribe data into sequences of characters or words while maintaining sequential information, LSTM networks are frequently used in handwriting and speech recognition applications. In this study, finger-writing data were collected by an array of ToF every 30 ms forming time-sequential information like that used in forecasting. As a result, the LSTM model is a good candidate for deep learning training.

In this study, the structure of this network is a single LSTM layer with 60 cells, with a time-step size of 3, followed by one fully connected layer with 26 neurons. The LSTM layer uses a tanh activation function while the fully connected layer uses the traditional SoftMax activation function for classification tasks. The structure can also be visualized in Figure 5.

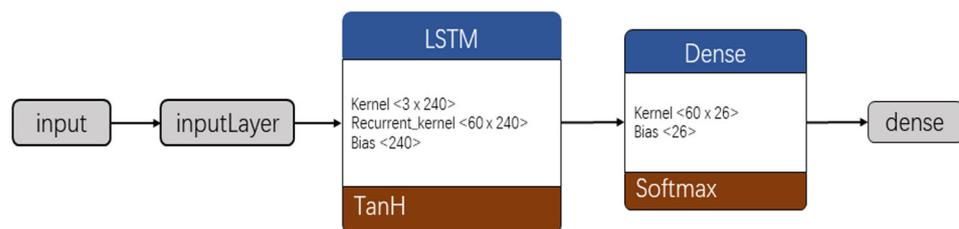


Figure 5. LSTM structure implemented for character recognition.

4.2. Convolutional Neural Network (CNN)

As mentioned in the previous section, CNN is a very popular neural network due to its excellent performance in image processing and human activity recognition.

For CNN’s input, the input datasets were rearranged to 60×1 , as indicated in the previous section. As shown in Figure 6, the model consists of six convolutional layers to extract relevant features from the input data. The first layer has 8 filters of dimensions 2×2 , the second and third layers both have 16 filters of dimensions 3×3 and the remaining convolutional layers have 16 filters of dimensions 3×3 . After these convolutional layers, a flattened layer and a fully connected layer with 16 neurons are included to classify the features into 26 lower-case letters. After every two convolutional layers, a “Max-Pooling” layer is applied to reduce the spatial dimensions of the input data. All the layers included in this network use the tanh activation function except for the last fully connected layer, which uses the SoftMax activation function for the final classification.

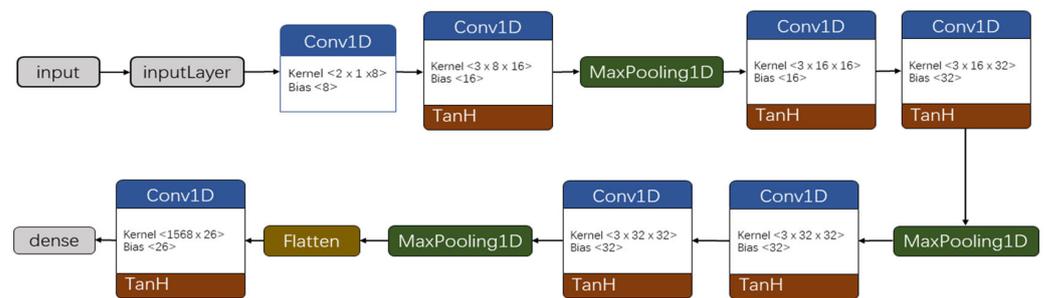


Figure 6. CNN structure implemented for character recognition.

4.3. BiLSTM

The BiLSTM is a special LSTM model. It is made up of two bi-directional LSTM layers in two directions, as shown in Figure 7. The first LSTM layer has 16 cells, with a time-step size of 3, and the return sequences are set to be True to return the full output sequence. The second LSTM layer uses 32 cells, with a time-step size of 3, followed by one fully connected layer with 26 neurons.

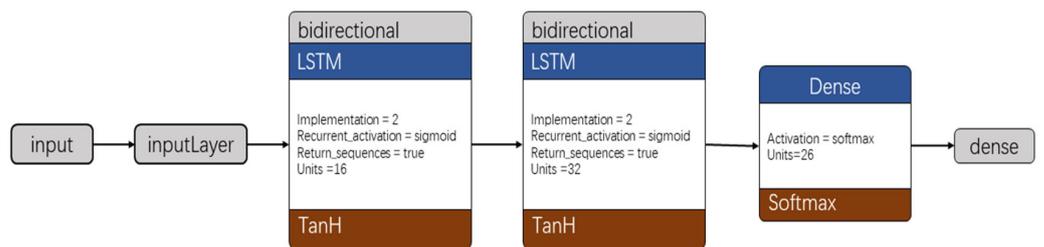


Figure 7. BiLSTM structure implemented for character recognition.

5. Experiment and Discussion

In this part of the paper, the training results of the three models will be discussed. On the computer side, all three models were compiled based on the categorical cross-entropy loss and a learning rate of 0.001, applying the Adam optimizer algorithm for about 200 epochs. The training was conducted using a computer with an NVIDIA RTX 3080 with 10 GB of GDDR6X memory, an 8704 CUDA core, an Intel Core i7-7700 HQ at 2.80 GHz and 32 GB of RAM. Python programming language was used for training and classification. All three deep learning models were trained with the same training, validation and testing dataset to obtain a fair training performance comparison. Then, STM X-Cube-AI expansion package is usable within the STM32CubeMX configuration tool. This package provides an automatic conversion of a pre-trained neural network and integration of generated optimized library into the user’s project.

On the MCU side, based on Keras library, we developed an optimized C code to run on MCU that can be used by X-CUBE-AI software to perform the automatic deployment on the STM32F401RE Nucleo board, which enables on-device inference on the STM32F401RE MCU.

Figures 8 and 9 show the three models' validation losses and validation accuracy, respectively. The validation dataset was an unknown dataset for the models. As a result, the validation loss graphs can also be used to observe training progress. In both figures, the LSTM performs poorly in the beginning and fluctuates greatly during the whole process. CNN achieves the best result at first and exhibits some ripples during its training, while BiLSTM maintains the most consistent performance throughout the entire procedure and reaches the best result in the end. This may mean that CNN is not good at extracting temporal features, while LSTM and BiLSTM are capable of keeping sequential information [16].

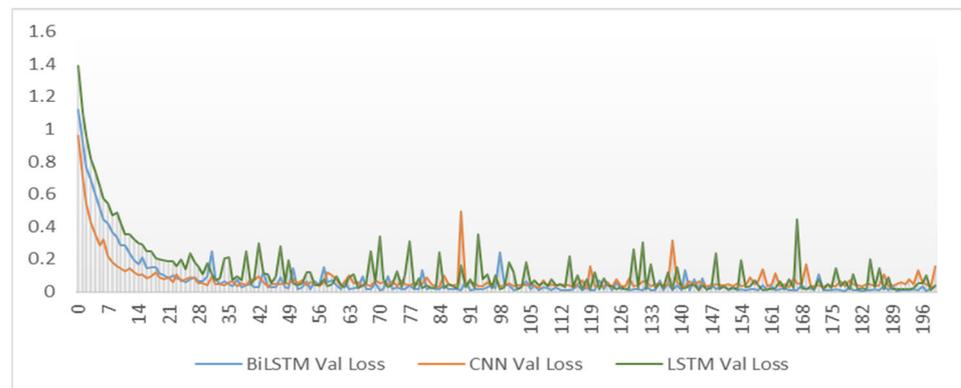


Figure 8. Loss of the validation datasets.

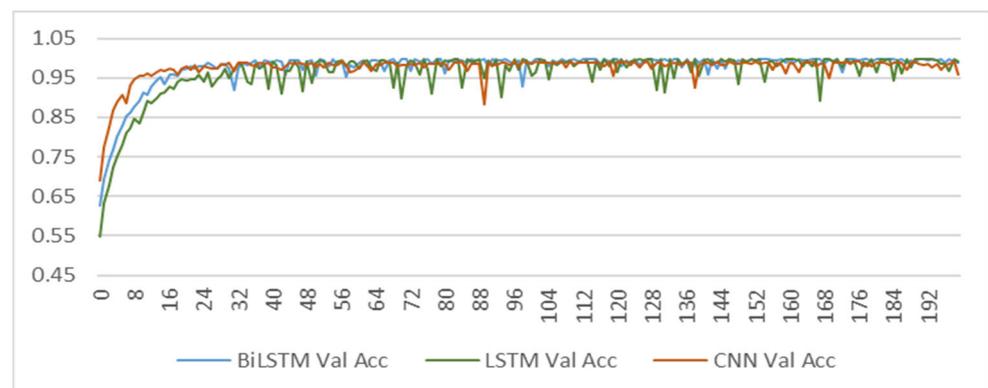


Figure 9. Accuracy graphs of the validation datasets.

The parameters of three classification algorithms for finger writing are presented in Table 2. These parameters primarily include the number of parameters, the latency and the accuracy of the models to indicate how well suited they are for the tasks as well as the complexity and size of the models, especially when operating on low-power edge devices. Table 2 shows that CNN, LSTM and BiLSTM all achieve satisfactory accuracy, above 95%. The BiLSTM model, with accuracy results of 99.82%, is the most accurate model. The LSTM and CNN achieved similar results, with a difference of 1.51% and 3.95% lower accuracy than BiLSTM, respectively. The CNN model achieved the lowest accuracy among the studied algorithms. To obtain a fair comparison result on latency, all latency measurements were carried out on the same device. The LSTM requires the lowest latency and has the fewest parameters. After considering parameters (accuracy, latency inference and resource consumption), LSTM still performs the best among the three.

Table 2. Comparison of the studied classification algorithms for finger-writing recognition.

Model	Number of Parameters	Maximum Latency (ms)	Accuracy (Test Data)	RAM (KB)	Flash (KB)
BiLSTM	20,890	71	99.82%	5.34	82.73
LSTM	16,946	50	98.31%	2.21	66.90
CNN	49,778	80	95.87%	7.90	194.45

In addition to the overall accuracy attained by the studied models, Figures 10–12 display the accuracy for each individual character. These figures show how the accuracy of the classes are well balanced in the studied algorithms, except for CNN, where the characters “c”, “n” and “v” achieved an accuracy under 90%, only 87.3%, 86.7% and 85.1%, respectively. At the same time, in Table 3, we can see that the three highest misclassification errors all come from CNN, and characters “c” and “a” occupy the first place, with a misclassification error of 8.9%. This may be explained by the fact that their trajectories of writing in the beginning are similar, as depicted in Figure 13. The second-highest misclassification error rate is contributed by “e” and “t”, which is 7.0%. The two letters do not appear to be related at all, but when “t” is written, a link connects its two portions (as seen in Figure 14), making “t” have curves similar to “e,” which can sometimes cause the algorithm to misclassify. The third position is taken by “v” and “u,” which are visually similar. In the rest of the studied algorithms, the error distribution among different characters does not indicate a clear misclassification across different classes, and errors are evenly distributed across all of them.

Table 3. The three pairs of characters with the highest misclassification rate.

Rank	Pairs	Rate	Classification Algorithm
1	a, c	8.9%	CNN
2	e, t	7.0%	CNN
3	v, u	6.3%	CNN

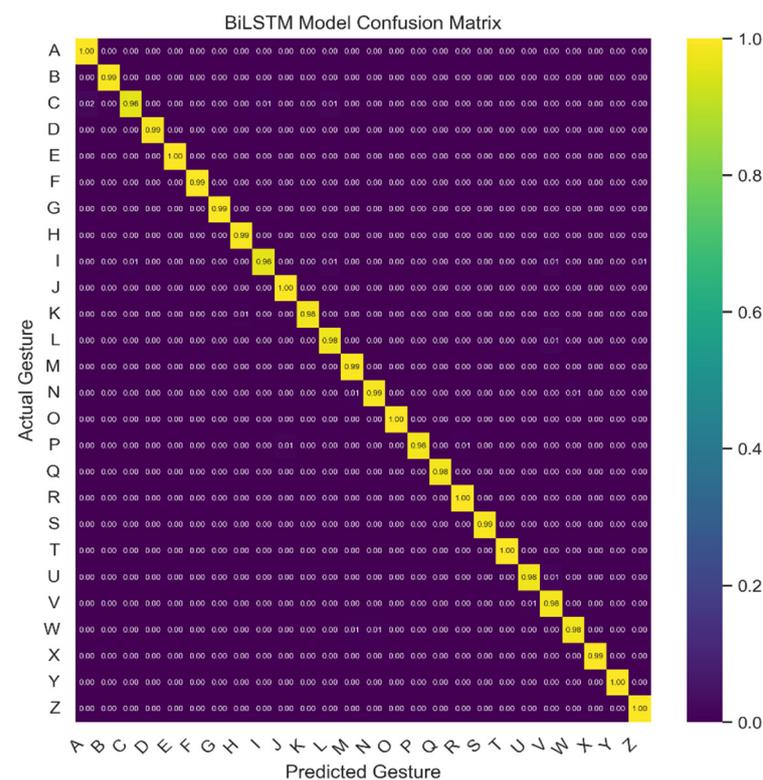


Figure 10. Confusion matrix generated using the BILSTM algorithm.

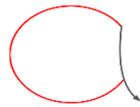


Figure 13. Similar trajectories of “a” and “c” in the beginning.



Figure 14. Link between the two portions of “t”.

In addition, to the best of our knowledge, at present, public datasets that have been acquired in a similar way to ours are not available. Hence, we could not find a relevant study to compare our study with.

6. Conclusions

A finger-writing character recognition system, based on an array of ToF sensors mounted on a low-power microcontroller STM32F401 equipped with deep learning algorithms, is presented in this work. The system aims to distinguish 26 English lower-case letters by users writing with their fingers in real time. Users are not required to wear additional devices but write on a flat surface (such as a desk). The ToF sensors are used to acquire pair distance values between each sensor to a writing finger within a 9.5×15 cm zone on a flat surface at specific time intervals.

The proposed method enables one to distinguish 26 English lower-case letters by users writing with their fingers and does not require one to wear additional devices. All data used in this work were collected from 21 subjects (12 males and 9 females) to evaluate the proposed system in a real scenario.

To test our system, a dataset containing 8190 finger-writing lower-case letter samples was recorded. Multiple algorithms were researched for this paper to study their performance in finger-writing recognition. It is shown in this paper how these algorithms provided high accuracy, whereby the best result was extracted from the LSTM, with 98.31% accuracy and 50 ms of maximum latency.

Despite the small benchmark datasets in this work, the study will serve as a foundation for further investigation into the automatic digitizing of handwritten characters, particularly from finger-writing motion. In the future, this approach will be expanded to include more symbols, such as writing commands and punctuation.

Author Contributions: Conceptualization, J.Z., G.P.; methodology, H.Y., and C.T.; validation, Y.T., and H.B.; writing—original draft preparation, H.B.; writing—review and editing, J.Z. and Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Opening Fund of Key Lab of Process Analysis and Control of Sichuan Universities of China (Grant number: 2018002).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nirjon, S.; Gummeson, J.; Gelb, D.; Kim, K.-H. Typingring: A wearable ring platform for text input. In Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services, Florence, Italy, 18–22 May 2015; pp. 227–239.

2. Chen, W.; Guan, M.; Huang, Y.; Wang, L.; Ruby, R.; Hu, W.; Wu, K. A low latency on-body typing system through single vibration sensor. *IEEE Trans. Mob. Comput.* **2019**, *19*, 2520–2532. [[CrossRef](#)]
3. Lakshmipathy, V.; Schmandt, C.; Marmasse, N. TalkBack: A conversational answering machine. In Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology, Vancouver, BC, Canada, 2–5 November 2003; pp. 41–50.
4. Muda, L.; Begam, M.; Elamvazuthi, I. Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. *arXiv preprint* **2010**, arXiv:1003.4083.
5. Kimura, N.; Kono, M.; Rekimoto, J. SottoVoce: An ultrasound imaging-based silent speech interaction using deep neural networks. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, 4–9 May 2019; pp. 1–11.
6. Juanpere, E.M.; Csapó, T.G. Ultrasound-based silent speech interface using convolutional and recurrent neural networks. *Acta Acust. United Acust.* **2019**, *105*, 587–590. [[CrossRef](#)]
7. Honarmandi Shandiz, A.; Tóth, L. Voice activity detection for ultrasound-based silent speech interfaces using convolutional neural networks. In Proceedings of the International Conference on Text, Speech, and Dialogue, Olomouc, Czech Republic, 6–9 September 2021; pp. 499–510.
8. Walker, J.W.; Clinnick, D.T.; Pedersen, J.B. Profiled hands in Palaeolithic art: The first universally recognized symbol of the human form. *World Art* **2018**, *8*, 1–19. [[CrossRef](#)]
9. Schrapel, M.; Stadler, M.-L.; Rohs, M. Pentelligence: Combining pen tip motion and writing sounds for handwritten digit recognition. In Proceedings of the 2018 CHI conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–11.
10. Zhang, H.; Chen, L.; Zhang, Y.; Hu, R.; He, C.; Tan, Y.; Zhang, J. A Wearable Real-Time Character Recognition System Based on Edge Computing-Enabled Deep Learning for Air-Writing. *J. Sens.* **2022**, *2022*, 1–12. [[CrossRef](#)]
11. Jing, L.; Dai, Z.; Zhou, Y. Wearable handwriting recognition with an inertial sensor on a finger nail. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; pp. 1330–1337.
12. Khan, N.A.; Khan, S.M.; Abdullah, M.; Kanji, S.J.; Iltifat, U. Use hand gesture to write in air recognize with computer vision. *IJCSNS* **2017**, *17*, 51.
13. Ogura, A.; Watanabe, H.; Sugimoto, M. Device-Free Handwritten Character Recognition Method Using Acoustic Signal. *J. Robot. Mechatron.* **2021**, *33*, 1082–1095. [[CrossRef](#)]
14. Lee, H.; Lee, Y.; Choi, H.; Lee, S. Digit Recognition in Air-Writing Using Single Millimeter-Wave Band Radar System. *IEEE Sens. J.* **2022**, *22*, 9387–9396. [[CrossRef](#)]
15. Molina, J.; Pajuelo, J.A.; Martínez, J.M. Real-time motion-based hand gestures recognition from time-of-flight video. *J. Signal Process. Syst.* **2017**, *86*, 17–25. [[CrossRef](#)]
16. Alemayoh, T.T.; Shintani, M.; Lee, J.H.; Okamoto, S. Deep-Learning-Based Character Recognition from Handwriting Motion Data Captured Using IMU and Force Sensors. *Sensors* **2022**, *22*, 7840. [[CrossRef](#)]
17. Shintani, M.; Lee, J.H.; Okamoto, S. Digital pen for handwritten alphabet recognition. In Proceedings of the 2021 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 10–12 January 2021; pp. 1–4.
18. Zhou, Z.; Tam, V.W.; Lam, E.Y. A Portable Sign Language Collection and Translation Platform with Smart Watches Using a BLSTM-Based Multi-Feature Framework. *Micromachines* **2022**, *13*, 333. [[CrossRef](#)]
19. Yanay, T.; Shmueli, E. Air-writing recognition using smart-bands. *Pervasive Mob. Comput.* **2020**, *66*, 101183. [[CrossRef](#)]
20. Sun, Y.; Fei, T.; Li, X.; Warnecke, A.; Warsitz, E.; Pohl, N. Real-time radar-based gesture detection and recognition built in an edge-computing platform. *IEEE Sens. J.* **2020**, *20*, 10706–10716. [[CrossRef](#)]
21. Gurbuz, S.Z.; Gurbuz, A.C.; Malaia, E.A.; Griffin, D.J.; Crawford, C.S.; Rahman, M.M.; Kurtoglu, E.; Aksu, R.; Macks, T.; Mdrafi, R. American sign language recognition using rf sensing. *IEEE Sens. J.* **2020**, *21*, 3763–3775. [[CrossRef](#)]
22. Breland, D.S.; Dayal, A.; Jha, A.; Yalavarthy, P.K.; Pandey, O.J.; Cenkeramaddi, L.R. Robust hand gestures recognition using a deep CNN and thermal images. *IEEE Sens. J.* **2021**, *21*, 26602–26614. [[CrossRef](#)]
23. Saez-Mingorance, B.; Mendez-Gomez, J.; Mauro, G.; Castillo-Morales, E.; Pegalajar-Cuellar, M.; Morales-Santos, D.P. Air-Writing Character Recognition with Ultrasonic Transceivers. *Sensors* **2021**, *21*, 6700. [[CrossRef](#)]
24. Boner, S.; Vogt, C.; Magno, M. Tiny TCN model for Gesture Recognition with Multi-point Low power ToF-Sensors. In Proceedings of the 2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS), Incheon, Republic of Korea, 13–15 June 2022; pp. 356–359.
25. Xing, T.; Yang, Q.; Jiang, Z.; Fu, X.; Wang, J.; Wu, C.Q.; Chen, X. WiFine: Real-time Gesture Recognition Using Wi-Fi with Edge Intelligence. *ACM Trans. Sens. Netw. TOSN* **2022**, *19*, 1–24. [[CrossRef](#)]
26. Rashid, N.; Demirel, B.U.; Al Faruque, M.A. AHAR: Adaptive CNN for energy-efficient human activity recognition in low-power edge devices. *IEEE Internet Things J.* **2022**, *9*, 13041–13051. [[CrossRef](#)]
27. Mendez, J.; Bierzynski, K.; Cuéllar, M.; Morales, D.P. Edge Intelligence: Concepts, architectures, applications and future directions. *ACM Trans. Embed. Comput. Syst. TECS* **2022**, *21*, 1–41. [[CrossRef](#)]
28. Zhao, Y.; Ren, X.; Lian, C.; Ma, R.; Zhang, X.; Sha, X.; Li, W.J. A Smart Wireless IoT Ring for Real-Time Keystroke Recognition Using Edge Computing. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–13. [[CrossRef](#)]

29. Yadav, S.; Pandey, A.; Aggarwal, P.; Garg, R.; Aggarwal, V. Handwriting Recognition using LSTM Networks. *Int. J. New Technol. Res.* **2018**, *4*, 263101.
30. Alam, M.S.; Kwon, K.-C.; Alam, M.A.; Abbass, M.Y.; Imtiaz, S.M.; Kim, N. Trajectory-based air-writing recognition using deep neural network and depth sensor. *Sensors* **2020**, *20*, 376. [[CrossRef](#)] [[PubMed](#)]
31. Gyongy, I.; Dutton, N.A.; Henderson, R.K. Direct time-of-flight single-photon imaging. *IEEE Trans. Electron. Devices* **2021**, *69*, 2794–2805. [[CrossRef](#)]
32. Choudhury, A.; Sarma, K.K. A CNN-LSTM based ensemble framework for in-air handwritten Assamese character recognition. *Multimed. Tools Appl.* **2021**, *80*, 35649–35684. [[CrossRef](#)]
33. Lee, B.-C.; Choi, B.-C.; Bang, H.-S.; Koh, Y.N.; Han, K.-Y. Study on Measurement Error Reduction using the Internal Interference Light Reduction Structure of a Time-of-Flight Sensor. *IEEE Sens. J.* **2022**, *22*, 12967–12975. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.