Handwritten Recognition via Index-Finger or Wrist Point Analysis using Mediapipe and EMNIST-Letters Database

Seyed Mojtaba Mirzadeh  
line 2: *dept. name of organization   
(of Affiliation)*Tehran, Iran  
msd.mirzadeh4@gmail.com

line 1: 2nd Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

line 1: 5th Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

line 1: 3rd Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

line 1: 6th Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

*Abstract*— Handwritten Recognition is a compelling branch within image processing, bearing numerous applications across mobile devices, tablets, personal computers, and interactive platforms. This technology aims to enhance human interaction with electronic devices efficiently [1]. Our paper introduces a novel approach to handwritten recognition, specifically analyzing index finger and wrist movements on a digital screen. These movements are systematically connected to form an image, subsequently identified using the EMNIST-letters image database.

Moreover, our approach contributes significantly by mitigating the inherent uncertainty associated with handwritten recognition. This uncertainty often arises from additional user movements, handwriting font, and size variations. Our technique effectively addresses these challenges, reducing tension and enhancing recognition accuracy.

We validate the effectiveness of our approach through real-world experiments, demonstrating its practical success and utility in various applications.

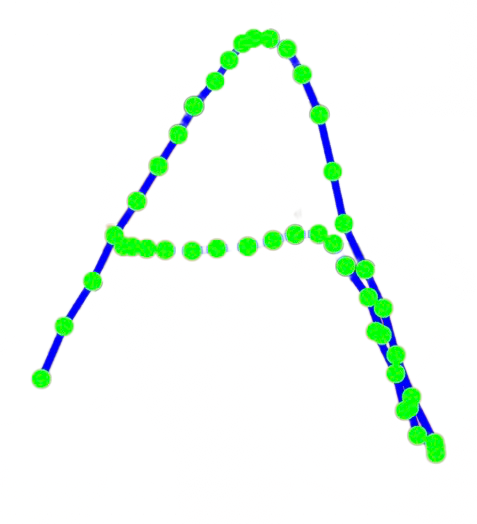
Keywords—Handwritten recognition, Index-finger analysis, Wrist point analysis, Mediapipe, EMNIST-Letters Database

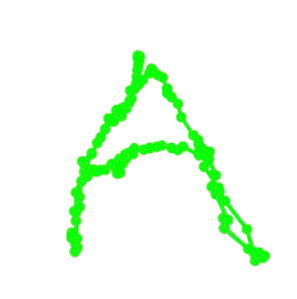
# Introduction (work on figure 1)

In an era where human-computer interaction continues to evolve, the capability to recognize and interpret handwritten input stands as a fundamental pillar of technological advancement. Handwritten recognition finds widespread applications in our daily lives, from mobile devices and tablets to personal computers and interactive platforms. However, it is not without its challenges. The intricacies of capturing the nuances of human handwriting, whether through stylus, touchscreen, or mouse input, have long been a focal point of research in image processing. This paper addresses a critical aspect of this research: reducing uncertainty in handwritten recognition. As we embark on this journey, we aim to revolutionize how we perceive and interact with digital handwriting.

Inherent uncertainties often plague the recognition of handwritten input. Users may vary the size and style of their writing, and additional movements can introduce complexities in accurately interpreting their intent. These challenges have motivated our investigation into a novel approach that leverages index-finger and wrist-point analysis to enhance the accuracy and efficiency of handwritten recognition [1, 2].

The primary purpose of this study is to propose a pioneering methodology for handwritten recognition, drawing on the capabilities of the Mediapipe framework in conjunction with the EMNIST-Letters Database. Our approach aims to improve recognition accuracy and streamline user interaction and electronic devices. We envision a future where digital handwriting is as intuitive and precise as its pen-and-paper counterpart. Our research holds significant promise in addressing the longstanding issues of uncertainty in handwritten recognition. We tackle the source of many recognition challenges by delving into the index finger and wrist movements analysis. Furthermore, incorporating the Mediapipe framework and the EMNIST-Letters Database introduces innovation at the intersection of computer vision and handwriting analysis, promising breakthroughs in multiple application domains. Figure 1.





As we chart this course, we draw inspiration from an array of related works in the fields of image processing, machine learning, and handwritten recognition. Notable contributions have come from a variety of researchers, each building on foundational concepts and techniques. Lecun et al. introduced the notion of gradient-based learning applied to document recognition [3], while Graves et al. presented a connectionist system for unconstrained handwriting recognition [4]. Recent advancements, such as the hybrid CNN-BLSTM network by Namdeo et al. [5], have further propelled the field forward. Our research builds upon these foundations, striving to reduce uncertainty in handwritten recognition by exploring novel avenues.

The remainder of this paper is organized as follows. In the subsequent section, we delve into a comprehensive literature review surrounding handwritten recognition, drawing insights from seminal works. We present the methodology, detailing our index-finger and wrist point analysis approach. Then, we showcase experimental results that validate the effectiveness of our technique. Finally, we conclude with a discussion of our findings and the broader implications of our research. With this paper, we invite readers to embark on a journey into the future of handwritten recognition, which merges technology seamlessly with human expression and promises to reshape our interactions with electronic devices.

# Related works

The field of handwritten recognition has witnessed significant advancements driven by the ever-growing demand for natural and intuitive human-computer interaction. This section reviews existing research in handwritten recognition techniques, gesture recognition, and database utilization, setting the stage for our novel approach to enhance handwritten recognition through index finger and wrist point analysis [1, 2].

Numerous studies have delved into handwritten recognition, exploring diverse techniques and algorithms. Convolutional Neural Networks (CNNs) have gained prominence for their ability to capture intricate patterns in handwritten text [2, 3]. Recurrent Neural Networks (RNNs) have also been employed to model the temporal aspects of handwriting [4]. Recent research by Namdeo et al. [5] introduced a hybrid approach that combines CNNs and RNNs, yielding impressive recognition accuracy.

Gesture recognition, closely related to our work, has been extensively explored. Researchers have developed methods for recognizing hand gestures and finger movements using depth sensors [6]. Fu et al. [7] proposed a real-time hand gesture recognition system that analyzes finger and wrist movements. These studies emphasize the relevance of gesture recognition techniques to our index finger and wrist point analysis approach.

Practical training and testing of handwritten recognition models heavily rely on comprehensive databases. The EMNIST-Letters Database [8] has gained traction for its extensive collection of handwritten letters in various fonts and sizes. However, some studies have identified limitations in existing databases, such as a lack of diverse handwriting styles [9]. These shortcomings highlight the need for innovative approaches, like ours, to address the challenges of various handwriting. The literature review underscores a significant knowledge gap in the field: the absence of a comprehensive approach that harnesses the combined power of index-finger and wrist point analysis, the Mediapipe framework [10, 11], and the EMNIST-Letters Database.

Our paper aims to bridge this gap by presenting a novel methodology that addresses the limitations of existing techniques and databases. With a solid understanding of the existing research landscape, we describe our proposed method, which leverages index-finger and wrist point analysis, the Mediapipe framework, and the EMNIST-Letters Database to enhance handwritten recognition.

# OUR APPROACH (Should add some images)

This section presents our approach to enhancing handwritten recognition through two key components: "Wrist Point Analysis" and "Index-Finger Point Analysis." Each part addresses specific challenges and contributes to improving recognition accuracy.

## Wrist Point Analysis

### Real-Time Wrist Point Capture

Our approach begins with the real-time capture of wrist movements during the handwriting process using the Mediapipe framework. We track the position and orientation of the wrist in a 3D space, providing valuable insights into handwriting dynamics.

### Connecting Wrist Movements

To recognize the drawn letter, we capture and connect the points of wrist movement. This approach initially aimed to streamline the recognition process. However, as we discuss in the next section, we encountered challenges related to accuracy and point alignment.

## Index-Finger Point Analysis

### Precise Index-Finger Tracking

Recognizing the limitations of wrist point analysis, we shifted our focus to index-finger point analysis. We improved our understanding of the Mediapipe framework to track the index finger more accurately. This adjustment proved to be crucial in capturing precise data for handwritten recognition.

1. *Line Drawing and Word Prediction*

With index-finger tracking, we can efficiently capture the points and draw lines between them to create the handwritten word. This approach allows for more accurate recognition of characters based on the drawn lines, significantly improving recognition results.

During our experimentation, we observed that the thickness of the drawn lines between points significantly impacted recognition accuracy. We tested the optimal line thickness, balancing improvements and potential issues arising from excessively thick lines. We conducted extensive experiments to validate our approach's effectiveness. We tested various handwriting styles and sizes, comparing our system against baseline methods. This section presents the experimental results and discusses the challenges encountered during testing. Incorporating both "Wrist Point Analysis" and "Index-Finger Point Analysis," our approach represents a comprehensive solution to enhance handwritten recognition. By leveraging the capabilities of the Mediapipe framework and the EMNIST-Letters Database, we address the challenges of accuracy and uncertainty, ultimately improving recognition accuracy and user interaction.

# Mediapipe(need edit and review and add refrences)

Mediapipe, a versatile framework designed for building pipelines to perform inference over sensory data, has emerged as a pivotal tool in the realm of computer vision and machine learning. This framework allows practitioners, including researchers, students, and software developers, to construct perception pipelines as interconnected graphs of modular components. These components encompass model inference, media processing algorithms, data transformations, and more, forming a robust ecosystem for processing diverse sensory data types. An example is shown in Figure 1.

In the context of our research, which focuses on enhancing handwritten recognition through index-finger and wrist point analysis, we leverage the capabilities of Mediapipe to enable real-time hand movement tracking. Our perception pipeline captures and processes critical information, such as index-finger and wrist movements, with a level of precision that enhances the recognition of handwritten characters.

Mediapipe empowers us to abstract and connect individual perception models into maintainable pipelines, eliminating excessive coupling between steps. This abstraction simplifies the development process and allows for the rapid prototyping of perception pipelines, which is invaluable for researchers and developers striving to create production-ready ML applications and technology prototypes. Additionally, by providing a rich configuration language and evaluation tools, Mediapipe facilitates the deployment of perception technology across a wide range of hardware platforms.

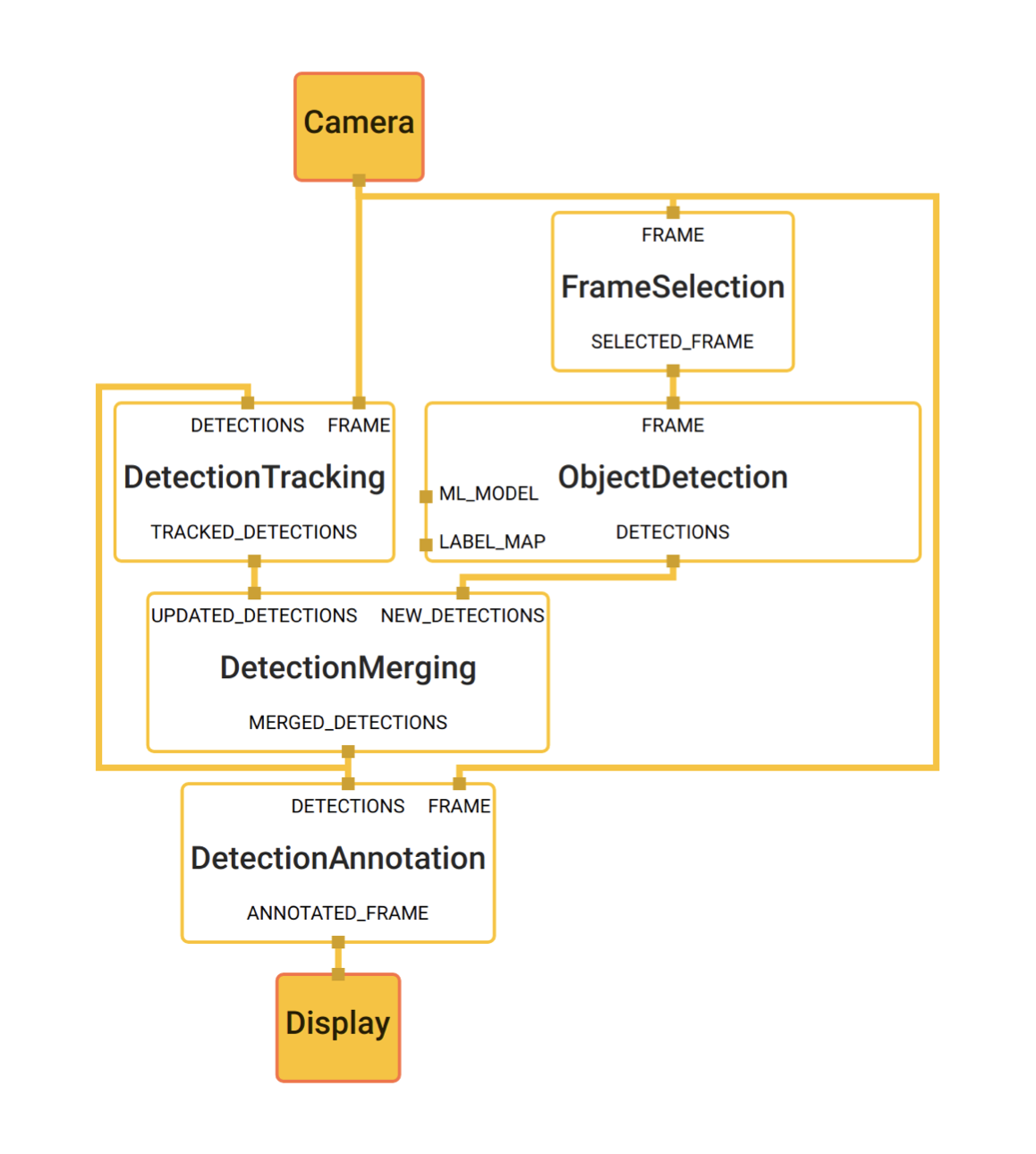


Figure 1(need edit): Object detection using MediaPipe. The trans- parent boxes represent computation nodes (calculators) in a MediaPipe graph, solid boxes represent external in- put/output to the graph, and the lines entering the top and exiting the bottom of the nodes represent the input and out- put streams respectively. The ports on the left of some nodes denote the input side packets. See Section 6.1 for details.

# EMNIST-Letters Database (add figures and refrences)

The EMNIST (Extended Modified National Institute of Standards and Technology) dataset family offers a rich resource for character recognition, and among its variants, the EMNIST-Letters dataset stands as a cornerstone in our research. This dataset presents a unique opportunity to train and evaluate our handwritten recognition model, consolidating a balanced set of uppercase and lowercase letters into a unified 26-class classification task.

The EMNIST Balanced dataset is meticulously curated to provide an equal number of samples per class, enhancing its utility for various recognition tasks. The EMNIST Letters dataset builds upon this balance by amalgamating the uppercase and lowercase classes, thus forming a harmonized classification framework representing all 26 letters of the English alphabet.

In contrast to its predecessors, such as the EMNIST Digits and EMNIST MNIST datasets, which cater to balanced handwritten digit recognition tasks, the EMNIST Letters dataset tackles the complexities of letter recognition. This consolidation minimizes the challenges associated with case confusion, further simplifying the classification task.

Our research leverages the EMNIST 'Letters' dataset to train and evaluate our handwritten recognition model. In our accompanying model, the dataset played a pivotal role, contributing to the impressive performance achieved. The model, trained with dedication and precision, reached a training accuracy of 95% and demonstrated robustness with a test accuracy of 91%.

To facilitate the classification task, we assign labels to each character in accordance with their English alphabet order. For reference, 'A' is labeled as 1, 'B' as 2, 'C' as 3, and so on, up to 'Z' labeled as 26. This mapping allows for consistent and intuitive character recognition.

# IMPLEMENTATION AND RESULTS

The technique presented in this paper is implemented in the Visual Studio Code Jupyter Notebook programming environment, Python programming language, and the recognized percentage of each letter in 100 tests (for each letter) is given in Figure ... . As can be seen, the most difficult letters to recognize are Q and O due to the similarity of writing, where they are both realized at the rate of 45%. In contrast, with 97% accuracy, the letter V is the easiest to recognize with our system. The average rate of recognizing English alphabet letters is 74.19%. The implementation of recognizing the letters “A” and “N” that form the “AN” word is shown in Figures ... and ... .

# Conclusion and future works

According to the results obtained, because our proposed method performs handwriting recognition operations in real-time and with good speed, it has an acceptable efficiency. This method is currently only available for capital English letters and can be extended to lowercase letters, numbers, and other languages. In the proposed form, we provide a system (Mediapipe Library) to recognize the Wrist and Index Finger movement points and minimize the uncertainty in the problem using an MNIST-Letter dataset. Also, solutions are proposed to reduce the error of incorrect identification, increasing the existing technique's accuracy. The implementation of the proposed approach leads to achieving more favorable results. Therefore, in the proposed method, it is possible to quickly implement recognition systems of the handwriting letters based on a hand point such as Index finger or wrist movement. The problem of hand movement recognition is essential because of its high utilization and numerous academic and executive institutions support it. Therefore, there are many areas of development. Among them, it is possible to point to the development of letter recognition with less error rate and the glory of small letters and letters of other languages.

##### References

1. M. Zare, M. Jampour, A. S. Arezoomand and M. Sabouri, "Handwritten Recognition based on Hand Gesture Recognition using Deterministic Finite Automata and Fuzzy Logic," *2019 4th* International Conference on Pattern Recognition and Image Analysis (IPRIA), Tehran, Iran, 2019, pp. 93-99, doi: 10.1109/PRIA.2019.8786038.
2. J. Y. Chang, G. Moon and K. M. Lee, "V2V-PoseNet: Voxel-to-Voxel Prediction Network for Accurate 3D Hand and Human Pose Estimation from a Single Depth Map," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 5079-5088, doi: 10.1109/CVPR.2018.00533.
3. Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
4. A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke and J. Schmidhuber, "A Novel Connectionist System for Unconstrained Handwriting Recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, pp. 855-868, May 2009, doi: 10.1109/TPAMI.2008.137.
5. R. K. Namdeo, C. Gupta and R. Shrivastava, "Offline Handwritten Text Recognition Using Hybrid CNN–BLSTM Network," *2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT)*, Indore, India, 2022, pp. 318-323, doi: 10.1109/CSNT54456.2022.9787587.
6. G. Lu, Z. Wang and Q. Zhu, "Real-Time 3D Gesture Interaction from Single RGB Images," *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, Chongqing, China, 2022, pp. 1197-1201, doi: 10.1109/ITAIC54216.2022.9836960.
7. J. Fu, L. Xiong, X. Song, Z. Yan and Y. Xie, "Identification of finger movements from forearm surface EMG using an augmented probabilistic neural network," *2017 IEEE/SICE International Symposium on System Integration (SII)*, Taipei, Taiwan, 2017, pp. 547-552, doi: 10.1109/SII.2017.8279278.
8. Cohen, G., Afshar, S., & Tapson, J. (2017). EMNIST: An extension of MNIST to handwritten letters. arXiv preprint arXiv:1702.05373.
9. A. Baldominos, Y. Saez, and P. Isasi, “A Survey of Handwritten Character Recognition with MNIST and EMNIST,” Applied Sciences, vol. 9, no. 15, p. 3169, Aug. 2019, doi: 10.3390/app9153169.
10. <https://doi.org/10.48550/arXiv.1906.08172>
11. https://doi.org/10.48550/arXiv.2006.10214