Dean&Francis

ISSN 2959-6157

Hand-Written Number Classification by Hardware Neural Network

Jiaxiang Wang

School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Soochow, China

Jiaxiang.Wang21@student.xjtlu.edu.cn

Abstract:

This study explores the application of hardware neural networks for handwritten digit recognition, focusing on the implementation of the perceptron model. Handwritten digit recognition is a key technology in fields such as automatic postal sorting and bank cheque processing. In this work, a hardware neural network based on Frank Rosenblatt's perceptron model is employed to efficiently process and classify handwritten digits. The system utilizes Verilog and VLSI technology to handle inputs in a parallel and pipelined fashion, greatly reducing computation time compared to traditional serial methods. Specifically, the designed system can recognize all ten digits in just 392 clock cycles, demonstrating significant improvements in efficiency while maintaining accuracy. Parallel processing and pipelining techniques are used to optimize performance, and experimental results show accurate identification of the test images. This research highlights the potential of hardware-based neural networks in applications such as embedded systems and IoT, emphasizing the role of hardware acceleration in advancing artificial intelligence technologies.

Keywords: Hardware acceleration; Pipeline; VLSI; Verilog; Hand-Written Number Classification.

1. Introduction

Handwritten digit recognition has become an essential tool in various industries, including automatic postal sorting and banking systems. The technology plays a vital role in converting handwritten numerical data into digital form, aiding in automating numerous tasks. As information exchange continues to grow in the digital age, handwritten digits remain a widely used form of communication, emphasizing the need for efficient recognition systems. Handwritten digit recognition, therefore, has emerged as a critical area of research in artificial intelligence and machine learning [1]. Over time, several approaches have been developed to enhance the performance of handwritten digit recognition systems. Traditional software-based neural networks have shown great promise in achieving high accuracy. However, the shift toward hardware-based solutions has garnered interest due to the potential for faster processing and reduced latency. Neural networks, particularly the perceptron model developed by Frank Rosenblatt, have been widely used for classification tasks, in-

cluding handwritten digit recognition. Advances in VLSI technology and hardware description languages such as Verilog have made it possible to implement efficient hardware neural networks capable of handling complex computations in parallel and pipelined architectures [2].

This paper presents the design and implementation of a hardware neural network system based on the perceptron model to achieve high-speed handwritten digit recognition. The system processes 28x28 pixel images and utilizes parallel processing and pipelining techniques to significantly reduce clock cycles and improve computation speed. Through Verilog and VLSI technology, the system is capable of recognizing all ten digits in only 392 clock cycles, demonstrating notable efficiency improvements. Additionally, this study explores the potential of hardware neural networks in embedded systems and Internet of Things (IoT) applications, suggesting future directions for optimizing the design and performance of such systems [3].

2. Relavant Theory

Handwritten number recognition technology refers to the training of learning models, is the machine will be handwritten numbers (0-9) will be correctly classified into the corresponding numbers. The implementation of such systems can be divided into four modules. handwritten numbers 2. image acquisition 3. image pre-processing 4. number recognition Among them, in the module of digital recognition, with the development of the times, different algorithms are constantly found to be applied to this technology [3]. The most common ones are. Deep Convolutional Self-Coding Neural Networks:A convolutional neural network is combined with an autoencoder to form a deep convolutional self-coding neural network, and the parameters of the convolutional kernel are set according to the characteristics of the handwritten digital image to perform feature extraction. The extracted features not only match the output features of the autoencoder, but also have the advantages of unsupervised and fast feature extraction. K-Nearest Neighbors (KNN) Algorithm:A frequently utilised machine learning algorithm for classification and regression tasks. The fundamental premise is that a given sample can be classified as belonging to a specific category if the majority of its nearest neighbours in the feature space also belong to that same category [4].

Perceptron model: One of the earliest artificial neural network models, the single-layer perceptron, was proposed by Frank Rosenblatt in 1957. It is widely regarded as the foundational work on neural networks. The single-layer perceptron was originally designed to solve binary classification problems, that is, to separate samples of different classes. Its structure is relatively simple, comprising a single output node and several input nodes. By performing linear weighting and thresholding operations on the input signals, the single-layer perceptron is able to produce classification results. algorithm principle: In this experiment, we utilise the perceptron model algorithm to implement the function of digital handwriting recognition. In this experiment, a 2828 pixel image is selected as the object, and the objective is for the designed element to recognise the digits within it. The perceptron model recognises the input, calculates its weighted sum, and passes the result to a linear function. The structural flow of the algorithm is illustrated in Fig 1. In this experiment, the input layer is an image comprising 28 by 28 pixels. The intensity of each pixel is represented by a greyscale value between 0 and 255, which are then multiplied by pre-designed values, referred to here as weights. Each neuron receives input from the pixel and generates a single output. The resulting value is therefore a numerical probability. The predicted number is the most probable outcome. The sequence of operations in this experiment is illustrated in Fig 2.

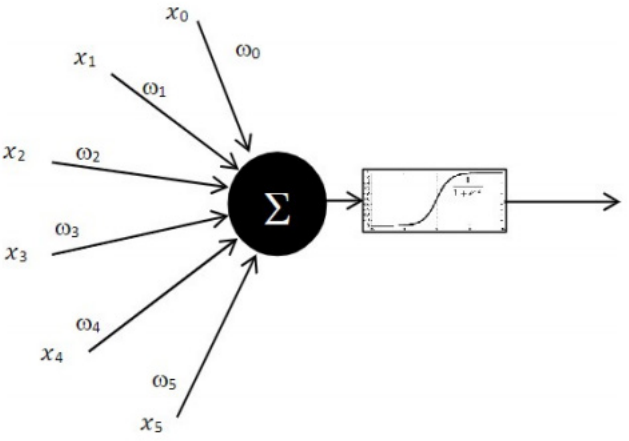

Fig. 1 Algorithm Structure Flowchart (Photo credit: Original).

ISSN 2959-6157

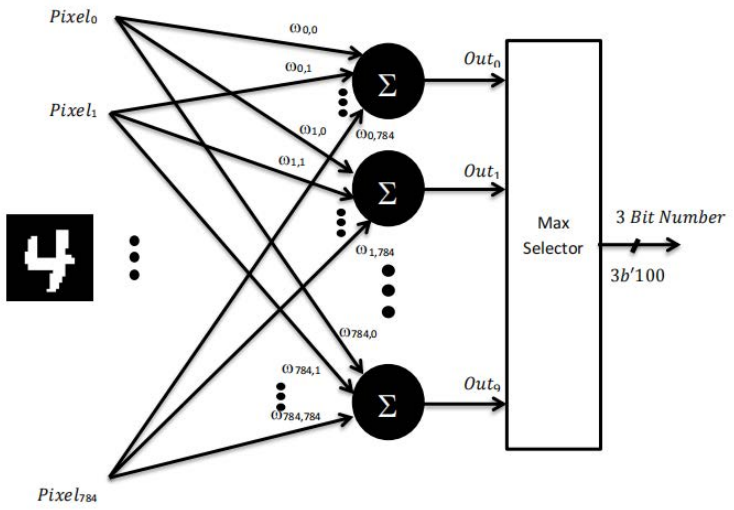

Fig. 2 Algorithm Structure Flowchart In This Experiment (Photo credit: Original).

3. Optimizing Handwritten Digit Recognition

To illustrate, in this study, one of the output values is designated as Out.

$Out_0 Out_1 Out_2 Out_3 Out_4 Out_5 Out_6 Out_7 Out_8 Out_9$] =

[0.1,0.08,0.004,0.022,0.87,0.034,0.134,0.007,0.0017,0.03] (1)

where the number '4' is larger. The corresponding output value of '4' is larger, thus indicating that the original figure should be represented by the number '4'. Experimental Design: In this study, two key techniques were explored: parallel processing and pipelining. Parallel Processing: Each pixel's grayscale value is stored in an adder, with 764 adders required to accommodate the 28x28 pixel grid. Additionally, since each pixel has a unique weight, 784 multipliers are needed to compute the final weighted sum. While parallel processing minimizes clock delay, the circuit complexity and required hardware resources present significant challenges, which diverge from the primary goal of reducing costs. As a result, alternative approaches must be considered to strike a balance between performance and hardware efficiency.Pipelining: To accelerate the recognition process, the number of multipliers and adders working in parallel must be increased, while also adhering to cost constraints. In this experiment, the weights of the 28 numbers across the 784 pixels were first calculated. The 28 cycles were then processed in a pipelined serial manner [5-7]. Assuming that each adder requires two clock cycles and each multiplier six cycles, the total number of cycles required for the entire process is reduced to 1736 (calculated as 228 + 6 = 1736). This pipelined approach is demonstrably more efficient than traditional serial processing, significantly reducing the overall computation time.Experimental Results: The experimental results are depicted in Fig 3.

```
------------Starting to Test Image Number 1 -------------------------
Actual Number is 0001
              795 clks of processing done
Estimatede Output Number is : 0001
Correct Output!! The Classifier has sucessfully classified this image
----------------------------------------------------------------------
------------Starting to Test Image Number 2 -------------------------
Actual Number is 0010
              795 clks of processing done
Estimatede Output Number is : 0010
Correct Output!! The Classifier has sucessfully classified this image
----------------------------------------------------------------------
------------Starting to Test Image Number 3 -------------------------
Actual Number is 0111
              795 clks of processing done
Estimatede Output Number is : 0111
Correct Output!! The Classifier has sucessfully classified this image
----------------------------------------------------------------------
```

Fig. 3 Experimental Results (Photo credit: Original).

The experimental results indicate that the numerical data in the three test images were accurately identified by the system. This demonstrates the effectiveness of the hardware neural network design, showcasing the ability of the perceptron model to process handwritten digits with high precision. The system's success in accurately classifying the test images supports the assertion that this hardware-based approach offers a viable alternative to traditional software-driven methods for handwritten digit recognition, particularly in applications where speed and efficiency are critical [8-10]. Experiment Improvement: Building upon the initial design, improvements can be made at the neuronal level by incorporating parallel processing to optimize the adder operation. In the original design, a sequential adder, depicted in Fig 4, was used to perform the necessary calculations. However, this method, while functional, introduced delays due to the sequential nature of the operations, which limited the overall speed of the system. To address this issue and enhance performance, the sequential adder was replaced with a more advanced 4-stage adder, as shown in Fig 5. The 4-stage adder introduces a pipeline structure that allows for simultaneous processing of multiple inputs, thereby significantly reducing the number of clock cycles required for computation. In the original design, the sequential adder required a substantial number of cycles to complete its operations. With the new design, the system requires only \((2 \times 4 + 6) \times 28 = 392 \) clock cycles to process the data, which represents a significant improvement in computational efficiency. This reduction in clock cycles leads to faster overall performance without a corresponding increase in hardware complexity or cost. The introduction of this 4-stage adder also maintains the system's low-cost objectives. By balancing the number of adders and multipliers used in the pipeline, the design achieves a high level of arithmetic efficiency while keeping the cost and resource requirements within acceptable limits. This is particularly important for real-world applications, where both speed and cost are critical factors in system design. In summary, the replacement of the sequential adder with a 4-stage adder not only improves the speed and efficiency of the system but also demonstrates the potential for further optimization of hardware-based neural networks. This improvement paves the way for future advancements in the field of hardware acceleration for machine learning tasks, particularly in the context of embedded systems and IoT applications. The ability to process handwritten digit recognition tasks in a fraction of the time required by traditional methods highlights the advantages of hardware neural networks in terms of speed, efficiency, and scalability.

Fig. 4 Sequential Adder (Photo credit: Original).

ISSN 2959-6157

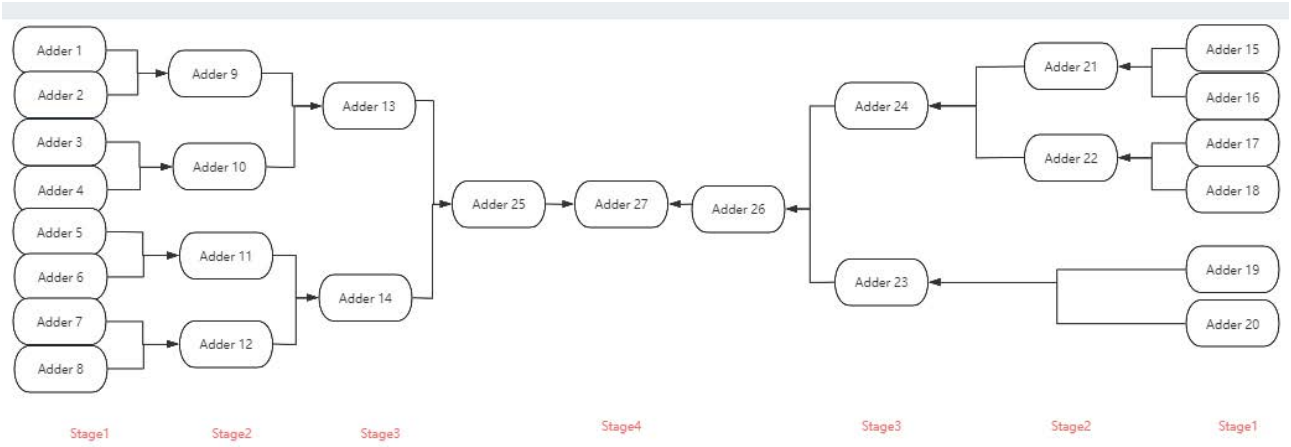

Fig. 5 4-stage adder (Photo credit: Original).

4. Conclusion

This study successfully demonstrates the implementation of a hardware neural network system for handwritten digit recognition using a perceptron model. By employing Verilog and VLSI technologies, the system was designed to process handwritten digits in a highly efficient manner, reducing the overall computation time significantly. The key advancements include the use of parallel processing and pipelining, which allowed the system to recognize all ten digits in only 392 clock cycles, a marked improvement over traditional serial methods. The experimental results confirmed that the system was able to accurately classify the test images, further validating the effectiveness of hardware-based neural networks in this domain. This research highlights the potential of hardware neural networks for real-time applications in fields like embedded systems and IoT, where both speed and accuracy are paramount.While this study focused on the perceptron model and its application in handwritten digit recognition, there are numerous opportunities for further research and improvement. Future studies could explore the incorporation of more advanced neural network architectures, such as multi-layer perceptrons or convolutional neural networks, into hardware systems to handle more complex recognition tasks. Additionally, further optimization of the hardware design, including the reduction of power consumption and the use of more efficient components, could enhance the applicability of hardware neural networks in a wider range of real-world scenarios. As the demand for faster, more efficient AI solutions continues to grow, the role of hardware acceleration in artificial intelligence is likely to expand, making this an exciting area for future exploration and development.

References

[1] Giardino, D., Matta, M., Silvestri, F., Spanò, S., & Trobiani, V. (2019). FPGA implementation of hand-written number recognition based on CNN. International Journal on Advanced Science, Engineering and Information Technology, 9(1), 167-171.

[2] Denker, J., Gardner, W., Graf, H., Henderson, D., Howard, R., Hubbard, W., ... & Guyon, I. (1988). Neural network recognizer for hand-written zip code digits. Advances in neural information processing systems, 1.

[3] Jagtap, V. N., & Mishra, S. K. (2014). Fast efficient artificial neural network for handwritten digit recognition. International Journal of Computer Science and Information Technologies, 5(2), 2302-2306.

[4] Thomas, W., & Singh, V. (2024). Accelerating Neural Network Inference in Handwritten Digit Recognition—Comparative Study. Parallel Processing Letters, 2430001.

[5] Zhu, X., Zhang, Y., Zhao, Z., & Zuo, J. (2019, July). Radio frequency sensing based environmental monitoring technology. In Fourth International Workshop on Pattern Recognition (Vol. 11198, pp. 187-191). SPIE. https://doi.org/10.1117/12.2536905

[6] Zhang, Y., Xu, H., Zhu, X., & Zuo, J. (2020). Detection and quantization technique of optical distributed acoustic coupling based on φ-OTDR. Journal of Shanghai Jiaotong University (Science), 25, 208-213. https://doi.org/10.1007/s12204-019-2128-5

[7] Zhao, Z., Peng, Y., Zhu, X., Wei, X., Wang, X., & Zuo, J. (2020). Research on prediction of electricity consumption in smart parks based on multiple linear regression. In 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC) (pp. 812-816). IEEE. https://doi.org/10.1109/ITAIC49862.2020.9339070

[8] Xu, H., Zhu, X., Zhao, Z., Wei, X., Wang, X., & Zuo, J. (2020). Research of pipeline leak detection technology and

application prospect of petrochemical wharf. In 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC) (pp. 263-271). IEEE.

[9] Zhu, X., Zhao, Z., Wei, X., Wang, X., & Zuo, J. (2021, February). Action recognition method based on wavelet transform and neural network in wireless network. In Proceedings of the 2021 5th International Conference on Digital Signal Processing (pp. 60-65). ACM.

[10] Wang, R., Zhu, J., Wang, S., Wang, T., Huang, J., & Zhu, X. (2024). Multi-modal emotion recognition using tensor decomposition fusion and self-supervised multi-tasking. International Journal of Multimedia Information Retrieval, 13(4), 39-51. https://doi.org/10.1007/s13735-023-00212-3