Handwritten digit recognition using quantum convolution neural network

Ravuri Daniel¹, Bode Prasad², Prudhvi Kiran Pasam³, Dorababu Sudarsa⁴, Ambarapu Sudhakar⁴, Bodapati Venkata Rajanna⁵

¹Department of Computer Science and Engineering, Prasad V. Potluri Siddhartha Institute of Technology, Vijayawada, India
²Department of Information Technology, Vignan’s Institute of Information Technology (A), Visakhapatnam, India
³Department of Computer Science and Engineering (IoT), R.V.R. & J. C. College of Engineering (A), Guntur, India
⁴Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, India
⁵Department of Electrical and Electronics Engineering, MLR Institute of Technology, Hyderabad, India

Article Info

ABSTRACT

The recognition of handwritten digits holds a significant place in the field of information processing. Recognizing such characters accurately from images is a complex task because of the vast differences in people’s writing styles. Furthermore, the presence of various image artifacts such as blurring, intensity variations, and noise adds to the complexity of this process. The existing algorithm, convolution neural network (CNN) is one of the prominent algorithms in deep learning to handle the above problems. But there is a difficulty in handling input data that differs significantly from the training data, leading to decreased accuracy and performance. In this work, a method is proposed to overcome the aforementioned limitations by incorporating a quantum convolutional neural network algorithm (QCNN). QCNN is capable of performing more complex operations than classical CNNs. It can achieve higher levels of accuracy than classical CNNs, especially when working with noisy or incomplete data. It has the potential to scale more efficiently and effectively than classical CNNs, making them better suited for large-scale applications. The effectiveness of the proposed model is demonstrated on the modified national institute of standards and technology (MNIST) dataset and achieved an average accuracy of 91.08%.

Keywords:
Convolutional neural network
Hand digit recognition
Modified national institute of standards and technology (MNIST) dataset
Quantum computing
Quantum convolutional neural network

Corresponding Author:
Ravuri Daniel
Department of Computer Science and Engineering, Prasad V. Potluri Siddhartha Institute of Technology
Vijayawada, India
Email: danielravuri@gmail.com

1. INTRODUCTION

Handwriting is considered the most conventional and structured way of documenting facts and information. Individuals have unique and idiosyncratic handwriting. A system that is able to recognize and analyze human handwriting in any language is referred to as a handwritten character recognition (HCR) system [1]–[3]. Handwriting recognition can be carried out from both online and offline sources. In recent times, the application of handwriting recognition has become increasingly prevalent and is now used in various domains, including but not limited to, reading postal addresses, language translation, bank forms and check amounts, digital libraries, keyword spotting, and traffic sign detection.

Recognizing human handwritten digits is a challenging task for computer applications since they come in various shapes and sizes, making them imperfect. Handwritten digit recognition refers to a computer’s ability to identify and classify human-written numbers from various sources, including images, papers, and touch...
screens, into ten predefined classes (0-9). Handwritten number recognition poses several challenges due to the varying styles of writing among individuals, making it different from optical character recognition. The handwritten digit recognition system utilizes digit images to overcome this challenge and recognize the digit present in the image. To achieve this, the present research work based on Support Vector Machine, Multilayer Perceptron, convolutional neural network (CNN), and other deep learning methods [4], [5] are applied over Modified National Institute of Standards and Technology (MNIST) dataset to recognize handwritten digits. But CNN [6], [7] require a large amount of labeled data to train, obtaining such datasets can be costly and time-consuming. It can also be prone to overfitting, which means that the model performs well on the training data but poorly on new, unseen data. In certain real-world applications, the usefulness of CNNs may be limited due to their struggle to generalize to new, unseen data beyond the scope of the training data. In this paper, a new method is proposed to overcome the aforementioned limitations by incorporating a quantum convolutional neural network algorithm (QCNN) [8], [9]. Because the QCNN can perform certain operations much faster than classical convolutional neural networks, especially those involving matrix multiplication and Fourier transforms. It can require fewer resources and can be more efficient than classical CNNs, especially when dealing with large datasets, working with noisy, and incomplete data. QCNNs have the potential to scale more efficiently and effectively than classical algorithms, making them better suited for large-scale applications. The effectiveness of the proposed method, conducted tests on the MNIST dataset and achieved an average accuracy of 91.08%.

This paper is organized as follows. Section 2 presents the related work. In section 3, the proposed method is described. In section 4, describe in detail the results and analysis. Finally, we summarized the main conclusions about the advantage and disadvantage of the proposed method in section 5.

2. RELATED WORK

Research works have presented numerous methods for categorizing handwritten characters and digits. Handwriting recognition has been previously demonstrated encouraging outcomes utilizing shallow networks [10], [11]. The accuracy rate attained by the MNIST dataset was 91.08% in Hinton et al.’s research on deep belief networks (DBN), which consist of three layers and incorporate a learning algorithm [12]. To recognize unconstrained handwriting, Pham et al. utilized a regularization technique called dropout to enhance efficiency of recurrent neural networks (RNNs) and lower the rates of word error (WER) and character error (CER) [13]. The performance of handwriting recognition (HCR) was significantly transformed by the introduction of the convolutional neural network (CNN), which achieved state-of-the-art accuracy [14], [15]. Simard et al. introduced a common CNN architecture for visual document analysis in 2003, which simplified the training of complex neural network methods [16]. Multilayer CNNs were utilized by Wang et al. to achieve excellent outcomes in performing end-to-end text recognition has been demonstrated on benchmark datasets, including street view text and ICDAR 2003 [17]. CNN has demonstrated exceptional performance in offline handwritten character recognition for various studies on handwritten text recognition in regional and international languages, including Chinese, have been conducted and carried out by researchers [18]–[20]. Arabic language [21]; handwritten Tamil character recognition [22]; handwritten character recognition on Indic scripts [23]; recognition of handwritten Urdu text. [24], [25]; Telugu character recognition [26]. In their model, Gupta and colleagues utilized CNN-derived features and identified informative local regions in recent character images, achieving better accuracy in recognition. They employed a novel multi-objective optimization framework for HCR. Ptucha et al. presented a conventional neural network-based intelligent character recognition (ICR) system in a logical manner [27]. The model was evaluated using IAM datasets and RIMES lexicon datasets in French language, and it reported a commendable result. Tapotosh Ghosh et al. utilized the CMATERdb dataset to convert images into 28×28 black-and-white forms with white as the foreground color in their study, and effectively designed CNN parameters using InceptionResNetV2, DenseNet121, and InceptionNetV3 to improve recognition performance.

A quantum convolutional neural network (QCNN) is a type of neural network that leverages the principles of quantum mechanics to perform computations. By using qubits instead of classical bits, QCNNs can potentially provide faster processing times and higher accuracy than classical neural networks for handwritten digit recognition. However, QCNN is an emerging technology and proposed QCNN based handwritten digit recognition to improve accuracy and reduce the processing time.

3. METHODOLOGY

The proposed methodology includes data collection, pre-processing, building the model, feature prediction, and visualization of results. The data collection involves gathering the relevant data that will be used
to train and evaluate the model. It’s important to ensure the data is representative and of good quality to produce reliable results. After collecting the data, preprocessing of the data needs to be prepared for the modeling phase. The tasks involved in this phase are data cleaning, handling missing values, removing duplicates, and dealing with outliers. Data pre-processing also included feature scaling, normalization, or transformation to make the data suitable for the selected model. The model is built using quantum convolutional neural network algorithm. The model is trained using the prepared data from the previous phase. After the model is trained, it can be used to make predictions on new, unseen data. The input features are provided to the model, and it generates predictions based on the learned patterns and relationships in the training data. After obtaining predictions from the model, the results are visualized and interpreted, the detailed process of the proposed method has shown in Figure 1 and detailed description as follows.

![Diagram of the proposed method based on QCNN](image)

**Figure 1. The proposed method based on QCNN**

### 3.1. Data collection

The data set is collected from the MNIST database is a substantial collection of handwritten numbers often utilized for various image processing techniques. This database consists of gray scale images of 60,000 training images and 10,000 test images and normalized to fit a 28×28 pixel bounding box. The sample handwritten digit recognition images are shown in Figure 2.

![Sample dataset of handwritten digit from MNIST](image)

**Figure 2. The sample dataset of handwritten digit from MNIST**
3.2. Pre-processing

Pre-processing is a crucial stage in hand digit recognition. The first is image normalization, a frequent pre-processing step that entails rescaling the values assigned to each pixel in the image such that they fall within a specific range, typically between 0 and 1. This can lessen the effect of changes in the input image or variations in illumination. The second method involves shrinking the input image to a fixed size, which can assist to cut down on the model’s parameter count and speed up training. Additionally, by lessening the effect of minute differences in the input image size, scaling can help to increase the resilience of the model. The third technique is called data augmentation, and it entails creating new training samples out of existing ones by rotating, translating, and scaling them. This can help to expand the training dataset and enhance the model's generalizability. The fourth method is feature extraction, which can help distinguish between several classes by extracting significant features from the input image. In QCNN, certain picture properties can be extracted by encoding them into the amplitudes of quantum states using quantum circuits. The final step is quantum circuit optimization, once the image has been transformed into a quantum circuit, it is crucial to refine the circuit to lessen the depth and the quantity of gates. This can aid in reducing the circuit's total runtime and enhancing the model's functionality.

3.3. Build the model using QCNN

The fundamental concept of quantum computing is that quantum bit or qubit that can be represented as a state $|0\rangle$ or $|1\rangle$, similar to a classical binary bit 0 or 1. However, a qubit can also exist in a superposition state $\alpha|0\rangle + \beta|1\rangle$ where the amplitudes $(\alpha, \beta) \in \mathbb{C}$ and satisfy the condition $|\alpha|^2 + |\beta|^2 = 1$. The process of quantum computation involves the utilization of quantum gates. These gates are unitary matrices that operate on either one or two qubits. For instance, the Hadamard gate is an example of a quantum gate that maps $|0\rangle$ to $1/\sqrt{2}(|0\rangle + |1\rangle)$ and $|1\rangle$ to $1/\sqrt{2}(|0\rangle - |1\rangle)$. After the computation, the result is a quantum state that can be measured to extract classical information. When a qubit is measured, i.e., the state $\alpha|0\rangle + \beta|1\rangle$ is observed, the result is either 0 or 1, and the probability of obtaining a particular outcome is proportional to the square of its corresponding amplitude.

The Hadamard gate and the CNOT (Controlled NOT) gate are two different types of gates in the quantum circuit seen in Figure 3. Except for the final qubit, which was initialised in state 1, all of the qubits were in state 0. With the exception of that final qubit, which needs to be switched from state 0 to state 1, states in Cirq are initialised in state 0 by default. Examining the circuit, we see that qubits 4 and 5 are connected by a series of CNOT gates, as are qubits 1 and 5 and qubits 0 and 5.

![Figure 3. Quantum circuit with Hadamard gate and CNOT gate](image_url)

Figure 3 illustrates a quantum neural network circuit. The input data is encoded into the qubits' state, and a sequence of quantum gates are applied to the qubits to process the input data. The readout qubit is then measured, and a prediction is made using the measurement data.

The procedure entails turning binary images made up of black and white pixels, such as the training and test data sets as shown in Figure 5 into quantum circuits. Figure 5(a) shows the training data sets (2,2) and (3,1) for CNOT Gate as a quantum circuit. Figure 5(b) shows the test data set (2,1) for CNOT Gate as a quantum circuit. Additionally, a threshold is applied to the pixel values and the qubits are only rotated through an $X$ gate if the pixel value is higher than the threshold. By doing this, noisy pixels' negative effects are lessened, and it is ensured that the resulting quantum circuit is efficient and robust.
Handwritten digit recognition using quantum convolution neural network (Ravuri Daniel)

In quantum computing systems, the quantum neural network (QNN) is used for learning tasks with quantum data more quickly. Figure 6 depicts the QNN architecture, the image is rescaling to 4×4 dimensions before inputting it into the unitary matrix for feature extraction across various channels. The extracted features are then utilized to develop a quantum circuit model, which is optimized a loss function combined with an optimizer. For binary classification issues, a 2-layer circuit design was adopted and was improved by hyper-parameter testing at several epochs. The final model, which resembled a tiny recurrent neural network stretched across pixels, was created using two layers, preparation, and readout processes. Every data qubit in every layer had an effect on the readout qubit since n repeats of the same gate were used. The stages of designing a quantum convolutional neural network as follows:

Stage 1: A quantum circuit is built to accept an input image with a 2×2 square region of focus and a limited field of view.
Stage 2: The unitary matrix (U) is applied to the gate, operations, and circuit in the form of a quantum operation, which is a common visual representation for quantum operations in Circuit.
Stage 3: The system is quantized by gathering a number of conventional values that are anticipated.
Stage 4: The predicted values for each channel of a single output pixel match to the conventional convolution layer in a similar way.
Stage 5: The procedure repeats the execution in different areas of the image and by relocating the image with more than one channel output object, a full scan of the input image can be accomplished.
Stage 6: Either a quantum or a classical layer would be compliant with the quantum convolution layer.

3.5. Prediction and visualization
In terms of visualization, QCNNs can be used to generate feature maps that highlight the regions of an image that are important for classification. This can help us understand how the network is making its predictions and identify any potential weaknesses or biases. The output of the QCNN is a set of probabilities for each possible digit class. The predicted class is the digit with the highest probability. After that to generate a visualization of the results, we can use feature maps to highlight the regions which are the portions of the input image that is important for classification. This can be done by identifying the quantum gates and circuits that were activated during the convolution operations and using these to generate a map of the input image.
RESULTS AND DISCUSSION

The current study has presented two performance indicators, namely accuracy and loss to assess the performance of the models when applied to the test set. These indicators are used to evaluate how well the models have performed. The accuracy of a model represents its ability to correctly identify the positive and negative classes. The accuracy in QCNN is more when compared to CNN. The accuracy in CNN at point 1 is between 0 to 25 percentages whereas in QCNN it is between 75 to 100 percentages. At point 4 in CNN the accuracy is between 50 to 75 percentages whereas in QCNN the accuracy is between 75 to 100 percentages. At point 10 in CNN the accuracy is at 75 percentages whereas in QCNN the accuracy is between 75 to 100 percentages. The Figure 7 shows the drastically increase in accuracy in CNN whereas there is a constant change in QCNN.

The loss in CNN is more when compared to QCNN. The loss in CNN at point 1 is between 20 to 25 percentages whereas in QCNN it is between 5 to 10 percentages. At point 4 in CNN the accuracy is between 15 to 20 percentages whereas in QCNN the accuracy is between 0 to 5 percentages. At point 10 in CNN the accuracy is between 5 to 10 percentages whereas in QCNN the accuracy is between 0 to 5 percentages. The Figure 8 shows the drastically decrease in loss in CNN whereas there is a constant change in loss in QCNN.

The Figure 9 shows that QCNN achieved higher accuracy (91.07%) than CNN (84.68%) for handwritten digit recognition. The loss of the QCNN (3.3%) is lower than CNN (7.33%) in terms of handwritten digit recognition. However, the specific difference in accuracy and loss between CNN and QCNN are 6.39% and 4.07% respectively.
5. CONCLUSION

In this paper, a novel approach to using a quantum neural learning model for handwriting recognition is provided. Using a sample set of over 60,000 handwritten digit images, the experimental comparison of the model showed a high level of efficiency with an overall accuracy of 91.07%. The model’s computation-based training process took much less time than more conventional classical CNN model development with comparable sample sizes while using quantum hardware. The model’s overall speed and effectiveness were also demonstrated by the fact that the inference time for each image was measured at one minute. With the use of QCNN, several of the drawbacks of CNN such as over-fitting and disappearing gradients, have been addressed, leading to higher accuracy rates. Additionally, the proposed technique has shown resilient to image distortions and changes in handwriting styles. The generalizability and scalability of QCNN for bigger datasets and trickier recognition tasks, however, require more investigation. The usage of QCNN has a lot of potential to advance the science of handwritten digit recognition overall.

REFERENCES


---

**BIOGRAPHIES OF AUTHORS**

**Dr. Ravuri Daniel** working as associate professor in the department of computer science and engineering, Prasad V. Potdar Siddhartha Institute of Technology, Vijayawada, India. He has industrial, national, and international teaching experience. Awarded Ph.D. from Jawaharlal Nehru Technological University Kakinada (JNTU Kakinada), M. Tech from Andhra University, and B. Tech from Jawaharlal Nehru Technological University Hyderabad. His area of research in the field of Computer Networks, Wireless Sensor Networks, Internet of Things, Data Analytics, AI and Embedded Systems. Presently, he is exploring innovative applications in IoT with data analytics. I am having hands-on experience on IoT hardware platforms (Arduino, Raspberry Pi) and software platforms (C, Python, R, Matlab). He can be contacted at email: danielravuri@gmail.com.
Dr. Bode Prasad received his master's degree from KoneruLakshmaiah Engineering College in Guntur, India, and obtained his Ph.D. degree from Andhra University College of Engineering, Visakhapatnam, India in the year 2015. He currently holds the position of Professor in the Department of Information Technology at Vignan’s Institute of Information Technology (A) in Visakhapatnam, India. His research interests span a wide range of topics, including computer networks, mobile ad-hoc networks, cloud computing, wireless networks, machine learning, data science, artificial intelligence, human action analysis, and sign language machine translation. He can be contacted at email: arjunprasad.bode@gmail.com.

Mr. Pasam Prudhvi Kiran having M. Tech degree in Information Technology and pursuing his Ph.D. degree in CSSE Dept. from Andhra University College of Engineering, Visakhapatnam, India. He has 8 years of teaching experience and he currently holds the position of Assistant Professor in the Department of Computer Science and Engineering (IoT) at R.V.R. & J. C. College of Engineering (A)in Guntur, India. His research area moves around IoT and Artificial Intelligence. He can be contacted at email: pasamprudhvi@gmail.com.

Mr. Dorababu Sudarsa M. Tech in Information Technology from Sathyabama University, Chennai, 2009. JNT University, Hyderabad, awarded him a B.Tech. in Computer Science and Information Technology in 2002. He received the Best Teaching Faculty Award from KoneruLakshmaiah University, Vadeswaram, Guntur, India, and Audi Sankara College of Engineering and Technology, Gudur, India. He taught 18 years total. Cloud computing, data mining and warehousing, data science, and IoT interest him. 15 international publications and conferences published him. He's an ISTE lifer. He can be contacted at email: dorababu.sudarsa@gmail.com.

Dr. Ambarapu Sudhakar since 2019 as Professor and Head of the Electrical and Electronics Engineering Department at MLR Institute of Technology, Hyderabad, Telangana, India. He taught engineering colleges for 17 years. He likes electric drives, clever controllers, etc. He has several patents and research articles in indexed journals. Research earned him the 2016 INDUS Research Excellence Award. He wrote a textbook and offered expert webinars and conferences. He organised state and national conferences with Indian government subsidies. He can be contacted at email: sudhakar.a@mlrinstitutions.ac.in.

Dr. Bodapati Venkata Rajanna received B. Tech degree in Electrical and Electronics Engineering from Chirala Engineering College, JNTU, Kakinada, India, in 2010, M. Tech degree in Power Electronics and Drives from KoneruLakshmaiah Education Foundation, Guntur, India, in 2015 and Ph.D. in Electrical and Electronics Engineering at KoneruLakshmaiah Education Foundation, Guntur, India, in 2021. Currently, he is working as an Associate Professor at MLR Institute of Technology, Hyderabad. His Current Research includes, Dynamic Modeling of Batteries for Renewable Energy Storage, Electric vehicles and Portable Electronics Applications, Renewable Energy Sources Integration with Battery Energy Storage Systems (BESS), Smart Metering and Smart Grids, Micro-Grids, Automatic Meter Reading (AMR) devices, GSM/GPRS and PLC (Power Line Carrier) Communication and Various modulation techniques such as QPSK, BPSK, ASK, FSK, OOK and GMSK. He can be contacted at email: rajannabv2012@gmail.com.

Handwritten digit recognition using quantum convolution neural network (Ravuri Daniel)