The Journal of Supercomputing (2025) 81:564
https://doi.org/10.1007/s11227-025-07083-3

A survey on quantum deep learning

Huaiguang Wu[1] · Jiahui Zhang[1] · Lijie Wang[1] · Daiyi Li[1] · Delong Kong[1] · Yucan Han[1]

Accepted: 17 February 2025

Abstract
Quantum deep learning (QDL), which combines the unique strengths of quantum computing and deep learning, is gradually becoming a focal point. It offers new ideas for addressing the many challenges currently faced. In this survey, we review the representative algorithms that have combined quantum computing and deep learning in recent years. Firstly, we categorize the discussion based on data types into three areas: text, image, and multimodal data. We focus on QDL algorithms within these categories and explore their characteristics. Secondly, this paper compares the performance of the QDL model with the traditional model. By comparison, QDL not only demonstrates enhanced feature extraction capabilities but is also able to handle more complex data. In addition, the unique properties of quantum computing, such as quantum superposition and quantum entanglement, can accelerate calculations and improve model performance. These advantages demonstrate its potential efficiency over traditional methods. Finally, a summary and outlook on the prevailing research conditions in QDL have been given. This article integrates current research findings in QDL, providing a clear research background for subsequent researchers.

Keywords Quantum computing · Machine learning · Quantum deep learning · Quantum machine learning

1 Introduction

Facing increasingly complex problems and data, traditional machine learning methods have encountered challenges. With little need for human intervention, deep learning, through the construction of deep neural network models, is able to automatically learn and extract features from data.

Jiahui Zhang, Lijie Wang, Daiyi Li, Delong Kong, and Yucan Han have contributed equally to this work.

Extended author information available on the last page of the article

Published online: 03 March 2025

Hinton et al. [1] put forward the concept of deep learning in 2006, and the expansion of processing capabilities and the support of massive training samples in 2012, deep learning became a hot topic in machine learning research. In 2014, the Google team led by Sutskever et al. [2] presented the Seq2Seq model, which combined with recurrent neural networks (RNNs), achieved a leap in machine translation and opened a latest chapter in the field of text processing for neural networks. In 2015, the AlphaGo program from the Google DeepMind team [3] demonstrated the extraordinary capabilities of deep learning in complex decision-making and strategy by defeating the world champion of Go, Lee Sedol. In 2016, Paszke et al. [4] developed the PyTorch framework, which is recognized for its dynamic computational graph, making it more flexible and efficient in model development and debugging. In 2017, Google introduced the TensorFlow framework [5], providing researchers with a powerful and flexible platform that promotes rapid iteration and optimization of deep learning algorithms. In 2017, Vaswani et al. [6] put forward the Transformer model, which has been highly successful in the field of Natural Language Processing (NLP) by achieving efficient processing of sequence data through self-attention mechanisms. Over time, deep learning has been applied more and more widely in various fields, and its development is closely related to computing technology, data resources, and algorithmic innovation. Since 2018, deep learning has achieved notable advancements in areas such as NLP and computer vision, fueling automation across various industries. Despite significant progress in the field of deep learning (DL), some challenges remain. As model complexity increases, the demand for computing resources in deep learning rises dramatically, especially when processing large-scale data. Training deep neural networks usually requires a lot of computing and hardware resources, which makes the training process more expensive and time-consuming. In addition, deep learning models perform well on training datasets, but in the real world, especially when the data distribution changes, the models often have difficulty adapting to unseen data, resulting in insufficient generalization ability and affecting their performance in practical applications.

As data volume continues to swell and computational demands escalate, deep learning is confronting significant challenges regarding computational resources and time efficiency. Quantum computing, with its potential for parallel processing and the unique characteristics of quantum entanglement, is exploring new possibilities within the realm of machine learning. The primary objective of quantum machine learning (QML) is to leverage qubits for storing and processing data, rather than traditional bits used in traditional machine learning. This method intends to boost the competence and robustness of machine learning, enabling more efficient handling of high-dimensional data and issues involving uncertainty. The main developmental milestones are depicted in Table 1.

Building on QML, quantum deep learning (QDL) further combines quantum computing with deep learning. It uses its unique parallel computing power to solve complex problems. In 1994, Shor proposed the Shor algorithm [17], a quantum algorithm for factoring large numbers into their prime components. It utilizes quantum Fourier transform to reduce the problem to finding the order. This algorithm is significant in the history of computer science as it demonstrates the potential of quantum computing to solve certain problems, such as

Table 1 The development of QML

Time	Character	Contribution
1995	Kak [7]	The notion of "quantum neural computing" was first proposed, establishing a basis for the integration of quantum computing and machine learning
2003	Anguita et al. [8]	The notion of quantum support vector machine was first introduced
2009	Harrow et al. [9]	The HHL quantum algorithm was proposed, which demonstrated unique advantages in resolving systems of linear equations and is one of the important algorithms in QML
2014	Lloyd [10]	Quantum principal component analysis (QPCA) was proposed, which can exponentially accelerate classical principal component analysis algorithms
2014	Rebentrost et al. [11]	A QSVM was proposed, which essentially utilizes quantum optimization algorithms to accelerate the inner product calculation problem in SVM
2015	Li et al. [12]	A four-qubit QSVM was implemented in a real environment using nuclear magnetic resonance quantum computers, and handwritten digit recognition was achieved based on this algorithm
2017	Farhi et al. [13]	The quantum approximate optimization algorithm (QAOA) was proposed, which made new progress in quantum optimization problems and became an important research direction in the field of quantum computing
2019	Benedetti et al. [14]	The parameterized quantum circuit (PQC) was proposed as a machine learning model, and the adjustable parameters of quantum gates were used to optimize the performance of quantum circuits
2022	Zhang et al. [15]	An improved circuit implementation of the HHL algorithm was proposed and simulated on the QISKIT platform, which promoted the practical application of quantum linear equation solving algorithms
2024	Khot et al. [16]	A quantum-inspired machine learning framework based on the physical Ising solver chip was proposed, and a QSVM was developed and its feasibility was verified

integer factorization, exponentially faster than classical algorithms, posing a major challenge to classical cryptographic systems. In 1996, Grover [18] introduced the Grover algorithm, a quantum search strategy that offers a square-root acceleration compared to classical algorithms. Under the condition that the database is sufficiently disordered and without specific data structure constraints, the Grover algorithm can efficiently address the issue of locating a specific item from an unsorted set of N items. Subsequently, Harrow et al. [19] introduced the HHL algorithm, a quantum approach designed to solve systems of linear equations. It boasts a lower time complexity in certain scenarios, which can solve a group of linear equations in N dimensions in O(log N) time. Following this, many corresponding quantum algorithms have been introduced, such as the Quantum Support Vector Machine (QSVM) algorithm [20], which utilizes the superposition and parallelism of quantum computing to handle larger datasets and provide more accurate predictive results in certain situations. The QPCA algorithm [21] utilizes the characteristics of quantum computing to extract the primary components from data, thereby better understanding and interpreting the structure of the data. Quantum simulation algorithms [22] use quantum computers to imitate complicated quantum architectures, such as molecular structures, materials, and biological systems. Lloyd et al. [23] proposed the Quantum Boltzmann Machine (QBM) algorithm. The QBM simulates the states of neurons through the states of quantum bits and uses quantum gate operations to implement connections and interactions between neurons, boosting the model's expressive capabilities and learning efficiency. Subsequently, the introduction of Quantum Clustering algorithms [24] provided more accurate clustering results. Parameterized Quantum Circuits (PQC) [25] utilize quantum gates with parameters to engineer quantum circuits and optimize the performance of the quantum circuit by adjusting these parameters. However, QDL still faces many challenges. In the NISQ era, the hardware limitations of quantum computing, such as the number and quality of quantum bits and noise issues, directly affect the reliability and scalability of the QDL model. Although hardware issues are currently difficult to change, at this stage, how to design quantum circuits that are more suitable for the NISQ era has become a key challenge to improve QDL performance.

As shown in Fig. 1, this article is divided into three core modules, which comprehensively explain the cutting-edge applications of quantum deep learning algorithms. The first module introduces text-based quantum deep learning, focusing on its specific implementation and potential advantages in text generation and classification. The second module focuses on image-based quantum deep learning, covering image generation, classification, and detection. The third module delves into multimodal-based quantum deep learning, analyzing its performance and application prospects in text processing, as well as image processing. Through in-depth discussions in these three modules, this article aims to reveal the application potential and future development direction of quantum deep learning under different data modalities.

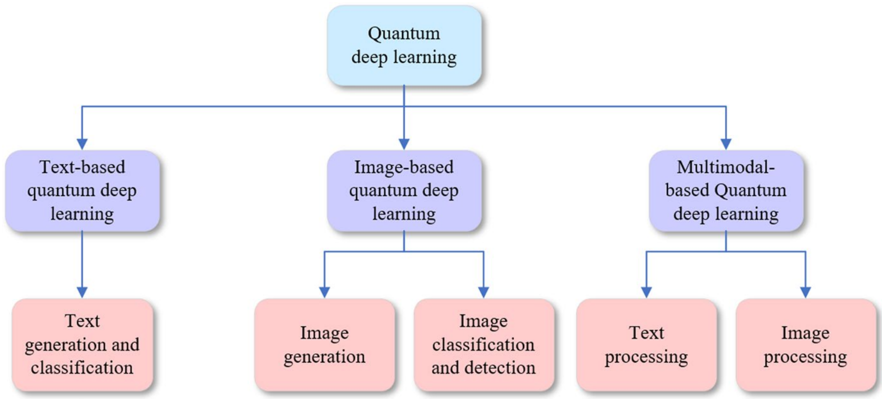


Fig. 1 Overall structure of the paper

2 Quantum foundations and commonly used quantum frameworks

The computational unit of classical computers is bits, while the fundamental unit in quantum computers is quantum bits. Quantum bits can be represented by quantum states $|0\rangle$ and $|1\rangle$, where $|0\rangle$ and $|1\rangle$ are a pair of orthogonal ground states corresponding to 0 and 1 in classical bits, respectively. Unlike classical bits, a quantum system in quantum mechanics can exist in several possible states, not just a single basic state. This is known as quantum superposition. The state of a quantum bit can be represented as follows:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{1}$$

where α and β are complex numbers. And it satisfies the normalization condition $\langle\psi|\psi\rangle = 1$, meaning $|\alpha|^2 + |\beta|^2 = 1$. $|\psi\rangle$ simultaneously contains a fusion of fundamental states $|0\rangle$ and $|1\rangle$, with the values of and describing the weights and phases of the superposition.

2.1 Quantum gates

With the rapid development of quantum computing and quantum neural networks (QNNs), gate-based quantum computers, as one of the core architectures, have attracted significant attention. In the optimization of QNN training, a new method has been proposed that reduces quantum circuit depth and optimizes quantum gradient descent, significantly improving training efficiency, stability, and accuracy, thereby laying the foundation for QNN applications [26]. Additionally, the optimization of circuit depth has long been a bottleneck in quantum computing performance. An innovative strategy, which simplifies circuit structures and reduces the number of layers, has enhanced computational efficiency while lowering noise and error rates, providing a key pathway for quantum hardware optimization [27]. In the application of quantum machine learning (QML), research combining QML techniques has significantly increased the capacity and access speed of quantum memory, supporting the efficient operation of near-term quantum devices [28]. Meanwhile, the scalability of quantum computing has also been addressed. A study on distributed quantum computing architecture, through the collaboration of multiple quantum computing units, has improved overall computational power and scalability, offering a feasible solution for large-scale quantum computing tasks [29]. Finally, research on the quantum internet has promoted the integration of quantum communication and computing, proposing key technologies for realizing efficient, secure, and scalable quantum communication networks, thus providing a clear direction for the application of quantum networks [30].

Quantum gates are fundamental operations in quantum computing, capable of transforming the states of one or more qubits. Unlike classical logic gates, quantum gates can handle quantum superposition states and follow specific rules such as reversibility. Commonly used single-qubit gates include the Pauli-X gate, Pauli-Y gate, Pauli-Z gate, and Hadamard gate. The Pauli-X gate flips the state of a qubit, the Pauli-Z gate changes its phase, and the Hadamard gate puts the qubit into a superposition of states. The controlled gate (CNOT gate) is a commonly used two-qubit quantum gate, which means that it operates on two qubits simultaneously. One of the qubits is known as the control qubit, usually positioned at the lower bit, while the other is called the target qubit, situated at the higher bit. The operation of the CNOT gate is very intuitive: When the control qubit is in the state $|0\rangle$, the CNOT gate does not perform any operation on the target qubit, meaning the state of the target qubit remains unchanged; when the control qubit is in the state $|1\rangle$, the CNOT gate performs a NOT operation on the target qubit, thus flipping the state of the target qubit. Quantum circuits are represented by parallel horizontal lines, with each line corresponding to a qubit. Quantum gates are denoted by boxes placed on these lines to indicate their action on the qubits. The circuits for the Hadamard gate and the CNOT gate are shown in Fig. 2.

In quantum mechanics and quantum technology, conducting measurements on quantum states is an essential and vital component of the procedure. Measurements are the sole means by which we can ascertain the precise current state of a quantum system. Quantum measurement is characterized by a collection of operators represented as M_m, where the measurement operators satisfy the normalization condition:

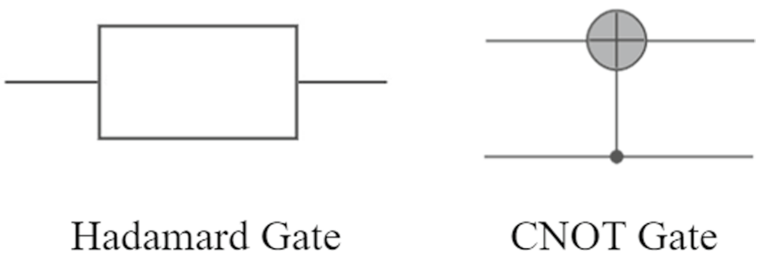


Fig. 2 Some commonly used quantum gates

$\sum_m M_m^\dagger M_m = I$. Here M_m is the measurement operator corresponding to the measurement result m, $M_m^\dagger$ is the conjugate transpose of M_m (commonly referred to as the Hermitian conjugate in quantum mechanics), and I is the unit operator. This condition ensures that when a quantum state $|\psi\rangle$ is measured, the aggregate of probabilities for every potential outcome is 1. The chance of getting the result m is indicated by $p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle$. Note that after the measurement, the system's state typically transitions from $|\psi\rangle$ to $\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}$ (assuming $p(m) \neq 0$), which is referred to as the "collapse" or "projection" following the measurement. From machine learning to deep learning, and then to QML and QDL, we can witness the continuous development and progress of computational technologies. These innovative technologies offer fresh perspectives for tackling the complexities of large-scale data processing and are poised to have a substantial impact on future scientific studies and industrial uses.

2.2 Quantum circuits

A quantum circuit is a structure composed of a series of quantum gates and qubits combined in a certain order and rule. Similar to circuit diagrams in classical computing, quantum circuits provide an intuitive way to describe how quantum gates interact with qubits to implement defined quantum algorithms or computational tasks. Common types of parametric quantum circuits include: non-entangled circuits, Bellman circuits, etc. In non-entangled circuits, since there is no interdependence among qubits, the behavior of the circuit is relatively more predictable and controllable. These circuits are typically used to perform simple quantum operations or in scenarios where the effects of quantum entanglement need to be avoided. Bellman circuits are widely used in the design of quantum circuits, primarily for implementing nonlinear functions in feedback circuits. The core of a Bellman circuit is a nonlinear element, usually a special diode known as a Bellman diode, which consists of two pn junctions and has asymmetric and nonlinear characteristics as shown in Fig. 3.

There is a significant positive correlation between a circuit's expressive power and its performance, while the relationship between entanglement and performance is relatively weak. More importantly, excessive entanglement can adversely affect circuit performance. Based on Hubregtsen et al.'s research, circuit design tends to

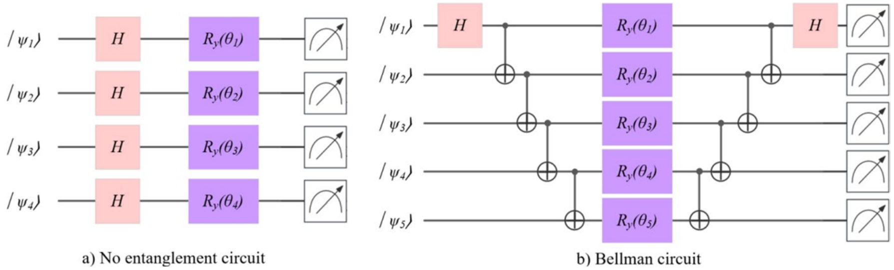
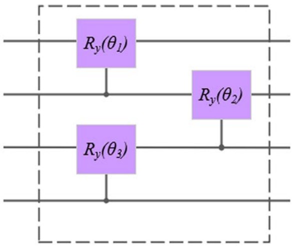


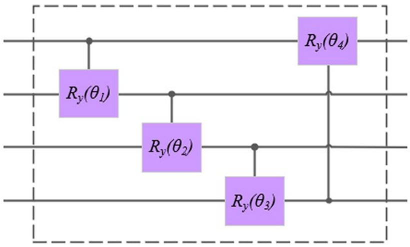
Fig. 3 Parameterized quantum circuits

favor structures with high expressiveness while also addressing the issue of excessive entanglement, particularly by limiting the use of CNOT gates. This led to the development of parametric quantum circuits (PQCs) with a ring structure. PQCs consist of a sequence of quantum gates with adjustable variables that define the state and function of the quantum circuit, thereby affecting the network's performance and learning capabilities.

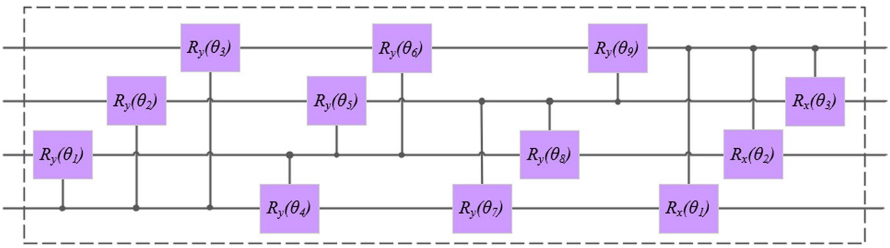
Figure 4 showcases the common topologies of Parameterized Quantum Circuits (PQCs), which include recently connected, annular, and fully connected structures. Each topology is distinguished by its unique qubit connectivity and circuit performance characteristics. The recently connected topology restricts interactions to adjacent qubits, resulting in circuits that are simple and have low-noise levels, albeit with limited expressive power. This makes it suitable for scenarios where noise minimization is crucial. The annular topology expands the expressive power and entanglement control by including interactions between non-adjacent qubits. This design offers a balanced approach, making it a popular choice in practical applications. The quantum gate sections within this topology can be customized to meet specific algorithmic requirements, providing a flexible framework for various quantum computing tasks. The fully connected topology allows for the most extensive interactions among all qubits, offering the highest degree of expressive power and entanglement. The design of the quantum gate sections in this topology is highly adaptable, enabling the construction of circuits tailored to diverse computational needs and problem domains. However, this flexibility comes with an increased number of CNOT gates, which can lead to greater circuit depth and a higher susceptibility to noise and errors. As such, this topology is most effective in advanced quantum hardware



(a) Recently connected type



(b) Annular type



(c) Fully connected type

Fig. 4 Common PQC topologies

environments capable of supporting complex circuitry. The design of PQCs, particularly the arrangement and selection of quantum gates within these topologies, significantly impacts the expressive power, training efficiency, and hardware performance of quantum models.

2.3 Common quantum frameworks

As quantum computing advances, a growing number of quantum computing frameworks have emerged. These frameworks have been instrumental in fostering the integration of quantum computing with deep learning. They not only aid in the growth and validation of quantum algorithms but also open up new avenues for research in QML. Table 2 shows the five main quantum computing frameworks.

(1) IBM's Qiskit, a freely available framework for quantum computing, provides a complete platform for crafting, simulating, and executing quantum algorithms. It also facilitates user access to the IBM Quantum Experience. Qiskit is highly influential in quantum computing simulation, quantum algorithm research, and running machine learning tasks on quantum hardware, but its efficiency can be limited in certain high-performance computing scenarios.
(2) QPanda was developed by Origin Quantum, a Chinese company, in 2018. It is designed for creating and executing complex quantum circuits and algorithms, making it easier for researchers and developers to utilize quantum computing resources. QPanda has widespread applications in quantum computing research and education, particularly within the Chinese quantum computing ecosystem.
(3) PennyLane, launched by the Canadian quantum computing company Xanadu in 2018, supports seamless integration with classical deep learning tools such as TensorFlow and PyTorch. It facilitates the construction of models that combine quantum and classical computing. PennyLane provides powerful tools for tasks such as quantum chemistry and quantum optimization, supporting automatic differentiation and running on various hardware platforms.
(4) TensorFlow Quantum (TFQ), developed by Google in collaboration with NASA Ames, combines TensorFlow with quantum computing. Released in 2020, TFQ focuses on the maturation of quantum data processing and QML models. It allows users to train QML models using TensorFlow's advanced APIs and integrate them with classical machine learning models.
(5) TorchQuantum, a quantum computing framework based on PyTorch, was released in 2022, concentrating on merging quantum computing capabilities with deep learning models. Developers can leverage PyTorch's automatic differentiation capabilities to train quantum neural networks and support hybrid computing with classical deep learning models. TorchQuantum is designed to provide developers familiar with PyTorch workflows, reducing the complexity of developing quantum neural networks.

These frameworks have their own characteristics and provide broad support for the blend of quantum computing with deep learning by providing different tools and

Table 2 Quantum frameworks

Framework	Release time	Publisher	Implementation principle and technical characteristics	Applicable scenarios and advantages
Qiskit	2017	IBM	An open-source framework based on Python, offering a platform for quantum programming, simulation, and execution. It supports access to quantum hardware, such as IBM Quantum Experience	Suitable for research and validation of quantum algorithms, quantum circuit design, and quantum machine learning tasks. However, it may face efficiency limitations in high-performance computing scenarios
QPanda	2018	Origin quantum	An open-source platform based on C++, supporting multiple programming languages, including Python and C++. It allows users to integrate it into different application scenarios through interfaces	Widely used in quantum computing research and education, especially in China's quantum computing ecosystem. Suitable for developing complex quantum circuits and algorithms
PennyLane	2018	Xanadu	Supports quantum circuit training via automatic differentiation and is compatible with multiple quantum hardware platforms, focusing on quantum chemistry and quantum optimization	Ideal for quantum machine learning tasks, including quantum chemistry simulations and quantum optimization problems. Its seamless integration with existing machine learning frameworks simplifies the development of Quantum Machine Learning (QML)
Tensor flow quantum	2020	Google and NASA ames	A framework based on TensorFlow, focusing on quantum data processing and QML model training, integrating quantum computing with classical machine learning models	Suitable for quantum machine learning models built with TensorFlow, particularly for tasks requiring the combination of classical and quantum data processing
Torch quantum	2022	MIT han lab	A quantum computing framework based on PyTorch, supporting automatic differentiation, facilitating the development of quantum neural networks and hybrid quantum-classical deep learning models	Suitable for developers familiar with PyTorch, especially for tasks that integrate deep learning models with quantum neural networks

platforms. Their joint efforts have promoted the rapid development of this field, enabling researchers to dig into the potential of the prospects of quantum computing in deep learning and develop more innovative QML models.

3 Quantum deep learning

QDL is the integration of quantum computing and deep learning, combining their strengths to create a more efficient and powerful learning framework. The foundation of QDL lies in the principles of quantum mechanics, which provide exponential speedup and parallelism, while leveraging the hierarchical feature extraction capabilities of deep learning. By combining the multi-layer structure of deep learning with the advantages of quantum computing, QDL forms a more powerful and efficient learning framework. In 2008, Giovannetti et al. [31] proposed Quantum Random Access Memory (QRAM), which provides an effective solution for data storage and retrieval, laying the foundation for certain QDL applications. QRAM itself is not inherently associated with QDL or QML, but rather serves as a supporting infrastructure. In 2016, Wiebe et al. [32] from Microsoft first combined quantum computing with deep learning, accelerating the refinement of deep networks by quantum sampling for gradient estimation of Restricted Boltzmann Machines. In recent years, researchers have proposed an architecture of Quantum Convolutional Neural Networks (QCNN), which is used to solve the classification problems of quantum data [33]. The research team led by Huang Heli from the University of Science and Technology of China [34] also proposed a Hybrid Quantum-Classical Convolutional Neural Network (QCCNN), which combines the characteristics of quantum and traditional CNNs and has shown learning accuracy surpassing classical CNNs in classification tasks. With the continuous betterment of quantum computing technology and the increasingly ubiquitous application of deep learning, QDL is anticipated to achieve momentous breakthroughs in the future.

In order to collate the recent literature in the field of quantum computing and deep learning convergence, this paper draws on the connectedpapers and Origin tools in order to dig deeper and demonstrate the intrinsic connections, relevance, and their academic impact among these important papers. During the review, we categorize the literature based on task types into three major categories: text-based quantum deep learning, image-based quantum deep learning, and multimodal-based quantum deep learning. For text-based quantum deep learning, we focus on text generation and classification. For image-based quantum deep learning, we further divide it into two task categories: image generation and image classification and detection. For multimodal-based quantum deep learning, we analyze its applications in text processing and image processing. We conduct in-depth discussions on specific tasks in each category and systematically review and evaluate the relevant literature.

3.1 Text-based quantum deep learning

3.1.1 Text generation and classification

Text generation and text classification are two key tasks in natural language processing (NLP). Text generation aims to automatically generate relevant text content based on input information, typically by providing partial information (such as text fragments or questions) to generate complete or relevant output. Text classification involves analyzing the content of a given text and categorizing it into different classes, thereby achieving automatic classification and processing of information. Both tasks are based on a deep understanding and modeling of text, in order to achieve more efficient and accurate data processing.

Figure 5 shows the general processing flow of text-based quantum deep learning. The main methods of data preprocessing here are: data cleaning (removing noise and irrelevant characters), word segmentation, standardization, removing stop words, building vocabulary, TF-IDF, word embedding, etc.; quantum coding is to convert the processed data into quantum states. In quantum deep learning, data encoding is an important link between classical data and quantum computing, and different encoding methods are suitable for different task scenarios. Basis Encoding directly maps classical data to the ground state of quantum bits, but it requires more quantum bits and is suitable for low-dimensional data. Amplitude Encoding uses the amplitude of quantum states to represent data, which can efficiently compress high-dimensional data, but realizing complex quantum circuits is one of the challenges. Phase Encoding represents data through the phase of quantum states and is often used for tasks that need to capture data periodicity or phase relationships. Angle Encoding maps data to rotation angles, which is easy to implement through simple revolving gates and is suitable for processing low-dimensional and periodic data. These encoding methods have their own characteristics. Choosing the right encoding method according to the task requirements is the key to improving the performance of quantum models. The next steps, including model selection and construction, training, evaluation, and testing, are similar to traditional methods. The difference is that in the algorithm implementation and performance evaluation stages, we introduced quantum algorithms and special evaluation criteria to optimize the performance of the model by taking

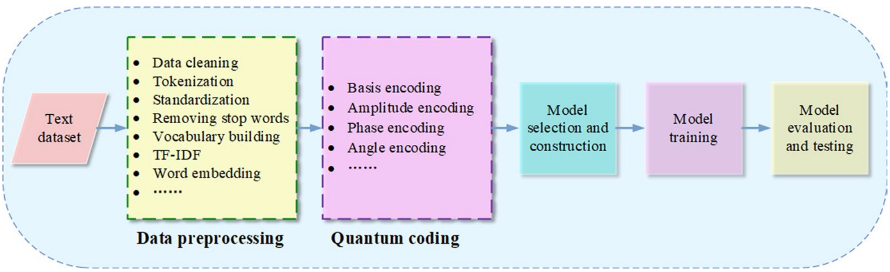


Fig. 5 General flowchart of text-based QDL

advantage of quantum computing and using customized indicators to accurately measure its effect.

The core technology of text processing is to enable computers to understand, interpret, and generate human language, thereby achieving more efficient human–computer interaction. Words, as the basic units of natural language, have a major influence on the capability of NLP tasks through their representation. Quantum physics, particularly quantum computing, offers new ideas for word representation.

Word representation involves word ambiguity. Word ambiguity involves applying quantum mechanics concepts and mathematical frameworks to represent and process the polysemy of words. By modeling this ambiguity, NLP systems can more accurately understand language. The polysemy of words can be modeled through quantum superposition states. Suppose we want to model the two meanings of the word "mouse": "rodent" and "device" as shown in Fig. 6.

In the quantum model, we can define two orthogonal state vectors $|rodent\rangle$ and $|device\rangle$. The word "mouse" can be expressed as the superposition of these two vectors:

$$| \text{mouse}\rangle = \alpha| \text{rodent}\rangle + \beta| \text{device}\rangle \tag{2}$$

Among them, α and β fulfill the condition $|\alpha|^2 + |\beta|^2 = 1$. This representation method allows “mouse” to contain both “rodend” and “device” meanings, and the specific meaning will be determined through “measurement” in a specific context, similar to the collapse of quantum states. In this way, quantum models provide a powerful tool for understanding and processing semantic ambiguity in natural language.

When a word has three meanings, such as m_1, m_2, and m_3, quantum models can be extended. The state of the word w can be represented as the superposition of three orthogonal state vectors $|m_1\rangle$, $|m_2\rangle$, and $|m_3\rangle$.

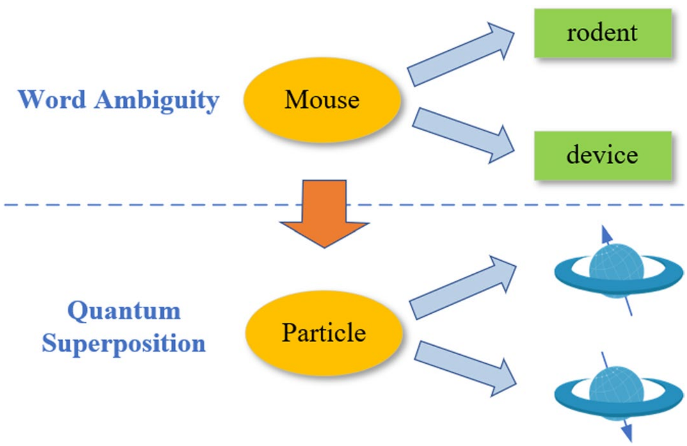


Fig. 6 Quantum processing of word ambiguity

$$|w\rangle = c_1|m_1\rangle + c_2|m_2\rangle + c_3|m_3\rangle \tag{3}$$

Among them, c_1, c_2, and c_3 are complex coefficients that satisfy the normalization condition: $|c_1|^2 + |c_2|^2 + |c_3|^2 = 1$. This representation method allows the word w to simultaneously contain these three meanings in a probabilistic manner. And the specific meaning will be determined through the measurement process based on the context. The measurement process will collapse the superposition state to one of the ground states $|m_1\rangle$, $|m_2\rangle$ or $|m_3\rangle$, and the probabilities of the measurement results are $|c_1|^2$, $|c_2|^2$, and $|c_3|^2$.

When a word has more meaning (such as more than three), the quantum model can still be extended. If the word w has n possible meanings (n>3), the state of the word can be expressed as follows: $|w\rangle = \sum_{i=1}^{n} c_i |m_i\rangle$. Among them, $|m_i\rangle$ is the orthogonal basis vector corresponding to the i-th meaning, and c_i is the probability amplitude coefficient of each meaning, satisfying the normalization condition: $\sum_{i=1}^{n} |c_i|^2 = 1$. This representation can flexibly model words with high polysemy and select the specific meaning of words through context or contextual "measurement."

By matching words with their meanings, we can construct the meaning of sentences. Quantum-inspired algorithms can simulate this comprehension process, helping us to grasp the deep structure of language more profoundly. In this process, words w get mapped to projectors.

$$w \to \Pi_w = |e_w\rangle\langle e_w| \tag{4}$$

Here $w \in$ V, and $|e_w\rangle$ is the single encoding of the word w. For example, if $V = \{\text{quantum, computing}\}$, then $\Pi_{\text{computing}} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. The vocabulary list $K = \{w_1, w_2, \ldots, w_n\}$ establishes the connections among words.

$$\kappa \to |\kappa\rangle\langle\kappa|, \quad |\kappa\rangle = \sum_{i=1}^{n} a_i |e_{w_i}\rangle \tag{5}$$

Among them, $\{a_i\}_{i=1}^{n})$ is a complex number, and $\sum_{i=1}^{n} a_i^2 = 1$. For example, we establish a dependency relationship between "quantum" and "computing," $K_{\text{gc}} = \{\text{quantum, computing}\}$, through $K_{\text{gc}} = |K_{\text{gc}}\rangle\langle K_{\text{gc}}|$, where

$$|k_{qc}\rangle = \sqrt{\frac{2}{5}}|e_{\text{quantum}}\rangle + \sqrt{\frac{3}{5}}|e_{\text{computing}}\rangle \tag{6}$$

So, $k_{qc} = \begin{bmatrix} \frac{2}{5} & \frac{\sqrt{6}}{5} \\ \frac{\sqrt{6}}{5} & \frac{3}{5} \end{bmatrix}$, where k_{qc} is the density matrix, and $|k_{qc}\rangle$ is the superposition state. In the framework of quantum mechanics, the components of the density matrix k_{qc} reveal the relationship between the quantum states $|\text{quantum}\rangle$ and $|\text{computing}\rangle$. Based on this, the interdependence between nature and language has been formally described.

We searched the literature using "NLP; Quantum Deep Learning" as key search terms through academic databases and indexing services such as IEEE Xplore, Springer, Web of Science, ACM Digital Library, Google Scholar, and Science Direct. We then applied the following inclusion criteria: (1) Articles explicitly addressing Quantum Deep Learning in the context of text generation and classification. (2) Studies published in peer-reviewed journals or presented at reputable conferences. (3) Papers that provided experimental validation or theoretical insights into QDL methods. (4) Research that offered novel contributions, such as innovative algorithms or unique applications. Ultimately, we selected 23 articles from recent years. These articles not only innovate theoretically but also demonstrate the immense potential of QDL algorithms in solving NLP problems in practical applications.

In terms of word embedding optimization, Zhang et al. [35] developed a neural network-based quantum language model (NNQLM) in 2018, which improved the efficiency and accuracy of question–answering systems. In 2019, Zhang et al. [36], considering the interactions between words with multiple meanings, proposed a language modeling method inspired by quantum multi-world functions (QMWF), using tensor products to precisely model the complex interactions between vocabulary. In 2019, Panahian et al. [37] proposed two methods, word2ket and word2ketXS, which reduce the storage space of word embeddings while maintaining NLP accuracy, but have a high computational complexity for inner product calculations, increasing computational costs. In 2020, Meichanetzidis et al. [38] proposed the DisCoCat model, which uses compositional distributional semantics, simplifies syntax into diagrams, encodes specific interactions between words, to construct sentence meaning. In 2021, Kartsaklis et al. [39] launched lambeq, the inaugural advanced Python library dedicated to QNLP, simplifying the process of sentence transformation to quantum circuits. Despite its powerful capabilities, its learning curve is somewhat steep for newcomers to quantum computing and advanced NLP. In 2022, Zhang et al. [40] proposed a model named C-NNQLM. The C-NNQLM uses complex-valued word vectors to uniformly represent words and compound words, reducing the expansion of the word space. However, to capture phase information, the complex-valued model requires more parameters, although parameter sharing can reduce the overall number of parameters. In 2024, Fan et al. [41] proposed a model named QLM-UT that eliminates the need for complex word embeddings, thereby reducing the model's parameters. By simulating the shifting semantics of sentences, QLM-UT can capture the importance of keywords in sentences, enhancing the model's effectiveness in tasks such as text classification. Quantum methods have demonstrated innovation in word embedding and language modeling, particularly in reducing storage space and improving efficiency. However, computational complexity and model complexity remain key issues that need to be addressed.

In terms of semantic modeling, Zhang et al. [42] proposed the QSR model in 2019, which integrates emotion and semantics and trains the density matrix using an improved global convergence algorithm. However, it is limited to emotional phrases of adjectives and adverbs and ignores the emotional effects of verbs and nouns. In 2019, Guo et al. [43] introduced a novel model that integrates GRU and DMATT, referred to as DMATT-BiGRU, which has significantly improved upon models that utilize traditional attention mechanisms. In 2020, Meichanetzidis et al. [44] integrated syntax and statistics through the DisCoCat model to precisely encode and compute semantics. However, the experimental data were synthetic and did not fully map the real challenges of NLP. In 2023, Meichanetzidis et al. [45] proposed an enhanced DisCoCat model, which combines syntactic structure and distributional semantics, using PQCs to represent sentences; however, the generalization ability remains to be verified due to the limited types of test sentences. These studies provide new directions for semantic modeling, but the generalization ability of models to more complex corpora and the processing efficiency of large-scale datasets still need to be explored.

In text classification, Shi et al. [46] proposed two quantum-inspired deep neural network models, ICWE-QNN and CICWE-QNN, in 2021. Among them, quantum-inspired complex-valued word embedding (ICWE) is a novel word embedding method that utilizes the concepts of superposition states and quantum particles in quantum mechanics to represent multiple semantics of words through complex values. In 2022, Yang et al. [47] combined the traditional BERT model with the Quantum Temporal Convolution (QTC) learning framework to enhance text classification performance. QTC processes sequential data compatible with quantum hardware and CPUs, enhancing practicality; however, the security of quantum models still needs to be addressed in actual deployment. In 2024, Li et al. [48] used GPQSA as the quantum version of self-attention. The proposed QSANN demonstrated efficiency and scalability on large datasets, with text classification performance surpassing existing best quantum NLP models and conventional classical self-attention neural networks. These methods perform well in temporal text classification, but the robustness and resource requirements of actual deployment still need to be addressed. The generalization ability of the model and the ability to capture emotions for different parts of speech still need further optimization.

In terms of cross-language learning, Wu et al. [49] proposed a three-dimensional classification system based on quantum theory usage, modeling objectives, and downstream applications in 2021, which has assisted research in the QNLP field. They also introduced their self-developed quantum NLP software toolkit, providing researchers with a resource platform for experimentation and in-depth exploration. In 2023, Li et al. [50] introduced a highly portable quantum language model (PQLM), capable of effortlessly transferring data to traditional devices, protecting private data learning, while securely providing external use. In 2024, Zhao [51] proposed a quantum neural network model QPEN for cross-linguistic sentiment analysis, which shares language knowledge through quantum entanglement without explicit transmission. Quantum methods have demonstrated innovation in cross-linguistic applications; however, their performance and practical effectiveness on large-scale datasets still need further validation.

In the field of quantum-enhanced models and tools, Yan et al. [52] proposed a graph neural network model named Doc2Ket in 2021. This model is inspired by quantum probabilities to capture global structural information between documents. Experiments show that the Doc2Ket model can perform well even with limited training data. In 2022, Qi et al. [53] introduced a hybrid model combining traditional and quantum technologies for voice command recognition, using VQC to enhance QNN and minimize the interference caused by quantum noise. On the Google dataset, the hybrid learning improved the effect, but did not surpass the classical CNN-DNN in noisy environments. In 2022, Di et al. [54] proposed a Quantum-Enhanced Transformer model for sentiment analysis, marking the advent of the Quantum-Enhanced Transformer model. This model combines VQC to reinforce traditional neural networks, especially showing advantages in sequence data processing. However, the training of Quantum LSTM networks takes a long time, indicating that quantum models still need optimization in efficiency. In 2023, Lorenz et al. [55] proposed quantum versions of composite models, creating sentence representations with natural quantum circuit mapping; however, experiments were based on a medium-sized dataset, limiting the scale for large NLP tasks. In 2023, Metawei et al. [56] introduced a topic-aware classifier that utilizes a quantum-conventional hybrid model. The training accuracy of the QNLP model increased by 45%, and the validation accuracy increased by 35%; however, the scale of the experiment was small, and large datasets are needed for verification. In 2024, Munshi et al. [57] compared the performance of QSVC and VQC algorithms in the prediction of chronic heart disease. The QSVC was more robust to noise after tuning but did not detail the specific performance of each algorithm under different noise levels. Quantum-enhanced models have shown great potential in performance improvement, especially when dealing with complex tasks and large-scale data. However, the training efficiency, scalability, and security in actual deployment of the model are still urgent issues that need to be addressed.

The main information of the above 23 papers is shown in Fig. 7. The X-axis marks the timeline, the Y-axis distinguishes academic levels, and the Z-axis stands for the number of citations. Each paper is represented as a sphere, with the size of the sphere directly corresponding to the JCR level of the paper, with larger volume indicating a higher JCR level. The color of the spheres enhances the visual distinction of paper categories for quick identification and comparison. It can be seen from the figure that even though some papers were not published in top journals or conferences, they have received a high number of citations, indicating their unique insights and significant contributions to specific fields and issues.

Quantum computing has shown significant advantages in text generation and classification tasks, but it also faces many challenges. On the one hand, quantum computing can bring revolutionary improvements in computing efficiency and performance. For example, through the SWAP-test algorithm, the similarity between two high-dimensional vectors can be directly obtained in one calculation, and its time complexity is only O(logN), achieving exponential acceleration. In addition, quantum machine learning algorithms use the parallelism and exponential acceleration characteristics of quantum computing to train more efficient classification models in a shorter time. This deep integration and innovation of quantum algorithms and natural language processing technology provides new solutions for text generation and classification.

On the other hand, quantum computing also faces many challenges in text generation and classification tasks. First, the development of quantum hardware technology still faces bottlenecks, and its practical application faces many difficulties, which limits the widespread deployment of quantum computing in practical applications.

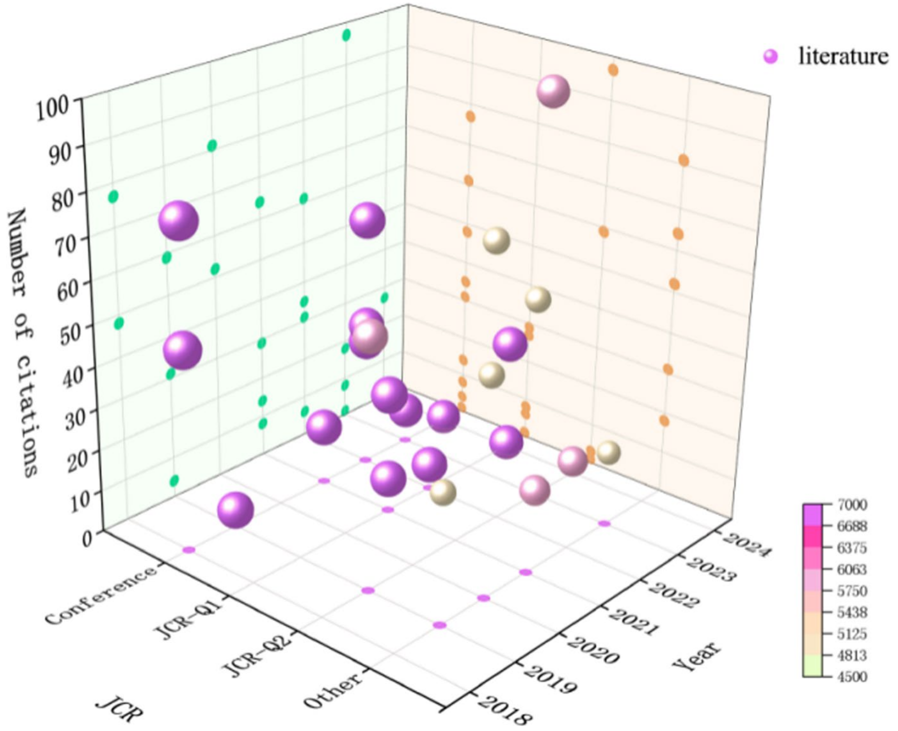


Fig. 7 Text generation and classification literature influence map

Secondly, there are also challenges in the adaptation between quantum algorithms and natural language processing data characteristics. How to efficiently process large-scale and complex text data in a quantum computing environment still needs further research.

3.2 Image-based quantum deep learning

3.2.1 Image generation

In QDL, the image processing workflow typically involves multiple steps, as shown in Fig. 8. Quantum image preprocessing methods are numerous and mainly include image reading and conversion, pixel value mapping, quantum state representation, quantum gate operations, and quantum circuit construction. To convert images to Tensors, one must first use an image processing library (such as Pillow or OpenCV) to load the image file, then perform format conversion, and finally convert the image data into a Tensor object. During the model selection and training phase, whether it is classical methods or deep learning, the process involves picking an appropriate model based on the task and dataset characteristics, and adjusting parameters through iterative optimization algorithms to minimize the

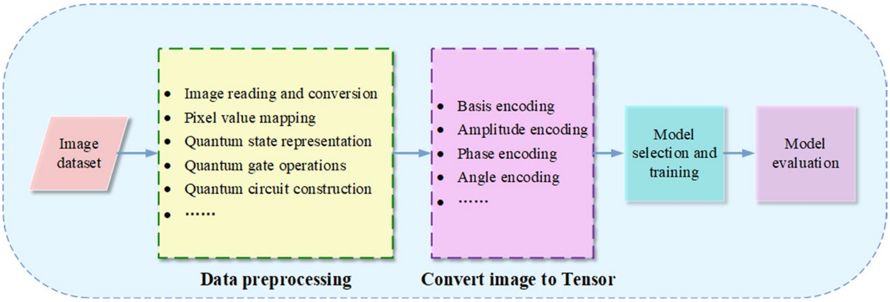


Fig. 8 General flowchart of image-based QDL

loss function, while also employing techniques to prevent overfitting. Model evaluation uses an independent test set to assess model performance, ensuring its generalization ability on unseen data, and often employs cross-validation to provide more stable results.

Within the domain of image generation, QDL has brought notable changes. New types of algorithms such as Quantum Generative Adversarial Networks (QGANs), integrated with the unique advantages of quantum computing, can generate more realistic, delicate, and diverse image samples. The application of algorithms like quantum random walks has further enhanced the creativity and artistry of image generation. These advancements have not only expanded the application scenarios of image generation, such as artistic creation and virtual reality, but also provided richer and more personalized visual experiences.

The following Fig. 9 illustrates the basic structure and operational process of a Quantum Generative Adversarial Network (Quantum GAN).

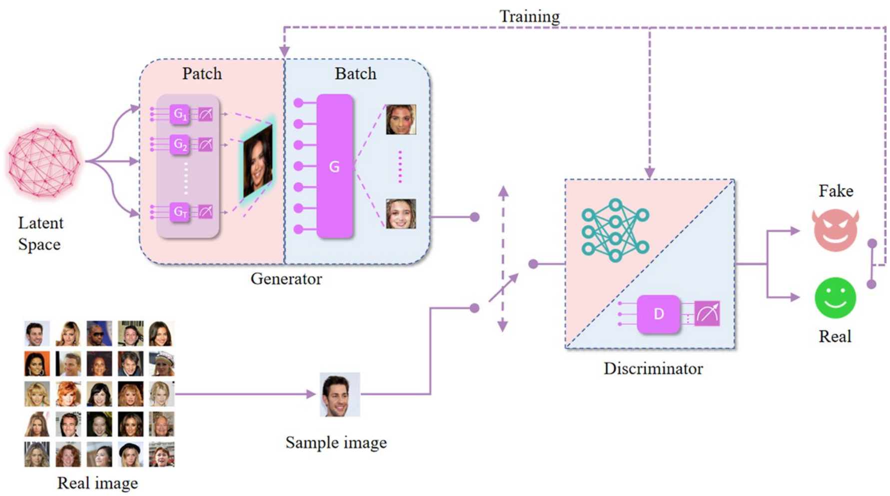


Fig. 9 Quantum generative adversarial network architecture

The quantum generator in the diagram is responsible for generating fake images from random vectors in the latent space. This generation process is implemented through multiple sub-generators (G1, G2...GT), each representing different layers or modules within the network. They collectively work on the input latent vector to produce images with certain dimensional features. The quantum state generated by the generator is shown in Eq. (7).

$$|\Phi_t^{(k)}(z)\rangle = U(\theta_t^{(k)})|z^{(k)}\rangle \tag{7}$$

Here $\theta_t^{(k)}$ is a set of parameters that are optimized during training so that the quantum circuit U can generate the desired quantum state. $|z^{(k)}\rangle$ represents the input quantum state of the kth iteration, usually a simple initial state. $\left|\Phi_t^{(k)}(z)\right\rangle$ represents the quantum state of the given input z at time step t and the k-th iteration (or training round).

The synthesized counterfeit images are then sent to the discriminator, where they are combined with real images. The task of the discriminator is to differentiate between the authenticity of these images. The training process of the hybrid quantum GAN is similar to that of the classical GAN, going through k iterations and gradually adjusting and optimizing the generative and discriminative models with the help of the loss function l.

$$l(\theta,\gamma) = \log\left(D_\gamma(x)\right) + \log\left(1 - D_\gamma\left(G_\theta(z)\right)\right) \tag{8}$$

The discriminator improves the accuracy of its judgment by learning the features that distinguish true and false images. The parameter updates of the generator and discriminator are the following two formulas, respectively.

$$\theta^{(k+1)} = \theta^{(k)} - \eta_G \times \frac{\partial_\theta l_G\left(\theta^{(k)},\gamma^{(k)}\right)}{\partial\theta^{(k)}} \tag{9}$$

Here $\theta^{(k)}$ represents the parameters of the generator at the k-th iteration, η_G is the learning rate for the generator, and l_G is the generator's loss function. The gradient $\frac{\partial L_G}{\partial\theta}$ measures how changes in θ affect the generator's loss, guiding the parameter updates to minimize l_G. This process enables the generator to produce more realistic synthetic images by learning from the feedback provided by the discriminator.

$$\gamma^{(k+1)} = \gamma^{(k)} - \eta_D \times \frac{\partial_\gamma l_D\left(\theta^{(k)},\gamma^{(k)}\right)}{\partial\gamma^{(k)}} \tag{10}$$

Similarly, $\gamma^{(k)}$ denotes the parameters of the discriminator at the k-th iteration, η_D is the learning rate for the discriminator, and l_D is the discriminator's loss function. The gradient $\frac{\partial L_D}{\partial\gamma}$ indicates how changes in γ influence the discriminator's loss, enabling the discriminator to improve its ability to distinguish real images from synthetic ones.

This section primarily reviews key literature on QDL related to image generation. A thorough examination of existing literature was performed, leveraging scholarly databases and indexing services such as Google Scholar, ScienceDirect, Springer, IEEE Xplore, and ACM Digital Library. The search was focused on articles related to "image generation" and "quantum deep learning." The search yielded 17,600 results. We then applied the following inclusion criteria: (1) Articles explicitly addressing Quantum Deep Learning in the context of image generation. (2) Studies published in peer-reviewed journals or presented at reputable conferences. (3) Papers that provided experimental validation or theoretical insights into QDL methods. (4) Research that offered novel contributions, such as innovative algorithms or unique applications. After further screening based on relevance and citation rates, this article reviews 18 papers.

In 2018, Lloyd et al. [58] demonstrated the application of quantum information processors in the generator and discriminator of GANs, introducing the concept of quantum computing to GANs for the first time; the paper assumed that quantum networks could flexibly track necessary gradients, but the validity of this assumption in large networks remains to be verified. In 2018, Dallaire-Demers et al. [59] discussed methods for constructing GANs with quantum circuits, provided a general form for quantum optimization and machine learning gradient computation, and proposed a quantum gradient computation method. In 2019, Hu et al. [60] first demonstrated a QGAN in a superconducting quantum circuit, showing that the quantum state generator (G), after adversarial learning, could replicate the statistical properties of quantum data output by a quantum channel simulator. In 2019, Zeng et al. [61] used quantum circuits to efficiently generate data with specific distributions and verified their learning and inference capabilities on the Bars-and-Stripes dataset; pattern collapse occurred during training, where the generator did not cover all data patterns, a challenge also faced by classical GANs. In 2019, Zoufal et al. [62] used QGANs to learn and approximately map general probability distributions onto quantum states, tested on quantum simulators and experimented on quantum processors provided by IBM Q Experience, and provided public code. In 2020, Situ et al. [63] addressed the vanishing gradient problem by using a QGAN with PQCs to produce separate data points, avoiding the input–output bottleneck of traditional quantum learning algorithms, it directly processes classical inputs and outputs, thus being more feasible in practice and helping to mitigate the vanishing gradient issue. In 2020, Lu et al. [64] implemented various adversarial attacks, verified the vulnerability of quantum classifiers to such attacks, and demonstrated the effectiveness of adversarial training. In 2021, Romero et al. [65] introduced a VQG architecture suitable for generation and classification tasks, designing a groundbreaking approach for merging quantum and classical computational systems. In adversarial learning, it is necessary to ensure that the competitive relationship promotes effective learning. In 2021, Huang et al. [66] used quantum GANs to implement handwritten digit learning generation for the first time on a superconducting processor, supporting high-dimensional image parallel training; but currently, the verification is limited to handwritten digits and grayscale images, and the generalization to complex images remains to be examined. In 2021, Ahmed et al. [67] applied Conditional Generative Adversarial Networks (CGANs) to quantum state tomography (QST) for the first time, and QST-CGAN was able to reconstruct high-fidelity quantum states with much less data than traditional methods. In 2021, Li et al. [68] applied quantum generative models to the field of drug discovery for the first time, the QGAN model effectively learned molecular distributions, and the proposed QGAN-HG model significantly accelerated the training procedure and avoided the prospective vanishing gradient problem of deep neural networks; although the synthetic drug-like molecules can be evaluated using the RDKit package, more experiments are required to verify the precision and dependability of these assessments. In 2022, Rudolph et al. [69] provided the first practical quantum-classical hybrid algorithm for generating high-resolution handwritten digit images, implemented quantum algorithm hardware experiments on a quantum device based on 171Yb+ ion trap, demonstrating the potential of quantum devices in generation tasks. In 2023, Zhou et al. [70] improved the architecture of the quantum generator and introduced an image generation approach that learns discrete distributions. This method simplifies the image generation process by employing a remapping technique to transform the complex distribution with multiple modes of the initial image into a straightforward single-peaked distribution. In 2024, Altares-López et al. [71] implemented the self-generating and training of quantum heuristic classifiers, which can automatically adjust the configuration of the quantum circuit according to the grayscale image dataset, improving the model's adaptability and predictive ability. In 2024, Gong et al. [72] proposed a QCNN rooted in pristine variational quantum circuits, which attained a 98.54% accuracy rate with the MNIST dataset after PCA dimensionality reduction to 8 dimensions, significantly higher than the traditional method's 86.32%; however, they did not discuss in detail the resource consumption of the QCNN in actual operation. In 2024, Feng et al. [73] introduced a Quantum-Inspired Local Aggregation (QILA) module, which mimics the feature mixing capability of quantum networks to improve feature reconstruction performance. By generating a seam mask, the model is guided to perform feature reconstruction within the detected fusion area, reducing artifacts related to misalignment. In 2024, Zhang et al. [74] introduced a Quantum Self-Attention Model (QSAM), which for the first brought into play for image classification tasks, demonstrating its potential as a general-purpose machine learning classifier; the stability under noise or imperfect hardware was not discussed in detail. In 2024, He et al. [75] introduced a method to simplify the preparation process of quantum image representation (BRQI) by optimizing the quantum circuit, significantly reducing the count of auxiliary qubits, thereby reducing complexity and memory space requirements. The paper did not discuss in detail the implementation details and ease of use of the algorithm.

The specific information of the aforementioned literature is depicted in Fig. 10. This three-dimensional plot ingeniously presents the three key attributes of the literature using the X, Y, and Z axes: the timeline (reflected through the year of publication), the academic level (JCR), and the citation frequency. In the illustration, each document is represented as a sphere, the size of which directly reflects the academic level of the literature—that is, the larger the sphere, the higher the academic level; the color of the sphere is categorized according to a preset color scale, helping readers quickly identify different types of the literature.

Quantum image generation technology has shown great potential. Its core advantages include the use of QGAN to efficiently process high-dimensional image data, the ability to learn complex distributions, and even exponential acceleration on certain problems. At the same time, VQG and hybrid quantum-classical algorithms

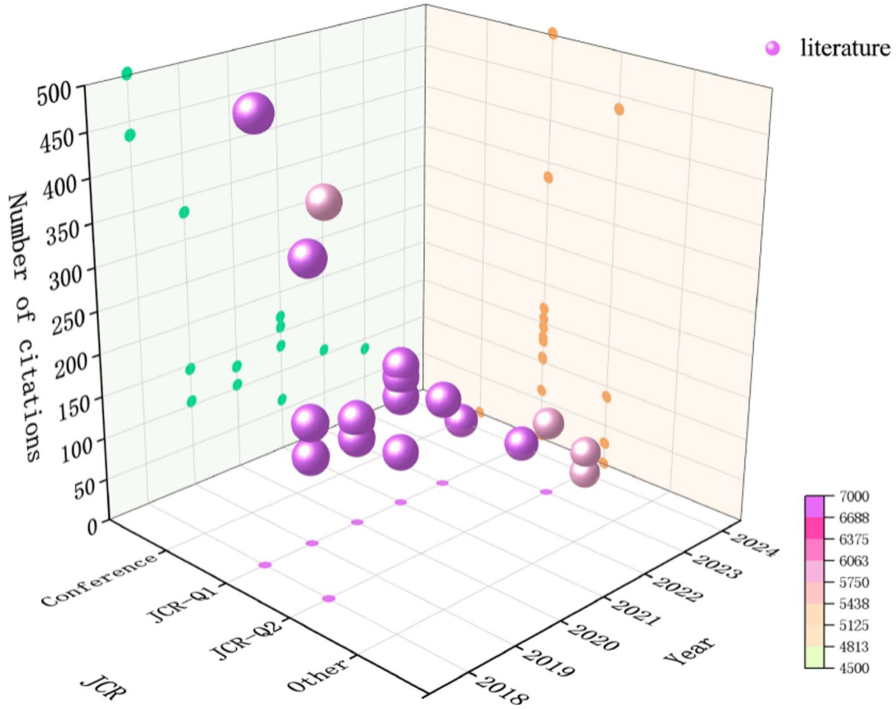


Fig. 10 Image generation literature impact chart

have further expanded the application scenarios of generative models and have made positive progress in experiments from handwritten digits to high-resolution image generation. However, this field still faces many challenges. The imperfections of quantum hardware (such as quantum gate errors and coupling problems between quantum bits) limit the performance of the model; the high cost and technical barriers of building complex experimental environments also hinder large-scale applications. In addition, current methods still need to overcome problems in classical generative adversarial networks such as gradient vanishing and mode collapse, and optimize resource consumption and model stability. In particular, more exploration is needed in the generalization of complex images and adaptation to real-world scenarios.

3.2.2 Image classification and detection

As artificial intelligence technology continues to advance rapidly, from autonomous driving to medical diagnosis to security monitoring, image recognition and detection have become an essential branch. Within the scope of image recognition and detection, QDL achieves efficient processing and high-precision identification of image data through algorithms such as QNNs and QSVMs. The complex nonlinear mapping capability of QNNs allows the algorithm to more accurately capture deep features in images, enhancing the accuracy and robustness of identification.

Quantum image classification and detection primarily use quantum convolution to improve the clarity and contrast of images. The quantum bit encoding method is shown in the following formula:

$$U_E : x \in \mathbb{R} \rightarrow |\phi(x)\rangle = \bigotimes_{i=0}^{n-1} \left(\cos \alpha_i |0\rangle + \sin \alpha_i |1\rangle\right) \tag{11}$$

Among them, the encoding unitary matrix is:

$$U_E = \bigotimes_{i=0}^{n-1} U_i \tag{12}$$

$$U_i = R_Y\left(2\alpha_i\right) = \begin{pmatrix} \cos \alpha_i & -\sin \alpha_i \\ \sin \alpha_i & \cos \alpha_i \end{pmatrix} \tag{13}$$

The encoding method presented in Eq. (14, 15, 16) represents one of the possible approaches for quantum bit encoding. This specific encoding utilizes a parameterized rotation gate $R_Y(2\alpha_i)$ to map classical data into quantum states, enabling the algorithm to capture features effectively. However, it is important to note that alternative encoding schemes, such as amplitude encoding, basis encoding, or other parameterized quantum circuit-based methods, can also be employed depending on the application and computational requirements. The choice of encoding method should align with the specific goals and constraints of the task at hand.

The parameterized quantum circuit of the quantum convolution kernel consists of the following series of unitary operators:

$$U(\vec{\theta}) = U_{l-1}(\theta_{l-1})U_{l-2}(\theta_{l-2}) \cdots U_0(\theta_0) \tag{14}$$

Here $\vec{\theta}$ represents the entire parameter vector, which is a certain parameter. Let the observed expectation be:

$$E(\vec{\theta}) = \left\langle \varphi \middle| U(\vec{\theta})^{\dagger} Z^{\otimes n} U(\vec{\theta}) \middle| \varphi \right\rangle \tag{15}$$

Then, the loss function that needs to be optimized is:

$$\text{Loss}(\vec{\theta}) = f\left(\left\langle \varphi \middle| U^{\dagger}(\vec{\theta}) Z^{\otimes n} U(\vec{\theta}) \middle| \varphi \right\rangle\right) \tag{16}$$

This section mainly reviews the important literature on quantum deep learning related to image generation. Based on the keywords "Image recognition; Detection; Quantum deep learning," a detailed literature search was conducted through academic databases and indexing services such as Google Scholar, IEEE Xplore, ACM Digital Library, Springer, and Science Direct, and 17,000 results were obtained. In the same way as in the previous section, 25 literature sources were selected for review in this article.

In 2019, Li et al. [76] introduced an automatic evolution approach for the architecture of deep CNNs that combine quantum particle swarm optimization (QPSO) with binary data representation, marking the first fully automatic algorithm that reduces manual adjustments, but did not elaborate on parameter selection and performance impact. In 2019, Pang et al. [77] utilized the Quantum Discrete Cosine Transform (QDCT) for signal and image compression, employing the principles of Grover's algorithm to solve complex unstructured search problems through two oracle operations, enhancing search efficiency. In 2021, Liu et al. [78] investigated a geologic interpretation method based on Quantum-Enhanced Deep Learning (QEDL), particularly for lithological analysis from well-logging data. This study marked the first utilization of quantum computing within the realm of lithology interpretation, offering a new perspective and methodology for geophysical data processing. In 2021, Alam et al. [79] introduced a model, using PQCs as filters to extract features from images. The use of trainable quantum filters in Quanvolution could learn optimal features. However, the method proposed in the paper did not demonstrate a quantum superiority in outperforming classical techniques. In 2022, Yang et al. [80] proposed a Hybrid Classical-Quantum Deep Learning Model (HCQDL) to improve the efficiency and accuracy of semiconductor defect detection, exploring PQCs with different expressive and entanglement capabilities, and evaluating their training performance to assess the anticipated advantages. In 2022, Azevedo et al. [81] improved the accuracy of breast cancer detection by utilizing a quantum computing-based transfer learning approach, employing transfer learning to tackle the categorization of complete mammograms into carcinogenic and non-carcinogenic classes. In 2022, Easom-McCaldin et al. [82] proposed a quantum image classification method based on encoding of a single qubit, which attained impressive accuracy scores of 94.6%, 89.5%, and 82.5% on the MNIST, FMNIST, and ORL datasets, respectively, with minimal resource consumption. In 2022, Shahwar et al. [83] employed ResNet34 as a feature extractor and further optimized features through quantum circuits, which not only automatically detected Alzheimer's disease but also significantly accelerated the diagnostic process in clinical settings. In 2022, Ovalle-Magallanes et al. [84] targeted the detection of stenoses in X-ray coronary angiography, enhancing image processing capabilities with the characteristics of quantum computing, with H-CQN achieving significant improvements in accuracy, recall, and F1 score. In 2022, Gong et al. [85] introduced a network intrusion identification framework utilizing the Variational Quantum Neural Network (VQNN), with the VQNN-based Intrusion Detection System (IDS) model showing a precision of 97.21%, superior to other classical IDS models. The KDD CUP99 dataset, while commonly used, may have issues with irrelevant or redundant features. In 2023, Wei et al. [86] used QML to provide efficient image encryption technology for the confidentiality and safety of medical images, leveraging the tenets of quantum mechanics to elevate the unpredictability and intricacy of the encryption scheme, demonstrating superior performance in the classification and prediction tasks of medical images. In 2023, Rashmi et al. [87] introduced a VQC to explore the association between children's BAT temperature and obesity for obesity screening, with customized CNN models achieving accuracy rates of 91.6% and 89.3% in abdominal and neck areas, respectively. In 2023, Amin et al. [88] implemented brain tumor classification with a QCNN, with accuracies of 0.96 on local images and 0.98 on the BRATS-2020 dataset; the study focused on technology but lacked clinical validation. In 2023, Amin et al. [89] proposed two models, J. DCNN and J. QCNN, for analyzing video anomalies, with J. QCNN achieving 99% accuracy and J. DCNN at 97%; although it improved the efficiency of surveillance anomaly detection, it did not elaborate on robustness against issues such as lighting changes and occlusions. In 2023, Kulkarni et al. [90] targeted the diagnosis of pneumonia through chest X-ray analysis using a hybrid model integrating VQC into a classical network, with significant results verified after 30 rounds of experiments, which is rare in hybrid networks. In 2023, Fan et al. [91] introduced a model applying quantum computing to obtain advanced key traits from Earth Observation (EO) data for image classification, studying the model's performance under different noise conditions, including data noise and model noise, enhancing robustness for practical applications. In 2024, Veeragandham et al. [92] introduced a model based on Coot Political Optimization (CPO), using the Pyramid Scene Parsing Network (PSPNet) for image preprocessing and segmentation, enhancing the accuracy of peanut weed identification, with excellent performance on the dataset. But the model did not consider soil properties, which may affect the classification accuracy under complex soil conditions. In 2024, Das et al. [93] proposed an innovative use of a named SQNN for surface crack detection, introducing entanglement feature mapping in SQNNs to improve performance, surpassing traditional methods and reducing training time. In 2024, Hassan et al. [94] integrated a QCNN with an adapted pre-trained ResNet (50) model, resulting in MQCNN achieving 99.6% accuracy on the medical MNIST dataset, outperforming both the pre-trained ResNet (50) and the QCNN models, thereby improving medical image classification performance. In 2024, Sinha et al. [95] discussed various applications of quantum computing in machine learning, including KNN, SVM, clustering algorithms, neural networks, and decision trees, and explored their equivalent technologies in the quantum domain. In 2024, Çavşi et al. [96] applied QDL models for gender recognition, with a slightly higher accuracy rate than classical CNN models but longer processing time. In 2024, Rao et al. [97] proposed MMS and MSMS quantum classifiers for detecting respiratory lung diseases based on chest radiographs, with the presented models achieving a training accuracy of 98.9% and a test accuracy of 98.1% on the CRD dataset; however, the diversity of the dataset was not fully discussed. In 2024, Chen et al. [98] discussed the implementation of QML in image recognition, drug efficacy prediction, and network security, outlining the difficulties encountered by QML. In 2024, Munshi et al. [99] applied QML technology to heart disease prediction, comparing the performance of QSVC and VQC, with a comprehensive evaluation of model performance through various assessment indicators. In 2024, Mir et al. [100] introduced an efficient quantum bit-plane delegate for grayscale images, EQBRGI, and for color images, EQBRCI, enhancing the efficiency of quantum computing for image processing, and validated the quantum circuit for 2x2 images through QISKIT simulation. In 2024, Venkatesh et al. [101] proposed a QCNN model (HQCM) for detecting multiple nutrient deficiencies and detection of stress in foliage, which analyzed plant leaf images to identify poly-nutrient stress and levels of nutrient stress, demonstrating high accuracy on the Xception, DECM, DWC, and ResNet50V2 datasets, surpassing traditional models.

Figure 11 cleverly utilizes three coordinate axes to display the three core features of the literature: publication year, academic level (JCR classification), and citation

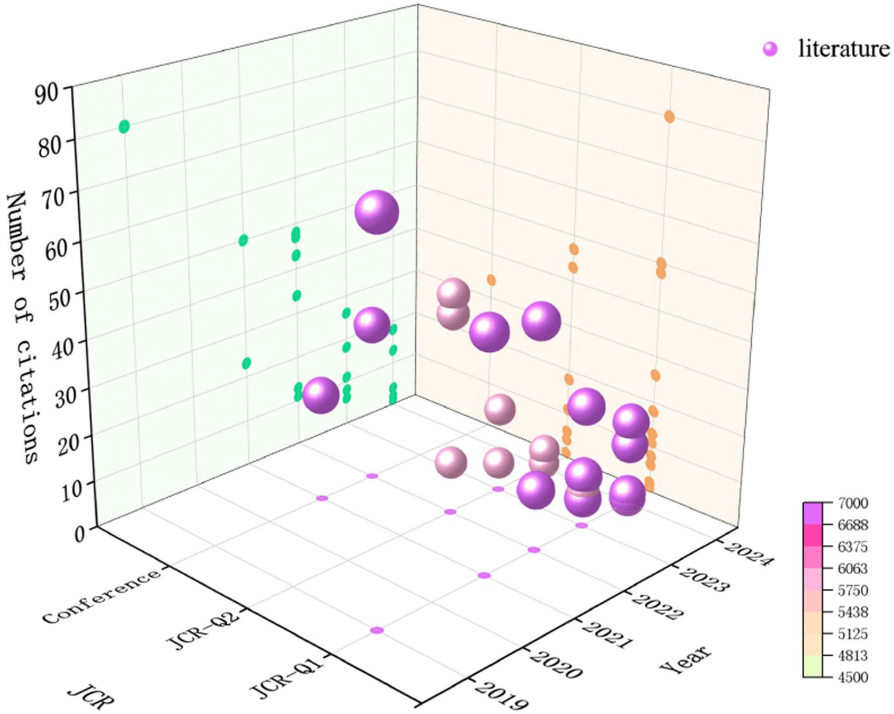


Fig. 11 Image classification and detection literature influence chart

count. Each document is presented in the form of a sphere, where the size of the sphere represents academic level, and color coding helps to quickly distinguish document categories.

Quantum computing provides significant advantages in image classification and detection, especially when dealing with complex data and high-dimensional features. Quantum computing utilizes the properties of quantum superposition and entanglement to process large amounts of data in parallel, thereby improving computational efficiency. Some quantum optimization algorithms, such as QPSO, can automatically adjust the structure of deep neural networks, reduce the workload of manual parameter tuning, and thus improve the accuracy of image classification models. In addition, QCNN and QEDL have achieved good results in feature extraction and image classification tasks, especially in specific fields such as medical images and remote sensing images, demonstrating potential.

Although quantum computing has shown potential in image classification and detection, it also faces many challenges. Firstly, the hardware limitations of quantum computers, such as the number of qubits and noise issues, still constrain the breadth and stability of their applications. In addition, most of the existing quantum image processing algorithms have not been fully validated in practical applications, especially in complex environments such as lighting changes, occlusion, etc. The robustness and generalization ability of the models have not been effectively evaluated. More importantly, the computational complexity of quantum algorithms is high, and the training process consumes a lot of time. However, existing quantum computer resources are limited, which still makes it difficult to process large-scale datasets. Therefore, how to solve the noise problem of quantum computing, improve training efficiency, and achieve practical applications of quantum computing are still urgent issues that need to be addressed.

3.3 Multimodal-based quantum deep learning

As application scenarios become increasingly complex, data from a single modality often falls short of meeting the intelligent system's need for a comprehensive understanding of the world. Multimodal learning, by integrating various types of data, provides machines with richer and more integrated information. QDL, as a cutting-edge field that combines quantum computing with deep learning, has demonstrated immense potential for processing high-dimensional data and executing complex algorithms.

As follows Fig. 12, multimodal sentiment analysis systems typically adopt a comprehensive framework. The framework initially processes textual data through a text feature extractor, which identifies emotion-related vocabulary and semantic features. Subsequently, a visual feature extractor analyzes visual elements in videos, such as facial expressions and body language, to gather non-verbal emotional cues. An auditory feature extractor then processes audio data, capturing vocal characteristics such as pitch and volume that reflect the speaker's intonation and emotional state. These diverse modal features are input into a fusion model that integrates the information using specific strategies to elevate the exactness and reliability of sentiment detection. Finally, the system provides a prediction of the emotional tendency of the input across text, visuals, and audio, classifying it as positive, neutral, or negative. This multimodal approach more comprehensively captures human emotional expression as it takes into account not only textual information but also the rich content from visual and auditory channels.

Multimodal information fusion falls into three categories: early, middle, and late fusion. Early fusion, or feature-level fusion, combines data across various forms after extraction but before classification. It leverages complementary data for

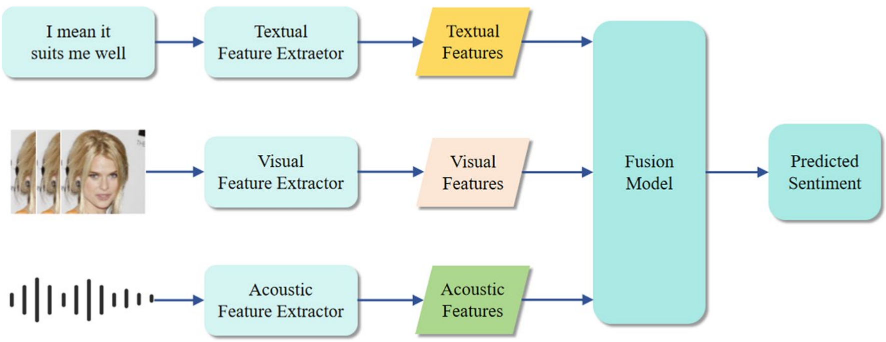


Fig. 12 General framework of a typical multimodal sentiment analysis system

enhanced model expression but can result in high-dimensional vectors with redundant information. Middle fusion, or intermediate-layer fusion, integrates data at the model's mid-level, offering flexibility but requiring complex implementation. Late fusion, or decision-level fusion, aggregates decisions made from each modality separately. It is robust and scalable, as system performance is not severely impacted by the loss or degradation of data from one modality. However, it might miss capturing intermodal relationships and complementarity.

In the fusion process, voting mechanisms are widely used in the merging strategy of multimodal features. Voting mechanisms can be used alone or combined with other merging strategies to form more complex fusion methods, as shown in Fig. 13. The parameters Rx, Ry, etc., in the figure can be changed as needed.

Assuming the text classifier predicts the probability of a category as P_T with weight λ_T, and the image classifier predicts the probability of a category as P_I with weight λ_I, then the fused prediction result is:

$$P = \lambda_T P_T + \lambda_I P_I \tag{17}$$

When fusing multimodal probability information, different weights should be assigned based on the reliability and accuracy of each modality. For instance, if image data are more accurate, it should be given a higher weight. Weights can either remain constant or be adjusted dynamically depending on the model's performance.

Multimodal data fusion, as a technology that integrates multiple sources of information, has shown its unique advantages in many fields. In this article, we conducted an exhaustive literature search using the keywords "Multimodal data fusion; Quantum deep learning" through academic databases and indexing services such as

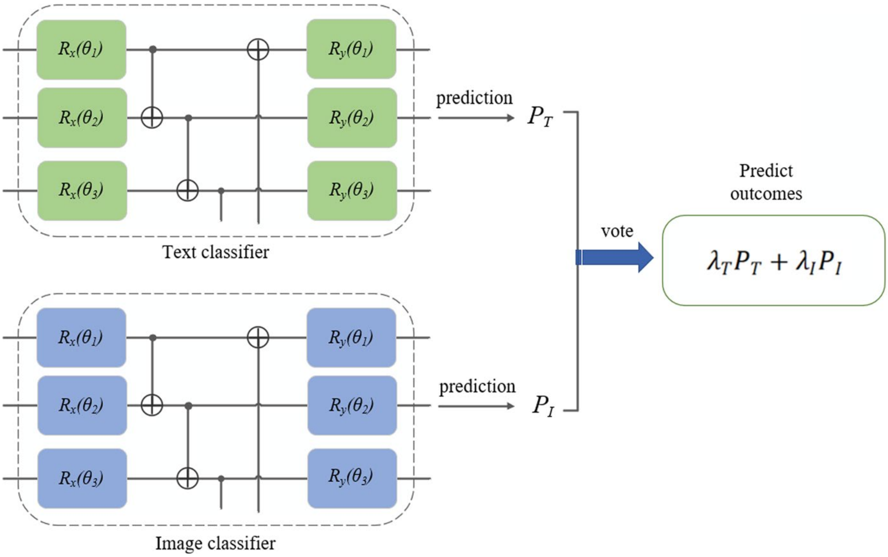


Fig. 13 Multimodal information fusion

Google Scholar, IEEE Xplore, ACM Digital Library, Springer, and Science Direct, and obtained 18,200 results. Then, we applied the following inclusion criteria: (1) Articles that explicitly involve quantum deep learning in the context of multimodality. (2) Research published in peer-reviewed journals or presented at well-known conferences. (3) Papers that provide experimental verification or theoretical insights for QDL methods. (4) Research that provides novel contributions, such as innovative algorithms or unique applications. Eighteen references were selected, and the following categories are reviewed for each literature.

3.3.1 Text processing

Text processing serves as the foundation for social media analysis by extracting meaningful information from textual data. By merging diverse modalities such as text and images and utilizing multimodal QDL algorithms, user emotional states and needs can be analyzed more accurately. In 2018, Zhang et al. [102] introduced a system for analyzing emotions across different media types that employ quantum theory to process and analyze emotional information across different modalities, such as text and images. The model uses density matrices to represent multimodal content, better encoding semantic dependency relationships and their probability distribution information. In 2020, Yu et al. [103] proposed a multimodal knowledge source representation method that describes images through visual graphs, semantic graphs, and factual graphs, which unify graph-structured representations of various modalities for easy cross-modality relational reasoning. In 2020, Zhang et al. [104] utilized a Quantum Multimodal Network (QMN) combined with quantum theory and LSTM to model the dynamics of conversational emotions. Tested on the MELD and IEMOCAP datasets, the performance surpassed baseline and advanced models. In 2021, Li et al. [105] introduced a quantum theory-based multimodal fusion platform for video emotion analysis, implementing a neural network utilizing complex numbers to simulate frameworks in quantum theory. In 2021, Li et al. [106] offered a quantum-inspired dialog emotion recognition framework, analogous to quantum measurement and emotion recognition, using complex-valued operations to construct a quantum layer that supports context modeling and multimodal fusion. In 2021, Gkoumas et al. [107] introduced a quantum cognitive theory-based decision fusion model for predicting emotional judgments in videos, considering "quantum-like" biases in human decision-making processes, and using the concept of incompatibility in quantum theory to handle conflicts in emotional judgments between different modalities, such as language, vision, and audition. In 2021, Gkoumas [108] proposed a quantum probability neural model capable of simulating human decision-making processes under uncertainty, including rational and irrational decisions. The model provided a comprehensive empirical comparison of existing multimodal fusion strategies and proposed more effective fusion models based on quantum-inspired components. In 2021, Liu et al. [109] offered a multitask learning framework based on quantum probability theory (QPM) to address the issues of multimodal sarcasm and sentiment detection. By extracting multimodal features through quantum incompatible measurements, this helps capture the interdependencies between detecting sarcasm and emotional tone tasks. In 2023, Gandhi et al.

[110] thoroughly discussed different multimodal fusion techniques, including their advantages and limitations, assisting researchers in choosing appropriate fusion strategies for their investigations. They proposed interdisciplinary applications and future research directions, such as mental health prediction and deception detection. In 2023, Zhu et al. [111] offered an extensive review of the multimodal sentiment analysis field. In 2023, Qu et al. [112] offered a model named QMFND, which showed excellent performance when tested on social media datasets; however, the model's capacity to handle different forms of fake news, such as deepfakes and misleading headlines, remains to be verified. In 2023, Al-Hawawreh et al. [113] introduced a novel privacy-conscious framework for storing, processing, and analyzing IoMT data without compromising data privacy. He suggested using QDL networks for intrusion detection. Experiments conducted using the WUSTL-EHMS-2020 and ICU datasets validated the effectiveness of the introduced framework in detecting network attacks. In 2024, Tiwari et al. [114] addressed multimodal sentiment and sarcasm detection by proposing a novel QFNN. The network employs a Fuzzifier and Defuzzifier to perform complex-valued fuzzy encoding and decoding of sentiment and sarcasm features, enhancing the expressive power of the features.

This section reviews the application of multimodal quantum deep learning (QDL) in text processing, covering tasks such as sentiment analysis, conversation emotion recognition, and sarcasm detection. Studies have shown that QDL has potential advantages in cross-modal information fusion and uncertainty modeling, but it still has problems such as high computational cost, insufficient interpretability, and limited generalization ability. In addition, the model's adaptability and adversarial attack defense capabilities in real social media environments still need to be improved. Future research can further optimize the quantum information fusion mechanism, improve computational efficiency, enhance model transparency, and explore more robust QDL architectures to improve its practicality.

3.3.2 Image processing

Image processing plays a vital role in the medical field by extracting and analyzing visual information from medical images. By combining medical imaging with electronic health record text data and leveraging multimodal QDL algorithms, disease diagnosis can be conducted more accurately, providing patients with more effective treatment plans. In 2018, Rajalingam et al. [115] introduced a deep learning-based multimodal medical image amalgamation method that uses a medical image pyramid for multi-scale fusion, enhancing the reliability of the fusion process and better aligning with human visual perception. In 2019, Liu et al. [116] leveraged the parallelism and entanglement traits of quantum computing to enhance the effectiveness of image processing tasks, capturing key features of source images through Quantum Wavelet Transform (QWT), and using the FRQI model to represent source digital images for efficient unified operations. In 2020, Tan et al. [117] introduced a novel boundary measurement-based PCNN model that optimizes high-frequency subband fusion and enhances image gradient extraction, adapting to multi-scale structures. The efficiency of this fusion algorithm was comprehensively verified through evaluation metrics such as entropy, standard deviation, NMI, SS, and VIF. In 2022, Dogra et al. [118] introduced a multi-tier filtering-based image fusion technique that efficiently combines data from CT and MR images into a unified image. Experimental findings indicated that the suggested approach was subjectively more effective than current leading methods. In 2023, Cui et al. [119] summarized various multimodal fusion techniques, discussed the application of these multimodal fusion methods in disease diagnosis and prognosis, and provided corresponding performance evaluations. Special attention was given to how to handle heterogeneous data from different modalities, which is a very important issue in multimodal learning.

In recent years, the integration of multimodal medical imaging with electronic health record data using deep learning and quantum computing has significantly improved disease diagnosis and treatment plans. Methods such as medical image pyramids for multi-scale fusion, Quantum Wavelet Transform (QWT), and boundary measurement-based models have enhanced image reliability and feature extraction, thus improving diagnostic accuracy. However, challenges remain in handling the heterogeneity of data from different modalities, as effectively fusing diverse image types (e.g., CT and MRI) and non-image data are complex. Despite advancements, there is still a need for more robust algorithms that can seamlessly integrate these data types, maintain consistency across varying scales, and optimize computational efficiency for real-time clinical applications.

Figure 14 shows the relevant information for all the literature in this section. The three-dimensional plot uses the X, Y, and Z axes to correspond to the three core

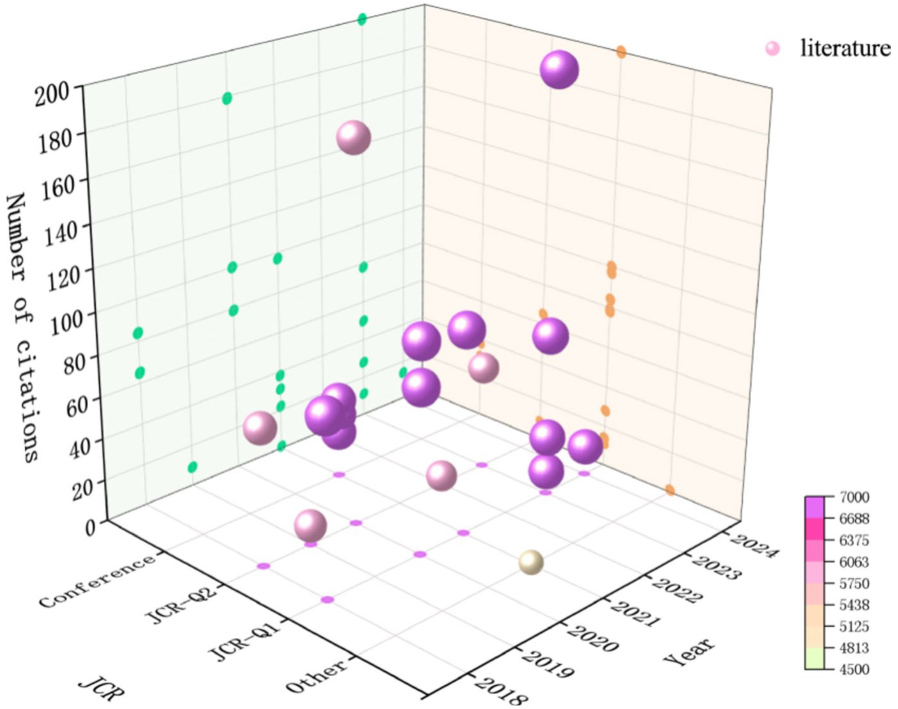


Fig. 14 Multimodal literature influence map

characteristics of the literature: the timeline (year of publication), academic level (JCR), and academic impact. In the illustration, each document is vividly transformed into a sphere, with the size of the sphere directly mapping to the academic level of the literature—the larger the size, the higher the academic level. The spheres are color-coded according to preset color schemes to allow readers to quickly distinguish and identify different categories of the literature.

4 Comparative analysis of model results

4.1 Introduction to classic datasets

Classic public text datasets include SQuAD, QQID, Yelp Reviews Dataset, IMDB, IWSLT, TREC, TAC, etc. These datasets have their own characteristics and play an important role in different tasks. In particular, the SQuAD dataset from Wikipedia has become one of the standard benchmark test sets for question–answering systems and reading comprehension tasks. In the quantum computing environment, the use of these datasets needs to consider the quantum state encoding method of text data by quantum algorithms and how to efficiently use quantum computing characteristics for optimization. For example, quantum state representation can provide exponential acceleration for the processing of large-scale text data, but how to effectively reduce the complexity of quantum state preparation remains a research focus.

Commonly used image datasets include MNIST handwriting dataset, CIFAR-10 dataset, ImageNet dataset, RetinaMNIST, etc. The MNIST handwritten digit recognition dataset has become the cornerstone for verifying the performance of new algorithms and models in the field of quantum deep learning with its simple structure and broad application basis, and has promoted the initial exploration of quantum image processing technology. However, in the quantum environment, the use of the MNIST dataset needs to consider the limitations of the number of quantum bits on image representation, as well as how quantum algorithms deal with feature extraction and classification problems of low-resolution images. The CIFAR-10 dataset provides richer image categories and complexity, and is an important evaluation benchmark for quantum deep learning models in complex image classification tasks. However, compared with MNIST, CIFAR-10 requires higher quantum state expression capabilities in quantum computing, especially how to optimize quantum circuit design when processing high-dimensional data, which is a key challenge facing researchers. In addition, the application of medical-related datasets such as RetinaMNIST in quantum computing also needs to consider the high fidelity and privacy protection of medical images, which provides a new research direction for quantum technology.

Commonly used multimodal datasets include CMU-MOSEI, IEMOCAP, Multi-Modal Sarcasm Detection in Twitter, MELD, MedMNIST, etc. CMU-MOSEI is one of the largest sentiment analysis and sentiment recognition datasets at present, which contains two labels: sentiment and emotion. The sentiment label is seven categories from negative to positive, while the emotion label contains six basic emotion categories. The emotion intensity of each category varies in the range of [-3, 3], providing more detailed annotations for sentiment analysis. In a quantum environment, the application of these multimodal datasets needs to solve the problem of consistency in feature representation between different modalities and enhance the efficiency of cross-modal feature fusion through quantum superposition and entanglement. For example, the fusion of speech and text modalities in the IEMOCAP dataset requires the use of quantum states to represent the correlation between audio signals and language features and optimize the context modeling capabilities of quantum models. At the same time, the application of medical-related multimodal datasets such as MedMNIST in quantum computing must comprehensively consider data heterogeneity, privacy protection, and the limitations of quantum computing resources.

4.2 Evaluation metrics

Model evaluation is a vital step in measuring the impact of an algorithm. Different task characteristics often require different assessment criteria to comprehensively and objectively reflect the performance of a model. This article organizes and analyzes the model evaluation metrics mentioned in a series of research literature, focusing particularly on the four core metrics: MRR, MAP, F1 score, and accuracy.

MRR is obtained by taking the reciprocal of the position of the initial relevant outcome for each query and then averaging the results across all queries. A higher MRR value suggests that the model is more effective at positioning relevant results toward the top of the search list. The formula is as follows:

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\mathrm{rank}_i} \tag{18}$$

where $|Q|$ is the size of the query set, and $rank_i$ is the ranking of the first correct answer for the i-th query.

MAP is a metric for evaluating the effectiveness of retrieval or classification tasks. It is obtained by calculating the average precision (AP) for each query and then taking the average of all queries. MAP effectively assesses the model's precision in predicting all categories. A high MAP value usually indicates that the model has stronger discrimination ability and higher predictive accuracy overall. The formula is as follows:

$$MAP = \frac{1}{|Q|} \sum_{i=1}^{|Q|} AP_i \tag{19}$$

where $|Q|$ is the size of the query set, and AP_i is the average precision of the i-th query.

The F1 score represents the harmonic average of precision and recall. It is crucial for balancing the relationship between the two, especially in classification tasks where both false optimistic and false pessimistic are sensitive. A higher F1 score indicates a more robust model performance in classification tasks.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{20}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{21}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{22}$$

Accuracy directly reflects the degree of match between the model's predicted outcomes and the actual labels.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{23}$$

4.3 Comparison of model results

In this section, I will roughly divide the cited literature into three categories: models that process text, images, and multimodal data, and organize the traditional models, quantum models, data sets, and evaluation indicators used by each type of model to form Tables 3, 4, and 5. Through these tables, the performance of traditional deep learning models and quantum models can be clearly compared. For example, in the "Models" column in the table, CNM represents the traditional model, and QMWF-LM (Zhang et al., 2018) is the quantum model, followed by the relevant authors and publication years.

The quantum deep learning models have shown good performance on multiple text datasets, and some of them are particularly outstanding on specific datasets, as shown in Table 3. From the table, we can see that the QLM-UT model has significant advantages in complex semantic modeling and keyword weight capture, especially in tasks that require accurate understanding of contextual semantics. In the text classification task, the quantum model QSANN has better accuracy than the classical model CSANN on the Yelp, IMDb, and Amazon datasets, especially on the Yelp dataset, with an accuracy of 84.79%. This shows that the quantum model has stronger capabilities in feature extraction and multi-semantic information processing.

Based on the above analysis, future research should further focus on the following points: (1) Explore more efficient quantum circuit design and parameter optimization methods to reduce the training complexity of quantum models and improve their stability. (2) Verify the performance of quantum models on larger and more diverse datasets and apply them to real-world scenarios to evaluate their practicality. (3) Verify the performance of quantum models on larger and more diverse datasets and apply them to real-world scenarios to evaluate their practicality.

The performance of quantum deep learning models on image datasets demonstrates the efficiency and accuracy of these models as shown in Table 4. On the MNIST dataset, quantum models such as QSAM and QCNN achieved 99.91%

Table 3 Results of QDL model based on text dataset

Models	Dataset	MRR	MAP
C-NNQLM	TREC-QA	0.712	0.820
CNM	TREC-QA	0.859	0.770
QMWF-LM (Zhang et al. 2018)	TREC-QA	0.814	0.752
QLM-UT (Fan et al. 2024)	TREC-QA	0.774	0.857
C-NNQLM	WIKI-QA	0.671	0.687
CNM	WIKI-QA	0.675	0.686
QMWF-LM (Zhang et al. 2018)	WIKI-QA	0.695	0.710
QLM-UT (Fan et al. 2024)	WIKI-QA	0.697	0.712
C-NNQLM	Yahoo-QA	0.671	0.805
QLM-UT (Fan et al. 2024)	Yahoo-QA	0.697	0.865
Models	Dataset	F1	Acc
CSANN	Yelp	N/A	83.11%
QSANN (Li et al. 2024)	Yelp	N/A	84.79%
CSANN	IMDb	N/A	79.67%
QSANN (Li et al. 2024)	IMDb	N/A	80.28%
CSANN	Amazon	N/A	83.22%
QSANN (Li et al. 2024)	Amazon	N/A	84.25%
QIN	MELD	0.662	67.90%
QMN (Zhang et al. 2020)	MELD	0.729	75.60%
CE-Sup	SST	N/A	82.60%
CE-Mix	SST	N/A	83.30%
ICWE-QNN	SST	0.8390	84.20%
CICWE-QNN (Shi et al. 2021)	SST	0.8360	85.00%
CaptionRep BOW	SUBJ	0.7740	N/A
DictRep BOW	SUBJ	0.9070	N/A
ICWE-QNN	SUBJ	0.9260	90.90%
CICWE-QNN (Shi et al. 2021)	SUBJ	0.9320	92.00%

and 98.66% accuracy, respectively, significantly higher than the 91.76% of the classical CNN. In addition, in the RetinaMNIST dataset, QCNN achieved an accuracy of 94.05%, far exceeding the 81.40% of the classical CNN. These quantum models use the superposition and entanglement properties of quantum states to greatly improve the ability and efficiency of image processing. Especially in medical datasets such as BRATS-2020 and MedNIST, quantum models such as J. Qnet and MQCNN performed well, with accuracies of 100.0% and 99.60%, respectively, demonstrating their potential in high-complexity tasks. In addition, methods such as quantum generative adversarial networks (QGAN) showed near-perfect performance in synthetic datasets, further demonstrating the advantages of quantum methods in optimization and generalization capabilities.

Table 4 Results of QDL model based on image dataset

Models	Dataset	F1	Acc
CNN	MNIST	N/A	91.76%
QCNN (Gong et al. 2024)	MNIST	N/A	98.66%
QCNN (Kerenidis et al. 2019)	MNIST	N/A	95.90%
QNNs (Henderson et al. 2022)	MNIST	N/A	95.00%
Noise-free QCNN (Wei et al. 2022)	MNIST	N/A	96.30%
EQIC-SQE (Easom-McCaldin et al. 2022)	MNIST	N/A	94.60%
QSAM (Zhang et al. 2024)	MNIST	N/A	99.91%
QAML (Lu et al. 2019)	MNIST	N/A	98.00%
CNN	RetinaMNIST	N/A	81.40%
QSAM (Zhang et al. 2024)	RetinaMNIST	N/A	82.22%
QCNN (Gong et al. 2024)	RetinaMNIST	N/A	94.05%
EQIC-SQE (Easom-McCaldin et al. 2022)	RetinaMNIST	N/A	89.50%
CAE + QNN (Alam et al. 2021)	RetinaMNIST	N/A	92.33%
CNN	POF Hospital	0.8721	87.20%
QML (Amin et al. 2022)	POF Hospital	0.9400	94.00%
J. Qnet (Amin et al. 2023)	POF Hospital	0.8800	89.00%
CNN	AT&T	N/A	80.32%
EQIC-SQE (Easom-McCaldin et al. 2022)	AT&T	N/A	82.50%
QCNN (Cavsi et al. 2024)	AT&T	N/A	92.00%
CNN	Bars-and-Stripes	N/A	90.10%
GAQC (Zeng et al. 2018)	Bars-and-Stripes	N/A	99.97%
QGAN (Situ et al. 2020)	Bars-and-Stripes	N/A	99.99%
QTML (Shahwar et al. 2022)	Alzheimer’s disease	N/A	97.00%
HCQNN-AD (Shahwar et al. 2022)	Alzheimer’s disease	N/A	99.10%
ETL-QVC (Amin et al. 2022)	BRATS-2020	N/A	90.91%
J. Qnet (Amin et al. 2023)	BRATS-2020	N/A	100.00%
CNN	Groundnut	0.8932	90.23%
HQCM (Venkatesh et al. 2024)	Multi-nutrient stress	0.9779	97.79%
CNN	MedNIST	0.8947	90.61%
MQCNN (Altares-Lopez et al. 2024)	MedNIST	0.9970	99.60%

Although quantum deep learning models have shown excellent performance in image classification and detection tasks, future research still needs to solve some key problems to promote their development. First, in order to improve the scalability of quantum models on large-scale real datasets, researchers should further optimize quantum circuit design and reduce the limitations of quantum hardware on computational complexity. Second, although existing models perform well under noise-free or low-noise conditions (such as Noise-free QCNN (Wei et al., 2022)), their robustness in high-noise environments still needs to be further explored. In addition, in the field of medical image analysis, how to better integrate quantum models with multimodal data has become an important direction

Table 5 Results of QDL model based on multimodal dataset

Models	Dataset	F1	Acc
LMF	CMU-MOSEI	0.8057	80.64%
QMF (Li et al., 2021)	CMU-MOSEI	0.8079	47.88%
QCMDF (Gkoumas et al., 2021)	CMU-MOSEI	0.9110	80.69%
MRRF	CMU-MOSI	N/A	77.46%
QMF (Li et al., 2021)	CMU-MOSI	0.7962	33.53%
QCMDF (Gkoumas et al., 2021)	CMU-MOSI	0.8450	84.60%
DM-QIMF model	IEMOCAP	0.6020	62.50%
QMN (Zhang et al., 2020)	IEMOCAP	0.6230	64.80%
QMNN (Li et al., 2021)	IEMOCAP	0.5988	60.84%
DM-QIMF model	MELD	0.7040	72.00%
QMNN (Li et al., 2021)	MELD	0.5800	60.81%
QMSA (Zhang et al., 2018)	Flickr	0.9301	93.14%
QMSA (Zhang et al., 2018)	Getty images	0.8969	88.24%
FNN	MUStARD	0.6700	67.26%
QPM (Liu et al., 2021)	MUStARD	0.7753	N/A
QFNN (Tiwari et al., 2024)	MUStARD	0.6888	68.88%
A-MTL	Memotion	0.6026	N/A
QPM (Liu et al., 2021)	Memotion	0.6139	N/A
QFNN (Tiwari et al., 2024)	Memotion	0.5177	51.77%
XLNet	Gossip	0.9290	88.40%
QMFND (Qu et al., 2023)	Gossip	0.9280	87.90%

for improving diagnostic accuracy and efficiency. Finally, future research should also strengthen the combination of classical and quantum models, further improve performance through hybrid architectures (such as HCQNN-AD), and explore the application potential of quantum models on resource-limited hardware. By solving these problems, quantum deep learning is expected to be widely used in more practical scenarios.

As shown from Table 5, quantum deep learning models based on multimodal datasets have shown strong performance in multiple tasks. QCMDF achieved an F1 score of 0.9110 and an accuracy of 80.69% on the CMU-MOSEI dataset, which is excellent. In addition, QMSA also performed very well on the Flickr dataset, reaching an accuracy of 93.14%. These results demonstrate the potential of quantum deep learning in multimodal data processing, especially in the image and text fusion task.

Future research can further focus on improving the stability and generalization ability of quantum deep learning models on various datasets, especially in complex tasks such as multimodal sentiment analysis and social media analysis. Some models (such as QMF and QMN) still have room for improvement in accuracy, and their performance may be improved in the future by optimizing the quantum network architecture and improving the training algorithm. In addition, with the development of quantum computing hardware, the efficiency and scalability of quantum deep learning models in processing large-scale data will become an important research direction.

5 Summary and outlook

This paper reviews the application of QDL in text-based, image-based, and multimodal data processing, demonstrating the potential of combining quantum computing with deep learning. In text-based QDL, applications include text generation and classification, where quantum algorithms leverage superposition and entanglement to process complex language structures efficiently. For image-based QDL, research has explored image generation and image classification and detection, utilizing quantum computing to enhance feature extraction and improve model performance. In multimodal-based QDL, advancements focus on text processing and image processing. These applications highlight the capability of QDL in integrating information from different data sources, improving decision-making accuracy, and enhancing fusion techniques for better data representation. Despite these advancements, challenges remain in effectively handling heterogeneous data, optimizing quantum algorithms for large-scale applications, and ensuring practical implementation in real-world scenarios. However, this field still faces technical challenges and open issues, and requires continuous exploration and development.

(1) **The limitations of quantum hardware.**
Implementing QDL algorithms requires substantial hardware support. However, the manufacturing costs of quantum computing hardware are high, and they require highly specialized technologies and equipment for maintenance. For instance, superconducting quantum chips need to operate at extremely low temperatures to maintain the coherence of qubits. These factors make the hardware implementation of QDL algorithms challenging and limit their promotion and popularization in practical applications. Currently, quantum computers mainly include IBM's quantum computers, China's "Origin Wukong," and China's "Jiuzhang" optical quantum computing prototype. The number of qubits available in current quantum computers is still very limited, and there is a relatively high error rate. Due to the diversity and complexity of quantum computing hardware, existing QDL algorithms often struggle to perfectly match specific hardware platforms, preventing the full performance of the algorithms. Current quantum computing hardware still faces issues with poor stability and high noise, which also affect the implementation effects of QDL algorithms on hardware. Future research should focus on implementing QNNs on larger-scale quantum hardware platforms to develop more complex and influential practical applications.

(2) **Quantum noise processing.** Quantum noise processing has made progress in quantum computing, but its complexity, error correction difficulties, and technical limitations still hinder performance improvement. Noise reduction strategies are diverse, but the variety and complexity of noise make it difficult to process, affecting computational accuracy and algorithm application. Noise comes from both internal and external sources, intertwined and difficult to separate. Quantum error correction codes are effective but consume a lot of resources and have limited efficiency. Current technology relies on hardware optimization, with limited

effects. This situation directly affects the architectural design of QNNs, as the lack of unified standards and a solid theoretical foundation makes the design of QNNs highly dependent on specific hardware platforms, making it difficult to ensure the universality and scalability of algorithms.

(3) **Quantum neural network architecture design.**

QNNs are a hot topic at the convergence of quantum computing and AI research. The architectural design of QNNs is like building a bridge, connecting the basic hardware of quantum computing with advanced AI applications. Currently, there is no unified standard for the architectural design of QNNs, with different researchers adopting various qubit layouts, quantum gate selections, and quantum circuit optimization strategies, making it difficult to compare the performance of different architectures. The theoretical foundation of QNNs is still imperfect, especially research on the representation and learning capabilities of QNNs is still in its infancy. Unlike classical neural networks, the design of QNNs needs to fully consider the distinctive characteristics of qubits, such as superposition and entanglement. For example, designing a QNN architecture that can utilize quantum parallelism can significantly accelerate tasks such as image recognition or NLP. However, the architectural design of QNNs is still in its early stages, with many design choices depending on experimental validation and theoretical innovation. Existing deep learning algorithms and frameworks are mostly designed for classical computers. How to effectively migrate these algorithms to quantum computing platforms while maintaining their performance advantages is an issue that urgently needs to be addressed.

(4) **Barren Plateaus.**

One of the most significant obstacles in QDL is the barren plateau phenomenon, where the gradient of the loss function vanishes exponentially as the size of the quantum system increases. This issue poses a severe limitation on the scalability of VQAs, which are commonly employed in QDL models. Researchers have proposed techniques such as layer-wise training, local cost functions, and improved initialization strategies to mitigate this issue, but a general and scalable solution is still lacking.

(5) **Classical Simulability.**

Another important question is whether QDL models provide genuine quantum advantages or if their behavior can be effectively simulated by classical methods. For example, classical tensor-network methods have shown surprising success in simulating certain quantum circuits. This raises concerns about the uniqueness and scalability of QDL approaches in real-world applications. A deeper theoretical understanding is needed to define the boundary between quantum and classical computational capabilities in machine learning tasks.

QDL remains an emerging interdisciplinary field with significant room for development in areas such as quantum noise processing, quantum neural network architecture design, the hardware implementation of QDL algorithms, and the integration of quantum computing with deep learning. As research continues to delve deeper and technology advances, there is reason to believe that in the near future, researchers will continue to break through existing boundaries, adding more innovative theoretical and practical achievements to the area of QDL. This will paint a vibrant and colorful landscape of scientific research, adding a bold and significant stroke to the field of QDL.

Acknowledgements This work is supported by Major Science and Technology Research Projects in Henan Province (221100210400), China, National Natural Science Foundation of China (61672470). Major Public Welfare Projects in Henan Province, China (201300210200). Henan Provincial Science and Technology Key Research and Development Program (252102211034). Henan Provincial Key Scientific Research Projects of Higher Education Institutions (25B520001). Doctoral Research Fund of Zhengzhou University of Light Industry (2024BSJJ014).

Author contributions W.H.G. conceived this study and provided overall guidance. Z.J.H. extensively collected literature related to the topic of the paper, covering various aspects of quantum deep learning, including fundamental theory, algorithm design, experimental validation, and more. W. H.G., Z.J.H., W.L.J., L.D.Y., K.D.L., and H.Y.C. conducted algorithm analysis on the collected literature. And various quantum deep learning algorithms were classified and organized. The discussion is divided into three areas based on data types: text, images, and multimodal data. A comprehensive summary of quantum deep learning was also provided. All authors, W.H.G., Z.J.H., W.L.J., L.D.Y., K.D.L., and H Y. C. actively participated in the discussion and contributed to the writing of the manuscript, ensuring comprehensive coverage and coherence of the manuscript

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

1. Hinton GE, Osindero S, Teh Y-W (2006) A fast learning algorithm for deep belief nets. Neural Computation 18(7):1527–1554
2. Sutskever I (2014) Sequence to sequence learning with neural networks. arXiv preprint arXiv:1409.3215
3. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M et al (2016) Mastering the game of go with deep neural networks and tree search. Nature 529(7587):484–489
4. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, et al (2019) Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems **32**
5. Team T (2015) Tensorflow: Large-scale machine learning on heterogeneous systems, 2015. URL http://tensorflow. org/. Software available from tensorflow. org
6. Vaswani A (2017) Attention is all you need. Adv Neural Inf Process Syst 30:1
7. Kak SC (1995) Quantum neural computing. Adv Imaging Electron Phys 94:259–313
8. Anguita D, Ridella S, Rivieccio F, Zunino R (2003) Quantum optimization for training support vector machines. Neural Netw 16(5–6):763–770
9. Harrow AW, Hassidim A, Lloyd S (2009) Quantum algorithm for linear systems of equations. Phys Rev Lett 103(15):150502
10. Lloyd S, Mohseni M, Rebentrost P (2014) Quantum principal component analysis. Nature Phys 10(9):631–633
11. Rebentrost P, Mohseni M, Lloyd S (2014) Quantum support vector machine for big data classification. Phys Rev Lett 113(13):130503

12. Li Z, Liu X, Xu N, Du J (2015) Experimental realization of a quantum support vector machine. Phys Rev Lett 114(14):140504
13. Farhi E, Goldstone J, Gutmann S (2014) A quantum approximate optimization algorithm. arXiv preprint arXiv:1411.4028
14. Benedetti M, Lloyd E, Sack S, Fiorentini M (2019) Parameterized quantum circuits as machine learning models. Quantum Sci Technol 4(4):043001
15. Zhang M, Dong L, Zeng Y, Cao N (2022) Improved circuit implementation of the hhl algorithm and its simulations on qiskit. Sci Rep 12(1):13287
16. Khot A, Kim T, Akash AR, Kim CH, Moy W (2024) Quantum-inspired machine learning framework using a physics-based ising solver chip. In: 2024 IEEE 7th International Conference on Industrial Cyber-Physical Systems (ICPS), pp. 1–6. IEEE
17. Shor PW (1994) Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings 35th Annual Symposium on Foundations of Computer Science, pp. 124–134. Ieee
18. Jozsa R (1999) Searching in grover's algorithm. arXiv preprint quant-ph/9901021
19. Zhang M, Dong L, Zeng Y, Cao N (2022) Improved circuit implementation of the hhl algorithm and its simulations on qiskit. Sci Rep 12(1):13287
20. Ahmed S (2019) Pattern recognition with quantum support vector machine (qsvm) on near term quantum processors. PhD thesis, Brac University
21. Lloyd S, Mohseni M, Rebentrost P (2014) Quantum principal component analysis. Nature Phys 10(9):631–633
22. Daley AJ, Bloch I, Kokail C, Flannigan S, Pearson N, Troyer M, Zoller P (2022) Practical quantum advantage in quantum simulation. Nature 607(7920):667–676
23. Amin MH, Andriyash E, Rolfe J, Kulchytskyy B, Melko R (2018) Quantum boltzmann machine. Phys Rev X 8(2):021050
24. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S (2017) Quantum machine learning. Nature 549(7671):195–202
25. Benedetti M, Lloyd E, Sack S, Fiorentini M (2019) Parameterized quantum circuits as machine learning models. Quantum Sci Technol 4(4):043001
26. Gyongyosi L, Imre S (2019) Training optimization for gate-model quantum neural networks. Sci Rep 9(1):12679
27. Gyongyosi L, Imre S (2020) Circuit depth reduction for gate-model quantum computers. Sci Rep 10(1):11229
28. Gyongyosi L, Imre S (2020) Optimizing high-efficiency quantum memory with quantum machine learning for near-term quantum devices. Sci Rep 10(1):135
29. Gyongyosi L, Imre S (2021) Scalable distributed gate-model quantum computers. Sci Rep 11(1):5172
30. Gyongyosi L, Imre S (2022) Advances in the quantum internet. Commun ACM 65(8):52–63
31. Giovannetti V, Lloyd S, Maccone L (2008) Quantum random access memory. Phys Rev Lett 100(16):160501
32. Wiebe N, Kapoor A, Svore KM (2014) Quantum deep learning. arXiv preprint arXiv:1412.3489
33. Cong I, Choi S, Lukin MD (2019) Quantum convolutional neural networks. Nat Phys 15(12):1273–1278
34. Liu J, Lim KH, Wood KL, Huang W, Guo C, Huang H-L (2021) Hybrid quantum-classical convolutional neural networks. Sci China Phys, Mech & Astron 64(9):290311
35. Zhang P, Niu J, Su Z, Wang B, Ma L, Song D (2018) End-to-end quantum-like language models with application to question answering. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32
36. Zhang P, Su Z, Zhang L, Wang B, Song D (2018) A quantum many-body wave function inspired language modeling approach. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 1303–1312
37. Panahi A, Saeedi S, Arodz T (2019) word2ket: Space-efficient word embeddings inspired by quantum entanglement. arXiv preprint arXiv:1911.04975
38. Meichanetzidis K, Gogioso S, De Felice G, Chiappori N, Toumi A, Coecke B (2020) Quantum natural language processing on near-term quantum computers. arXiv preprint arXiv:2005.04147
39. Kartsaklis D, Fan I, Yeung R, Pearson A, Lorenz R, Toumi A, Felice G, Meichanetzidis K, Clark S, Coecke B (2021) lambeq: An efficient high-level python library for quantum nlp. arXiv preprint arXiv:2110.04236

40. Zhang P, Hui W, Wang B, Zhao D, Song D, Lioma C, Simonsen JG (2022) Complex-valued neural network-based quantum language models. ACM Trans Inf Syst (TOIS) 40(4):1–31
41. Fan Z, Zhang J, Zhang P, Lin Q, Li Y, Qian Y (2024) Quantum-inspired language models based on unitary transformation. Inf Process & Manag 61(4):103741
42. Zhang Y, Song D, Zhang P, Li X, Wang P (2019) A quantum-inspired sentiment representation model for twitter sentiment analysis. Appl Intell 49:3093–3108
43. Guo P, Zhang J, Hou Y, Gong X, Wang P, Zhang Y (2019) Quantum-inspired dmatt-bigru for conversational sentiment analysis. In: 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), pp. 1602–1606. IEEE
44. Meichanetzidis K, Toumi A, Felice G, Coecke B (2020) Grammar-aware question-answering on quantum computers. arXiv preprint arXiv:2012.03756**7**
45. Meichanetzidis K, Toumi A, Felice G, Coecke B (2023) Grammar-aware sentence classification on quantum computers. Quantum Mach Intell 5(1):10
46. Shi J, Li Z, Lai W, Li F, Shi R, Feng Y, Zhang S (2021) Two end-to-end quantum-inspired deep neural networks for text classification. IEEE Trans on Knowl Data Eng 35(4):4335–4345
47. Yang C-HH, Qi J, Chen SY-C, Tsao Y, Chen P-Y (2022) When bert meets quantum temporal convolution learning for text classification in heterogeneous computing. In: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8602–8606. IEEE
48. Li G, Zhao X, Wang X (2024) Quantum self-attention neural networks for text classification. Sci China Inf Sci 67(4):142501
49. Wu S, Li J, Zhang P, Zhang Y (2021) Natural language processing meets quantum physics: A survey and categorization. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp. 3172–3182
50. Li SS, Zhang X, Zhou S, Shu H, Liang R, Liu H, Garcia LP (2023) Pqlm-multilingual decentralized portable quantum language model. In: ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5. IEEE
51. Zhao X, Wan H, Qi K (2024) Qpen: Quantum projection and quantum entanglement enhanced network for cross-lingual aspect-based sentiment analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, pp. 19670–19678
52. Yan P, Li L, Jin M, Zeng D (2021) Quantum probability-inspired graph neural network for document representation and classification. Neurocomputing 445:276–286
53. Qi J, Tejedor J (2022) Classical-to-quantum transfer learning for spoken command recognition based on quantum neural networks. In: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8627–8631. IEEE
54. Di Sipio R, Huang J-H, Chen SY-C, Mangini S, Worring M (2022) The dawn of quantum natural language processing. In: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8612–8616. IEEE
55. Lorenz R, Pearson A, Meichanetzidis K, Kartsaklis D, Coecke B (2023) Qnlp in practice: running compositional models of meaning on a quantum computer. J Artif Intell Res 76:1305–1342
56. Metawei MA, Taher M, ElDeeb H, Nassar SM (2023) A topic-aware classifier based on a hybrid quantum-classical model. Neural Computing Appl 35(25):18803–18812
57. Munshi M, Gupta R, Jadav NK, Polkowski Z, Tanwar S, Alqahtani F, Said W (2024) Quantum machine learning-based framework to detect heart failures in healthcare 4.0. Softw : Pract Exp 54(2):168–185
58. Lloyd S, Weedbrook C (2018) Quantum generative adversarial learning. Phys Rev Lett 121(4):040502
59. Dallaire-Demers P-L, Killoran N (2018) Quantum generative adversarial networks. Phys Rev A 98(1):012324
60. Hu L, Wu S-H, Cai W, Ma Y, Mu X, Xu Y, Wang H, Song Y, Deng D-L, Zou C-L et al (2019) Quantum generative adversarial learning in a superconducting quantum circuit. Sci Adv 5(1):2761
61. Zeng J, Wu Y, Liu J-G, Wang L, Hu J (2019) Learning and inference on generative adversarial quantum circuits. Phys Rev A 99(5):052306
62. Zoufal C, Lucchi A, Woerner S (2019) Quantum generative adversarial networks for learning and loading random distributions. npj Quantum Inf 5(1):103
63. Situ H, He Z, Wang Y, Li L, Zheng S (2020) Quantum generative adversarial network for generating discrete distribution. Inf Sci 538:193–208

64. Lu S, Duan L-M, Deng D-L (2020) Quantum adversarial machine learning. Phys Rev Res 2(3):033212
65. Romero J, Aspuru-Guzik A (2021) Variational quantum generators: Generative adversarial quantum machine learning for continuous distributions. Adv Quantum Technol 4(1):2000003
66. Huang H-L, Du Y, Gong M, Zhao Y, Wu Y, Wang C, Li S, Liang F, Lin J, Xu Y et al (2021) Experimental quantum generative adversarial networks for image generation. Phys Rev Appl 16(2):024051
67. Ahmed S, Sánchez Muñoz C, Nori F, Kockum AF (2021) Quantum state tomography with conditional generative adversarial networks. Phys Rev Lett 127(14):140502
68. Li J, Topaloglu RO, Ghosh S (2021) Quantum generative models for small molecule drug discovery. IEEE Trans Quantum Eng 2:1–8
69. Rudolph MS, Toussaint NB, Katabarwa A, Johri S, Peropadre B, Perdomo-Ortiz A (2022) Generation of high-resolution handwritten digits with an ion-trap quantum computer. Phys Rev X 12(3):031010
70. Zhou N-R, Zhang T-F, Xie X-W, Wu J-Y (2023) Hybrid quantum-classical generative adversarial networks for image generation via learning discrete distribution. Signal Process: Image Commun 110:116891
71. Altares-López S, García-Ripoll JJ, Ribeiro A (2024) Autoqml: automatic generation and training of robust quantum-inspired classifiers by using evolutionary algorithms on grayscale images. Expert Syst Appl 244:122984
72. Gong L-H, Pei J-J, Zhang T-F, Zhou N-R (2024) Quantum convolutional neural network based on variational quantum circuits. Opt Commun 550:129993
73. Feng C-B, Zhang J, Li J, Zhou Y (2024) Seam mask guided partial reconstruction with quantum-inspired local aggregation for deep image stitching. In: ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2430–2434. IEEE
74. Zhang H, Zhao Q, Chen C (2024) A light-weight quantum self-attention model for classical data classification. Appl Intell 54(4):3077–3091
75. He J, Zhu H, Zhou X (2024) Quantum image encryption algorithm via optimized quantum circuit and parity bit-plane permutation. J Inf Secur Appl 81:103698
76. Li Y, Xiao J, Chen Y, Jiao L (2019) Evolving deep convolutional neural networks by quantum behaved particle swarm optimization with binary encoding for image classification. Neurocomputing 362:156–165
77. Pang C-Y, Zhou R-G, Hu B-Q, Hu W, El-Rafei A (2019) Signal and image compression using quantum discrete cosine transform. Inf Sci 473:121–141
78. Liu N, Huang T, Gao J, Xu Z, Wang D, Li F (2021) Quantum-enhanced deep learning-based lithology interpretation from well logs. IEEE Trans Geosci Remote Sens 60:1–13
79. Alam M, Kundu S, Topaloglu RO, Ghosh S (2021) Quantum-classical hybrid machine learning for image classification (iccad special session paper). In: 2021 IEEE/ACM International Conference on Computer Aided Design (ICCAD), pp. 1–7. IEEE
80. Yang Y-F, Sun M (2022) Semiconductor defect detection by hybrid classical-quantum deep learning. In: proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2323–2332
81. Azevedo V, Silva C, Dutra I (2022) Quantum transfer learning for breast cancer detection. Quantum Mach Intell 4(1):5
82. Easom-McCaldin P, Bouridane A, Belatreche A, Jiang R, Al-Maadeed S (2022) Efficient quantum image classification using single qubit encoding. IEEE Trans Neural Netw Learning Syst 35(2):1472–1486
83. Shahwar T, Zafar J, Almogren A, Zafar H, Rehman AU, Shafiq M, Hamam H (2022) Automated detection of alzheimer's via hybrid classical quantum neural networks. Electronics 11(5):721
84. Ovalle-Magallanes E, Avina-Cervantes JG, Cruz-Aceves I, Ruiz-Pinales J (2022) Hybrid classical-quantum convolutional neural network for stenosis detection in x-ray coronary angiography. Expert Syst with Appl 189:116112
85. Gong C, Guan W, Gani A, Qi H (2022) Network attack detection scheme based on variational quantum neural network. J Supercomputing 78(15):16876–16897
86. Wei L, Liu H, Xu J, Shi L, Shan Z, Zhao B, Gao Y (2023) Quantum machine learning in medical image analysis: A survey. Neurocomputing 525:42–53

87. Rashmi R, Snekhalatha U, Krishnan PT, Dhanraj V (2023) Fat-based studies for computer-assisted screening of child obesity using thermal imaging based on deep learning techniques: a comparison with quantum machine learning approach. Soft Computing 27(18):13093–13114
88. Amin J, Anjum MA, Ibrar K, Sharif M, Kadry S, Crespo RG (2023) Detection of anomaly in surveillance videos using quantum convolutional neural networks. Image Vis Computing 135:104710
89. Amin J, Anjum MA, Ibrar K, Sharif M, Kadry S, Crespo RG (2023) Detection of anomaly in surveillance videos using quantum convolutional neural networks. Image Vis Computing 135:104710
90. Kulkarni V, Pawale S, Kharat A (2023) A classical-quantum convolutional neural network for detecting pneumonia from chest radiographs. Neural Computing Appl 35(21):15503–15510
91. Fan F, Shi Y, Guggemos T, Zhu XX (2023) Hybrid quantum-classical convolutional neural network model for image classification. IEEE Trans Neural Netw Learning Syst 35(12):18145–18159
92. Veeragandham S, Santhi H (2024) Optimization enabled deep quantum neural network for weed classification and density estimation. Expert Syst with Appl 243:122679
93. Das M, Naskar A, Mitra P, Basu B (2024) Shallow quantum neural networks (sqnns) with application to crack identification. Appl Intell 54(2):1247–1262
94. Hassan E, Hossain MS, Saber A, Elmougy S, Ghoneim A, Muhammad G (2024) A quantum convolutional network and resnet (50)-based classification architecture for the mnist medical dataset. Biomed Signal Process Control 87:105560
95. Sinha PK, Marimuthu R (2024) Conglomeration of deep neural network and quantum learning for object detection: Status quo review. Knowl -Based Syst 288:111480
96. Çavşi Zaim H, Yılmaz M, Yolaçan EN (2024) Design of gender recognition system using quantum-based deep learning. Neural Computing and Appl 36(4):1997–2014
97. Rao GE, Rajitha B, Srinivasu PN, Ijaz MF, Woźniak M (2024) Hybrid framework for respiratory lung diseases detection based on classical cnn and quantum classifiers from chest x-rays. Biomed Signal Process Control 88:105567
98. Chen L, Li T, Chen Y, Chen X, Wozniak M, Xiong N, Liang W (2024) Design and analysis of quantum machine learning: a survey. Connect Sci 36(1):2312121
99. Munshi M, Gupta R, Jadav NK, Polkowski Z, Tanwar S, Alqahtani F, Said W (2024) Quantum machine learning-based framework to detect heart failures in healthcare 4.0. Softw : Pract Exp 54(2):168–185
100. Mir MS, Bhat HA, Khanday FA (2024) Efficient representation of bit-planes for quantum image processing. Multim Tools Appl 83(31):1–18
101. Venkatesh K, Naik KJ, Shankar A (2024) Quantum convolution neural network for multi-nutrient detection and stress identification in plant leaves. Multim Tools Appl 80(24):1–23
102. Zhang Y, Song D, Zhang P, Wang P, Li J, Li X, Wang B (2018) A quantum-inspired multimodal sentiment analysis framework. Theor Computer Sci 752:21–40
103. Yu J, Zhu Z, Wang Y, Zhang W, Hu Y, Tan J (2020) Cross-modal knowledge reasoning for knowledge-based visual question answering. Pattern Recognit 108:107563
104. Zhang Y, Song D, Li X, Zhang P, Wang P, Rong L, Yu G, Wang B (2020) A quantum-like multimodal network framework for modeling interaction dynamics in multiparty conversational sentiment analysis. Inf Fusion 62:14–31
105. Li Q, Gkoumas D, Lioma C, Melucci M (2021) Quantum-inspired multimodal fusion for video sentiment analysis. Inf Fusion 65:58–71
106. Li Q, Gkoumas D, Sordoni A, Nie J-Y, Melucci M (2021) Quantum-inspired neural network for conversational emotion recognition. Proceedings of the AAAI Conference on Artificial Intelligence 35:13270–13278
107. Gkoumas D, Li Q, Dehdashti S, Melucci M, Yu Y, Song D (2021) Quantum cognitively motivated decision fusion for video sentiment analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 827–835
108. Gkoumas D (2021) Quantum Cognitively Motivated Context-aware Multimodal Representation Learning for Human Language Analysis. Open University (United Kingdom), Milton Keynes, U.K
109. Liu Y, Zhang Y, Li Q, Wang B, Song D (2021) What does your smile mean? jointly detecting multi-modal sarcasm and sentiment using quantum probability. Find Assoc Computational Linguist : EMNLP 2021:871–880
110. Gandhi A, Adhvaryu K, Poria S, Cambria E, Hussain A (2023) Multimodal sentiment analysis: a systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions. Inf Fusion 91:424–444

111. Zhu L, Zhu Z, Zhang C, Xu Y, Kong X (2023) Multimodal sentiment analysis based on fusion methods: a survey. Inf Fusion 95:306–325
112. Qu Z, Li Y, Tiwari P (2023) Qnmf: a quantum neural network based multimodal fusion system for intelligent diagnosis. Inf Fusion 100:101913
113. Al-Hawawreh M, Hossain MS (2023) A privacy-aware framework for detecting cyber attacks on internet of medical things systems using data fusion and quantum deep learning. Inf Fusion 99:101889
114. Tiwari P, Zhang L, Qu Z, Muhammad G (2024) Quantum fuzzy neural network for multimodal sentiment and sarcasm detection. Inf Fusion 103:102085
115. Rajalingam B, Priya R (2018) Multimodal medical image fusion based on deep learning neural network for clinical treatment analysis. Int J Chem Tech Res 11(06):160–176
116. Liu X, Xiao D (2019) Multimodality image fusion based on quantum wavelet transform and sum-modified-laplacian rule. Int J Theor Phys 58:734–744
117. Tan W, Tiwari P, Pandey HM, Moreira C, Jaiswal AK (2020) Multimodal medical image fusion algorithm in the era of big data. Neural computing and applications, 1–21
118. Dogra A, Ahuja CK, Kumar S (2022) A multi-modality paradigm for ct and mri fusion with applications of quantum image processing. Concurrency Comput : Pract Exp 34(20):6610
119. Cui C, Yang H, Wang Y, Zhao S, Asad Z, Coburn LA, Wilson KT, Landman BA, Huo Y (2023) Deep multimodal fusion of image and non-image data in disease diagnosis and prognosis: a review. Prog Biomed Eng 5(2):022001

Authors and Affiliations

Huaiguang Wu[1] · Jiahui Zhang[1] · Lijie Wang[1] · Daiyi Li[1] · Delong Kong[1] · Yucan Han[1]

✉ Lijie Wang
wlj01618@163.com

✉ Daiyi Li
lidaiyiyy@163.com

Huaiguang Wu
2011008@zzuli.edu.cn

Jiahui Zhang
z1921206288@163.com

Delong Kong
k17637831420@163.com

Yucan Han
zzulihyc@163.com

[1] Zhengzhou University of Light Industry, School of Computer Science and Technology, Zhengzhou 450000, China