



Quantum-AI Empowered Intelligent Surveillance: Advancing Public Safety Through Innovative Contraband Detection

Vigilancia inteligente potenciada por IA cuántica: desarrollando la seguridad pública mediante la detección de contrabando innovadora

Vigilância inteligente alimentada por IA quântica: desenvolvimento da segurança pública por meio da detecção inovadora de contrabando

Syed Atif Ali Shah^{1, 3 *}, Nasir Algeelani², Najeeb Al-Sammarraie^{2, 3}

Received: Aug/19/2023 • Accepted: Aug/19/2024 • Published: Nov/30/2025

Abstract

[Objective] This research aims to develop an intelligent surveillance model, Quantum-RetinaNet, by integrating a RetinaNet model with Quantum Convolutional Neural Networks (QCNN) to enhance accuracy and processing speed, thus addressing limitations of conventional CNN-based approaches. The study evaluates Quantum-RetinaNet's performance in real-time scenarios to determine its potential as a practical and scalable solution for intelligent monitoring in densely populated areas. **[Methodology]** This research integrates a RetinaNet model with Quantum Convolutional Neural Networks (Quantum CNN or QCNN), designating the resulting framework as Quantum-RetinaNet. By harnessing the quantum capabilities of QCNN, Quantum-RetinaNet achieves a balance between accuracy and processing speed. This innovative integration positions it as a game-changer, addressing the challenges of active monitoring in densely populated scenarios. As demand for efficient surveillance solutions grows, Quantum-RetinaNet offers a compelling alternative to existing CNN models, upholding accuracy standards without sacrificing real-time performance. **[Results]** The unique attributes of Quantum-RetinaNet have far-reaching implications for the future of intelligent surveillance. Its enhanced processing speed is poised to revolutionize the field, addressing the critical demand for systems that provide both rapid and precise monitoring **[Conclusions]** As Quantum-RetinaNet becomes the new standard, it ensures public safety and security while pushing the boundaries of AI in surveillance.

Keywords: Quantum AI; Deep Learning; Quantum Deep Learning; CNN; QCNN; intelligent surveillance; weapon detection.

* Corresponding author

Syed Atif Ali Shah, ✉ welcometatif@yahoo.com,  <https://orcid.org/0000-0002-4275-9731>

Nasir Algeelani, ✉ n.a.a.al.geelani@pl.hazne.nl,  <https://orcid.org/0000-0002-9354-3412>

Najeeb Al-Sammarraie, ✉ dr.najeeb@mediu.edu.my,  <https://orcid.org/0000-0003-2982-2211>

1 School of Engineering and Applied Sciences, Bahria University, Islamabad, Pakistan.

2 Hanze University of Applied Science, Institute of Engineering, Electrical and Electronic Engineering, Groningen, Netherlands.

3 Faculty of Computer and Information Technology, Al-Madinah International University, Malaysia.



Resumen

[Objetivo] Los sistemas de vigilancia han surgido como elementos cruciales para mantener la paz y la seguridad en el mundo moderno: su ubicuidad ayuda a monitorear actividades sospechosas de manera eficaz. Sin embargo, en entornos densamente poblados, el monitoreo activo continuo se vuelve poco práctico, lo que obliga al desarrollo de sistemas de vigilancia inteligentes. La integración de IA en la vigilancia fue una gran revolución; sin embargo, los problemas de velocidad han impedido su implementación generalizada. No obstante, la inteligencia artificial cuántica ha llevado a un avance significativo. Se ha demostrado que los sistemas de vigilancia basados en inteligencia artificial cuántica son más precisos y capaces de funcionar bien en escenarios en tiempo real nunca vistos. **[Metodología]** Esta investigación integra un modelo RetinaNet con Quantum CNN, denominándolo Quantum-RetinaNet. Al aprovechar las capacidades cuánticas del QCNN, Quantum-RetinaNet logra un equilibrio entre precisión y velocidad. Esta innovadora integración lo posiciona como un elemento innovador, que aborda los desafíos del monitoreo activo en escenarios densamente poblados. A medida que crece la demanda de soluciones de vigilancia eficientes, Quantum-RetinaNet ofrece una alternativa convincente a los modelos CNN existentes, manteniendo los estándares de precisión sin sacrificar el rendimiento en tiempo real. **[Resultados]** Los atributos únicos de Quantum-RetinaNet tienen implicaciones de largo alcance para el futuro de la vigilancia inteligente. Su velocidad de procesamiento mejorada está lista para revolucionar el campo, con lo cual se satisface la necesidad apremiante de un monitoreo rápido pero preciso. **[Conclusiones]** A medida que Quantum-RetinaNet se convierte en el nuevo estándar, garantiza la seguridad pública al tiempo que amplía los límites de la IA en la vigilancia.

Palabras clave: IA cuántica; aprendizaje profundo; aprendizaje profundo cuántico; CNN; QCNN; vigilancia inteligente; detección de armas.

Resumo

[Objetivo] Os sistemas de vigilância surgiram como elementos essenciais para manter a paz e a segurança no mundo moderno: sua onipresença ajuda a monitorar atividades suspeitas de forma eficaz. No entanto, em ambientes densamente povoados, o monitoramento ativo contínuo torna-se pouco prático, exigindo o desenvolvimento de sistemas de vigilância inteligentes. A integração da IA na vigilância foi uma grande revolução; no entanto, problemas de velocidade impedem sua implementação generalizada. No entanto, a inteligência artificial quântica tem levado a um avanço significativo. Os sistemas de vigilância baseados em inteligência artificial quântica têm se mostrado mais precisos e capazes de apresentar um bom desempenho em cenários em tempo real jamais visto. **[Metodologia]** Esta pesquisa integra um modelo RetinaNet com o Quantum CNN, chamando-o de Quantum-RetinaNet. Ao aproveitar os recursos quânticos do QCNN, o Quantum-RetinaNet alcança um equilíbrio entre precisão e velocidade. Essa integração inovadora o posiciona como um elemento revolucionário, abordando os desafios do monitoramento ativo em cenários densamente povoados. À medida que cresce a demanda por soluções de vigilância eficientes, a Quantum-RetinaNet oferece uma alternativa atraente aos modelos CNN existentes, mantendo os padrões de precisão sem sacrificar o desempenho em tempo real. **[Resultados]** Os atributos exclusivos da Quantum-RetinaNet têm implicações de longo alcance para o futuro da vigilância inteligente. Sua velocidade de processamento aprimorada está pronta para revolucionar o campo, atendendo à necessidade urgente de monitoramento rápido, mas preciso. **[Conclusões]** À medida que a Quantum-RetinaNet se torna o novo padrão, ela garante a segurança pública e, ao mesmo tempo, amplia os limites da IA na vigilância.



Palavras-chave: IA quântica; aprendizagem profunda; aprendizagem profunda quântica; CNN; QCNN; vigilância inteligente; detecção de armas.

Introduction

Intelligent surveillance research has significant implications across various domains, including public safety, security, and law enforcement efforts. It can be employed in public spaces, transportation hubs, schools, and other crowded areas to identify and prevent potential threats, mitigating risks associated with armed violence and terrorist activities. By integrating such systems into law enforcement operations, authorities can identify and apprehend individuals transporting contraband or harmful objects, reducing the occurrence of violent crimes. Mass shooting prevention can be achieved by detecting contraband in locations like schools and public events, while border security is essential for securing borders and preventing contraband trafficking (United Nations, 2006). Military settings can benefit from systems in identifying and neutralizing enemy threats, providing added protection for troops. Airport and aviation security can be enhanced by detecting concealed contraband in carry-on luggage and other areas. Prisons and correctional facilities can also benefit from technologies, enhancing overall security for staff and inmates. AI-based image identification, infrared imaging, millimeter-wave scanners, and improved sensor systems are examples of technological advances in contraband detection.

Turning to quantum computing, this represents a field of computer science and physics that explores the principles and technologies underlying quantum mechanics to develop new types of computers. Traditional computers use bits representing a 0 or 1 to

process and store information, while quantum computers use quantum bits or qubits, which can exist simultaneously in multiple states, enabling quantum computers to perform certain computations much faster than the classical ones, especially for complex problems such as factorization and optimization. Based on ideas like quantum superposition, entanglement, and interference, quantum computing can transform several industries, including banking, drug research, encryption, and AI. However, problems remain to be solved, such as creating dependable quantum hardware and new quantum computing algorithms.

An emerging technology called quantum deep learning uses deep learning methods to solve complex issues in fields such as image recognition, natural language processing, and drug development. Carrying out deep learning tasks entails creating quantum algorithms and quantum computing hardware. The quantum circuits use quantum gates to process and alter the quantum states created by data encoding into them for deep learning tasks. Despite these difficulties, there is significant interest in quantum deep learning, and organizations and academics worldwide are looking at how it may lead to new developments in AI.

In a hybrid CNN-QCNN, QCNN is used for later layers, such as classification or detection, while traditional CNN is used for earlier levels, like feature extraction and reduction. While quantum-inspired optimization algorithms like the Variational Quantum Eigensolver (VQE) or Quantum Approximate Optimization Algorithm (QAOA) can be used, this method can also be learned using



conventional deep learning approaches like backpropagation. Quantum deep learning is a potential technology area for the future since it can improve performance on some sorts of issues when combined with the classical CNN and quantum methods. For instance, the data may be preprocessed using the traditional CNN to extract relevant features, which can then be fed into the QCNN for additional processing and analysis.

While these technical advancements hold immense potential for enhancing intelligent systems, their implementation may raise ethical concerns, including privacy risks and biases in surveillance technology. These aspects warrant further investigation in future studies. Real-time warning and integration with surveillance systems provide a speedy reaction to possible threats. Overall research has far-reaching consequences for the safety and security of individuals and communities. However, achieving a balance between security measures and individual rights and privacy needs careful consideration of ethical and legal factors. Collaborations between researchers, law enforcement agencies, and policymakers are critical for developing and applying responsible intelligent surveillance technologies.

Literature review

Countries with large stocks of contraband, such as firearms, often experience elevated crime rates. Research indicates that illegal activities contribute to severe consequences, including murder, theft, destruction of infrastructure, and loss of billions of dollars (Atif et al., 2023). These findings underscore the critical need for effective measures to combat contraband-related crimes.

Traditional CCTV cameras are used to monitor specific areas, but the surveillance

technique relies largely on manual observation (Gopinath & Krishna, 2014). However, deep learning has modified this scenario, and researchers have developed more sophisticated models to identify multiple forms of contraband. This article examines these advancements, ranging from early contraband detection systems to the latest designs (Lai & Maples, 2017). This journey began with manual methods and ended with fully automated intelligent systems (Uddin et al., 2020).

The initial methods relied on manual techniques that incorporated weighted Gaussian mixtures, polarization signal-based methodologies, multiresolution mosaicism, three-dimensional (3D) computed tomography (CT), and the Haar cascade method (Uddin et al., 2020). Later, machine learning approaches emerged, introducing tools such as the Visual Background Extractor Algorithm, 3-layer ANNs used in conjunction with active mmWave radar, X-ray-based methodologies, and ANN-based MPEG-7 classifiers (Ho et al., 2019).

With the introduction of Speeded Up Robust Features (SURF), Harris Interest Point Detector (HIPD), and Fast Retina Keypoint (FREAK), the era of convolutional neural networks began (Grega et al., 2016). The arrival of various models was also announced to improve efficiency. Related to this, many scientific publications have explored the application of deep convolutional networks and transfer learning technology (Wang et al., 2019). For example, X-rays are used for classification purposes, while infrared imaging is employed to detect concealed contraband.

Recently, a special CNN model that is accurate and significantly speeds up contraband detection in streaming video was announced. It includes R-CNN, Fast R-CNN, Faster R-CNN, Inception, YOLO, VGG-Net,



ZF-Net, and YOLO-V3. In addition to speed and accuracy, another critical factor is complexity, which allows the model to run more smoothly on small devices such as smartphones and can be used in IoT apps to study such areas (Agurto et al., 2007). In addition to these advancements, numerous other models have been developed, such as Regional Proposal Network (RPN), GoogleNet, SqueezeNet, HyperNet, RetinaNet, LeNet, AlexNet, ZFNet, GoogleNet, VGGNet, ResNet, Startup Model, ResNeXt, SENet, MobileNet V1/V2, X, NASNet, PNASNet, ENASNet, EfficientPointNet, MobiliNet2, Inception ResNetV2, and ResNet50. This study also examines the terminology frequently used in the literature, covering concepts like “Complex Backbone vs. Light Backbone,” “Two-Phase Indicator vs. One-Phase Indicator,” and “Pros and Cons” of Different Model (Tiwari & Verma, 2015).

It is important to note that this study not only presents research models and their evolution over time but also aims to help researchers establish a solid foundation for future studies (Uddin et al., 2020). This review discusses the evolution of intelligent surveillance models and their performance in detecting guns and pistols (Piyadasa, 2020). Traditional methods for firearm detection, such as X-ray technology or millimetric wave imaging, are expensive and impractical due to their reaction to all metallic items, including different categories of contraband (Jain et al., 2020).

However, deep learning has proven to be the most effective learning method, with convolutional neural networks (CNNs) outperforming traditional methods (Zhao et al., 2019). Transfer learning, which re-utilizes information from one domain to another related domain, is becoming popular. Other popular techniques include Scale-Invariant Feature Transform (SIFT), Rotation-Invariant Feature

Transform (RIFT), and Fast Retina Keypoint (FREAK) (Ali Shah et al., 2020).

Uddin et al. (2022) proposed a reinforcement learning-based framework to optimize policy decisions for COVID-19 prevention. Their study highlights how machine learning can dynamically adapt public health policies in response to evolving pandemic data. By modeling different policy scenarios, the system learns optimal actions to minimize infection rates while balancing social and economic factors. This work is relevant in demonstrating the potential of AI in public safety and crisis management.

Initial efforts have been done to detect pistols in images, but these approaches struggle to identify multiple pistols within a single image (Olmos et al., 2019). The Bag of Words Surveillance System (BoWSS) (Danilov et al., 2021) algorithm was used to detect guns in images, and Faster R-CNN deep learning was used to detect a hand-held gun (Ashraf et al., 2022). However, the model can only detect and locate pistols and often fails to detect other types of contraband, such as machine guns (Uddin et al., 2020). Fernandez et al. presented a new CNN model for detecting guns and knives in video surveillance systems and conducted a comparative analysis with GoogleNet and SqueezeNet (Khan et al., 2020). The results indicated that GoogleNet performed better in knife detection (Dong et al., 2023), while SqueezeNet demonstrated superior performance in gun detection. A sequential deep neural network-based approach was developed to resolve three major problems in automated Ballistic and Concealed Guns (BCG) detection and learning (Gelana & Yadav, 2019). Nevertheless, real-time detection of pistols remains a challenge due to factors such as distance, visibility, type of pistol, scale, rotation, and shape. Advancements in these domains are being made to improve accuracy



and performance in contraband detection (Atif et al., 2023). Ali Shah et al. (2021) conducted a comprehensive review of existing weapon detection techniques, specifically focusing on their applicability to street-crime scenarios. The study evaluates various methods, including image processing, machine learning, and sensor-based systems, highlighting their strengths and limitations. This review provides a foundational understanding of the technological landscape in weapon detection, aiding future research in developing more robust and context-aware security systems. Khalid et al. (2023) developed a weapon detection system aimed at enhancing surveillance and security in real-time environments. Utilizing advanced computer vision and machine learning algorithms, the system was designed to identify weapons in live video feeds with high accuracy. The research contributes significantly to intelligent surveillance solutions by addressing the growing need for automated threat detection in public and private security settings. Kambhatla and Ahmed (2024) explored advanced weapon detection techniques using YOLO and Faster R-CNN, optimizing speed and accuracy through model pruning and ensembling, achieving high average precision (AP) scores. This highlights the significance of refining architectures for practical applications (Kambhatla & Ahmed, 2024). Xie and Wang (2023) introduced a deep learning pipeline incorporating multiple base models (BMs) to mitigate false positives and negatives. Their ensemble approach outperformed individual architectures, achieving robust detection capabilities (Xie & Wang, 2023).

Dataset

For this research, we have designed our dataset to include the most used type of contraband in third-world countries.

Other datasets include a variety of weaponry but usually contain various types of contraband used in Hollywood movies or in developed countries. When the same model is implemented in third-world countries, it is less accurate. To make our work acceptable internationally, we developed our dataset with an armory used in all parts of the world. This research primarily focuses on the types of weaponry used in street crimes, such as short-range rifles, shotguns, pistols, and knives, collectively referred to as the Street Crimes Arms Dataset (SCAD). Each category contains approximately 5,000 images, bringing the total to about 20,000 images. To optimize the dataset for effective model training, various alterations have been applied, such as augmentation and normalization.

Methodology

This research integrates the power of Quantum computing with conventional Deep Learning, specifically Convolutional Neural Networks. Both technologies have their strengths and weaknesses. Quantum computing is famous for its tremendous speed, but due to the limited availability of hardware, its implementation is still uncommon. On the other hand, although Deep learning applications are widely used, they lack speed in many real-time scenarios. Combining both technologies reveals a new dimension and thus suppresses their weaknesses. There is currently no prior work that integrates QCNN with RetinaNet for intelligent surveillance applications. In RetinaNet, the task of feature extraction is replaced by QCNN. Thus, it achieves the speed of Quantum Computing and accuracy of RetinaNet.



RetinaNet

RetinaNet is a state-of-the-art object detection algorithm that was introduced by researchers at Facebook AI Research in 2017. It is designed to solve the problem of detecting objects in an image, where the location and type of objects can vary widely. The key innovation of RetinaNet is the use of a novel loss function called Focal Loss, which addresses the issue of class imbalance in object detection. Object detection typically has many more negative examples (background) than positive examples (objects of interest). RetinaNet utilizes Focal Loss to reduce bias toward negative examples and focuses on hard examples in the positive class, which helps the model identify rare and important positive cases. RetinaNet's feature pyramid network (FPN) design mixes low-resolution and high-resolution characteristics to recognize objects of varying sizes and resolutions. The model has gained cutting-edge performance on benchmarks like COCO and Pascal VOC, making it a popular choice for object identification tasks in industrial and commercial applications.

Quantum-RetinaNet

The core functionality of the RetinaNet is divided into two main parts. The first is called feature extractor, which deals with

the extraction of features. The second one is called task-specific networks, responsible for Classification and Bounding Box. The first part, i.e., the feature extractor, uses convolution and pooling layers to extract features, which is a time-consuming process. In Quantum-RetinaNet, this time is reduced using Quantum-Convolutional Neural Network.

Fig. 1 summarizes how Q-RetinaNet works. This model converts an image into a vector of corresponding bits for quantum computing. Qubits are fabricated, and feature extraction is performed using QCNN. These features are then forwarded to Task Specific Networks, which draw bounding boxes and detect desired objects.

Feature Extraction using Quantum Convolutional Neural Network A Quantum Convolutional Neural Network or Quantum Convolutional Neural Network uses quantum circuits to conduct convolutional operations on input data. The procedure entails converting the incoming data into a quantum state through methods like amplitude encoding or qubit encoding. Convolutional filters are used to identify specific characteristics, such as edges or textures. After that, the data is pooled using a pooling layer, which lowers the dimensionality while maintaining crucial properties, as shown in Fig. 2.

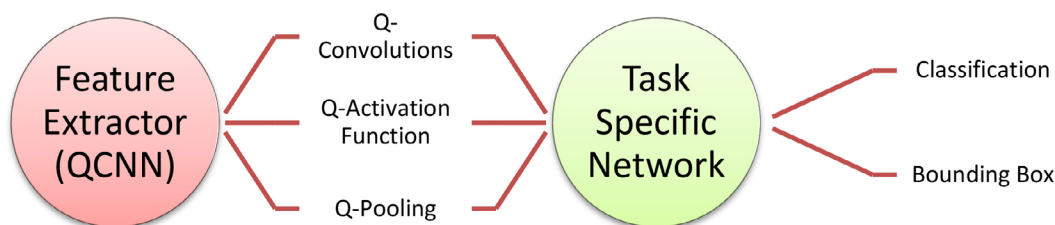


Figure 1. *The Basic Architecture of the RetinaNet.*

Note: the first part deals with the extraction of features, and the second one deals with task-specific networks. Derived from research.

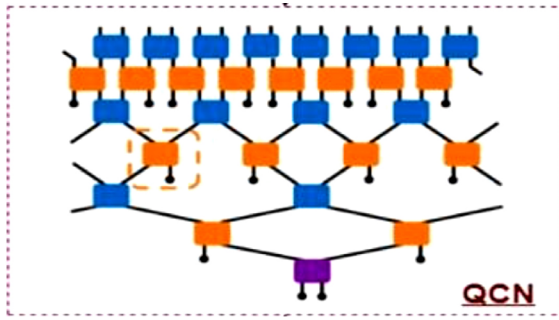


Figure 2. *Quantum Convolutional Neural Network or Qonvolutional Neural Network*
 Note: QCNN uses quantum circuits to conduct convolutional operations on input data, using amplitude or qubit encoding to convert incoming data into a quantum state. Derived from research.

For classification tasks, the output is processed utilizing additional layers of quantum circuits, employing methods such as variational quantum classifiers or quantum SVMs. To improve network performance, the parameters of the quantum circuits are improved using strategies like quantum gradient descent or variational approaches.

Bits to Qubits. The data in a bit is encoded into the quantum state of a single qubit by transforming it from a classical bit to a quantum bit. A qubit is a type of quantum bit that may exist in several states at once as opposed to a classical bit, which can only exist in one state at a time. For example, to encode a classical bit into a qubit, initialize the qubit in the $|0\rangle$ state, which corresponds to a 0 state of a classical bit. If the classical bit is 1, apply the X gate to change the qubit from its $|0\rangle$ state to its $|1\rangle$ state. The qubit's final state corresponds to the encoded form of the conventional bit. Several additional techniques can convert conventional bits into qubits depending on the exact application.

Quantum Convolutions. Quantum convolutions are implemented using

quantum circuits, gates, and quantum algorithms. While gates, such as the Hadamard-Walsh transform gate for discrete cosine transform (DCT) and discrete wavelet transform (DWT) operations, apply the convolutional filter directly to a quantum state, circuits employ quantum Fourier transforms (QFTs) to change a quantum state. Quantum algorithms are created to work on quantum data and perform various tasks, including quantum convolution, such as the quantum singular value transformation (QSVD) or the quantum Fourier transform (QFT).

Quantum Activation Functions. Quantum activation functions are quantum analogs of classical activation functions in neural networks. They are mathematical functions that transform the output of a quantum circuit into a new quantum state, introducing nonlinearity into the output of a neural network. There are several types of quantum activation functions, including Quantum ReLU (Q-ReLU), Quantum Sigmoid (Q-Sigmoid), and Quantum Softmax (Q-Softmax). Quantum activation functions can be used in quantum machine learning algorithms to improve their performance on specific problems.

Quantum Pooling

The quantum pooling approach is employed in quantum machine learning to decrease the spatial dimensions of a quantum feature map while keeping the most critical data. Since quantum states are inherently uncertain, identifying the most important details can be challenging. Numerous quantum pooling strategies have been suggested to overcome this issue, including quantum maximum pooling, quantum mean pooling, and quantum median pooling. Quantum amplitude pooling is implemented using quantum gates and circuits, such as quantum amplitude pooling.



The input quantum state is transformed into the frequency domain using a quantum Fourier transform, and the amplitude of each frequency component is squared using a quantum circuit containing Hadamard gates and Controlled- NOT (CNOT) gates. Through a measurement operation, the squared amplitudes are then used to construct a new quantum state with fewer qubits.

Qubit to bit. Quantum computing involves measuring the state of a qubit to convert it to a classical bit, with the probability determined by the quantum state of the qubit. Quantum convolution operations are used in quantum neural networks, specifically in Qonvolutional Neural Networks (QCNNs). In Qonvolutional, input data is encoded into a quantum state, usually using qubits, and quantum gates are applied to the state to perform the convolution operation. The output is then decoded from the resulting quantum state.

Task Specific Networks. For each of the X anchors and Y object classes, the classification subnet forecasts the likelihood of an object's presence at each spatial position. The subnet is a fully connected network interconnected to all levels, sharing features throughout. Its design includes a channel

input feature map, four convolutional layers with filters, ReLU activations, filters, and sigmoid activations to output binary predictions for each spatial location. The object classification subnet and box regression subnet share a structure, but the box regression subnet uses a class-agnostic bounding box regression with fewer parameters. This strategy is equally successful, as both subnets employ distinct parameters for regressing the offset from anchor boxes to neighboring ground-truth objects.

The detailed process architecture of the Q-RetinaNet is shown in Fig. 3. The process begins with an image being fed into the model; then, it is converted to a vector of corresponding bits. For quantum computing, bits are transformed into Qubits. Once Qubits are fabricated, feature extraction can be performed using QCNN that is an equivalent of CNN and only works on Qubits. After extracting the features, the next step is to forward these feature maps to Task Specific Networks. Due to compatibility issues, these Qubits must now be transformed back into conventional bits. Finally, these bits are transferred to Task Specific Networks, which are responsible for drawing bounding boxes and detecting the desired objects from the image.

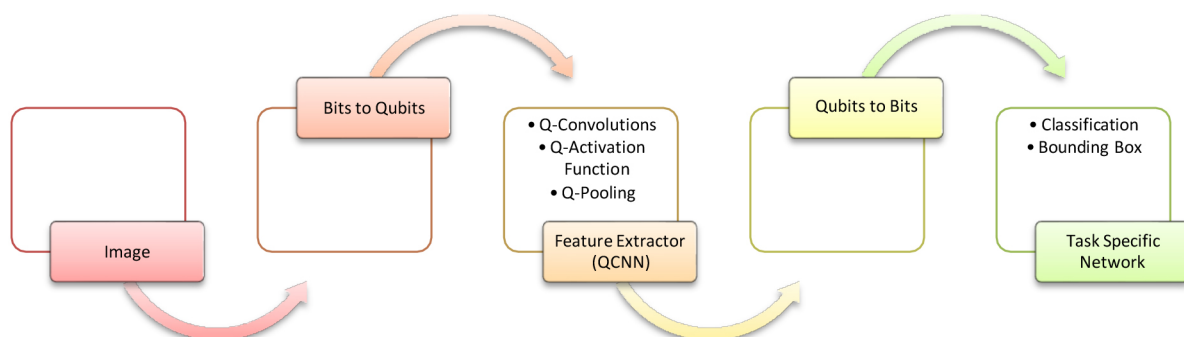


Figure 3. *Process Flow Diagram of RetinaNet*

Note: An image is converted into a vector of corresponding Qubits for quantum computing, and feature extraction is performed using QCNN. Moreover, Task Specific Networks draw bounding boxes and detect desired objects. Derived from research.



Analysis and Results

Various comparative analysis approaches are utilized to compare the performance of the different models employed in the research.

Accuracy

Fig. 4 illustrates the accuracy of LeNet, AlexNet, VGG, RetinaNet, and

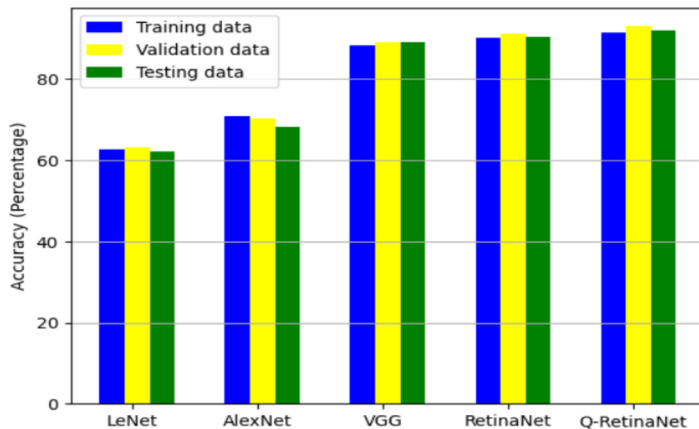


Figure 4. *The Comparison of the Accuracy of Lenet, Alexnet, VGG, RetinaNet, and Quantum-RetinaNet.*
Note: derived from research.

Quantum- RetinaNet. Although the Q-RetinaNet contains fewer Quantum layers, it still shows sufficient accuracy; although it partially implements the QCNN compared to CNN, the results are quite impressive.

Confusion Matrix

In machine learning and deep learning, a classification model's performance is assessed using a confusion matrix. It is used to assess the classification model's accuracy by contrasting the anticipated and actual class labels. An explanation of how the confusion matrix calculates the percentage of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) predictions can be found in the confusion matrix. The instances in each row of the matrix correspond to a predicted class, whereas the examples in each column correspond to an actual class, as shown in Fig. 5.

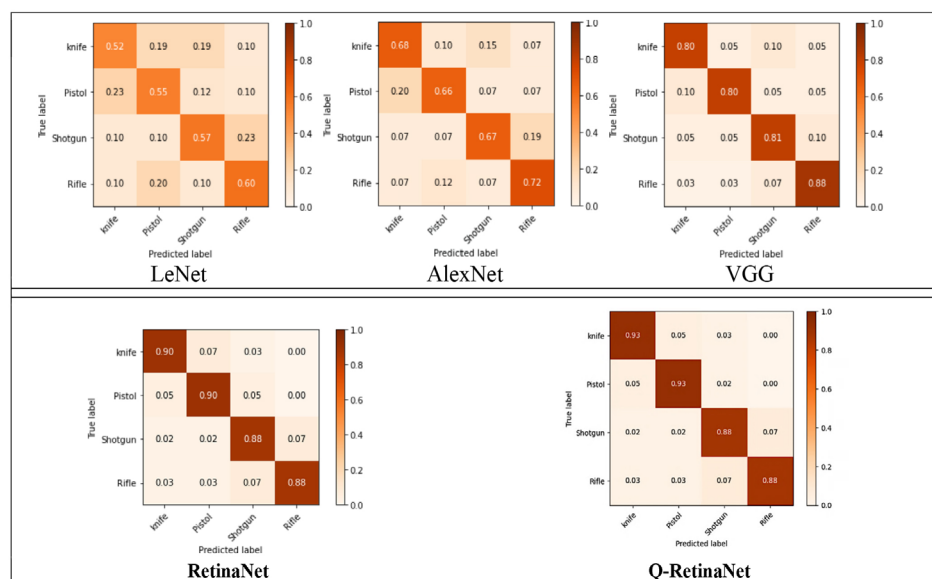


Figure 5. *The Comparison of the Confusion Matrix of Lenet, Alexnet, VGG, RetinaNet, and Quantum-RetinaNet*
Note: derived from research.

Several performance indicators are calculated using the confusion matrix, including accuracy, precision, recall, and F1-score. It offers a thorough analysis of the categorization model and aids in determining the model's advantages and disadvantages.



F1-SCORE

Deep learning often uses the F1-score to track performance, particularly in binary classification. It is the harmonic mean of precision and recall, where recall is the ratio of true positives to the total of both true and false positives, and precision is the sum of both true and false positives (FN). These steps are used to determine the F1 score:

$$F1\text{-score} = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$$

The F1 score is crucial for evaluating a classification model's accuracy and recall. It ranges from 0 to 1, with 0 indicating no predictive capacity and 1 indicating perfect precision and recall. It can be computed independently for each class and averaged to provide a performance indicator. See Fig. 6 for detailed information.

ROC

A ROC curve in deep learning measures the accuracy of binary classifier models by plotting the true positive rate (TPR) against the false positive rate (FPR) at different threshold settings, as shown in Fig. 7. It helps determine the optimal threshold setting and compares the performance of different classifier models. The area under the ROC curve (AUC) is a commonly used metric, with scores ranging from 0 to 1, indicating poor classification and random guessing.

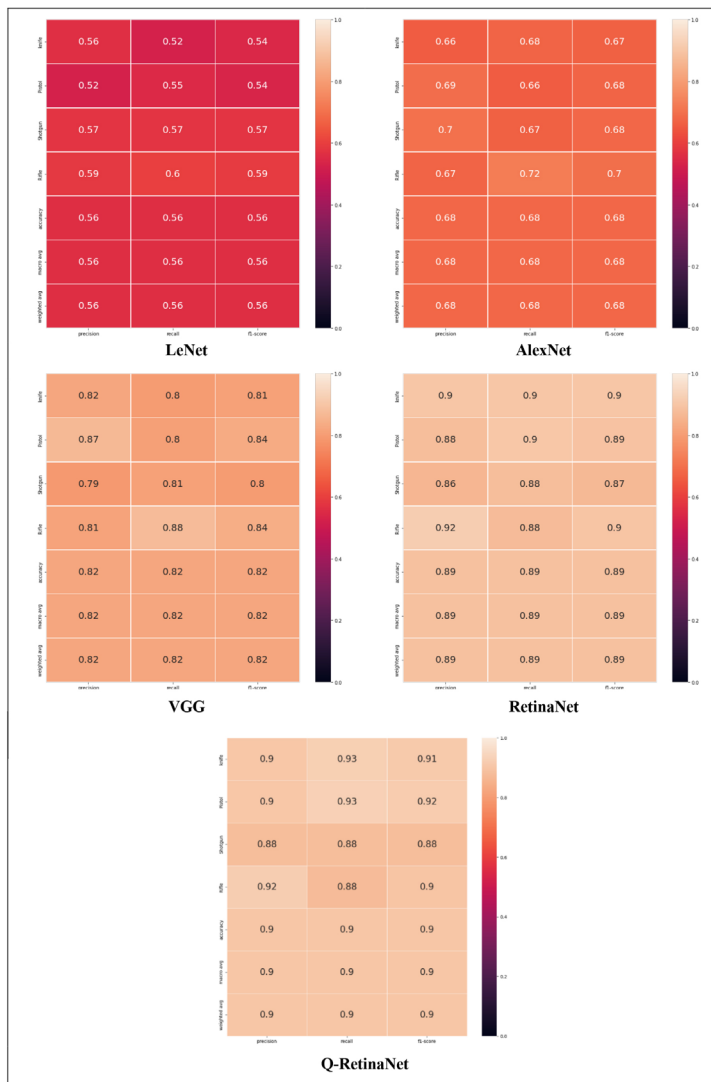


Figure 6. The Comparison of Accuracy of Lenet, Alexnet, VGG, RetinaNet, and Quantum-RetinaNet
Note: derived from research.

Conclusion

Hybrid neural networks (QCNN-RetinaNet) for object detection integrating conventional and quantum computing methods have attracted much interest. RetinaNet is utilized for object detection, while the QCNN layers are responsible for feature extraction and reduction. Quantum gates and circuits can be used to implement these layers, potentially resulting in exponential speedup in some applications. Traditional deep learning methods like backpropagation

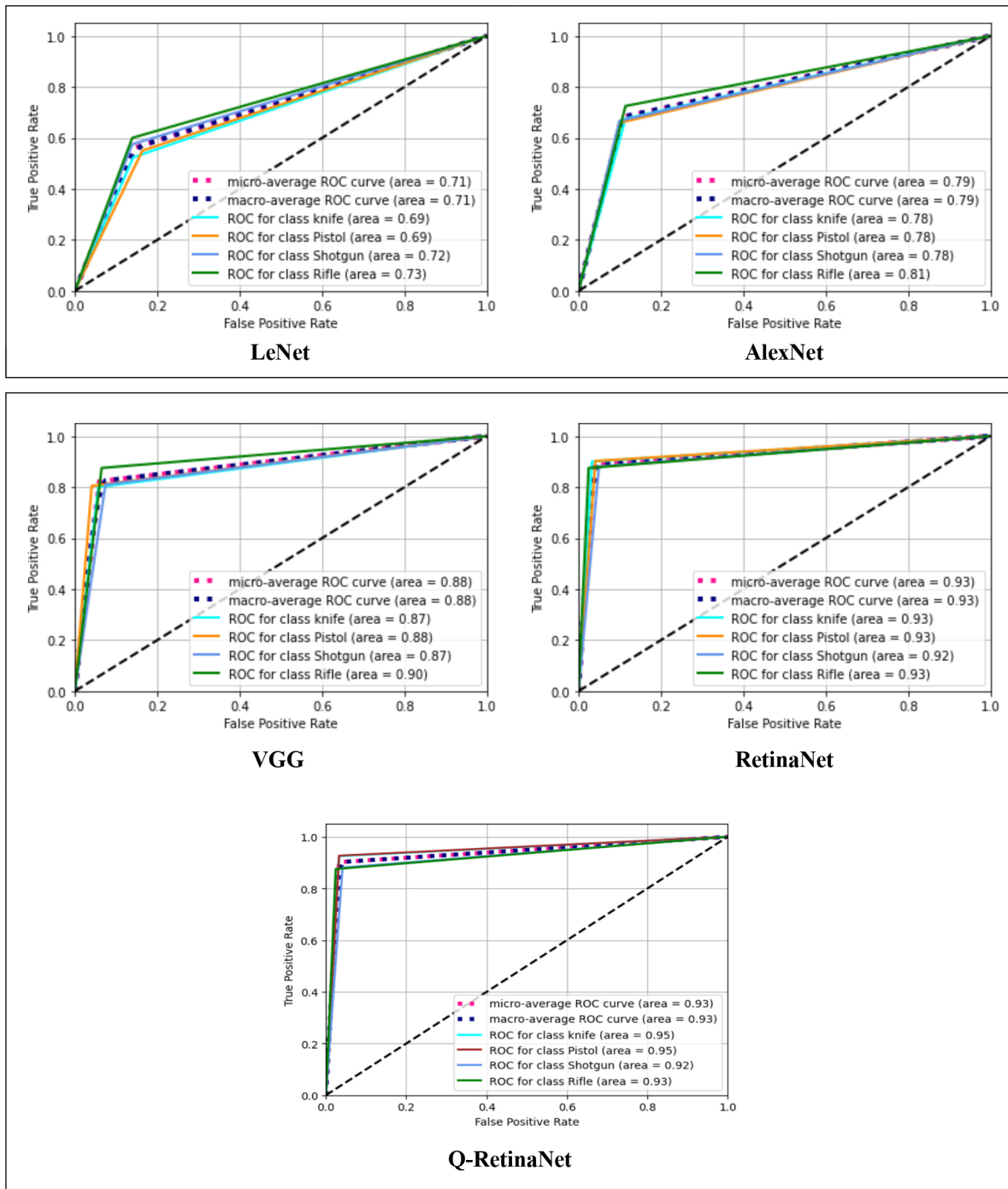


Figure 7. The Comparative Analysis of ROC of LeNet, AlexNet, VGG, RetinaNet, and Quantum-RetinaNet
 Note: derived from research.



and focus loss function may be used to train RetinaNet, which was created to overcome the class imbalance in object recognition. When solving specific object detection issues, the hybrid QCNN-RetinaNet performs better than a QCNN or RetinaNet. However, in addition to competence in deep learning and neural networks, constructing a hybrid QCNN-RetinaNet combines the benefits of both classical and quantum computing methods.

Ethics

This research generated a dataset designed to include a range of armories such as pistols, shotguns, short-range rifles, knives, and Kalashnikovs. It was defined as a dataset that contained all types of armories most commonly seen in the streets. The dataset shows armories frequently found in third-world countries, as the project focuses on surveillance in developing countries, which makes the dataset relevant for countries in such regions. Our investigation involved consulting various sources, including law enforcement, police, private security agencies, websites, the Internet, and the crime branch of a major news channel, among other security departments.

Conflict of interest

The authors declare no competing interests.

Author contribution statement

Dr. Syed Atif Ali Shah: Conceptualization idea, methodology, investigation, experiment implementation, data collection, and writing (80%).

Dr. Nasir Ageelani: Supervision and review (15%).

Dr. Najeeb Al-Sammurrai: Supervision and validation (5%).

Data availability statement

The data supporting the results of this study will be made available by the corresponding author, Dr. Syed Atif Ali Shah, upon reasonable request.

Preprint A Preprint version of this paper was deposited in: <https://arxiv.org/abs/2309.03231>

References

- Agurto, A., Li, Y., Tian, G. Y., Bowring, N., & Lockwood, S. (2007). A review of concealed weapon detection and research in perspective. In *2007 IEEE International Conference on Networking, Sensing and Control* (pp. 443–448). <https://doi.org/10.1109/ICNSC.2007.372819>
- Ali Shah, S. A., Ahmad Al-Khasawneh, M., & Uddin, M. I. (2021). Review of Weapon Detection Techniques within the Scope of Street-Crimes. 2021 2nd International Conference on Smart Computing and Electronic Enterprise (ICSCEE), 26-37. <https://doi.org/10.1109/ICSCEE50312.2021.9498007>
- Ali Shah, S. A., Uddin, I., Aziz, F., Ahmad, S., Al-Khasawneh, M. A., & Sharaf, M. (2020). An enhanced deep neural network for predicting workplace absenteeism. *Complexity*, 1–12. <https://doi.org/10.1155/2020/5843932>
- Ashraf, A.H., Imran, M., Qahtani, A.M., Alsufyani, A., Almutiry, O. et al. (2022). Weapons detection for security and video surveillance using CNN and yolo-v5s. *Computers, Materials & Continua*, 70(2), 2761–2775. <https://doi.org/10.32604/cmc.2022.018785>
- Atif, S., Shah, A., Abdel-Wahab, A., Ageelani, N., & Najeeb, N. (2023). Street-crimes modelled arms recognition technique employing deep learning and quantum deep learning. *Indonesian Journal of Electrical Engineering and Computer Science*, 30(1), 528–544. <http://doi.org/10.11591/ijeecs.v30.i1.pp528-544>



- Danilov, V. V., Klyshnikov, K. Y., Gerget, O. M., Kutikhin, A. G., Ganyukov, V. I., & Frangi, A. F. (2021). Real-time coronary artery stenosis detection based on modern neural networks. *Scientific Reports*, 11(1), 1–14. <https://doi.org/10.1038/s41598-021-87174-2>
- Dong, S., Liu, X., Zhong, S., Shi, K., & Zhu, H. (2023). Practical synchronization of neural networks with delayed impulses and external disturbance via hybrid control. *Neural Networks*, 157, 54–64. <https://doi.org/10.1016/j.neunet.2022.09.025>
- Gelana, F., & Yadav, A. (2019). Firearm detection from surveillance cameras using image processing and machine learning techniques. In *Advances in Intelligent Systems and Computing* (Vol. 851, pp. 89–98). http://dx.doi.org/10.1007/978-981-13-2414-7_3
- Gopinath, B. Y., & Krishna, V. S. (2014). Concealed weapon detection using image processing. *International Journal of Electronics and Communication Technology*, 5(6), 13–17.
- Grega, M., Mاتیolański, A., Guzik, P., & Leszczuk, M. (2016). Automated detection of firearms and knives in a CCTV image. *Sensors*, 16(1), 1–20. <https://doi.org/10.3390/s16010047>
- Ho, N., Yang, C.-Y., & Jordan, M. I. (2019). Convergence rates for Gaussian mixtures of experts. *Journal of Machine Learning Research*, 23(1), 1–81. <http://arxiv.org/abs/1907.04377>
- Jain, H., Vikram, A., Mohana, Kashyap, A., & Jain, A. (2020). Weapon detection using artificial intelligence and deep learning for security applications. *International Conference on Electronics and Sustainable Communication Systems*. pp. 193–198. <http://dx.doi.org/10.1109/ICESC48915.2020.9155832>
- Kambhatla, A., & Ahmed, K. R. (2024). Real-time deep learning weapon detection techniques for mitigating lone wolf attacks. *arXiv Preprint*. <https://doi.org/10.48550/arXiv.2405.14148>
- Khalid, S., Waqar, A., Ain Tahir, H. U., Edo, O. C., & Tenebe, I. T. (2023). Weapon detection system for surveillance and security. *2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, 1–7. <https://doi.org/10.1109/ITIKD56332.2023.10099733>
- Khan, A., Gul, M. A., Uddin, M. I., Ali Shah, S. A., Ahmad, S., & Al Firdausi, M. D., Zaindin, M. (2020). Summarizing online movie reviews: A machine learning approach to big data analytics. *Scientific Programming*, 2021, 1. <https://doi.org/10.1155/2020/5812715>
- Lai, J. & Maples S. (2017). *Developing a real-time gun detection classifier*. <https://cs231n.stanford.edu/reports/2017/pdfs/716.pdf>
- Olmos, R., Tabik, S., Lamas, A., Pérez-Hernández, F., & Herrera, F. (2019). A binocular image fusion approach for minimizing false positives in handgun detection with deep learning. *Information Fusion*, 49, 271–280. <https://doi.org/10.1016/j.inffus.2018.11.015>
- Piyadasa, T. D. (2020). Concealed weapon detection using convolutional neural networks. *ResearchGate*. https://www.researchgate.net/publication/346974570_Concealed_Weapon_Detection_Using_Convolutional_Neural_Networks
- Tiwari, R. K., & Verma, G. K. (2015). A computer vision-based framework for visual gun detection using Harris interest point detector. *Procedia Computer Science*, 54, 703–712. <https://doi.org/10.1016/j.procs.2015.06.083>
- Uddin, M. I., Ali Shah, S. A., Al-Khasawneh, M. A., Alarood, A. A., & Alsolami, E. (2022). Optimal policy learning for COVID-19 prevention using reinforcement learning. *Journal of Information Science*, 48(3), 336–348. <https://doi.org/10.1177/0165551520959798>
- Uddin, M. I., Shah, S. A. A., & Al-Khasawneh, M. A. (2020). A novel deep convolutional neural network model to monitor people following guidelines to avoid COVID-19. *Journal of Sensors*, 1–10. <https://doi.org/10.1155/2020/8856801>
- Uddin, M. I., Zada, N., Aziz, F., Saeed, Y., Zeb, A., Ali Shah, S. A., & others. (2020). Prediction of future terrorist activities using deep neural networks. *Complexity*. 1–16. <https://doi.org/10.1155/2020/1373087>
- United Nations [UN]. (2006). *Small arms review conference 2006*: United Nations conference to review progress made in the implementation of the programme of action to prevent, combat and eradicate the illicit trade in small arms and light weapons in all its aspects. *United Nations*. <http://www.un.org/smallarms2006>
- Wang, Y., Wang, C., Zhang, H., Dong, Y., & Wei, S. (2019). Automatic ship detection based on RetinaNet using multi-resolution Gaofen-3 imagery. *Remote Sensing*, 11(5), 1–20. <https://doi.org/10.3390/rs11050531>



- Xie, J., & Wang, L. (2023). Novel deep learning pipeline for automatic weapon detection. *arXiv Preprint*. <https://doi.org/10.48550/arXiv.2309.16654>
- Zhao, Y., Han, R., & Rao, Y. (2019). A new feature pyramid network for object detection. In *Proceedings of the 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS 2019)* (pp. 428–431).



An Innovative Framework for Intelligent Computer Vision Empowered by Deep Learning
(Thulasi Bikkur, Srinivasarao Thota, Abayomi Ayotunde Ayode) [Uniciencia](#) is protected by
[Attribution-NonCommercial-NoDerivs 3.0 Unported \(CC BY-NC-ND 3.0\)](#)