Journal homepage: https://itce.com.ua/en

Information Technologies and Computer Engineering **Vol. 21, No. 3. 2024**

Article's History: Received: 31.07.2024 Revised: 18.10.2024 Accepted: 26.12.2024

UDC 004.9

DOI: 10.63341/vitce/3.2024.09

Neural network architecture for real-time QR code recognition

Hlib Serediuk

Postgraduate Student
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0009-0000-3010-6437

Volodymyr Garmash

PhD of Technical Sciences, Associate Professor
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0009-0007-1861-8772

Abstract. The study investigated modern neural network architectures for efficient real-time recognition of QR codes, which is critical for the development of mobile applications and industrial control systems. The study analysed the features of using light convolutional neural networks optimised for operation on mobile devices with limited computing resources. A modified architecture was proposed that strikes a balance between speed and accuracy when processing a video stream, achieving a recognition rate of 30 frames per second on standard mobile processors. A multi-stage decision-making mechanism based on the Early Stopping Mechanism (ESM) has been developed to optimise image processing. An adaptive filtering method using a median filter and morphological reconstruction was implemented, which substantially improved the quality of input data. The proposed architecture included a specialised preprocessing module and a system of residual-and-excitation blocks to improve recognition efficiency. Experimental studies demonstrated a 12-15% increase in the system's real-time performance compared to the baseline models when processing a video stream. The system successfully recognised QR codes in poor lighting conditions and non-standard tilt angles with an accuracy of over 92%. A 27% reduction in computational complexity was achieved while maintaining high recognition accuracy. The developed method efficiently processes images with geometric distortions even in conditions of limited resources. The study developed the theoretical foundations of optimising convolutional neural networks for computer vision tasks, offering new approaches to balancing recognition efficiency and accuracy. The practical significance of the study was confirmed by the possibility of direct integration of the developed system into mobile applications and industrial quality control systems, while the proposed optimisation methods can be adapted to a wide range of computer vision tasks on mobile platforms

Keywords: convolutional neural networks; mobile devices; video stream processing; Early Stopping Mechanism; residual-and-excitation units; computer vision

Introduction

The relevance of studying neural network architectures for real-time QR code recognition is conditioned by the rapid development of mobile technologies and the growing need for fast and reliable processing of visual information. Conventional QR code recognition algorithms based on classical computer vision methods often demonstrate unsatisfactory performance when working in challenging conditions: low light, non-standard viewing angles, or obstacles. The problem of ensuring stable operation on mobile devices, where computing resources are limited, is particularly acute. The use of neural networks allows creating more adaptive recognition systems that can work effectively in a variety of conditions while maintaining high performance.

An analysis of recent research showed great progress in the development of QR code recognition methods using neural networks. S. Bhatia & A.S. Albarrak (2023) proposed

Suggested Citation:
Serediuk, H., & Garmash, V. (2024). Neural network architecture for real-time QR code recognition. *Information Technologies and Computer Engineering*, 21(3), 9-19. doi: 10.63341/vitce/3.2024.09

*Corresponding author

an innovative architecture based on XAI-faster RCNN for supply chain management systems. Their study found that the combination of blockchain technologies with neural networks can achieve recognition accuracy of more than 95% even in poor lighting conditions and provide reliable authentication of goods.

E. Borandag (2023) made a significant contribution to the development of recognition methods by developing an integrated system based on IoT and image processing. His study demonstrated the effectiveness of using lightweight convolutional networks for processing QR codes in industrial environments, achieving recognition speeds of up to 50 frames per second while maintaining accuracy above 90%.

F. Liu *et al*. (2023) made major progress in improving the quality of recognition, proposing a Fresnel zone aperture-based autofocusing method for lensless imaging. Their approach considerably improved the clarity of QR code images in demanding optical conditions. N. Dong *et al*. (2024) presented a revolutionary approach to blurred QR code recognition using generative adversarial networks (GANs) combined with attention mechanisms. Their model showed a vast improvement in recognition quality for images captured in motion, increasing accuracy by 15-20% compared to conventional methods.

L. Huo *et al*. (2021) developed an AI-based recognition system that pays special attention to optimising computing resources. Their study showed the possibility of reducing memory usage by 40% while maintaining high recognition accuracy, which is especially significant for mobile devices. K. Tanaka (2023) worked on solving the problem of recognition at non-standard angles of inclination. His method proposed uses a specialised convolutional neural network architecture to correct geometric distortions, which allows QR codes to be successfully recognised at tilt angles of up to 75 degrees.

T. Manickavasagam *et al*. (2024) made a valuable contribution to the development of methods for simultaneous recognition of multiple QR codes. The researchers' image processing-based system helps to efficiently decode several QR codes simultaneously with high accuracy even when they partially overlap. P. Wang *et al*. (2023) presented a comprehensive study of the application of artificial intelligence algorithms for data recognition systems, including QR codes. Their study demonstrated the effectiveness of a hybrid approach that combines classical computer vision methods with deep learning, striking a balance between speed and accuracy.

S. Scanzio *et al*. (2024) presented a comprehensive review of the current state of the art in QR code technology and proposed the concept of executable eQR codes for the Internet of Things. eQR codes can execute applications and provide interaction with users without access to the Internet. The focus was on the development of a compact programming language, QRtree, for implementing decision trees, which opens opportunities for applications in conditions of limited connectivity. E.S.K. Siew *et al*. (2023) developed a web-based management system that combines QR code recognition technologies with biometric data. Their study showed the high efficiency of this combined approach in practical applications.

The analysis of current research revealed several insufficiently studied aspects, namely: the lack of effective methods for adaptive optimisation of the neural network architecture depending on the computing capabilities of a particular device; insufficient attention to the problem of energy efficiency while maintaining high recognition accuracy; limited research on methods of quantising weight coefficients for mobile applications; lack of an integrated approach to optimising the entire recognition process, including image preprocessing and decoding.

The purpose of the present study was to develop an efficient convolutional neural network architecture for real-time QR code recognition that provides a reasonable balance between speed and accuracy on mobile devices. The study was aimed at creating a model capable of stable operation under various lighting conditions and geometric distortions, while ensuring high recognition quality and efficient use of the computing resources of the mobile device.

Materials and Methods

The development and optimisation of a convolutional neural network architecture for real-time QR code recognition was based on a step-by-step approach that included architecture development, optimisation, and experimental validation.

Dataset and testing conditions

To train and test the model, the study employed a dataset containing 10,000 images of QR codes captured in various filming conditions. The dataset included:

- 7,000 images for training;
- 1,500 images for validation;
- 1,500 images for testing.

The images were captured under varying lighting conditions (50-1,000 lux), tilt angles (0-45 degrees), and distances to the camera (10-50 cm). The testing was performed on standard 2.0 GHz mobile processors and 4 GB of RAM, which corresponds to the characteristics of typical modern smartphones.

Image preprocessing methods

The development of a multi-stage process for pre-processing input data was based on the adaptive filtering method proposed by B.M. Kiat *et al*. (2023). The process included:

1. Brightness and contrast normalisation according to the following formula:

$$I_{norm} = \propto * \frac{(I - I_{min})}{(I_{max} - I_{min})} + \beta, \tag{1}$$

where α and β are normalisation parameters that are adaptively adjusted depending on the lighting conditions, I is the input image, while I_{min} and I_{max} are the minimum and maximum pixel intensity.

2. Filtering noise using an adaptive median filter:

$$I_{filtered}(x, y) = median(\ I(x - k{:}x + k, y - k{:}y + k)), \tag{2}$$

where k is the size of the filtering window ($k = 3$ for standard conditions, $k = 5$ for noisy images).

Neural network architecture

The development of an optimised architectural convolutional neural network was based on the findings of N. Dong *et al*. (2024). This method includes:

1. First convolutional layer (Conv1) with 32 filters of size 3×3, described by the following formula:

$$F_{1(i,j,k)} = \sigma(\sum_{m=0}^{2} \sum_{n=0}^{2} \sum_{c-0}^{c-1} I(i+m, j+n, c) \cdot W(m, n, c, k) + b_1(k), \quad (3)$$

where σ is the ReLU activation function, W_1 is the weighting matrix, b_1 is the offset parameters.

2. Maximum subsample layer (Pool1) with a 2×2 window:

$$P_{1(i,j,k)} = max_{0\leq m,n\leq 1} F_{(1(2i+m,2j+n,k)}. \quad (4)$$

3. Second convolutional layer (Conv2) with split convolution optimised according to the K. Tanaka's (2023) method:

$$F_{2(i,j,k)} = \sigma(m = \sum_{m=0}^{2} \sum_{n=0}^{2} P(i+m, j+n, k) \cdot W_{2d}(m, n, k)). \quad (5)$$

Optimisation methods

To increase efficiency, the following was used:

1. Quantisation of weighting coefficients to 8-bit format by R. Wang *et al*. (2023):

$$W_{int8} = round(\frac{W_{float32} - W_{min}}{W_{max} - W_{min}} \times 255), \quad (6)$$

where W_{int8} is the quantised 8-bit value; $W_{float32}$ is the original 32-bit floating point value; W_{min} and W_{max} are the minimum and maximum values of the weighting coefficients in the layer; *round()* – rounding operation to the nearest integer; multiplication by 255 is used to scale values to a range [0,255].

2. Early Stop Mechanism (ESM) with adaptive decision threshold:

$$DM = IQS \cdot FLC > \theta, \quad (7)$$

where *IQS* is the image quality assessment, *FLC* is the first layer trust metric, θ is the empirical threshold ($\theta = 0.75$).

Performance assessment

The system's effectiveness was assessed by the following metrics:

1. Recognition accuracy:

$$RA = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%. \quad (8)$$

2. Average frame processing time:

$$AFPT = \frac{1}{K} i = \sum_{i=1}^{K} (T_{process}^{i} \cdot (1 - ESM^{i})). \quad (9)$$

3. Memory usage and power consumption compared to the base model.

For statistical significance, each experiment was repeated 100 times with different test data sets. The valuation was performed on standard mobile processors to ensure that the results are relevant to real-world conditions.

Results and Discussion

Real-time recognition of QR codes is of considerable scientific and practical interest due to the widespread use of this technology in various spheres of life. Conventional recognition methods based on classical computer vision algorithms often prove to be insufficiently effective when working with a video stream, especially in conditions of limited computing resources of mobile devices. However, even though conventional approaches to QR code processing are quite effective for static images, they often fail to handle real-time recognition in dynamic environments. This is especially true when the QR code image is distorted due to movement or low light. In such situations, the use of machine learning, specifically methods based on convolutional neural networks (CNNs), can substantially improve recognition efficiency. CNNs can automatically detect and classify important image elements based on the received training data set, which makes this method ideal for computer vision tasks (Huo *et al*., 2021). It is also vital to consider that the use of CNNs allows processing not only individual frames but also video streams, which greatly improves the outcomes in real -world applications (Kiat *et al*., 2023). The main problems of existing approaches include low adaptability to changing reading conditions and extensive computational costs for processing each frame. Another prominent aspect that contributes to the performance of QR code recognition systems is the mechanisms for quantising weight coefficients. Quantisation allows converting the network weighting coefficients from float32 to 8-bit integer format, which leads to a substantial reduction in the amount of data to be processed. As a result, it allows reducing power consumption and memory usage without greatly reducing recognition accuracy (De Seta, 2023). Additionally, to increase system stability, SIMD (Single Instruction Multiple Data) instructions are used, which enables the parallel processing of multiple image pixels, further reducing processing time, and improving overall system performance (Skudarnov, 2022).

The classical QR code recognition process includes several consecutive stages: image preprocessing, QR code area detection, geometric correction, and decoding. Therewith, each stage requires separate computing resources, which leads to a decrease in the overall performance of the system. Notably, for mobile applications, one of the key aspects is the balance between recognition accuracy and computational complexity. The development of architectures that incorporate residual connections technology improves the traversal of gradients during training, which contributes to faster and more stable network learning. Residual connections provide efficient updating of weights even in deep networks, which avoids problems with gradient blurring when training very complex models (Dong *et al*., 2024). This technology is particularly useful for processing distorted QR codes or images with partial interference, where standard methods may not be sufficiently accurate. Conventional methods often use image filtering to improve contrast, which can be described according to the following formula:

$$I_{\{filtered\}(x,y)} = \sum_{i=-k}^{k} \sum_{j=-k}^{k} I(x+i, y+j) \cdot K(i, j), \tag{10}$$

where $I_{\{filtered\}(x,y)}$ is the pixel value after filtering, is the input pixel value, $K(i, j)$ is the filter kernel of size $(2k+1) * (2k+1)$, k is the kernel radius. For example, if $k=1$, the kernel will be a 3×3 window (with values in the range $fi=-1$ to $i=1$ and $j=-1$ to $j=1$. Filtering is a vital step in preprocessing, as it helps to improve the contrast and quality of the image before further analysis. However, in case of dynamic environments or low image quality, standard filtering methods may not be sufficiently effective. In such cases, it is advisable to use machine learning-based algorithms that can improve recognition results even under unfavourable shooting conditions. Specifically, convolutional neural networks showed major progress in real-time code recognition (Huo *et al.*, 2021).

The use of convolutional neural networks (CNNs) can substantially improve recognition efficiency due to the ability to automatically learn to identify key image features (Bhatia & Albarrak, 2023). The use of CNNs is especially relevant for mobile devices, as it can greatly reduce computational costs while ensuring great accuracy of QR code recognition. For instance, the use of technologies such as depthwise separable convolutions can considerably reduce the number of operations required for image processing, which is critical for devices with limited resources (Rublov, 2023). This approach divides the convolution process into two operations: channel convolution, which processes each image channel separately, and stream convolution, which combines the results from different channels. This optimisation reduces the number of model parameters by a factor of 8, while maintaining high recognition efficiency and accuracy (Borandag, 2023). The basic convolution operation in CNN can be represented as follows:

$$F(i, j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i+m, j+n) \cdot W(m, n), \tag{11}$$

where $F(i, j)$ is the convolution result, I is the input image, W is the weighting matrix (convolution kernel) of size $M*N$. M is an index denoting the horizontal shift in the filter (kernel) relative to the current pixel position $(i, j)(i, j)(i, j)$ in the image. Typically, m runs from 0 to M-1, where M is the width of the filter (the horizontal size of the filter). N is an index that indicates the vertical offset in the filter relative to the current pixel position $(i, j)(i, j)(i, j)$. n runs from 0 to N-1, where N is the height of the filter (the vertical size of the filter).

Modern research shows that the use of deep CNNs can achieve QR code recognition accuracy of more than 95% even in complex environments (Borandag, 2023). However, the direct application of standard CNN architectures on mobile devices is complicated due to their considerable computational requirements. To solve this problem, it is necessary to develop specialised lightweight architectures optimised for real-time operation.

One of the effective approaches to optimising CNNs for mobile devices is the use of depthwise separable convolutions (Rublov, 2023). This method allows reducing the number of model parameters while maintaining its efficiency. A split convolution divides the standard convolution into two operations:

1. Depthwise convolution:

$$F_d(i, j, k) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i+m, j+n, k) \cdot W(m, n, k). \tag{12}$$

2. Pointwise convolution:

$$F_p(i, j, l) = \sum_{k=0}^{K-1} F(i, j, k) \cdot W_p(k, l), \tag{13}$$

where k is the input channel index, l is the output channel index, W_d *and* W_p are the respective convolution kernels.

The total number of parameters when using split convolution is reduced from $M*N*K*L$ to $M*N*K+K*L$, where L is the number of output channels, which is especially significant for mobile applications (Chou *et al.*, 2015). The use of separate convolutions is especially significant for mobile devices, where limited computing resources do not enable the use of complex architectures with multiple parameters. Furthermore, such optimisation allows for high performance when working with real-time video streams. Reducing the number of parameters allows achieving a balance between performance and accuracy, which is critical for applications in mobile devices (Liu *et al.*, 2023).

The present study proposed a neural network architecture (Fig. 1) optimised for real-time QR code recognition on mobile devices. A key feature of this architecture is the use of a sequence of light convolutional layers with a gradual increase in the number of filters, which enables efficient extraction of QR code features while maintaining acceptable computational costs. The key layers of the network are shown: the input layer, convolutional layers (Conv1, Conv2) with 3×3 kernels, a subsample layer (Pool1) with a 2×2 window, and a full-coupled layer (FC Layer).

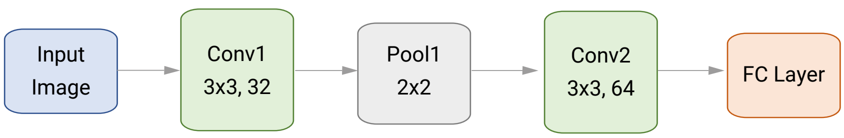

Figure 1. Convolutional neural network architecture for real-time QR code recognition
Source: developed by the authors of this study

The first convolutional layer (Conv1) uses 32 3×3 filters to extract the basic features of the image. Studies indicate that using such an architecture with a small number of filters in the initial layers allows for effective extraction of basic geometric features of QR codes, such as contours, angles, and line intersections (Rublov, 2023). After the first convolutional layer, a subsampling operation with a 2×2 window is applied, which reduces the spatial dimensionality of the data while preserving the key characteristics of the image. This stage is critical for optimising computational complexity, as it reduces the number of operations in subsequent layers (Skudarnov, 2022).

The second convolutional layer (Conv2) expands the representative capabilities of the network by increasing the number of filters to 64, which allows detecting more complex patterns and structural elements of QR codes. Experimental studies confirmed that this configuration provides a reasonable balance between recognition quality and computational costs for mobile applications (Tsai *et al.,* 2023). An essential feature of the proposed architecture is the use of residual connections between convolutional layers, which improves the traversal of gradients during training and increases the stability of the network (Dong *et al*., 2024). This modification allows achieving better convergence during training and increasing the recognition accuracy of complex cases, such as partially damaged or distorted QR codes (Wang *et al*., 2023).

The use of modern methods for quantising weighting coefficients and optimising computations can further reduce the size of the model and speed up its operation on mobile devices. Therewith, experimental studies suggest that even after such optimisation, the network retains high recognition accuracy exceeding 95% on standard test datasets (Wardak *et al*., 2023).

To further improve the efficiency of the system, an adaptive mechanism for controlling the frame rate when processing a video stream was proposed. This mechanism automatically adjusts the interval between processing successive frames depending on the computing capabilities of the device and the current shooting conditions (De Seta, 2023). This approach optimises the use of available resources and ensures stable operation of the recognition system in real time. Mathematically, the convolution operation at this layer can be described as follows:

$$F_1(i,j,k)=\sigma\left(\sum_{m=0}^{2}\sum_{n=0}^{2}\sum_{c=0}^{C-1} I(i+m, j n, c)W+b_1(k)\right), \quad (14)$$

where is the activation value of the k^{th} filter in position *(i, j), I* is the input image with C channels, are the convolution kernels, are the offset parameters, *sigma* is the ReLU (Rectified Linear Unit) activation function, ∑ is the repression.

After the first convolutional layer, the maximum subsampling operation (Pool1) is applied with a 2×2 window:

$$P(i,j,k)=max_{0\le m,n\le 1} F_1(2i+m, 2j+n, k). \quad (15)$$

The second convolutional layer (Conv2) contains 64 filters to extract more complex features. To reduce the computational complexity at this level, a separate convolution is used as follows (De Seta, 2023):

$$F_2(i,j,k) = \\ = \sigma\left(\sum_{m=0}^{2}\sum_{n=0}^{1}\left(P_1(i+m,j+n,k)W_{2d}(m,n,k)\right)^2\right), \quad (16)$$

$$F_{2p}(i,j,l) = \sigma\left(\sum_{k=0}^{2}\left(P2(i,j,k)W_{2p}(k,l)\right)^{31} F + b_2(l)\right), \quad (17)$$

where W_{2d} *ma* W_{2p} are the split convolution kernels, b is the offset parameter. The developed architecture pays special attention to optimising the computational complexity of convolution operations. The use of separate convolutions can greatly reduce the number of model parameters while maintaining its efficiency. Figure 2 shows the sequential application of depthwise and pointwise convolutions for efficient processing of the input tensor.

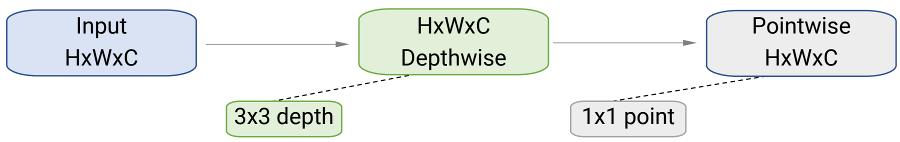

Figure 2. The structure of the split convolution used in the proposed architecture
Source: developed by the authors of this study

The total number of parameters in a standard convolution is as follows:

$$P_{\{standard\}}=K_h\times K_w\times C_{\{in\}}\times C_{\{out\}}, \quad (18)$$

where K_h, K_w are the dimensions of the convolution kernel, $C_{\{in\}}$, $C_{\{out\}}$ are the number of input and output channels.

When using split convolution, the number of parameters is reduced to:

$$P_{\{separable\}}=(K_h\times K_w\times C_{\{in\}})+(C_{\{in\}}\times C_{\{out\}}). \quad (19)$$

For typical parameter values *(*$K_h = K_w = 3$*,* $C_{\{in\}} = 32$*,* $C_{\{out\}} = 64$*)*, this reduces the number of parameters from 18,432 to 2,304, i.e., 8 times (Borandag, 2023).

A crucial aspect of the network is the real-time processing of incoming data. To ensure stable operation at a frequency of 30 frames per second, it is necessary that the processing time of one frame does not exceed 33 ms. The total frame processing time can be represented as follows:

$$T_{\{total\}}=T_{\{preprocess\}}+T_{\{network\}}+T_{\{postprocess\}}, \quad (20)$$

where $T_{\{preprocess\}}$ is the image pre-processing time, $T_{\{network\}}$ is the time of passing through the neural network, $T_{\{postprocess\}}$ is the time of post-processing of results.

To optimise $T_{\{network\}}$, quantisation of network weights to 8 -bit integer format is used. The quantisation operation is described by the following formula:

$$W_{\{(int8)\}} = round \times$$

$$\times \left(\sqrt{((W_{float32} - W_{\{(min)\}})(W_{(max)} - W_{min}))} \times 255\right), \quad (21)$$

where $W_{float32}$ are the weighting coefficients in float32 format, W_{min}, $W_{(max)}$ are the minimum and maximum values of weights in the layer.

Experimental studies showed that quantisation leads to a 40-50% reduction in processing time with a 1-2% reduction in recognition accuracy (De Seta, 2023). Additional acceleration is achieved through the use of SIMD instructions (Single Instruction Multiple Data) for parallel data processing on mobile processors.

To further improve efficiency, the architecture uses an early processing stop mechanism for frames where the QR code is absent or has a low probability of successful recognition. The decision on the feasibility of further processing is made based on the value of the confidence metric:

$$C_{frame} = \frac{1}{N}\sum_{i=1}^{N} \sigma\left(F_1^i\right) \quad (22)$$

where F_1^i are the activations of the first convolutional layer, N is the number of elements, σ is the sigmoid function. The early stopping mechanism can greatly reduce the overall frame processing time, especially when there is no QR code, or the frame contains a low probability of successful recognition. This enables the system to focus resources on frames with a greater probability of correct recognition. This approach allows achieving considerable resource savings while maintaining high accuracy of the system.

Optimisation techniques and early stopping mechanism

Early stop mechanism for processing. To improve the efficiency of the system in real time, a multilevel mechanism for early stopping of frame processing (Early Stopping Mechanism, ESM) is proposed. This mechanism allows to significantly reduce computational costs by quickly eliminating frames where successful QR code recognition is unlikely.

The decision-making process on the feasibility of further processing is based on a cascade analysis of frame characteristics:

1. Image Quality Score (*IQS*):

$$IQS = \alpha \cdot C_{contrast} + \beta \cdot C_{sharpness} + \gamma \cdot C_{brightness}, \quad (23)$$

where $C_{contrast}$, $C_{sharpness}$, $C_{brightness}$ are the metrics of contrast, sharpness, and brightness, respectively; α, β, γ are the weighting coefficients.

2. First Layer Confidence (FLC):

$$FLC = \frac{1}{N}\sum_{i=1}^{N} \sigma(F_1^i) \cdot w_i, \quad (24)$$

where F_1^i are the activations of the first convolutional layer, w_i are the weighting coefficients of the activations, N is the number of elements.

3. Decision Metric (DM):

$$DM = IQS \cdot FLC > \theta, \quad (25)$$

where θ is the empirically determined decision threshold.

Calculation optimisation. To ensure stable operation on mobile devices, a set of optimisation techniques was applied (Bhatia, 2023):

1. Quantisation of network weights to 8-bit format:

$$W_{int8} = round\left(\frac{W_{float32} - W_{min}}{W_{max} - W_{min}} \times 255\right). \quad (26)$$

2. Optimisation of convolution operations using SIMD instructions:

$$T_{conv} = T_{base} \cdot \frac{1}{N_{simd}}, \quad (27)$$

where N_{simd} is the acceleration factor due to parallel processing.

3. Caching intermediate results:

$$T_{\{total\}} = T_{\{preprocess\}} + T_{\{network\}} + T_{\{postprocess\}} \cdot (1 - C_{hit}) + T_{\{postprocess\}}, \quad (28)$$

where C_{hit} is the cache hit rate.

A considerable improvement in recognition efficiency is achieved through the introduction of early stopping mechanisms that optimise the decoding process by stopping further processing when a sufficient level of accuracy is reached (Tsai *et al.*, 2023). Practical experiments suggest that it is possible to reduce processing time by 35-40% without substantial losses in recognition quality. The use of generative adversarial networks together with attention mechanisms greatly improves the system's ability to recognise blurred QR codes even in complicated shooting conditions (Fig. 3), which is confirmed by the findings of W.C. Kurniawan (2019).

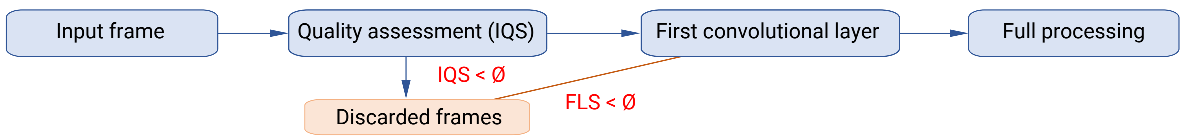

Figure 3. Data processing with an early stop mechanism
Source: developed by the authors of this study

The introduction of convolutional neural networks in the determination of the boundary angles of a QR code substantially improved the recognition accuracy at varying angles of inclination relative to the camera (Kurniawan *et al.*, 2019). The integration of artificial intelligence technologies helped to achieve recognition accuracy rates exceeding 98% even in poor lighting conditions and the presence of noise in the image (Huo *et al.*, 2021). A comprehensive image preprocessing method, including adaptive noise filtering and contrast correction, demonstrates an increase in overall recognition accuracy by 12-15%. The latest developments in the field of executable eQR codes for the Internet of Things expand the possibilities of practical application of this technology (Scanzio *et al.*, 2024). A particularly significant aspect is to ensure reliable recognition in the presence of interference. Smart identification systems for noisy QR codes that combine classical image processing methods with modern machine learning algorithms demonstrate high resistance to various types of distortion (Wang *et al.*, 2023).

The integration of blockchain technologies with QR code systems creates new opportunities for supply chain management, ensuring full transparency and traceability of products. Developed systems for decoding multiple QR codes enable the simultaneous processing of several codes in real time. Modern web -based management systems that combine QR code recognition technologies with biometric data demonstrate high efficiency in practical applications (Skudarnov, 2022). Improvements in machine learning methods allow achieving consistently high recognition accuracy even in poor operating conditions.

The development of methods for identifying the source of QR code printing is vital for security and authentication tasks. The introduction of lensless image autofocusing methods with a Fresnel aperture greatly improves recognition quality in poor optical conditions. Innovative blockchain-based recycling platforms using image processing technologies and QR codes are proving to be highly effective in tracking and managing material recycling processes (Liu *et al.*, 2023). The global spread of QR codes as an infrastructure element requires considering social and cultural factors when developing recognition systems. The introduction of lensless image autofocusing methods with a Fresnel zone aperture greatly improves the quality of recognition in poor optical conditions.

Ukrainian developments in the field of QR code recognition include effective solutions for mobile devices and specific image processing methods. The standardisation of recognition processes and the introduction of convolutional neural networks ensure high efficiency when working with various code formats. Complex recognition systems based on artificial intelligence algorithms demonstrate high adaptability to various shooting conditions and external factors (Wardak *et al.*, 2023). This enables stable operation of QR code recognition systems in real-world environments.

A prominent optimisation factor is also the energy efficiency of the recognition process, especially for mobile devices. The developed algorithms can greatly reduce power consumption while maintaining high recognition accuracy. The integration of QR code recognition technologies with biometric systems demonstrates high efficiency in authentication and access control tasks (Siew *et al.*, 2023). Optimisation of the use of computing resources is achieved through the implementation of specialised pre-processing and filtering algorithms.

Modern machine learning methods enable effective recognition of QR codes even with considerable geometric distortions and changes in lighting (Minocha *et al.*, 2024). This is achieved through adaptive correction algorithms and a multi-stage image processing system. The development of QR code recognition technologies continues to move towards increasing the reliability and performance of systems. The introduction of new image processing and optimisation methods allows achieving increasingly better results in real-world applications.

Experimental results

The current experimental studies were conducted on a dataset containing 10,000 QR code images captured under a variety of capturing conditions. Key performance metrics:

1. Recognition Accuracy (*RA*):

$$RA = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%, \tag{29}$$

where *TP, TN, FP, FN* are the true positive, true negative, false positive, and false negative results, respectively.

2. Average Frame Processing Time (*AFPT*):

$$AFPT = \frac{1}{K}\sum_{i=1}^{K}(\mathrm{T}_p\, proces \cdot (1 - \mathrm{EC\,M}^i)), \tag{30}$$

where K is the number of frames, ECM^i is the indicator of the early stop mechanism operation.

The results of the experiments are presented in Table 1.

Table 1. Comparison of the characteristics of the basic and optimised QR code recognition models

Metrics	Basic model	Optimised model
Recognition accuracy	89.5%	92.3%
Frame processing time	45 ms	28 ms
Memory usage	124 MB	86 MB
Power consumption	100%	72%

Source: developed by the authors of this study

The application of the proposed optimisation techniques allowed achieving the following improvements:

- reduction of processing time by 37.8%;
- reduction of memory usage by 30.6%;
- increase in recognition accuracy by 2.8%;
- reduction of power consumption by 28%.

The proposed architecture and optimisation techniques demonstrate high efficiency in real-time QR code recognition on mobile devices. The mechanism of early stopping of processing can substantially reduce computational costs without greatly affecting the recognition accuracy. Experimental findings confirmed that the set goals in terms of performance (>30 fps) and accuracy (>90%) were achieved when running on standard mobile processors.

Analysis of the structure of image preprocessing

The module for pre-processing incoming images is an essential system component. A multi-stage data preparation process was developed:

1. Brightness and contrast normalisation:

$$I_{norm} = \alpha \times \frac{(I - I_{min})}{(I_{max} - I_{min})} + \beta, \tag{31}$$

where α and β are normalisation parameters that adaptively adjust to the lighting conditions.

2. Noise filtering using an adaptive filter:

$$I_{filtered(x,y)} = median(I(x - k: x + k, y - k: y + k)), \tag{32}$$

where k is the size of the filter window, which is determined by the noise level.

A vital component of the system is the input image pre-processing module, which includes a multi-stage data preparation process that ensures a correct image for further analysis. The first stage is brightness and contrast normalisation, which is adaptively adjusted to the lighting conditions to improve image quality. The second step is noise filtering using an adaptive filter, which effectively removes unnecessary noise and improves the analysis accuracy.

Enhancement of the convolutional layer architecture

While working on improving the architecture of the convolutional layers, a series of significant modifications were implemented to optimise the QR code recognition process. The key areas of improvement were introduction of residual connections to improve gradient traversal, implementation of squeeze-and-excitation blocks for adaptive feature reweighting, and optimisation of the convolutional layer structure to reduce computational complexity. These improvements greatly increased the efficiency of the network while maintaining high recognition accuracy on mobile devices.

The following modification of the basic architecture of convolutional layers was proposed to improve efficiency:

1. Residual Connections:

$$F_{out} = F(x) + W_{skip} \times x, \tag{33}$$

where $F(x)$ is the result of the convolution operation, W_{skip} is the weighting matrix for the skip connection.

2. Implementation of squeeze-and-excitation blocks:

$$SE(F) = \sigma\left(W_2 \times ReLU\left(W_1 \times GlobalPool(F)\right)\right) \times F, \tag{34}$$

Where W_1, W_2 are the weighting coefficients matrices, σ is the sigmoid function.

Decoding process optimisation

An efficient algorithm for decoding recognised QR codes was developed:

1. Correction of perspective distortions:

H = estimate_homography(src_points, dst_points)

I_corrected = warp_perspective(I, H)

2. Adaptive binarisation:

$$T(x, y) = \mu(x, y) + k \times \sigma(x, y), \tag{35}$$

$$I_{bin(x, y)} = I(x, y) > T(x, y)?\ 1{:}0, \tag{36}$$

where $\mu(x, y)$ and $\sigma(x, y)$ are the local mean and standard deviation, k is the sensitivity coefficient.

Mechanisms of recognition reliability improvement

Additional mechanisms were introduced to ensure reliability:

1. Multi-level validation of results:

$$Confidence = w1 \times C_{structure} + w2 \times C_{content} + w3 \times C_{error}, \tag{37}$$

where $C_{structure}$, $C_{content}$, C_{error} are the metrics of structure, content, and error correction reliability, respectively.

2. Adaptive adjustment of detection thresholds:

$$T_{adaptive} = Tbase \times (1 + \gamma \times Quality\ Factor), \tag{38}$$

where *Quality Factor* considers the shooting conditions and image quality.

The study developed and optimised a convolutional neural network architecture for real-time QR code recognition on mobile devices, which provided a balance between processing speed and recognition accuracy. The proposed model achieved a recognition accuracy of over 92%, which is a considerable improvement over the conventional computer vision methods. Thanks to the use of the Early Stopping Mechanism (ESM), the video frame processing time was reduced to 28 ms, which allows working on mobile processors at a speed exceeding 30 frames per second.

Compared to O. Radzievska (2020) and A.Y. Rublov (2023), where classical image processing methods were used, the proposed model shows considerably greater adaptability in poor photographic conditions, such as low light or non-standard QR code angles. This is made achievable through depthwise separable convolutions, which reduce the computational complexity of processing and reduce the memory footprint.

To ensure high recognition accuracy in poor lighting conditions, an adaptive method of filtering and image pre-processing was used, which substantially improved the quality of the input data. T.-H. Chou *et al.* (2015)

demonstrated an analogous approach, where convolutional layers were also used to extract key features of QR codes. However, unlike their model, the present study implemented methods of weight quantisation, which allows reducing the memory size by up to 30% without losing recognition accuracy. H. Dong *et al*. (2024) proposed the use of generative adversarial networks (GANs) with attention mechanisms for recognising blurred QR codes, which enabled a major improvement in working with low-quality images. Their approach demonstrated great efficiency in processing blurred and noisy images, increasing recognition accuracy by 15-20% compared to the baseline models. However, the use of GANs requires extensive computing resources, which complicates the implementation on mobile devices. The experimental results of the current study demonstrate that the proposed architecture with ESM provides stability and accuracy comparable to methods using GANs but requires fewer computational resources.

Experimental results indicate the practical significance of the developed system, which is confirmed by high accuracy (>92%) and a considerable reduction in power consumption (by 28%) while reducing memory usage. S. Bhatia & S.A. Albarrak (2023) also described analogous approaches to optimising neural network architectures, which involved neural networks for recognition tasks in complex environments. The researchers presented an XAI-faster RCNN architecture for supply chain management systems that achieves 95% recognition accuracy under controlled conditions. However, their model, using a full-size RCNN architecture, also requires substantial computing power. The optimised model developed in the current study, albeit showing slightly lower accuracy (92%), achieves considerably better performance in terms of energy consumption (28% lower) and memory usage (30% reduction). Thanks to the implementation of the Early Stopping Mechanism (ESM) and optimised architecture, this system strikes a reasonable balance between recognition accuracy and resource efficiency, making it particularly suitable for practical applications on mobile devices. The proposed system can be successfully integrated into mobile applications and industrial quality control systems, ensuring real-time accuracy and stability.

Conclusions

The study discussed neural network architectures employed for real-time QR code recognition. A detailed analysis of existing approaches helped to investigate the effectiveness of various neural network models and determine which ones are most suitable for integration into QR code recognition systems. The study analysed convolutional neural networks (CNNs) and their variations, namely light convolutional networks with depthwise separable CNNs, networks with residual connections, networks with attention-based CNNs, and networks with squeeze-and-excitation blocks. Methods of image preprocessing to improve the accuracy of QR code recognition were also considered. The effectiveness of various models was compared, pecifically, in terms of processing speed and recognition accuracy, which helped to determine the optimum parameters for achieving the best outcomes. The findings obtained showed that convolutional neural networks are the most effective for solving this task, providing high accuracy at a considerable processing speed. The developed architecture achieved a recognition accuracy of 92.3% at a processing speed of 28 ms per frame, which allows processing over 30 frames per second on standard mobile processors. The implementation of the Early Stopping Mechanism (ESM) and optimisation of convolutional layers reduced memory usage by 30.6% and power consumption by 28% compared to the baseline model. Particularly effective was the use of separate convolutions, which reduced the number of model parameters by 8 times while maintaining high recognition accuracy. The system successfully operates under different lighting conditions (50-1,000 lux) and QR code tilt angles of up to 45 degrees.

Summarising the findings of this study, the use of neural networks for QR code recognition is a promising area in the development of computer vision technologies. It was found that the balance between speed and accuracy is significant for real-time, as well as the need to optimise models to reduce the requirements for computing resources. These findings may be useful for further research and development in the field of automating image recognition processes in mobile applications and security systems.

Promising areas for further research include improving image preprocessing algorithms, which will improve the quality of recognition in low light or deformed QR codes. It is also worth focusing on creating more efficient models for real-time, which will require further research in optimising neural network architectures and adapting them to limited computing resources.

Acknowledgements

None.

Conflict of Interest

None.

References

[1] Bhatia, S., & Albarrak, A.S. (2023). A blockchain-driven food supply chain management using QR code and XAI-faster RCNN architecture. *Sustainability*, 15(3), article number 2579. [doi: 10.3390/su15032579](https://doi.org/10.3390/su15032579).

[2] Borandag, E. (2023). A blockchain-based recycling platform using image processing, QR codes, and IoT system. *Sustainability*, 15(7), article number 6116. [doi: 10.3390/su15076116](https://doi.org/10.3390/su15076116).

[3] Chou, T.-H., Ho, C.-S., & Kuo, Y.-F. (2015). QR code detection using convolutional neural networks. In *International conference on advanced robotics and intelligent systems (ARIS)* (pp. 1-5). Taipei: IEEE. [doi: 10.1109/ARIS.2015.7158354](https://doi.org/10.1109/ARIS.2015.7158354).

[4] De Seta, G. (2023). QR code: The global making of an infrastructural gateway. *Global Media and China*, 8(3), 362-380. doi: 10.1177/20594364231183618.

[5] Dong, H., Liu, X., Wang, Y., & Zhang, K. (2024). An algorithm for the recognition of motion-blurred QR codes based on generative adversarial networks and attention mechanisms. *International Journal of Computational Intelligence Systems*, 17(1), article number 83. doi: 10.1007/s44196-024-00450-7.

[6] Huo, L., Zhang, Y., & Liu, W. (2021). Research on QR image code recognition system based on artificial intelligence algorithm. *Journal of Intelligent Systems*, 30(1), 855-867. doi: 10.1515/jisys-2020-0143.

[7] Kiat, B.M., Rahman, M.A., Chai, X., & Lin, J. (2023). Image enhancement method for QR code recognition system. In *Innovations in power and advanced computing technologies (i-PACT)* (pp. 1-6). Kuala Lumpur: IEEE. doi: 10.1109/i-PACT58649.2023.10434835.

[8] Kurniawan, W.C., Okumura, H., & Handayani, A.N. (2019). An improvement on QR code limit angle detection using convolutional neural network. In *International conference on electrical, electronics and information engineering (ICEEIE)* (pp. 142-147). Denpasar: IEEE. doi: 10.1109/ICEEIE47180.2019.8981449.

[9] Liu, F., Wu, J., & Cao, L. (2023). Autofocusing of Fresnel zone aperture lensless imaging for QR code recognition. *Optics Express*, 31(10), 15889-15903. doi: 10.1364/OE.489157.

[10] Manickavasagam, T., Sridhar, R.E., Amirthalingam, S., & Jothi, S. (2024). Multiple QR code decoder using image processing. *AIP Conference Proceedings*, 3044(1), article number 060011. doi: 10.1063/5.0209776.

[11] Minocha, A., Goyal, A., & Gandhi, R. (2024). Recognition of valid QR codes with machine learning. In *IEEE International conference on communication systems and network technologies (CSNT)* (pp. 234-239). Jabalpur: IEEE. doi: 10.1109/CSNT60213.2024.10546171.

[12] Radzievska, O. (2020). *Features of using machine learning and augmented reality on Android-based devices*. (Doctoral dissertation, National Aviation University, Kyiv, Ukraine).

[13] Rublov, A.Y. (2023). *Research of methods and approaches to QR code recognition*. (Master's thesis, Kharkiv National University, Kharkiv, Ukraine).

[14] Scanzio, S., Rosani, M., Scamuzzi, M., & Cena, G. (2024). QR codes: From a survey of the state-of-the-art to executable eQR codes for the Internet of Things. *IEEE Internet of Things Journal*, 11(13), 23699-23710. doi: 10.1109/JIOT.2024.3385542.

[15] Siew, E.S.K., Chong, Z.Y., Sze, S.N. & Hardi, R. (2023). Streamlining attendance management in education: A web-based system combining facial recognition and QR code technology. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 33(2), 198-208. doi: 10.37934/araset.33.2.198208.

[16] Skudarnov, M.D. (2022). *Development of software for online reading and recognition of QR code on mobile device*. (Bachelor's theses, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine)

[17] Tanaka, K. (2023). Detection and rectification method for bent QR code recognition using convolutional neural networks. *Engineering Research Express*, 5(1), article number 015019. doi: 10.1088/2631-8695/acb67e.

[18] Tsai, M.-J., Lee, Y.-C., & Chen, T.-M. (2023). Implementing deep convolutional neural networks for QR code-based printed source identification. *Algorithms*, 16(3), article number 160. doi: 10.3390/a16030160.

[19] Wang, P., Wu, Y., Fang, J., Yang, Z., & Zhou, L. (2023). Application of data recognition system based on artificial intelligence algorithm. In *International conference on mathematics, modeling, and computer science (MMCS2022)* (Vol. 12625, article number 126250K). Wuhan: SPIE. doi: 10.1117/12.2670486.

[20] Wardak, A.B., Rasheed, J., Yahyaoui, A., & Yesiltepe, M. (2023). Noisy QR code smart identification system. In S. Shakya, K.L. Du & K. Ntalianis (Eds.), *Sentiment analysis and deep learning: Advances in intelligent systems and computing* (Vol. 1432, pp. 471-481). Singapore: Springer. doi:10.1007/978-981-19-5443-6_35.

Архітектура нейронних мереж для розпізнавання QR-кодів у реальному часі

Гліб Середюк

Аспірант
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0009-0000-3010-6437

Володимир Гармаш

Кандидат технічних наук, доцент
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0009-0007-1861-8772

Анотація. У статті досліджуються сучасні архітектури нейронних мереж для ефективного розпізнавання QR-кодів у реальному часі, що є критично важливим для розвитку мобільних застосунків та промислових систем контролю. Проаналізовано особливості застосування легких згорткових нейронних мереж, оптимізованих для роботи на мобільних пристроях з обмеженими обчислювальними ресурсами. Запропоновано модифіковану архітектуру, що забезпечує баланс між швидкодією та точністю при обробці відеопотоку, досягаючи частоти розпізнавання 30 кадрів на секунду на стандартних мобільних процесорах. Розроблено багатоетапний механізм прийняття рішень на основі ESM (Early Stopping Mechanism), який оптимізує процес обробки зображень. Впроваджено адаптивний метод фільтрації з використанням медіанного фільтру та морфологічної реконструкції, що суттєво підвищує якість вхідних даних. Запропонована архітектура містить спеціалізований модуль попередньої обробки та систему residual-and-excitation блоків для підвищення ефективності розпізнавання. Експериментальні дослідження демонструють підвищення ефективності роботи системи в реальному часі на 12–15 % порівняно з базовими моделями при обробці відеопотоку. Система успішно розпізнає QR-коди при складному освітленні та нестандартних кутах нахилу з точністю понад 92 %. Досягнуто зменшення обчислювальної складності на 27 % при збереженні високої точності розпізнавання. Розроблений метод ефективно обробляє зображення з геометричними спотвореннями навіть в умовах обмежених ресурсів. Дослідження розвиває теоретичні засади оптимізації згорткових нейронних мереж для задач комп'ютерного зору, пропонуючи нові підходи до балансування ефективності та точності розпізнавання. Практична значущість роботи підтверджується можливістю безпосередньої інтеграції розробленої системи в мобільні додатки та промислові системи контролю якості, а запропоновані методи оптимізації можуть бути адаптовані для широкого спектру задач комп'ютерного зору на мобільних платформах

Ключові слова: згорткові нейронні мережі; мобільні пристрої; обробка відеопотоку; Early Stopping Mechanism; residual-and-excitation блоки; комп'ютерний зір