IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 5 September 2023, accepted 15 October 2023, date of publication 18 October 2023, date of current version 25 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3325670

RESEARCH ARTICLE

Inheritance and Innovation Development of Sports Based on Deep Learning and Artificial Intelligence

RUI ZHOU AND FAN WU
Nanchang University College of Science and Technology, Gongqingcheng 330029, China
Corresponding author: Fan Wu (13576995856@sohu.com)

This work was supported by the 2018 Humanities and Social Science Research Project of Universities in Jiangxi Province ''Research on Accelerating the Inheritance and Innovation Construction of Sports Towns in the New Era'' under Grant TY18116.

ABSTRACT This work aims to speed up the intelligent construction of Sports and Leisure Characteristic Towns (SLCT) in the new era. AI and robot technologies have become the pillar of scientific research. First, the MobileNetV2 is simplified and improved using robot technology to design a lightweight Image Recognition (IR) system. The proposed IR system integrates the framework of Facial Recognition (FR) and Automatic License Plate Recognition (ALPR). Then, a hardware acceleration scheme for Convolution Neural Network (CNN) is designed to accelerate the calculation of the improved MobileNetV2-based IR system. Finally, an experiment is designed to verify the performance of the proposed IR system on FR and ALPR tasks. The results are summarized as follows: 1) FR performance, the proposed IR system achieves 99.46% and 99.23% recognition accuracy for CelebFaces and CASIA-Web face data sets on the Xilinx VC707 platform, respectively. Thus, it can recognize facial images efficiently and timely on devices without Graphics Processing Unit (GPU); 2) regarding ALPR performance, the proposed IR system uses NVIDIA Titan X GPU to process 960 × 640-pixel images with 30/Frames Per Second (FPS). Hence, it meets the requirements of real-time processing; (3) proposed hardware acceleration scheme is tested on LeNet to generalize the acceleration effect. Consequently, the proposed convolution acceleration scheme can lend to common CNN. Compared with Eyeriss, the performance of the LeNet convolution layer is significantly improved in efficiency and resource consumption. Therefore, the proposed acceleration scheme is excellent and feasible. In conclusion, the lightweight neural network can obtain high accuracy and good generalization. The proposed hardware acceleration scheme can accelerate CNN operation. The findings lay the foundation for developing robot technology and constructing the intelligent SLCT. It contributes to the inheritance and innovative development of intelligent towns.

INDEX TERMS Sports and leisure characteristic towns (SLCTs), convolutional neural network (CNN), artificial intelligence (AI), image recognition (IR), facial recognition (FR), automatic license plate recognition (ALPR).

I. INTRODUCTION

A. RESEARCH BACKGROUND AND MOTIVATION

With the continuous progress of society, scientific and technological research is showing a rapidly changing trend. Especially with the rapid emergence of robot technology, it provides strong technical support for the vigorous development of various industries and promotes the comprehensive progress of society. In this context, fully utilizing robot technology and optimizing social infrastructure has become one of the main goals leading the research trend. At the same time, as the core element of urban construction, the optimization and popularization of sports are particularly urgent. However, there is a general lack of sufficient attention and planning to

The associate editor coordinating the review of this manuscript and approving it for publication was Abdel-Hamid Soliman.

sports in urban construction, and there is an urgent need to deepen the integration of sports elements in urban development strategies. Urban construction should actively expand the scale and coverage of sports projects, creating favorable conditions for residents' physical and mental health, which will further consolidate the city's vision toward modernization.

In the construction direction of Sports and Leisure Characteristics Towns (SLCTs), the General Administration of Sport of China has issued a work notice, clearly proposing to actively support provinces (cities) and autonomous regions in building a batch of SLCTs with distinct sports characteristics, rich cultural connotations, sustainable ecological environment development, and high industrial integration. The construction goals of these SLCTs must be centered around meeting people's health and sports needs [1]. The advocacy of the General Administration of Sport of China occupies an important position in national policies, highlighting the important direction of national sports in the new era. The Chinese government hopes to promote innovation, transformation, and upgrading of the sports industry by building an alliance, thereby comprehensively promoting economic and social progress. In the new historical period, to achieve the grand strategic goal of SLCTs, it is necessary to deeply integrate urban construction with the Artificial Intelligence (AI) industry. AI devices, scenarios, and technologies will be widely used in all levels of the construction of SLCTs [2]. In recent years, deep learning (DL) technology has influenced almost every field. With the continuous advancement of DL technology, more advanced application scenarios and characteristic technologies have emerged. The intelligent construction of SLCTs must be people-oriented, reflecting the core concept of paying attention to the well-being of the people. Among the many basic services, technologies such as Facial Recognition (FR) and Automatic License Plate Recognition (ALPR) must be closely integrated with AI. Only by solving these basic technical problems can the sports and leisure service level of SLCTs obtain higher quality inheritance and innovation [3].

B. RESEARCH OBJECTIVES

This work focuses on optimizing the SLCTs based on AI and robot technology to improve their inheritance and innovative development. Robot technology is a high-tech integration with the computer, cybernetics, mechanism, information, sensing technology, AI, and bionics. It is a very active and increasingly widely used field in contemporary research. Applying robots is an important symbol of a country's industrial automation level. Robots do not replace human labor in a simple sense. For example, it can be a humanoid electronic and mechanical device that combines human and machine expertise. It can imitate people to respond to environmental conditions, analyze, and judge quickly. It inherits the advantages of machines working continuously and tirelessly, with high accuracy and robustness in harsh environments. In a sense, it is also the product of the evolution of machines, important production and service equipment in industry and non-industry, and indispensable automation equipment in advanced manufacturing technology. AI is a new technological science that studies and develops theories, methods, technologies, and application systems for simulating, extending and expanding human intelligence. AI is a branch of computer science aiming to understand the essence of intelligence and produce a new intelligent machine that can respond similarly to humans. The research in the AI field includes robots, language recognition, Image Recognition (IR), Natural Language Processing (NLP), and Expert Systems (ESs). Since its birth, AI theory and technology have become increasingly mature, and the application field has also been expanding. It will eventually become the main force of future scientific and technological construction. Here, an SLCT is defined as a form of Characteristic Town under the guidance of the state's dual policies of promoting the development of Characteristic Towns and the sports industry. Furthermore, the SLCT is also an important carrier and starting point for developing the sports industry. It follows the national policy of the gradual normalization of integrating leisure, fitness, and sports in urban construction. The SLCT development plays an important role in national health.

According to the above research, this work applies AI and robot technologies in SLCT research, which is of breakthrough significance. First, an IR system is proposed based on the Deep Convolution Neural Network (DCNN) and considers the most basic application scenarios of FR and ALPR in SLCT. The proposed IR system is verified on different data sets. Simultaneously, a hardware acceleration scheme is proposed against the problem of slow hardware operation caused by complex algorithm parameters in IR. The innovation is to propose an IR system that integrates FR and ALRP algorithms with a hardware acceleration scheme. The difficulty of the research is to fuse different basic algorithms with DCNN. The find extends the application of IR based on AI and provides an important reference for realizing intelligent and comprehensive services and developing and inheriting various tourism cultures in SLCT.

II. LITERATURE REVIEW

The inheritance and innovative development of SLCTs is imperative, driving the development of AI and robot technology in future urban construction. It also provides an impetus for the comprehensive development of society. Although the current application of AI and robot technologies in the inheritance and innovative development of SLCTs is not perfect, many studies have provided technical support.

FR and ALPR are essentially image processing technology. Many scholars have studied image processing technology. To name a few, Hu et al. proposed a brain image processing and brain disease diagnostic model. They verified the model through the simulation based on the brain Nuclear Magnetic Resonance Images (NMRIs) in hospitals [4]. Xie et al. constructed a DL method to process hospital pathological images and predict diseases [5]. Uttaro et al. designed an image analysis method for identifying marbling on the surface of bare boneless pork loin. They solved the problem of the uneven color of lean meat [6]. Kar and Neogi built a new local triangular loop binary mode on understanding the local variation of face radius of local rotation [7]. Some scholars' AIe research on Smart Cities can also provide a reference for constructing the SLCTs. For example, Lv et al. modeled the Smart City vertical market system from an economic perspective. They analyzed the security and confidentiality of the model through simulation. The Smart City vertical market details were discussed from an economic perspective [8]. DL algorithm is the most applied research in Machine Learning (ML), and many researchers have applied it in related research. Yu et al. found that applying DL had promoted the Factor decomposition Machine (FM) in the recommendation system. DL and FM could learn high-order and low-order feature combinations, respectively. By comparison, with a hidden vector system, FM could learn information from sparse data to solve some classical models [9]. Al-Antari et al. used a fast DL computer-aided diagnosis method based on digital chest X-ray images for the Corona Virus Disease 2019 (Covid-19) [10].

In summary, the current AI and robot technology have developed very maturely. However, their application in urban construction is not perfect, so more research is needed to provide a reference for their development.

III. RESEARCH METHODOLOGY

A. SLCT

With the release of new urbanization policies and social development, sports tourism gives birth to the concept of SLCT [11]. Scholars classify the SLCT from target customers and industrial attributes into four types: event-oriented, industry-oriented, fitness-oriented, and leisure-oriented SLCT [12]. This paper mainly studies the combination of SLCT and AI technology [13].

A sound ecological environment lays a solid foundation for the SLCT infrastructure. Generally, SLCT flourishes on a tourism industry chain based on specific scenic spots. The SLCT holds various modern, high-participatory, and popular sporting events with a tourism chain and diversified industrial agglomeration services around the Double Line of sports and tourist attractions. Therefore, in the initial stage of construction, high standards are implemented in infrastructure construction, advantageous regional factors, and sports culture construction. Fig. 1 shows the agglomeration diagram of SLCT.

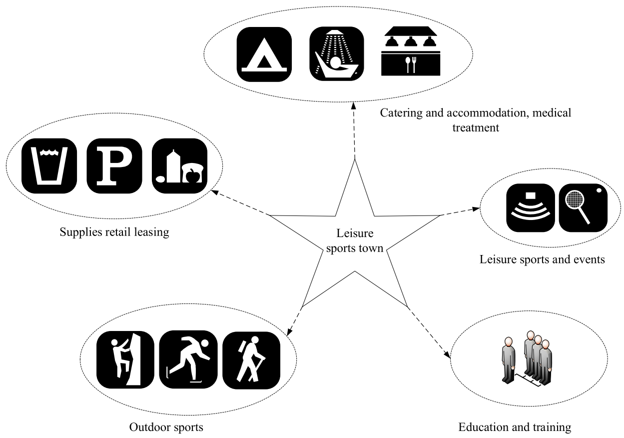

FIGURE 1. Schematic diagram of SLCT agglomeration.

B. THEORETICAL BASIS OF SLCT CONSTRUCTION

(1) Industrial convergence theory

"Industrial convergence" is a phenomenon in that technology extends and integrates to the edge or center. Its motivation is the continuous evolution and development of industrial tools [14], which is also the main content of industrial economics theory. Industrial convergence concerns the country's economic transformation and development and industrial up-gradation, such as constructing new towns, particularly SLCT.

The SLCT industry increasingly fuses with modern service industries, such as culture, tourism, and the Internet, as an emerging and promising industry. This integration is mainly reflected in technology and business. Therefore, it is necessary to establish relevant technical standards timely and understand the importance of introducing compound talents of culture, tourism, and sports products. This paper integrates sports with the Internet and AI technologies. It employs the DL algorithm of ML to realize the convenient, data-based, standardized, and intelligent construction of various service items in the leisure industry of SLCT, thereby developing a new sustainable "sports +" model [15].

(2) New urbanization theory

New urbanization theory can combine with the sports industry and jointly promote SLCT development. Site selection around the city and rational development of natural resources can link the city and the countryside [16]. New urbanization emphasizes people-oriented, green, coordinated, and sustainable development. SLCT requires innovative construction concepts to protect the ecological environment and supplement new characteristic towns' construction.

C. APPLICATION OF ROBOT TECHNOLOGY

Robot technology was born in the United States in 1963 with the world's first industrial robot. After years of development, America has become one of the world's top robot powers. Throughout the history of the development of robot technology, the road has been tortuous. Robotics is an interdisciplinary branch of computer science and engineering. Robot technology involves robot design, construction, operation, and use. Robotics aims to design machines that can help human beings. Robots integrate mechanical, electrical, and information engineering. It also involves mechatronics, electronics, bioengineering, computer, control, and software engineering, as well as mathematics.

Machines developed by robotics can replace humans and replicate human behaviors. Robots can be used for many purposes in different cases. Nowadays, many robots are used in dangerous environments (checking radioactive substances, bomb detection, and deactivation). Sometimes, they can be found in manufacturing processes. Places where human beings cannot survive (space, underwater, high-temperature scenes, and places that need cleaning up and curbing harmful substances and radiation) can employ robots. Robots can take any form, and some look similar to humans. It is believed that humanoid design helps robots accept certain replication behaviors that humans usually perform. This kind of robot replicates walking, weightlifting, speaking, cognition, or other human activities. Many robots today are inspired by nature and have contributed to the field of bionic robots.

Although robot technology has made great progress, in the future, human beings will be more suitable for some skills than machines. The argument is about achieving the best combination of human and robot skills. The advantages of robot technology include heavy work with accuracy and repeatability. In contrast, human advantages include creativity, decision-making ability, flexibility, and adaptability. Combining the best of both has led to collaborative robots and humans sharing a common workspace and developed new methods and standards to ensure the safety of Man-Machine Integration. Therefore, this work integrates robot technology and AI to optimize SLCT construction, which also plays an essential role in the development of robots in the future.

D. AI TECHNOLOGY

AI is an intelligence that is realized in machines by artificial means [17]. As early as 1955, international researchers proposed the concept of AI, which aims to simulate human thought processes through machines, so that machines can perform specific human activities [18]. ML, pattern recognition, and NLP are the core contents of the AI field. However, in recent years, with the enhancement of computing power and the large-scale application of data, DL has gradually become an important branch in the field of AI, leading the further development of AI.

DL is a vital branch of ML, and its core is the neural network model. The DL model simulates the connection between neurons in the human brain to construct multi-layer neural networks, thereby achieving high-level abstraction and representation learning of data. This makes DL perform well in processing complex and large-scale data, becoming a key technology in fields such as IR and NLP [19].

DL technology has shown a wide range of application prospects in today's practical applications. For example, tasks such as translation, speech recognition, and IR have all benefited from breakthroughs in DL. DL technology can automatically extract features from images in image processing, thereby achieving efficient image classification and recognition. At the same time, it has also demonstrated excellent performance in NLP, such as text sentiment analysis and semantic understanding.

In addition to DL, AI technology has also presented widespread applications in other fields. For example, infrared and wireless positioning technologies strongly support object recognition and positioning. The continuous development and innovation of these technologies have brought new possibilities and opportunities to multiple fields, while also providing a solid foundation for the intelligent construction of SLCTs.

E. DL ALGORITHM

The core of DL lies in the Deep Neural Network (DNN). The neural network layers can now increase significantly thanks to the development of electronic equipment, especially the upgrading of the Graphics Processing Unit (GPU); meanwhile, various data sets can find training ideas in the newly born neural networks [20]. In particular, CNN can process comprehensive data sets with spatial depth, such as images and videos [21].

DL is a summary of ML methods; the essence of learning is algorithm design. The pixels and audio signals in image data or sound data are the original data of the algorithm, and the algorithm is the process of abstract display of the main features. Different presentations need different algorithms, and effective algorithms benefit learning [22]. The Activation Function (AF) can control the algorithm range and has distinct functions in various application environments. Here are three common AFs.

(1) Rectified Linear Unit (ReLU) function

ReLU function is simple and easy to understand, as shown in Eq. (1).

$$y = \max(x, 0) \tag{1}$$

Eq. (1) shows that ReLU is a simple maximum function. If x<0, output=0. ReLU AF can calculate the maximum real-time result in DL to provide a powerful reference for adjusting the model. In particular, Parametric Rectified Linear Unit (PReLU) improves and popularizes ReLU by solving the model overfitting problem without significantly increasing the calculation. Also, it has proven to have a faster convergence and lower error rate. Moreover, it proposes a more robust initialization method, which fully considers the nonlinearity of the rectifier unit. This method allows training deeper networks directly from scratch. In essence, PReLU is based on PReLU networks (PReLU-nets). It has achieved a top-5 error of 4.94% on the ImageNet-2012 classification dataset, thus surpassing humans in computer vision recognition tasks for the first time [8].

(2) Sigmoid function

The Sigmoid function features strong derivability and is applied to Binary Classification (BC) because its output falls within [0,1] and can be described by probability,

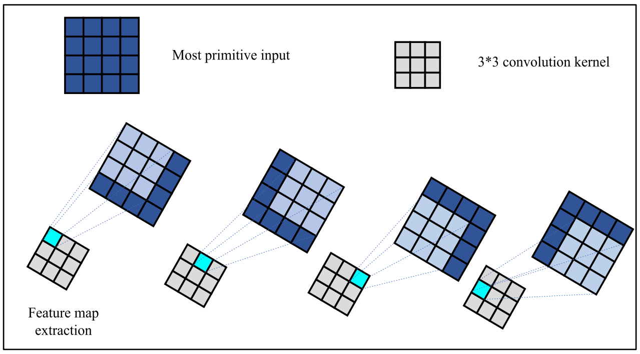

FIGURE 2. CNN convolution process.

as shown in Eq. (2).

$$y = \frac{1}{(1+e^{-x})} \tag{2}$$

The saturation region of the Sigmoid function exists when the x is infinitely away from 0. In that case, its derivative is close to 0.

(3) SoftMax function

The SoftMax function is often applied to multi-classifications and is necessary for the network output layer function in Reinforcement Learning (RL), as shown in Eq. (3).

$$y = \frac{e^{a}}{\sum_{i=0}^{k} e^{i}} \tag{3}$$

SoftMax function has a unique feature to apply to classification problems; that is, all classification probabilities sum up to 1.

F. CNN

CNN is a typical feed-forward neural network used to extract data containing spatial information, such as image and video recognition and classification [23]. The CNN comprises the convolutional, pooling, and fully connected layers. Fig. 2 depicts the schematic diagram of the convolution process.

The CNN algorithm is mainly composed of training and reasoning links. Large training sets can determine network parameters. The trained model can output the test set through the convolution model algorithm. Often, people design CNN's reasoning link as a hardware acceleration link. The convolution operation multiplies the convolution kernel by the input data elements at the corresponding position of the input matrix and then adds them to obtain the output. Table 1 shows the common CNN and the fixed parameters [24].

Depthwise-Separable Convolution (DSC) is a convolution operation. From the perspective of image processing, the convolution kernel is a Three-Dimensional (3D) filter that can process the depth and space dimensional data simultaneously. Generally, the CNN structure can simultaneously reflect the input data's depth and space dimensions. The design idea of DSC is: to convolute the received information layer by layer, modify the convolution process in general, and divide it into Depthwise and Pointwise processes. Fig. 3 depicts the schematic diagram.

TABLE 1. Parameter analysis of common CNN (Serial numbers a-d represent Yolo v3, ResNet-50, AlexNet, and VGG-16 neural networks, respectively).

	Convolution kernel size	Convolutional layer parameters	Multiply and add the number of fully connected layers	Number of channels except for the first layer	Fully connected layer parameters
a	1,3	61.8M	-	32-1024	-
b	1,3,7	23.4M	100.4M	64-2048	100.4M
c	3,5,11	2.3M	58.6M	64-256	58.6M
d	3	14.7M	130M	2-512	124M

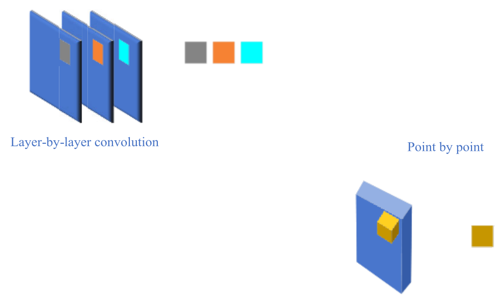

FIGURE 3. Evaluation system of economic quality development level.

The essence of DSC is the analytical separation of the initial (N * H * W * C)-dimensional convolution kernel, where N represents the number of input channels of convolution. H and W denote the size of convolution. C refers to the number of output channels. Convolution decomposition first decomposes into N * H * w * 1 and performs the convolution operation according to the DSC method in Fig. 3 to obtain the N-feature map. Then, the convolution is decomposed into N * 1 * 1 * C. The Pointwise operation is performed on the acquired N-feature map. Each 1 * 1 convolution kernel performs a weighted fusion of N features in the feature dimension to compensate for the loss of depth information caused by the first convolution step. Finally, the number of feature maps is C [25].

Eq. (4) presents the calculation of traditional convolution (D_F represents the size of the feature map):

$$F_{normal} = N \times H \times W \times C \times D_F \times D_F \tag{4}$$

Eq. (5) manifests the calculation of DSC:

$$F_{DW} = N \times H \times W \times D_F \times D_F + N \times C \times D_F \times D_F \tag{5}$$

Therefore, the ratio of Eq. (4) and Eq. (5) can obtain the compression ratio of DSC to convolution calculation,

TABLE 2. **Comparison of LeNet and the streamlined MobileNetV2 (t represents network expansion coefficient, and c stands for the number of convolution kernels).**

Input	MobileNetV2	224*3	112*16	56*24	14*96	7*1280
	LeNet	112*3	56*32	28*24	14*64	1*512
Operator	MobileNetV2	Conv2d	Bottleneck3	Bottleneck3	Bottleneck3B	Avgpool 7*7
	LeNet	Conv2d	Bottleneck5	Bottleneck5	Bottleneck3-SD	Conv2d 1*1
t	MobileNetV2	-	6	6	6	-
	LeNet	-	1	1	4	-
c	MobileNetV2	32	24	32	160	1
	LeNet	32	64	64	128	128

as expressed in Eq. (6):

$$\delta = \frac{F_{DW}}{F_{normal}} \tag{6}$$

The analysis of the above equations suggests that the DSC network reduces parameters and simplifies the calculation. The compression ratio (δ) of DSC to traditional convolution depends on the size of the convolution kernel. The larger the convolution kernel is, the greater the compression ratio is. Thus, a large convolution kernel can obtain a wider convolution range, and the comprehensive performance of the network is improved accordingly.

G. LIGHTWEIGHT AND EFFICIENT IR TECHNOLOGY

The performance analysis of the DCNN characteristics implies that the essence of the network operation is to filter the input image data step by step. The high-level image features will be more significant. At the same time, the high-level convolutional layer's processing effect will reduce due to repeated accumulation, with which many interfering convolution kernels are generated. To this end, this paper optimizes the network interception of MobileNetV2 and cuts off some high-level networks. Table 2 compares the specific network structures [26].

In Table 2, Input represents the dimension of input data, such as 224*3; that is, the original data input by the current layer is 224*224*3. Operator refers to the Block attribute of the network. For example, Conv2dl*1 represents the 2d convolutional layer, the current layer. The convolution kernel size is 1*1, and an unmarked convolution kernel is generally 3*3 by default. Due to the unique expansion residual structure of MobileNetV2, the selection of expansion coefficient must be considered in the network setting [27]. Therefore, using MobileNetV2 in lightweight and efficient IR can obtain more accurate results, reduce errors in FR and ALPR, and provide a more stable intelligent recognition system for constructing an intelligent SLCT. MobileNetV2 is based on the inverted residual structure. The ordinary residual structure first passes through the 1 × 1 convolution core to downsize the feature map channels, then passes through the 3 × 3 convolution

TABLE 3. **Parameter comparison of MobileNetV2.**

Network	Top 1	Params	MAdds	CPU
MobileNetV2(streamlined)	72.0	3.4M	300M	75ms
MobileNetV2(original)	70.6	4.2M	575M	113ms
ShuffleNet(1.5)	71.5	3.4M	292M	-
ShuffleNet(x2)	73.7	5.4M	524M	-
GhostNet 1.0×	73.9	5.2M	164M	-
GhostNet 1.3×	75.7	7.3M	226M	-
NasNet	74.0	5.3M	564M	183ms

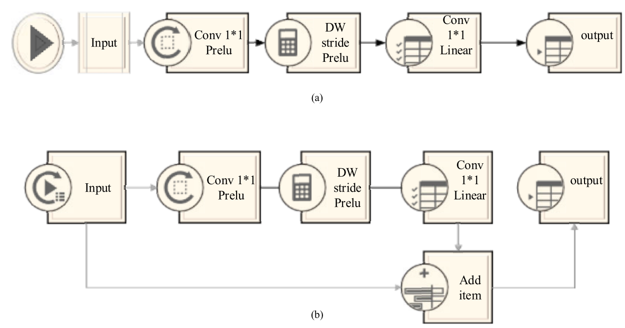

FIGURE 4. **Evaluation standard of high economic development level in S city in 2022.**

core, and finally uses the 1 × 1 convolution core to expand the channel number back. That is, it compresses first and then expands. In contrast, the inverted residual structure of MobileNetV2 expands first and then compresses. Additionally, it is vital to remove layers with few channels for linear activation [28]. Table 3 compares the MobileNetV2 and other network structures.

As listed in Table 3, the Params value is smaller for the MobileNetV2 structure. The MAdds value is also lower. Most importantly, the CPU has the shortest response time and faster calculation speed. Therefore, this paper uses MobileNetV2 as the primary network model [29]. It might help to remove downsampling from the Block at the low level to eliminate the accuracy decline caused by the continuous downsampling of the network because the amount of feature information of the low-level network is relatively large. Such operations can make the image background information monotonous and straightforward and, thus, downsize the image. Reducing the convolution kernel dimension can avoid incomplete learning and convolution kernel involution caused by rapid expansion. The above is the optimization idea of the MobileNetV2 network to make the network concise and efficient. Further, this section selects the parallel expansion mode to design the basic network structure to improve the model's parallel performance during channel expansion and reduce the difference in network input and output channels, as illustrated in Fig. 4.

As shown in Fig. 4, compared with the traditional CNN, the residual structure adds more calculation processes to the MobileNetV2 model. The residual structure can screen the images through the calculation to obtain more recognition accuracy from blurred images. The proposed improved

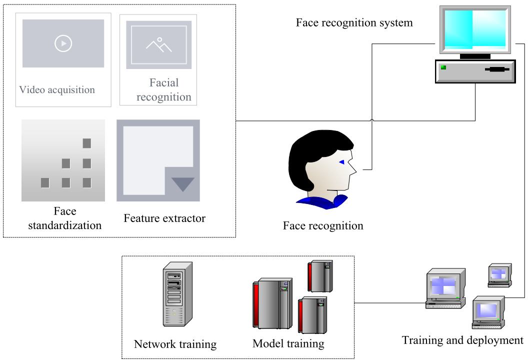

FIGURE 5. Schematic diagram of the MxNet-based FR framework.

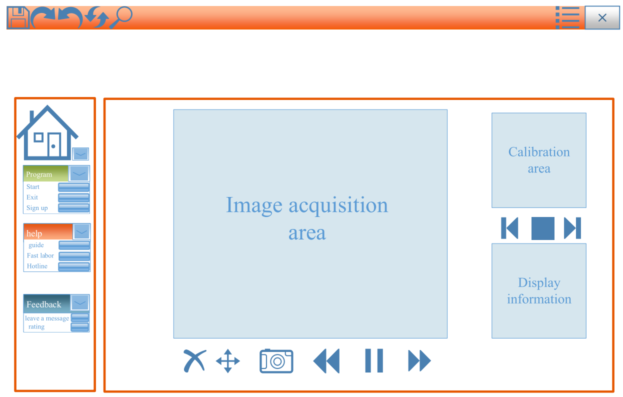

FIGURE 6. Interface of the MxNet-based FR framework.

MobileNetV2-based IR system has greater practical significance than the traditional CNN-based IR system.

IV. EXPERIMENTAL DESIGN AND PERFORMANCE EVALUATION

A. EXPERIMENTAL MATERIALS

With the development of the times, traveling has become very convenient, especially thanks to vehicle popularization. Thus, vehicle control has become one of the main tasks of social construction. At the same time, as professional and recreational sports develop fast, human sports activities also show more diversity. Therefore, in intelligent SLCT construction, vehicle control, and human activity detection have become the main concerns. For this reason, this section proposes an improved MobileNetV2-based IR system by fusing the FR and ALPR frameworks and implementing the two frameworks based on different DCNN models. The FR framework chooses the MxNet, and the ALPR framework employs the LeNet in MobileNetV2. Then, the proposed improved-MobileNetV2-based IR system will be tested through experimental design on different datasets for FR and ALPR. The FR framework includes network training deployment and face verification according to the CNN architecture. Firstly, the CNN model must be trained until the optimal model parameters are determined and are robust enough. Secondly, the well-trained model will be deployed for practical FR tasks. The model trained by Python can be flexibly called in a C++ environment [30]. In particular, the FR framework consists of face detection, facial data standardization, feature extraction, comparison, registration program, and system setting functional modules. Meanwhile, the FR framework is composed of a multi-layer computing structure. It can carry out multi-level computing on the face to accurately identify the user's identity. Additionally, it can also accelerate the image recognition speed through technological improvement. Fig. 5 gives the overall framework of the MxNet-based FR framework.

As in Fig. 5, the user can confirm his identity through the MxNet-based FR framework with simple operations. Thus, the MxNet-based FR work greatly reduces the user's operation process and saves the user time and resource costs. Noticeably, the MxNet-based FR framework's interface includes function selection and setting key area, acquisition display area, acquisition adjustment area, and help and suggestions, as drawn in Fig. 6.

From Fig. 6, the MxNet-based FR framework provides users with reliable computing technology. It also offers users a minimalist operation interface with a reliable computing process and excellent services. This section selects FR tasks to verify the MxNet-based FR framework's efficiency. Specifically, the robust and well-trained CNN model extracts the image features. Then, a 1:1 facial image verification is performed against the MxNet-based FR framework on the LFW (Labeled Faces in the Wild) dataset. The specific experimental steps are as follows:

(1) Detect and crop preprocessed LFW data set.

(2) Extract the features of the verification image in turn based on LeNet.

(3) The cosine distance of the feature of each pair of images is extracted by Eq. (7):

$$\cos\theta = \frac{X_1 \times X_2}{\|X_1\|\,\|X_2\|} \tag{7}$$

(4) Set the judgment condition: a value greater than the set range indicates a correct identification. Otherwise, it is judged as an incorrect identification.

(5) Operation of overall recognition rate

Receiver Operating Characteristic (ROC) curve can intuitively determine the network's performance on LFW data under different judgment ranges. True Positive Rate (*TPR*) and False Positive Rate (*FPR*) will affect ROC. The expressions of *TPR* and *FPR* read:

$$FPR = \frac{FP}{FP + TN} \tag{8}$$

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

In Eqs. (8) and (9), TP represents True Positive; FN stands for False Negative; FP refers to False Positive; TN indicates True Negative.

Here, the FR test refers to the DL framework of the Amazon Deep Learning library, namely, the MxNet. Then,

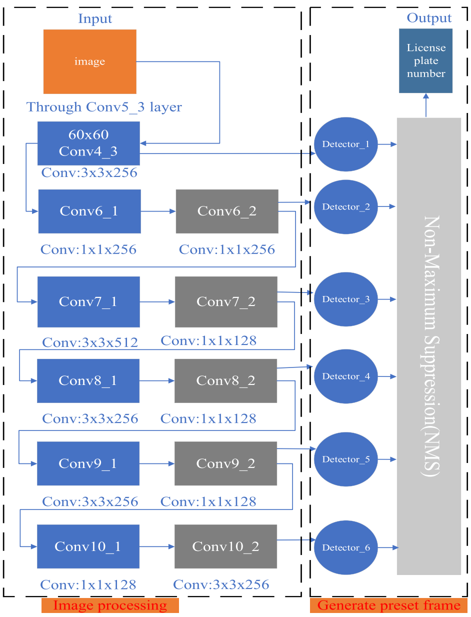

FIGURE 7. Schematic diagram of ALPR framework.

it uses Telsa v100 GPU (16G) for training. After systematic evaluation, the FR test obtains the network accuracy results under different amounts of data. The experiment mainly measures the ROC curve and compares the accuracy of different training stages. According to Eq. (9), FPR and TPR are the abscissa and ordinate of the ROC curve. The closer the upper left corner of the ROC curve is to the coordinate axis, the better the FR performance is. Lastly, the experiment analyzes the validation data and the multi-dimensional calculation process, providing technical support for the FR tasks. The results suggest that the MxNet-based FR framework's calculation accuracy has improved. Overall, the MxNet-based FR framework improves the comprehensive service efficiency in recognizing facial images.

B. EXPERIMENTAL ENVIRONMENT BASED ON DL

DL technology is a significant branch in the field of ML, which simulates the connection between neurons in the human brain by constructing multi-layer neural network models to achieve high-level abstraction and feature learning of data. In recent years, with the improvement of computing power and the widespread application of data, DL has gradually become one of the core technologies in AI field, offering strong impetus for the development of AI. For ALPR tasks, based on DL technology, this section integrates the Single Shot Multibox Detector (SSD) algorithm into the ALPR framework to adjust and optimize MobileNetV2.

The SSD model includes image processing and preset box generation based on Non-Maximum Suppression (NMS). Image processing calculates the collected images to extract inherent features and use them for ALPR. The preset box classifies the image features to prepare for subsequent analysis. In contrast, NMS filters the extracted image features. The filtering conditions are based on the maximum calculated image analysis results. Through multi-dimensional input calculation and two-level comparison and filtering process, the comprehensive calculation ability and the comprehensive calculation accuracy for ALPR are improved. The results provide technical support for the ALPR tasks. Fig. 7 sketches the whole network flowchart.

Fig. 7 gives the two functionalities of the ALPR framework: image processing and preset box generation. Apparently, the ALPR framework adopts multiple steps to recognize images. The results can be filtered by recognizing multiple processes, and the uncertain images can be processed multiple times. As a result, the final results are more accurate and have greater practical significance. Importantly, the SSD algorithm is used to simplify the LeNet structure in MobileNetV2 by cutting off the fully connected layer and retaining only five convolutional layers and five pooling layers. The feature images of each convolutional layer correspond to the feature extraction results of different layers. After feature extraction, the algorithm uses the preset box detector for detection. Then, it selects many possible desired target areas in advance. Next, the algorithm will filter and identify these areas again, judge and identify the actual target, and adjust the border with the preset box. Compared with Faster RCNN, the SSD-streamlined MobileNetV2 model has obvious speed advantages. Compared with YOLO, the SSD-streamlined MobileNetV2 model has advantages over Mean Average Precision (mAP). SSD has the following main characteristics: (I) It inherits the idea of transforming detection into expression from YOLO to complete target positioning and classification at one time. (II) A similar Prior box is proposed based on Anchor in Faster RCNN. (III) The detection method is added based on the Pyramid Feature Hierarchy. That is, the target is predicted on the feature map of different receptive fields.

Next, the Chinese City Parking Dataset (CCPD)-Challenge dataset is selected to train the proposed SSD-streamlined MobileNetV2 model. The dataset contains 50,003 license plate images [31]. The experiment randomly selects images to speed up the learning process. Suppose that the license plate is ready and the license plate position matches the original image. In that case, the image in the data set is processed into 50,003 pieces of 960×640 RGB images by resizing and clipping. Table 4 tabulates the overall situation of the proposed FR and ALPR systems.

As shown in Table 4, the FR and ALPR frameworks need to use different datasets for training due to technical differences. At the same time, there are significant differences in image proportion and the DL framework. Therefore, the innovation of this paper is to design an improved MobileNetV2-based IR system by integrating FR and ALPR frameworks for common use. The face images in the LFW dataset are all from real-life scenes, which increases the difficulty of recognizing them. Mainly, due to the influence of multiple poses, illumination, expression, age, occlusion, and other factors, even the images

TABLE 4. Comparison of FR and ALPR characteristics.

System	Data set	Dataset characteristics	Batch size	DL framework	Learning rate	Optimizer method
FR	LFW dataset	13,233 images	250×250 face images	MxNet DL framework	0.005	AdaGrad
ALPR	Challenge dataset	50,003 license plates	960×640 RGB images	LeNet structure of Streamlined MobileNetV2	0.003	AdaGrad
EasyPR		3,000 license plates	136×36 RGB images	-	0.007	-

of the same person are very different. Moreover, some images contain more than one face. In such cases, only the face in the central coordinate is selected as the recognition target. In contrast, the faces in other areas are background interference. The CCPD-Challenge dataset contains 50,003 license plate images with different illumination and environmental conditions, which provides essential technical support for the training of the proposed model. At the same time, compared with the traditional ALPR frameworks, the proposed LeNet-based ALPR framework increases the illumination and angle analysis. As such, it comprehensively reduces the impact of illumination on image recognition and expands the influencing factors. The proposed LeNet-based ALPR framework can be installed at any angle through analysis, which improves its practicability [32].

C. PARAMETERS SETTING

In the context of DL, parameter settings have a crucial impact on algorithm performance. Convolutional hardware acceleration mainly focuses on multiplication and addition operations. Convolution hardware acceleration focuses on multiplication and addition operations. Overall, there are three main constraints on acceleration: computational efficiency, technology, and internal storage. The convolution hardware can be accelerated from loop expansion, parallel computing, and data reuse [33].

The hardware architecture of the CNN acceleration structure based on the Field Programmable Gate Array (FPGA) is as follows. AXI4 bus is responsible for connection and communication between peripheral modules; AXI4 bus is responsible for data transmission control, and AXI Lite bus is responsible for signal transmission. The external control unit is usually the Central Processing Unit (CPU), which controls the FPGA to sequentially convolute by accelerating the internal registers' reading and writing operations. The hardware structure of the proposed convolution acceleration scheme is composed of five parts: register, data interface, convolution area, merging area, and control area, as displayed in Fig. 8.

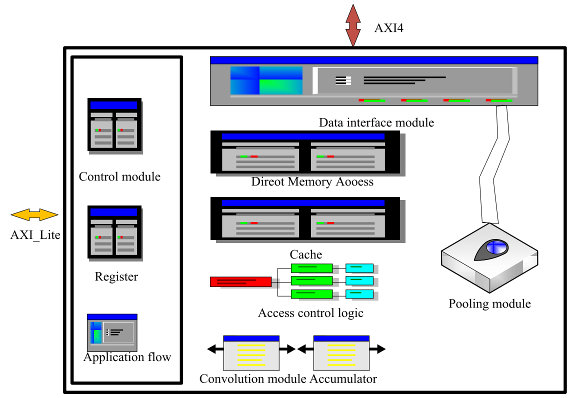

FIGURE 8. Hardware architecture of accelerator.

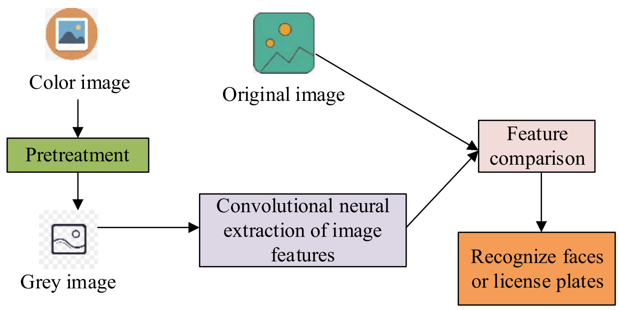

FIGURE 9. FR and ALPR process.

Subsequently, the experiment verifies the proposed convolution acceleration scheme on the LeNet structure. LeNet can have a convolution kernel of 3×3, 5×5, or 11×11 to test the trial conditions of the acceleration scheme for other more complex CNNs. According to the proposed hardware acceleration structure, the quantized 8-bit feature map and weight input will generate the 8-bit output feature map. Then, the hardware acceleration results are compared with the software calculation to verify the logical correctness of the acceleration structure. Fig. 9 signifies the flow of FR and ALPR.

As displayed in Fig. 9, the FR and ALPR process first collects the color images of the face and license plates, preprocesses and converts the color images into gray images, and then extracts the image features through CNN. Finally, it compares the image features in the image database with the collected features and recognizes the face and license plates. Eq. (10) calculates the detection of IR signals (human faces or license plates):

$$\mathrm{Verif}\left(f_i, f_j, y_{ij}, \theta_{ve}\right) = \begin{cases} \frac{1}{2}\left\|f_i - f_j\right\|_2^2 & \text{if } y_{ij} = 1 \\ \frac{1}{2} max\left(0, m - \left\|f_i - f_j\right\|_2\right)^2 & \text{if } y_{ij} = -1 \end{cases} \tag{10}$$

In Eq. (10), m represents the feature range of the image. In Eq. (10), f_i and f_j represent the image images. If the image is recognized correctly, y_{ij} is 1. If the image recognition is

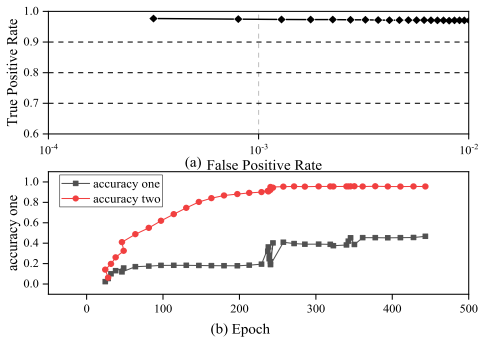

FIGURE 10. (a) ROC curve of FR effect; (b) variation of recognition accuracy with iterations.

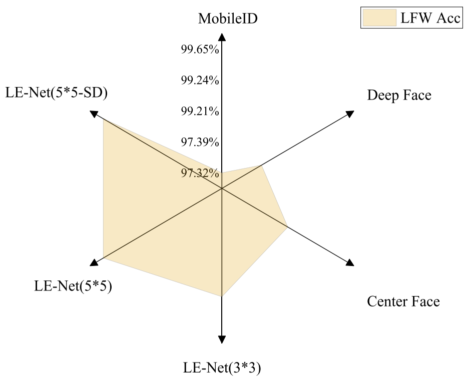

FIGURE 11. Accuracy comparison of different FR models on LFW test data set.

FIGURE 12. License plate images collected for training.

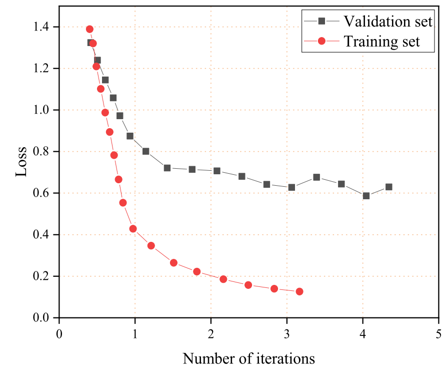

FIGURE 13. Schematic diagram of ALPR experimental training curve.

incorrect, y_{ij} is -1. Then, Eq. (11) can calculate the IR loss:

$$L = \sum_{i}^{N} \left[\left\| f\left(x_i^a\right) - f\left(x_i^p\right) \right\|_2^2 - \left\| f\left(x_i^a\right) - f\left(x_i^n\right) \right\|_2^2 + \alpha \right]_+ \tag{11}$$

In Eq. (11), $\boldsymbol{x_i^a}$, $\boldsymbol{x_i^p}$, and $\boldsymbol{x_i^n}$ represent the original features, feature correction values, and feature changes of the image. Eq. (11) calculates the difference between the specific deviation of the image within a particular space and the images, and the loss function can correct the IR result. Here, the loss function is used in both FR and ALPR.

D. EXPERIMENTAL RESULTS AND ANALYSIS

1) Experimental results of FR and its application in SLCTs

Fig. 10 compares the ROC curve between the MxNet-based FR framework and other classical FR models and the variation curve of FR accuracy.

According to the ROC curve in Fig. 10, the proposed MxNet-based FR framework performs excellently on the LFW test dataset. Further, to highlight the advantages of the proposed MxNet-based FR framework, Fig. 11 compares its recognition accuracy with other classical FR models on the LFW dataset.

The accuracy of the proposed MxNet-based FR framework on the LFW test set is more than 99%. Thus, the testing effect is relatively good. Then, CelebFaces and CASIA-WebfaceFace datasets are selected to test the progress of the proposed MxNet-based FR framework, resulting in a test accuracy of 99.46% and 99.23%, respectively.

The outstanding performance of the MxNet-based FR framework mentioned above in FR has brought positive impacts to the development of SLCTs. The experimental results demonstrate the high accuracy and stability of this technology on multiple datasets, providing effective tools for town managers to achieve more intelligent management and high-quality services. By applying FR technology to areas such as entrance control, crowd monitoring, and user experience improvement in small towns, the town can achieve more efficient resource utilization and better community management, thus promoting the inheritance and innovative development of the town.

2) The impact of experimental results on ALPR on the development of SLCTs

According to the experimental design description above, Fig. 12 illustrates the collected sample images of the license plate for training.

This section selects 50,003 images in the CCPD-Challenge dataset as training samples, 20,000 images as test samples, and 7,000 images as a verification set for model evaluation. Fig 13 signifies the training curve.

After several rounds of operation, the loss of the verification set tends to stabilize. Then, the experiment obtains the overall accuracy by comprehensively evaluating the ALPR process.

The results on the verification set prove that the recall rate negatively correlates with the set Intersection Over Union

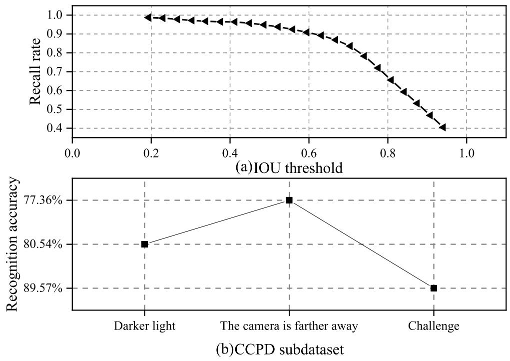

FIGURE 14. Intersection parallel ratio-recall rate curve (a) and robustness analysis accuracy diagram (b).

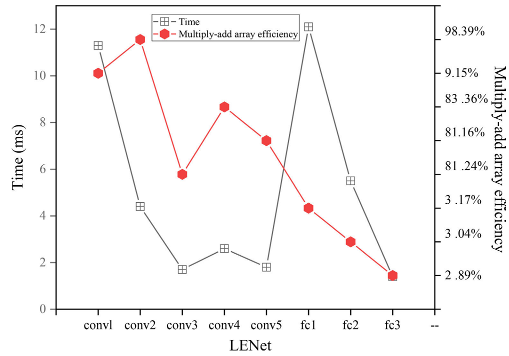

FIGURE 15. Schematic diagram of LeNet hardware acceleration results.

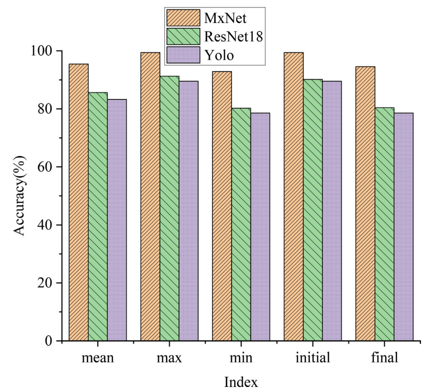

FIGURE 16. Comparison between the proposed MxNet model and other models.

(IOU) threshold. The CCPD-Challenge dataset trains the model. The model evaluation chooses the CCPD-DB (dark light) and CCPD-FN (far from the camera) datasets. Fig. 14 shows the statistical recall rate curve and robustness accuracy evaluation analysis.

Fig. 14 counts the character recognition accuracy of the model in the verification set. The recognition is deemed accurate only when all the seven characters (an ordinary blue license plate has seven characters) are output correctly. Then, it counts the recognition results when the IOU threshold is 0.6 after all license plates have been detected. Finally, the statistical results prove that the TPR of ALPR in 20,000 images reaches 89.57%.

The application of ALPR technology in parking management in the experiment has greatly improved the efficiency of parking management in small towns. Accurately identifying vehicles and automatically recording their entry and exit effectively reduces congestion during the parking process. This not only makes the parking process more convenient, but also provides residents and tourists with a better parking experience. At the same time, ALPR technology can also be used for functions such as electronic payment and booking parking spaces, further promoting the intelligent level of parking management.

The experimental results of ALPR technology reveal that it also has a significant impact on traffic flow control and urban planning. By accurately recording the driving path and dwell time of vehicles, town managers can better understand the situation of traffic congestion and conduct intelligent traffic scheduling based on data. This helps to optimize road use, improve traffic efficiency, and create a smoother travel environment for residents and tourists.

The application of ALPR technology in small towns also provides opportunities for intelligent innovation. By integrating with other intelligent facilities and systems, such as intelligent parking systems and urban monitoring systems, small towns can achieve higher levels of intelligent management and services. This intelligent innovation not only enhances the technological image of the town, but also provides strong support for its future sustainable development.

3) Experimental results of DL algorithm acceleration module and its significance in the construction of intelligent towns

Fig. 15 plots hardware acceleration results on LeNet.

Fig. 15 indicates that the proposed hardware acceleration scheme has high acceleration efficiency except for the first convolutional layer. The proposed acceleration scheme can be generalized to the proposed lightweight MobileNetV2 model. Because input channels are 3, the first convolutional layer runs in an arithmetic unit with an input channel parallelism of 32. The multiplication and addition matrix is inefficient and takes a long time. The data transmission volume of the fully connected layer far exceeds the calculation scale. In the state of multiplexing weights, the speed of reading weights restricts the operation efficiency, reducing the calculation efficiency in the acceleration structure. Thanks to the convolution acceleration, the efficiency of matrix multiplication and the addition of other convolutional layers exceeds 80%. The difference is mainly due to the different output efficiency of the feature maps of different convolutional layers classified by the output module.

Fig. 16 compares the results between the proposed MxNet model and other DL models, in which 1,000 iterative training tests are used.

According to Fig. 16, the proposed MxNet model is tested and compared with ResNet18 and YOLO. Here, the mean represents the average value. The max is the maximum value;

the min denotes the minimum value; the initial is the initial value; the final is the final output. Apparently, the accuracy of the proposed MxNet model is higher than the other two models. The highest and lowest accuracy of the proposed MxNet model reaches about 99% and 95%. By comparison, the highest and lowest accuracy of ResNet18 is 92% and 84%. Yolo model reaches about 89% at best and about 79% at worst. Thus, the proposed MxNet model is better than other models in vehicle recognition.

The construction of intelligent towns needs to support various intelligent applications, such as intelligent traffic management, environmental monitoring, safety monitoring, etc. These applications typically require many computing resources to process and analyze data. In this case, the hardware acceleration scheme in this work can play an important role. By providing faster and more efficient computing resources, the implementation of intelligent systems can be supported, thereby improving small towns' operational efficiency and quality of life.

E. DISCUSSION

This work is oriented toward the social level of the new era and studies the inheritance and innovative development of SLCT based on DL technology and AI technology. It is found that, firstly, the FR framework based on the proposed MxNet performs well on the LFW test data set, reaching an accuracy of over 99%. At the same time, CelebFaces and CASIA-Web face datasets are selected to test the progress of the proposed FR framework based on MxNet, and the test accuracy reaches 99.46% and 99.23%, respectively. Then, 50,003 images in the CCPD-Challenge dataset are selected as training set, 20,000 images are test set, and 7,000 as verification set. After several rounds of operation, the loss of the verification set tends to be stable. Then, the experiment obtains the overall accuracy by comprehensively evaluating the ALPR process. The results on the validation set prove that the recall rate is negatively correlated with the Intersection Over Union (IOU) threshold. The model is trained on the CCPD-Challenge dataset. CCPD-DB (dark light) and CCPD-FN (away from the camera) datasets are selected for model validation. The results show that the recognition is accurate only when all seven characters (seven characters for an ordinary blue license plate) are correctly output. After all license plates are detected, the recognition result is calculated when the IOU threshold is 0.6. Finally, the statistical results find that the TPR of the ALPR framework reaches 89.57% in 20,000 images. Lastly, the proposed hardware acceleration scheme can be extended to the lightweight MobileNetV2 model. Because the input channel is 3, the first convolution layer runs in an arithmetic unit with an input channel parallelism of 32. Multiplication and addition matrices are inefficient and time-consuming. The data transmission volume of the fully connected layer far exceeds the computing scale. Thanks to convolution acceleration, the matrix multiplication and addition efficiency of other convolution layers reach over 80%. The difference is mainly due to the different output efficiency of the feature map of different convolution layers classified by the output module. The comparison results with other DL models show that the accuracy of the proposed MxNet model is higher than the other two models. Of these, the highest and lowest accuracy of the proposed MxNet model is 99% and 95%. The highest and lowest accuracy of the ResNet18 model is about 92% and 84%. By comparison, the highest and lowest accuracy of the YOLO model is about 89% and 79%. Therefore, the proposed MxNet model is better than other models in vehicle recognition. Compared with the study of Rademeyer et al. (2020). The model proposed in this work can better reflect the technology of ALPR and provide a more accurate and valuable solution for SLCT construction [34]. Compared with the study of Hassan et al. (2022), the model proposed in this work integrates AI and DL more deeply in image analysis, and the design of the image analysis algorithm is more comprehensive [35].

V. CONCLUSION

With the rapid development of AI and robot technology, the intelligent construction of cities and towns is imperative. Based on this and the development idea of SLCTs in the new era, this work starts with cutting-edge technology: AI and robot technologies. It designs a lightweight IR system using an optimized DL algorithm to complete FR and ALPR tasks in the context of the SLCT construction. Simultaneously, a hardware acceleration scheme is designed for the proposed IR system to accelerate the image processing algorithm. The experimental results show that the proposed IR system can achieve lightweight and efficient FR and ALPR. The designed image processing hardware acceleration also achieved good results. It is concluded that the lightweight network can also achieve excellent accuracy and generalization. The designed hardware acceleration model is viable in the operation acceleration of the CNN. The research is based on integrating AI and DL algorithms. The purpose is to build the technical foundation of characteristic services of SLCTs, inherit towns' culture and tourism economy, and innovate towns' digital construction. Because of the depth of AI technology and the complexity of DL algorithms, the research process also faces limitations in the algorithm display and classification, the model establishment and training, and the dataset selection. Not all scenarios are considered. Therefore, it is hoped that more researchers will participate in the follow-up research on the inheritance and innovative development of intelligent SLCTs. It is expected to contribute to the SLCT construction of the new era.

A. RESEARCH CONTRIBUTION

The contribution of this work lies in:

1) It provides technical support for AI and robot technology innovation.

2) It provides a reference for developing AI and robot technology in future urban construction.

3) It has contributed to the inheritance and innovative development of SLCTs in the future.

B. FUTURE WORKS AND RESEARCH LIMITATIONS

The future research direction will focus on the improvement of the modernization of urban construction. Also, AI and robot technologies will be strengthened and innovated. It is hoped to adapt to the times and provide the impetus for social development.

The specific limitations of this work are:

1) This work only involves the innovation and verification of technology. The effect in practical application is not reflected;

2) The optimization effect of the integration of AI technology and robot technology is not deep enough;

3) AI and robot technologies are developing very rapidly in the future. It is necessary to continue updating the model's design concept to adapt to social and technological progress.

REFERENCES

[1] L. Y. Zhou and X. Y. Song, "Coordinated development mechanism of sports characteristic town construction and intangible cultural heritage protection," *J. Guangxi Normal Univ. Nationalities.*, vol. 37, no. 1, pp. 19–22, 2020.

[2] S. Mao and J. F. Zhu, "Analysis of sports industry agglomeration points in Jiangsu Province," *Liaoning Sports Sci. Technol.*, vol. 42, no. 3, pp. 11–17, 2020.

[3] J. Zhang and L. P. Zhong, "Research on the cultivation of sports characteristic towns under the background of rural revitalization strategy," *Hubei Sports Sci. Technol.*, vol. 39, no. 4, pp. 293–296, 2020.

[4] M. Hu, Y. Zhong, S. Xie, H. Lv, and Z. Lv, "Fuzzy system based medical image processing for brain disease prediction," *Frontiers Neurosci.*, vol. 15, Jul. 2021, Art. no. 714318.

[5] S. Xie, Z. Yu, and Z. Lv, "Multi-disease prediction based on deep learning: A survey," *Comput. Model. Eng. Sci.*, vol. 127, no. 3, pp. 1–34, 2021.

[6] B. Uttaro, S. Zawadski, I. Larsen, and M. Juárez, "An image analysis approach to identification and measurement of marbling in the intact pork loin," *Meat Sci.*, vol. 179, Sep. 2021, Art. no. 108549.

[7] A. Kar and P. Neogi, "Triangular coil pattern of the local radius of gyration face for heterogeneous face recognition," *Appl. Intell.*, vol. 50, no. 3, pp. 698–716, 2020.

[8] Z. Lv, D. Chen, and J. Li, "Novel system design and implementation for the smart city vertical market," *IEEE Commun. Mag.*, vol. 59, no. 4, pp. 126–131, Apr. 2021.

[9] Z. Yu, S. U. Amin, M. Alhussein, and Z. Lv, "Research on disease prediction based on improved DeepFM and IoMT," *IEEE Access*, vol. 9, pp. 39043–39054, 2021.

[10] M. Al-Antari, C. H. Hua, and J. Bang, "Fast deep learning computer-aided diagnosis of COVID-19 based on digital chest X-ray images," *Appl. Intell.*, vol. 51, no. 5, pp. 2890–2907, 2021.

[11] Q. Li, "Sports + tourism characteristic sports town construction Hunan Province," *Sports Sci. Technol. Literature Bulletin.*, vol. 28, no. 9, pp. 141–142, 2020.

[12] X. Wang and P. L. Cao, "Research on the integrated development of Hainan's tourism industry and sports towns," *Contemp. Tourism*, vol. 18, no. 10, pp. 67–68, 2020.

[13] J. L. Gao and R. S. Shi, "SWOT analysis of the development of leisure fishing industry in Wuhan from the perspective of sports towns. Journal of the Hubei university of economics," *Humanities Social Sci. Ed.*, vol. 17, no. 11, pp. 37–39, 2020.

[14] H. Caferoglu, D. Elsner, and M. G. Moehrle, "The interplay between technology and pre-industry convergence: An analysis in the technology field of smart mobility," *IEEE Trans. Eng. Manag.*, vol. 70, no. 4, pp. 1504–1517, Apr. 2023.

[15] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 18–44, Mar. 2021.

[16] M. X. Chen, X. R. Huang, and G. Z. Huang, "New-type urbanization and informal employment: Scale, pattern and social integration," *Adv. Geographical Sci.*, vol. 40, no. 1, pp. 50–60, 2021.

[17] D. Zeltner, B. Schmid, G. Csiszár, and O. Csiszár, "Squashing activation functions in benchmark tests: Towards a more eXplainable artificial intelligence using continuous-valued logic," *Knowl.-Based Syst.*, vol. 218, Apr. 2021, Art. no. 106779.

[18] J. Borenstein, F. S. Grodzinsky, A. Howard, K. W. Miller, and M. J. Wolf, "AI ethics: A long history and a recent burst of attention," *Computer*, vol. 54, no. 1, pp. 96–102, Jan. 2021.

[19] M. Mckendrick, S. Yang, and G. A. Mcleod, "The use of artificial intelligence and robotics in regional anesthesia," *Anaesthesia*, vol. 76, no. 1, pp. 171–181, 2021.

[20] J. W. Fan, R. Zhang, and M. Lu, "Application of deep learning methods in the diagnosis of diabetic retinopathy," *Acta Autom. Sinica*, vol. 47, no. 5, pp. 985–1004, 2021.

[21] K. Lin, Y. Zhao, L. Tian, C. Zhao, M. Zhang, and T. Zhou, "Estimation of municipal solid waste amount based on one-dimension convolutional neural network and long short-term memory with attention mechanism model: A case study of Shanghai," *Sci. Total Environ.*, vol. 791, Oct. 2021, Art. no. 148088.

[22] C. Tohill, L. Ferreira, and C. J. Conselice, "Quantifying non-parametric structure of high-redshift galaxies with deep learning," *Astrophysical J.*, vol. 916, no. 1, pp. 4–18, 2021.

[23] R. Chen, S. Zhao, W. Wu, Z. Sun, J. Wang, H. Wang, and G. Han, "A convolutional neural network algorithm for breast tumor detection with magnetic detection electrical impedance tomography," *Rev. Sci. Instrum.*, vol. 92, no. 6, Jun. 2021, Art. no. 064701.

[24] Y. Jin, X. Jiang, Z. Wei, and Y. Li, "Chest X-ray image denoising method based on deep convolution neural network," *IET Image Process.*, vol. 13, no. 11, pp. 1970–1978, Sep. 2019.

[25] D. Chen, P. Wawrzynski, and Z. Lv, "Cyber security in smart cities: A review of deep learning-based applications and case studies," *Sustain. Cities Soc.*, vol. 66, Mar. 2021, Art. no. 102655.

[26] P. N. Srinivasu, J. G. Sivasai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of skin disease using deep learning neural networks with MobileNet v2 and LSTM," *Sensors*, vol. 21, no. 8, p. 2852, Apr. 2021.

[27] Z. Lv, D. Chen, R. Lou, and Q. Wang, "Intelligent edge computing based on machine learning for smart city," *Future Gener. Comput. Syst.*, vol. 115, pp. 90–99, Feb. 2021.

[28] H. Karimi-Rouzbahani, F. Ramezani, A. Woolgar, A. Rich, and M. Ghodrati, "Perceptual difficulty modulates the direction of information flow in familiar face recognition," *NeuroImage*, vol. 233, Jun. 2021, Art. no. 117896.

[29] D. Sáez Trigueros, L. Meng, and M. Hartnett, "Generating photo-realistic training data to improve face recognition accuracy," *Neural Netw.*, vol. 134, pp. 86–94, Feb. 2021.

[30] Z. Chen, L. R. Li, and Z. Li, "Research on license plate recognition technology based on machine learning," *Comput. Technol. Develop.*, vol. 30, no. 6, pp. 19–24, 2020.

[31] Z. Cheng, X. Zhu, and S. Gong, "Face re-identification challenge: Are face recognition models good enough?" *Pattern Recognit.*, vol. 107, Nov. 2020, Art. no. 107422.

[32] S. Wu, W. Zhai, and Y. Cao, "PixTextGAN: Structure aware text image synthesis for license plate recognition," *IET Image Process.*, vol. 13, no. 14, pp. 2744–2752, Dec. 2019.

[33] J. Li, S. Jiang, S. Gong, J. Wu, J. Yan, G. Yan, and X. Li, "SqueezeFlow: A sparse CNN accelerator exploiting concise convolution rules," *IEEE Trans. Comput.*, vol. 68, no. 11, pp. 1663–1677, Nov. 2019.

[34] M. C. Rademeyer, A. Barnard, and M. J. Booysen, "Optoelectronic and environmental factors affecting the accuracy of crowd-sourced vehicle-mounted license plate recognition," *IEEE Open J. Intell. Transp. Syst.*, vol. 1, pp. 15–28, 2020.

[35] M. R. Hassan, M. F. Islam, M. Z. Uddin, G. Ghoshal, M. M. Hassan, S. Huda, and G. Fortino, "Prostate cancer classification from ultrasound and MRI images using deep learning based explainable artificial intelligence," *Future Gener. Comput. Syst.*, vol. 127, pp. 462–472, Feb. 2022.