

A Deep Learning-based Intelligent Medicine Recognition System for Chronic Patients

Wan-Jung Chang, *Member, IEEE*, Liang-Bi Chen, *Senior Member, IEEE*,
Chia-Hao Hsu, Cheng-Pei Lin, and Tzu-Chin Yang

Abstract—This paper proposes an intelligent medicine recognition system based on deep learning techniques, named ST-Med-Box. The proposed system can assist chronic patients in taking multiple medications correctly and in avoiding taking the wrong medications, which may cause drug interactions, and can provide other medication-related functionalities such as reminders to take medications on time, medication information, and chronic patient information management. The proposed system consists of an intelligent medicine recognition device, an app running on an Android-based mobile device, a deep learning training server, and a cloud-based management platform. Currently, eight different medicines can be recognized by the proposed system. The experimental results show that the recognition accuracy reaches 96.6%. Therefore, the proposed system can effectively reduce the problem of drug interactions caused by taking incorrect drugs, thereby reducing the cost of medical treatment and giving patients with chronic diseases a safe medication environment.

Index Terms—artificial intelligence over the Internet of Things (AIoT), chronic diseases, deep learning, Internet of Things (IoT), medicine recognition.

I. INTRODUCTION

CURRENTLY, the world's society is aging. Among the 7.5 billion people in the world, the elderly population accounts for 600 million, including 480 million people with chronic diseases. According to statistics from the World Health

Organization (WHO), the average elderly person suffers from 1.4 chronic diseases, and the typical medication dosage of an elderly person is five times that of a younger person.

Elderly people are also seven times more likely to take the wrong medicine because of declining physiological functions. The WHO also indicates that one-third of the world's deaths are caused not by diseases themselves but by the incorrect use of drugs, and the costs associated with such improper drug use amount to nearly 28.5 billion U.S. dollars every year.

Due to the abovementioned problem of deaths caused by the improper use of drugs, the smart medicine pillboxes available on the market are constantly being updated. For example, the PillDrill [1], [2], a smart medicine pillbox, can help users conveniently store and distribute medication and has a reminder function to remind users who forget to take their medicine.

However, a patient cannot know whether the medications he or she has taken are correct. Due to the wide variety of drugs used for patients with chronic diseases, their classification is obviously a complex task, and misidentification of medications caused by negligence may lead to the possibility of taking the wrong medicine.

Taking the wrong medicine may result in harmful interactions or offset the intended effects of the drugs, leading to further serious consequences such as acute complications. To address this problem, this paper proposes a deep-learning-based intelligent medicine recognition system, named ST-Med-Box, that can recognize medications and remind patients with chronic diseases when to take their medications.

By using the proposed system, patients with chronic diseases can know whether a drug is taken correctly the first time, thus reducing the probability of taking the wrong medicine and the cost of social medical care.

The remainder of this paper is organized as follows. Section II reviews related works. The proposed system is described in Section III. Section IV introduces deep learning preliminaries and techniques that are applied in the proposed system. A prototype of the proposed system and its related experiments are demonstrated and described in Section V. Finally, Section VI concludes this work.

II. RELATED WORKS

In the medical field, adverse events (AEs) are the leading cause of morbidity and mortality [3], according to data from the

Manuscript received March 9, 2019; accepted March 25, 2019.

This work was supported in part by the Ministry of Science and Technology (MoST), Taiwan, R.O.C., under Grants MOST 107-2637-E-218-003, MOST 107-2632-E-218-001, and MOST 108-2622-8-218-004-TE2. This work was also supported in part by the Ministry of Education (MoE), Taiwan, R.O.C., under Grant 1300-107P735, through the Advanced Intelligent Biomedical Allied Research Center under the Higher Education Sprout Project. (Corresponding author: Wan-Jung Chang.)

W.-J. Chang is with the Department of Electronic Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan, R.O.C. (e-mail: allenchang@stust.edu.tw)

L.-B. Chen is with the Department of Electronic Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan, R.O.C. (e-mail: liangbi.chen@gmail.com)

C.-H. Hsu is with the Department of Electronic Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan, R.O.C. (e-mail: 4a337906@stust.edu.tw)

C.-P. Lin is with the Department of Electronic Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan, R.O.C. (e-mail: 4a40h020@stust.edu.tw)

T.-C. Yang is with the Department of Electronic Engineering, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan, R.O.C. (e-mail: 4a43a045@stust.edu.tw)

U.S. Food and Drug Administration (FDA). Since 1997, the number of all AE cases reported by the FDA has increased by nearly a factor of five, and between 2006 and 2014, the number of deaths due to AE cases increased by 232%. To address such problems, many pharmaceutical companies are trying to deal with the increased number of cases by means of manual case records, but this is not a sustainable solution because of the considerable growth in the incidence of AE cases [3].

To address AE-related problems, Dev *et al.* [4] identified an increasing number of adverse drugs and new drugs. They used deep learning based on a word2vec module for AE case classification. The results showed increased robustness to unseen AE cases and medical words, but due to FDA restrictions, this work lacked interpretability, which led to difficulties in application.

With the rapid development of contemporary medicine, one of the key AE-related problems is adverse drug events. To avoid adverse drug events, a high-accuracy automatic medicine recognition system is needed to assist people in recognizing various kinds of drugs. To meet this requirement, many related drug recognition technologies and systems have been addressed, discussed, and developed.

To develop pill image recognition techniques, the U.S. National Library of Medicine (NLM) held a challenge competition in 2016. Yaniv *et al.* [5] reported the results of this challenge competition. Ushizima *et al.* [6] investigated pill recognition approaches on a new pill image dataset for the NLM competition. They also discussed some data planning strategies for effective content-based image retrieval.

Ribeiro *et al.* [7] proposed a medicine box recognition system that adopted a three-stage (barcode recognition, text recognition, and feature matching) approach. Their proposed system used a camera mounted on a device and used an Android system to correctly recognize medicine packages to provide people experiencing difficulties (such as elderly individuals and individuals with visual impairments) with related medication packaging information. Bar code detection and optical character recognition (OCR) were used to recognize the names on the medicine packages. Their proposed system achieved a recognition success of up to 80%. However, the system had a recognition blind spot in the case of medicine packages bearing the same name but with different contents (number of tablets, dosage, and/or route of administration). Wang *et al.* [8] also presented a deep-learning-based recognition methodology for recognizing drug blister packages.

Yu *et al.* [9] presented an accurate and automatic pill recognition system that combined imprint extraction and description to make use of imprint information. Moreover, a loopy-belief-propagation-based image segmentation approach was applied to the imprint on the pill to solve the problem of incoherent and coarse strokes. The experimental results showed that this pill recognition system achieved accuracies of 90.46% on the top rank and 97.16% on the top five ranks.

Neto *et al.* [10] proposed a pill feature extractor based on shape and color, called CoforDes. In this work, K-nearest neighbor (K-NN)-, support vector machine (SVM)-, and Bayes-based techniques were adopted to evaluate the extracted

features. Moreover, CoforDes can be executed as a part of real-time embedded applications.

Calix *et al.* [11] presented a method for medical personnel to use in tracking and monitoring the safe use of drugs, called the Deep Gramulator, which is designed to automatically extract tweets related to personal health experience from social media. In this work, classifiers were built based on a variety of machine learning algorithms, such as deep neural networks (DNNs), to help detect such data. However, when the Deep Gramulator was applied to an independent test set of 3,156 samples, the classifier did not perform well in terms of accuracy.

Another important issue is medication adherence. Poor medication adherence can threaten people's health. Hence, many medication adherence monitoring and tracking systems have been proposed [12]-[15].

Kalantarian *et al.* [12], [13] proposed a smartwatch-based system in which built-in 3-axial accelerometers and gyroscopes were adopted to recognize several motions for medication adherence detection. Ma *et al.* [14] also developed a smartwatch-based adherence monitoring system, which used machine learning and distributed computing techniques. Moreover, data were collected from built-in inertial sensors to monitor the sequence of actions occurring during a person's medication intake.

Aldeer *et al.* [15] designed a noninvasive medication adherence monitoring system based on integrated collaborative sensing technology in a wireless sensor network (WSN) environment. Several sensors were mounted on the pill bottle to monitor medication adherence behavior regarding pill intake.

Other related issues have also been addressed [16]-[31] in attempts to provide immediate medication records, cloud-based online medicine management systems, and personalized telemedicine services. Given several instances, Chen *et al.* [16] proposed a comprehensive medicine management system that integrated medical information from various sources. The proposed system could automatically detect inappropriate drugs. Moreover, every participant could fully track the patients' most recent medicine use online in real time.

Kim [17] developed a medication compliance monitoring system for checking medication compliance in clinical trials. The proposed system was composed of a clinical trial database management platform and a PDA with a barcode scanner for each clinical trial participant. Related healthcare ecosystems have also been addressed, discussed and proposed [19]-[31] to enable new personalized, predictive, preemptive, pervasive, and precise e-health services.

Jiménez-Fernández *et al.* [32] discussed possible issues of adaptation usability and sensor device interoperability for chronic disease management systems. They noted that a large percentage of chronic patients are also elderly patients. Their experimental results showed that sensor-network-based systems could play an important role in monitoring patients. Hence, the usability of such a sensor-network-based system must be adapted to meet the patients' needs.

Although many solutions have been published for adverse drug events and systems for providing immediate medication

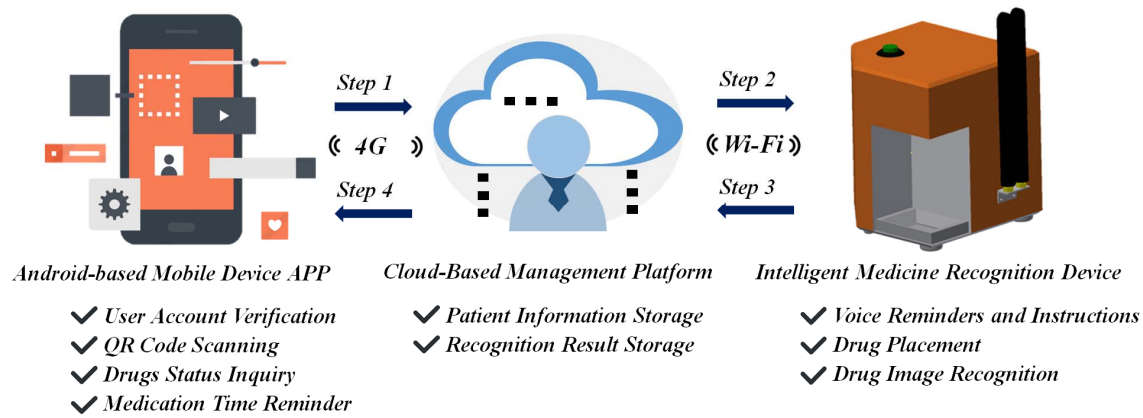


Fig. 1. Overview of the proposed system.

records and personalized services have been developed for patients, their functions are often not comprehensive or easy to use. Therefore, in this paper, we develop an intelligent medicine recognition system based on deep learning technology. The proposed system has many benefits, such as its capabilities of perfect recording and high-precision recognition, its simple operation, and its friendly interface. Compared to other machine learning approaches, deep learning is the most effective recognition approach, achieving high recognition rates [33]. The proposed system incorporates edge computing [34] capabilities to reduce delay and to make the connected application more sensitive, thus enabling real-time computing capabilities and the performance optimization of applications based on artificial intelligence over the Internet of Things (AIoT).

III. THE PROPOSED SYSTEM

A. Design Concept of the Proposed System

To address the problems posed by patients with chronic diseases taking multiple medications for those diseases, we propose an intelligent medicine recognition system based on deep learning technology. This system can automatically identify pills and assist patients with chronic diseases in understanding the dosage of their medications and other related information, thus mitigating the problem of patients taking the wrong medications.

Currently, smart medicine pillboxes are often used to organize drugs and help patients take medicine and combine a variety of drugs, but they ignore the potential adverse effects of placing different drugs together. To address this problem, we use deep-learning-based image recognition technology to achieve immediate multiple drug placement and instant recognition and to provide voice explanations of medication information.

Fig. 1 shows an overview of the proposed system. The proposed system consists of an intelligent medicine recognition device, an app running on an Android-based mobile device, a deep learning training server, and a cloud-based management platform. As shown in Fig. 1, the proposed system is designed as a personalized service for patients with multiple chronic

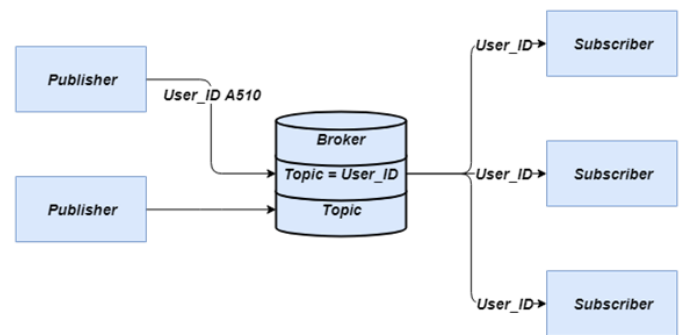


Fig. 2. The MQTT architecture.

diseases taking multiple medications. The proposed system is introduced as follows.

Step 1: First, a user logs in to his or her account through the account verification mechanism of the Android-based mobile device app. After successfully logging in, the user can click on the QR code option in the Android mobile device app to scan the QR code on a medicine package to obtain the medication information. Then, that medication information is transmitted over the 4G network to a cloud-based management platform for storage. Drug information (such as the drug name, medication time, and dosage) can be checked via a website.

Step 2: The cloud-based management platform transmits the medication information obtained from the QR code on the medicine package to the proposed intelligent medicine recognition device over a Wi-Fi network. The proposed intelligent medicine recognition device issues a voice prompt to remind the patient to take his or her medicine. Then, the patient places the medicine in the recognition region of the proposed intelligent medicine recognition device and presses the button to recognize the pills. After the recognition process is complete, the current medication status (the medication is correct, the medication is incorrect, more medicine needs to be taken, less medicine needs to be taken, or other related medication information) will be announced to the patient by voice.

Step 3: The proposed intelligent medicine recognition system transmits the recognition results back to the cloud-based management platform over the Wi-Fi network. Thus, family members or the patient can check the patient's medication

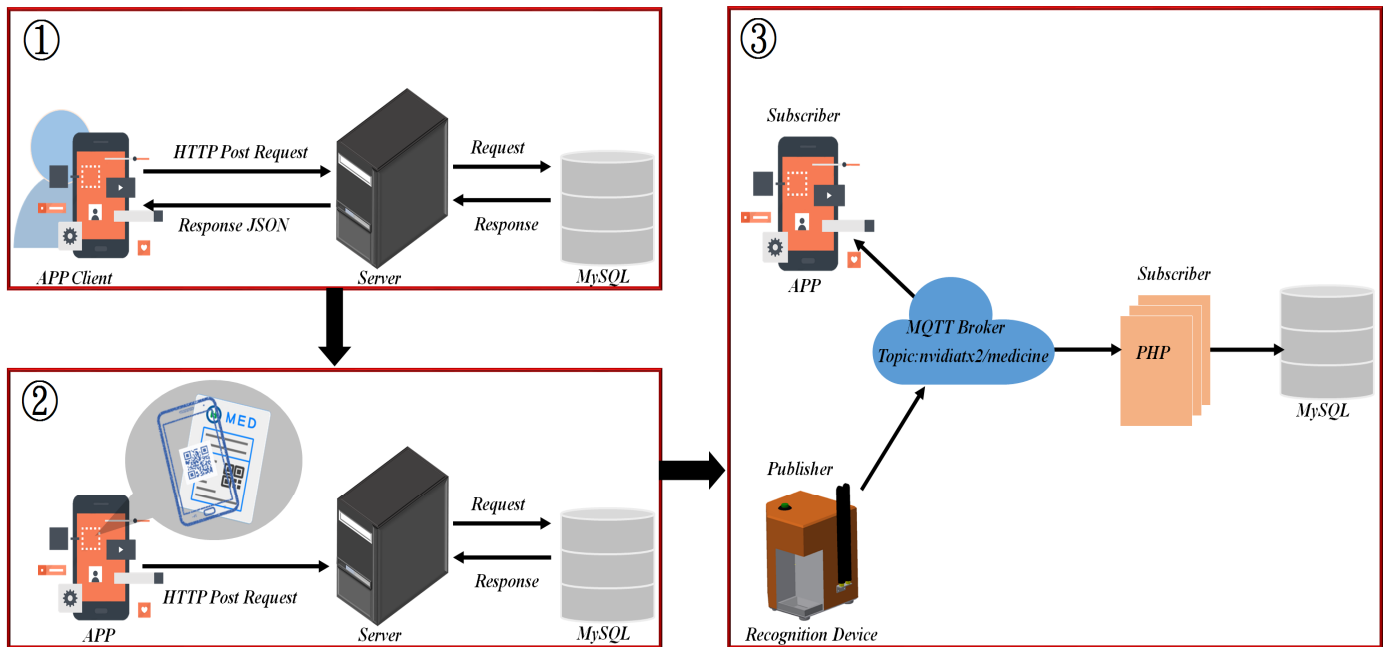


Fig. 3. Data transfer process between the components of the proposed system via the MQTT and HTTP protocols.

records (drug name, dosage, and actual medication time) through the website.

Step 4: The Android-based mobile device can receive medication records from the cloud-based management platform over the 4G network. In this way, family members or the patient can also instantly view the patient's medication records (drug name, dosage, and actual medication time) through the Android mobile device app to ensure proper management of chronic diseases.

For intercommunication among the intelligent medicine recognition device, the cloud-based management platform, and the Android-based mobile device, in addition to using the HTTP protocol, the message queuing telemetry transport (MQTT) protocol [35], [36] is also adopted. The MQTT protocol is used to quickly transfer data between the Android mobile device app and the cloud-based management platform. The MQTT protocol is a publish/subscribe-based message transmission protocol. The MQTT architecture is shown in Fig. 2.

The MQTT architecture consists of publishers, subscribers, and a broker. As shown in Fig. 2, a publisher is the source of a message. Messages are sent to particular topics. If a subscriber is subscribed to a topic, that means that he or she wants to receive messages related to that topic.

For example, if a publisher sends message "A510" to a particular topic, then all subscribers subscribed to that topic will receive this message.

Fig. 3 shows the data transfer process between the components of the proposed system via the MQTT and HTTP protocols. The server and the MySQL database are both part of the cloud-based management platform. The data transfer process is described as follows.

1) First, the Android mobile device app requests one of two things from the server through the HTTP protocol. The first is to register a user account. The server passes the password entered by the user to the MySQL database for comparison. If the information is correct, the server will return an account registration success message in the JSON format to the Android mobile device app display; otherwise, it will return an account registration failure message, also in the JSON format.

The second possible type of request is for a user to log in to his or her account. Similarly, if the login is successful, the server will return a JSON-format account login success message for display via the Android mobile device app; otherwise, a JSON-format account login failure message is returned.

2) After a successful login, the QR code option on the Android mobile device app is used to transfer the medication package information to the MySQL database via an HTTP POST request.

3) After the proposed intelligent medicine recognition device performs drug image recognition, the drug information is released to the broker via MQTT communication. The selected topic is nvidiatx2/medicine. The server is subscribed to this topic through PHP; thus, it obtains the data, which are then transmitted to the MySQL database. The Android mobile device app is also subscribed to this topic and displays the recognition result.

The proposed system is composed of an intelligent medicine recognition device, an Android-based mobile device app, and a cloud-based management platform. The designs of these constituent components are introduced in the following subsections.

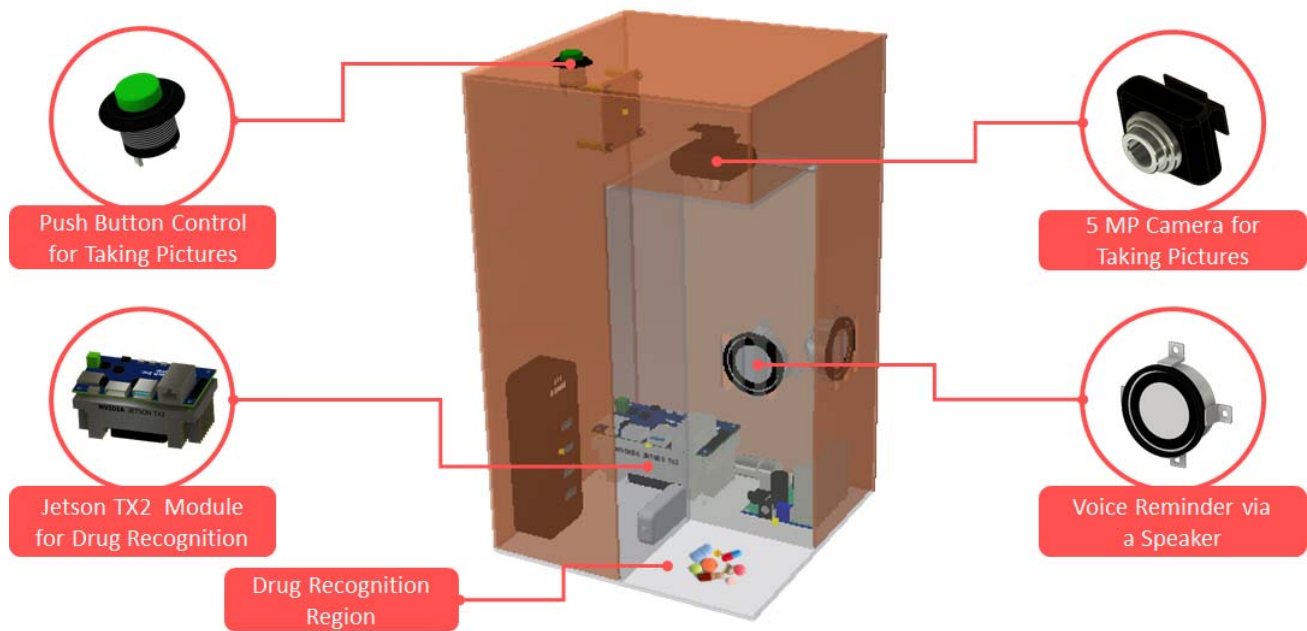


Fig. 4. Structure of the proposed intelligent medicine recognition device.

B. Intelligent Medicine Recognition Device

Chronic diseases have become the largest health concern for modern people. These chronic diseases often lead to the use of many different types of drugs, resulting in confusion and difficulty in storage. Therefore, we have designed an intelligent medicine recognition device, as shown in Fig. 4. The internal structure of the proposed device consists of a 5 MP camera, a GPU-based embedded computing system module (NVIDIA Jetson TX2 module [37]), a push button, and a 15 W amplifier-speaker unit.

The GPU-based embedded computing system module controls the speaker to send voice reminders to remind patients to take their medicine. This module also controls a 5 MP camera to capture pictures of the drugs to be recognized and instantly recognizes them. Finally, the recognized results are stored in the database of the cloud-based management platform.

The NVIDIA Jetson TX2 module has some of the top specifications among the embedded GPUs in embedded computing system modules that are currently on the market. The embedded GPU in the NVIDIA JETSON TX2 module has a Pascal architecture, which has features that greatly accelerate recognition and enable more accurate recognition based on richer DNNs, as the data structure of the human cerebral cortex becomes the basis of deep-learning-related research.

Deep learning has become a popular recognition methodology in recent years. It is based on the principle of imitating the nervous system of the human brain for training. Because it can automatically capture the characteristics of images, it is also called “feature learning”. The main advantage of deep learning, compared to other machine learning methodologies, is that its high-accuracy and scalable databases can be used in many ways, such as language learning, image

recognition, and speech recognition.

In our testing process, we first collected images of drugs taken by patients with chronic diseases, input them into the module with adjusted parameters for training on the cloud server, and then transferred the trained module to the embedded computing system module for drug image recognition. The deep learning techniques and the development framework adopted in this paper will be introduced in detail in Section IV.

C. Android-based Mobile Device App

To communicate with the proposed intelligent medicine recognition device and instantly comprehend a chronic patient's medication status, we have also developed an Android-based mobile device app, which allows chronic patients to communicate with the proposed intelligent medicine recognition device through the cloud-based management platform. Its functions include QR code scanning for connecting to the intelligent medicine recognition device, drug package information storage, medication status querying, medication time reminders, and other related functions. The steps of operation of the proposed Android-based mobile device app are described as follows.

First, the user scans the QR code on a drug package using the QR code scan option to transfer the drug package information, such as the drug name and medication time, to our cloud-based management platform for storage.

The drug package information is also important for enabling the proposed intelligent medicine recognition device to make drug comparisons and to check whether drugs should be taken by patients with chronic diseases, as shown in Fig. 5.

Next, the user scans the QR code on the proposed intelligent medicine recognition device using the QR code scan option to connect to the proposed intelligent medicine recognition device

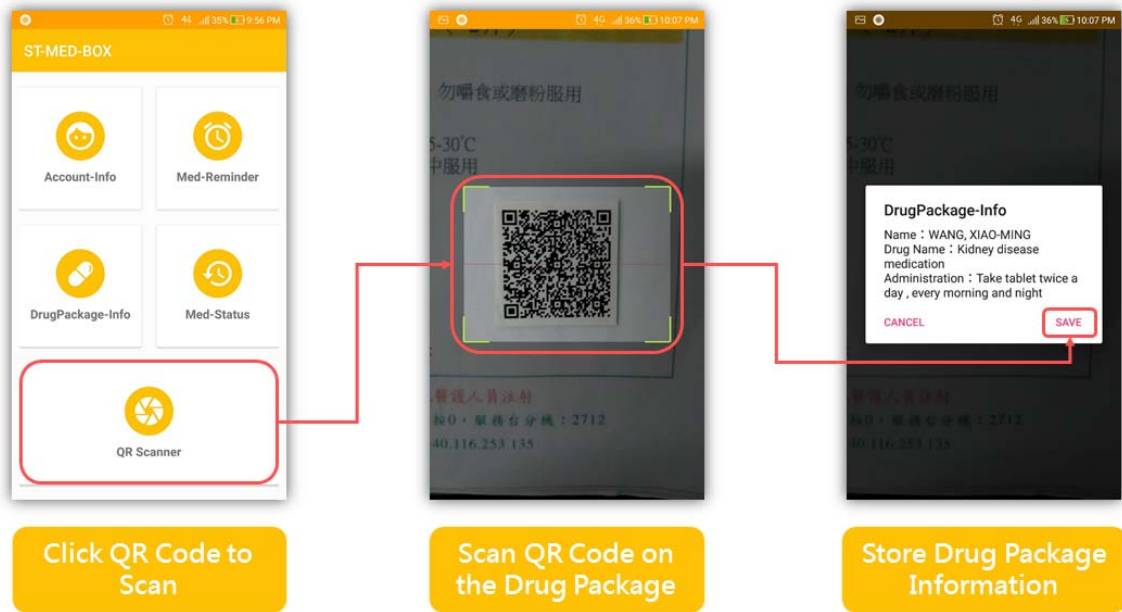


Fig. 5. Operation screen: QR code scanning—drug package.

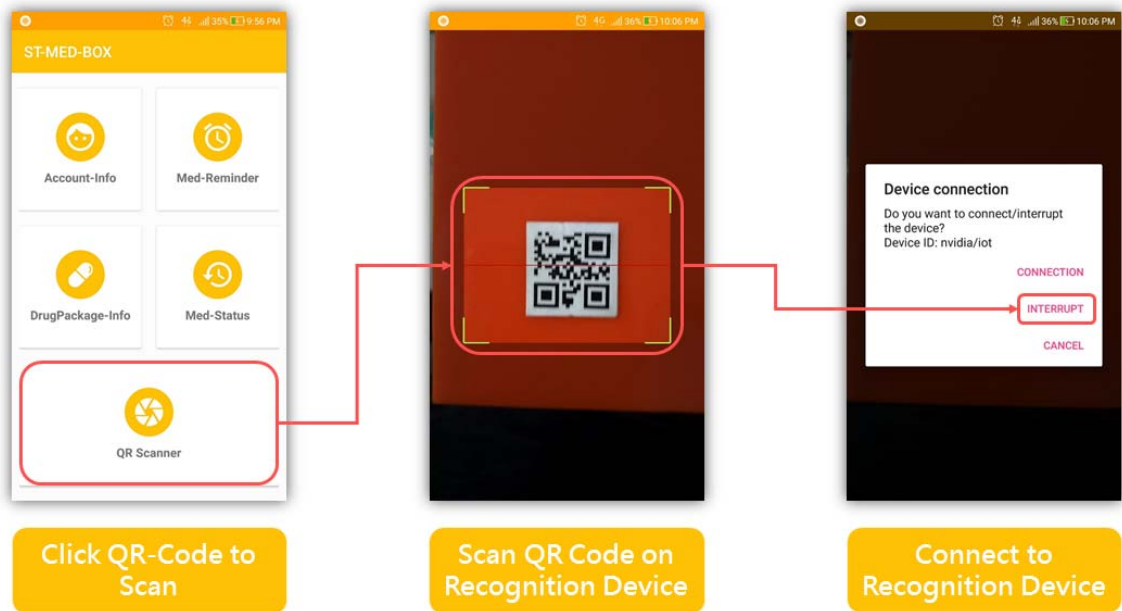


Fig. 6. Operation screen: QR code scanning—the proposed intelligent medicine recognition device.

and initiate data transmission between the Android-based mobile device and the proposed intelligent medicine recognition device, as shown in Fig. 6.

Hence, this Android-based mobile device app can be used to view the information corresponding to a scanned medicine package and the recognition results generated by the proposed intelligent medicine recognition device to ensure proper medication usage for patients with chronic diseases, as shown in Fig. 7.

In addition to medication status checking, this app also provides medication time reminders. The medication reminder

setting screen is used to enter the time when a patient needs to take his or her medicine. The app will count down until the specified time arrives; the mobile device will then send out a notification and issue a ringtone to remind the patient to take his or her medicine, as shown in Fig. 8.

Finally, the basic information of the current user, such as the current user account, the total amount of medication required, the number of times the proposed intelligent medicine recognition device has recognized medicine, the connection ID, the connection status and the number, can also be viewed in the app.

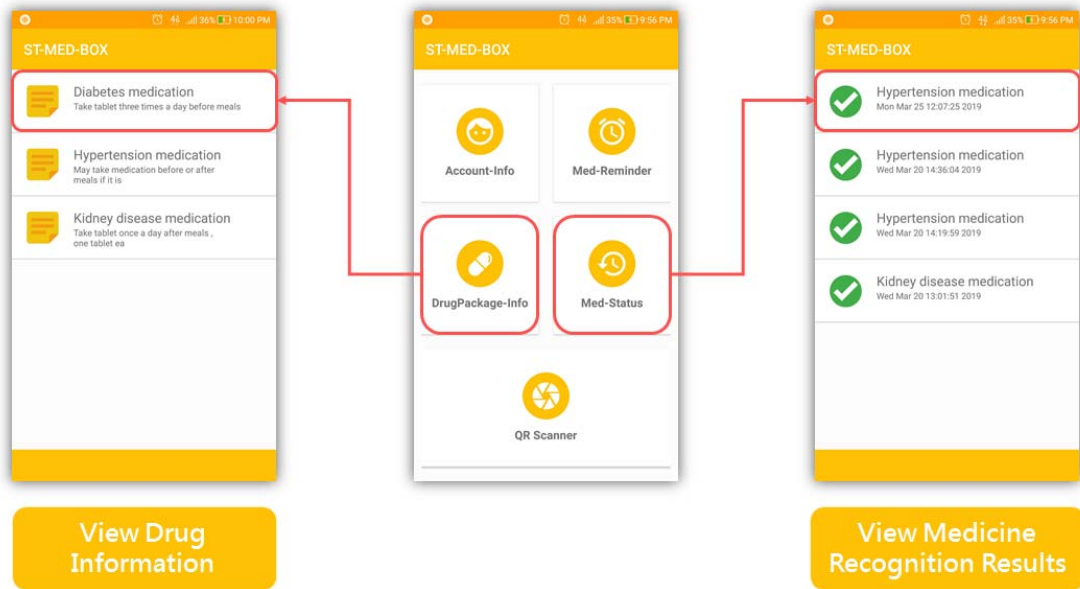


Fig. 7. Operation screen: viewing the information from scanned medicine packages and the recognition results generated by the proposed intelligent medicine recognition device to ensure proper medication usage for patients with chronic diseases.

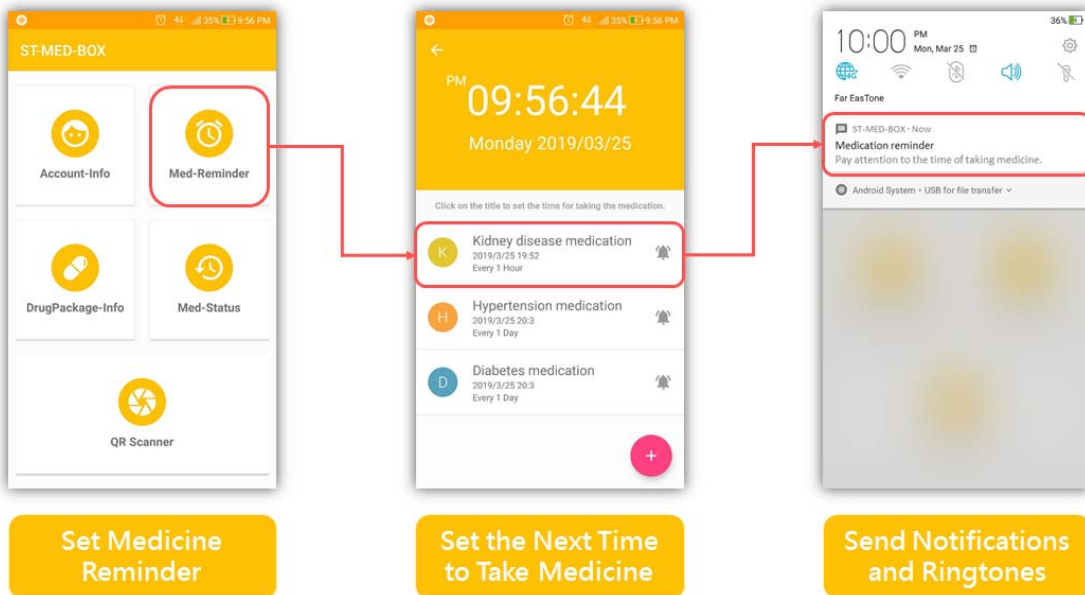


Fig. 8. Operation screen: medication time reminder.

D. Cloud-based Management Platform

In addition to viewing medication records on the proposed Android-based mobile device app, a patient's medication records stored on the proposed cloud-based management platform can also be viewed via a website, as shown in Fig. 9.

The back-end information system of the proposed cloud-based management platform is implemented with PHP and MySQL. The related information stored in the MySQL database includes the mobile app client data, the QR code information obtained via the mobile app by scanning medication packages, and the drug recognition results

generated by the proposed intelligent medicine recognition device.

IV. DEEP LEARNING TECHNIQUES

In this paper, Google TensorFlow is adopted as our deep learning development framework. Google TensorFlow compares poorly with other frameworks (such as Caffe, MXNet, and Torch) in terms of CPU performance, but its ability to use GPU computing resources is superior to that of other frameworks.

Based on the TensorFlow framework, an open-source

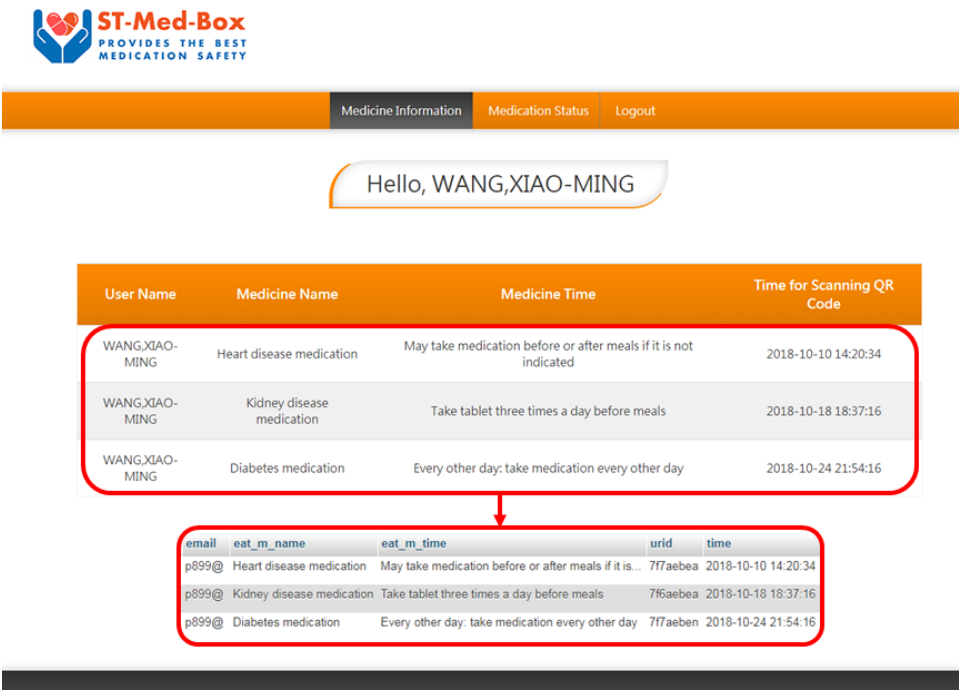


Fig. 9. Viewing a patient’s medication information on the proposed cloud-based management platform.

module combining Faster R-CNN and Inception V3 [38], [39] is used as the basis for the system presented in this paper. For training, XML files containing drug information and corresponding drug images are converted into the “TFRecord” format, and an adjusted “Config” file is loaded, as shown in Fig. 10.

The API provided by Google TensorFlow is used for image recognition, and the related parameters are adjusted specifically for drug image recognition in the proposed system. The flowchart of the deep-learning-based drug image recognition process in the proposed system is shown in Fig. 11.

As shown in Fig. 11, our deep-learning-based drug image recognition process combines Faster R-CNN and Inception V3 modules. The deep-learning-based drug image recognition process performed in the proposed system will be introduced in the following subsections.

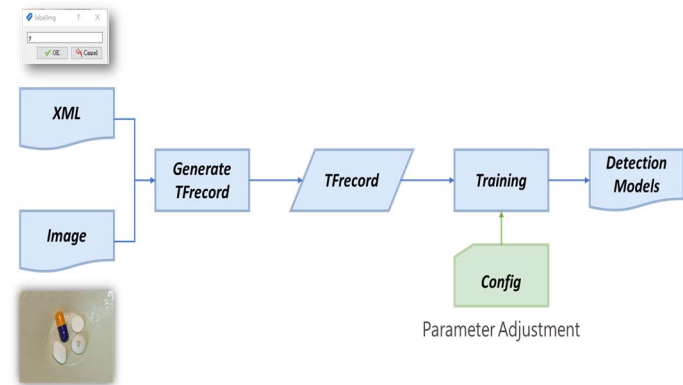


Fig. 10. Flowchart of the training process.

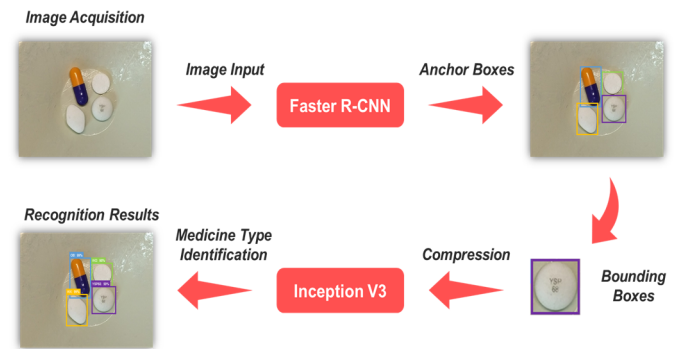


Fig. 11. Flowchart of the image recognition process.

A. Faster R-CNN [40]

As shown in the above flowchart (Fig. 11), a medicine image is first input into the first module (*namely*, the Faster R-CNN module). Faster R-CNN is a CNN-based object detection methodology, as shown in Fig. 12.

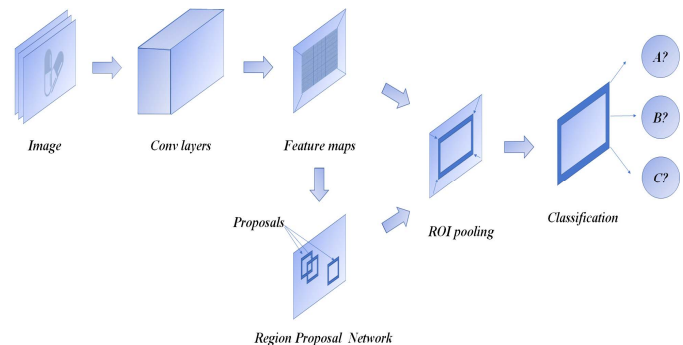


Fig. 12. Flowchart of the Faster R-CNN module.

First, feature maps are extracted through the Conv layers [40]. Then, the feature maps are transferred to the region proposal network (RPN) layer, which generates frames around the possible object-containing regions.

Then, the various output framed areas are scaled to a fixed size by the region of interest (ROI) pooling layer and finally sent to the classifier to determine which of these framed areas contain objects.

1) Conv Layers

These layers include convolutional layers, activation functions, and pooling layers. For all convolutions in the Faster R-CNN Conv layers, the dimensions of the original image are initially expanded to $(M+2) \times (N+2)$; then, after a 3×3 convolution, an $M \times N$ output is generated. Hence, the convolutional layers in the Conv layers do not change the matrix size between the input and output, as shown in Fig. 13.

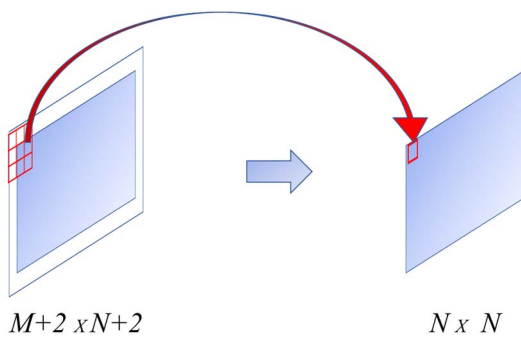


Fig. 13. Edge expansion to maintain the original size.

The kernel size of each pooling layer in the Conv layers is 2×2 , and the stride is 2. These settings reduce the dimensions of the input by half, and reduce the computations by 75%, such that an $M \times N$ matrix passing through a pooling layer will be reduced to a size of $(M/2) \times (N/2)$. In summary, among the entire set of Conv layers, the convolutional and activation function layers do not change the size of the input; only the pooling layers cause a change in size, with the output length and width each being $1/2$ of the corresponding dimension of the input. Thus, an $M \times N$ matrix is ultimately reduced in size to $(M/16) \times (N/16)$ by the Conv layers. Therefore, the feature maps generated by the Conv layers can be associated with the original image as shown in Fig. 14.

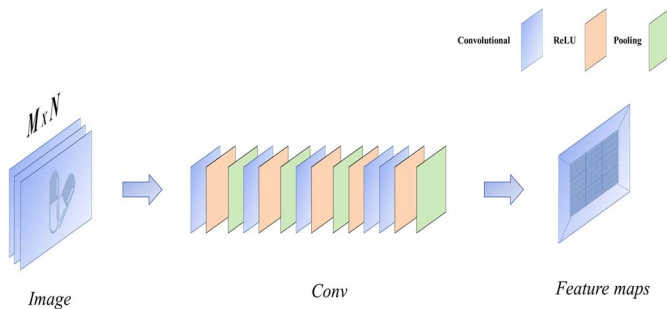


Fig. 14. Feature map generation.

2) Region Proposal Network (RPN)

An RPN uses a CNN to directly generate regions containing

possible objects. The foreground and background are obtained by means of softmax classifying anchors (the detection target is the foreground) because the anchor mechanism and frame regression can be used to generate multiscale region proposals with various aspect ratios. The RPN is also a fully connected convolutional network that can be trained on the frames of the generated areas to simultaneously predict the boundary and scope of an object, and it contains 2 additional convolutional layers (Reg and Cls layers), as shown in Fig. 15. The Reg layer predicts the (x, y, w, h) coordinates of a proposal, corresponding to the anchor of the proposal. The Cls layer determines whether the proposal is foreground or background.

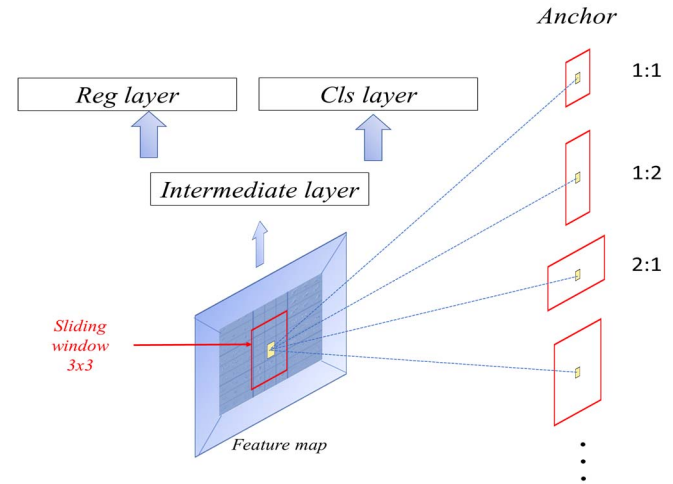


Fig. 15. Additional convolutional layers (Reg and Cls layers).

3) Anchors

The center point of the 3×3 convolution kernel corresponds to a position on the original image, and this point is used as the center point of each anchor. Based on this point, multiscale anchors with various aspect ratios are arranged on the original image. Therefore, the anchors refer not to the feature maps generated by the Conv layers but to the original image, as shown in Fig. 16.

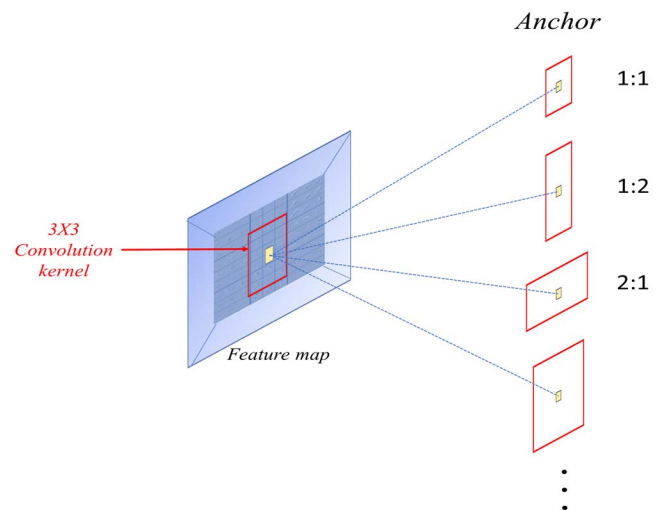


Fig. 16. Multiscale frames generated by anchors.

4) Bounding Box Regression

In Fig. 17, the yellow frame represents the ground truth (GT) of the target object, and the red frame represents the extracted foreground anchor. Even if the red frame is recognized as the target object by the classifier, if the red frame is not positioned accurately, the detection result will be considered incorrect, and the target object will be out of the frame. Thus, we need a method to fine-tune the red box to bring the foreground anchor closer to the GT, as shown in Fig. 17.

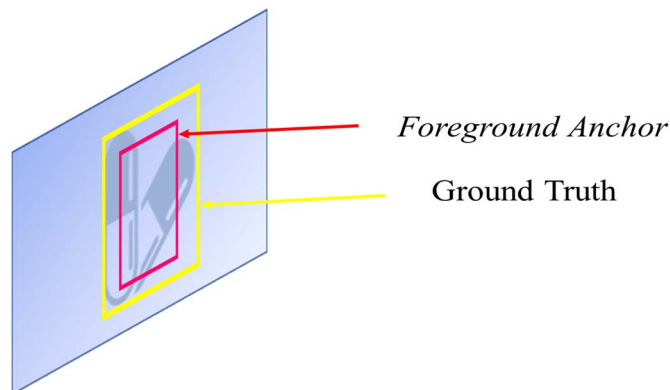


Fig. 17. Bounding box regression.

5) ROI Pooling

The ROI pooling layer has two main functions: (1) align an object proposal in the image with the corresponding patch in the feature map and (2) collect the sample corresponding to this feature map patch and transfer it to the fully connected layer.

The ROI pooling layer evenly divides each candidate area into $M \times N$ blocks and performs max pooling on each block. Thus, proposals of different sizes on the feature maps are converted into data of uniform size and sent to the next layer, as shown in Fig. 18.

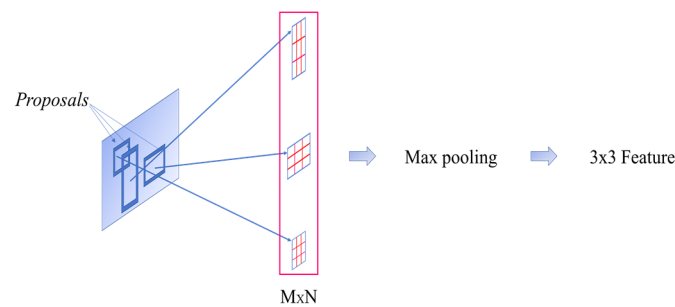


Fig. 18. ROI pooling process.

6) Classification

Classification is performed using the feature map regions obtained for each proposal to determine the category to which each proposal belongs by means of the fully connected layer and the softmax layer. Bounding box regression is used to correct each anchor box to a more precise position, as shown in Fig. 19.

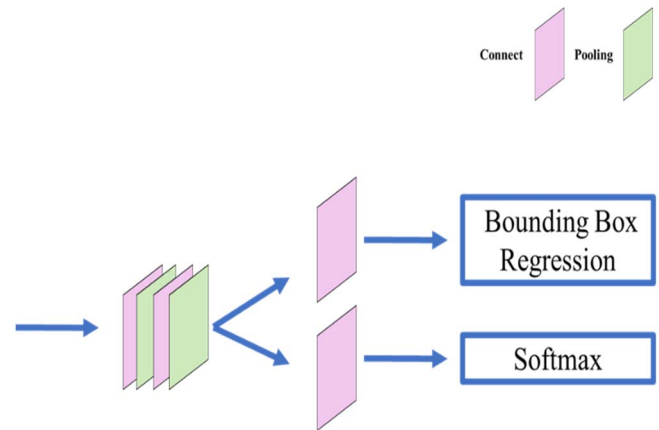


Fig. 19. Classification flowchart.

B. Inception V3 [41]

The Inception V3 module (see Fig. 20) has a relatively simple architecture compared to other high-performance neural networks, a more complete architecture and a modest computational cost. Thus, the Inception V3 module is appropriate to use when large amounts of data need to be processed at a reasonable cost or when memory and computing power are limited.

The architecture of the Inception V3 module [41] also further improves the classification effect on ImageNet, which is the main reason the Inception V3 module was selected for the recognition process in the proposed system.

In the recognition process of the proposed system, the main function of the Inception V3 module is to receive the compressed image regions in the generated anchor boxes and to predict the possible object types in those regions. Finally, the complete prediction/recognition result is output as an image.

1) Asymmetric Convolutions

The design concept of the Inception V3 architecture is to find the optimal local sparse structure in a convolutional visual network. This structure needs to be covered and predicted by the available dense components; to this end, convolutional building blocks are spatially reused to create the network [41], as shown in Fig. 21.

In this layer-by-layer structure, the relevant statistics for each layer are analyzed to identify clusters with high correlation; these clusters then form the input to the next layer and are transferred through connected with the neurons of the previous layer [42].

To avoid patch alignment issues, the filter sizes in the Inception structure are limited to 1×1 , 3×3 , and 5×5 , mostly for convenience rather than necessity [42]. Thus, in a reasonable network structure, the output filters of each layer should be combined into a single vector to form the input to the next layer.

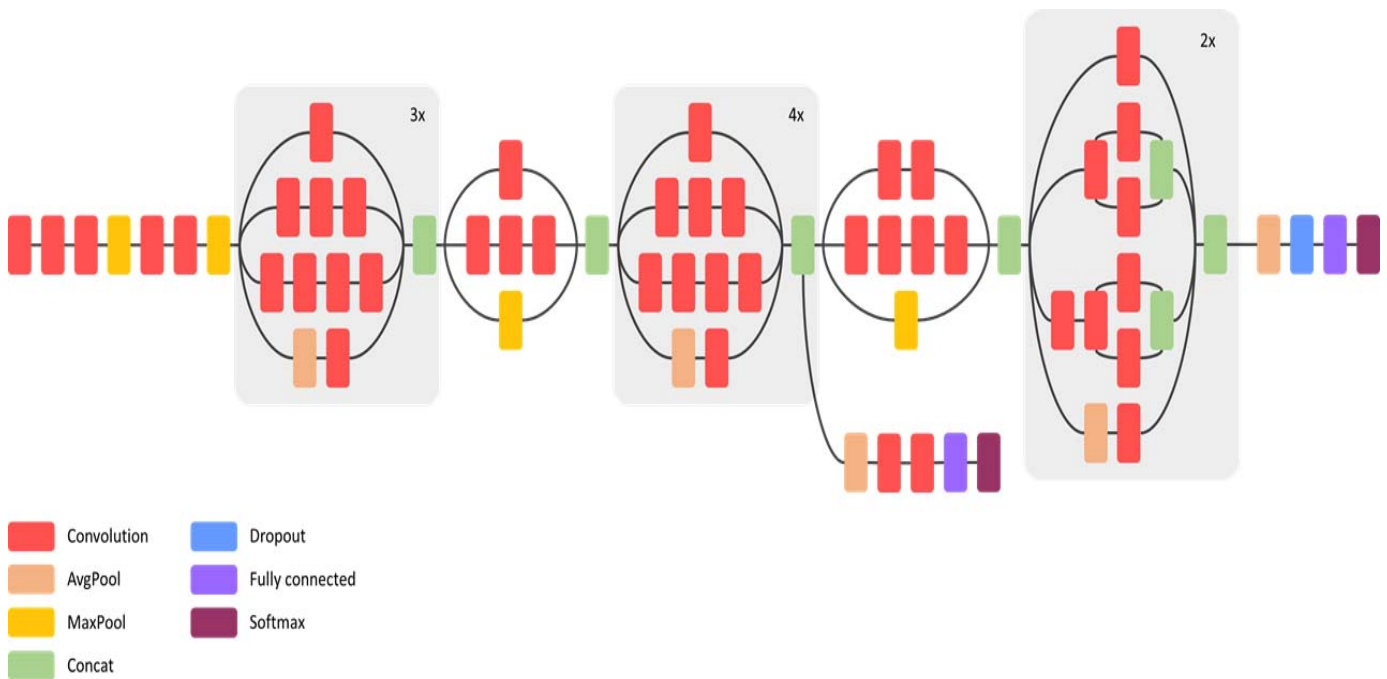


Fig. 20. Architecture of the Inception V3 module [41].

Hence, we assume that each cell of the previous layer corresponds to some area of the input image and that these cells are grouped into filters. At the level closest to the input, the associated units are focused on local areas, and a large number of units that are focused on the same area are obtained and are covered by the 1×1 convolution of the next layer [43].

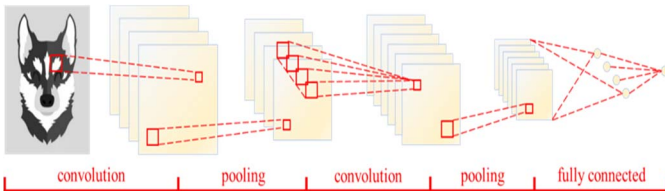


Fig. 21. Spatial reuse of convolutional building blocks to create the network.

In addition, pooling operations are very important for Inception's convolutional network; hence, a parallel pooling path is added to reduce the amount of data and the computation time, as shown in Fig. 22.

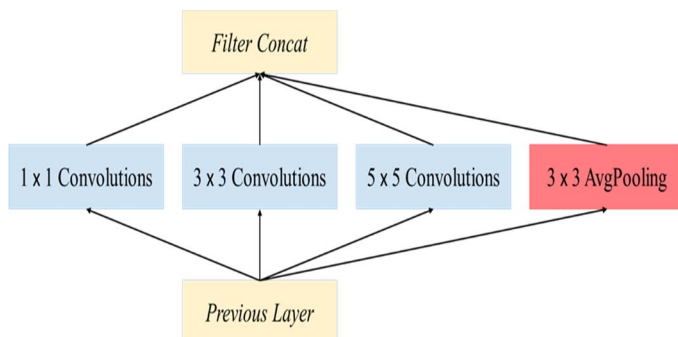


Fig. 22. Initial version of Inception.

However, the above model has a large problem in that the number of output filters is the same as the number of filters in the previous process. Thus, the combination of the pooling layer output and the convolutional layer output leads to a sharp increase in the size of the output of each step.

To solve this problem, a 1×1 convolution is added before each of the 3×3 and 5×5 convolutions for dimensionality reduction [42], and a 1×1 convolution is also used as a linear activation function after rectification. The resulting architecture is shown in Fig. 23.

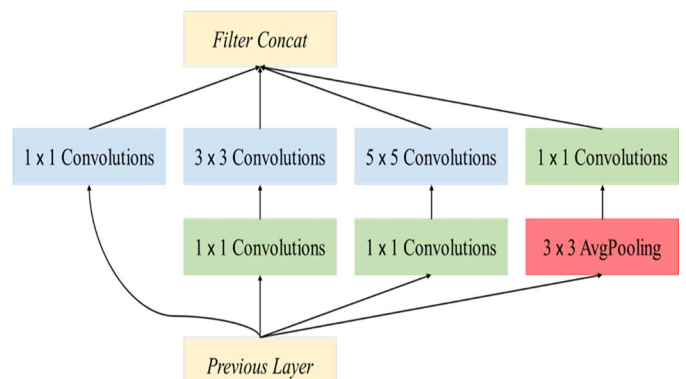


Fig. 23. Inception with size compression.

In general, dimensionality reduction can be applied to prevent a large number of filters from a previous layer from flowing into the next layer of filters by reducing the size of the input before convolution is performed on large blocks. An advantageous aspect of this structure is that it allows a large increase in the number of neurons per step without causing an

increase in computational effort.

However, a convolution with large spatial filters is usually very computationally expensive. For the same number of filters, a 5×5 convolution is $25/9 = 2.78$ times more expensive than a 3×3 convolution, and using a filter larger than 3×3 for convolution is usually not very useful because it can be reduced to a sequence of 3×3 convolutions [41].

With asymmetric ($n \times 1$) convolution (Fig. 24), even better efficiency can be achieved. For example, a 3×1 convolution followed by two 1×3 convolutions can be 33% less computationally expensive than 3×3 convolution, and as the value of n increases, the computational savings are more significant.

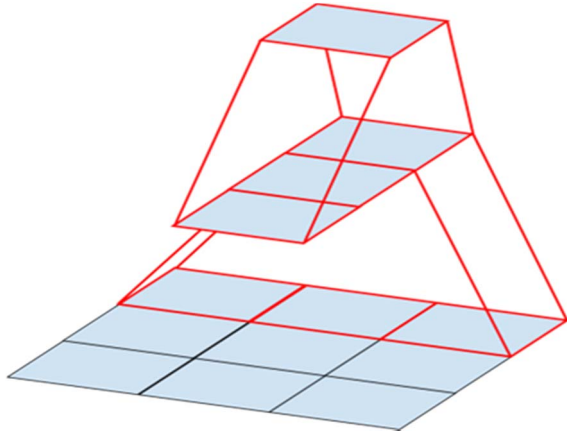


Fig. 24. Mininetwork replacing a 3×3 convolution [41].

Hence, in the Inception V3 architecture, each 5×5 convolution is replaced with two 3×3 convolutions (Fig. 25), and each 3×3 convolution is then replaced with a 3×1 ($n \times 1$) convolution followed by two 1×3 ($1 \times n$) convolutions (Fig. 26). Dividing the architecture into a multilayer structure in this way can save computations [41], especially for a medium grid size on an $m \times m$ feature map. When m is between 12 and 20, the effect is particularly obvious.

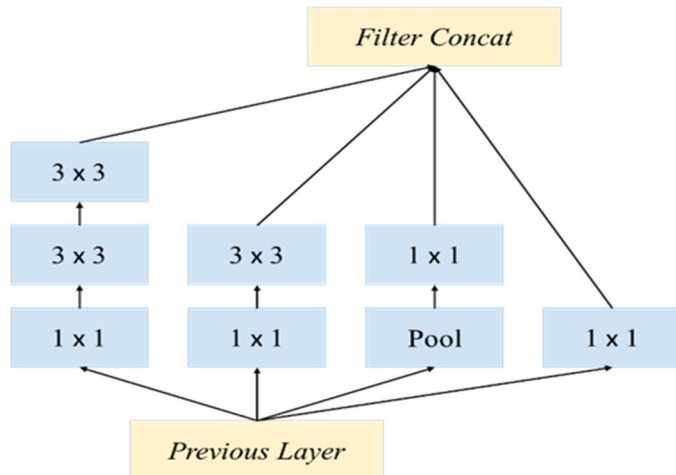


Fig. 25. Inception module with each 5×5 convolution replaced with two 3×3 convolutions [41].

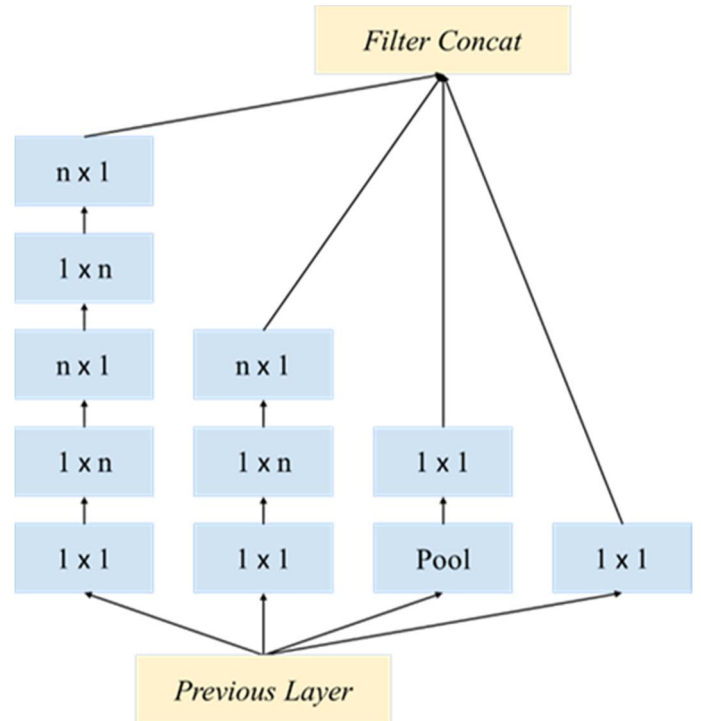


Fig. 26. Inception module after the factorization of the $n \times n$ convolutions [41].

2) Extending the Filter Bank Output

The profile of the Inception V3 network architecture is shown in Table I. Based on the method described above, in Inception V3, the traditional 7×7 convolutions are decomposed into three 3×3 convolutions, while the overall network effectively applies a 35×35 convolution.

Three traditional Inception modules (Fig. 27), each with 288 filters, are then reduced to a 17×17 grid, with 768 filters, using mesh reduction techniques (Fig. 28), and that grid is then further reduced to an $8 \times 8 \times 1280$ grid. The Inception module with extended filter output (Fig. 27) is applied to this coarsest 8×8 grid to facilitate high-dimensional representation [41].

3) Performance Comparison

Table II shows the results reported for the recognition performance of the Inception V3 architecture for multicropped images and multiple combined models and compares these results with those published for other networks based on the ILSVRC 2012 classification benchmark dataset [44].

Compared to other modern deep learning architectures, the Inception V3 architecture is relatively simple, has a more complete architecture, and has a modest computational cost. The error rates achieved through a combination of four models and multiple evaluations are 17.2% (top 1) and 3.58% (top 5), and the network has fewer than 25 million parameters and requires approximately 5 billion operations.

TABLE I
PROFILE OF THE INCEPTION V3 NETWORK ARCHITECTURE [41]

Type	Patch size/stride or remarks	Input size
conv	$3 \times 3 / 2$	$299 \times 299 \times 3$
conv	$3 \times 3 / 1$	$149 \times 149 \times 32$
conv padded	$3 \times 3 / 1$	$147 \times 147 \times 32$
pool	$3 \times 3 / 2$	$147 \times 147 \times 64$
conv	$3 \times 3 / 1$	$73 \times 73 \times 64$
conv	$3 \times 3 / 2$	$71 \times 71 \times 80$
conv	$3 \times 3 / 1$	$35 \times 35 \times 192$
3xInception	As in figure 5	$35 \times 35 \times 288$
5xInception	As in figure 6	$17 \times 17 \times 768$
2xInception	As in figure 7	$8 \times 8 \times 1280$
pool	8×8	$8 \times 8 \times 2048$
linear	logits	$1 \times 1 \times 2048$
softmax	classifier	$1 \times 1 \times 1000$

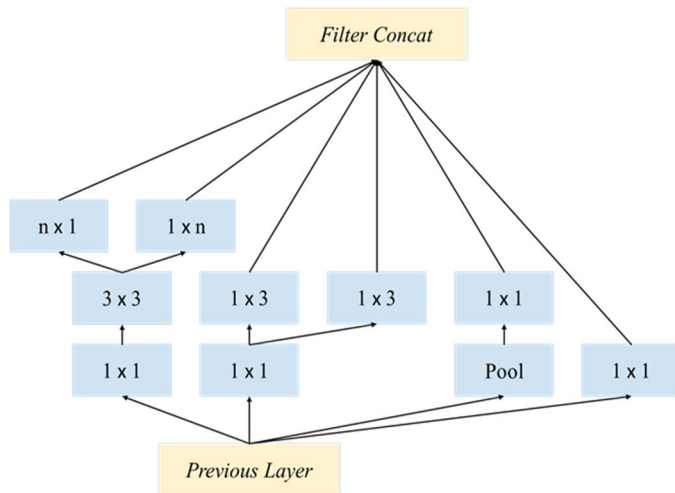


Fig. 27. Inception module with extended filter bank output [41].

TABLE II
COMPARISON OF REPORTED MULTIMODEL, MULTICROP RESULTS ON THE
ILSVR 2012 CLASSIFICATION BENCHMARK [41]

Network	Models Evaluated	Crops Evaluated	Top-1 Error	Top-5 Error
Simonyan & Zisserman (VGGNet) [45]	2	-	23.7%	6.8%
Szegedy <i>et al.</i> (GoLeNet) [42]	7	144	-	6.67%
He <i>et al.</i> (PReLU) [46]	-	-	-	4.94%
Ioffe & Szegedy (BN-Inception) [43]	6	144	20.1%	4.9%
Szegedy <i>et al.</i> (Inception V3) [41]	4	144	17.2%	3.58%

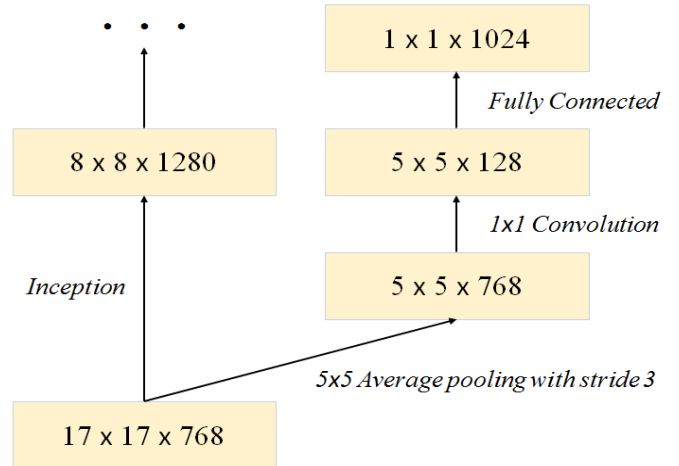


Fig. 28. Auxiliary classifier on top of the last 17×17 layer [41].

V. PROTOTYPE AND EXPERIMENTS

A. Prototype

Fig. 29 shows a photograph of a prototype of the proposed system. The proposed system is composed of an intelligent medicine recognition device, an app running on an Android-based mobile device, a deep learning training server, and a cloud-based management platform.

To make the system more useful in daily life, patients with multiple medications can scan the QR codes on medicine packages using our proposed mobile app, and the related medicine usage information is then uploaded to the proposed cloud-based management platform for storage.

Patients can also view the actual status of their medicine recognition results via our proposed mobile device app or our proposed cloud-based management platform for immediate drug safety monitoring.

The proposed mobile device app and the proposed cloud-based platform can provide complete records, high-accuracy recognition, and simple operation to meet the needs of patients with multiple medications for ensuring medication safety.

Fig. 30 presents an example of the steps of operation of the proposed system. First, a hospital or pharmacy can provide a QR code corresponding to a patient's medication information.

The patient or one of his or her family members can then scan the QR code on the medicine package and transmit the corresponding medication information, such as the medication time, over Wi-Fi/4G networks to the proposed intelligent medicine recognition device for storage. This capability is important because when chronic patients or elderly patients are prescribed a wide variety of drugs, they may take the wrong medicine or forget to take their medicine.



Fig. 29. A photograph of a prototype of the proposed system.

At the correct time, the proposed intelligent medicine recognition device will send a ringtone notification to remind the patient to take his or her medication. Then, the patient should place the medicine into the recognition region of the proposed device and press the push button to initiate the recognition process, after which the device will provide a voice description to explain the name, use, side effects and dosage of the drug and state whether the selected drug is correct.

Finally, via the proposed app, family members can check whether chronic patients or elderly patients are taking the correct medication to ensure safe medication behavior.

B. Experiments

The deep learning framework of the proposed system is based on the TensorFlow framework, which was developed by Google. Our training module for image recognition is based on the `faster_rcnn_inception_v2_coco` module, which combines open-source Faster R-CNN and Inception V3 modules.

Classification learning is executed 200,000 times on the deep learning training server. The learning time is approximately 26 hours.

A medicine recognition module is generated after learning is finished, and this module is uploaded to the proposed intelligent medicine recognition device for medicine recognition.

For this experiment, we prepared 8 types of image data for drug image recognition training. A total of 5,000 640×480 images of medications were obtained, of which 4,000 were images of single drugs and 1,000 were images of multiple drugs. The code names were simplified for the recognition process, as shown in Fig. 31.

Next, we converted the drug images and the corresponding XML files specifying the drug scopes into the “TFRecord” format and accelerated the packaged binary files to the speed of the training module.

In this study, we used the `faster_rcnn_inception_v2_coco` module to generate a “Config” file to adjust the parameters and then carried out 200,000 training steps on the deep learning training server; the training time was approximately 26 hours.

The specifications of the deep learning training server can be seen in Table III, and the resources used by the deep learning training modules are listed in Table IV.

Table V shows the recognition rates for the 8 types of drugs used in the test reported in this paper. The experimental results show that a recognition accuracy of up to 96.6% was achieved. Fig. 32 presents examples of actual photographs captured during the actual drug recognition process and the recognition results generated by the proposed intelligent medicine recognition device.

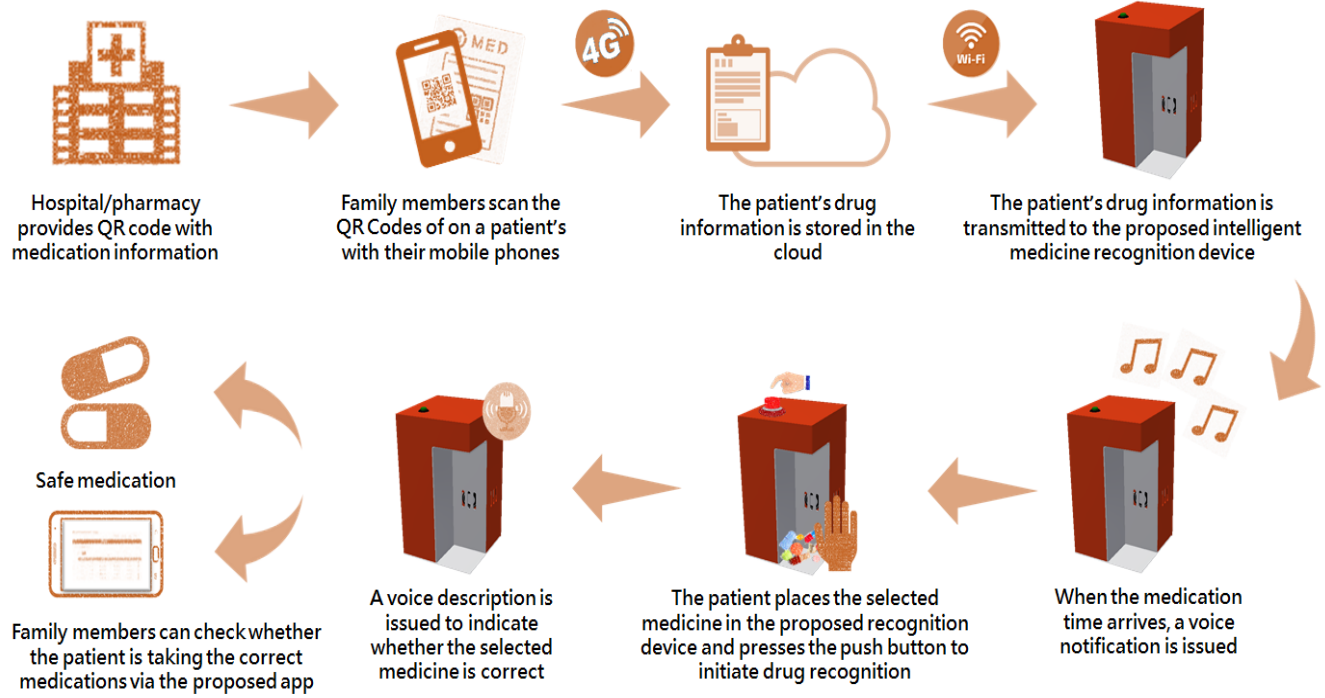


Fig. 30. Steps of operation of the proposed system in a typical scenario.



Fig. 31. Code names for 8 types of drugs.

TABLE III
SPECIFICATIONS OF THE DEEP LEARNING TRAINING SERVER

Item	Specifications
CPU	Intel i7 7800X
GPU	Nvidia GTX 1080
RAM	16 GB * 4
SSD	512 GB M.2 PCIe
HDD	Seagate 1TB 7200 RPM Enterprise Disk

TABLE IV
RESOURCES USED BY THE DEEP LEARNING TRAINING MODULES

	RAM (MB)	CPU (%)	GPU (%)	Detection Time (s)
ssd_mobilenet_v1_coco	1,716	22.8	4	1.4
ssd_inception_v2_coco	2,847	33.8	4	1.7
faster_rcnn_inception_v2_coco	2,454	96.5	45	2.9
faster_rcnn_resnet50_coco	3,802	98.3	53	3.5
rfcn_resnet101_coco	4,393	92.1	26	4.0

TABLE V
RECOGNITION RATES FOR 8 TYPES OF DRUGS

Drug Code	Y	YSP68	RT927	WO
Recognition Rate	98.2%	98.5%	96.6%	99.1%
Drug Code	OB	WX	G	O
Recognition Rate	98.3%	99.2%	99.4%	99.5%

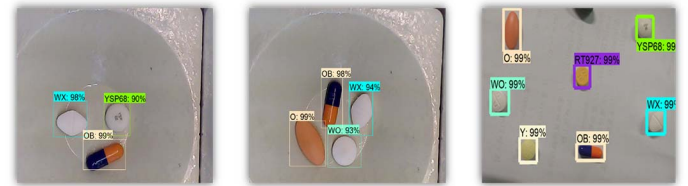


Fig. 32. Photographs captured during the actual drug recognition process and the recognition results generated by the proposed intelligent medicine device.

VI. CONCLUSIONS

Chronic patients, including 480 million elderly people in the world today, suffer from a variety of diseases. In the treatment of multiple chronic diseases, many drugs are needed, and physiological functions decline. Cognitive ability is reduced, possibly causing patients to take the wrong medicine. Therefore, elderly people have become a high-risk group for adverse drug events.

To solve the problem of taking the wrong medicine, in this paper, we have successfully developed an intelligent medicine recognition system named ST-Med-Box based on deep learning technology. This system can recognize drugs and deliver recognition results in a systematic and practical way. The

chronic disease drug recognition rate of the proposed system reaches 96.6% or higher; thus, it can help patients to take their medications more safely and securely.

The proposed system can automatically provide notifications stating the names of drugs and indicating medication times to address the problem of lapses in human judgment. Moreover, the proposed system incorporates a cloud-based database to provide patients with additional integrated information services.

As a result, when using the proposed system, chronic patients do not need to worry about forgetting to take their medicine. They need only download the proposed Android-based mobile device app and scan the QR codes on their medicine packages to store the corresponding medication information. Then, they can access related services, such as medication reminders and records. Consequently, the proposed system can effectively reduce the problem of drug interactions caused by taking incorrect drugs, thereby reducing the cost of medical treatment and giving patients with chronic diseases a safe medication environment.

In our future work, we will cooperate with a pharmacy to train the system on more drugs. We will first ask pharmacies to apply to participate in our research and identify a suitable pharmacy with which to collaborate. As the pharmacy continues to provide us with chronic disease drugs for testing, we will continue to perform deep learning training to continuously improve the recognition accuracy of the system.

ACKNOWLEDGMENTS

The authors would like to thank Ms. Yu-Ching Mo from the Artificial Intelligence over the Internet of Things Applied Research Center (AIoT Center) at the Southern Taiwan University of Science and Technology in Tainan, Taiwan, R.O.C., for supporting and assisting with the cloud-based database management platform.

REFERENCES

- [1] "PillDrill-Smart medication tracking that simply works," 2018. [Online]. Available: <https://www.pilldrill.com/>.
- [2] M. Brown, "PillDrill review: A high-tech medication-tracking system," 2018. [Online]. Available: <https://www.techhive.com/article/3270733/health-fitness/pilldrill-review.html>
- [3] "Preventable adverse drug reactions: a focus on drug interactions," U.S. Food and Drug Administration, 2018. [Online]. Available: <https://www.fda.gov/drugs/developmentapprovalprocess/developmentresources/druginteractionslabeling/ucm110632.htm>
- [4] S. Dev, S. Zhang, J. Voyles, and A. S. Rao, "Automated classification of adverse events in pharmacovigilance," in *Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1562-1566.
- [5] Z. Yaniv, J. Faruque, S. Howe, K. Dunn, D. Sharlip, A. Bond, P. Perillan, O. Bodenreider, M. J. Ackerman, and T. S. Yoo, "The national library of medicine pill image recognition challenge: an initial report," in *Proceedings of the 2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2016, pp. 1-9.
- [6] D. Ushizima, A. Carneiro, M. Souza, and F. Medeiros, "Investigating pill recognition methods for a new national library of medicine image dataset," in *Proceedings of the International Symposium on Visual Computing (ISVC)*, Part II, LNCS 9475, 2015, pp. 410-419.
- [7] B. Riheiro, N. Vlves, M. Guevara, and L. Magalhães, "A three-staged approach to medicine box recognition," in *Proceedings of the 2017 24th Encontro Português de Computação Gráfica e Interação (EPCGI)*, 2017, pp. 1-7.
- [8] J.-S. Wang, A. M. Ambikapathi, Y. Han, S.-L. Chung, H.-W. Ting, and C.-F. Chen, "Highlighted deep learning based identification of pharmaceutical blister packages," in *Proceedings of the 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2018, pp. 638-645.
- [9] J. Yu, Z. Chen, S.-i Kamata, and J. Yang, "Accurate system for automatic pill recognition using imprint information," *IET Image Processing*, vol. 9, no. 12, 1039-1047, 2015.
- [10] M. A. V. Neto, J. W. M. de Souza, P. P. R. Filho, and A. W. de O. Rodrigues, "CoforDes: an invariant feature extractor for the drug pill identification," in *Proceedings of the 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, 2018, pp. 30-35.
- [11] R. A. Calix, R. Gupta, M. Gupta, and K. Jiang, "Deep gramulator: Improving precision in the classification of personal health-experience tweets with deep learning," in *Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1154-1159.
- [12] H. Kalantarian, N. Alshurafa, E. Nemati, Tuan Le, and M. Sarrafzadeh, "A smartwatch-based medication adherence system," in *Proceedings of the 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2015, pp. 1-6.
- [13] H. Kalantarian, N. Alshurafa, and M. Sarrafzadeh, "Detection of gestures associated with medication adherence using smartwatch-based inertial sensors," *IEEE Sensors Journal*, vol. 16, no. 4, pp. 1054-1061, Feb. 2016.
- [14] J. Ma, A. Ovalle, D. M.-K. Woodbridge, "Medhere: A smartwatch-based medication adherence monitoring system using machine learning and distributed computing," in *Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 4945-4948.
- [15] M. Aldeer, R. P. Martin, and R. E. Howard, "PillSense: designing a medication adherence monitoring system using pill bottle-mounted wireless sensors," in *Proceedings of the 2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018, pp. 1-6.
- [16] L.-B. Chen, W.-J. Chang, K.-M. Lee, C.-W. Huang, and K. S.-M. Li, "A comprehensive medicine management system with multiple sources in a nursing home in Taiwan," *IEICE Transactions on Information and Systems*, vol. E99-D, no. 6, pp. 1447-1454, Jun. 2016.
- [17] E.-Y. Kim, "Development and application of direct data capture for monitoring medication compliance in clinical trials," *Healthcare Informatics Research*, vol. 23, no. 4, pp. 249-254, Oct. 2017.
- [18] H. F. Badawi, F. Laamarti, and A. E. Saddik, "ISO/IEEE 11073 personal health device (X73-PHD) standards compliant systems: a systematic literature review," *IEEE Access*, vol. 7, pp. 3062-3073, 2019.
- [19] M. D. Sanctics, "Personalized ubiquitous health and wellness care: the knowledge-care vision," *Wireless Personal Communications*, vol. 88, no. 1, pp. 53-62, Mar. 2016.
- [20] T. Mugammed, R. Mehmood, A. Albeshri, and I. Katib, "UbeHealth: a personalized ubiquitous cloud and edge-enabled networked healthcare system for smart cities," *IEEE Access*, vol. 6, pp. 32258-32285, 2018.
- [21] E. C. Schiza, T. Kyprianou, N. Petkov, and C. N. Schizas, "Proposal for an health based ecosystem serving national healthcare," *IEEE Journal of Biomedical and Health Informatics*, pp. 1-13, 2018. (Early Access Articles)
- [22] M. M. Alam, H. Malik, M. I. Khan, T. Pardy, A. Kuusik, and Y. L. Moullec, "A survey on the roles of communication technologies in IoT-based personalized healthcare applications," *IEEE Access*, vol. 6, pp. 36611-36631, 2018.
- [23] S. B. Baker, W. Xiang, and I. Atkinson, "Internet of Things for smart healthcare: technologies, challenges, and opportunities," *IEEE Access*, vol. 5, pp. 26521-26544, 2017.
- [24] A. Alamri, "Monitoring system for patients using multimedia for smart healthcare," *IEEE Access*, vol. 6, pp. 23271-23276, 2018.
- [25] X. Li, X. Huang, C. Li, R. Yu, and L. Shu, "EdgeCare: leveraging edge computing for collaborative data management in mobile healthcare systems," *IEEE Access*, vol. 7, pp. 22011-22025, 2019.
- [26] J. Yoon, C. Davtyan, and M. van der Schaar, "Discovery and clinical decision support for personalized healthcare," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 4, pp. 1133-1145, 2017.
- [27] Y. Gao, X. Xiang, N. Xiong, B. Huang, H. J. Lee, R. Alrifai, X. Jiang, and Z. Fang, "Human action monitoring for healthcare based on deep learning," *IEEE Access*, vol. 6, pp. 52277-52285, 2018.

- [28] S. Kim, S. Yeom, O.-J. Kwon, D. Shin, and D. Shin, "Ubiquitous healthcare system for analysis of chronic patients' biological and lifelog data," *IEEE Access*, vol. 6, pp. 8909-8915, 2018.
- [29] P. Verma and S. K. Sood, "Fog assisted-iot enabled patient health monitoring in smart homes," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1789-1796, 2018.
- [30] G.-W. Kim and D.-H. Lee, "Intelligent health diagnosis technique exploiting automatic ontology generation and web-based personal health record services," *IEEE Access*, vol. 7, pp. 9419-9444, 2019.
- [31] L. Yu, W. M. Chan, Y. Zhao, and K.-L. Tsui, "Personalized health monitoring system of elderly wellness at the community level in Hong Kong," *IEEE Access*, vol. 6, pp. 35558-35567, 2018.
- [32] S. Jiménez-Fernández, P. de Toledo, and F. del Pozo, "Usability and interoperability in wireless sensor networks for patient telemonitoring in chronic disease management," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3331-3339, 2013.
- [33] W.-C. Wang, L.-B. Chen, W.-J. Chang, "Development and experimental evaluation of machine-learning techniques for an intelligent hairy scalp detection system," *Applied Sciences*, vol. 8, no. 5, article 853, pp. 1-28, May 2018.
- [34] H. Li, K. Ota, M. Dong, "Learning IoT in edge: deep learning for the internet of things with edge computing," *IEEE Network*, vol. 32, no. 1, pp. 96-101, Jan.-Feb. 2018.
- [35] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: a survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347-2376, Jun. 2015.
- [36] D. Locke, "MQ telemetry transport (MQTT) v3. 1 protocol specification," IBM, Aug. 2010, [online] Available: <https://www.ibm.com/developerworks/library/ws-mqtt/index.html>.
- [37] Nvidia Developer, "Nvidia Jetson TX2 Module," Nvidia Embedded Computing, [online] Available: <https://developer.nvidia.com/embedded/buy/jetson-tx2>
- [38] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, and K. Murphy, "Speed/accuracy trade-offs for modern convolutional object detectors," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 3296-3305.
- [39] Chardch and Pkulze, "Tensorflow Object Detection API," GitHub. [online] Available: https://github.com/tensorflow/models/tree/master/research/object_detector
- [40] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, Jun. 2017.
- [41] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception architecture for computer vision," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818-2826.
- [42] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1-9.
- [43] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 2015, pp. 1-9.
- [44] ImageNet: Large Scale Visual Recognition Challenge 2012 (ILSVRC2012), held in conjunction with PASCAL Visual Object Classes Challenge 2012 (VOC2012), [online] Available: <http://www.image-net.org/challenges/LSVRC/2012/index>
- [45] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [46] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1026-1034.



Wan-Jung Chang (M'16) obtained a B.S. in electronic engineering from Southern Taiwan University of Science and Technology, Tainan, Taiwan, R.O.C., in 2000; an M.S. in computer science and information engineering from National Taipei University of Technology, Taipei, Taiwan, R.O.C., in 2003; and a Ph.D. in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, R.O.C., in 2008. He is currently an Associate Professor in the Department of Electronic Engineering at Southern Taiwan University of Science and Technology, Tainan, Taiwan, R.O.C. He is also a director of the Artificial Intelligence over the Internet of Things Applied Research Center (AIoT Center) and the Internet of Things Laboratory (IoT Lab) at Southern Taiwan University of Science and Technology, Tainan, Taiwan, R.O.C. He received the Best Paper Award at IEEE ChinaCom '09, the Best Paper Award at ICCPE '16, the 1st Prize Excellent Demo! Award at IEEE GCCE '16/GCCE '17/GCCE '18, and the First Prize Best Paper Award at IEEE ICASI '17. His research interests include cloud/IoT/AIoT systems and applications, protocols for heterogeneous networks, and WSN/high-speed network design and analysis. He is a member of IEEE and the IEEE Consumer Electronics Society.



Liang-Bi Chen (S'04-M'10-SM'16) received B.S. and M.S. degrees in electronic engineering from the National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan, in 2001 and 2003, respectively. From 2004 to 2010, he was enrolled in the computer science and engineering Ph.D. program at the National Sun Yat-Sen University, Kaohsiung, Taiwan. He is currently pursuing a Ph.D. degree in electronic engineering at Southern Taiwan University of Science and Technology, Tainan, Taiwan. From August to September 2008, he had an internship with the Department of Computer Science, National University of Singapore, Singapore. He was also a Visiting Researcher with the Department of Computer Science, University of California, Irvine, CA, USA, from September 2008 to August 2009 and with the Department of Computer Science and Engineering, Waseda University, Tokyo, Japan, from July to August 2010. In May 2012, he joined BXB Electronics Co., Ltd., Kaohsiung, Taiwan, as an R&D Engineer. In May 2013, he was transferred to the position of Executive Assistant to the Vice President, which he held until September 2016. In October 2016, he joined Southern Taiwan University of Science and Technology, Tainan, Taiwan, as a senior technical manager and a part-time lecturer. He has also served as a TPC member, an IPC member, and a reviewer for many IEEE/ACM international conferences and journals. He was a recipient of the 2014 IEEE Education Society Student Leadership Award. He received the First Prize Best Paper Award at IEEE ICASI '17, the Honorable Mention Award at ICAAI '16, the 1st Prize Excellent Demo! Award at IEEE GCCE '16/GCCE '17/GCCE '18, and the 2nd Prize Excellent Poster Award at IEEE GCCE

'16. He is a senior member of IEEE and a member of the IEICE and the PMI.



Chia-Hao Hsu obtained a B.S. degree in electronic engineering from Southern Taiwan University of Science and Technology, Tainan, Taiwan, R.O.C., in 2017. He is currently pursuing an M.S. degree in electronic engineering from Southern Taiwan University of Science and Technology, Tainan, Taiwan, R.O.C.

He received the First Place Award in the Healthcare Applications group at the 13th Digital Signal Processing Creative Design Contest, Ministry of Education (MoE), Taiwan, R.O.C. His current research interests include deep learning and computer vision for the IoT and big data.



Cheng-Pei Lin is currently pursuing a B.S. degree in electronic engineering from Southern Taiwan University of Science and Technology, Tainan, Taiwan, R.O.C. He received the First Place Award in the Healthcare Applications group at the 13th Digital Signal Processing Creative Design Contest, Ministry of Education (MoE),

Taiwan, R.O.C. His current research interests include cloud-based database implementation, web-based platform development, and IoT-based communications protocol design.



Tzu-Chin Yang is currently pursuing a B.S. degree in electronic engineering from Southern Taiwan University of Science and Technology, Tainan, Taiwan, R.O.C. He received the First Place Award in the Healthcare Applications group at the 13th Digital Signal Processing Creative Design Contest, Ministry of Education (MoE),

Taiwan, R.O.C. His current research interests include AIoT systems implementation, IoT-based communications protocol design, and mobile device app development.