Real-Time Pill Identification with Prescription Confirmation Using Deep Learning on Embedded System

Roy Chaoming Hsu
Electrical Engineering Department,
National Chiayi University
Chiayi City, Taiwan
rchsu@mail.ncyu.edu.tw

Po-Cheng Su
Electrical Engineering Department,
National Chiayi University
Chiayi City, Taiwan
bsses970128@gmail.com

***Abstract*—In the study, a real-time pill identification with prescription confirmation within medical packaging using deep learning on embedded system is proposed. In this system, multiple pills within a medical packaging for inpatients can be identified using deep learning and confirmed with the doctor's prescription in QR code to avoid error in the packaging process. The system is implemented in an embedded system with GPU for accelerating the identification and confirmation processes. Experimental results of the proposed system using Nvidia's Jetson Nano shown that multiple pills with different color and shapes can be identified and confirmed with doctor's prescription in real time.**

Keywords*—*pill identification, deep learning, embedded system, prescription confirmation

I. Introduction

With the rapidly increasing of human longevity and elder population, the number of in-patients within a hospital requiring caring and monitoring also increases. Incorrect medication, inaccurate dosage, and misplaced pill in a medical packaging often lead to the crisis of safety for in-patients. Health care and monitoring systems have evolved rapidly over the past two decades and have the potential to change the way healthcare is currently delivered. smart health monitoring systems automate patient monitoring tasks, thereby improving the management for patients health and safety [1]. In the case of increased medical demand, the quality of medical care is affected. In order to improve the quality of medical care and establish smart medical care, information and communication technology and systems have been introduced to establish smart medical services. Through medical big data and artificial intelligence, precision medicine can be promoted, medical efficiency and medical quality can be improved, and the goal of intelligent medicine can hence be achieved.

When a medication error results in harmful to a patient is called an adverse drug event (ADE), and ADEs associated with medication errors are preventable. Those errors were most likely to occur during the prescribing (56%) and administration (34%) phases in a hospital setting, where there is often no one between the nurse (or someone else administering the drug) and the patient to detect or predict the error [2]. The similarity of drug appearance or drug name leads to insufficient identification, and the harm of giving the wrong drug to the patient. Look-alike and sound-alike (LASA), in particular, is the major error at the level of pharmacists or physicians. It will put the patients safety in danger when the LASA occurs. Human eyes recognize medicines as well as artificial and machine-packed medicines may produce certain errors in many cases. Examples include long hours at work and high workload of pharmacists, or mis-packaging of medications due to other factors. Wrong medications can be harmful to the patient's safety and in some cases can cause injury or even death to the patient. Therefore, there is a need for a definitively correct drug that is precisely identified for the patient.

Drug errors were the most common cause of adverse events. The cost of ADE is $2 billion per year in the US. Electronic prescription systems can prevent serious doctor order errors by up to 55%, and bar-codes can reduce dispensing errors in hospital pharmacies, for both techniques are well developed, cost-effective, and have low variability. The advanced QR code technique can also be alternatively used with cost-effective, yet the resources of individual healthcare institutions are still limited in employing QR-code [3].

In the past, traditional image recognition method needed to be specially designed for a specific image, which took a lot of time. Thanks to the vigorous development of deep learning method, as long as the algorithm is designed, a large number of pictures are given to the computer to learn by itself such that the solution can be obtained. Therefore, this study uses Jetson Nano of Nvidia for running deep learning based identification and precise confirmation for drug dispensing. The detailed description of the proposed pill identification and conformation system is described in the following.

II. Literatur Review

Traditional image recognition methods often process a single type of image, and it takes times to train an algorithm over images. Thanks to the vigorous development of deep learning techniques and the strong computation power of GPU, a large number of images can be used to fast train a deep learning algorithm and the computer can learn by itself in obtaining an solution for image recognition and detection task. LeCun, Y. et al. [4] applied Convolutional Neural Networks (CNN) to Graph Transformer Networks (GTN) for handwriting recognition, and compare the various partial blocks, which are called features, in the two pictures. The CNNs can compare two images for similarities by comparing general features in similar and relative locations. The YOLO series (You only look once, Yolo) [5] is a neural network-like algorithm about object detection. It is proposed by Joseph Redmon and implemented with a niche architecture darknet. YOLO did not use any famous deep learning framework, yet with its light weight, less dependencies, high algorithm efficiency, it is very valuable in industrial applications, such as pedestrian detection, industrial image detection and so on. Dinh, D. L. et al. [6] used YOLO and SSD algorithms running on Nvidia Jetson Nano and Google

Coral Dev Board embedded systems for vehicle tracking and counting. In their method, the observation area, i.e., region-of-observation (ROO) is set to reduce the amount of input data such that the clarity of the image can be provided to identify the vehicle, and the performance of the embedded system can also be improved at the same time. Because the original resolution of the entire 4K image is too large, when it is compressed into the 608x608 size that YOLO can use, the algorithm will not be able to identify objects in an image, because the details of the image will be distorted. Hence, Růžička, V. et al. [7] cut and adjust the size of the original image, called cropping, to improve the prediction in image recognition. YOLOv3 is the last version of the YOLO method proposed by Redmon, J. [8]. The YOLOv3 introduces the pyramid network, and uses the better basic network architecture, named darknet-53, which has two classifications of the loss function such that recognition accuracy and the ability to detect smaller objects can be guaranteed. Because YOLOv3 uses information fusion and does not make full use of low-level information, this weakness limits the development of industrial applications. Hence, Peng, et al. [9] proposed the YOLO-Inception method, which uses a variety of initial structures to provide richer semantic information and improve the performance of small object detection. Tian proposed the YOLOv3-dense method [10]. It uses the DenseNet method to process low-resolution feature layers, effectively enhances feature propagation, improves feature reuse, and improves network performance. Alexey, et al. [11] proposed the YOLOv4, which uses the CSPDarknet53 as backbone, the spatial pyramid pooling model, and the PANet path aggregation in combining with the YOLOv3 as the YOLOv4 structure. Chang, W. J. et al. [12] implemented Fast-RCNN combined with Inception V3 on NVIDIA Jetson TX2 to identify drugs and combined with mobile phone APP to scan QR code for drug comparison and inspection [12]. In this study, a pill identification and confirmation system for medical packaging that can identify the medicines in the medicine pack and facilitate the verification of the medicines by scanning the QR Code to in the medicine pack in comparison with the doctor's prescription. The detailed mechanism of the system is described below.

III. The Proposed Pill Identification and Verification Sysytem Flow

In this proposed pill identification and confirmation system, an embedded system with a camera is used to read image of medical packaging. YOLOv4 is used in the embedded system to perform pill identification by extracting features from pill pictures of different shapes and colors, locating pill coordinates, and then identifying pill names in obtaining pill information. The QR code printed on the medical packaging will also be read by the camera of the embedded system and the pill information of the medical packaging will be extracted from the QR code and is compared with the identified pill to confirm whether they are consistent or not. If they are not consistent, the name of the wrong pill within the package will be obtained and the result is output with visual or audio warning effect.

A. System Flow

System flow of the proposed pill identification and confirmation system is shown in Figure 1 below and each step is explained in the following..

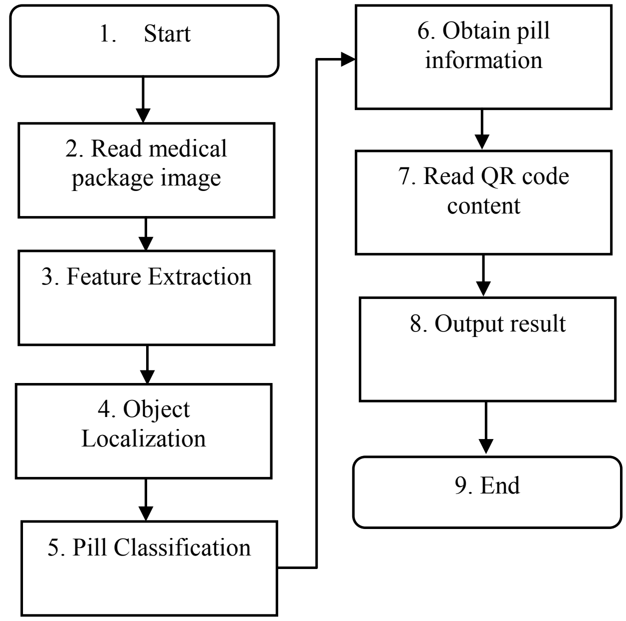


Figure 1 System flow of the proposed pill identification and confirmation system.

1. Start the pill identification program to identify the name and coordinates of the pill(s) in the medical packaging.
2. Read in the videos to identify pills by YOLOv4.
3. Obtain pill characteristics, shape, and color.
4. Locate pill coordinates, including x, y coordinate positions
5. Identify the pill name and compare it with the pill information encoded in QR code
6. Display the obtained information on the screen for manual interpretation
7. Read the QR code information, including the name of the pill prescribed by the doctor
8. Compare pill identification results with the pill information in the QR code for confirmation and output the differences with audio warning/visual alarm.
9. End of pill identification program

B. Training Procedure of the YOLOv4

The training procedure of the YOLOv4 is shown in Figure 2. Description of each step is in the following.

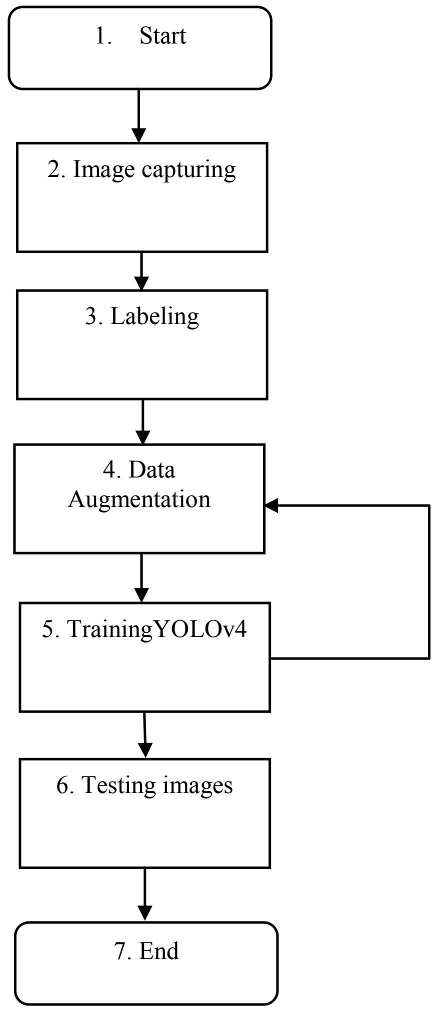


Figure 2 Training procedures for the YOLOv4

1. Enable Camera of the embedded system to capture images
2. Use the camera module to take pictures and send the pictures to the deep learning training machine
3. Label each pill image and enter the pill name
4. Rotating 0~359 degrees of the pill image for augmentation of the pill images
5. Use deep learning to train the machine and YOLOv4
6. Test YOLOv4 training results
7. End of training

In the training of YOLOv4, a photo of the pill image will be first taken and labelled with its respective name. The pill image will be rotate from 0 degree to 359 degree and the photo be taken as augmentation to increase the numbers of pill images, which is about 500 photos per pill for YOLOv4 training, and for testing the training effect of YOLOv4.

IV. Experimental Results

The dataset used and the configuration of the experiment is described in the following.

A. Medical Pill Dataset

Figure 3 shows the medical pill dataset used in this study. Totally 20 kind of pills are used for the pill identification and confirmation experiment. The original picture of the pill is provided by the hospital in association. Each pill image is consisted of the front and back of the pill, as well as the size with blue background. There are pills of various shapes, such as capsules, round, transparent, pentagonal and other irregular shapes, and various colors in this dataset.

Figure 3 The dataset used in this experiment.

Figure 4 shows each pairs of pill with the background removed by pre-processing of the pill images. The blue background is removed after the image pre-processing of the pill image. The dataset in Figure 4 is used to generated the training set and test data.

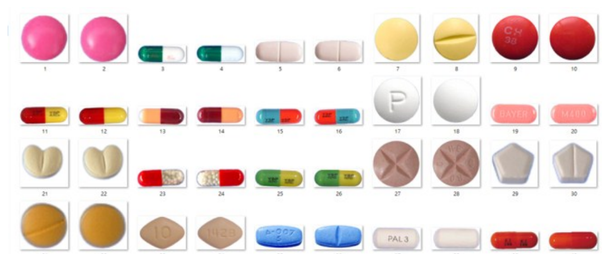


Figure 4 Each pair of pills after removing the background.

B. Configuration of the experiment

The hardware and software of the deep learning server used for this experiment is shown in Table 1.

Table 1 Configuration of the experiment

Deep Learning Server	**Specification**
CPU	Xeon W2245
GPU	NVIDIA 3080
RAM	16GB
Software in the	**Description**
CUDA 11.1	allows software to use certain types of graphics processing units (GPUs) for general purpose processing
cuDNN 7.6	Deep neural network & GPU-accelerated library

Tensorflow-gpu: 2.3.1	Machine learning library
Python: 3.6.12	Development program

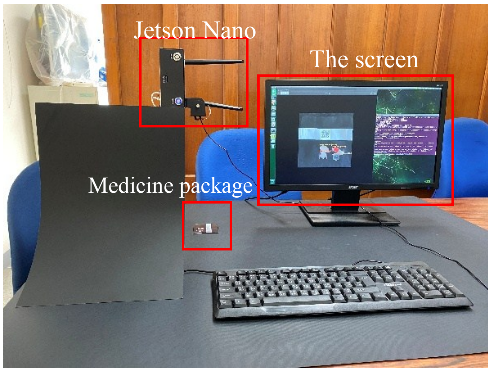


Figure 5 Experimental setup for the training and testing

Figure 5 shows the Jetson Nano from Nvidia® connected with a Sony IM219 (8 million pixel CSI Camera) for taking the pill images is mounted in the upper left. The monitor is connect to the Jetson Nano for displaying the content from the Jetson Nano. The medicine package is in the center of the Figure 5. The training parameter for the YOLOv4 is shown in Table 2.

Table 2 Parameters for the YOLOv4

items	parameters
Image size	608x608
Validation ratio	10%
Momentum	0.949
Decay	0.0005
Backbone	CSPDarkNet53
angle	1
Saturation	1.5
Exposure	1.5
Hue	0.1
Learning rate	0.001

Since deep learning requires a lot of data for effective training. The pill in Figure 4 are rotated with 0~359 degree to increase the number of training set. Figure 6 shows the rotated pills with photos taken by the Sony IM219 camera.

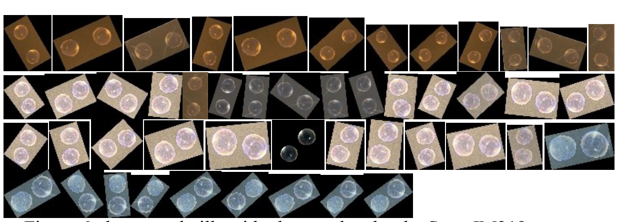


Figure 6 the rotated pills with photos taken by the Sony IM219 camera.

C. Training and Testing Results

The pill images used for training and identification include capsules, circles or ellipses, and various shapes, and the colors include white, red, transparent and various colors. The training image contains only one pill each, while the medical packaging for pill identification contain multiple pills. To test the pill identification and confirmation system, a QR Code is generated by encoding the pill information and added to the synthesized medical packaging image with pill filled in as shown in Figure 7 below. The pills in the medical packaging is consisted with the pill information encoded in the QR code.

Figure 7 A synthesized medical packaging with QR code

Figure 8 shows the loss of the training with respect to the training times. It can be seen that the loss decreases as the number of training times increases, and the red line represents mean average precision (mAP), which is the average value of the average precision (AP) of each object after calculation.

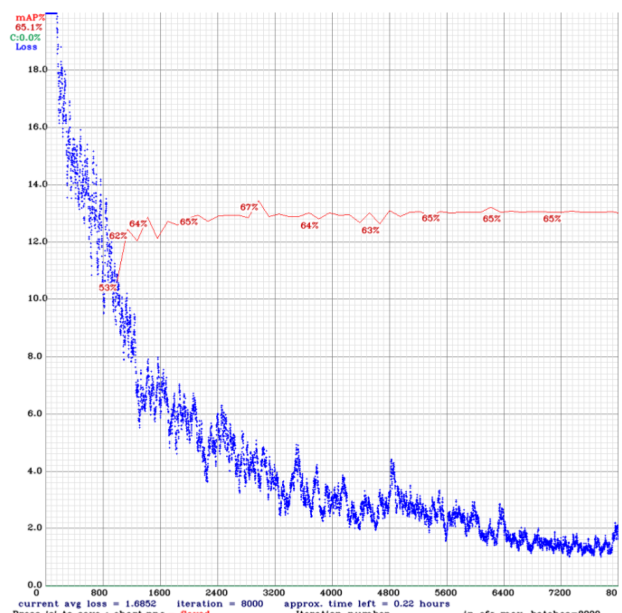


Figure 8 The loss of the training with respect to the training times.

Figure 9 shows the identification result of the medical packaging of Figure 7. The recognition results of the Jetson Nano with YOLOv4 contain confidence and location information of the 4 pills, which are Bensau_100mg_cap

(79.05%), Voren_cap (98.98%), Biocal_Plus_tab (95.58%), and Bonamin (99.28%).

Figure 9 The identification result of the medical packaging of Figure 7

The comparison between the identification and the QR code information is show in Figure 10, where the identified pill names are the same as QR code information shown in the last three lines from the buttom as Bensau_100mg_cap,Voren_cap,Biocal_Plus_tab,Bonamin (x,y)[(466,355),(610,355),(609,498),(466,499)].

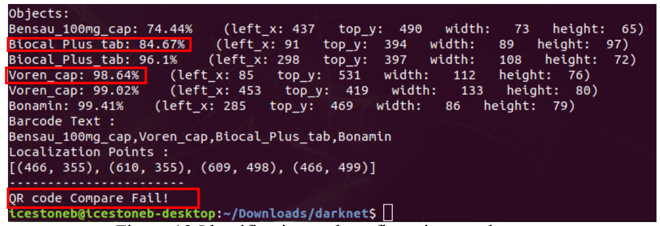
```
Objects:
Bensau_100mg_cap: 79.05%    (left_x: 437   top_y:  490   width:   71   height:  63)
Biocal_Plus_tab: 95.58%    (left_x: 298   top_y:  396   width:   110   height:  72)
Voren_cap: 98.98%    (left_x: 453   top_y:  420   width:   133   height:  80)
Bonamin: 99.28%    (left_x: 285   top_y:  470   width:   84   height:  79)
4.137749910354614
Barcode Text :
Bensau_100mg_cap,Voren_cap,Biocal_Plus_tab,Bonamin
Localization Points :
[(466, 355), (610, 355), (609, 498), (466, 499)]
icestoneb@icestoneb-desktop:~/Downloads/darknet$
```

Figure 10 The comparison between the identified pills and the QR code information

In order to test the confirmation function of this system, two pills are pre-assumed mistakenly inserted into the medical packaging, which are shown in Figure 11.

Figure 11 An assumed mistakenly filled medical packaging.

Figure 11 also shows that the YOLOv4 identification results contain pill confidence and location information, which are Bensau_100mg_cap (74.44%), Biocal_Plus_tab (84.67%), Biocal_Plus_tab (96.1%), Voren_cap (98.64%), Voren_cap (99.02%), Bonamin (99.41%) %). Because the medical packaging contained two extra pills which is assumed mistakenly inserted and the confirmation result obtained by comparine the information in the QR code and the identified pills by the YOLOv4 is inconsistent as show in the Figure 12.



```
Objects:
Bensau_100mg_cap: 74.44%    (left_x: 437   top_y:  490   width:   73   height:  65)
Biocal_Plus_tab: 84.67%    (left_x: 91   top_y:  394   width:   89   height:  97)
Biocal_Plus_tab: 96.1%    (left_x: 298   top_y:  397   width:   108   height:  72)
Voren_cap: 98.64%    (left_x: 85   top_y:  531   width:   112   height:  76)
Voren_cap: 99.02%    (left_x: 453   top_y:  419   width:   133   height:  80)
Bonamin: 99.41%    (left_x: 285   top_y:  469   width:   86   height:  79)
Barcode Text :
Bensau_100mg_cap,Voren_cap,Biocal_Plus_tab,Bonamin
Localization Points :
[(466, 355), (610, 355), (609, 498), (466, 499)]
------------------------
QR code Compare Fail!
icestoneb@icestoneb-desktop:~/Downloads/darknet$
```

Figure 12 Identification and confirmation results.

In Figure 12, the identification result shows there are 6 pills in the medical packaging, however, the information from the QR code contains only 4 pills. The confirmation result comes out with the “QR code compare Fail!” along with a warning sound.

V. Conclusion

In this study, a real-time pill identification and confirmation system for pill in a medicine package is proposed. This system can identify the pills in the medicine pack and compare them with the prescription information encoded in the QR code. An embedded system with GPU, Jetson Nano from Nvidia®, is used to execute the YOLOv4 for identifying the pills in an medical packaging. And the result is compared with Doctor’s order encoded in the QR code to verify the identification is the same as doctor’s order. Experimental results show that the proposed system not only can detect and identify the pill is the medicine package, but it also can verify the Doctor’s order encoded in the QR code for confirmation purpose.

References.

[1] Baig, M. M., & Gholamhosseini, H. (2013). Smart health monitoring systems: an overview of design and modeling. *Journal of medical systems*, 37(2), 1-14.

[2] Bates, D. W., & Slight, S. P. (2014, August). Medication errors: what is their impact?. In M*ayo Clinic Proceedings*, Vol. 89, No. 8, pp. 1027-1029. Elsevier.

[3] Maviglia, S. M., Yoo, J. Y., Franz, C., Featherstone, E., Churchill, W., Bates, D. W., ... & Poon, E. G. (2007). Cost-benefit analysis of a hospital pharmacy bar code solution. *Archives of internal medicine*, 167(8), 788-794.

[4] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.

[5] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[6] Dinh, D. L., Nguyen, H. N., Thai, H. T., & Le, K. H. (2021). Towards AI-Based Traffic Counting System with Edge Computing. *Journal of Advanced Transportation*, 2021.

[7] Růžička, V., & Franchetti, F. (2018, September). Fast and accurate object detection in high resolution 4K and 8K video using GPUs. In 2018 IEEE High Performance extreme Computing Conference (HPEC) (pp. 1-7). IEEE.

[8] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

[9] P. Du, X. Qu, T. Wei, C. Peng, X. Zhong and C. Chen, "Research on Small Size Object Detection in Complex Background," 2018 Chinese Automation Congress (CAC), 2018, pp. 4216-4220, doi: 10.1109/CAC.2018.8623078.

[10] Tian, Y.; Yang, G.; Wang, Z.; Wang, H.; Li, E.; Liang, Z. Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput Electron Agr*, 2019,157:417-426.

[11] Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv 2020, arXiv:2004.10934.

[12] Ting, H. W., Chung, S. L., Chen, C. F., Chiu, H. Y., & Hsieh, Y. W. (2020). A drug identification model developed using deep learning technologies: experience of a medical center in Taiwan. *BMC Health Services Research*, 20(1), 1-9.