

Received March 27, 2020, accepted April 19, 2020, date of publication April 22, 2020, date of current version May 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2989673

# Automated Firearm Classification From Bullet Markings Using Deep Learning

PATTRANIT PISANTANAROJ<sup>1,2</sup>, PIMLAPUS TANPISUTH<sup>1,3</sup>, PIYAWUT SINCHAVANWAT<sup>1,2</sup>, SIRIPORN PHASUK<sup>1,2</sup>, (Member, IEEE), PHONGPHAN PHIENPHANICH<sup>1,2</sup>, (Member, IEEE), PARINTON JANGTAWEE<sup>4</sup>, KITTISAK YAKOOMPAI<sup>5</sup>, MONTRI DONPHOONGPI<sup>6</sup>, SANONG EKGASIT<sup>4</sup>, AND CHARTURONG TANTIBUNDHIT<sup>1,2</sup>, (Member, IEEE)

<sup>1</sup>Center of Excellence in Intelligent Informatics, Speech and Language Technology, and Service Innovation (CILS), Thammasat University, Rangsit Campus, Pathum Thani 12120, Thailand

<sup>2</sup>Department of Electrical and Computer Engineering, Faculty of Engineering, Thammasat University, Rangsit Campus, Pathum Thani 12120, Thailand

<sup>3</sup>Ruamrudee International School, Bangkok 10510, Thailand

<sup>4</sup>Sensor Research Unit, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

<sup>5</sup>Firearms and Ammunition Subdivision, Central Police Forensic Science Division, Bangkok 10830, Thailand

<sup>6</sup>Professional Level Firearm and Physical Evidence Examination Section, Ministry of Justice, Bangkok 10210, Thailand

Corresponding author: Charturong Tantibundhit (tchartur@engr.tu.ac.th)

**ABSTRACT** Firearm violence is one of the leading causes of death in many countries around the world, including Thailand. This work proposes a fast and accurate automated method to classify firearm brands from bullet markings. Specifically, a panoramic image of a bullet collected from a crime scene was captured using a developed mobile phone application and custom-built portable hardware. The top three state-of-the-art CNNs pretrained on ImageNet—DenseNet121, ResNet50, and Xception—were further trained on the same training set, which was composed of 718 bullets collected from eight different firearm brands—Beretta, Browning, CZ, Glock, Norinco, Ruger, Sig Sauer, and Smith & Wesson—using a five-fold cross validation technique. DenseNet121 provided the highest AUC of 0.99 for CZ classification (the most common registered firearm brand in Thailand) and the highest average AUC for the eight firearm brands ( $0.9780 \pm 0.0130$  SD), which was significantly higher than those of ResNet50 and Xception. In addition, there were no interaction effects between the CNN model and firearm brand on AUC. DenseNet121, which had the highest AUC, was evaluated on the test set (72 bullets), and the results showed that the Beretta and CZ classifications had the lowest accuracy (91.18%), followed by the Browning and Norinco classifications (96.88%), whereas the Glock, Ruger, Sig Sauer, and Smith & Wesson classifications had the highest accuracy (98.41%). These results suggest that the developed mobile phone application based on a deep learning algorithm and the custom-built portable hardware have promising potential for use at crime scenes to classify firearms from bullet markings. By narrowing down the list of suspects, this convenient approach can potentially accelerate bullet identification processes for many forensic science examiners.


**INDEX TERMS** Forensic science, automated firearm classification, 9 mm bullet marking, densely connected convolutional network.

## I. INTRODUCTION

The right to bear arms is one of the most contentious issues worldwide. In 2016, there were 250,000 deaths worldwide as a result of firearm attacks, making firearm violence one of the leading causes of death in many countries, including Thailand. In 2017, Thailand had a gun violence rate of 3.71 victims in every 100,000 people; many of these events occurred in southern Thailand [1]. Thailand has the 19<sup>th</sup> highest gun

violence rate in the world. Moreover, among the countries in East Asia, Southeast Asia, and Australasia, Thailand has the second highest firearm violence rate after the Republic of the Philippines. In fact, Thailand has a higher gun violence rate than other countries with a notoriety for violence, such as Iraq, which has a gun violence rate of 3.54 victims in every 100,000 people [2].

Apart from the high number of registered firearms, the high rate of gun violence in Thailand is predominantly a consequence of the prolonged process of firearm identification. There are only two laboratories in Thailand: the Central

The associate editor coordinating the review of this manuscript and approving it for publication was Kim-Kwang Raymond Choo .

Institute of Forensic Sciences (CIFS) and the Office of Police Forensic Sciences. Most investigations can take up to 30 days, not including the number of days taken to collect and transport physical evidence. In consideration of the death rates attributed to firearm violence, forensic investigation processes should be improved and accelerated to help solve current crimes and prevent future crimes.

### A. BACKGROUND

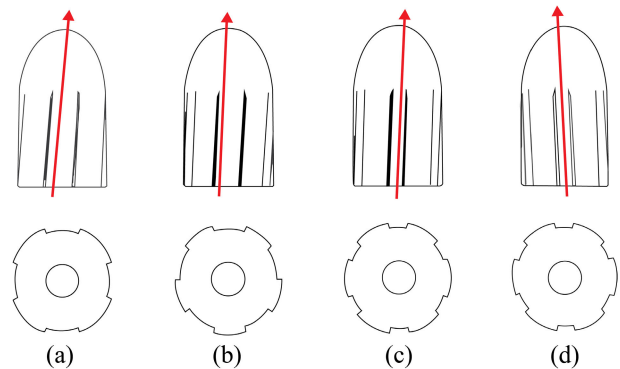
Currently, forensic investigation processes are dependent on the physical evidence collected from crime scenes [3]; however, some types of evidence are easily destroyed. These types of easily destroyed physical evidence include fingerprints, blood, footwear impressions, tire tracks, and gunshot residue [4]. Criminals can carry out precautionary acts to avoid leaving clues to the crime, their connection to the crime, or their identity, which can confuse and interfere with crime investigation or forensic science efforts [5]. Since discharged bullets are more difficult to destroy than most physical evidence, bullets provide forensic scientists essential information about the identities of criminals [6].

When a bullet is fired, it is ignited within the chamber and travels through the barrel of the firearm. Since bullets are intentionally designed to be wider than the barrel of the firearm, as the bullet passes through the barrel, the barrel compresses, thereby leaving toolmarks on the bullet [7], [8]. These toolmarks become a “ballistic fingerprint” that examiners can use to identify specific characteristics of the firearm that discharged the bullet [9].

In semiautomatic firearms, apart from the discharged bullets, an empty cartridge case is also ejected. Similar to the phenomenon observed in bullets, the ejection process leaves distinct marks on the empty cartridge case, making the cartridge case another “ballistic fingerprint” that can be used to identify the firearm [10], [11]. However, cartridge identification complicates investigation processes, since the identification of one empty cartridge case from numerous cartridge cases in a crime scene does not necessarily suggest that it was expelled from the firearm used to shoot the victim. Hence, in this paper, the classification of firearms will be made by a collection of data from bullet markings.

The toolmarks imprinted on bullets consist of various striations. The width, depth, and pitch of the striations depend on the material of the bullet [12]. These toolmarks are created by the barrel as the bullet passes through it, and there are four types of striations: (a) 4-groove right-hand twist, (b) 5-groove right-hand twist, (c) 6-groove right-hand twist, and (d) 6-groove left-hand twist, as shown in Fig. 1. When a bullet is fired, nitrocellulose-based propellants increase the temperature and pressure of the bullet, causing the bullet to move through the barrel at a high velocity [13]. This action creates the striations on the bullet.

Rifling in barrels is dependent on two characteristics: class characteristics and individual characteristics [14]. The class characteristics of a bullet are the rifling specifications of the barrel, including the caliber, number, direction of twist, and



**FIGURE 1.** Side views and cross-section of fired bullet with (a) 4-groove right-hand twist, (b) 5-groove right-hand twist, (c) 6-groove right-hand twist, and (d) 6-groove left-hand twist (modified from Fig. 1 in [12]).

**TABLE 1.** Class characteristics of eight commonly registered firearms in Thailand.

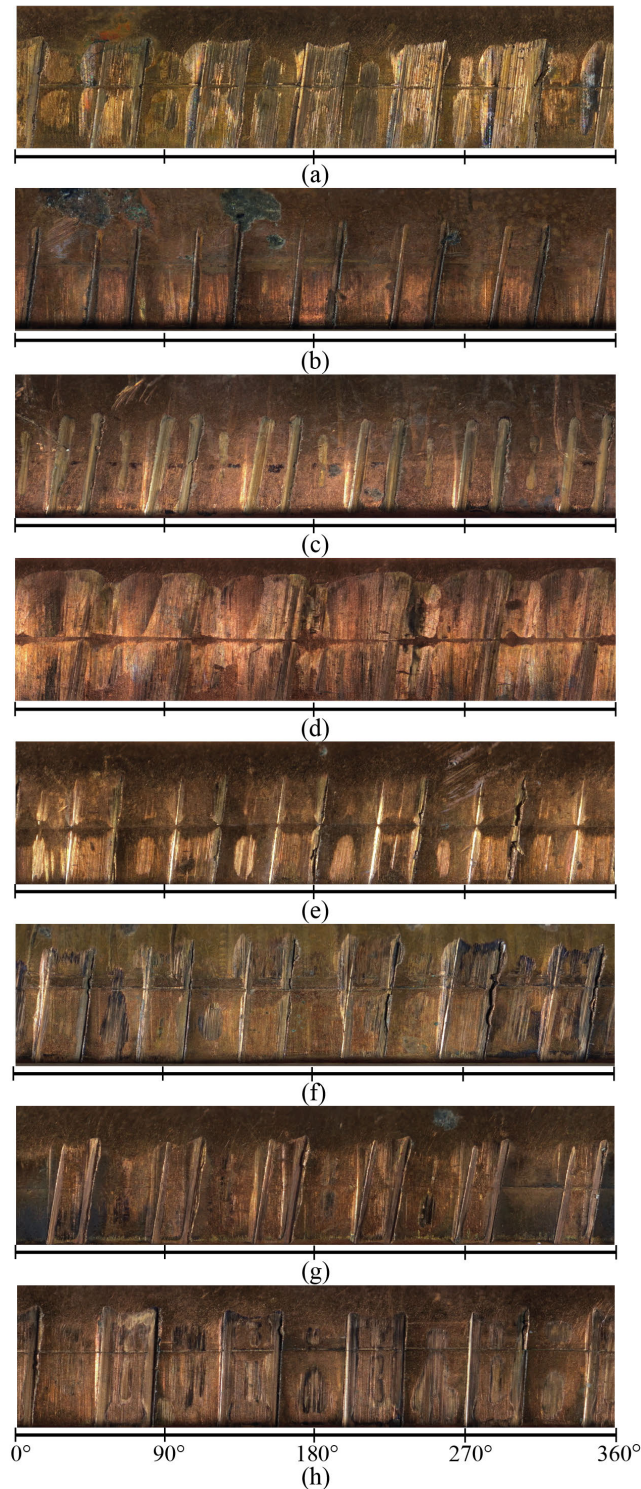
Twist Number	Twist Side	Twist Characteristics	Brand
5	Right	Convolutional	Smith & Wesson
6	Right	Convolutional	Beretta, Browning, CZ, Norinco, Ruger, and Sig Sauer
6	Right	Polygonal	Glock

widths of the lands and grooves [8], [14]. Table 1 summarizes class characteristics of eight commonly registered firearms in Thailand. On the other hand, the individual characteristics are the random imperfections in the barrels that create the toolmarks on the bullets [14]. The class and individual characteristics of all the firearm brands used in this paper are shown in Fig. 2. Individual characteristics are more specific and are normally used for bullet identification; however, the individual characteristics of firearms can be switched by exchanging the barrel with that from another firearm from the same manufacturer. For this reason, analyzing the individual characteristics of a discharged bullet does not necessarily indicate the identity of the criminal. Hence, class characteristics can prevent incorrect associations with unrelated suspects.

### B. RELATED WORK

To date, many traditional bullet identification methods have been developed to help identify unknown firearms, including microscopic detection [8], [15], [16], the continuous shooting method [17], and roughness measurements with a stylus [18]. These identification methods are time consuming and are likely to be subjective [19]–[21]. Although there are accepted methodologies that allow qualitative comparisons of bullets [22], there has not been much quantitative proof to make morphological deductions. Furthermore, professional examiners must be present during the process for accurate conclusions to be drawn.

Since the abovementioned identification methods are time consuming and have certain limitations, many forensic science examiners have used integrated ballistics identification



**FIGURE 2.** Markings on bullets fired by eight commonly registered firearms in Thailand: (a) Beretta, (b) Browning, (c) CZ, (d) Glock, (e) Norinco, (f) Ruger, (g) Sig Sauer, and (h) Smith & Wesson.

systems (IBIS Heritage) [19] to digitally compare images of ballistic evidence to real bullet markings [19], [20]. However, using the IBIS Heritage can lead to loss of information because the correspondence and the surface topography are not one to one, possibly reducing the accuracy of the

conclusions. Furthermore, the list of firearm rankings generated by the IBIS Heritage is limited to the firearms that are inside its database [23], [24], thereby requiring a high level of expertise for analysis. Most importantly, the price for adopting the IBIS Heritage is quite high for developing countries including Thailand, as the system costs approximately 98,000,000 Baht (approximately 3,250,000 USD), not including annual maintenance costs. For this reason, the IBIS Heritage is available in only six of the seventy-seven provinces in Thailand.

Most studies have used image processing for bullet identification, but image processing has drawbacks. Chu *et al.* proposed an algorithm to perform correlation calculations to identify 48 bullets fired from six different barrel manufacturers, which were further classified into their class characteristics according to their land-engraved areas [19]. After further applying the correlation function for automatic selection of the effective correlation area and the extraction of a signature bullet profile, the correlation results showed a 9.3% higher accuracy rate than current commercial systems. Although this work improved the ranking of correct matches, the developed software could only distinguish the impression quality along the longitudinal (bullet axis) direction.

Lu *et al.* implemented a 3D laser color scanner (3DLCS) to scan bullets and map a 3D model of the bullet on a 2D plane [17]. They fired nine bullets—three bullets from three different barrels—and further classified the bullets into different groups. The unwrapped images were then preprocessed by various method, including image enhancement, edge detection, binarization, thinning and denoising. Then, different algorithms were applied to extract the individual characteristics of the bullets. Their results showed that there was a high similarity between the bullets fired from the same barrel, one group exceeding 90% similarity [17]. However, their experimental results were based on a few data points, suggesting the possibility of an inaccurate algorithm if implemented in larger data samples.

Xie *et al.* developed a measurement system based on a Talyrond 365 roundness/cylindricity system [25]. Their 3D system consists of a rotary stage that allows the collection of all facets of 3D surface topography information; moreover, this system is composed of a small and shapeable smart stylus sensor that can measure form deviations [18]. In addition to the use of a novel measurement system, feature extraction of bullet marks based on surface topography techniques has been implemented [22], [25]. They extracted class characteristics through surface segmentation, effectively separating the bullets into regions. On the other hand, individual feature extraction was completed through surface abstraction, wavelet filtering, and comparison. After initially identifying the bullets, firearm examiners suggested the best matches out of the whole list [22].

In addition to image processing, recent works have implemented machine learning methods for bullet identification. Petraco and Chan used multivariate statistical analysis and machine learning for toolmark impression pattern recognition

and bullet identification [26]. Striation patterns were viewed as mean profiles, a form of multivariate feature vectors. Along with the use of standard multivariate machine learning methods, these mean profiles were used to estimate identification error rates by using a combination of principal component analysis (PCA), canonical variate analysis (CVA), and support vector machines (SVMs). Their experimental results showed low identification error rate estimates, with a general error rate of 1% with 95% confidence intervals. Many pieces of software were developed for the visualization of toolmark surfaces in the database; however, the analysis of 3D impression patterns and incomplete toolmarks from striation patterns proved to be too complicated for software analysis.

Banno also used a machine learning approach, wherein they developed a neural network to process binary signals obtained from striation images of ten unidentified bullets and ten database bullets [27]. After inputting two signals into the network, the network evaluates the similarity of the signals and produces a score that indicates the similarity of the bullet striations. Even though the neural network was able to correctly pair the unidentified bullets and database bullets, the number of bullets in the dataset used to train the algorithm proposed was small, and a final decision must still be made by a forensic scientist. Changmai applied the  $k$ -nearest neighbor ( $k$ -NN) machine learning algorithm to classify data based on similar features of the sample data [28]. Six real bullets were tested and classified into three classes. The  $k$ -NN algorithm can correctly classify 86.67% of firearms on average.

Although previous studies have used image processing and machine learning to identify bullets from bullet markings, the corresponding feature selections are time consuming and require a high level of expertise to accurately compare the bullets. In this paper, we have incorporated the implementation of deep learning for firearm classification, since deep learning has resulted in significant improvements in classifications such as object detection, face recognition, and speech recognition [29]–[32]. To the best of our knowledge, deep learning has never been used for bullet identification or firearm classification. Therefore, in this paper, we will develop a deep learning algorithm to classify panoramic images of markings on 9 mm bullets (the most commonly used bullet size in Thailand) collected from crime scenes into eight classes: Beretta, Browning, CZ, Glock, Norinco, Ruger, Sig Sauer, and Smith & Wesson (the most commonly used firearm brands in Thailand). Table 2 summarizes the advantages and disadvantages of previous and our approaches.

First, we developed portable hardware and used a mobile phone to record an immersive video from which a panoramic image of the bullet markings can be obtained. Second, all of the data from the discharged 9 mm bullets were collected at the Firearms and Ammunition Subdivision of the Central Police Forensic Science Division in Thailand, a division that issues firearm-carrying permits and investigates physical evidence from crime scenes. Then, the collected data were pre-processed and augmented prior to deep learning development. These steps will be explained in Section II. The top three

performing deep learning algorithms tested on ImageNet [33] were used in this paper. These algorithms will be summarized in Section III. The hyperparameter settings of these three deep learning algorithms, five-fold cross validation, and evaluation will be explained in Section IV. The experimental results and statistical significance of these algorithms are presented in Section V. The findings of this paper will be concluded and discussed in Section VI. Finally, future work will be discussed in Section VII.

## II. DATA ACQUISITION AND PREPARATION

The cylindrical-based surface of a bullet contains characteristic marks, such as grooves, thread patterns, and microscopic details, made by the specific gun barrel used to fire the bullet. These unique characteristic marks are crucial inputs for bullet identification, so close-up panoramic detail of the bullet surface is required as raw digital data for further computational analysis. The instrument incorporates a bullet rotating mechanism and an illuminating system to capture high-quality images of the bullet markings with a smartphone. Moreover, the aspherical plano-convex lens was attached over the smartphone camera to shorten its focal length; as a result, the smartphone was able to capture close-up particular details on the bullet surface.

### A. HARDWARE DEVELOPMENT

Our developed hardware was mainly made of 10-mm thick polymethyl methacrylate (PMMA), a clear plastic, as shown in Fig. 3 (a)–(c). This hardware was prototyped by laser cutting, 3D printing, and a computer numerical control (CNC) machine, and all the parts were fastened together with bolts and nuts to enhance precision and avoid blurring effects from the evaporated solvent of the adhesive. Two white light LED lamps were placed on both sides of the bullet to illuminate the surface details without contrast generated by grooves and thread patterns. They were set at 30–45 degrees to the horizontal plane (the angle depends on the material, e.g., copper and lead) to obtain the highest intensity and the least reflection on the surface where the lens is focused (along the green line shown in Fig. 3 (a)). Additionally, the illumination system provides stable light to obtain a similar condition of white balance in each captured image.

A 9.0 mm diameter BK-7 glass plano-convex lens was used (Fig. 3 (b)) and embedded into the smartphone case with the planar side facing the smartphone camera; its back focal length was 15.0 mm. Consequently, the distance between the lens and the bullet surface was fixed at 15.0 mm, and the design of the whole instrument was mainly dependent on this parameter. During use, the smartphone is placed on a stationary platform that is parallel to the horizontal plane, where the camera faces the bullet and the smartphone screen faces the user. The user is able to place the smartphone on-off the platform without clamping the smartphone, making this platform compatible for every smartphone model. This design decreases the amount of stray light that can interfere with the white balance of the camera. A driving thread is rotated by

**TABLE 2.** Summary of the advantages and disadvantages of previous and our approaches.

Author	Method	Advantages	Disadvantages
IBIS Heritage [19]	<ul style="list-style-type: none"> <li>- Data acquisition station (DAS)</li> <li>- Signature analysis station (SAS)</li> </ul>	<ul style="list-style-type: none"> <li>- Bullet markings can be digitized and automatically sorted.</li> <li>- Bullet markings can be automatically compared with other bullets previously entered in the system.</li> <li>- Produces the numerical probability for each match.</li> </ul>	<ul style="list-style-type: none"> <li>- Correspondence and surface topography are not one to one.</li> <li>- Does not produce a short-list of matches.</li> <li>- Requires judgment by an examiner.</li> <li>- Expensive (98 million baht per adoption).</li> </ul>
Chu et al. [19]	<ul style="list-style-type: none"> <li>- Striation density</li> </ul>	<ul style="list-style-type: none"> <li>- Has strong sensitivity to image quality except when the edge detector is selected.</li> <li>- Quantitative criterion to assess whether there are sufficient striations for bullet identification by showing the number and distribution of striation marks.</li> <li>- Determines and predicts bullet identifiability by evaluating image quality.</li> </ul>	<ul style="list-style-type: none"> <li>- More parameters can be considered to evaluate the striation marks.</li> <li>- No validation that the proposed methodology can be a standard indication of the reliability of using bullet images for automatic bullet signature correlation.</li> </ul>
Lu et al. [17]	<ul style="list-style-type: none"> <li>- 3D laser color scanner</li> </ul>	<ul style="list-style-type: none"> <li>- Similarity sorting results of data from a dataset has a high similarity (greater than 90%).</li> <li>- Bullet images are in 3D.</li> <li>- Enhances image quality by removing noise.</li> <li>- Based on order and distance of the bullet striation feature, which allows the unique features of the bullet to be extracted.</li> </ul>	<ul style="list-style-type: none"> <li>- Only nine bullets from three different gun barrels are classified.</li> </ul>
Xie et al. [25]	<ul style="list-style-type: none"> <li>- Automated bullet-identification system based on surface topography techniques</li> </ul>	<ul style="list-style-type: none"> <li>- 3D surface topography measurement</li> <li>- Applied advanced surface topography techniques (i.e. surface data segmentation, wavelet filtering, visualization technology) to improve extraction of class and individual characteristics</li> </ul>	<ul style="list-style-type: none"> <li>- Sample of bullets used is too small</li> </ul>

a knob, as shown in Fig. 3 (b), to bring the sliding jaw into close contact with the bullet; as a result, the jaw is able to move along with the motor.

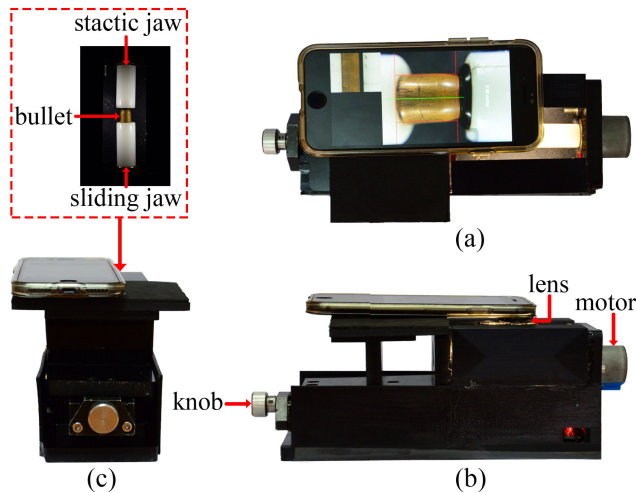
A 28BYJ-48 stepper motor combined with its ULN2003 driver (Amazon, United States) was selected for the bullet rotating system because it operates with 5 V DC, which

**TABLE 2. (Continued) Summary of the advantages and disadvantages of previous and our approaches.**

Author	Method	Advantages	Disadvantages
Petraco and Chan [26]	<ul style="list-style-type: none"> <li>- Application of machine learning to toolmarks</li> </ul>	<ul style="list-style-type: none"> <li>- Mean profiles are used with standard multivariate machine learning methods (i.e. principal component analysis, canonical variate analysis, support vector machine methodology) to estimate identification error rates of striation patterns.</li> <li>- Toolmark identification error rates are approximately 1% with 95% confidence intervals</li> </ul>	<ul style="list-style-type: none"> <li>- The database is small, consisting only of impression patterns on Glock cartridge cases and screwdriver and chisel striation patterns.</li> </ul>
Banno [27]	<ul style="list-style-type: none"> <li>- Neural networks</li> </ul>	<ul style="list-style-type: none"> <li>- Determines the similarity of striations from binary signals.</li> <li>- Able to identify deformed striations.</li> <li>- Shortens time required to find the similarity of bullets because the location of bullets does not need to be extracted from an image of land impression.</li> </ul>	<ul style="list-style-type: none"> <li>- The database is very small, consisting of 10 unidentified bullets that are matched with ten database bullets.</li> <li>- The final decision must still be made by a forensic scientist.</li> </ul>
Ours	<ul style="list-style-type: none"> <li>- Automated firearm classification from bullet markings using DenseNet121</li> </ul>	<ul style="list-style-type: none"> <li>- Larger sample of bullets used to train and test the model.</li> <li>- Can be used for both lead and copper-jacketed bullets.</li> <li>- The mobile application is portable and convenient for use.</li> <li>- Provides real time results, without depending on the judgment of examiners.</li> <li>- Tested and trained state-of-the-art deep learning models that have a small number of layers and high accuracy, including ResNet50, DenseNet121, and Xception.</li> </ul>	<ul style="list-style-type: none"> <li>- Only nine mm bullets can be evaluated.</li> <li>- Can evaluate only 8 commonly used firearm brands in Thailand.</li> </ul>

can be directly controlled by an Arduino UNO microcontroller (Amazon, United States). The motor consists of 4 magnetic stepping poles and a driving gear system, as shown in Fig. 3 (b), which takes 512 steps to complete one full rotation. The speed of the motor was fixed by the delay time (time

taken for the stepper motor to move from the current step to the next step) of 121 milliseconds; hence, one full rotation took 62 seconds. On the other hand, the motor is attached to the “static jaw”. An O-ring is used to connect the motor shaft to a plastic cylindrical jaw, which has a 6 mm inner



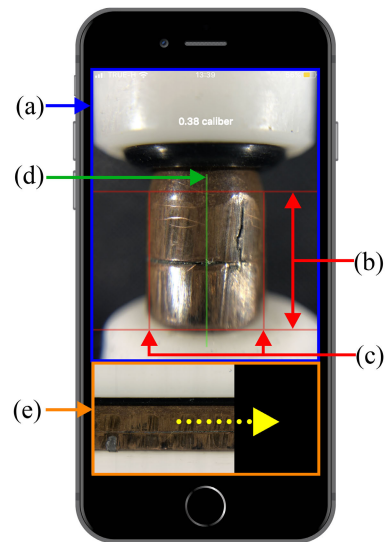
**FIGURE 3.** The developed hardware used to take panoramic images of fired bullets: (a) top view, (b) side view, and (c) front view.

diameter and a 2 mm round thickness at the end. This jaw was designed to fit the bullet head tightly because the bullet has to be rotated along with the motor without slippage; therefore, this jaw is called the “static jaw”, as shown in Fig. 3 (c). The “sliding jaw” is a plastic cylinder, with one flat end designed to push the bottom of the bullet and the other end designed to connect to a 15 mm bearing that is aligned concentrically to the motor shaft, as shown in Fig. 3 (c). The “sliding jaw” can translate along the Y-axis (parallel to the axis of the center of the bullet).

## B. MOBILE APPLICATION

An iPhone8 was used to record the video of the markings on the rotating bullet. The 9 mm bullet rotates at 62 seconds per revolution. Thus, the angular velocity of the rotating bullet is 5.81 degrees per second, and the translational speed at the surface is 0.46 mm per second. Taking into account the shutter speed of the iPhone8 camera, the experimental results indicate that rotation at this speed produces images at the optimal quality. The 80-second long videos cover 30% more than one revolution, which ensures that all the surface details were recorded. The videos were recorded at a resolution of 1,080p with a frame rate of 30 frames per second.

To further aid forensic science examiners, we designed a mobile application to increase the convenience and efficiency of bullet identification. The mobile application can be installed on various smartphones. Using the smartphone camera and the developed hardware, the mobile application can record a video of a bullet specimen. The video recorded has a resolution of 1,080p (columns) at 30 frames per second, meaning that 30 pictures of the rotating bullet specimen are taken per second. This is the standard number of frames per second for most videos captured on a smartphone. Each frame is cropped from  $1,920 \times 1,080$  pixels (1,920 rows  $\times$  1,080 columns) to  $1,920 \times 2$  pixels (1,920 rows  $\times$  2 columns), i.e., one pixel from the middle of



**FIGURE 4.** Illustration of each component of our developed bullet classification on the mobile application: (a) video capture preview, (b) horizontal grids, (c) vertical grids, (d) middle grid, and (e) real-time panoramic image preview.

the frame to the left and to the right, respectively. This is the size that fits with the rotational speed of the bullet specimen.

Figure 4 illustrates each component of our developed bullet classification on the mobile application. The smartphone displays the video of the bullet specimen as it is recorded, as shown in (a). The red grid lines in the mobile application are shown on the smartphone screen, which allows the user to check whether the height of the bullet specimen can be recorded. The top red line is placed below the top of the bullet because the parts that were excluded did not have any useful striations that can be used to identify the firearm brand. Since the mobile application is created to classify 9 mm bullets, the size of the red grid lines can be used to check whether the size of the bullet is appropriate for classification. The user can check whether the height of the bullet specimen can be recorded, as shown in (b). The red line allows the user to check the width of the bullet specimen, as shown in (c). The user can also check if the bullet striations are placed in the middle, which is the correct position for evaluation, as shown in (d). While a video of the bullet specimen is recording, the video is converted into a panoramic image, as shown in (e), where the yellow arrow indicates the direction of the panoramic image being taken. This process was performed to reduce the file size for each bullet specimen, decreasing the time taken to transfer data to obtain the predicted firearm brand. Then, the developed algorithm will analyze the panoramic images of the bullet specimen to identify the firearm brand in 62 seconds.

## C. DATA COLLECTION

The data were collected at the Firearms and Ammunition Subdivision of the Central Police Forensic Science Division in Thailand, a division that issues firearm-carrying

permits and investigates physical evidence collected from crime scenes. The data collected for this paper were bullets acquired from real crime cases, which are presented in Table 3. The table shows that the number of copper bullets is significantly higher than the number of lead bullets. As the bullet passes through the barrel, bullets that are not covered with a copper jacket cannot withstand the increased friction, resulting in fouling [13]. For this reason, most people prefer bullets with a copper jacket in semiautomatic guns. Thus, most of the bullets collected from crime scenes are copper-jacketed bullets.

**TABLE 3. Number of collected copper and lead bullets fired from eight commonly registered firearms in Thailand.**

Brand	Number of collected data		
	Copper	Lead	Total
Beretta	108	8	116
Browning	46	2	48
CZ	196	30	226
Glock	193	23	216
Norinco	65	7	72
Ruger	52	5	57
Sig Sauer	71	8	79
Smith & Wesson	77	7	84
Total	808	90	898

#### D. PREPROCESSING

To process the images of bullet markings, a panoramic image must be obtained; however, not every smartphone can take panoramic images. To solve this problem, we developed a mobile application that allows videos to be taken from smartphones that lack the ability to capture panoramic images. By converting these videos and generating panoramic images in our mobile application, the control points of the sequence of images taken as a subset of a single video can be extracted; this technique is similar to panoramic stitching in various software packages [34]. As shown in Fig. 5, the recorded video was first split into 1,860 frames (62 seconds  $\times$  30 frames), then the images were cropped at the middle of the frame (1,920  $\times$  2 pixels), as shown in the calculation below, and the background images were removed.

#### E. DATA AUGMENTATION

Data augmentation is a technique to increase the size of a dataset to avoid overfitting. The features of the dataset are spiral grooves with a degree of rotation and left-hand/right-hand circular rotation; therefore, a flip or rotational method cannot be implemented. Thus, the data were augmented by implementing the shifting method. The original image, a 360-degree panoramic image of a fired bullet, was shifted by 10 degrees until the same panoramic image as in the beginning was obtained (10  $\times$  36 = 360 degrees). As shown in Fig. 5, the shifting method augmented the dataset from 898 to 32,328 panoramic images in total (4,176 Beretta,

1,728 Browning, 8,136 CZ, 7,776 Glock, 2,592 Norinco, 2,052 Ruger, 2,844 Sig Sauer, and 3,024 Smith & Wesson).

### III. DEEP LEARNING ALGORITHMS

The goal of this section is to investigate the performance of different deep learning algorithms in classifying panoramic images of bullet markings into eight different firearm brands. To accomplish this aim, we selected deep learning algorithms that have one of the smallest sizes and produce one of the most accurate results among the deep learning algorithms using Keras, which were tested on the ImageNet dataset. The three deep learning algorithms with the highest performance and the lowest number of layers were chosen because we would like our developed model to be lightweight yet accurate. This study used three models, residual neural network (ResNet50) [35], densely connected convolutional network (DenseNet121) [36], and Xception [37], which were pre-trained on the ImageNet dataset. Specifically, the pretrained weights were used as the initial weights. Then, these weights were further updated in the training set until the local optimum was reached for each model. These three deep learning algorithms are summarized hereinafter.

#### A. RESIDUAL NEURAL NETWORK

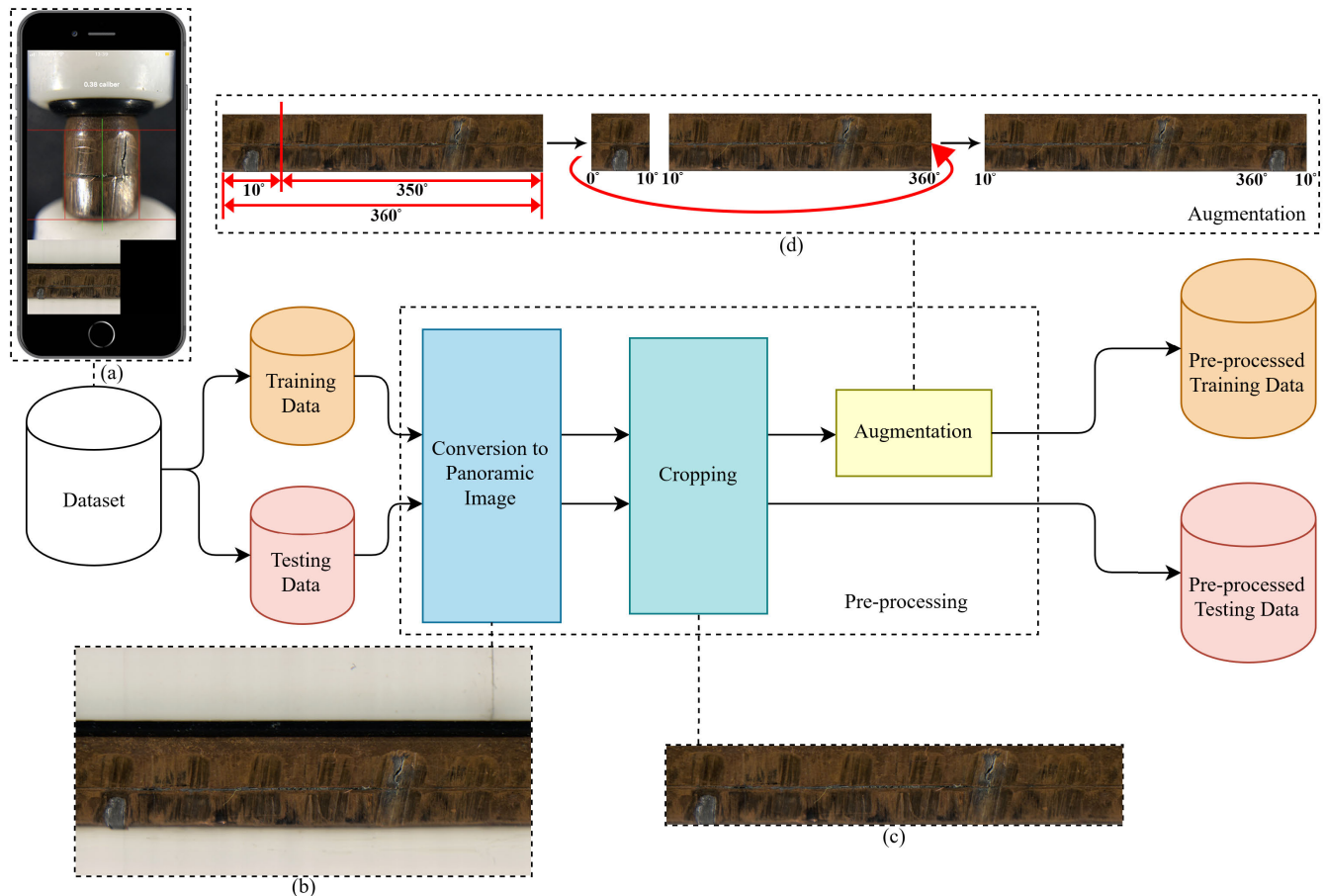
Residual neural network (ResNet) [35] is one of the first architectures that can handle sophisticated deep learning tasks. In the past, attempts have been made to increase the number of layers of deep neural networks for extracting high-level features to obtain a higher understanding of the data and make better predictions. As the number of neural network layers increases, the gradient of the loss function decreases and becomes too small for effective training, since the weight and biases of the initial layers will not be able to update effectively in each training session. Although renormalization and the rectified linear unit (ReLU) activation function might be able to resolve the vanishing gradient problem, they are not the best alternatives when the depth increases due to the emergence of the degradation problem. As neural networks start to converge, the accuracy becomes saturated, leading to a higher training error [35].

He *et al.* [35] proposed ResNet to alleviate the vanishing gradient and degradation problems using skip connections, allowing the model to successfully train many layers of a network by feeding the output of one layer as the input of the subsequent layers to reconstruct information required from previous layers. Consequently, the skip connection forms an alternate shortcut path that reduces the vanishing gradient problem and degradation problem. Furthermore, the skip connection allows the model to learn an identity function that ensures that higher layers work as good or better than the lower layers [35]. In this paper, ResNet50 was used.

#### B. DENSELY CONNECTED CONVOLUTIONAL NETWORK

The densely connected convolutional network (DenseNet) has a convolutional neural network architecture that is the state-of-the-art according to the classification results with





**FIGURE 5.** Preprocessing workflow. (a) Diagram of the portable hardware developed to allow users to directly capture a video of a discharged bullet found at a crime scene directly from a smartphone. (b) A panoramic image of the discharged bullet was generated ( $1,920 \times 3,720$  pixels) and then cropped. (c) After generating and cropping the panoramic images ( $1,200 \times 3,720$  pixels), they were divided into two datasets: training data (90%) and testing data (10%). (d) Each training panoramic image was circularly augmented  $10^\circ$ , providing a 36-fold increase in the number of training data for each bullet.

the ImageNet validation dataset. Huang *et al.* [36] used a direct connection from each layer to every other layer in a feed-forward direction. Each layer in the network receives concatenation of the feature maps produced in the previous layers as inputs and implements nonlinear functions such as batch normalization, ReLU, and convolution or pooling.

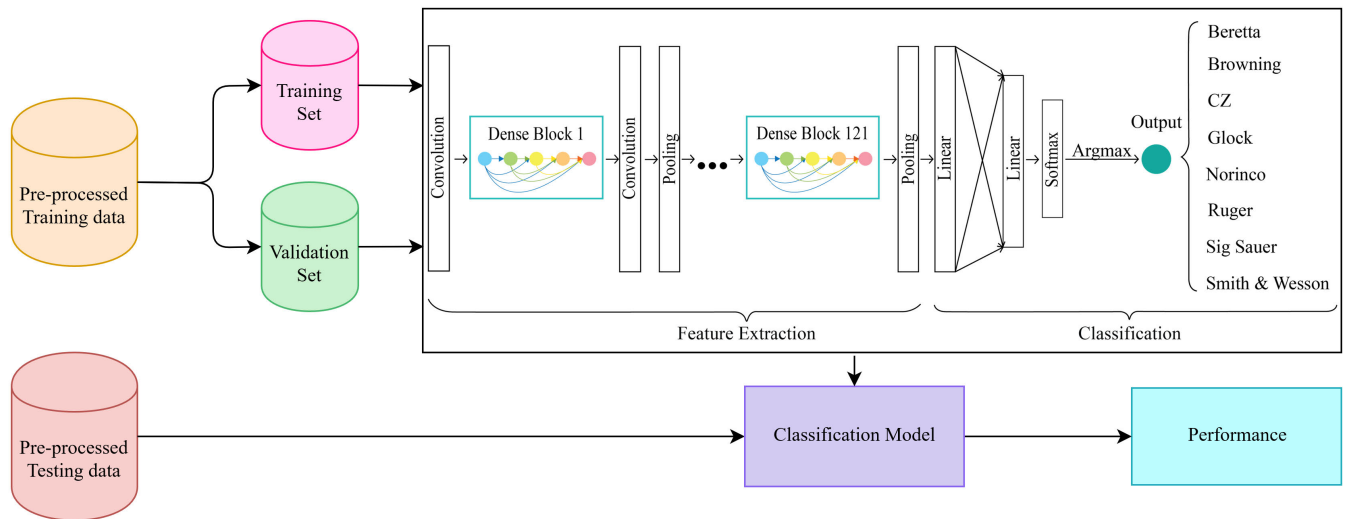
After the nonlinear function operation, the product feature maps of each layer are used as inputs to every subsequent layer. The concatenation operation is not effective when the size of the feature maps varies; therefore, the pooling operation is important by changing the size of the feature maps. To facilitate the pooling operation, the architecture is divided into multiple blocks, i.e., densely connected dense blocks, and the layers between dense blocks are transition layers that perform batch normalization, convolution, and pooling operations.

Generally, each function produces  $k$  feature maps, a hyperparameter called the growth rate. The growth rate determines how many feature maps in each layer contribute to the network. The feature maps can be accessed anywhere in the network once they are contributed. Unlike traditional

architecture, there is no need to replicate one layer to another. Each layer in the network produced  $k$  feature maps and typically has many inputs. For this reason, a  $[1 \times 1]$  convolution was used in the bottleneck layer to reduce the number of input feature maps to  $4k$ . Another merit of DenseNet is compactness: the ability to reduce the number of feature maps at transition layers. If a dense block contains  $m$  feature maps, the number of feature maps will change to  $\theta m$  after the transition layer, where  $0 < \theta \leq 1$  is referred to as the compression factor. In addition to compactness, the architecture of DenseNet has several advantages: alleviating the vanishing gradient problem, strengthening feature propagation, and reducing the number of parameters [36]. In this paper, DenseNet121 was used.

### C. XCEPTION

Xception [37] is a convolutional neural network that was developed by Google, and the model was the 1<sup>st</sup> runner up in ILSVRC 2015. Xception has been shown to outperform VGGNet [38], ResNet, Inception-v3 and Inception-v4 [39].



**FIGURE 6.** Classification workflow. Diagrams of panoramic images derived from preprocessed training data and preprocessed testing data (Fig. 5) were fed as inputs to the DenseNet121 model to automatically extract features and classify bullet markings into 8 different classes: Beretta, Browning, CZ, Glock, Norinco, Ruger, Sig Sauer, and Smith & Wesson.

Xception works using depthwise separable convolution and shortcuts between convolution blocks. Depthwise separable convolution involves two main processes: depthwise convolution and pointwise convolution. This approach has the advantages of efficiency in terms of computation time. In depthwise convolution, a convolution of size  $n \times n$  is applied. Unlike depthwise convolution, pointwise convolution is a  $[1 \times 1]$  convolution that is applied to change the dimension. Another process that is a part of Xception is the skip connections found in ResNet; therefore, Xception is an improvement of ResNet [37].

#### IV. EXPERIMENTAL SETUP

After data augmentation, all of the marking images from the discharged bullets in the training dataset were used to train DenseNet121, ResNet50, and Xception. The Adadelta optimization algorithm [40] was used to iteratively update the network with the training data, and the hyperparameters used in the training process were set as follows: a dropout rate of 0.50, a minibatch size of 32, and 30 epochs of training. The initial weights were the weights of a model pretrained with the ImageNet dataset [41]. After passing a panoramic image of the discharged bullet as an input to the models, the features were automatically extracted, resulting in a feature vector with 1,024 dimensions; the model weights were updated at the end of each learning rate cycle. The feature vector was fed as an input to the softmax classifier [42] and classified into one of eight classes: Beretta, Browning, CZ, Glock, Norinco, Ruger, Sig Sauer, and Smith & Wesson. The experiments were run on a system with the following specifications: a 44-core CPU, 64 GB of RAM, an Nvidia Telsa P40 GPU, TensorFlow 1.4.0, Keras 2.1.5, and Python 3. Figure IV illustrates one of three classification workflows (DenseNet121), where panoramic images derived from preprocessed training data

and preprocessed testing data (Fig. 5) were fed as an input to the DenseNet121 model to automatically extract features and classify bullet markings into 8 different classes: Beretta, Browning, CZ, Glock, Norinco, Ruger, Sig Sauer, and Smith & Wesson.

#### A. FIVE-FOLD CROSS-VALIDATION

There are a total of 898 bullets from 266 guns in the dataset: 116 Beretta bullets, 48 Browning bullets, 226 CZ bullets, 216 Glock bullets, 72 Norinco bullets, 57 Ruger bullets, 79 Sig Sauer bullets, and 84 Smith & Wesson bullets. These bullets provide a total of 32,328 images in the training dataset: 4,176 images of Beretta bullets, 1,728 images of Browning bullets, 8,136 images of CZ bullets, 7,776 images of Glock bullets, 2,592 images of Norinco bullets, 2,052 images of Ruger bullets, 2,844 images of Sig Sauer bullets, and 3,024 images of Smith & Wesson bullets. Panoramic images of the bullets were distributed such that all images from bullets fired from the same gun have to be in the same fold. It is worth mentioning that in each fold, there are approximately the same number of images.

Panoramic images of bullets from the development dataset were split into five folds, where each fold is composed of 53 guns, corresponding to 6,408–6,552 images (median 6,444). Subsequently, four folds were taken as the training dataset, and one fold was taken as the validation dataset. The testing dataset consists of 12 images of Beretta bullets (12 bullets from four guns), three images from Browning bullets (three bullets from one gun), 20 images from CZ bullets (20 bullets from six guns), 16 images from Glock bullets (16 bullets from five guns), seven images from Norinco bullets (seven bullets from one gun), four images from Ruger bullets (four bullets from one gun), six images from Sig Sauer bullets (six bullets from two guns), and seven

images from Smith & Wesson bullets (seven bullets from two guns). After the five-fold cross-validation was performed, the overall sensitivity, specificity, accuracy, and area under the receiver operating characteristic (ROC) curve (AUC) for each experiment were calculated.

**B. EVALUATION**

The efficiency of the models in classifying fired bullet images as Beretta, Browning, CZ, Glock, Norinco, Ruger, Sig Sauer, and Smith & Wesson were evaluated by graphing the ROC curves. The performance of each model was evaluated by calculating the AUC. Because CZ has the most registered firearms in Thailand, we focused on the classification of CZ. The model with the highest AUC was selected to be tested on the testing dataset.

**C. SIMILARITY SCORES**

The experimental setup was designed for classification among 8 firearm brands. Then, the confusion matrix was obtained. The similarity score and the perceptual distance for each brand pair were systematically derived from the confusion matrix based on a method proposed by [43]. The similarity score between each pair of firearm brands was calculated from the confusion matrix:

$$S_{ij} = \frac{P_{ij} + P_{ji}}{P_{ii} + P_{jj}}, \tag{1}$$

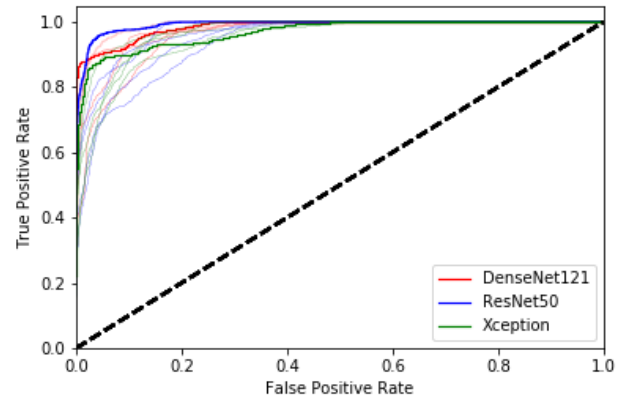
where  $S_{ij}$  is the similarity between brand  $i$  and brand  $j$  and  $P_{ij}$  is an element of the confusion matrix when given with brand  $i$  (row) and classified as brand  $j$  (column). Finally, the perceptual distance ( $D_{ij}$ ) is derived from the similarity score as follows [43]:

$$D_{ij} = -\ln S_{ij}. \tag{2}$$

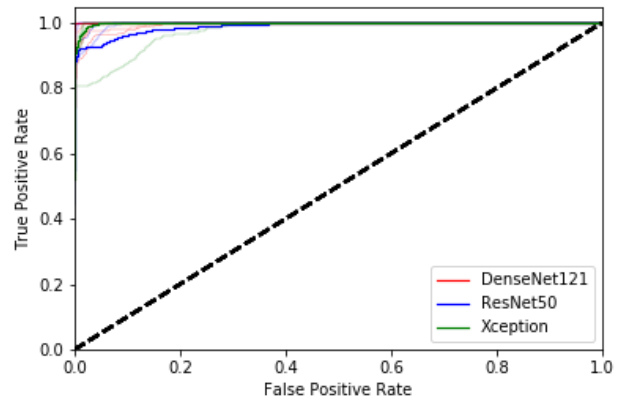
**V. EXPERIMENTAL RESULTS**

After performing 5-fold cross-validation, ROCs were plotted, and AUCs were calculated. Figure 7 illustrates all 15 ROCs of DenseNet121 (five red curves), ResNet50 (five blue curves), and Xception (five green curves) for CZ classification evaluated on the validation dataset. The thick red, blue, and green curves represent the ROCs with the highest AUCs of DenseNet121, ResNet50, and Xception, respectively. Similarly, the results of Glock and Smith & Wesson are shown in Figs. 8 and 9. The ROCs of Beretta, Browning, Glock, Norinco, Ruger, and Sig Sauer are not shown here due to the space limitations in this paper.

As previously stated, we focused on CZ classification. From the results shown in Table 4, DenseNet121 yielded the highest average AUC for CZ classification (the most registered firearm brand in Thailand) of  $0.9780 \pm 0.0130$  SD; the corresponding values for ResNet50 and Xception are  $0.9640 \pm 0.0195$  SD and  $0.9640 \pm 0.0084$  SD, respectively. Furthermore, the chosen DenseNet121 model with the highest AUC for the CZ classification has an AUC on the validation set of 0.99 and provided an AUC of 0.94 for



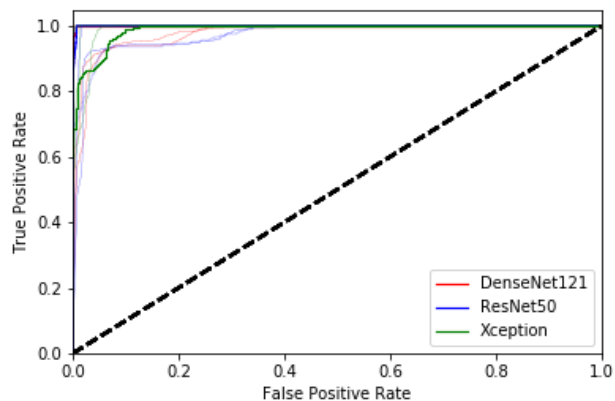
**FIGURE 7.** Fifteen receiver operating characteristic curves (ROCs) of DenseNet121 (five red curves), ResNet50 (five blue curves), and Xception (five green curves) for CZ classification evaluated on the validation set, where the thick red, green, and blue curves represent the ROC with the highest AUC of DenseNet121, ResNet50, and Xception, respectively.



**FIGURE 8.** Fifteen receiver operating characteristic curves (ROCs) of DenseNet121 (five red curves), ResNet50 (five blue curves), and Xception (five green curves) for Glock classification evaluated on the validation set, where the thick red, green, and blue curves represent the ROC with the highest AUC of DenseNet121, ResNet50, and Xception, respectively.

Beretta classification, 0.81 for Browning classification, 1 for Glock classification, 0.91 for Norinco classification, 0.92 for Ruger classification, 0.96 for Sig Sauer classification, and 1 for Smith & Wesson classification. Moreover, the best performance of ResNet50 for CZ classification has an AUC of 0.99 and provided an AUC of 0.97 for Beretta classification, 0.85 for Browning classification, 0.99 for Glock classification, 0.86 for Norinco classification, 0.89 for Ruger classification, 0.93 for Sig Sauer classification, and 1 for Smith & Wesson classification. Finally, the best performance of Xception for CZ classification has an AUC of 0.97 and provided an AUC of 0.9 for Beretta classification, 0.89 for Browning classification, 1 for Glock classification, 0.89 for Norinco classification, 0.92 for Ruger classification, 0.89 for Sig Sauer classification, and 0.99 for Smith & Wesson classification.

Two-factor balanced ANOVA was applied to determine whether there were any significant differences among the three CNN models and any interaction effect between the CNN model and the firearm brand. Therefore, we considered



**FIGURE 9.** Fifteen receiver operating characteristic curves (ROCs) of DenseNet121 (five red curves), ResNet50 (five blue curves), and Xception (five green curves) for Smith & Wesson classification evaluated on the validation set, where the thick red, green, and blue curves represent the ROC with the highest AUC of DenseNet121, ResNet50, and Xception, respectively.

**TABLE 4.** Average AUC ± SD for DenseNet121, ResNet50, and Xception (highest average AUC of each firearm brand shown in bold).

Brand	Average AUC ± SD		
	DenseNet121	ResNet50	Xception
Beretta	<b>0.9420 ± 0.0268</b>	0.9120 ± 0.0377	0.9180 ± 0.0303
Browning	0.8380 ± 0.0622	0.8240 ± 0.0550	<b>0.8680 ± 0.0239</b>
CZ	<b>0.9780 ± 0.0130</b>	0.9640 ± 0.0195	0.9640 ± 0.0089
Glock	<b>0.9980 ± 0.0044</b>	0.9980 ± 0.0045	0.9980 ± 0.0134
Norinco	<b>0.8980 ± 0.0216</b>	0.8260 ± 0.0835	0.8860 ± 0.0288
Ruger	<b>0.9300 ± 0.0400</b>	0.8820 ± 0.0606	0.9140 ± 0.0279
Sig Sauer	<b>0.9620 ± 0.0164</b>	0.9280 ± 0.0286	0.9000 ± 0.0361
Smith & Wesson	0.9920 ± 0.0110	0.9860 ± 0.0152	<b>0.9960 ± 0.0055</b>

two independent variables, the CNN model (DenseNet121, ResNet50, and Xception) and the firearm brand (Beretta, Browning, CZ, Glock, Norinco, Ruger, Sig Sauer, and Smith & Wesson), and one dependent variable: AUC. The experimental results showed that the three CNN models are significantly different [ $F(2, 96) = 6.28, p = 0.0027$ ]. DenseNet121 provided the highest average AUC, and post hoc analysis showed a statistically significant higher AUC than that of ResNet50 ( $p = 0.0018$ ) and a higher AUC than that of Xception ( $p = 0.2547$ ). In addition, there are no interaction effects between the two independent variables (CNN model and firearm brand) on AUC [ $F(14, 96) = 1.45, p = 0.1447$ ]. When considering the AUC of DenseNet121 in classifying firearm brands from bullet markings, the average AUC from six firearm brands—Beretta, CZ, Glock, Ruger, Sig Sauer, and Smith & Wesson—are significantly higher than that of Browning.

DenseNet121 was the model chosen for firearm classification from bullet markings in this paper. The accuracy, sensitivity, and specificity of each firearm brand tested on the test set are shown in Table 5. From the results, DenseNet121 has very high accuracy for all gun brands, ranging from 91.18% (Beretta and CZ), 96.88%

**TABLE 5.** Sensitivity, specificity and accuracy for the testing dataset.

Classes	Sensitivity	Specificity	Accuracy
Beretta	83.33%	92.86%	91.18%
Browning	33.33%	100%	96.88%
CZ	88.24%	92.16%	91.18%
Glock	100%	97.87%	98.41%
Norinco	71.43%	100%	96.88%
Ruger	75.00%	100%	98.41%
Sig Sauer	83.33%	100%	98.41%
Smith & Wesson	100%	98.21%	98.41%

(Browning and Norinco), and 98.41% (Glock, Ruger, Sig Sauer, Smith & Wesson). DenseNet121 has 33.33% sensitivity for Browning; 71.43% and 75.00% sensitivity for Norinco and Ruger, respectively; and 83.33%, 83.33%, and 88.24% sensitivity for Beretta, Sig Sauer, and CZ, respectively. Interestingly, DenseNet121 has 100% sensitivity for both Glock and Smith & Wesson. Finally, DenseNet has very high specificity: 92.16% for CZ, 92.86% for Beretta, 97.87% for Glock, 98.21% for Smith & Wesson, and 100% for Browning, Norinco, Ruger, and Sig Sauer. More specifically, Glock and Smith & Wesson both have 100% sensitivity and 97.87% and 98.21% specificity, respectively, and both have 98.41% accuracy, which are the highest performance metrics among the other firearm brands. On the other hand, Browning has only 33.33% sensitivity but has 100% specificity, resulting in 96.88% accuracy.

The confusion matrix of DenseNet121 on the testing dataset is shown in Table 6, where the row is the ground truth and the column is the predicted class. Each number represents raw data, whereas its normalized percentage is represented in parentheses. For Beretta, ten of twelve fired bullets were correctly classified, but two fired bullets were misclassified as CZ. For Browning, only three fired bullets were available in the test set: one fired bullet was correctly classified, but two were misclassified as Beretta. For CZ, 15 of 17 fired bullets were correctly classified, whereas one fired bullet each was misclassified as Glock and another was misclassified as Smith & Wesson. Impressively, bullets fired from Glock and Smith & Wesson were all correctly classified. For Norinco, five fired bullets were correctly classified, but one fired bullet was misclassified as Beretta and another was misclassified as CZ. For Ruger, three fired bullets were correctly classified, and one fired bullet was misclassified as CZ. For Sig Sauer, five fired bullets were correctly classified, and one bullet was misclassified as Beretta. Among all misclassified fired bullets, fired bullets were misclassified as Beretta and CZ most frequently (four bullets each). Specifically, among the four fired bullets misclassified as Beretta, two bullets were fired from Browning, one bullet was fired from Norinco and one from Sig Sauer. For the four bullets misclassified as CZ, two bullets were fired from Beretta, one bullet was fired from Norinco and one from Ruger. In addition, one fired bullet

**TABLE 6. Confusion matrix of DenseNet121 to classify firearms from bullet markings in the testing dataset.**

Ground truth	Predicted class								Total
	Beretta	Browning	CZ	Glock	Norinco	Ruger	Sig Sauer	Smith & Wesson	
Beretta	10(83.33%)	0	2(16.67%)	0	0	0	0	0	12(100%)
Browning	2(66.67%)	1(33.33%)	0	0	0	0	0	0	3(100%)
CZ	0	0	15(88.24%)	1(5.88%)	0	0	0	1(5.88%)	17(100%)
Glock	0	0	0	16(100%)	0	0	0	0	16(100%)
Norinco	1(14.29%)	0	1(14.29%)	0	5(71.43%)	0	0	0	7(100%)
Ruger	0	0	1(25%)	0	0	3(75%)	0	0	4(100%)
Sig Sauer	1(16.67%)	0	0	0	0	0	5(83.33%)	0	6(100%)
Smith & Wesson	0	0	0	0	0	0	0	7(100%)	7(100%)
Misclassification	4	0	4	1	0	0	0	1	

**TABLE 7. Distance scores between pairs of eight firearm brands.**

Brand	Beretta	Browning	CZ	Glock	Norinco	Ruger	Sig Sauer	Smith & Wesson
Beretta	0	0.2513	2.3316	$\infty$	2.3826	$\infty$	2.3026	$\infty$
Browning		0	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$
CZ			0	3.4657	2.4138	1.8763	$\infty$	3.4657
Glock				0	$\infty$	$\infty$	$\infty$	$\infty$
Norinco					0	$\infty$	$\infty$	$\infty$
Ruger						0	$\infty$	$\infty$
Sig Sauer							0	$\infty$
Smith & Wesson								0

was misclassified as Glock and another was misclassified as Smith & Wesson. These will be discussed in the next section.

Table 7 presents the distance scores between pairs of eight firearm brands, which were calculated from Eq. (2) using the percentage values in Table 6. The shorter the distance score is, the higher the similarity between a pair of firearm brands (based on bullet markings). From the results, Beretta and Browning are most similar with the shortest distance score (0.2513). This can be visualized from the bullet markings shown in Fig. 2 (a) and (b). CZ and Ruger are ranked second most similar, with a distance score of 1.8763. In addition, Beretta and Sig Sauer (2.3026), Beretta and CZ (2.3316), Beretta and Norinco (2.3826), and CZ and Norinco (2.4138) have higher distance scores. Finally, CZ and Glock, CZ and Smith & Wesson have the same distance score of 3.4657.

## VI. DISCUSSION AND CONCLUSION

We have developed a highly accurate automated firearm classification method. The proposed method uses computer-aided approaches to provide fast and accurate firearm brand classifications of eight firearm brands from panoramic images of fired bullets. As the first firearm brand classifier on a mobile application to use deep learning, the classification of the most common registered firearms in Thailand can bring convenience and accelerate bullet identification processes for many forensic science examiners by narrowing down the list of suspects.

In Thailand, many criminals often use the same brand of high-quality hand guns because many firearms are expensive. Therefore, the classification of firearm brands can help identify criminals that have been involved in repeated crimes, regardless of where the crime is committed. Most importantly, the use of the mobile application is relatively cheap and is accessible for all officers without a high level of expertise in examining physical evidence collected from crime scenes. The proposed firearm brand classifier decreases the time required for forensic science investigations, increasing the likelihood of identifying and arresting criminals for gun violence.

To avoid overfitting of bullets with similar features that were fired from the same barrel, five-fold cross-validation was applied during data partitioning so that every bullet image fired from the same firearm was categorized as one subject and every fold was partitioned by subject. As a state-of-the-art classification approach, deep learning results benefit from the transfer of learning from pretrained networks that have a large dataset. For this reason, we chose the top three CNN models that have the highest accuracy and the fewest number of layers. According to a two-factor balanced ANOVA performed on the average AUC for all three models, as shown in Table 4, DenseNet121 outperformed ResNet50 and Xception, among which the performance differences were statistically significant. No interaction effects were found between the CNN model and firearm brand, suggesting that the performance

metrics of all three CNN models were not influenced by any firearm brands.

The proposed models using DenseNet121, ResNet50, and Xception involved the Keras and TensorFlow backends. The reason behind the model selection is that we selected models have high accuracy scores and as few layers as possible; thus, DenseNet121 and ResNet50 were selected herein. We suggest that several advantages of DenseNet, including alleviating the vanishing gradient problem, strengthening feature propagation, and reducing the number of parameters [36], are the main reasons that it outperformed ResNet50 and Xception.

In addition, the dense connectivity pattern of DenseNet requires fewer parameters than ResNet and does not need to relearn redundant feature maps. In fact, ResNet contributes to the information that passes from its preceding layer to the subsequent layer but needs very little preservation and can be randomly dropped during training. Normal deep learning architectures use depthwise convolution followed by a pointwise convolution method, whereas the Xception architecture uses pointwise convolution followed by a depthwise convolution method, a distinction from other architectures that is worth exploring.

Out of all firearm brand classifications, Glock classification achieved the highest results, mainly because of its difference in rifling from the other seven firearm brands. Glock is the only firearm brand in the dataset that has a 6-groove right-hand twist and polygonal rifling, as shown in Fig. 2(d). On the other hand, other firearm brands have a 6-groove right-hand twist and conventional rifling, as shown in Fig. 2. Similarly, Smith & Wesson bullets were rarely misidentified because the Smith & Wesson barrel has a 5-groove right-hand twist, which is different from other the firearm brands used in this experiment, as shown in Fig. 2(h).

There were a few drawbacks to the dataset, since there was a limited number of bullets available for data collection at the Firearms and Ammunition Subdivision of the Central Police Forensic Science Division in Thailand. First, the collected discharged bullet samples were composed of different materials because the dataset was collected from crime cases, which is related to the popularity of firearm brands used by criminals in Thailand. The Browning training dataset has 48 fired bullets from 14 guns, whereas the CZ has 226 fired bullets from 68 guns; consequently, the differences in the number of firearms and bullets fired from each firearm brand may affect the learning of the models.

In the experiment, an unequal number of lead bullets with and without copper jackets were used to train and test the model. The number of lead bullets without copper jackets collected was considerably lower than the number of lead bullets with copper jackets. Second, the number of discharged bullet samples from each firearm brand was unbalanced, which may lead to bias in the models. In our dataset, there was an overwhelming number of CZ bullets: 226 bullet fired bullets from 68 CZ guns. This could potentially affect the classification of other bullets fired from other firearm brands

that have similar bullet markings as CZ, such as Beretta, Norinco, Ruger, and Sig Sauer. On the other hand, bullets with different markings from CZ, such as those produced by the barrel rifling types and number of twists of Glock and Smith & Wesson, seemed unaffected by the unbalanced data, i.e., the lead bullet samples in the Browning brand were included from only two firearms containing lead bullets (1 already partitioned into the test dataset), which cannot be partitioned 5-fold. However, this drawback was addressed by putting the mentioned firearm in every training set five-fold while leaving the validation set with no lead bullets.

## VII. FUTURE WORK

Since the iPhone 8 was used to take videos of the rotating bullets, the optimal rotational speed of the motor was found based on trial-and-error. In our future work, we will use position markers placed on the two concentric shafts of our hardware to not only identify the width of the bullet specimen but also to calculate the optimal speed of the rotating motor so that it corresponds with the shutter speed of the smartphone used to take the videos.

In the future, we hope to develop our algorithm further to enable classification of more class characteristic subjects, such as the firearm model and serial number from bullet markings. Furthermore, we also aspire to expand our dataset so that damaged bullets can be classified, since our current dataset consists of undamaged bullets. Taking this limitation into consideration, we will crop out sections of a fewer number of pairs of lands and grooves from bullet images and then use the cropped images to train our model instead of using the whole panoramic image as our input.

## REFERENCES

- [1] T. Parks, N. J. Colletta, and B. Oppenheim, *The Contested Corners of Asia: Subnational Conflict and International Development Assistance*. San Francisco, CA, USA: Asia Foundation, 2013.
- [2] N. Aizenman, "Gun violence: How the US compared with other countries," NPR, Seattle, WA, USA, Tech. Rep. 10, Oct. 2017.
- [3] Y. Yusoff, R. Ismail, and Z. Hassan, "Common phases of computer forensics investigation models," *Int. J. Comput. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 17–31, Jun. 2011.
- [4] R. Saferstein, *Criminalistics: An Introduction to Forensic Science*, 12th ed. London, U.K.: Pearson, 2017.
- [5] B. E. Turvey, *Criminal Profiling: An Introduction to Behavioral Evidence Analysis*. Academic Press, 2011.
- [6] N. R. Council, *Forensic Analysis: Weighing Bullet Lead Evidence*. National Academies Press, 2004.
- [7] A. F. Sedda and G. Rossi, "Bullets fragments identification by comparison of their chemical composition obtained using instrumental neutron activation analysis," *Forensic Sci. Int.*, vol. 206, nos. 1–3, pp. e5–e7, Mar. 2011.
- [8] A. Carriquiry, H. Hofmann, X. H. Tai, and S. VanderPlas, "Machine learning in forensic applications," *Significance*, vol. 16, no. 2, pp. 29–35, Apr. 2019.
- [9] D. Li, "Ballistics projectile image analysis for firearm identification," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2857–2865, Oct. 2006.
- [10] E. Nadeau, M. Pages, D. Roberge, A. Roy, and M. Léger, "3D bullet and cartridge case analysis," U.S. Patent 7 796 810, Sep. 14, 2010.
- [11] *How Good a Match is it? Putting Statistics Into Forensic Firearms Identification*. Accessed: Nov. 10, 2019. [Online]. Available: <https://www.nist.gov/news-events/news/2018/02/how-good-match-it-putting-statistics-forensic-firearms-identification>

- [12] N. Lummis. *Excerpts from Basic Firearms Identification—PPT Video Online Download*. Accessed: Nov-2018. [Online]. Available: <https://slideplayer.com/slide/1695040>
- [13] W. Harris. *10 innovations that led to the modern bullet*. Accessed: Jun. 27, 2019. [Online]. Available: <https://science.howstuffworks.com/10-bullet-innovations8.htm>
- [14] J. E. Davis, *An Introduction to Tool Marks, Firearms and The Striagraph*. Springfield, IL, USA: Charles C Thomas Publisher, 1968.
- [15] R. M. Thompson, “Firearm identification in the forensic science laboratory,” Nat. District Attorneys Assoc., Alexandria, VA, USA, 2010.
- [16] S. Vanderplas, M. Nally, T. Klep, C. Cadevall, and H. Hofmann, “Comparison of three similarity scores for bullet LEA matching,” *Forensic Sci. Int.*, vol. 308, Mar. 2020, Art. no. 110167.
- [17] J. Lu, S.-H. Wu, K.-C. Yang, and M. Xia, “Automated bullet identification based on striation feature using 3D laser color scanner,” *Optik*, vol. 125, no. 10, pp. 2270–2273, May 2014.
- [18] J. De Kinder and M. Bonfanti, “Automated comparisons of bullet striations based on 3D topography,” *Forensic Sci. Int.*, vol. 101, no. 2, pp. 85–93, Apr. 1999.
- [19] W. Chu, J. Song, T. Vorburger, J. Yen, S. Ballou, and B. Bachrach, “Pilot study of automated bullet signature identification based on topography measurements and correlations,” *J. Forensic Sci.*, vol. 55, no. 2, pp. 341–347, Mar. 2010.
- [20] T. V. Vorburger, J. H. Yen, B. Bachrach, T. B. Renegar, L. Ma, H. G. Rhee, X. A. Zheng, J. F. Song, and C. D. Foreman, “Surface topography analysis for a feasibility assessment of a national ballistics imaging database,” Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. 7362, May 2007.
- [21] S. Bigdeli, H. Danandeh, and M. Ebrahimi Moghaddam, “A correlation based bullet identification method using empirical mode decomposition,” *Forensic Sci. Int.*, vol. 278, pp. 351–360, Sep. 2017.
- [22] T. V. Vorburger, J. Song, and N. Petraco, “Topography measurements and applications in ballistics and tool mark identifications,” *Surf. Topography, Metrol. Properties*, vol. 4, no. 1, 2015, Art. no. 013002.
- [23] A. A. Braga and G. L. Pierce, “Reconsidering the ballistic imaging of crime bullets in gun law enforcement operations,” *Forensic Sci. Policy Manage., Int. J.*, vol. 2, no. 3, pp. 105–117, Aug. 2011.
- [24] T. B. Brinck, “Comparing the performance of IBIS and BulletTRAX-3D technology using bullets fired through 10 consecutively rifled barrels,” *J. Forensic Sci.*, vol. 53, no. 3, pp. 677–682, May 2008.
- [25] F. Xie, S. Xiao, L. Blunt, W. Zeng, and X. Jiang, “Automated bullet-identification system based on surface topography techniques,” *Wear*, vol. 266, nos. 5–6, pp. 518–522, Mar. 2009.
- [26] N. Petraco and H. Chan, *Application of Machine Learning to Toolmarks: Statistically Based Methods for Impression Pattern Comparisons*. Mannheim, Germany: Bibliographisches Institut AG, 2012.
- [27] A. Banno, “Estimation of bullet striation similarity using neural networks,” *J. Forensic Sci.*, vol. 49, no. 3, pp. 1–5, 2004.
- [28] P. Changmai, K. Bora, R. Suresh, N. Deb, and L. B. Mahanta, “On the study of automated identification of firearms through associated striations,” in *Proc. 31st Int. Symp. Ballistics*, 2019, doi: [10.12783/ballistics2019/33156](https://doi.org/10.12783/ballistics2019/33156).
- [29] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [30] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, “Scalable object detection using deep neural networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2147–2154.
- [31] O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep face recognition,” in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, 2015, vol. 1, no. 3, p. 6.
- [32] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.
- [33] *Imagenet*. 2016. Accessed: Oct. 19, 2019. [Online]. Available: <http://www.image-net.org/index>
- [34] *Panorama | Create Beautiful High-Resolution Panoramas | iOS*. Accessed: Oct. 23, 2019. [Online]. Available: <http://www.limitpointstore.com/products/panorama/iOS.html>
- [35] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [36] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
- [37] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258.
- [38] L. Wang, S. Guo, W. Huang, and Y. Qiao, “Places205-VGGNet models for scene recognition,” *CoRR*, vol. abs/1508.01667, 2015.
- [39] C. Szegedy, S. Ioffe, and V. Vanhoucke, “Inception-v4, inception-resnet and the impact of residual connections on learning,” *CoRR*, vol. abs/1602.07261, 2016. [Online]. Available: <http://arxiv.org/abs/1602.07261>
- [40] M. D. Zeiler, “ADADELTA: An adaptive learning rate method,” 2012, *arXiv:1212.5701*. [Online]. Available: <http://arxiv.org/abs/1212.5701>
- [41] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [42] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [43] N. Shepard, “Psychological representation of speech sounds,” in *Human Communication: A Unified View*. New York, NY, USA: McGraw-Hill, 1972, pp. 67–113.



**PATTRANIT PISANTANAROJ** received the B.E. degree in software engineering from Thammasat University, Bangkok, Thailand, in 2018. Her research interests include machine learning and deep learning in forensic science.



**PIMLAPUS TANPISUTH** is currently pursuing the degree with the Ruamrudee International School, Bangkok, Thailand. She is currently a Researcher with the Center of Excellence in Intelligent Informatics, Speech and Language Technology.



**PIYAWUT SINCHAVANWAT** received the B.E. degree in computer engineering from Thammasat University, Bangkok, Thailand, in 2018. His research interests include machine learning and deep learning in forensic science.



**SIRIPORN PHASUK** (Member, IEEE) received the B.E. degree in computer engineering from Thammasat University, Bangkok, Thailand, in 2018. Her research interests include machine learning and deep learning in medicine.



**PHONGPHAN PHIENPHANICH** (Member, IEEE) received the B.E. degree (Hons.) in computer engineering from the Suranaree University of Technology, Nakhon Ratchasima, Thailand, in 2009, and the M.E. degree in electrical engineering from Thammasat University, Bangkok, Thailand, in 2012, where he is currently pursuing the Ph.D. degree in computer engineering. He worked as a Co-Researcher with the National Electronics and Computer Technology Center (NECTEC),

Thailand, from 2010 to 2012. His research interests include signal and speech processing, pattern recognition, and machine learning.



**MONTRI DONPHONGPI** received the B.Sc. degree in chemistry from Ramkhamhaeng University, Bangkok, Thailand, in 1997. He has been working as a Firearms and Toolmark Examiner with the Central Institute of Forensic Science (CIFS), Ministry of Justice, Bangkok, since 2004.



**PARINTON JANGTAWEE** received the B.Sc. degree in chemistry from the Prince of Songkla University, Songkla, Thailand, in 2012, and the M.Sc. degree in physical chemistry from Chulalongkorn University, Bangkok, Thailand, in 2015, where he is currently pursuing the Ph.D. degree in physical chemistry. He is currently a Co-Founder of Lens and Smart Classroom Company, Ltd., that manufactures and sells smartphone microscopes, lab kits, and learning aids for education, a company established in 2016. He has been interested in innovative projects.



**SANONG EKGASIT** received the B.Sc. degree in chemistry and the M.Eng. in polymer technology from Chulalongkorn University, Bangkok, Thailand, in 1989 and 1992, respectively, and the Ph.D. degree in polymer science and engineering from Case Western Reserve University, Cleveland, OH, USA, in 1996. He has been working with the Department of Chemistry, Faculty of Science, Chulalongkorn University, since 1999. His research interests include molecular spectroscopy, nanomaterials (synthesis, modification, and commercialization), cellulose nanocrystals (synthesis and applications), and lab-on-phone research.



**CHARTURONG TANTIBUNDHIT** (Member, IEEE) received the B.E. degree in electrical engineering from Kasetsart University, Bangkok, Thailand, in 1996, and the M.S. degree in information science and the Ph.D. degree in electrical engineering from the University of Pittsburgh, Pittsburgh, PA, USA, in 2001 and 2006, respectively. From 2007 to 2008, he was a Postdoctoral Researcher with the Signal Processing and Speech Communication Laboratory (SPSC), Graz University of

Technology, Graz, Austria. Since 2006, he has been working with Thammasat University, Thailand, where he is currently an Associate Professor with the Department of Electrical and Computer Engineering, and the Head of the Speech and Language Technology Cluster, Center of Excellence in Intelligence Informatics, Speech and Language Technology, and Service Innovation (CILS). His research interests include handcrafted machine learning and deep learning in medicine, biomedical signal processing, and speech processing. He was the IEEE ICASSP Student Paper Contest Winner, in 2006. He led a team that won the Grand Prix from the 45th International Exhibition of Inventions of Geneva, in 2017.



**KITTISAK YAKOOMPAI** received the B.E. degree in physics and the M.A. degree in jurisprudence from Thammasat University, Bangkok, Thailand. He currently works as a Scientist (Level 4) with the Firearms and Ammunition Identification Subdivision, Central Scientific Crime Detection Division, Bangkok.

...