MJFAS
Malaysian Journal of Fundamental and Applied Sciences

Chuan *et al.* | Malaysian Journal of Fundamental and Applied Sciences, Vol. 19 (2023) 525-538

RESEARCH ARTICLE

A Comparative of Two-Dimensional Statistical Moment Invariants Features in Formulating an Automated Probabilistic Machine Learning Identification Algorithm for Forensic Application

Zun Liang Chuan[a]*, David Chong Teak Wei[b], Connie Lee Wai Yan[a], Muhammad Fuad Ahmad Nasser[a], Nor Azura Md Ghani[c], Abdul Aziz Jemain[d], Choong-Yeun Liong[d]

[a]Centre for Mathematical Science, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuh PersiaranTun Khalil Yaakob, 26300 Kuantan, Pahang, Malaysia; [b]Ever Al Holdings Sdn Bhd, 12, Jalan Anggerik Aranda 31/170c, Kota Kemuning, 40460 Shah Alam, Selangor DE, Malaysia; [c]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor DE, Malaysia; [d]Department of Mathematical Sciences, Faculty of Science & Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor DE, Malaysia

Abstract IBIS, ALIS, EVOFINDER, and CONDOR are the massive ballistics computerised technological machines that have typically been utilised in forensic laboratories to automatically locate similarities between images of cartridge cases and bullets. However, it imposed a long execution time and requires physical interpretation to consolidate the analysis results when employing these market-available technologies to accomplish ballistics matching tasks. Therefore, the principal objective of this study is to propose an improvised automated probabilistic machine learning identification algorithm by extracting the two-dimensional (2D) statistical moment invariants from the segmented region of interest (ROI) corresponding to the cartridge case and bullets images. To pursue this principal objective, several 2D statistical moment invariants have been compared and tested to determine the most suitable feature set applied in the proposed identification algorithm. The 2D statistical moment invariants employed include Orthogonal Legendre moments (OLM), Hu moments (HM), Tsirikolias-Mertzois moments (TMM), Pan-Keane moments (PKM), and Central Geometric moments (CGM). Moreover, the proposed identification algorithm is also tested in different scenarios, including based on the classification of strength association measurements between the extracted feature sets. The empirical results in this article revealed that the proposed identification algorithm applied with the CGM comprising the weak association classification yielded the best identification accuracy rates, which are >96.5% across all the sample sizes of the training set. These empirical results also conveyed that the superior proposed identification algorithm in this research could be developed as a mobile application for ballistics identification that can significantly reduce the time taken and conveniently perform the ballistics identification tasks.

Keywords: Ballistics identification, automated, machine learning identification algorithm, statistical moment invariants.

***For correspondence:**
chuanzl@ump.edu.my

Received: 25 Feb. 2023
Accepted: 27 July 2023

Introduction

Figure 1 depicts the surface of a fired cartridge case has several distinctive key impressions, including

the ejector, breech face, and firing pin impressions. These impressions are created during the cartridge case's loading, firing, and disposal processes [1-6]. These impressions are of paramount vital in a criminal investigation since they allow judicatory ballistics experts to identify the type and model of a firearm employed by the criminals. The traditional ballistics identification method employed in forensic laboratories is performed by comparing the characteristic impressions of the cartridge case found at the crime scene with test specimens based on a low-powered optical comparison microscope [7]. However, the traditional identification method has inherent difficulties because it is heavily dependent on the expertise and experience of judicatory ballistics experts. Moreover, the traditional identification method is time-consuming, imposing several weeks for a single analysis [8].

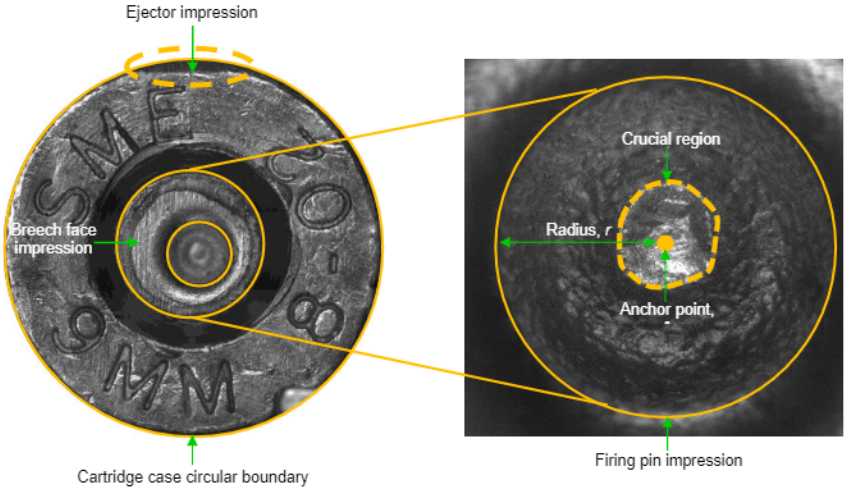

Figure 1. Characteristic of key impressions left on the fired cartridge case

Since the 1990s, several market-available ballistics identification systems have been developed and widely employed in forensic laboratories, such as IBIS, ALIS, EVOFINDER, and CONDOR systems. Although the execution time for a single analysis can be significantly reduced from several weeks to several hours by utilising these market-available ballistics identification systems, analysing the specimens remains dependent on physical interpretation in order to authenticate the analytical results. In the past few decades, several semi-automated and automated feature-based probabilistic machine learning identification algorithms have been proposed to address this limitation [1-5, 8-15]. Comparing these semi-automated and automated probabilistic machine learning identification algorithms to the market-available ballistics identification systems, the principal benefits of these probabilistic machine learning identification algorithms do not depend on physical interpretation and imposed a short execution time ranging from seconds to minutes. The general schematic architecture of the semi-automated and automated feature-based probabilistic machine learning identification algorithms in practice is depicted in Figure 2.

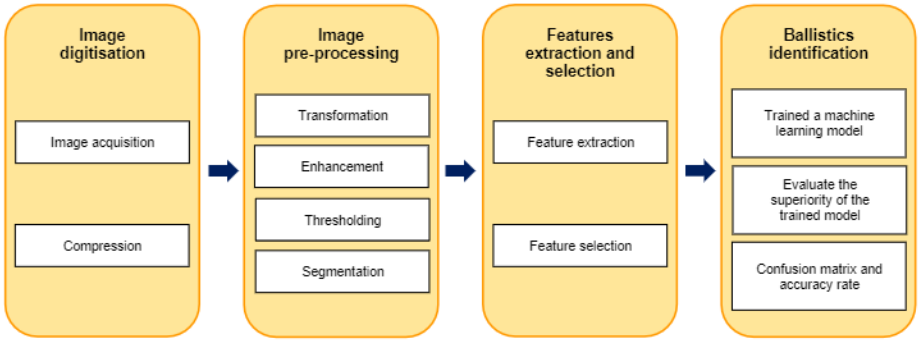

Figure 2. The general architecture in formulating a semi-automated and automated feature-based probabilistic machine learning identification algorithm based on data science methodology

In particular, there are four principal stages in architecting the feature-based probabilistic machine learning identification algorithm, namely image digitisation (data understanding[1]), image pre-processing (data preparation[1]), feature extraction and selection (data preparation[1]), and ballistics identification (modelling and evaluation[1]). For the primary stage of image digitisation, the digital images of fired bullets and cartridge cases are frequently captured utilising the market-available ballistics identification system. Meanwhile, for the image pre-processing stage, the input digital images acquired utilising the market-available ballistics identification system are subjected to several appropriate image processing operators, including image enhancement, noise reduction, and segmentation. Image pre-processing is essential in most pattern recognition algorithms since blurring and noise may be introduced into the images during the digitising process. The superior operation for image pre-processing tasks such as image enhancement and segmentation is required to select utilising a trial and error approach because there is no general theory to guide this decision [6, 12, 16]. In this research, the position of unique impressions on fired cartridge cases is also required to determine before segmentation. The segmented image or region of interest (ROI) is subsequently represented and aggregated in quantitative measurement during the feature's extraction stage so that it might be utilised more easily during the ballistics identification stage. On the other hand, not all features extracted from the key impressions are useful in the ballistics identification stage. Hence, the selection of a set of relevant and informative features is indeed much needed. The selected features required have two properties, namely the selected features required to have a minimum variation within pattern classes and vice versa for the between pattern classes. Furthermore, the selected features should comprise the invariant properties corresponding to translation, rotation, and scaling (TRS) [16, 17]. By inputting the selected informative features into the trained superior probabilistic machine learning identification algorithm, the unknown classes of the ballistics can be identified.

In order to segment the ROI from the images, determining the position of crucial distinctive key impressions on fired cartridge case images is indeed much needed. In literature [1, 9, 14], the circle Hough Transform (CHT) is the method utilised to determine the position of crucial distinctive key impressions on fired cartridge case images in the previously proposed probabilistic machine learning identification algorithms. However, CHT required a lot of time and effort due to its mathematical complexity rather than the least square-fitting circle that was employed in this study [18]. Therefore, the first principal objective of this article is to develop an improvised automated machine learning identification algorithm for ballistics that utilises the unweighted least square-fitting circle to detect the position (anchor point, **A**) of the centre-firing pin impression circular boundary with radius *r*. In this research, the firing pin impression is the principal focus key impression in developing the improvised machine learning identification algorithm for ballistics. This is because the previous empirical studies conveyed the invariant properties of this impression in ballistics identification rather than others' key impressions on the fired cartridge cases [2-6,11-15].

Moreover, the principal limitation of the existing market-available ballistics identification systems is the lack of good separability feature extraction techniques and related intelligent methods [10]. In other words, feature extraction remains an unresolved issue from a practical perspective as detailed in Section 2. Therefore, the second principal objective of this article is to identify the superior two-dimensional (2D) statistical moment invariants extracted from the segmented ROI corresponding to the firing pin impression images. The 2D statistical moment invariants taken into account in this article are including Orthogonal Legendre moments (OLM), Hu moments (HM), Tsirikolias-Mertzois moments (TMM), Pan-Keane moments (PKM), and Central Geometrics moments (CGM), which were tested utilising different scenarios in term of classification of strength association measurements and the sample sizes of training sets. Meanwhile, the principal reason the 2D statistical moment invariants is preferable is that it provides sufficient discrimination power to identify the ballistics belonging to different types and models [19].

In order to pursue the aforementioned principal objectives, the rest of this article is organised as followed. In Section 2, an overview of related works in formulating the semi-automated and automated machine learning identification algorithms for ballistics based on previous studies is presented, while Section 3 provided an overview of the theoretical background for the 2D statistical moment invariants are taken into account in this study. Meanwhile, the schematic architecture of the proposed improvised automated machine learning identification algorithm for ballistics in this study is detailed in Section 4. Moreover, the empirical results and discussion of this research are detailed in Section 5. Finally, the concluding remarks and future work in this study are provided in Section 6.

[1] Characterisation of the activity in the stage for the data science methodology.

Related Works

In the early 2000s, Xin ***et al***. [1] proposed an automated machine learning identification algorithm for ballistics by extracting and selecting several good separability features from the key impressions, such as the firing pin, bullet bottom, and ejector impressions on the fired cartridge cases. In particular, the firing pin impression yields features such as position, radius, depth, shape, tongue impression, and tiny impression, whereas the bullet bottom impression yields texture features. Moreover, shape, position, and direction are among the features extracted from the ejector impression. In this study, they employed CHT to detect the position of the bullet bottom and firing pin impressions. The empirical results of this study revealed that the proposed automated machine learning identification algorithm for ballistics is competent to yield an identification accuracy rate of >85% for the top 5%, 10%, and 20%, respectively. Zhou ***et al***. [9] also proposed an automated machine learning identification algorithm for ballistics. In this research, they proposed to extract the position of the firing pin impression, the shape description of the firing pin impression employing an active contour model, and the texture pattern of the breech face impression as the features, which the position of the bullet bottom and firing pin impressions are detected utilising CHT. Based on their empirical results, they have concluded that the proposed automated machine learning identification algorithm for ballistics employing a support vector machine (SVM) is performed superior rather than the existing identification algorithms.

On the other hand, Leng and Huang [10] proposed another automated machine learning identification algorithm for ballistics by extracting the statistical moment invariants from the centre-firing and rim-firing mechanism of fired cartridge case images. In particular, they proposed to extract circle moment invariants (CMIs) from the circle-centralised images as features set after applying a variety of image pre-processing operators such as power-law transformation, Sobel sharpening spatial filter, and Otsu's threshold selection method. The empirical results of identification based on the three-layer backpropagation neural network (BPNN) revealed that the CMIs are efficient and effective to be utilised as the features set for an automated ballistics identification algorithm. However, Chuan ***et al***. [12] authenticated that extracting the features set from the square-window segmented ROI from the firing-pin impression images is superior rather than the circle-window segmented ROI in terms of the identification accuracy rate and the imposed execution time. As a consequence, CMIs do not take into account in this article.

In Malaysia, Ghani ***et al***. [2, 3, 4, 15] proposed a series of semi-automated machine learning identification algorithms for ballistics principally focusing on feature extraction and selection. Ghani ***et al***. [2] have proposed their first machine learning identification algorithm for ballistics by extracting several numerical descriptive statistics measurements such as the first four statistical moments and the measures of position from three distinctive segmented centre-firing pin impression images. These include the segmented centre-firing pin with a half radius of the whole centre-firing pin impression, the segmented ring firing pin, and the segmented whole firing pin impressions. In the late 2000s and early 2010s, Ghani ***et al***. [3,4] proposed another machine learning identification algorithm for ballistics by extracting the geometric moments as the features set. In these articles, they also proposed to extract the features from three distinctive segmented centre-firing pin impression images.

Meanwhile, Liong ***et al***. [20] proposed another semi-automated machine learning identification algorithm for ballistics, which extended the research works from Ghani ***et al***. [2, 3, 4]. In particular, they proposed to extract both numerical descriptive statistics measurements and geometric moments from the three distinctive segmented centre-firing pin impression images. Owing to their study involved the high-dimensional features which comprise 68 statistical numerical features, therefore they have proposed to reduce the feature-dimensionality utilising the principal component analysis (PCA). Among this series of semi-automated machine learning identification algorithms for ballistics, Ghani ***et al***. [2, 3, 4] and Liong ***et al***. [20] proposed to utilise the powerful machine learning identification algorithm, namely Fisher's Linear Discriminant Analysis (LDA) algorithm [21] in performing the ballistics identification tasks.

In contrast, Kamarrudin ***et al***. [11] proposed another semi-automated machine learning identification algorithm with the principal objective to improve the classifier applied in previously proposed identification algorithms in Malaysia. In specific, their research proposed to employ the two-layer feed-forward backpropagation neural network (FBPNN) utilising the tansig-tansig activation function rather than the LDA classifier, while the extracted feature set from the segmented ring firing pin images. A few years later, Ghani ***et al***. [15] extended their research from Ghani ***et al***. [4] by replacing the two-layer FBPNN utilising the tansig-purelin activation function, while the similar selected geometric features set from the segmented firing pin impressions also applied in their proposed improvised semi-automated machine learning identification algorithm. However, the limitation of this series of proposed probabilistic machine learning identification algorithms is these algorithms are remaining required a physical interpretation,

which is to detect the position of the centre-firing pin impressions utilising the naked eye. Furthermore, the extracted numerical features set also does not have invariant properties corresponding to TRS.

Due to the limitation of the machine learning identification algorithm for the ballistics in the previous studies, Chuan ***et al***. [5, 12, 13] have proposed an improvised automated machine learning identification algorithm for the ballistics from the series of research works of Ghani ***et al***. [2, 3, 4] and Liong ***et al***. [20]. In this article, they proposed to extract the selected OLM as a feature set applied in their proposed identification algorithm, in which OLM comprises the translation and scale invariant properties. In their article, they highlighted that the Zernike moment invariants are inappropriate to be employed as these moment invariants imposed high complexity computational costs and were merely applicable for a unit disk, while this research principally focuses on the square-window segmented ROI. In order to authenticate the invariant properties corresponding to the noisy images, Chuan ***et al***. [6] have further tested the proposed automated machine learning identification algorithm based on the simulated noisy images contaminated with random-valued impulse noise with noisy levels as high as 70%. Their empirical analysis results indicated that their proposed automated machine learning identification algorithm is efficient to yield identification accuracy rates of >90% regardless of the investigated noise levels after applying the adequate window size of the median smoothing spatial filter.

Moreover, Razak ***et al***. [14] also proposed another effective automated machine learning identification algorithm for ballistics with an achieved identification accuracy rate is 93%. In this study, they proposed to employ the Canny edge detection operator and CHT as the sharpening spatial filter and centre-firing pin impression position detection technique, respectively. In addition, they proposed to extract selected geometric moments as the feature set for ballistics identification from the segmented firing pin impression. However, the comparative empirical results of the sharpening spatial filter carried out by Chuan [22] conveyed the Laplacian edge detection operator is superior rather than Sobel, Canny [23], and Marr-Hildreth's [24] sharpening spatial filter.

Recently, Liong ***et al***. [25] carried out another research on developing a mobile application for a ballistics identification algorithm. In this study, they have carried out a comparison of the effectiveness of the selected geometric moments feature set, which was also extracted from three distinctive segmented centre-firing pin impression images as the research works of Ghani ***et al***. [2, 3, 4, 15]. Despite the analysis results based on the two-layer FBPNN utilising sigmoid-linear activation function revealed the identification accuracy achieved 98% when extracted the feature set from the segmented centre-firing pin impression. However, the segmentation of the centre-firing pin impression in this article remains required physical interpretation, and the extracted features also do not comprise invariant properties such as TRS. In summary, feature extraction from the key impressions on the fired cartridge case images remains an unresolved issue from the practical perspective. This can be authenticated that distinctive feature sets have been employed in the proposed semi-automated and automated ballistics identification algorithms in the previous studies, and it remains in research for the recent study.

Theoretical Backgrounds

This section provides an overview of the theoretical background for the statistical moments of polynomials for images. In literature, statistical moment invariants extracted as an insightful feature for recognition have been successfully utilised in a variety of pattern recognition applications. These include ship and aircraft identification [26, 27], optical character recognition (OCR) [17, 28], passenger position recognition [29], and ballistics identification [2-5, 10, 12, 13, 14, 20, 25]. Consider $\mathbf{R} = [f(x_4, y_4)]_{r/4 \times r/4};\ x_4, (y_4) = 0, 1, \ldots, {}^{r}/_{4} - 1, ({}^{r}/_{4} - 1)$ represents the ROI segmented from the centre-firing pin impression circular boundary image, $\mathbf{P} = [f(x_3, y_3)]_{r \times r};\ x_3, (y_3) = 0, 1, \ldots, r-1, (r-1)$, which the size of $\mathbf{R}$ is ${}^{r}/_{4} \times {}^{r}/_{4}$ with an intensity value of each pixel, $f \epsilon \left\{ {}^{i}/_{255}; i = 0,1, \ldots, L = 255 \right\}$ located at the coordinate point (x_4, y_4). Therefore, $\mathbf{R}$ can be aggregated in quantitative measurement by using several statistical moments of polynomials for images that comprise invariant properties corresponding to TRS.

In specific, the statistical moments of polynomials for the images detailed in this article include OLM, HM, TMM, PKM, and CGM. In addition, this article merely considered the order of the extracted statistical moment invariants from $\mathbf{R}$ up to the sixth order except for HM (up to the third order). This is because the higher order of statistical moment invariants typically imposed high complexity and is not robust to the noise. Moreover, the best feature set employed in this article is selected based on the classification of the strength's association measurements between extracted features, the stepwise selection technique, and the multicollinearity effects, and tested by employing a variety of sample sizes of the training sets.

Orthogonal Legendre Moment Invariants

There are several well-known orthogonal moments of polynomials for images such as the OLM [30] and Tchebichef moments (TM) [31], which are competent to provide an efficient orthogonal representation and appropriate to extract as a feature set for recognition. This is because the orthogonal moments are more efficient in minimising attribute redundancy, robustness to noise, and comprise the invariant properties corresponding to TRS [30]. However, the preliminary empirical analysis results of this research revealed that OLM is superior rather than TM after the proposed automated machine learning identification algorithm is extracted from both selected OLM and TM provided as features for ballistics identification. In particular, the high identification accuracy rate ($> 90\%$) has resulted when the selected OLM is extracted as a feature set and vice versa for TM ($< 90\%$). As a consequence, TM does not take this into account in this article. In mathematics, the discrete form of OLM of order $(p+q)$, λ_{pq}; $p,(q) = 0, 1, \ldots, 3, (3)$ in the expression of order $(a+b)$ geometric moments, m_{ab} can be defined as follows.

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{4} \sum_{a=0}^{p} \sum_{b=0}^{q} c_{pa} c_{qb} m_{ab} \tag{1}$$

where $m_{ab} = \sum_{x_4=0}^{r/4-1} \sum_{y_4=0}^{r/4-1} \left(\frac{8x_4}{r} - 1\right)^a \left(\frac{8y_4}{r} - 1\right)^b f(x_4, y_4)$, and the coefficients for the Legendre polynomial generating functions, c_{pa} and c_{qb} respectively given as

$$c_{pa} = \begin{cases} \dfrac{(-1)^{(p-a)/2}(p+a)!}{2^p \left(\frac{p-a}{2}\right)! \left(\frac{p+a}{2}\right)! a!}; & (p-a)\ even \\ 0 \qquad\qquad\qquad\qquad\quad ; & (p-a)\ odd \end{cases} \tag{2}$$

$$c_{qb} = \begin{cases} \dfrac{(-1)^{(q-b)/2}(q+b)!}{2^q \left(\frac{q-b}{2}\right)! \left(\frac{q+b}{2}\right)! b!}; & (q-b)\ even \\ 0 \qquad\qquad\qquad\qquad\quad ; & (q-b)\ odd \end{cases} \tag{3}$$

Hu Moment Invariants

The moment's theory has been established after Hu [32] initially proposed seven moments invariants comprise invariant properties corresponding to TRS utilising ordinary geometric moments, which is well-known as HM. Since the crucial region is randomly distributed on segmented $\mathbf{R}$ and this leads to the orientation of the crucial region varying with every acquired centre-firing pin image in this research. Therefore, HM which comprises orientation invariant is also employed to deal with the varying orientation. This article highlighted that there are other three-moment invariants expended from HM (Eqs. (6), (9) and (13)) also taken into account in this study other than the seven origins HM. In particular, HM up to the third order in terms of M_{ab}^* is mathematically defined as follows.

$$R_1 = M_{20}^* + M_{02}^* \tag{4}$$

$$R_2 = (M_{20}^* - M_{02}^*)^2 + 4(M_{11}^*)^2 \tag{5}$$

$$R_3 = M_{20}^* M_{02}^* - (M_{11}^*)^2 \tag{6}$$

$$R_4 = (M_{30}^* - 3M_{12}^*)^2 + (3M_{21}^* - M_{03}^*)^2 \tag{7}$$

$$R_5 = (M_{30}^* + M_{12}^*)^2 + (M_{21}^* + M_{03}^*)^2 \tag{8}$$

$$R_6 = (M_{30}^* M_{03}^*)^2 - 6(M_{30}^* M_{21}^* M_{12}^* M_{03}^*) + 4M_{30}^*(M_{12}^*)^3 + 4M_{03}^*(M_{21}^*)^3 - 3(M_{21}^* M_{12}^*)^2 \tag{9}$$

$$R_7 = (M_{20}^* - M_{02}^*)((M_{30}^* + M_{12}^*)^2 - (M_{21}^* + M_{03}^*)^2) + 4M_{11}^*(M_{30}^* + M_{12}^*)(M_{21}^* + M_{03}^*) \tag{10}$$

$$\begin{aligned} R_8 = {} & (M_{30}^* - 3M_{12}^*)(M_{30}^* + M_{12}^*)((M_{30}^* + M_{12}^*)^2 - 3(M_{21}^* + M_{03}^*)^2) \\ & + (M_{03}^* - 3M_{21}^*)(M_{03}^* + M_{21}^*)((M_{03}^* + M_{21}^*)^2 - 3(M_{12}^* + M_{30}^*)^2) \end{aligned} \tag{11}$$

$$\begin{aligned} R_9 = {} & (M_{03}^* - 3M_{21}^*)(M_{30}^* + M_{12}^*)(3(M_{21}^* + M_{03}^*)^2 - (M_{30}^* + M_{12}^*)^2) \\ & - (M_{30}^* - 3M_{12}^*)(M_{03}^* + M_{21}^*)((M_{03}^* + M_{21}^*)^2 - 3(M_{12}^* + M_{30}^*)^2) \end{aligned} \tag{12}$$

$$R_{10} = (M_{30}^* M_{03}^* - M_{21}^* M_{12}^*)^2 - 4(M_{03}^* M_{12}^* - (M_{21}^*)^2)(M_{30}^* M_{21}^* - (M_{12}^*)^2) \tag{13}$$

where the ordinary geometric moments of order $(a+b)$, M_{ab}^* is given as

$$M_{ab}^* = \sum_{x_4=0}^{r/4-1} \sum_{y_4=0}^{r/4-1} x_4^a y_4^b f(x_4, y_4) \tag{14}$$

Tsirikolias-Mertzois Moment Invariants

In general, TMM is a two-dimensional orthogonal moment, which is a modified form of the one-dimensional standardised moments that normalises corresponding to the standard deviation. Likewise, HM, TMM proposed by Tsirikolias and Mertazios [17] is another two-dimensional statistical orthogonal moment that comprises invariant properties corresponding to TRS. Moreover, TMM is also less sensitive to noise. Consequently, a better ballistics identification performance could have resulted. In statistical theory, the general form of TMM of order $(p+q)$ can be expressed in a mathematical equation, such that:

$$\gamma_{pq} = \frac{16}{r^2} \sum_{x_4=0}^{r/4-1} \sum_{y_4=0}^{r/4-1} \left(\frac{x_4 - \mu_{x_4}}{\sigma_{x_4}} \right)^p \left(\frac{y_4 - \mu_{y_4}}{\sigma_{y_4}} \right)^q f(x_4, y_4) \tag{15}$$

where $\mu_{x_4} = \frac{M^*_{10}}{M^*_{00}}$ and $\mu_{y_4} = \frac{M^*_{01}}{M^*_{00}}$ represent the coordinates of the image's centroid of the segmented $\mathbf{R}$ respectively corresponding to coordinates x_4 and y_4. Meanwhile, the standard deviations of the segmented $\mathbf{R}$ respectively correspond to coordinates x_4 and y_4.

$$\sigma_{x_4} = \sqrt{\frac{16}{r^2} \sum_{x_4=0}^{r/4-1} \sum_{y_4=0}^{r/4-1} (x_4 - \mu_{x_4})^2 f(x_4, y_4)} \tag{16}$$

$$\sigma_{y_4} = \sqrt{\frac{16}{r^2} \sum_{x_4=0}^{r/4-1} \sum_{y_4=0}^{r/4-1} (y_4 - \mu_{y_4})^2 f(x_4, y_4)} \tag{17}$$

Pan-Keane Moment Invariants

In statistical theory, PKMs are orthogonal moments which are invariant corresponding to scale or size. These moments are proposed by Pan and Keane [33] initially for optical character recognition (OCR). However, the principal reason of PKMs employed in this study is the comparative effectiveness with other orthogonal moment invariants due to the size of the segmented $\mathbf{R}$ and the crucial region is both not the same for each image. Therefore, PKMs are advantageous in this condition. In mathematics, PKMs of order $(p+q)$ in terms of T_{pq} (Eq. (19)) can be expressed as Eq. (18).

$$S_{pq} = \frac{T_{00}^{(p+q+2)/2}}{T_{20}^{(p+1)/2} T_{02}^{(q+1)/2}} T_{pq} \tag{18}$$

Central Geometric Moment Invariants

The limitation of the extraction of geometric moments as the feature set in the previously proposed machine learning identification algorithms for ballistics [2, 4, 11, 14, 15, 20, 25] is non-orthogonal moments so that the extracted moments do not comprise the invariant properties corresponding to TRS. To overcome this limitation, this study proposed to employ central moments, also known as central geometric moments (CGM) as the comparison. From a statistics perspective, CGMs are the statistical moments that comprise invariant properties corresponding to translation or position. In this research, the position of the crucial region on $\mathbf{R}$ varies with every acquired centre-firing pin image despite that the cartridge cases are fired by utilising the same firearm. Therefore, CGM is very useful in shifting the crucial region such that its centroid coincides with the origin of the coordinate system. Mathematically, the CGM of order $(p+q)$, T_{pq} can be expressed as follows.

$$T_{pq} = \sum_{x_4=0}^{r/4-1} \sum_{y_4=0}^{r/4-1} (x_4 - \mu_{x_4})^p (y_4 - \mu_{y_4})^q f(x_4, y_4) \tag{19}$$

where μ_{x_4} and μ_{y_4} represent the coordinates of the image's centroid of the segmented $\mathbf{R}$ respectively corresponding to coordinates x_4 and y_4, which M^*_{ab} is defined in Eq. (14).

The Architecture of the Proposed Automated Probabilistics Machine Learning Identification Algorithm

This section provided the schematic procedures to implement the proposed automated probabilistic machine learning identification algorithms for ballistics. The following are the procedures for architecting the proposed identification algorithm tailored based on Figure 2.

STEP 1: Input 125 centre-firing pin impression circular boundary images (training set) (Figure 1-Right) with an equal sample size for each class of pistol taken into account in this study. Mathematically, the input image can be expressed in array form as in Eq. (20).

$$\mathbf{P_0} = [f(x_0, y_0)]_{M\times N};\ x_0, (y_0) = 0, 1, \dots, M, (N) \tag{20}$$

of size $M \times N$ with intensity value of each pixel, $f \epsilon \left\{ {}^{i}/_{255}; i = 0,1, \dots, L = 255 \right\}$, located at the coordinate point (x_0, y_0).

STEP 2: Enhance the edge of image $\mathbf{P}$ using a Laplacian sharpening spatial filter. The Laplacian sharpening spatial filter is selected in this research since this operator presented the best empirical result after comparing it with other operators such as Sobel, Canny, and Marr Hildreth sharpening spatial filters. Specifically, the empirical result is evaluated based on the ballistics identification accuracy rates and the execution time. Consider $\mathbf{M_L}$ represents the kernel of the Laplacian sharpening spatial filter, and $\mathbf{S_p}$ represents the sub-image for the image $\mathbf{P}$, which $\mathbf{M_L}$ and $\mathbf{S_P}$ can be defined in the following array, respectively.

$$\mathbf{M_L} = \begin{bmatrix} 0 & +1 & 0 \\ +1 & -4 & +1 \\ 0 & +1 & 0 \end{bmatrix} \tag{21}$$

$$\mathbf{S_P} = \begin{bmatrix} f(x-1, y+1) & f(x+0, y+1) & f(x+1, y+1) \\ f(x-1, y+0) & f(x+0, y+0) & f(x+1, y+0) \\ f(x-1, y-1) & f(x+0, y-1) & f(x+1, y-1) \end{bmatrix} \tag{22}$$

A new image, $\mathbf{P_L} = [f_{\mathbf{L}}(x_2, y_2)]_{(M-2)\times(N-2)};\ x_2, (y_2) = 1, 2, \dots, M-1, (N-1)$ has resulted after the convolution process completed, where the intensity value of each pixel, $f_{\mathbf{L}}$ is computed based on Eq. (23).

$$f_{\mathbf{L}}(x_2, y_2) = \sum_{x_2-1, y_2-1}^{x_2+1, y_2+1} 255 vec(\mathbf{M_L})' vec(\mathbf{S_P}) \tag{23}$$

where $f_{\mathbf{L}} \in \{-\infty, \dots, -1, 0, 1, \dots, \infty\}$, and $vec(\cdot)$ represents the vectorisation function.

STEP 3: The resulting $f_{\mathbf{L}}$ in **STEP 2** is not within the desired range of $[0, 1]$. As a result, there is a requirement to scale the range of $f_{\mathbf{L}}$ using mix-max normalisation. In mathematical, the function can be utilised to normalise the intensity value of each pixel, $f_{\mathbf{N}}$ can be expressed as

$$f_{\mathbf{N}}(x_2, y_2) = \frac{1}{255}\left\lfloor \frac{f_{\mathbf{L}}(x_2,\ y_2) - min\ \mathbf{P}_{f_{\mathbf{L}}}}{max\ \mathbf{P}_{f_{\mathbf{L}}} - min\ \mathbf{P}_{f_{\mathbf{L}}}} \times 255 \right\rfloor \tag{24}$$

where $\lfloor \cdot \rfloor$ represents the floor function. Meanwhile, $min\ \mathbf{P}_{f_{\mathbf{L}}}$ and $max\ \mathbf{P}_{f_{\mathbf{L}}}$ represent the minimum and maximum intensity values of each pixel for the image $\mathbf{P_L}$. Consequently, a normalised image, $\mathbf{P_N} = [f_{\mathbf{N}}(x_3, y_3)]_{(M-2)\times(N-2)};\ x_3, (y_3) = 1, \dots, M-2, (N-2)$ is resulted.

STEP 4: Otsu [34] threshold selection method is employed to binarise the image $\mathbf{P_N}$. Otsu's method is employed in this study because this method is widely and successfully utilised in multidisciplinary applications [5,6,12,13,22,35,36,37] even though this threshold method has been introduced over four decades ago. An optimal threshold value, $\tilde{w}$, is attained by maximising the function defined in Eq. (25). Consequently, a binary image, $\mathbf{P_B} = [f_{\mathbf{B}}(x_3, y_3)]_{(M-2)\times(N-2)}$, is resulted in intensity values "1" when $f_{\mathbf{N}} \geq \tilde{w}$ and "0" when $f_{\mathbf{N}} < \tilde{w}$.

$$\tilde{w} = \frac{1}{255}\arg\left(\max_{w}\left((\mu_L - \mu_w)^2\left(\frac{\pi_w}{1-\pi_w}\right)\right)\right) \tag{25}$$

where $\mu_L = \sum_{\theta=0}^{L} 255\theta k_{\mathbf{N}}(\theta)$, $\mu_w = \sum_{\theta=0}^{w} 255\theta k_{\mathbf{N}}(\theta)$, $\pi_w = \sum_{\theta=0}^{w} k_{\mathbf{N}}(\theta)$, $k_{\mathbf{N}}(\theta) = \sum_{x_3, y_3}^{M-2, N-2} I(f_{\mathbf{N}}(x_3, y_3) = 255\theta) \Big/ (M-2)(N-2)$, and $I(\cdot)$ represents the indicator functions.

STEP 5: This study is compared to the proposed automated probabilistic machine learning identification algorithm which is applied with the weighted least square estimator proposed by Albano [38] and the unweighted least square estimator proposed by Moura and Kitney [18]. The preliminary analysis results indicated the proposed identification algorithms which

are applied with an unweighted least square estimator are more effective rather than the identification algorithms which are applied with a weighted least square estimator in terms of ballistics identification accuracy rate and execution time. As a result, an unweighted least square estimator has been employed in research in order to estimate the coordinate of the anchor point, $\mathbf{A}^T = [X_\mathbf{A}, Y_\mathbf{A}]$ and radius, r, for centre-firing pin impression circular boundary (Figure 1-Right) based on Eqs. (26) and (27), respectively.

$$\mathbf{A} = \begin{bmatrix} 2(\alpha\beta_{20} - \beta_{10}^2\beta_{00}) & 2(\alpha\beta_{11} - \beta_{10}\beta_{01}) \\ 2(\alpha\beta_{11} - \beta_{10}\beta_{01}) & 2(\alpha\beta_{02} - \beta_{01}^2\beta_{00}) \end{bmatrix}^{-1} \times \begin{bmatrix} \alpha(\beta_{12} + \beta_{30}) - \beta_{10}(\beta_{20} + \beta_{02}) \\ \alpha(\beta_{21} + \beta_{03}) - \beta_{01}(\beta_{20} + \beta_{02}) \end{bmatrix} \tag{26}$$

$$r = \sqrt{\frac{1}{(M-1)(N-2)} \sum_{x_3, y_3} ((X_\mathbf{A} - x_3)^2 + (Y_\mathbf{A} - y_3)^2)} \tag{27}$$

where $\alpha = \sum_{x_3, y_3} f_\mathbf{B}(x_3, y_3)$ and $\beta_{uv} = \sum_{x_3, y_3} x_3^u y_3^v f_\mathbf{B}(x_3, y_3)$.

STEP 6: Segment the $\mathbf{R}$ by utilising the estimated $\mathbf{A}$ and r from **STEP 5**.

STEP 7: Extract the OLM as features with values of p and q respectively up to the order of 3. In this research, the low-order moment invariants have been selected rather than the high-order moment invariants. This is due to the low-order moments imposing low mathematical complexity and more stability to noise [19].

STEP 8: Conduct the correlation analysis to obtain the Pearson correlation coefficient among the OLM features set extracted from $\mathbf{R}$.

STEP 9: Categorise the pairs of the extracted OLM features set corresponding to the following classification of strength association measurements.

Weak association classification	$: 0.0 \leq \lvert\hat{\rho}\rvert < 0.3$
Moderate association classification	$: 0.3 \leq \lvert\hat{\rho}\rvert < 0.7$
Strong association classification	$: 0.7 \leq \lvert\hat{\rho}\rvert < 1.0$

STEP 10: Selects a set of informative features for the weak classification of strength association measurements by employing the stepwise selection technique and multicollinearity effects measured in terms of tolerance (< 0.1). The selected informative features set required does not meet the multicollinearity condition due to the multicollinearity effects being able to decrease the statistical efficiency.

STEP 11: Trains a supervised paradigm of probabilistic machine learning model, namely Fisher's Linear Discriminant Analysis (LDA) [21] model based on the selected informative features in **STEP 10**.

STEP 12: Identify the classes of the unknown pistols (The unselected centre-firing pin images in **STEP 1** which is known as the test set) using the trained LDA model in **STEP 11**.

STEP 13: Summarise the identification accuracy rate in a confusion matrix. In addition, record the imposed computational time in running the proposed identification algorithm.

STEP 14: Compute the average of the identification accuracy rate among the classes of pistols diagonally based on the resulting confusion matrix in **STEP 13**.

STEP 15: Repeat **STEPS 1-14** for the distinctive sample sizes for the training sets, classification of strength association measurements and the statistical moments invariant employed in this study.

Results and Discussion

There are 747 centre-firing pin impression images utilised in this study. These images were collected from five pistols model of Parabellum Vektor SP1 9 mm, namely Pistol A (150 images), Pistol B (150 images), Pistol C (150 images), Pistol D (149 images) and Pistol E (148 images). The five pistols employed in this study are selected by the judicatory ballistic expert of the Royal Malaysia Police and comprise similar age, caliber, made and model so it is difficult to distinguish among these five pistols. In addition, the pistols model of Parabellum Vektor SP1 9 mm is employed in this study because this is the pistol model typically utilised by criminals in Malaysia.

This study has allocated the acquired centre-firing pin images into a training set and a test set. The principal objective training set is utilised in this study to train the classifiers, whereas the test set is utilised to evaluate the efficiency of the proposed identification algorithm. A total of 100 images from each pistol

are randomly selected and assigned to the training set while the unselected images are assigned to the test set. In particular, four distinctive sample sizes of training sets such as 125, 250, 375 and 500 images, which these images are selected randomly from the 500 images assigned for training purposes with equivalent frequency among the five pistols. This article highlighted that the low ratio of the training set (sample size of 125 images) rather than the test set is also taken into account in this study to consolidate the effectiveness of the extracted moment invariants as features set despite that the LDA model trained using a small sample size of the training set. The identification accuracy rates and the execution time are shown in Table 1 and Table 2.

Table 1. Ballistics identification accuracy rates (in %) correspond to the weak, moderate, and strong classification of strength association measurements among the selected features by utilising the test set

Moment Invariants	The Number of Features Extracted	The Sample Size of the Training Set	Ballistics Identification Accuracy Rate		
			Weak	Moderate	Strong
OLM	16	125	**94.7(9)** [F1]	**94.7 (9)** [F1]	NA
		250	95.5 (13)	**96.0 (11)**	NA
		375	**96.0 (9)**	93.5 (9)	NA
		500	**96.8 (9)**	95.5 (9)	NA
HM	10	125	**71.7 (4)**	66.0 (3) [F2]	66.0 (3) [F2]
		250	**75.7 (7)**	73.3 (7) [F3]	66.4 (4)
		375	**77.3 (8)**	73.3 (7) [F3]	61.1 (5)
		500	**77.7 (8)**	73.3 (7) [F3]	69.6 (5)
TMM	12	125	**95.5 (7)**	92.3 (5)	92.7 (6) [F4]
		250	**95.1 (7)**	89.9 (5)	93.5 (8)
		375	**95.5 (8)**	92.3 (7)	92.7 (6) [F4]
		500	**94.3 (8)**	88.7 (7)	**94.3 (8)**
PKM	13	125	93.5 (6)	**95.5 (7)**	87.0 (4)
		250	**93.9 (7)** [F5]	**93.9 (7)** [F5]	89.5 (6) [F6]
		375	**93.9 (7)** [F7]	**93.9 (7)** [F7]	89.5 (6) [F6]
		500	**93.9 (8)** [F8]	**93.9 (8)** [F8]	89.5 (6) [F6]
CGM	14	125	**96.8 (7)** [F9]	**96.8 (7)** [F9]	**96.8 (7)**
		250	**97.2 (8)**	96.4 (7)	**97.2 (10)** [F10]
		375	**97.6 (10)**	95.5 (7)	96.8 (8)
		500	**98.0 (10)**	**98.0 (10)**	97.2 (10) [F10]

Note: The superscript "**F1-F10**" means that the selected feature set is the same for those sharing the same number in the superscript.

Table 2. The execution time (in seconds) correspond to the weak, moderate and strong classification of strength association measurements among the selected features by utilising the test set

Moment Invariants	The Number of Features Extracted	The Sample Size of the Training Set	Ballistics Identification Accuracy Rate		
			Weak	Moderate	Strong
OLM	16	125	82.0 [F1]	82.0 [F1]	NA
		250	86.8	84.8	NA
		375	82.3	82.2	NA
		500	82.3	82.5	NA
HM	10	125	78.1	74.1 [F2]	74.1 [F2]
		250	79.5	79.7 [F3]	78.3
		375	82.8	79.7 [F3]	78.4
		500	82.8	79.7 [F3]	78.4
TMM	12	125	80.4	77.1	78.8 [F4]
		250	79.8	77.4	81.1
		375	81.1	79.7	78.8 [F4]

Moment Invariants	The Number of Features Extracted	The Sample Size of the Training Set	Ballistics Identification Accuracy Rate		
			Weak	Moderate	Strong
PKM	13	500	80.9	79.6	81.2
		125	83.3	84.4	81.1
		250	84.7 ^{F5}	84.7 ^{F5}	83.9 ^{F6}
		375	84.2 ^{F7}	84.2 ^{F7}	83.9 ^{F6}
CGM	14	500	85.9 ^{F8}	85.9 ^{F8}	83.9 ^{F6}
		125	82.7 ^{F9}	82.7 ^{F9}	82.1
		250	83.6	82.3	85.6 ^{F10}
		375	85.7	82.1	83.5
		500	85.8	85.9	85.6 ^{F10}

Note: The superscript "**F1**-**F10**" means that the selected feature set is the same for those sharing the same number in the superscript.

Based on Table 1, the empirical analysis results indicated that for the weak classification of strength association measurements, the accuracy rate of the proposed automated probabilistic machine learning identification algorithm employing OLM, TMM, PKM and CGM is $> 90\%$ for all sample sizes of the training sets taken into account in this study, whereas the ballistics identification accuracy rates for the algorithm extracting HM is $< 80\%$. Meanwhile, for the moderate classification of strength association measurements, the identification accuracy rates of the proposed identification algorithm extracted OLM, PKM and CGM as the feature set are $> 90\%$ for all sample sizes of training sets and vice versa for the HM $(< 80\%)$. Conversely, the ballistics identification accuracy rates for the strong classification of strength association measurements of OLM are not available (NA). This is due to no pair of these extracted features existing, for which the Pearson correlation coefficient is within the range of $0.7 \leq |\hat{\rho}| < 1.0$. Moreover, Table 1 also illustrates that on average, the larger the sample size employed for the training set, the higher the ballistics identification accuracy rate for all classifications of strength association measurements. This empirical analysis results are consistent and can be consolidated based on the machine learning perspective.

On the other hand, the larger the number of features utilised in the classifier, the longer the required execution time. This statement is proven by the empirical analysis results of the execution time for the proposed identification algorithm depicted in Table 2. In particular, it can be observed that empirical analysis results of LDA with a larger number of selected features impose a longer execution time rather than a small number of selected features. In addition, the type of features extracted from $\mathbf{R}$ generally affects the execution time. Based on these theories, the type and the number of features utilised in the algorithm are factors which influence the execution time in order to perform the proposed identification algorithms. However, the empirical analysis results in Table 2 are unsupported by the aforementioned theories. In other words, there is not much difference between the execution times of the proposed identification algorithm when both factors are taken into account simultaneously.

According to Leng and Huang [10], the limitation of the semi-automated probabilistic machine learning identification algorithm of Ghani ***et al.*** [3,4,15] is ignoring the rotation of the circular firing pin impressions. Contrarily, the empirical analysis results of this research conveyed that the rotation of the centre-firing pin impression circular boundary plays a minor role in the proposed automated probabilistic machine learning identification algorithm in this study. In particular, the ballistics identification accuracy rates for the proposed identification algorithm that extracted HM as features for identification are frequently $< 80\%$ for all classification of strength association measurements and all sample sizes of training sets taken into account in this study. These lower identification accuracy rates are due to the characteristic of HM which is invariant corresponding to rotation. In fact, the identification accuracy rates for these identification algorithms are the lowest compared to OLM, TMM, PKM and CGM. As a result, this study concluded that the HM is inappropriate to be employed as features set in the proposed identification algorithm due to the low identification accuracy rates.

Conversely, Table 1 revealed that the identification accuracy rates for the proposed identification algorithm employed CGM as a feature set resulted in the highest for all classification of strength association measurements and all sample sizes of training sets rather than OLM, HM, TMM, and PKM. Among the three classifications of strength association measurements of CGM, the feature set selected from the weak classification of strength association measurements is superior rather than the moderate and strong classifications. Furthermore, Table 2 also presented that there is not much difference in

execution times among the classification of strength association measurements and all sample sizes of the training set for CGM. Therefore, this study concludes that the proposed automated probabilistic machine learning identification algorithm for ballistics which extracted the selected CGM as an informative feature set from the weak classification of strength association measurements is the superior identification algorithm rather than other classifications of strength association measurements and moments invariant taking into account this study.

Conclusions

In summary, there are two principal objectives in this article, namely to develop an improvised automated probabilistic machine learning identification algorithm for ballistic that utilises the unweighted least square-fitting circle to detect the position (anchor point) of the centre-firing pin impression circular border with radius r, and to identify the most appropriate statistical moment invariants for images extracted from the segmented ROI corresponding to the centre-firing pin impression images. In order to pursue these objectives, a total of 747 centre-firing pin impression images from five pistols of model Parabellum Vektor SP1 9mm are employed to test the proposed identification algorithm in distinctive scenarios. These include the distinctive classification of strength of association measurements and sample sizes of the training sets. The empirical results of this study revealed that the proposed probabilistic machine learning identification algorithm for ballistic which extracted a selected informative CGM based on weak classification of strength association measurements is superior with identification accuracy rates $> 96.5\%$ for all sample sizes of training set rather than the identification algorithm which extracted other moment invariants taking into account in this study. Moreover, this study also revealed that the proposed identification algorithm which extracted features of moment invariants merely comprise the rotation invariant properties is insufficient and ineffective in ballistics identification. Conversely, extracted features of moments invariant merely comprise the translation or position invariant properties that are superior in ballistics identification. The superior automated probabilistic machine learning identification algorithm for ballistics proposed in this study could be beneficial for the authority such as judicatory ballistics experts in Malaysia for ballistics identification. This is due to the proposed identification algorithm does not require physical interpretation and competently completing an identification task ranging from seconds to minutes. This study suggests extending this work in future by employing deep learning approaches such as artificial neural network (ANN) and convolutional neural network (CNN) approaches.

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Acknowledgement

The authors would like to acknowledge the Forensic Laboratory of the Royal Malaysia Police (RMP) for providing the equipment and assistance in data collection. Grateful appreciation is extended to the Ministry of Higher Education (MOHE) for providing Research Grant No. FRGS/1/2019/STG06/UPM/02/7 and UKM-ST-06-FRGS0183-2010. A word of appreciation also repeatedly goes to the Ministry of Higher Education for the scholarship of financial support belonging to the MyBrain15 program.

References

[1] Xin, L. -P., Zhou J., G. Rong. (2000). A cartridge identification system for firearms authentication. *Proceedings of the 5th International Conference on Signal Processing,* IEEE Press, Beijing, China, 1405-1408. https://doi.org/10.1109/icosp.2000.891807.

[2] N. A. M. Ghani, C. -Y. Liong, A. A. Jemain. (2009). Extraction and selection of basic statistical features for forensic ballistic specimen identification. *Sains Malaysiana*, *38*(2) 249-260.

[3] N. A. M. Ghani, C. -Y. Liong, A. A. Jemain. (2009). Extraction and selection of geometric moment features for forensic ballistic specimen identification. *Matematika*, *25*, 15-30. https://doi.org/10.11113/matematika.v25.n.256.

[4] N. A. M. Ghani, C. -Y. Liong, A. A. Jemain. (2010). Analysis of geometric moments as features for firearm identification. *Forensic Science International, 198*(1-3), 143-149. https://doi.org/10.1016/j.forsciint.2010.02.011.

[5] Z. L. Chuan, A. A. Jemain, C. -Y. Liong, N. A. M. Ghani, Effectiveness of cross-entropy and Tsallis entropy thresholding for automatic forensic ballistics identification system. *Journal of Quality Measurement and Analysis, 9*(1), (2013) 33-46.

[6] Z. L. Chuan, A. A. Jemain, C. -Y. Liong, N. A. M. Ghani, L. K. Tan. (2017). A robust firearm identification algorithm of forensic ballistics specimens. *Journal of Physics: Conference Series*, *890,* 012126. https://doi.org/10.1088/1742-6596/890/1/012126.

[7] C. L. Smith, D. Li. (2008). Intelligent imaging of forensic ballistics specimens for ID. *Proceedings of the 2008 Congress on Image and Signal Processing*, IEEE Press, Sanya, China, 37-41. https://doi.org/10.1109/CISP.2008.760.

[8] D. G. Li. (2003). Image processing for the positive identification of ballistics specimens. *Proceedings of the 6th International Conference on Information Fusion*, IEEE Press, Queensland, Australia, 1494-1498. https://doi.org/10.1109/ICIF.2003.177417.

[9] J. Zhou, L. -P. Xin, D. -S. Gao, C. -S. Zhang, D. Zhang. (2001). Automated identification for firearms authentication. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Press, Kauai, USA, 749-754. https://doi.org/10.1109/CVPR.2001.990551.

[10] J. Leng, Z. Huang. (2012). On analysis of circle moments and texture features for cartridge images recognition. *Expert Systems with Applications*, *39*(2), 2092-2101. https://doi.org/10.1016/j.eswa.2011.08.003.

[11] S. B. A. Kamaruddin, N. A. M. Ghani, C. -Y. Liong, A. A. Jemain. (2012). Firearm classification using neural networks on ring of firing pin impression images. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 1*(3), 27-34. https://doi.org/10.14201/ADCAIJ20121312734.

[12] Z. L. Chuan, N. A. M. Ghani, C. -Y. Liong, A. A. Jemain. (2013). Automatic anchor point detection approach for firearms firing pin impression. *Sains Malaysiana*, *42*(9), 1339-1344.

[13] Z. L. Chuan, C. -Y. Liong, A. A. Jemain, N. A. M. Ghani. (2014). An efficient automatic firearm identification system. *AIP Conference Proceedings*, *1602*(1), 1185-1189. http://dx.doi.org/10.1063/1.4882634.

[14] N. A. Razak, C. -Y. Liong, A. A. Jemain, N. A. M. Ghani, S. Zakaria, H. Sulaiman. (2017). Firing pin impression segmentation using Canny edge detection operator and Hough Transform. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(1) 23-26.

[15] N. A. M. Ghani, C. -Y. Liong, A. A. Jemain. (2018). Neurocomputing approach for firearm identification. *Pertanika Journal of Science & Technology*, *26*(1), 341-352.

[16] R. C. Gonzalez, R. E. Woods. (1993). *Digital image processing*, 3rd edn. Pearson Prentice Hall.

[17] K. Tsirikolias, B. G. Mertzios. (2017). Statistical pattern recognition using efficient two-dimensional moments with applications to character recognition. *Pattern Recognition*, *26*(6), 877-882. https://doi.org/10.1016/0031-3203(93)90053-Y.

[18] L. Moura, R. Kitney. (1991). A direct method for least-squares circle fitting. *Computer Physics Communications*, *64*(1), 57-63. https://doi.org/10.1016/0010-4655(91)90049-Q.

[19] J. Flusser, T. Suk, B. Zitová. (2009). *Moments and moment invariants in pattern recognition*. John Wiley and Sons.

[20] C. -Y. Liong, N. A. M., Ghani, S. B. A. Kamaruddin, A. A. Jemain. (2012). Firearm classification based on numerical features of the firing pin impression. *Procedia Computer Science*, *13*, 144-151. https://doi.org/10.1016/j.procs.2012.09.123.

[21] R. A. Fisher. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics, 7*(2) 179-188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

[22] Z. L. Chuan. (2014). Statistical firearm identification for forensic ballistics. Universiti Kebangsaan Malaysia. Unpublished Doctoral Thesis.

[23] J. Canny. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8, 6*, 679-698. https://doi.org/10.1109/TPAMI.1986.4767851.

[24] D. Marr, E. Hildreth. (1980). Theory of edge detection. *Proceedings of the Royal Society B, 207*(1167), 187-217. https://doi.org/10.1098/rspb.1980.0020.

[25] C. -Y. Liong, N. A. M. Ghani, S. B. A. Kamaruddin, A. A. Jemain. (2020). Conceptual design of firearm identification mobile application (FIMA). *AIP Conference Proceedings*, *2266,* 090014. https://doi.org/10.1063/5.0018445.

[26] F. W. Smith, M. H. Wright. (1971). Automatic ship photo interpretation by the method of moments. *IEEE Transactions on Computers*, *C-20*(9), 1089-1095. https://doi.org/10.1109/T-C.1971.223408.

[27] S. A. Dudani, K. J. Breeding, R. B. McGhee. (1977). Aircraft identification by moment invariants. *IEEE Transactions on Computers*, *C-26*(1), 39-46. https://doi.org/10.1109/TC.1977.5009272.

[28] Y. C. Chim, A. A. Kassim, Y. Ibrahim. (1999). Character recognition using statistical moments. *Image and Vision Computing*, *17*(3-4), 299-307. https://doi.org/10.1016/S0262-8856(98)00110-3.

[29] C. -Y. Liong, N. A. M. Ghani, A. A. Jemain, C. Thompson. (2012). Momen ortogon Legendre sebagai suatu fitur untuk pengecaman kedudukan penumpang, *Jurnal Teknologi, 48*(1), 41-58. https://doi.org/10.11113/jt.v48.224.

[30] P. Kaur, H. S. Pannu, A. K. Malhi. (2020). Comprehensive study of continuous orthogonal moments-a systematic review. *ACM Computing Surveys*, *52*(4), 1-30. https://doi.org/10.1145/3331167,

[31] C. Camacho-Bello, J. S. Rivera-Lopez. (2018). Some computational aspects of Tchebichef moments for higher orders. *Pattern Recognition Letters*, *112*, 332-339. https://doi.org/10.1016/j.patrec.2018.08.020.

[32] M. -K. Hu. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, *8*(2), 179-187. https://doi.org/10.1109/TIT.1962.1057692.

[33] F. Pan, M. Keane. (1994). A new set of moment invariants for handwritten numeral recognition. *Proceedings of the 1st International Conference on Image Processing,* IEEE Press, Texas, USA, 154-158. https://doi.org/10.1109/ICIP.1994.413294.

[34] N. Otsu. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics, 9*(1), 62-66. https://doi.org/10.1109/TSMC.1979.4310076.

[35] D. Indra, T. Hasanuddin, R. Satra, N. R. Wibowo. (2018). Eggs detection using Otsu thresholding method. *Proceedings of the 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT),* IEEE Press, Makassar, Indonesia, 10-13. https://doi.org/10.1109/EIConCIT.2018.8878517.

[36] C. V. V. S. Srinivas, M. V. R. V. Prasad, M. Sirisha. (2020). Remote sensing image segmentation using OTSU algorithm. *International Journal of Computer Application*, *178*(12), 46-50.
[37] M. Sholihin, M. Arif, M. H. Alfansury, N. M. Yuzi, Sumijan. (2019). Identification of palm using Otsu method and mathematical morphology to open house doors. *Jurnal KomtekInfo*, *7*(2), 101-109. https://doi.org/10.35134/komtekinfo.v7i2.70.
[38] A. Albano. (1974). Representation of digitized contours in terms of conic arcs and straight-line segments, *Computer Graphics and Image Processing*, *3*(1), 23-33. https://doi.org/10.1016/0146-664X(74)90008-2.