<u>15<sup>th</sup> May 2020. Vol.98. No 09</u> © 2005 – ongoing JATIT & LLS

ISSN: 1992-8645

www.jatit.org



# TEXT-INDEPENDENT CHINESE WRITER IDENTIFICATION USING HYBRID SLT-LBP FEATURE

<sup>1</sup>GLORIA JENIS TAN, <sup>2</sup>ROSELY KUMOI, <sup>2</sup>MOHD SHAFRY MOHD RAHIM,

<sup>3</sup>TAN CHI WEE, <sup>4</sup>GHAZALI SULONG

<sup>1</sup>Centre for IT Development and Services, Universiti Malaysia Sarawak (UNIMAS), Sarawak, Malaysia

<sup>2</sup>Faculty of Engineering, School of Computing, Universiti Teknologi Malaysia (UTM), Johor, Malaysia

<sup>3</sup>Faculty of Computing and Information Technology, Tunku Abdul Rahman University College, Malaysia.

<sup>4</sup>Management & Science University (MSU), Malaysia

E-mail: <sup>1</sup>jtgloria@gmail.com

## ABSTRACT

This study proposes a new hybrid method using texture features of input handwriting document image as global to overcome the limitation of data heterogeneity, which causing the ambiguity and leads to inconsistent results apart from problems of scale involve database size. The method first adopts Slantlet Transform (SLT) to bring out hidden texture details prior to feature extractions. Then, Local Binary Pattern (LBP) descriptor is applied on the SLT image to extract texture features. A new hybrid method Slantlet Transform based Local Binary Pattern (SLT-LBP), are experimented on an open and widely used HIT-MW Chinese database for performance evaluation. This study strengthens the idea that to unravel some of data heterogeneity and lead to improve identification performance, especially searching for relevant document from large complex repositories is an essential issue.

#### Keywords: Chinese Handwriting, Local Binary Pattern, Slantlet Transform, Text-independent, Texture, Writer Identification

## 1. INTRODUCTION

The history of research on off-line textindependent identification of the writer dates decades back but is still an active issue [1][2][3][4][5][6][7][8][9]. In details. textindependent is not limited by text content. It has received increased attention in recent years on extracting writing style features from global writing text, where a writer's handwriting is regarded as a texture. This global approach is based on texture analysis, commonly used and accepted method in practical applications. Therefore, in the case of text-independent approach, different handwriting images are considered as different textures. However, text-independent is a very challenging task [10][11].

Extensive research has led in this field due to its importance in forensic analysis and documents authorization [1][12][13][14][15][16][17]. The problem becomes more difficult especially in case of criminal investigation where writers are not determinable because did not having access to databases to train the classifier in off-line mode [18].

There has been a great deal of effort input by several researchers in improving the writer identification techniques. A comprehensive review on writer identification techniques is given in [19] [20]. It is observed that during that period, there are significant progresses achieved on English and Arabic. However, the growth on Chinese is rather slow and far from satisfactory in comparison to its wide usage [21]. Few of the recent techniques, stated for all languages, performed ambiguously when tried on different languages. In addition to the challenges presented by characteristics of different language scripts, data size negatively affects the identification rate.

Traditional method segments [22][23][24][25][26] texts into small square windows to model writing by small strokes as opposed to graphemes. The methods in this category mainly differ in how the handwriting is segmented into graphemes and how the graphemes are clustered. Graphemes and window-based

15th May 2020. Vol.98. No 09 © 2005 - ongoing JATIT & LLS



www.jatit.org

methods represents the writing samples for writer identification which produced a codebook of the fundamental units of graphemes. Current method involve codebook is to find and generate all documents features, then by comparing the feature vector distance between query and library image, this performance however, proven to effectively perform although in this context, window-based extraction calls for a tedious, challenging from a size-adjustable sliding window and the selection of directly window size affect identification performance. Undoubtedly, many achievements have been made and only focused on identification performance on this very subject but a major problem with this kind of traditional method is to search for the relevant document from large complex document image repositories.

Thus, this study focuses off-line text-independent writer identification in Chinese language. The proposed method is based on our previously presented work [27]. The method first adopts Slantlet Transform (SLT) to bring out hidden texture details prior to feature extractions. This paper presents an approach using SLT proposed by [33], which includes a new parameter orthogonal to Discrete Wavelet Transforms (DWT) for writer identification. It is an equivalent form of DWT with two-zero moments and better time localization. SLT as a filter bank is implemented in a parallel structure which is more time efficient in comparison with iterative approach of DWT.

Then, Local Binary Pattern (LBP) descriptor is applied on the SLT-transformed image or SLT image to extract texture features. LBP that was first introduced by [29] and has been quickly gained considerable attention since its publication [30] is applied to extract texture features. LBP is a local operator which discriminates different types of textures. It has been shown to be an effective descriptor in texture classification and less computational complexity. This is a robust method and does not being effected by rotation or noise in the image.

The rest of paper is organized as follows: Section 2, mainly discusses the proposed approach. Experimental results are presented in Section 3. Finally, conclusions and future work are presented in Section 4.

# 2. IMPLEMENTATION

We propose the use of Slantlet Transform based Local Binary Pattern (SLT-LBP), a textural based approach for off-line text independent Chinese writer identification. Figure 1 illustrates the whole writer identification phase.

The methods described in this chapter are based on heuristics strategies, which means to discover previous researches with similar problems and problem solving based on trial and error. It is characterized by repeated, varied attempts which are continued until the quality of the solution obtained.

The method mentioned in the flowchart involves four components: (1) Pre-processing steps in Section 2.1; (2) Partitioning of image in Section 2.2; (3) Image Decomposition using Slantlet Transform (SLT) in Section 2.3; (4) Computation of Local Binary Pattern (LBP) in Section 2.4.

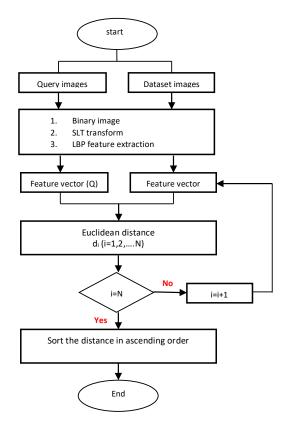


Figure 1: A proposed method block diagram

<u>15<sup>th</sup> May 2020. Vol.98. No 09</u> © 2005 – ongoing JATIT & LLS

ISSN: 1992-8645

www.jatit.org

#### 2.1 Pre-processing

Pre-processing involves the removal of noise and elimination of empty spaces from images before subjecting them to feature extraction. Since we work on various size to make the method become size independent of contemporary Chinese texts and the images are already binarized, the preprocessing in our case simply comprises removal of noise and the elimination of empty spaces. The final processing image produces different size of images depend on elimination of empty space.

#### 2.2 Partitioning of Images

In this stage, the width of dimension noise removed image is divided into two halves images which stored in a separate folder for feature extraction in next step as illustrated in Figure 2 for writer ID 241.

也留过 庸 年迁年7 文礼 反礼 4 收孔认佳区 幼幼礼佳灾肝 收了白玉 肘下运孔收孔的 £ 3 存 种如鲠在喉 不上不快的感见引通过 t# 支(建 孔门好 时节 这类电视∫丘好像見 埋 日人们 uba 1 10.n 反腐力反 1 #1 台口目相+的电化媒体 足片月 老白姓生 建队电化 好好上厂 近1 将る 内石等価 迎り 刮铘R

(a) Image to be partitioned of sample writer 241

「也想过管隅 收礼认住这 隅 收礼认住这 隅 收礼认住这 隅 收礼 5 斤 延礼收礼 小此 5 斤 种加翅在 邮 每是 + 女 (延 礼)门好 讨 鳞露了同人的 理 要是; 5 叶 从中来 」他们 反相 ; 5 5 元 在 经 生 奇 白 日相 使才是 华节将至 建以电 计 髓 内有导师 斑	年过年7 文化 段礼 前 二 50 功 物 礼 隹 変屏 7 土 7 快約 町 紀 7 通 节 足 本 电 化 「 二 好 像 是 9 國 人 这 「 二 戶 予 仟 、 周 方 反 1 晋 [ 力 + 1 行 八 ・ 内 电 化 媒体 史 F 円 . 礼 好 好 上 「 夏 川 「 劇 膠 民
(1) T <sup>1</sup> (1, 10)	() 0 11 10

(b) First half

(c) Second half

Figure 2: Image is divided into two halves for writer ID 241

#### 2.3 Image Decomposition using Slantlet Transform

In this study, the Slantlet Transform (SLT) is applied to decompose the binary images in order to bring out its texture details prior to feature extractions. The SLT algorithm is applied to decompose the image.

An algorithm called SLT proposed by [33], which includes a new parameter orthogonal to Discrete Wavelet Transforms (DWT). It is an equivalent form of DWT with two-zero moments and better time localization. SLT as a filter-bank is implemented in a parallel structure which is more time efficient in comparison with iterative approach of DWT. SLT filter banks involved are low pass filter hi(n), adjacent of low pass filter, fi(n) and remaining filter, gi(n). As shown in SLT formula, the size of SLT matrix ranging from 2×2 to 256×256 pixels. The process initiates with determining coefficient filter i.e. gi(n), fi(n) and hi(n) to obtain new SLT filter matrix. In this study, the size of Slantlet matrix is empirically chosen after a series of experiments performed with various sized in stages,  $32 \times 32$ ,  $64 \times 64$ ,  $128 \times 128$  and  $256 \times 256$ . Here, the coefficients of the filters are given, gi(n), fi(n) and hi(n) with filter size of 128x128 gives the best result.

In order to quickly grasp the SLT process, the SLT matrix operation with filter size of  $2 \times 2$  is obtained as shown in Figure 3. Following that, a new Slantlet image matrix named SLTimage is generated using the SLTfilter, image block, and SLTfilterT are multiplied as shown in Figure 4.

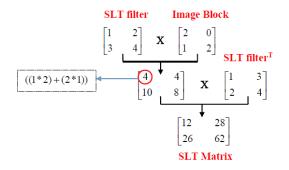


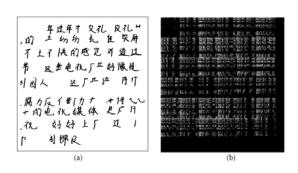
Figure 3: The SLT matrix operation

www.jatit.org

<u>15<sup>th</sup> May 2020. Vol.98. No 09</u> © 2005 – ongoing JATIT & LLS



E-ISSN: 1817-3195



ISSN: 1992-8645

Figure 4: Second half image of sample writer 241; (a) Partitioning image of 827x857 pixels, (b) Images with 128x128 size of Slantlet matrix

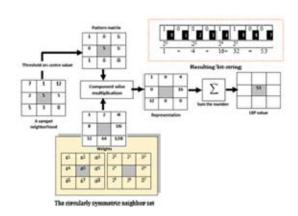


Figure 5: An example on how a LBP values has been generated

#### 2.4 Local Binary Pattern (LBP) Features Computation

In Then, Local Binary Pattern (LBP) descriptor is performed to extract texture feature on SLT image. The feature extraction process represents each handwriting image. These LBP texture features are used to compare handwriting image between training and other from testing image.

The LBP is an operator that was first introduced by [29] and has been quickly gained considerable attention since its publication [30]. It has been shown to be an effective descriptor in texture classification and less computational complexity. Local Binary Pattern (LBP) is a local operator which discriminates different types of textures. The original LBP operator [33] defines a label LBP code of each pixel of an image. This is a robust method and does not get effected by rotation or noise in the image. LBP is dense local texture descriptor that can be used to describe the local structure of images [30].

In order to quickly grasp the above process, let's work on a numerical example by giving an example to show how a SLT matrix of a LBP values has been generated is given in Figure 5. An example is given to show how the LBP is computed. The original LBP operator is a  $3\times3$ -pixel block. At a given pixel position in the image, the LBP is defined as an ordered set of binary comparator of pixel intensities between the centre pixel and its 8 neighbours to create an image of integer valued code, then pooling these codes into histograms.

In this study, we used  $LBP_{8,1}$  to obtain different values of unique local binary patterns which are independent of rotation. The  $LBP_{8,1}$ represent the occurrence statistics of the patterns and corresponds to certain features in the image. Thus, the binary patterns of unique rotation invariant can be considered as feature of texture LBP. Finally, a set of features is extracted. A feature vector describing the textual properties of the image is then obtained from a histogram of the LBP values of the image which illustrated in Figure 6.



Figure 6: LBP values has been generated in the sample writer 241 image with dimension 827x857 pixels

# Journal of Theoretical and Applied Information Technology

15th May 2020. Vol.98. No 09 © 2005 - ongoing JATIT & LLS

ISSN: 1992-8645

www.jatit.org

experiments [32]. This database contains natural handwriting images of 241 different writers. It contains 853 images of handwriting samples, out of which 254 images are labelled with 241 writer IDs. Among 241 writers, most of these writers have onepage writing sample, 10 writers contributed at least two pages and the remaining writers contributed more than two pages each. There is variety of sample size for each writer. The distribution of number of samples per writer is illustrated in Figure 8, while a sample form from the database is shown in Figure 9.

These feature vectors of query image and dataset images are then arranged as rows of two separate matrices, the matrix is generated. It is a row vector of 1 row and 256 columns. These feature vectors are then arranged into a matrix to determine the similar handwriting images. An LBP histogram is computed independently for each handwriting image. In previous process, before the colour of image is being inverted, all the white spaces are removed. This causing the processed image tend to have a very large region of white pixels (255) and produce mode of 255 in graph as illustrated in the rightmost of Figure 7. Then, all the resulting histograms are concatenated together into a single vector.

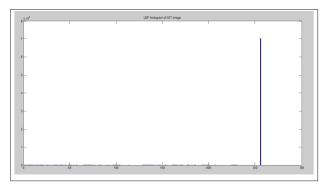


Figure 7 LBP histogram for sample writer 241 image with dimension 827x857 pixels

The next step entails the coarse matching between two images to provide an initial result. The Euclidean distances are used for matching between query and database image.

## 2.5 Creation of a Shortlist of Matching Images

After the textures were extracted from both images, query and dataset, they are concatenated into two distinct feature vectors, Q (Q1, Q2, ..., Qn) and Gk (G1, G2, ..., Gn) where k=1, 2, ..., N. Where Q represents feature vector of query image; Gk denotes feature vector of the kth image of the Afterward, for each dataset image, a dataset. distance is computed using Expression 1 to measure the difference between the image and the query image. The range of the distance is [0,1] – Where "0" indicates a perfect match, while "1" reflects a total stranger - The greater distance means the less the similarity. The Euclidean Distance is calculated as follows:

$$d'(x_1, x_2) = \sqrt{\sum_{i=1}^{i=1} (x_i(i) - x_i(i))^2}$$

where, x1 and x2 are the two feature vectors of both images

Once all the distances (d1, d2, ..., dN) have been computed for all images of the dataset, they are then sorted in ascending order.

#### **3. DISCUSSION AND EXPERIMENTAL EVALUATIONS**

This study focuses on accuracy to measure the quality of having high accuracy and consistency of the proposed method to the Chinese language. Experiments are carried out by using the abovedescribed method on HIT-MW Chinese dataset.

To quickly grasp the performance evaluation, the identification performance is evaluated based on of Top-N criterion. It is checked if Top-N documents are written by same writer and depends on the number of documents from the same writer as that of the query document in the dataset. Here, Top-1, Top-5 and Top-10 of identification rates is reported in this study. Top-1 means that writer of the query document is matched with the first ranked sample in the sorted list. Similarly, Top-10 means the query document is similar within the top 10 writers retrieved by the system.

## 3.1 HIT-MW Chinese Database

The experimental evaluations mainly conducted on the HIT-MW database HIT stand for Harbin Institute of Technology; MW stand for Multiple Writers [31]. The HIT-MW Chinese database which is based on 300dpi images of Chinese handwritten text documents is used in the



(1)

<u>15<sup>th</sup> May 2020. Vol.98. No 09</u> © 2005 – ongoing JATIT & LLS

ISSN: 1992-8645

www.jatit.org

E-ISSN: 1817-3195

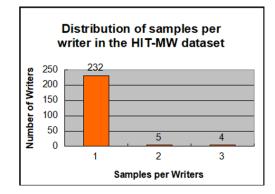


Figure 8: Distribution of samples per writer in the HIT-MW dataset

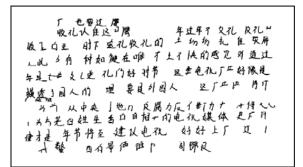


Figure 9: Sample writer ID 000241

# 3.2 Evaluation Criterion of Identification

The methods described are based on heuristics strategies, which means to discover previous researches with similar problems and problem solving based on trial and error. It is characterized by repeated, varied attempts which are continued until the quality of the solution obtained.

A series of experiment has been conducted by increasing the number of writers based on most of the previous studies are focused on the structurebased approaches extract features from handwriting images which involve segmentation and window selection. Therefore, a method without this step is proposed, thus making it independent of segmentation and window size. The performance of SLT-LBP based on matrix size. The proposed method using texture features of input handwriting document image as global or whole image to overcome the limitation. The identification performance for various database sizes and Top-N values are summarized in Table 1.

 Table 1: The accuracy rate (in %) for identification

 respect to the number of writer.

# of writers	SLT-LBP features		i
_	Top1	Top5	Top10
50	96	98	100
100	93	99	99
150	87.33	98.66	99.33
200	88	99	99
241	84.64	98.75	98.75

Based on Table 1, it is clearly snows that the identification performance over the entire textual handwriting image degrades with the increase in database size. Naturally, these results indicate that the writer identification without any segmentation on the query image and feature extraction for entire image as global is not reliable and hence downgrade the identification performance. This is the reason where most of existing structure-based are based on the contours or the allograph fragments of handwriting too depend on segmentation level process which involve images modification to fulfil the requirement with a certain amount of characters to aim good result in certain language. Hence, it leads ambiguous result when applied to another language with different of characteristics that can be missing in the classification.

In comparison, it can be noticed that the identification rate is good in hybrid SLT-LBP method. In this method, an experiment is conducted in Table 2 based on SLT matrix size and the best result accuracy is 84.64% with matrix size of 128x128.

© 2005 – ongoing JATIT & LLS

#### ISSN: 1992-8645

#### www.jatit.org

E-ISSN: 1817-3195

 Table 2: The accuracy rate (in %) for SLT-LBP in different SLT matrix size.

SLT matrix size	SLT-LBP features		
	Top1	Top5	Top10
32x32	82.57	98.34	98.75
64x64	79.25	98.34	98.75
128x128	84.64	98.75	98.75
256x256	78.83	98.75	99.17

Figure 10 illustrate the retrieved images using Euclidean Distance, in which here query image to be used is writer ID 241 was found and arbitrary chosen from the dataset. Retrieved images from dataset is arranged according to their similarity distance with the query document – the 3rd top left corner writer on the first row is matched with query document writer ID 241.

	生态 展了 中菜 村子 中書 展了 中菜书 心城 中音 及明寺 村子	電子に変は、前日前後 したし、税利、参約7月7月 打すたして、役利不利 打すたして、役利不利 打すたして、役利不利 行きする人での総定利(昭元 行きする人での総定利(昭元 行きする)、(本利) 行きの(名見)、	キオキィ 交化 P化 の よ かめ ち 生 気 気 オ よりめ ち 生 気 気 寿 オ よえ 人 れ の 死 え 可 逸 立 ち まま き ム デ = か 引 男 ム ま き ム デ = か 引 頭 ん ( す) テ ・ 村 へ 一 の 色 へ 縄 朱 足 デ 川 下 引 得良 Writer ID 241
1日、11 2年1月日 1日、11 2月1日 1日、11日、11日 1日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日 1日、11日、11日、11日、11日 1日、11日、11日、11日、11日 1日、11日、11日、11日、11日 1日、11日、11日、11日、11日 1日、11日、11日、11日、11日 1日、11日、11日、11日、11日、11日、11日 1日、11日、11日、11日、11日、11日、11日 1日、11日、11日、11日、11日、11日、11日 1日、11日、11日、11日、11日、11日、11日、11日、11日 1日、11日、11日、11日、11日、11日、11日、11日、11日、11日、	用 节目 4 起 勢 把化一甲 中方林美国化 法 死 中方水 美国 教育 同天空 大学 大学 野型的。 小 水 マーム 大 美 中心、 赤型 电池市田里的初期 : Writer ID 3	オ羽殺 5 夏をもらまま り み変 大家大湖 努力学 すけ んしけつれるだう メオオ ふわみか エ大心 き 内 但 したの によう 同 計解 ドート か たり 近 える テ むからび 二入地に 心にのあんう葉 菜 ジェッ Writer ID 217	$\{\frac{1}{2}, \frac{1}{2}, $
275200	1年史東東市新加生的安地。 1年史東東北市北京大学の天 18日本1日、北京大学の大学校 18日本1日、北京大学校 18日本1日、北京大学校 18日本1日、北京大学校 18日本1日、北京大学校 18日本1日、北京大学校 18日本1日、北京大学校 18日本1日 18日本11日 18日本111日 18日本111日 18日本11日 18日本11日 18日本11日 18日本11日 18日本11日 18日本11	15日 1時1 1日 5日 5日 用金 5 1 (1) 2 1 1 1 5 5 5 5 1 1 1 1 1 1 1 1 1 1 1 1	常校三篇の11时 大変 (1日本) は、 1時後又 金の原料 れえたかけ第年のよどが 日本の1 のの成本 日本の1 のの成本 とこの1支 スムの決失 許 前 の行び1支 行のよりあい。 、 工業には、 ので、よっご参加して、 して、よ、二部 他職本 での「参加して、 して、 しまの一支 スムの決失 許 前 の行び1支 スムの決失 許 前 の行び1支 スムの決失 許 前 の行いため、 のに、 のに、 のに、 のに、 のに、 のに、 のに、 のに

はキシ、 2月 9 7 8人 8連 第 成 大軍司百 4見 2 1 / 日表見 2 4 37 昭 2017 (日東京 4 37 昭 2017) (1) 2 4 4 1 1 2 2 3 4 4 1 2 3 4 1 1 2 3 4 1 1 2 3 4 1 2 3 4 1 1 2 3 4 1 1 2 3 4 1 1 2 3 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	12日 8 方面2 11月 日本的 林長 トッシー 3 天和 1 将 頃春 な高低計石 又城 2 移行者 字訳 2月1天山 2月6日 字訳 2月2日 2月2日 字文 美 夏 長 久 御心又 2011月 注 11号者 近月 数据文字	后来前从这些 结果见上在如白 于赤虾 酸蛇蛇病 和防止 重大类酸 从此行变又 丹 350 下足 转声。向泉子 24 下注 该约驾建人王 的,朋友 了的平安之子门,打罢夫子 累 盘 5 猪蛋 267 年月 年 展 库现出人是对面发	自关石行中达里了合称 电压 → 工与 辅助计估协论 阈 場合人 医诊察产展 正元并 理由 目覺禮 司 在广告合用基本进行任于 监 一时成分化 → 各 为等部路 建了 时成分化 → 各 为等部路 建了 时成分化 → 各 为等部路 与公子领导 一 无点 建式种物产 种文件 发展 展示 此无病 计 例相词
き数 お大 司火等型」研究の 法史司達 近計し上上1月のか	星的本内产用置的 十年九	以美和拉约为睦为分 包人美 建的内球质处理	X X X X X X X X X X X X X X X X X X X
Writer ID 177	Writer ID 45	Writer ID 112	Writer ID 51
は快 な まま」を日 「後 秋 前 慶 勤 石 う よ 、 」 う 保 勇 本 料 局 の 海 ト や ひ 、 ひ 「 胃 満 昇 料 局 の 海 ト や 丘 、 値 町 約 名 名 欠 内 伯 ( 打 み く 自 膠 力 久 氏 水 が か 車 思 ち 筆 予 ス 入気 水 所 ま 、 足 引 末 物 冠 も 夫 + 5 1	有如此的推行完全有关重要之的 ) 数字电缆 大学 安克酸 一致子电缆 大学 安克酸 ) 第二 生命 使于 安克酸 ) 第二 生命 使于 安克酸 ) 第二 生命 使于 安克酸 ) 第二 生命 使于 医子脑 一致 化学 化 这面 2 经基金 数字电子 特别历史 2 经基金 数字电子 特别分离完美 学生的 文面 数字电子 特别分离完美 学生的 文面 和学校研究者和	電史 章 ヘリスキキュキ 人 可有本市営業など大大大会にか、 可有本市営業など大大大会にか、 同者本市営業など研究に 展 が 首になった に 取得考え、の以下市 注 大大ス なご 引 」の豊思キ 目) 市 (1人人、 聖人 3) 町人人	件 主脑体使用 网络 K 3 用手 支金 加速力 F 20 分段 建以用石灰取成 用于 30 力扩力的标志 带式型 以只要加点 非清热 多只料力 精的过去式 40000 有过 40000 元 4000 有过 40000 元 4000 有过 40000 有达 40000 有达 40000 有达 40000 有达 40000 40000 40000 500000 500000 500000 500000 500000 500000 500000 500000 5000000 5000000 500000000
Writer ID 62	Writer ID 11	Writer ID 198	Writer ID 92
東直 (北南河)是由智」比 为? 司友之的建建,各小田尾行 - 句 / 起 步 / な為代田尾行 - 句 / 起 步 / な為代田尾行 - 句 読田(了此を武法者) 是司/ 反帮企理记并售簿 - 序 / 百 司/ 村田是居/前 イン是 / 年3司速表全直 書 了民会的社会 次尾	1.45 的工作,让邻分分割两式 1.450 年末 1.24 分支制两式 1.450 年末 1.430 大部分前的原始 天主 有值 - 21 元汇标做的 2.4 有值 - 21 元汇标做的 2.4 有值 - 21 元 1.25 计 直载3.47 文 2.5 27 应 开和文女的启行 ·壁文计应开示段交上 ()的关键	▲村市上 15 〒11月125日 15 〒11月15日 15 〒11月15日 15 〒11日 15 日 15 日	మ సెహ్రెస్ సిర్మాహం రై. కె. సిహ్రెస్ సిర్మా ) విష్ణిళి లెకెంగ్ సెక్రిస్తునే? ఉ. సై న గ్రోష్ లెకెంగ్ సెక్రిస్తునే? ఉ. సై న దిని ప్రతిశిశ్వం సంక్షి కి. సిర్మాహం కి సిర్మా సిక్రి సిక్రి సిక్రి సిక్రి సిక్రి సిక్రి పిరి సిక్రి పిని సిక్రి సిక్రి సిక్రి సిక్రి సిక్రి సిర్మా స్త్రి సిక్రి సిక్రి సిక్రి సిక్రి సిక్రి సిర్మా స్త్రి సిర్మా స్త్రి సిర్మా స్త్రి సిర్మా స్త్రి సిర్మా స్త్రి సిర్మా స్త్రి సిర్మా స్త్రి సిర్మా స్త్రి సిర్మా స్త్రి సిక్రి స్త్రి సిక్రి స్త్రి సిక్రి స్త్రి సిక్రి స్త్రి సిక్రి స్త్రి సిక్రి స్త్రి సిక్రి సిక్రి స్త్రి సిక్రి సిక్రి స్త్రి సిక్ సిక్రి సిక్ సిక్రి సిక్రి సిక్రి సిక్ సిక్రి సిక్రి సిక్రి సిక్రి సిక్ర
Writer ID 146	Writer ID 183	Writer ID 82	Writer ID 190
(1) 2013年、1013年、20135755555555555555555555555555555555555	2 下生在一个小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小	· 21番約 壁 現場之 人内 21番229	管里园是在水水水、 建国家产生品的大学生。 "你这个人们的一个人们的一个人们的一个人们的一个人们的一个人们的一个人们的一个人们的一
Writer ID 201	Writer ID 165	Writer ID 220	Writer ID 76
林病、1約人-2等後点、 「生生素」の一人で、後点、 「生」、「生」、「ない」、 「ない」、「ない」、 「ない」、「ない」、 「、 「ない」、 「、 「、 「、 「、 「、 「、 「」、 「、 「」、 「、 「」、 「、 「」、 「、 「」、 「、 「、 「、 「、 「」、 「、 「、 「、 「、 「、 「、 」、 「、 「、 「、 」、 「、 「」、 「、 「」、 「、 「、 「、 「」、 「、 「」、 「、 「、 「、 「、 「、 」、 「、 「、 「、 」、 「、 「、 「、 」、 「、 「、 「、 」、 「、 」、 「、 「、 」、 「、 「、 」、 「、 「、 「、 」、 「、 」、 「、 」、 「、 」、 「、 」、 「、 」、 「」、 「	■増加加20万 かん間容量 該。」後かですかかな 印炭 オ美皮人美術神古松設成 ら 命更に (かく部人部) 更硬経者的人の小文利代景 2夏下2、 K-+5 40 支船)	1 基础全接定 在对人广的门 王基留对 200 搜讨命 201 指数 12	3五 工程即電量五 東土山雅社展記術商 酸と打球化的加速 高は異处 水市石肥の後の 高大美物に小野型
	Writer ID 43	Writer ID 69	Writer ID 143
「蜀土丸 青米 起送不一个海点 例 引大、国际上が开放码 、3 定量前用炉 用 仁武 使 电化电 机晶晶的 动 成正生产五份 子面的 市局 市面 点 脚股上 众 例使中 电子子 发 面 二例建成中44 ℃ 「 代 五	聖總員了二人 一個一個人 一個一一一一一一一一	b/结款瓶夏 扇制新人 直度特虑痛精50度之1 局勢 給命度起今 30度 1 局壓 內邊行 煙的 10度 1 能展需求的了/依赖進 賣大司有天濟有石油200 另別所能源 得熱層低 十 1 納的 費加 引 2 厚印 1 日耗。由得达到	1.170 ~ 平力 節 天正中目 え血ア 州幕 来自巴西书 え取府机物境部列州大 石田七工 建氟化 男子 2.20月机物境部列州大 石田七工 建氟化 男子 2.20月机物境部 2.40月 一代中 进 一十打大两国 1.2011 印公十月 1.2011 印公十月 1.2011 印公十月 1.2011 印公 1.2011 印公 1.2011 印公 1.2011 印公 1.2011 印公 1.2011 印公 1.2011 印 1.2011 日 1.2011 印 1.2011 日 1.2011 日 1.20
Writer ID 6	Writer ID 216	Writer ID 109	Writer ID 66

ISSN: 1992-8645

www.jatit.org



10.3 科力量超其現一进的 圣子 如命中日王弟的"磨吧" 田昌 了 书林田桥 他是整洁着在水场。 为十分自然草方所提出了 3月66 大雄千天 桥水门 他 有书水处置 温静 天和地 和石窟 不知的 展览服之里是不凡 如此的指 行的情况之间是很为 引起了 天 形刻 (3年天) 是 天 引 美 ) 第五元 国家名法	1 積冊 英文商台镇 推出 2 雷龙频带 为 研络灰 村 巴 局站焊 通川博 七上 为名 7 外客床 亡 回知 雄立属 半月 天 回 的 市结晶 / 博 天 回 和 社社 上 小 家 医一种 短过 上 小 家 医一种 短过 上 小 家 医一种 一种 人 了 二 他 人 多足	284 30 14724-4-4 4997242224978 4997242224978 4967242729 566745729 567454 4967454 4967454 4967454 4967454 496746 49746 497666 497666 497666 4976666 4976666666666	2.36m以洛 将甲羟基6 (2) 将种野豆的 读了 (2) 将种野豆的 读了 (2) 米林等于豆的 读了 可能处理文理 也 形中均信 形中均层 2) 彩红小学业的 3) 彩红小学业的 4) 丁生生的有大量他 户超发文理
Writer ID 205	Writer ID 18	Writer ID 13	Writer ID 235
144年之一 大工程( 185年5月二次年代) 花 同日,4月20年月月 花 同日,4月20年月月 7 次,3月10日年月 年 支援後年5年月 10日,10日 金麗(1995年1月10日) 金麗(1995年1月11日) 金麗(1995年1月) 金麗(1995 1月) 金麗(1995年1月) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金麗(1995 1) 金 (1995 1) 金 (1995 1) 金 (1995 1) 金 (1995 1) 金 (1995 1) 〇 (1995 1) (1995	1.450 (2.46) · 2. (1.45) (1.46) (1.16) · 人 5 (2.9) (1.46) (1.46) · 人 5 (2.9) (1.46) (1.46) · 人 5 (2.9) (1.46) (1.47) · 人 4 (2.9) (1.47) (1.47) (1.47) · 人 4 (2.9) (1.47)		▲ 代表 F 走 1 10年 似戶 1定 上化 推动是为下外展的 1版 全县人大件長 受及下的 并 8位 更的前级下 这些 原则 提 医盐强 50 开放 久 1 代於各位 又重成 ス 1 代於各位 又重成 ス 1 代於各位 3 一年 (1 一年)
Writer ID 36	Writer ID 228	Writer ID 166	Writer ID 172
本美賀連今十方香2 40 2×建村美 小寿 50 40 1基文 山寿 50 第1年 第時 日本 240 年 12 24 連合 第十月 22 24 連合 第十月 24 24 連合 第十月 24 24 第一日 24 第一日 25 第 25	内部 急22.07分的小は 大 現美しく 小銀 20日 お わず快行 一次第四色二次 主次用があぶて知道、 との用する21万 になったした。 との目前の21万 になったした。 との目前の25万 になったした。 と、25日の10年のたかに、 ヴタ 力差別34人の指統101前	台 钟品女 经计 习代验告 繁荣 他 永远的名 雨 当前的 张永兴人子 2 5 36 其他 节点 即不不到一个 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	1. 1443-1442-1443-1443-1443-1443-1443-1443-
TTL: TD AL			
Writer ID 31	Writer ID 106	Writer ID 50	Writer ID 72
Writer ID 31	Writer ID 106 一 本市 竹正 配付 持命者が大田 (中本) 5年、一からまたをか 6年、一からまたをか 5年、一からまたをか 5年、一からまたを 5年、「小」、「茶一枚」 1月 10 - 10 1日 10 10 10 10 1	T这样的的名词复 有了研究和特别。 电视卡华 年轻 克观卡华 年轻 克罗东拍空母校 资本采档空母校 专了想的问题。 专了想的感觉,它看了 有一句意念。2016	Writer ID 72
(14-1) 25 57 5 25 57 1 5 5 25 57 1 5 5 2 5 57 1 5	1600001から1500000 160001から15000 たまっから30000 たまっから30000 たまいたい 25000000000 25000000000 250000000000	T这样的約8.60度 有了研查相關很 免犯計率 年轻 免犯事实推制接 資業後,他的客用 接 资業後,他的客用 考 5 累裕 在心候部代 支援 有勢水 奮了 所 何意 怎 2 恐始 日本 Writer ID 25	日十月 电槽风下口 时间有从地递到35天 地员 仁庵九唐大晨时 在广告化工业 再也 用 长户的 是了上 冬天巴开 25、岩境,39、山人来哈拉
(14-1) 25 かす。 通知) 送うかす。 通知) たい たまたす。 こうかす。 こうかす。 たったい ですれる でする でする でする でする でする でする でする です	1600001から11500000 160001から1150000 たました。 たまか、たたりまたまた。 たまか、たたりまたまた。 第回したのできたらく気管 2巻くすーム。 一部であたる、 100001000000000 100000000000000000 100000000	一支援 03 約8 約3 変 有了 07 会 相信 像 电视十坐 年轻 免犯要女推測接 資準尿 林的客用 按 5 架 招迎公報校 季 打腳的過 全行 5 新 好意 名 2 20 H 司 士 Writer ID 25 中秋時勤 差育 括力 了 5 新 所 向 雨 岸 行 迎 2 記載年 同 柳 時 1 回見載末年 1 / 妙 开 妙	智士 2 世際 凍 行上 日前 御水港 編明 5 天 地 定 一 地 思 地 現 王 本 思 思 北 思 思 王 孝 金 元 王 李 朱 肥 开 名 安 地 市 朱 十 李 朱 肥 开 名 安 地 市 永 平 山 林 永 4 5 9 地 3 9 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5
「かっか」ではるメスト 通いう」となっても、全 通いう」となっても、全 通いう」となっても、 たったいではるように、 たったいではるようになった。 たったいではるようになった。 たったいではるようになった。 たったいではるようになった。 たったいではるようにはない。 だったいではなったい。 ではたいではなったい。 ではたいではなったい。 ではたいではなったい。 ではたいではなったい。 ではたいではなったい。 ではたいではなったい。 ではたいではないではない。 ではたいではないではない。 ではたいではないではない。 ではたいではないではない。 ではないではないではない。 ではないではないではない。 ではないではないではない。 ではないではないではない。 ではないではないではない。 ではないではないではないではない。 ではないではないではないではない。 ではないではないではないではないではない。 ではないではないではないではないではない。 ではないではないではないではないではないではないではないではないではないではない		一世球的局容的宽 有了研查 相關係 电视行坐 年轻 免视行坐 年轻 免犯者女推測接 育軍家, 桃间客用 接 言軍家, 桃间客用 考 5 累裕 在心夜都低 专丁]期際局 之后 1. 前 何意 念 2 犯 續 引 走 Writer ID 25 中球的局部 是算 括於 3 1. 超時代書 時本為進 1. 超時代書 時本 1. 超時代書 時本 1. 通 一方 表示內面配為任 3 一方 表示內面配為錄 5 1. 利田 和 5 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	211 ( 11) 司以忠邀词(35) 建文 生原意正序(表) 定するした。 51 (20)
「小小小」は5 メコム 通知 通道 がする。 通知 近 えんにころもまれ。 通知 たいたころもまれ。 建 る で、それのないなん まったいそれないなん。 パット いうですれ、 なん パット しまいなんなん パット しまいたんなん ないたいました。 なんし、のおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのおい なん、のよいのよい なん、のよいのおい なん、のよいのよい なん、のよいのない なん、たい なん、たい なん、たい なん、のたい なん、たい なん、たい なん、のま、 なん、のよいのない なん、のよいのない なん、たい なん、たい なん、たい なん、たい なん、たい なん、たい なん、のまい なん、 なん、のまい なん、のま、 なん、のまい なん、		一世祥 63 約8 60)夏 有了 67点 相信很 免況十坐 年轻 免況十坐 年轻 免況 年 4 62 完 4 7 2 2 4 7 2 5 8 約 2 2 2 4 7 2 2 4 4 5 8 約 2 2 2 4 4 5 8 約 2 2 2 4 4 3 ± Writer ID 25 (4 5 6 6 4 4 3 3 1 2 5 4 5 1 5 1 5 1 5 1 5 1 5 1 5 1 5 1 5 1	(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)

近草有效 投票最终统计符 远。没有透出代表例代表的 29至 代的成两名1×36/进 深组纪和10岁 国家 底量的 2 47名6 我男别观超名对投
--

Writer ID 128

#### Figure 10: Retrieved images from dataset matched with the query image writer ID 241

#### 3.3 Discussion

Overall, these results can obtain positive results but its accuracy is not as high as expected and still have rooms for improvements. Despite the above described significance of this study, it must understand that it will not confine only to the enrichment of knowledge. It is self-evident of potentially capable of practical applications.

A series of experiments has been conducted by increasing the number of writers based on most of the previous studies are focused on the structure-based approaches extract features handwriting images which from involve segmentation and window size selection. It requires segmentation algorithm which may lead to wrong results in case of improper segmentation. Various window sizes lead to ambiguous results and various database sizes obvious that decrease the identification performance degrades with the increase in database size. Such a method segments texts into small square windows has been proven to effectively perform although in this context, window-based extraction calls for a tedious, challenging from a size-adjustable sliding window and the selection of window size directly affect identification performance.

Thus, making a comparison of this study with others would be appropriate due to the different experimental procedures, handwriting used, data collection process and the use of real offline handwritten word lead to ambiguous results. Manually intensive techniques are utilized by most handwriting identification expert nowadays. Due to unique of each language scripts, the performance of writer identification from different languages strictly depend the selection window size or codebook size, thus, posing a new challenge to the writer identification. Data heterogeneity and human interpret ability problems are there, which is involved in most approaches and manipulation of a constant window size or codebook size is

performed ambiguously. The types of error made by machine and human are quite different due to the fundamental differences between automatic and manual methods.

Therefore, a method without this step is proposed, thus making it independent of segmentation and window size selection. Greater efforts to use standard datasets are needed to ensure the evaluation among different researchers can be comparable and not lead to ambiguous results.

Considering the above issues, the research outcomes are expected to introduce hybrid method Slantlet Transform based Local Binary Pattern (SLT-LBP). SLT Transform to extract texture features by using SLT descriptor for images to represent the handwritten samples rather than using the whole writing. The different of this LBP method is computed reference images from SLT instead of original images which have variety of images dimension.

To make a contribution in achieving this goal by proposing reliable methods that can one size fit all any handwritten language especially for three major world language and not tailored to individual needs depend on language characteristic in future. It is the scenario of a real-world and can be effectively and efficiently fused in existing methods for all languages instead of being specific to a certain language. Added to this, the potential of size insensitive method can be explored for expanding its applicability to multiple languages, expanding the experiment to prove that proposed method is language invariant. This area, however, still has large room for research which can be taken by upcoming researchers.

In conclusion, expectation of this work would help researchers whose work is dedicated to writer identification and the related problems, specifically those working in document analysis and handwriting recognition.

# 3. CONCLUSION

The main contribution of this approach is that it is based on a new hybrid method to extract texture features using Slantlet Transform based Local Binary Pattern (SLT-LBP). The contribution of this study is that it highlights the importance of a size-independent writer identification mechanism which is can corroborate real-world application on a large image repository. The experimental results of proposed method were satisfactory and can be compared with earlier approaches. Future work focuses on testing on different handwriting databases, and real-world implementation.

# ACKNOWLEDGEMENT

This research under Vot No. R. J130000.7851.5F116 is managed by Research Management Centre (RMC), University of Technology Malaysia (UTM) and Ministry of Higher Education.

# **REFRENCES:**

- L. Schomaker, M. Bulacu, and S. Member, "Connected-Component Contours and Edge-Based Features of Uppercase Western Script," Pattern Anal. Mach. Intell. IEEE Trans., vol. 26, no. 6, pp. 787–798, 2004.
- [2] H. Cao, R. Prasad, and P. Natarajan, "Improvements in HMM adaptation for handwriting recognition using writer identification and duration adaptation," Proc. -12th Int. Conf. Front. Handwrit. Recognition, ICFHR 2010, pp. 154–159, 2010.
- [3] J. Chen, D. Lopresti, and E. Kavallieratou, "The impact of ruling lines on writer identification," Proc. - 12th Int. Conf. Front. Handwrit. Recognition, ICFHR 2010, pp. 439–444, 2010.
- [4] G. Louloudis, B. Gatos, N. Stamatopoulos, and a. Papandreou, "ICDAR 2013 Competition on Writer Identification," 2013 12th Int. Conf. Doc. Anal. Recognit., pp. 1397–1401, Aug. 2013.
- [5] R. Fernandez-de-Sevilla, F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "Forensic writer identification using allographic features," Proc. - 12th Int. Conf. Front. Handwrit. Recognition, ICFHR 2010, vol. 2010, pp. 308– 313, 2010.
- [6] A. Gordo, A. Fornés, E. Valveny, and J. Lladós, "A bag of notes approach to writer identification in old handwritten musical scores," Proc. 8th IAPR Int. Work. Doc. Anal. Syst. - DAS '10, pp. 247–254, 2010.
- Q. A. Bui, M. Visani, S. Prum, and J.-M. Ogier, "Writer Identification Using TF-IDF for Cursive Handwritten Word Recognition," 2011 Int. Conf. Doc. Anal. Recognit., vol. 844, no. 1, pp. 844–848, Sep. 2011.
- [8] A. Chaabouni, H. Boubaker, M. Kherallah, A. M. Alimi, and H. El Abed, "Multi-fractal modeling for on-line text-independent writer identification," Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, pp. 623–627, 2011.

# Journal of Theoretical and Applied Information Technology

<u>15<sup>th</sup> May 2020. Vol.98. No 09</u> © 2005 – ongoing JATIT & LLS



ISSN: 1992-8645

<u>www.jatit.org</u>

- [9] A. Chaabouni, H. Boubaker, M. Kherallah, A. M. Alimi, and H. El Abed, "Combining of Offline and On-line Feature Extraction Approaches for Writer Identification," 2011 Int. Conf. Doc. Anal. Recognit., pp. 1299–1303, Sep. 2011.
- [10] J. T. J. Tan, J.-H. L. J.-H. Lai, C.-D. W. C.-D. Wang, and M.-S. F. M.-S. Feng, "Off-Line Chinese Handwriting Identification Based on Stroke Shape and Structure," 2010 2nd Int. Conf. Inf. Eng. Comput. Sci., no. March, pp. 1– 4, Dec. 2011.
- [11] J. Tan, J. Lai, and W. Zheng, "Chinese Handwritten Writer Identification based on Structure Features and Extreme Learning Machine," Proc. 2nd ICDAR Int. Work. Autom. Forensic Handwrit. Anal. AFHA 2013, pp. 2–6, 2013.
- [12] E. N. Zois and V. Anastassopoulos, "Morphological waveform coding for writer identification," Pattern Recognit., vol. 33, no. 3, pp. 385–398, 2000.
- [13] H. E. S. Said, T. N. Tan, and K. D. Baker, "Personal identification based on handwriting," Pattern Recognit., vol. 33, pp. 149–160, 2000.
- [14] V. Pervouchine and G. Leedham, "Extraction and analysis of forensic document examiner features used for writer identification," Pattern Recognit., vol. 40, no. 3, pp. 1004–1013, 2007.
- [15] M. Bulacu, L. Schomaker, and A. Brink, "Textindependent writer identification and verification on offline arabic handwriting," Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, vol. 2, pp. 769–773, 2007.
- [16]Z. He, X. You, and Y. Y. Tang, "Writer identification of Chinese handwriting documents using hidden Markov tree model," Pattern Recognit., vol. 41, no. 4, pp. 1295– 1307, Apr. 2008.
- [17] A. Gordo, A. Fornés, and E. Valveny, "Writer identification in handwritten musical scores with bags of notes," Pattern Recognit., vol. 46, no. 5, pp. 1337–1345, 2013.
- [18] C. Liu, R. Dai, and Y. Liu, "Extracting individual features from moments for Chinese writer identification," Proc. 3rd Int. Conf. Doc. Anal. Recognit., vol. 1, pp. 438–441, 1995.
- [19] M. Sreeraj and S. M. Idicula, "A Survey on Writer Identification Schemes," Int. J. Comput. Appl., vol. 26, no. 2, pp. 23–33, 2011.
- [20] G. J. Tan, G. Sulong, and M. S. M. Rahim, "Writer identification: A comparative study across three world major languages," Forensic Sci. Int., vol. 279, pp. 41–52, 2017.

- [21] G. J. Tan, G. Sulong, and M. Rahim, "Off-Line Text-Independent Writer Recognition for Chinese Handwriting: A Review," J. Teknol., vol. 2, pp. 39–50, 2015.
- [22] Rakhmadi, A., Syazrah Othman, N. Z., Bade, A., Mohd Rahim, M. S., & Amin, I. M. (2010). Connected component labeling using components neighbors-scan labeling approach. Journal of Computer Science, 6(10), 1099-1107. doi:10.3844/jcssp.2010.1099.1107
- [23] Bashardoost, M., Mohd Rahim, M. S., Saba, T., & Rehman, A. (2017). Replacement attack: A new zero text watermarking attack. 3D Research, 8(1) doi:10.1007/s13319-017-0118-y
- [24] Rad, A. E., Mohd Rahim, M. S., Kolivand, H., & Mat Amin, I. B. (2017). Morphological region-based initial contour algorithm for level set methods in image segmentation. Multimedia Tools and Applications, 76(2), 2185-2201. doi:10.1007/s11042-015-3196-y
- [25] Kurniawan, F., Rahim, M. S. M., Daman, D., Rehman, A., Mohamad, D., & Shamsuddin, S. M. (2011). Region-based touched character segmentation in handwritten words. International Journal of Innovative Computing, Information and Control, 7(6), 3107-3120.
- [26] Taha, M. S., Mohd Rahim, M. S., Lafta, S. A., Hashim, M. M., & Alzuabidi, H. M. (2019). Combination of steganography and cryptography: A short survey. Paper presented at the IOP Conference Series: Materials Science and Engineering, 518(5) doi:10.1088/1757-899X/518/5/052003
- [27] G. J. Tan, G. Sulong, and M. S. Mohd Rahim, "Offline Text-Independent Chinese writer identification using GLDM features," J. Telecommun. Electron. Comput. Eng., vol. 9, no. 3-3 Special Issue, 2017.
- [28] I. W. Selesnick, "The Slantlet Transform," vol. 47, no. 5, pp. 1304–1313, 1999.
- [29] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognit., vol. 29, no. 1, pp. 51–59, 1996.
- [30] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971–987, 2002.
- [31] T. Su, T. Zhang, and D. Guan, "Corpus-based HIT-MW database for offline recognition of general-purpose Chinese handwritten text," Int.



ISSN: 1992-8645

www.jatit.org



J. Doc. Anal. Recognit., vol. 10, no. 1, pp. 27-38, Mar. 2007.

- [32] T. Z. Tonghua Su, "HIT-MW Dataset for Chinese Handwritten Offline Text Recognition," Proc. Tenth Int. Work. Front. Handwrit. Recognit., 2006.
- [33] G.Zhang, X. Huang, S. Z. Li, Y. Wang, and X.Wu, "Boosting Local Binary Pattern (LBP)-Based Face Recognition," pp. 179-189, 2004.