J. Electrical Systems 20-11s (2024): 1440-1451

**Naresh Purohit[1*]**
**Subhash Panwar[2]**

# A Methodological Review of an Offline Writer Identification Framework Utilizing Deep Learning and Handcrafted Approaches

Journal of Electrical Systems

**Abstract**: Handwriting symbolizes a prevalent form of inquired writing and often gains major interest in legal contexts. Since handwriting is a behavioural trait, no two matured handwritings are identical or can be replicated. Thus, it is a highly effective biometrics approach. This work proposes a thorough examination of techniques for recognizing writers. It aims to present an overview of several datasets, techniques for obtaining attributes and classification algorithms (both handcrafted and deep learning based) for writer recognition. This work contributes valuable insights and support to fresh scholars by concisely presenting several feature extraction methodologies and classification strategies necessary for writer recognition across English, Arabic, Western, and other language scripts. Ultimately, we emphasized the obstacles and untouched findings in the discipline of offline writer recognition. At last, we propose possibilities for future exploration.

***Keywords***- Convolution Neural Network, Handcrafted Features, Handwritten Document Analysis, Image Classification, Writer Recognition.

## 1. INTRODUCTION

Writer recognition pertains to the assessment of the author of a handwritten document. The handwriting of every person is distinctive, making it viable for personal identification reasons. This is a relatively novel field of investigation in contrast to authentication of signatures, which has been extensively studied over the years. This is a notable learning trait of an individual and serves a crucial function for forensic records professionals in establishing any individual's legitimacy. In recent times, automated evaluation of handwritten documents has garnered considerable interest from investigators, particularly in the area of vintage document analysis. Given the large volume of handwritten samples, it requires a considerable amount of period for forensic experts to manually evaluate and contrast the questioned document with all specimens in the database to identify forgeries. Consequently, developing an automated method for writer recognition could prove highly beneficial, significantly relaxing the efforts of forensic experts by accurately recognizing material authored by an accused writer among an extensive collection of documents [1].

This investigation provides a methodological review on writer recognition of handwritten content across several datasets in numerous dialects, employing deep learning methods and foundational handcrafted features. The writer identity problem in document image analysis and recognition exhibits a significant challenge due to the considerable variability in individual writing patterns. The category of work divided into two approaches: (a) online and (b) offline, according to the writing mode employed. Online writer recognition concerns to orientation, force, and pace of writing, while offline writer recognition addresses words, characters, lines, or paragraphs. In online mode, handwritten data is generated concurrently with its collection. Using this manner, the author often produces handwriting utilizing a mouse or an electronic stylus, with both temporal and spatial data recorded as output. In offline mode, static graphics serve as our data within a scanned handwritten manuscript. The handwritten text data is acquired by camera or scanner equipment in the form of an image [2]. Researchers have focused a substantial amount of work on tackling certain problems, such as handwriting recognition and writer identification, as seen in figure 1, which displays numerous handwriting assessments. In the past twenty years, there has been an acceleration in research in automatic offline write recognition. The approach includes pre-processing, feature extraction, and writer classification phases, as seen in figure 2. The pre-processing phase is employed for purification the handwriting, effectively eliminating noise. The phase of elimination may involve several actions: splitting the handwritten image into smaller regions, transforming the dimensions of these regions, and applying relevant morphological procedures for interpreting the features [3]. Subsequently, the characteristics of the sources are obtained and recorded in the information database. The same procedure is utilized for the requested document. Considering the information database, learned classifiers allocate the unclassified request to one of the established patterns throughout the classification phase.

This research examines the investigation of offline writer recognition mostly conducted in English, Chinese, and Arabic written content. This article comprises several of the comprehensively examined datasets pertaining to such scripts. While there are certain works available for Indic scripts such as Devanagari, Bangla, and Telugu, conventional datasets for such scripts are lacking. Two approaches exist for the writer recognition task: text-dependent and text-independent. Text-dependent approaches require identical information for comparison and composition by multiple sources, but text-independent approaches require no uniformity in information. Text-independent approaches have broader applications; nevertheless, they do not achieve the same level of efficiency as text-dependent methods [4]. The article is structured as

[1*]Department of Computer Science & Engineering, Government Engineering College Bikaner, Bikaner, Rajasthan, India. E-mail: iteng.naresh@gmail.com
[2]Department of Computer Science & Engineering, Government Engineering College Bikaner, Bikaner, Rajasthan, India. E-mail: panwar.subhash@gmail.com

follows: Section 2 delineates the many categories of datasets. Section 3 discusses various methods and approaches used for offline writer recognition. Section 4 explores the comparison analysis. Conclusions are presented in Section 5.

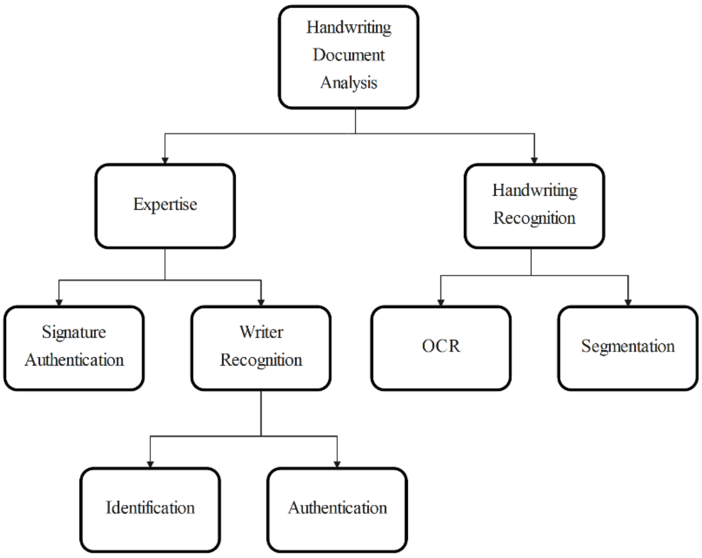


**Figure 1. Specific analysis of handwriting**

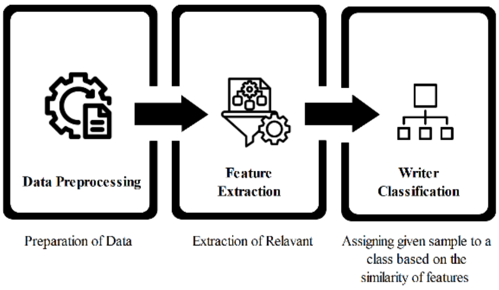


**Figure 2. Overview of writer recognition process**

## 2. DATASETS

The foundation of any investigation is the dataset. The presence of a dataset is a vital requirement over the establishment or assessment of any research area, as well as the equivalent pertains for handwriting and writer recognition. Various datasets for writer recognition, word detection, and recognition of characters have been released to the literature over the past few decades. We are compiling datasets that correspond with the languages in the forthcoming subsections. Several datasets utilized in this investigation is tabulated in table-1 with script and writer count.

### 2.1 English Datasets

Worldwide, more than 1.5 billion people—or 20% of the population—speak English, a common and ancient language. There are 360 million persons for whom English is their primary dialect. For numerous concerns, including optical character recognition (OCR), writer recognition, handwriting analysis, Parkinson's disease prediction (PDP), and many more, there has been substantial research done on handwritten text [5,6]. The following points discuss the well-known script datasets used for writer recognition in English.

• **IAM:** University of Bern produced a widely utilized IAM handwritten database for writer recognition. Originally consisting of 1066 forms created by 400 unique writers, the database was later expanded to incorporate 1539 forms developed by 657 distinct writers. Line, phrase, and word-level grouping details, as well as the writer's name and the ground truth text, are all included in the data set. It has 13,353 lines of labelled data with varying text, with around 14 lines of text each writer. Sixty percent of the lines of text served as a baseline, while forty percent were utilized to evaluate performance. Using the IAM dataset, several writer recognition and verification studies were conducted.

• **CEDAR:** A variety of datasets were generated at the University of Buffalo by the CEDAR, which refers to the Center of Excellence for Document Analysis and Recognition. The CEDAR-Letter database, which comprises monochrome and binary format images of text written by one thousand different authors, was the first and largest database developed for the purpose of author recognition and handwriting authentication.

### 2.2 Arabic Datasets

The 6$^{th}$ widely utilized language in the globe, Arabic, is an essential sematic language. Arabic is spoken by 420 million individuals worldwide. Investigators focused heavily on Arabic script for challenges including analyzing documents, identification of patterns, and processing of images since of the script's complications. The following points concentrate on several Arabic script datasets used for writer recognition and handwriting recognition.

- **AHDB:** AHDB, which denotes the Arabic Handwritten Database, is extensively utilized in the identification of handwritten text and in the recognition of writers in Arabic script. The database consists of the majority of prevalent written Arabic terms and texts from 105 authors. The dataset comprises around ten thousand phrases for Arabic cheque interpretation.
- **KHATT:** KHATT was created by investigators from the University of Braunschweig in Germany, the University of Dortmund in Germany, and the KFUPM in Saudi Arabia. Documentation authored in Arabic by one thousand different authors make up this collection. Six paragraphs were written by each writer, with around two thousand arbitrary paragraphs, fixed paragraphs, and free paragraphs also included.
- **IFN/ENIT:** The IFN/ENIT collection is a further blend of information sets from the Institute of Communications Technology (IFN) and the Ecole Nationale dIngenieurs de Tunis (ENIT) that is utilized for writer identification and handwriting recognition. A total of 411 authors that contributed 937 labels of towns and cities to the dataset. It generated almost 210,000 characters and 26,459 images. On each of the five sheets, twelve different city names were listed by the authors. The initial, character pattern order, and ground truth metadata are all encoded with each name. Numerous international contests and over a hundred investigators from over 30 countries utilize this dataset for Arabic handwritten text recognition.

### 2.3 Western languages dataset

For the purpose of author recognition, there are certain datasets available in western languages such as French and Dutch, among others. A couple of the following are:

- **Firemaker:** A Dutch language dataset employed for writer recognition is the Firemaker database. The dataset was developed by acquiring the handwritings of 1008 scanned documents belonging to Dutch scholars. Each scholar writes four pages, the first of which has five paragraphs written in standard handwriting. The author uses capital letters for two paragraphs on the second page, fabricated and artificial handwriting appears on the third page, and his own words are used to describe a cartoon on the fourth page. Therefore, in general, the first and fourth pages are utilized to identify the writer.
- **RIMES:** RIMES (Recon- naissance et Indexation de donnes Manuscrites et de fac similes) is yet another rather distinct dataset in writer recognition. This French script dataset consists of handwritten letters in French that individuals have submitted to businesses or government agencies. Over 1300 authors completed five letters each, for an aggregate of 5600 letters in over 12,000 pages including annotations. Additional databases containing characters, handwritten phrases totaling 300,000 fragments, and designs were also acquired.

## 3. APPROACHES AND METHODS

The following discussion highlights the strategies and methodologies employed in the data pre-processing, feature extraction, and classification phases utilized in writer recognition.

### 3.1 Data Pre-Processing

Data in its native binary format is understandable by machinery, while unstructured text or picture data is incomprehensible. Therefore, it is quite unlikely that simply feeding our system scanned handwritten images would be sufficient for training. Data pre-processing entails transforming or encoding the data such that it is easily interpretable by the system. In order to prepare scanned handwritten visuals of authors for writer recognition assignments, these visuals undergo several preprocessing methods:

- **Grayscale:** The process of converting color pictures into grayscale ones.
- **Smoothing and De-noising:** Primarily to eliminate unwanted spots and distortion.
- **Binarization:** A binary imagine consists just of two color levels, whereby each pixel contains a gray value of either 0 or 255, denoting black and white, respectively.
- **Normalization:** It mostly pertains to size normalization.
- **Refinement:** This involves the systematic elimination of edge points from the side shadow layer of the pre-binarized text, while maintaining the skeletal structure of the original text.
- **Geometric Transformations:** flipping, rotating, inverting, and center cropping an image (translation, transposition, mirroring, rotation, scaling), etc.

Moore's contour-following method is employed by Bulacu et al.[23] for obtaining the interior and exterior contours of binary written texts. The contours of all pixels situated precisely within the ink-background border comprise a series of coordinates. Rajiv Jain et al. [24] adopt a Canny edge detector on binary document images to acquire contours that derive geometry and boundary details. They breakdown the curves into an array of lines using a line fitting method. The majority of the suggested methods employ Otsu's approach for image binarization, as demonstrated in Stefan Fielet al. [8]. Initially,

they binarized the images of the CVL[6] dataset and subsequently applied line and word segmentation. In order to identify writers using convolutional neural networks, the input image must be square, meaning that its height and width are equal. The resultant line image width value is greater than the height when the line segmentation process is applied to a scanned image document. Linjie Xing et al. [9] resize the shorter edge (length) of the input image document to 113 while maintaining the aspect ratio. They then arbitrarily crop 113x113 image segments from the original text line image on the IAM[5] dataset. For the purpose of image extraction Keglevic et al.[11] employ the localization of SIFT key-points in their patches, which originated from the Harris corner detector utilized as the focal points of the patches.
Distorted images originate from inadequate image scanning. To rectify the distorted image, it is essential to figure out the skewed angle and subsequently correct it. Rehman et al. [10] utilized the Probabilistic Hough Transform (PHT) to identify lines, evaluate skew, and analyze accuracy on the QUWI Arabic dataset. Adak et al. [12] employed GOLESTAN-a, an approach designed around a 2D Gaussian filter, for separating words from a specified line image. Connected component tagging was employed to separate handwritten documents into text elements, further filtering out smaller elements such as points, slashes, punctuation, periods, and noise. Xiangqian et al. applied a Laplacian of Gaussian filter to separate phrases from handwriting images in the HIT-MW database. In [38], a computational morphological procedure known as closure is employed to separate words from segmented lines. The subsequent three steps are employed to adjust all word images: (b) slope adjustment; (c) slant rectification; (d) space minimization to reduce the impact of handwriting variance, as seen in figure 3.

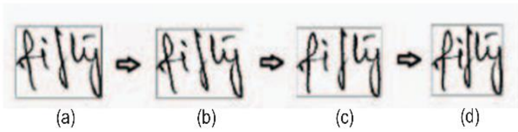


**Figure 3. Normalization of word images**

As seen in figure 4, methods for data enhancement were used on scanned handwriting images to enhance the efficiency of the AlexNet framework of the Convolution Neural Network (CNN) [10]. This is accomplished by removing the image data's structural components in order to extract the contours. Following the above, un-sharpening and masking the image are part of the data enhancement stage, which also incorporates the image's set differences and the erosion that follows.

**Figure 4. Data enhancement with erode and masking operations**

**3.2 Feature Extraction and Classification**

Various kinds of features were employed by investigators to recognize writers. The majority of these features are also employed in the machine detection of handwritten text. In the following section, the features that have been employed in recognizing of writers are described. The features employed by collaborating experts in writer recognition will be discussed in combination, followed by the findings of other investigators. The reader is able to observe a set of features in their suitable extent by organizing them by investigation team. It also illustrates the evolution of these features over time and the various applications or sets of data that these features are employed in. Two categories exist for the feature extraction techniques: conventional handcrafted-based and deep learning-based methods.

**Table 1. Datasets utilized for writer recognition**

| **Dataset** | **Language** | **Writers Count** |
|---|---|---|
| IAM | English | 657 |
| HWDB | Chinese | 300 |
| CVL | English/German | 310 |
| BFL | Portuguese | 315 |
| ICDAR 2011 | English/French/German/Greek | 26 |
| ICDAR 2013 | English/Greek | 50 |
| ICDAR 2017 | English/French/German/Greek | 720 |
| QUWI | Arabic/English | 1017 |

| | | |
|---|---|---|
| JEITA-HP | English | 580 |
| FIREMAKER | English | 250 |
| KHATT | Arabic | 1000 |
| NewISIdb: SoT | English/Bengali | 200 |
| CMATERdb1.1.1 | English-Bengali Mix | 40 |
| CVC-MUSCIMA Dataset | Music Scores | 50 |
| CEDAR | Signature Database | 50 |
| CERUG-EN | Chinese/English | 105 |
| MADCAT | Arabic | 325 |
| HIT-MW | Chinese | 241 |
| IFN/ENIT | Arabic | 411 |
| ICFHR-2012 | Arabic | 200 |

**3.2.1 Handcrafted-based methods**

For more than a decade, some computer vision applications are reliant on methodologies that are based on handcrafting. The techniques fall into two groups: (a) those that rely on texture and (b) those that rely on structure. Texture attributes store a set of handwritten texture characteristics, whereas Structure features are employed when handwriting is described as a collection of segmented shape. To enhance the reliability of models, Rajiv Jain et al. [24] presented a technique that makes use of K-adjacent segment (KAS) features. Figure 5 shows the line fitting techniques used to locate the image document's contours. The natural curves are transformed to a series of lines using these techniques. By grouping instances based on KAS attributes, a codebook is generated. After the vectors of features for the codebook are created and normalized, they may be compared using the Euclidean distance to classify writers. The results were 89.6% accurate using the IAM dataset, which included 350 authors. In addition, the similar author enhanced accuracy from 89.6% to 94.1% in 2014 by combining KAS features with SURF and contour gradient descriptors [25].

**Figure 5. K-adjacent segment (KAS) features**

The sleek curves are transformed into a series of lines using these techniques. Clustering exemplars based on KAS characteristics is used to build the codebook. After the vectors of features for the codebook are created and normalized, they may be compared using the Euclidean distance to classify writers. They were 89.6% accurate using the IAM dataset, which included 350 authors. In addition, the similar author enhanced accuracy from 89.6% to 94.1% in 2014 by combining KAS attributes with SURF and contour gradient descriptors [25]. Tang et al. introduced the Stroke Fragment Histogram (SFH) and Local Contour Pattern Histogram (LCPH) [26]. SFH is a segmentation an algorithm that extracts chunks from connected components in handwritten documents. For each connected components and SFH, the minimal bounding rectangle (MBR) with width and height is examined. The sliding panel is built with a set width=30. To create segments, the sliding window on the MBR slides from left to right using an interval break. To create the codebook, the Hierarchical Kohonen SOM clustering technique is employed. Last, they calculate Euclidian Distance between stroke fragments and code words from the code book by applying the following the following formula:

$$D = \sum (f - c)^2 \quad (1)$$

Where **f** stands for stroke fragments and **c** for code words, respectively. With the IAM dataset, they achieved an accuracy of 97.1%, while with the Firemaker dataset, it was 90.2%. Xiangqian et al. [27] employed the HIT-MW dataset, which is a Chinese dataset consisting of 241 authors. Before extracting SIFT descriptors and their associated scales and orientations, they use the LoG filter to obtain word areas from documents. Their accuracy rate was 95.4%. Additionally, Stefan Fiel et al. [28] utilize SIFT. They also came up with a vocabulary that explains typical handwriting patterns. In SIFT calculation, there are four stages. The initial stage involves using the initial images to build a Gaussian pyramid. An octave is one level of the pyramid, and the DoG filter further breaks it down into smaller layers. Figure 6 shows the processes to compute the scales, orientations, and positions of persistent spots that have been discovered. The final stage involves creating a SIFT description for every important point.

**Figure 6. The key points detected in a word region by SIFT**

For the CVL dataset's binary images, Salil Kanetkar et al. [29] suggested attributes that utilized Local Derivative Patterns (LDP). The distance required for writer recognition is calculated using LDP's histogram comparing. In order to avoid data duplication or inefficiencies, Christiein et al. [30] employed Contour-Zernike Moments (CZM) to express image attributes. The Zernike values were then encoded using VLAD (Vectors of Locally Aggregated Descriptors), that were taken using the image file later on. Using the ICDAR-2013 database, VLAD uses Nearest Neighbor classification to recognize writers with an accuracy of 97.5%; on the CVL dataset, it achieves 98.8%. For the KHATT Arabic dataset, the similar investigator used Root-SIFT features calculated sparsely at the script contour to reach an accuracy of 99.5%. An expansion of SIFT, Root-SIFT is defined as the employing of a Hellinger kernel for histogram distance measurements rather than the standard Euclidean distance. They achieved success rate of 99.0% on ICDAR-2013 and 93.4% on CVL using GMM super vector encoding [31]. Working with the Oriya script, Chanda et al. [32] used SVM classifiers to extract attributes using bi-quadratic interpolation techniques, and they achieved an accuracy rate of 94%. One feature encoding approach that Andrew et al.[33] suggested is orientated Basic Image Feature Columns, or oBIF Columns. These columns include Derivative-of-Gaussian (DoG) filters that can have seven different forms of symmetry.

Hannad et al. [34] extracted features from the IFN/ENIT dataset using three texture descriptors: LBP (Local Binary Pattern), LTP (Local Ternary Patterns), and LPQ (Local Phase Quantization). The LPQ method achieved an impressive accuracy of 94.89%. For labeling a pixel in the imagery, the LBP operator employs a 3x3 threshold with the center value. The threshold scores were then averaged and graded by powers of 2. Another approach, Cross multi-scale based Locally encoded Gradient Patterns (CLGP), developed by Chahi et al. [35], relies on the LETRIST descriptor. They employed transform feature building to get the writing image's texture information, and then they encapsulate it using the HOG function in non-overlapping units at several scales. On the IFN/ENIT dataset, they achieved an accuracy of 98.4%, resulting in is a 4.4% improvement above [34].

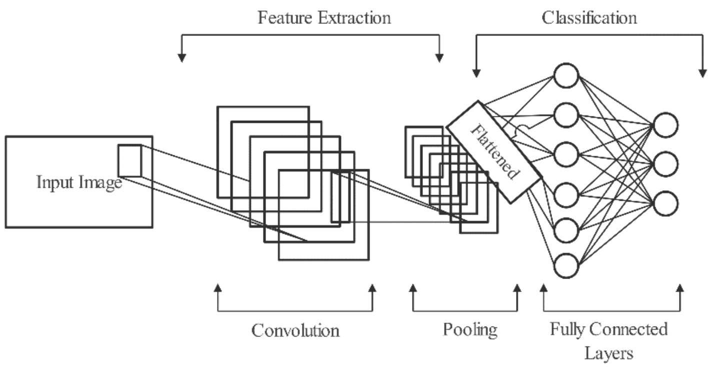


**Figure 7. Basic pipeline of CNN**

### 3.2.2 Deep Learning-based methods

This study addresses several deep learning strategies that relies on Convolution Neural Networks (CNN). Recognition effectiveness is often enhanced by using procedures or approaches based on deep learning as opposed to approaches that rely on hand-crafted features. The acquired attributes can make better use of data adaptation, which is the main reason. The most significant deep learning approach used in computer vision is CNN, that are capable of recognizing and classifying visual features. Figure 7 depicts the architecture, which consists of two primary components: feature extraction and classification. For obtaining features from an image, CNN uses a convolution layer (CL) to apply a filter to the entire image. This component extracts several characteristics by applying several filters to the image. Pooling Layer (PL) down samples CL-generated feature maps while retaining significant data. Alternating CL and PL in CNN models helps extract key characteristics. Vector patterns are flattened following the last pooling layer and sent to the fully connected layer for classification. A probability loss function predicts the writer class of the sample inquiry document. When it comes to classification, the majority of CNN-based algorithms utilize the SoftMax function. Several newly-proposed writer recognition models have achieved promising results by combining several techniques, such as support vector machines

(SVMs) and vector encodings (VLADs) with CNN. Xing et al.[9] extracted feature maps using a CNN framework similar to LeNet-5, with 5 CL and 3 PL. CL is employed with 96, 256, and 384 filters, each with a stride of 2. One further non-linear activation function that comes after CL is the Rectified Linear Unit (ReLU), which instructs the classifier SoftMax to round up values that are less than zero. Using the IAM dataset with labeled photos produced a 98.23% success rate. A different approach proposed by Christlein et al. [13] involves clustering images using a mini-batch version of k-means and using SIFT (Scale-invariant feature transform) key-points retrieved from the images. Figure 8 shows a CNN model learning with the cluster index as its image target. They make use of a 20-layer deep residual network (ResNet). The two-branch residual building components are that enable ResNet work. The first one contains numerous CL, while the second one just passes the result from the previous layer on to the next one. During ResNet's local extracted feature clustering, VLAD encoding is implemented.

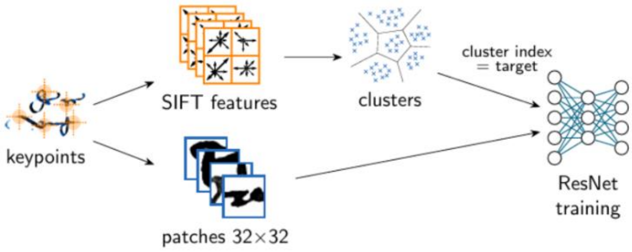


**Figure 8. Unsupervised feature learning using SIFT and ResNet**

As long as up to this point, tagged data has been the training set for deep learning methods. To train the algorithm with both labelled and un-labelled data, Shimingchen et al. [14] suggested semi-supervised method for learning features as depicted in figure 9. Data enhancement on un-labelled data is accomplished by means of weighted label smoothing regularization (WLSR), a weighted homogenous label dispersion approach. Afterwards, a vector of locally aggregated descriptors (VLAD) serves as an encoder that is used to create universal characteristics from local information. The classification stage makes further use of these global feature vectors. On the CVL dataset, they employed ResNet-50, a CNN variant that attains 99.2% accuracy. By calculating the average of all the values throughout the depth of n local feature vectors, Nguyen et al. [15] proposed average aggregation, an optimum localized feature integration approach. Using the IAM and Firemaker datasets, which combined 900 authors, they were able to attain a 91.81% accuracy rate. A deep adaptive learning approach suggested by Sheng et al. [16] in 2019 to employ the leaky-ReLU activation function following each fully connected or convolutional layer in their CNN model. Their study defines leaky ReLU as:

$$F(x) = \max(\alpha x, x) \quad where\ \alpha = 0.1 \tag{2}$$

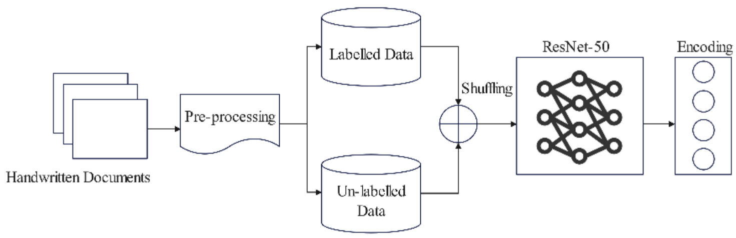


**Figure 9. Semi-supervised feature learning using ResNet-50 and VLAD encoding**

Most suggested systems encode CNN local features to create global characteristics. Christlein et al. [17] suggested GMM super vector encoding in 2015. Learning a Gaussian mixture model (GMM) from ZCA-whitened activation features is similar to the dictionary. Generating stats using this dictionary encodes localized descriptors. The regional feature maps derived from convolutional layers are encoded after ZCA whitening. This approach gets 97.6% and 99.4% accuracy with English text on ICDAR-2013 and CVL datasets. Also in 2018, similar author developed a solution with VLAD encoding on KHATT and attained 99.1% accuracy [18]. Using exemplar SVM as a classifier improved the efficiency of the model. Adak et al. [19] present an alternative CNN architecture known as SqueezeNet, which has a fire convolution component consisting of two layers: squeeze and expand. When deployed to an in-house Bengali dataset, it attains an accuracy of 90.43%. CNN also utilized for the writer classification problem in music score imagery by Leandro et al. [20] using the CVC-MUSCIMA dataset, that includes 1000 images across 50 groups. A mixture of deep learning and handcrafted descriptors is employed to extract similarities from handwritten images proposed by Sulaiman et al. [21]. Local Binary Pattern (LBP) technique served as a handcrafted feature descriptor, while the Alex-Net convolutional neural network

(CNN) architecture functioned as a deep descriptor. Both descriptors are consolidated and subsequently encoded using VLAD encoding, resulting in accuracies of 97.55% and 86.33% on the CVL and IAM datasets, respectively.

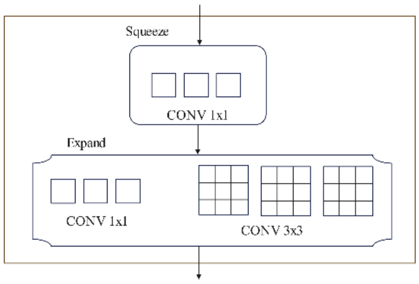


**Figure 10. Fire convolution component in SqueezeNet**

Dilara et al.[22] attained 97.2% efficiency using 64x64 image regions on the CEDAR dataset using a CNN alterations called a capsule network for signature detection and verification. By 2020, FragNet has been suggested by Sheng et al. [42]. It segments the semantic and a high degree aspects of the handwriting pattern included in the neural network's feature maps, those are referred to as fragments. Two paths exist in the FragNet. The first one, known as the feature pyramid, is employed for feature map extraction; the second one, known as the fragment route, is trained to anticipate the writer's identity using input imagine fragments. On the CVL dataset, it attains an accuracy of 90.2%; on the Firemaker dataset, 69%; and on the CERUG-EN dataset, 77.5%. Punjabi et al. [43] introduced ResNet-18 to examine various image patch dimensions from 100x100 to 1500x1500 for the purpose of writer classification. They employed a voting mechanism to consolidate all patch outcomes. The proposed system by Davood et al. [44] incorporates a combination of ResNet and Inception V3, achieving accuracies of 99.28% and 98.66% on the MCYT and UTSIG datasets, respectively. Macro et al. [45] developed a methodology employing ResNet on 32x32 picture patches from the Historical-WI dataset of ICDAR-2017, achieving an accuracy rate of 88.3%. Nabi et al. [46] proposed CNN methodologies based on VGG-16 for Urdu scripts. The suggested model has 16 layers succeeded by a completely linked layer. They attained a 98.71% accuracy rate on the proprietary Udru dataset.

The present investigation provides the reader a high-level overview of scripts, several feature extraction techniques, and classification approaches. The viewer also learned that larger databases tend to have better accuracy. Our survey results also attempt to demonstrate that studies pertaining to writer recognition have shown promising results in regards to accuracy rates. The study offered numerous examples of accurate and reliable results in a number of scripts, including Arabic, Chinese, French, Japanese, Latin, Urdu, and Devanagari. Another big issue is that there isn't a standardized database for different Indic scripts. Also included in the report is a brief summary of the writer identification system's attributes and classifiers. In order to achieve the highest possible accuracy in the coming years, it is possible to think about creating novel techniques for feature extraction and classification, as well as hybrid strategies that combine features and classifiers. Another fresh potential future angle in this manner is the issue of maintaining datasets with enough writers.

## 4. COMPARATIVE INVESTIGATION

Here, we provide a concise overview of the most recent, peer-reviewed approaches to offline writer recognition. In most cases, the outcomes are better with deep learnt features than with hand-crafted ones. Two of the most popular datasets utilized by researchers are IAM and CVL, as can be seen in Tables 2 and 3. Tang et al. [26] attained a performance of 97.1% using handmade techniques, whereas Xing et al. [9] achieved 98.23% using deep learning methodologies on IAM datasets. After an approach employing K-Adjacent Segments (KAS) was suggested by Rajiv Jain et al. [24] and achieved an accuracy of 89.6%, another study combining SIFT with KAS improved the performance even more in [25] and achieved an accuracy of 94.1% on the IAM dataset. The CVL collection includes around 300 authors work in English and German.

**Table 2. Comparative investigation of deep learning based methods**

| Author | Approaches | Classifier | Dataset | Accuracy (%) |
|---|---|---|---|---|
| Xing et al. [9] | CNN | SoftMax | IAM | 98.23 |
| Jija Das et al. [38] | CNN | SVM | ISIHWD | 85.70 |
| B. Kumar et al. [39] | CNN | SoftMax | Devanagari | 97.5 |
| Helal et al. [40] | CNN | SoftMax, SVM. | CVL | 99.80 |
| Fiel et al. [8] | Caffe-Net CNN | SoftMax | ICDAR 2011 | 97.40 |
| | | | ICDAR 2013 | 93.75 |
| | | | CVL | 99.18 |
| Tang et al. [41] | CNN | Joint Bayesian | ICDAR 2013 | 99.2 |
| | | | CVL | 99.8 |

| | | | | |
|---|---|---|---|---|
| Christlein et al. [13] | CNN, SIFT | VLAD, E-SVM | ICDAR 2017 | 88.9 |
| Chen et al. [14] | ResNet-50 | VLAD + Nearest Neighbor | ICDAR 2013 | 96.6 |
| | | | CVL | 99.2 |
| Razzak et al. [10] | CNN | Multi-Class SVM | Eng. QUWI | 92.78 |
| Nguyen et al. [15] | CNN | SoftMax | JEITA-HP | 99.97 |
| | | | IAM+Firemaker | 91.81 |
| Sheng et al. [16] | Adaptive CNN | SoftMax | CVL | 94.3 |
| | | | IAM | 85.2 |
| Christlein et al. [17] | CNN + GMM Supervector Encoding | SoftMax | ICDAR 2013 | 97.6 |
| | | | CVL | 99.4 |
| Christlein et al. [13] | CNN + VLAD Encoding | Exemplar SVM | ICDAR 2013 | 99.6 |
| | | | CVL | 99.5 |
| | | | KHATT | 99.6 |
| Keglevic et al. [11] | Dense-Net | VLAD + N. Neighbor | ICDAR 2013 | 98.9 |
| Adak et al. [12] | CNN | SVM | NewISIdb: SoT | 89.75 |
| Adak et al. [19] | Squeeze-Net CNN | SoftMax | In-house Bengali | 90.43 |
| Leandro et al. [20] | CNN | SoftMax | CVC-MUSCIMA | 84.0 |
| Sulaiman et al. [21] | CNN + LBP | VLAD + Nearest Neighbor | CVL | 99.69 |
| | | | IAM | 96.1 |
| | | | KHATT | 99.70 |
| Dilara et al. [22] | Capsule-Net CNN | SoftMax | CEDAR | 97.2 |
| Sheng et al. [42] | Frag-Net CNN | SoftMax | IAM | 85.1 |
| | | | CVL | 90.2 |
| | | | Firemaker | 69.0 |
| | | | CERUG-EN | 77.5 |
| Punjabi et al.[43] | ResNet-18 | Voting Scheme | IAM | 96.8 |
| | | | Firemaker | 99.2 |
| | | | ICDAR-17 | 83.6 |
| Davood et al. [44] | ResNet + Inception V3 | VLAD + N. Neighbor | UTSIG | 98.66 |
| | | | MCYT | 99.28 |
| Macro et al. [45] | ResNet +SIFT | VLAD | ICDAR-17<br>ICDAR-19 | 88.3<br>96.1 |
| Nabi et al. [46] | VGG-16 | Softmax | In-house Urdu dataset | 98.71 |

On CVL datasets, Helal et al. [40] achieved the top accuracy 99.80% using SVM while Tang et al. [41] achieved a score of 99.50 using Joint Bayesian techniques. Many researchers analyzed ICDAR 2013 (e.g., [8], [41], [11], [13], and [14]), and the evaluated studies also utilize the ICADR dataset collection. The best efficiency on the ICDAR 2013 dataset was 99.6 %, achieved by Christlein et al. [13] using a classifier that combines VLAD encoding with CNN and SVM. Tang et al. [41] achieved 99.2% accuracy using the same dataset and a Joint Bayesian classifier prior to [13]. Tables 2 and 3 demonstrate that in addition to English datasets, Arabic, Chinese, and a few in-house datasets are also evaluated.

## 5. CONCLUSION & FUTURE DIRECTIONS

Evaluating the development of offline writer recognition over the past 15 years, this article considers both traditional methods that usually used SIFT and LBP. New and improved network models for writer recognition tasks are being used by professionals in academia, whereas deep learning-based approaches focus on CNN reconstruction. These frameworks use networks such as ResNet, VGG-16, AlexNet, and they vary from basic CNN alterations to more advanced versions. There remains substantial research to be carried out on the subject of writer recognition since present methodologies fall lack of meeting society's real-world demands. Even though many other researchers have attempted to combine conventional feature extraction approaches with deep learning, the majority of future work will remain on combining deep learning with classical learning to get more robust recognition results. This report highlights areas for future research to improve offline writer recognition mechanisms. However, there are still imperfections such as the lack of a comprehensive review of each traditional approach for feature extraction and the absence of an examination of comparable papers employing the same classification technique. This paper, in our opinion, might be a valuable resource for scholars curious in learning more about existing offline writer recognition algorithms.

**Table 3. Comparative investigation of traditional handcrafted based methods**

| **Author** | **Approaches** | **Classifier** | **Dataset** | **Accuracy (%)** |
|---|---|---|---|---|
| Bulacu et al. [23] | Probability Distribution Function | Nearest Neighbor | Firemaker | 86.0 |
| | | | IAM | 89.0 |
| Jain et al. [24] | K-adjacent segment KAS | Proposed Distance Function | IAM | 89.6 |
| Jain et al. [25] | KAS, SIFT, Contour Gradient Descriptors | Fisher Vector + Cosine Distance | IAM | 94.1 |
| | | | ICDAR 2013 | 96.4 |
| | | | CVL | 97.0 |
| Tang et al. [26] | Stroke Fragment Histogram (SFH) and Local Contour Pattern Histogram (LCPH) | Chi-square distance | IAM | 97.1 |
| Tang et al. [27] | SIFT descriptors & corresponding scales and orientations | Manhattan distance, Chi-square distance | M-IAM | 98.5 |
| | | | Firemaker | 92.4 |
| | | | HIT-MW | 95.4 |
| Fielet al. [28] | SIFT, Fisher Kernels | Cosine Distance | ICDAR 2011 | 96.2 |
| | | | CVL | 95.6 |
| Kanetkar et al. [29] | Local Derivative Patterns | Chi-square distance | CVL | 92.1 |
| Christiein et al. [30] | Contour-Zernike Moments + VLAD | Cosine Distance | ICDAR 2013 | 97.5 |
| | | | CVL | 98.8 |
| Christlein et al. [31] | Root-SIFT, GMM Super Vector | Exemplar SVM | ICDAR 2013 | 99.0 |
| | | | CVL | 93.4 |
| | | | KHATT | 99.5 |
| Chanda et al. [32] | Oriented Basic Image Features | Nearest Neighbor | IAM | 89.6 |
| Hannadet al. [34] | LBP, LTP, LPQ | Dissimilarity Measure, Hamming distance | IFN/ENIT | 94.89 |
| | | | IAM | 89.54 |
| Chahiet al. [35] | Cross multi-scale locally encoded gradient patterns | Nearest Neighbor | IFN/ENIT | 98.54 |
| | | | IAM | 94.06 |
| | | | Firemaker | 97.60 |
| Hussein et al. [36] | SIFT, FAST | Naive Bayes + Nearest-Neighbor | ICFHR 2012 | 98.8 |
| | | | CVL | 99.8 |
| Bertolini et al. [37] | LBP, LPQ | SVM | IAM | 96.7 |

## REFERENCES


[1] Y.-J. Xiong, Y. Lu and P. S. P. Wang, “Off-line text-independent writer recognition: A survey”, Int. J. Pattern Recognition Artificial Intelligence, vol. 31, no. 5, 2017.

[2] Awaida, S. M., & Mahmoud, S. A. 2012. “State Of The Art In Off-Line Writer Identification Of Handwritten Text And Survey Of Writer Identification of Arabic text”. Educational Research and Reviews. 7(20):445-463

[3] Zhu, Y., Wang, T. T., & Yunhong. 2000. “Biometric Personal Identification Based on Handwriting”. Pattern Recognition Proceedings 15th International Conference. Barcelona. 3-7 September 2000. 2: 797-800

[4] Sreeraj.M, &Idicula, S. M. 201. “A Survey on Writer Identification Schemes”. International Journal of Computer Applications. 25(12): 23-33

[5] Marti, U. V., & H.Bunke. 1999. “A full English Sentence Database For Off-Line Handwriting”. 5th International Conference on Document Analysis and Recognition. Bangalore. 20-22 September 1999. 705-708.

[6] M. Diem, S. Fiel, F. Kleber, and R. Sablatnig, “CVL-Database: An Off-line Database for Writer Retrieval, Writer Identification and Word Spotting,” in2013 ICDAR (forthcoming), 2013

[7] G. Louloudis, N.Stamatopoulos, and B. Gatos, “ICDAR 2011 writer identification contest,” in Proc. Int. Conf. Document Anal. Recognit.,Beijing, China, 2011, pp. 1475–1479.

[8] Fiel, S. and Sablatnig, R. (2015). “Writer identification and retrieval using a convolutional neural network”. In Computer Analysis of Images and Patterns, pages 26–37. Springer.

[9] L. Xing and Y. Qiao, “Deepwriter: A multi-stream deep CNN for text-independent writer identification,” in2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, oct 2016,pp. 584–589

[10] Rehman, S. Naz, M.I. Razzak, H.A. Ibrahim, “Automatic visual features for writer identification: a deep learning approach”. IEEE Access7, 17149–17157 (2019)

[11] M. Keglevic, S. Fiel, and R. Sablatnig, “Learning Features for Writer Retrieval and Identification using Triplet CNNs,” in ICFHR, 2018, pp.211–216.

[12] C. Adak, B.B. Chaudhuri, M. Blumenstein, "Impact of Struck-out Text on Writer Identification", Proc. 30thInt. Joint Conference on Neural Networks (IJCNN), pp. 1465-1471, 2017.

[13] V. Christlein, M. Gropp, S. Fiel, A. K. Maier, "Unsupervised feature learning for writer identification and writer retrieval", in: International Conference on Document Analysis and Recognition (ICDAR), volume 1, 2017, pp.991–997.

[14] S. Chen, Y. Wang, C.-T. Lin, W. Ding, and Z. Cao. (2018). "Semi-supervised feature learning for improving writer identification."

[15] Nguyen, H.T., Nguyen, C.T., Ino, T., Indurkhya, B., Nakagawa, M.: "Text-independent writer identification using convolutional neural network". Pattern Recognition Letters121, 104–112 (2019)

[16] S. He and L. Schomaker, "Deep adaptive learning for writer identification based on single handwritten word images", Pattern Recognition, vol. 88,pp. 64 – 74, 2019.

[17] V. Christlein, D. Bernecker, A. Maier, and E. Angelopoulou, "Offline writer identification using convolutional neural network activation features," in Pattern Recognition: 37th German Conference, GCPR2015, Aachen, Germany, October 7-10, 2015, Proceedings. Springer International Publishing, 2015, vol. 9358, pp. 540–552

[18] Christlein and A. Maier, "Encoding CNN activations for writer recognition," in Proc. 13th IAPR Int. Workshop Document Anal. Syst. (DAS),Apr. 2018, pp. 169–174.

[19] C. Adak, B. B. Chaudhuri, M. Blumenstein, "A Study on Idiosyncratic Handwriting with Impact on Writer Identification" , ICFHR, pp. 193–198, 2018.

[20] L. T. Hattori, M. Gutoski, M. R. Aquino, and H. S. Lopes, "Patch-Based Convolutional Neural Network for the Writer Classification Problem in Music Score Images," in Proceedings of XIII BrazilianCongress on Computational Intelligence. Piscataway, NJ: IEEE Press,2017, pp. 1–6.

[21] Sulaiman, A., Omar, K., Nasrudin, M. F., &Arram, A. (2019). "Length Independent Writer Identification Based on the Fusion of Deep and Hand-Crafted Descriptors". IEEE Access, 7, 91772-91784.

[22] Gumusbas and T. Yildirim, "Offline signature identification and verification using capsule network," in Proc. IEEE Int. Symp. Innov. Intell. Syst.Appl. (INISTA), Jul. 2019, pp. 1–5.

[23] M. Bulacu and L. Schomaker, "Text-Independent Writer Identification and Verification Using Textural and Allographic Features," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 701-717, April 2007

[24] R. Jain and D. Doermann, "Offline Writer Identification Using K-Adjacent Segments," 2011 International Conference on Document Analysis and Recognition, Beijing, 2011, pp. 769-773

[25] R. Jain and D. Doermann, "Combining Local Features for Offline Writer Identification," 2014 14th International Conference on Frontiers in Handwriting Recognition, Heraklion, 2014, pp. 583-588

[26] Y. Tang, X. Wu and W. Bu, "Offline text-independent writer identification using stroke fragment and contour based features," 2013 International Conference on Biometrics (ICB), Madrid, 2013, pp. 1-6

[27] X. Wu, Y. Tang and W. Bu, "Offline Text-Independent Writer Identification Based on Scale Invariant Feature Transform," in IEEE Transactions on Information Forensics and Security, vol. 9, no. 3, pp. 526-536

[28] S. Fiel and R. Sablatnig, "Writer Identification and Writer Retrieval Using the Fisher Vector on Visual Vocabularies," 2013 12th International Conference on Document Analysis and Recognition, Washington, DC, 2013, pp. 545-549

[29] S. Kanetkar, A. Pathania, V. Venugopal and S. Sundaram, "Offline Writer Identification Using Local Derivative Pattern," 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, 2016, pp. 355-360

[30] V. Christlein, D. Bernecker and E. Angelopoulou, "Writer identification using VLAD encoded contour-Zernike moments," 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Tunis, 2015, pp. 906-910

[31] V. Christlein, D. Bernecker, F. H ̈onig, A. Maier, and E. Angelopoulou, "Writer identification using GMM supervectors and exemplar-SVMs," Pattern Recognition, vol. 63, pp. 258–267, 2017

[32] S. Chanda, K. Franke and U. Pal, "Text Independent Writer Identification for Oriya Script," 2012 10th IAPR International Workshop on Document Analysis Systems, Gold Cost, QLD, 2012, pp. 369-373

[33] A.J. Newell, L.D. Griffin, "Writer identification using oriented basic image features and the delta encoding", Pattern Recognition 47 (6) (2014) 2255–2265.

[34] Hannad, Y., Siddiqi, I., & El Kettani, M. E. Y. (2015). "Arabic writer identification using local binary patterns (lbp) of handwritten fragments." In Pattern recognition and image analysis (pp. 237–244)

[35] Abderrazak Chahi, Youssef Elmerabet,Yassine Ruichek, Raja Touahnia. "An effective and conceptually simple feature representation for off-line text-independent writer identification", 2019

[36] H. Mohammed, V. Mäergner, T. Konidaris and H. S. Stiehl, "Normalised Local Naïve Bayes Nearest-Neighbour Classifier for Offline Writer Identification," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, 2017, pp. 1013-1018

[37] D. Bertolini, L. Oliveira, E. Justino, R. Sabourin, "Texture-based Descriptors for Writer Identification and Verification," Expert Systems with Applications 40 (6) (2013) 2069–2080.

[38] J. D. Gupta and S. Samanta, “ISIHWD: A Database for Off-line Handwritten Word Recognition and Writer Identification,” 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR), Bangalore, 2017, pp. 1-6

[39] B. Kumar, P. Kumar and A. Sharma, “RWIL: Robust Writer Identification for Indic Language,” 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018, pp. 695-700

[40] L. G. Helal, D. Bertolini, Y. M. G. Costa, G. D. C. Cavalcanti, A. S. Britto and L. E. S. Oliveira, “Representation Learning and Dissimilarity for Writer Identification,” 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), Osijek, Croatia, 2019

[41] Y. Tang and X. Wu, “Text-Independent Writer Identification via CNN Features and Joint Bayesian,” 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, 2016, pp. 566-571

[42] S. He and L. Schomaker, “FragNet: Writer Identification Using Deep Fragment Networks,” in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3013-3022, 2020

[43] A. Punjabi, J. R. Prieto and E. Vidal, “Writer Identification Using Deep Neural Networks: Impact of Patch Size and Number of Patches,” 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021, pp. 9764-9771

[44] Keykhosravi, D., Razavi, S.N., Majidzadeh, K. et al. “Offline writer identification using a developed deep neural network based on a novel signature dataset.” J Ambient Intell Human Comput 14, 12425–12441 (2023).

[45] Peer, M., Kleber, F., Sablatnig, R. (2023). “Towards Writer Retrieval for Historical Datasets”. In: Fink, G.A., Jain, R., Kise, K., Zanibbi, R. (eds) Document Analysis and Recognition - ICDAR 2023. ICDAR 2023. Lecture Notes in Computer Science, vol 14187. Springer, Cham.

[46] Nabi, S.T., Kumar, M. & Singh, P. “DeepNet-WI: a deep-net model for offline Urdu writer identification.” Evolving Systems 15, 759–769 (2024).

[47] N. Purohit and S. Panwar, “State-of-the-Art: Offline Writer Identification Methodologies,” 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-8.

[48] N. Purohit and S. Panwar, “A Novel Multi-Input Hybrid Deep Convolution Neural Network Methodology for Offline Writer Recognition.”, Journal of Propulsion Technology, Vol. 44 No. 3(2023)