WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Writer Identification Using SVM and LBP

Krishna Narayanan[1], Mithun M P[2], Rahul A Jogalekar[3], Nithish B K[4], Kolluri SVVS Ganesh Guptha[5]

[1]Assistant Professor, Department of Artificial Intelligence and Machine Learning, MVJ College of Engineering, Bangalore, Karnataka, India.
[2,3,4]Undergraduate Scholar, Department of Computer Science and Engineering, MVJ College of Engineering, Bangalore, Karnataka, India.

A B S T R A C T

Owing to the rapid growth of touch screen mobile terminals and pen-based interfaces, handwriting-based writer identification systems are attracting increasing attention for personal authentication and digital forensics. The main goal is to analyze documents of different writing styles in order to identify the writers Outperforming the original algorithm, we evaluate a full 2D probability distribution that accounts for all potential combinations of angle pairs. A valuable biometric modality is the identification of a person based on scanned images of their handwriting.

1. Introduction

In this study, we offer a writer identification system that must recognise the author of a handwritten paragraph after receiving training on a variety of writers' handwriting. In our work, we used texture-based approach for writer identification, we used the potential texture descriptor, namely Local Binary Pattern (LBP), To distinguish between feature vectors from various writers, we employed Support Vector Machine (SVM) as a classifier. On the IAM handwritten dataset, where we tested our system, we were able to get a correct identification accuracy of roughly 99%. The field of image processing and pattern recognition has been actively researching signature verification for many years. Despite the continuous effort, signature verification remains a challenging issue. Signature verification provides a means of identifying the writer of a piece of handwriting in order to verify claimed identity in security and related applications. The writer requires to write the same fixed text. In this sense, signature verification may also be called text- dependent writer verification (which is a special case of text-dependent writer identification where more than one writer has to be considered). In practice there requirement for the use of fixed text makes writer Verification prone to forgery.

2. Literature Review

2.1 A texture-based approach for offline writer identification.

Offline handwriting has become a very important part in the field of behaviour biometrics. In this paper, we proposed an approach to identify writersfrom their handwritten documents. The foremost contribution of this study is to suggest the use of co-occurrence features to improve the performance of writer identification. A contour texture-based feature is extracted from pre-processed regions of interest (components or sub-images) and their exterior contours are based primarily on Modified Local Binary Pattern (MLBP) and Ink-trace Width and Shape Letters (IWSL) measurements. Considering the contour as a texture, and using MLBP and IWSL on different pixels are calculated in order to determine the similarities between different images of handwriting. Identification is carried out using the nearest neighbour rule and Chi- square distance. The proposed system has been evaluated on eight well-known handwriting databases (Arabic IFN/ENIT and KHATT, English IAM and CVL, Dutch Fire maker, Portuguese BFL, Chinese CERUG-CN, and English/Greek ICDAR2013). According to experimental findings, the proposed scheme outperforms all other schemes on the following databases: KHATT, CVL, Firemaker, BFL, CERUG-CN, and ICDAR2013. It also outperforms all other schemes on the following databases: IFN/ENIT and IAM.

2.2 Segmentation-free writer identification based on CNN

Writer identification is a technique used to positively identify the author of a given handwritten text sample. The applications of writer identification have changed significantly over the past few decades, ranging from the analysis of historical documents to contemporary forensic document analysis. Although the handwriting on old documents is rarely scrutinised, it is nonetheless useful to distinguish and identify distinct writers. Applications for forensic document analysis include the verification of signatures, the identification of document forgeries, and the authenticity of addressing records. The writer of a given handwritten text sample is predicted as part of the writer identification procedure. Under the document analysis paradigm, the handwriting recognition system is connected to writer identification and verification (WIV). a mechanism for author identification.

2.3 Writer identification system for pre-segmented offline hand written Devanagari characters using k-NN and SVM

A biometric identification system based on single and multiple modalities has been an evolving concept for solving criminal issues, security and privacy maintenance and for checking the authentication of an individual. The writer identification system is a type of biometric identification in which handwriting of an individual is taken as a biometric identifier. With this approach, the author's handwriting can be used to determine who they are. These systems use techniques for machine learning and pattern recognition to create a framework. The authors of this study have developed a novel approach for author identification based on pre-segmented Devanagari script characters as well as thorough state-of-the-art research. The experiment is performed on the corpus consisting of five copies of each character of Devanagari script written by 100 different writers, selected randomly at the publicplaces methodologies such as zoning, diagonal, transition and peak extent-based features and classification methods such as k-NN and linear SVM are used with identification accuracy of 91.53% when using zoning, transition and peak extent-based features with a linear SVM classifier

2.4 Deep adaptive learning for writer identification based on single handwritten word images

A common pattern-recognition challenge, writer identification seeks to identify the author of a handwritten paragraph from a photograph of it. A handwritten document's authorship is an implicit (indirect) attribute. A writer-identification system typically takes the handwriting-style data from the image of the analysed document and compares it to the data on the known writers' styles. The distribution of graphemes or the global statistics of ink trails are two geometric aspects that are frequently used to assess the handwriting style. The amount of text in handwritten photographs affects how well a conventional writer-identification system using handcrafted features performs. It was discovered that in order to identify a writer from a sample of Western handwriting using standard methods, roughly 100 letters are required.

2.5 Block Wise local binary count off-Line text writer identification

Handwriting has been found to be a very efficient distinguishing particular attribute of a person, which might be a helpful behavioural biometric modality. As compared to the electronic and printed texts, handwriting carries more information about the individual who wrote it. Also, each individual is distinguished by a distinctive writing style and the repetition of unconscious habits. As a result, no two texts will be written in exactly the same way. Because to these factors, the automatic writer identification field of study in the pattern recognition paradigm has become a vibrant and exciting one in recent years. Owing to technology advancements, handwriting analysis is attracting interest from psychologists, graphologists, forensic experts, and historians, who share some of their areas of expertise. Text-dependent and text-independent methods are the two basic categories into which writer identification methodologies are generally divided. While text-independent methods do not depend on the textual content and any writing samples can be used to identify the writer with a free choice of the pen, text-dependent approaches require each writer to produce exactly the same predefined text contents in both training and testing sets (for example, signature verification). Text-independent techniques can be further categorised and applied in one of two ways: I Online method, which makes use of dynamic writing data obtained from specialised acquisition tools to carry out writer identification at the moment of writing. They consist of writing speed and time, pen pressure and position, writing trajectory, and strokes. Therefore, the computerized algorithms for handwriting identification serve a great deal by assisting the human experts that otherwise have to manually compare an unknown sample within a large database. Writer identification also finds applications in offline/online verification of handwritten signatures (Kumar, Sharma, Chanda, 2012, Plamondon, Lorette, 1989), classification of ancient documents (arabadjis et al.., 2013), smart meeting room (Liwicki et al.., 2006)

2.6 Writer Identification Using GMM Super vectors and SVM.

Since handwritten text can be used as a biometric identifier like faces or speech, it plays an important role for law enforcement agencies in proving someone's authenticity. However, in such scenarios the decision is typically made by experts in forensic handwriting. On the other hand, a huge document database search for similar scribes highlights the need for an automated handwriting system (method).This topic has attracted significant attention recently, especially in the field of historical document analysis. In this application an automatic identification for particular writers can give new insights of life in the past. The focus of this paper is writer identification. Given a document, writer identification is the task of finding the specific writer (author) of the textfrom a set of writers which are known to the system. Depending on the data at hand, one has to differentiate between offline and online writer identification. In online writer identification the data contains temporal information about the text formation. Offline writer identification, in contrast, focuses solely on the handwritten text itself without taking into account any extra information. Textural approaches and allograph-based methods are further divided into two types for offline writer identification. Global data derived from the handwriting's style, such as measures of the ink width or the angles of the stroke directions, are used to define the handwriting in the first group. On the other hand, in all o-graph-based approaches, the distribution of features taken from tiny letter parts, or "allographs," is used to characterise the writer. It is necessary to train a vocabulary in advance using the training set's feature descriptors.

2.7 Writer Identification using texture description of handwritten.

The analysis of handwriting and hand-drawn shapes has been an interesting research area for many centuries that has attracted a significant number of psychologists, graphologists, palaeo-graphers and experts to solve a wide variety of problems. With the theoretical and practical advancements in computers and the availability of appropriate tools and technologies, automated systems for analysis of handwriting are being researched and developed for more than now. Although these computerized tools cannot completely replace the human analysis, they serve a great deal by assisting the human

experts through features like reduction of search space, visualization, segmentation, efficient processing, etc. In addition to the classical task of handwriting recognition, other relevant problems include classification of writing styles, keyword spotting of handwritten words, studies on correlation between handwriting and different neurological disorders, prediction of writer demographics from handwriting and verification and identification of writers from handwritten samples. The focus of this study lies on the last of the aforementioned problems i.e. identification of individuals (writers) from samples of their handwritten text.

3. Conclusion

In this paper, we present a novel technique to detect plant leaf images by combining LBP features and then classify the leaves using Multiclass Support Vector Machine (SVM). We carried out our experiment on a publicly available dataset called Flavia Leaf Dataset. Besides we can see that LBP features with SVM accuracy is 40.6%. Then we combined 8 x 8 cell size descriptor with LBP features and feed into SVM and the overall detection accuracy is 91.25%. The experimental shows the effectiveness of LBP for handwritten detection. In the future, it will be interesting to apply this method to the different Publicly available dataset and also in various domain

References

[1] A. Brink, J. Smit, M. Bulacu, L. Schomaker, Writer Identification Using Directional InkTrace Width Measurements, Pattern Recognition (2012)

[2] T. Gilliam, R. Wilson, J. Clark, Scribe Identification in Medieval English Manuscripts, in: Pattern Recognition (ICPR), 2010

[3] D. Fecker, A. Asit, V. M¨argner, J. El-Sana, T. Fingscheidt, Writer Identification for Historical Arabic Documents, in: Pattern Recognition (ICPR), 2014 22nd International Conference on, Stockholm,

[4] M. Bulacu, L. Schomaker, Text-Independent Writer Identification and Verification Using Textural and Allographic Features, Pattern Analysis and Machine Intelligence, IEEE Transactions on 29 (4) (2007)

[5] C. Djeddi, L.-S. Meslati, I. Siddiqi, A. Ennaji, H. E. Abed, A. Gattal, Evaluation of Texture Features for Offline Arabic Writer Identification, in: Document Analysis Systems (DAS), 2014 11th IAPR International Workshop on, Tours, 2014

[6] S. He, L. Schomaker, Delta-n Hinge: Rotation-Invariant Features for Writer Identification, in: Pattern Recognition (ICPR), 2014 22nd International Conference on, Stockholm, 2014

[7] S. Fiel, R. Sablatnig, Writer Identification and Writer Retrieval using the Fisher Vector on Visual Vocabularies, in: Document Analysis and Recognition (ICDAR), 2013 12th International Conference on, Washington DC, 2013

[8] S. He, M. Wiering, L. Schomaker, Junction Detection in Handwritten Documents and its Application to Writer Identification, Pattern Recognition

[9] R. Jain, D. Doermann, Writer Identification Using an Alphabet of Contour Gradient Descriptors, in: Document Analysis and Recognition (ICDAR), International Conference on, Buffalo, 2013

[10] R. Jain, D. Doermann, Combining Local Features for Offline Writer Identification, in: Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on, Heraklion, 2014