International Journal of Applied Sciences and Smart Technologies

Volume 4, Issue 2, pages 233 -240

p-ISSN 2655-8564, e-ISSN 2685-9432

Design of Someone's Character Identification Based on Handwriting Patterns Using Support Vector Machine

R. A. Kumalasanti[1,*]

[1]*Department of Informatics, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*

[*]*Corresponding Author: rosalia.santi@usd.ac.id*

(Received 01-11-2022; Revised 21-11-2022; Accepted 28-11-2022)

Abstract

Image processing has a fairly broad scope and is rich in innovation. Today, image processing has developed with various reliable methods in almost all aspects of life. One of the uses of technology in the field of image processing is biometric identification. Biometric is a system that utilizes specific data in the form of individual physical characters in the process of identifying and validating data. There is also a biometric attribute that will be developed in this study is handwriting. The handwriting pattern of each individual has a different character and uniqueness so that it can be used as an identity. The uniqueness of this handwriting will be studied with the aim of recognizing a person's character or personality. If someone's personality data has been obtained, this can help the process of recruiting prospective employees in a company by simply reading from handwriting patterns. Handwriting can be studied by combining the science of Psychology so that it can provide output in the form of a person's characteristics or personality. This research will be developed using the multi class Support Vector Machine (SVM) classification. The preprocessing stage in the form of

This work is licensed under a Creative Commons Attribution 4.0 International License

Volume 4, Issue 2, pages 233 -240
p-ISSN 2655-8564, e-ISSN 2685-9432

binarization, thinning and data extraction will also greatly affect the reliability of the system. Simulations with variations of variables and parameters are expected to obtain optimal accuracy.

Keywords: handwriting, biometric, SVM

1 Introduction

Digital image processing is digital manipulation and interpretation of images using a computer. Image processing has a fairly broad scope and is still being developed in various aspects of life. This shows that image processing has a good impact on technological developments in this era. Digitalization is something that is already familiar to hear because almost all daily activities utilize digital data. Information in the form of digital data has a flexible nature so that it can be used for certain purposes. when COVID-19 hit, activities & social interactions were increasingly restricted in order to break the chain of the virus. This has an impact on social sustainability which is increasingly limited. An example is the recruitment of prospective employees which is quite complex, starting from a psychological test to interviews. The Phsycotest aims to find out the character of the prospective employee and an interview test is conducted to see the gestures and how to answer each question given to the prospective employee to be assessed. Some of these tests are sometimes added to other tests which are quite time-consuming. This becomes less effective when it has to be done in the present because social needs are being restricted. Here image processing can be a solution, namely by using digital data to identify a person's character through handwriting. Handwriting is one of the biometric attributes that can be used to identify other than fingerprints, facial recognition, voice, and many more. A person's hand writing has its own uniqueness because it can be seen from the firmness of the strokes, the distance between one word and another, the distance between the lines / boundaries of the paper used [1]. These characteristics can be used to see explicitly the characteristics of a person as an individual. If a person's character has been obtained, the recruitment activities for prospective employees can run effectively and efficiently without having to meet face to face.

Volume 4, Issue 2, pages 233 -240
p-ISSN 2655-8564, e-ISSN 2685-9432

In this research, we will try to use SVM (support Vector Machine) with the same data set, namely handwriting. Handwriting will be retrieved from manual writing offline (on paper) and then scanned for pre-processing to the testing stage. Handwriting will be trained and tested to get the characteristics of each individual.

2 Research Methodology

Handwriting is one of the biometric attributes that can be used as authentic data because each individual's handwriting is unique. Emphasis when writing, the firmness of the strokes, the distance between one word and another, the distance between the lines and the margins of writing and much more. In this study, handwriting is used to read a person's character. Computer science and psychology will be collaborated to find optimal solutions.

2.1 Biometrics. Biometrics is a science and innovation to describe information from the human body naturally [2]. Biometrics refers to individual differences based on physiological or social attributes. Some of the biometrics on the human body are iris, hand geometry, signatures, fingerprints, face, handwriting and many more. Biometric values themselves can be in the form of individual physiological or social attributes that meet completeness needs.

2.1 Preprocessing. Preprocessing is a stage that aims to improve the quality of the images obtained so that they are easier to process at a later stage [3]. The data set in the form of handwriting will be scanned to obtain a digital image. The stages of this preprocessing are cropping, negative images, image binaryization, and thinning. Cropping is used to maximize the information retrieved so that there is not too much information that is not needed. Binaryization is used to process data from RGB to black and white. This aims to ease data computation and can provide optimal data sets in the next stage. Thinning is also done to thin out the writing so that the writing has the same thickness even though the thickness of the pen varies.

International Journal of Applied Sciences and Smart Technologies

Volume 4, Issue 2, pages 233 -240

p-ISSN 2655-8564, e-ISSN 2685-9432

2.2 SVM (Support Vector Machine). This research will utilize the Support Vector Machine (SVM) method in conducting data training and testing. SVM is a technique for finding hyperplanes that can separate two data sets from two different classes [4]. There are also advantages of SVM, namely being able to determine distances using support vectors so that the computation process becomes faster and more effective. SVM is one of the methods in supervised learning which is usually used for classification and regression [5]. In classification modeling, SVM has a more mature and clearer concept mathematically compared to other classification techniques [6]. SVM can also solve linear and non-linear classification and regression problems.

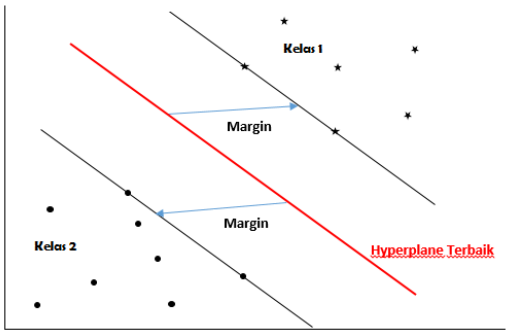


Figure 1. Illustration of Class Separation in SVM [4]

The best hyperplane as seen in figure 1 can be found with

$$\max L_D = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i=1,j=1}^{n} a_i a_i y_i y_j x_i x_j$$

In condition that

$$\sum_{i=1}^{n} a_i y_i = 0, a_i \geq 0$$

Not all data can be separated linearly, while SVM is basically only able to separate data linearly, so a development is needed to make SVM able to separate non-linear data, one of which is by adding a kernel function. By adding the kernel function to SVM, data x will be mapped to a higher vector space so that a hyperplane can be constructed. Hyperplane illustration can be seen in figure 2.

International Journal of Applied Sciences and Smart Technologies

Volume 4, Issue 2, pages 233 -240

p-ISSN 2655-8564, e-ISSN 2685-9432

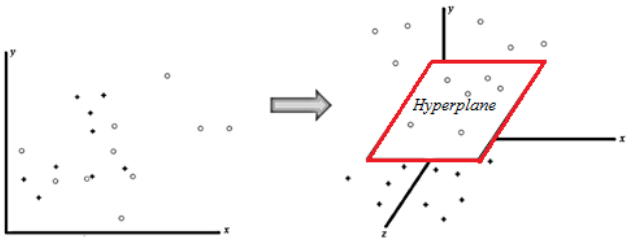


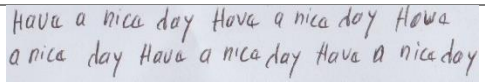
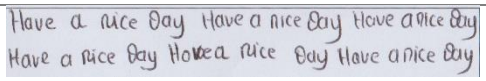
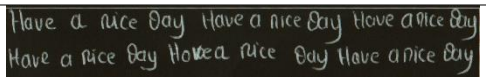
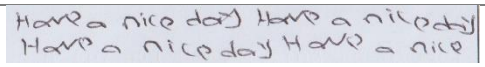
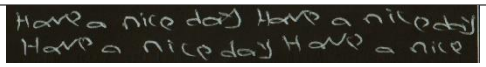
Figure 2. Hyperplane illustration in a higher dimension [4]

3 Results and Discussions

Identification of handwriting characteristics which is still being developed because the dataset is in the form of biometric data is an interesting thing to develop. In today's digital era, it is possible to model all forms of activity as concisely as possible and as effectively as possible. Moving on from instantaneous individual needs, the time-consuming process of recruiting prospective employees can be completed in a short and effective way using image processing.

The raw data in the form of handwriting after being scanned is then immediately given preprocessing to go to the next stage. Raw data and preprocessing results can be seen from table 1.

Table 1. Sample Image from Preprocessing results

ID	Raw Data	Preprocessing Result
1		
2		
3		

The appropriate input data will be calculated the distance, namely the distance between words in handwriting. Furthermore, if the distance results have been obtained, then the kernel calculation will be carried out with the SVM kernel function, namely the linear kernel. The results and predictions obtained after the SVM kernel calculation process will then be carried out by cross validation calculations to obtain accurate

Volume 4, Issue 2, pages 233 -240
p-ISSN 2655-8564, e-ISSN 2685-9432

results from the classification process in kernel calculations. The SVM flowchart can be seen in Figure 3 below.

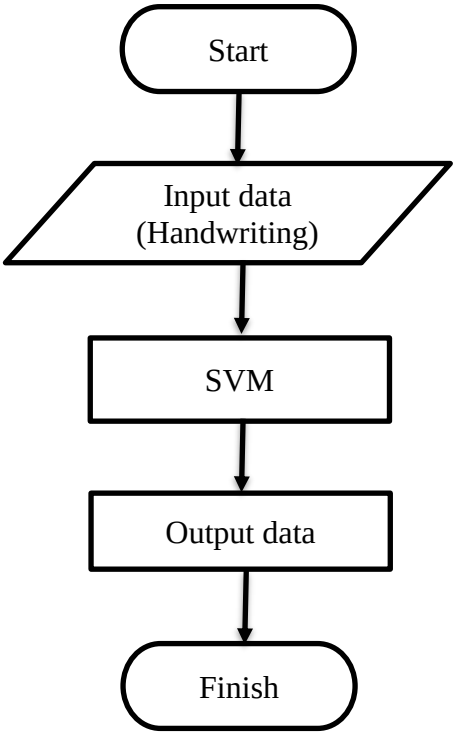


Figure 3. SVM Flow Chart

Calculation of handwriting patterns using word spacing patterns is given to training images and test images. This research uses computer science and psychology in application. Based on the results obtained from the SVM calculation, it will then be classified according to a person's character or personality. Here it takes a lot of attributes to compare one character with another so that the results obtained can be better.

4 Conclusions

The results and predictions of the handwritten character identification design using SVM and the accuracy obtained after calculations using cross validation are expected to provide optimal accuracy. The results of a person's character classification will be able to help make decisions regarding the recruitment of prospective employees. It is hoped that the results obtained will not only be utilized in the pandemic era but can also be developed in the future.

Volume 4, Issue 2, pages 233 -240

p-ISSN 2655-8564, e-ISSN 2685-9432

References

[1] G. Thilagavathi, G. Lavanya and N. K. Karthikeyan, "Tamil Handwritten Character Recognition Using Artificial Neural Network," *International Journal of Scientific & Technology Research* , **8**(12), 1611-1616, 2019.

[2] O. C. Ahuja, M. A. Mabayoje and R. Ajibade, "Offline Signature Recognition & Verification using Neural Network," *International Journal of Computer Applications*, **35**(2), 44-51, 2011.

[3] S. Saidah, M. B. Adinegara, R. Magdalena and N. K. Pratiwi, "Identifikasi Kualitas Beras Menggunakan Metode K-Nearest Neighbor dan Support Vector Machine," *Jurnal Telekomunikasi, Elektronika, Komputasi, dan Kontrol*, **5**(2), 114-121, 2019.

[4] M. Athoillah, "Pengenalan Wajah Menggunakan SVM Multi Kernel dengan Pembelajaran yang Bertambah," *JOIN (Jurnal Online Informatika)*, **2**(2), 84-91, 2017.

[5] S. Khedikar and U. Yadav, "Identification of Disease by Using SVM Classifier," *International Journal of Advanced Research in Computer Science and Software Engineering*, **7**(4), 81-86, 2017.

[6] P. A. Octaviani, Y. Wilandari and D. Ispriyanti, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *Jurnal Gussian*, **3**(4), 811-820, 2014.

International Journal of Applied Sciences and Smart Technologies

Volume 4, Issue 2, pages 233 -240

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank