

# Hiragana and Katakana minutiae based recognition system

Piotr Szymkowski<sup>\*1</sup>, Khalid Saeed<sup>1</sup>, Nobuyuki Nishiuchi<sup>2</sup>

<sup>1</sup>Bialystok University of Technology, Department of Computer Science, Bialystok, 15-351, Poland

<sup>2</sup>Tokyo Metropolitan University, Department of Computer Science, Tokyo, 191-0065, Japan

## ARTICLE INFO

Article history:

Received: 12 July, 2021

Accepted: 12 October, 2021

Online: 20 November, 2021

Keywords:

Handwriting recognition

Biometric recognition

Minutiae

## ABSTRACT

The Japanese writing system is unique due to the number of characters employed and the methods used to write words. It consists of three different 'alphabets', which may result in the methods used to process Latin script not being sufficient to obtain satisfactory results when attempting to apply them to a recognition of the Japanese script. The authors present an algorithm based on minutiae, i.e., feature points, to recognise the hiragana and katakana characters. A method using image processing algorithms is compared with a method using a neural network for the purpose of automating this process. Based on the distribution and type of minutiae, vectors of features have been created to recognise 96 different characters. The authors conducted a study showing the effect of the chosen segmentation method on the accuracy of the character recognition. The proposed solution has achieved a maximum accuracy at the level of 65.2%.

## 1 Introduction

This paper is an extension of work originally presented in The International Conference on Graphics and Signal Processing [1]. The Japanese writing system consists of three types of characters: kanji, hiragana and katakana. The hiragana and katakana characters are found only in Japan. They are syllabaries, which means that a single written character stands for a single spoken syllable. Hiragana is mainly used to write prefixes and suffixes of the Japanese words, particles and words of Japanese origin which for some reason are not written in the kanji characters. Hiragana is also used to write furigana, a reading aid for the difficult or unusual kanji characters. It is often found in books for children and teenagers.

Katakana is used to write first names, surnames and words of foreign origin. It is also applied for onomatopoeia (mimetic words), which frequently occur in Japanese. The two previously discussed syllabaries each contain 46 characters each.

The last group of characters are the kanji characters. These characters have been borrowed from the Chinese language. The number of the kanji characters in Japan is estimated to be approximately 50,000, but the average Japanese needs circa 3,000 characters for the use on an everyday basis. The kanji characters are used to write the subject of words. Unlike the hiragana and katakana, the kanji characters have different manners of reading. The way they are read depends on whether they stand alone or are accompanied by

other characters. One kanji character can have even more than 10 different ways of reading.

The quantity and complexity of the writing system in relation to the 26 characters of the Latin alphabet results in the fact that recognition methods and algorithms for the purpose of writing should be subject to a change.

## 2 State of the Art

For the purpose of the recognition of the Japanese writing system, publications on the on-line recognition are easier to find than the off-line ones. The key publications presenting a cross-section of ongoing studies on this very subject matter will be introduced below.

### 2.1 Deep Convolutional Recurrent Network

The first solution presented uses deep convolutional recurrent neural networks in order to recognise single lines of text written in Japanese [2]. The network used consists of three layers: a convolutional feature extractor, recurrent layers and transcription layers [3]. Prior to the recognition itself, a preprocessing of the images is required. This consists of normalising, i.e., homogenising the samples, and binarizing them with the use of the Otsu's method. The authors use the TUAT Kondate database. Another aspect presented in this paper constitute the artificially generated samples, developed on the basis

\*Corresponding Author: Piotr Szymkowski, [prszymkowski@gmail.com](mailto:prszymkowski@gmail.com)

of the previously mentioned database. The model presented here shows a character recognition rate at the level of 96.35% when using only the data from the database, and at the level of 98.05% when using also the artificially generated data. The generation of these data is based on the use of the local and global elastic distortions.

### 2.2 Extended Peripheral Direction Contributivity

The recognition method of the Japanese writing system presented by the authors [4] adopts the Extended Peripheral Direction Contributivity method. It is one of the more commonly used directional features. The feature vector is developed in the following manner. First of all, the image is scanned from their peripheral to inner part in the eight directions. Then, on each black pixel of the contour of the first, second and third line encountered in each direction, the run-length  $l_i$  in each of the eight directions is calculated and the direction contributivity  $d_i$  ( $i = 1, 2, \dots, 8$ ) is calculated for the pixel using (1) [5].

$$d_i = \frac{l_i}{\sqrt{\sum_{j=1}^8 l_j^2}} \quad (1)$$

Thirdly, the image is divided into several zones in each direction of the peripheral scan. Finally, a feature vector is obtained by recording the directional components for each zone, in each direction and in each order encountered. The information that can be obtained by the aforementioned method are as follows [6]:

- A quantity of lines and their orientation,
- A position of the lines,
- The relations between the lines.

### 2.3 Address Recognition

The algorithm presented, represents the use of a system that reads information from the letters and postal shipments concerning the information on a destination address [7]. For the purpose of reducing the information to be extracted, a priori knowledge about the structure of the Japanese address has been used. The introduced system is based on five processes: segmentation, segment combination, character recognition, character combination and path search. The system is supported by a glossary of the possible addresses.

## 3 Image Processing Algorithms

The algorithm presented by the authors uses many image processing methods. Some of the currently used solutions use a priori knowledge, which also reduces a probability of misclassification of the characters. The principles of operation of the most important of them will be specified below.

### 3.1 Morphological Closing

One of the essential operations performed on the images are the morphological ones. As their name suggests, their application results in a change in the appearance of the result image. In case of

the algorithm presented by the authors, the image on which these operations are employed is black and white. The two most basic morphological operations are the image erosion and image dilation. The result of applying the filter is shown in Figure 1. Their application is equivalent to the minimum and maximum of the statistical filters [8].

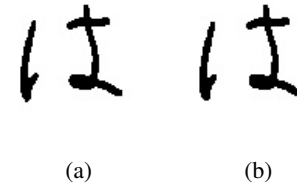


Figure 1: Original image (a) and image after morphological closing applied (b).

The closing filter consists in the performance of the dilation and erosion operations on the image using the same structural element. As a result of this operation, objects lying close to each other are merged, their edges are filled and smoothed.

### 3.2 Bilateral Filter

The bilateral filter is one of the blurring filters used to remove the initial noise present in the images. In contrast to the Gauss filter, it is not linear, which means that it blurs the image while retaining transparency of the edges [9].

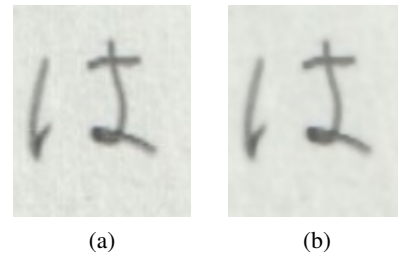


Figure 2: Original image (a) and image after bilateral filter application (b).

This is achieved by varying the blur rate, which depends on the size of the structural object and the value of the local gradient. It results in stronger blurring of homogenous areas and lesser blurring in the regions of sharp local changes, i.e. edges, what can be seen at Figure 2. The bilateral filter is presented in (2).

$$I_{filter} = \frac{1}{W_p} \sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|) \quad (2)$$

The normalising formula is presented in (3).

$$W_p = \sum_{x_i \in \Omega} f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|) \quad (3)$$

where:

- $I_{filter}$  – result image,
- $I$  – input image,

- $x$  - coordinates of pixel under analysis,
- $\Omega$  - structural object,
- $f_r$  - smoothing factor (intensity),
- $g_s$  - smoothing factor (distance).

### 3.3 Modified K3M Algorithm

Thinning is one of the operations used to simplify the description of the shapes of objects in the image. The operation is performed on a binarized image. As shown in Figure 3, all lines in the image are narrowed until they reach a width of 1 pixel. The modified K3M algorithm used in this publication has been created by one of the authors [10].

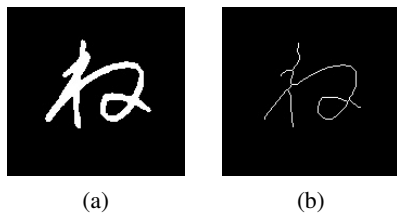


Figure 3: Original image (a) and sign thinned using Modified K3M algorithm (b).

### 3.4 Otsu's Method

The Otsu's method is one of the binarization techniques, which uses the thresholding what means that the algorithm uses a predetermined threshold to segment the entire input image. An example of how the method works is shown in Figure 4. It is a technique based on the discriminant analysis [11]. The optimal threshold is chosen based on one of the presented in (4) equivalent criterion functions.

$$\eta = \frac{\delta_B^2}{\delta_T^2}, \lambda = \frac{\delta_B^2}{\delta_W^2}, \kappa = \frac{\delta_T^2}{\delta_W^2} \quad (4)$$

Where:

- $\delta_B^2 = P_0(\mu_0 - \mu_T)^2 + P_1(\mu_1 - \mu_T)^2 = P_0P_1(\mu_0 - \mu_1)^2$  - between class variation,
- $\delta_W^2 = P_0\delta_0^2 + P_1\delta_1^2$  - within class variation,
- $\delta_T^2 = \sum_{i=0}^{L-1} (i - \mu_T)^2 P_i$  - global variation,
- $\delta_T^2 = \delta_W^2 + \delta_B^2$ .

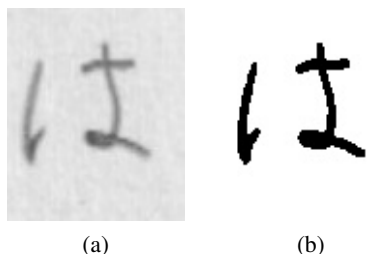


Figure 4: Original image (a) and image after Otsu method application (b).

## 4 Algorithm Description

To achieve the results, a Python application was prepared by the authors using keras library to apply artificial neural networks and sklearn to apply the SVM algorithm. The presented proprietary solution uses two artificial intelligence algorithms for the purpose of automating the recognition process of Japanese handwriting, therefore two types of feature vectors are created. It is similar to another used in the biometric algorithms, more specifically while recognising a user by their fingerprints. In order to create feature vectors for the character recognition, characteristic points - minutiae - are used. Figure 5 shows a diagram demonstrating how the presented solution works.

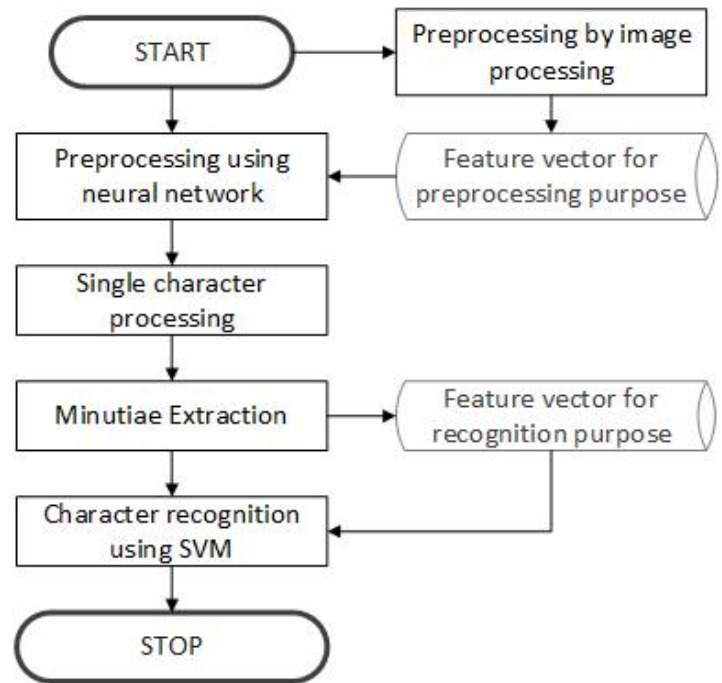


Figure 5: Algorithm diagram

### 4.1 Preprocessing Using Image Processing

This step is followed by the creation of feature vectors further used during the learning of the neural network in order to perform an automatic segmentation. The steps of this process are shown at Figure 6. The first stage consists in the use a bilateral filter [12]. As a consequence of the aforementioned, some of the noise that appeared in the image is eliminated. However, this has failed to affect the edge condition.

The next step has been to apply the background subtraction algorithm. It is usually applied when segmenting a series of photographs or films. After subtracting two consecutive photographs or films frames, the result image shows changes in the image, movement of animals, objects. In case of the presented solution, an image on which an averaging filter has been applied using a structural object with a diameter of 5 pixels is subtracted from the original image upon removing noise. The result image shows the differences in the image frequency.

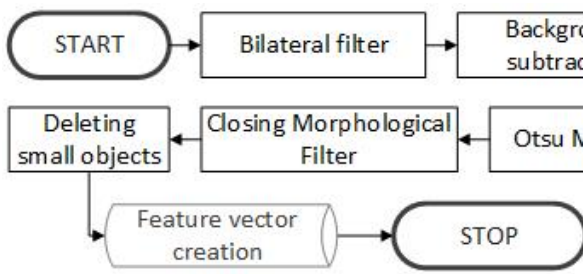


Figure 6: Image processing approach to preprocessing.

Binarization is the third preprocessing step. For this purpose, the algorithm according to the Otsu’s method has been used. As in the event of the biometric algorithms, the green channel has been used for this purpose. Upon completing this stage, the image shows a partially or fully completed character. Then a closing morphological filter is applied. After this operation, the edges in the image are smoothed and some of the noise is deleted. Finally, the noise is removed by applying the last algorithm.

An algorithm has been developed that calculates the area of each object appearing in the image. If it is less than a threshold one, the object is deleted (Figure 12).

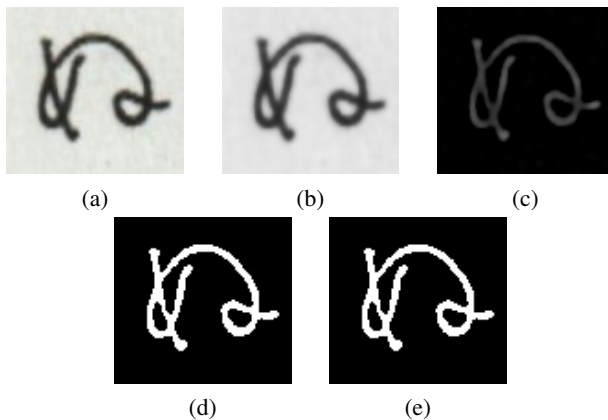


Figure 7: Character after each operation of preprocessing: (a) original image, (b) image after bilateral filter application, (c) image after background subtraction, (d) binarized image and (e) image after small object deletion and closing filter application.

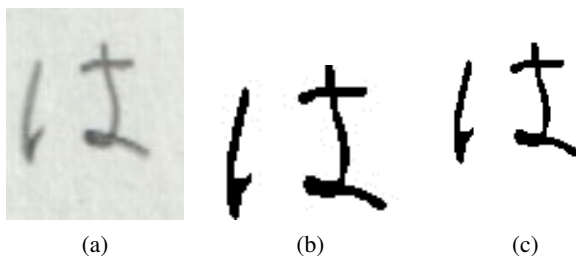


Figure 8: Sample (a), sample preprocessed by hand (b) and sample preprocessed using image processing approach (c).

The final step is to create a feature vector based on the output image. A square structural object with a length and a width of 11 pixels is used. Each vector possesses information about the central pixel, its 120 neighbouring ones, as well as information about its

classification upon preprocessing (1 - background, 0 - object). The result of these operations can be seen in Figure 8.

#### 4.2 Segmentation Using a Neural Network

A neural network has been used to automate the preprocessing process. The large illustrations (2900 x 1200 pixels) representing Japanese writing system have been used to develop the sufficient feature vectors. The principle of operation of this part of the algorithm is presented in Figure 9.

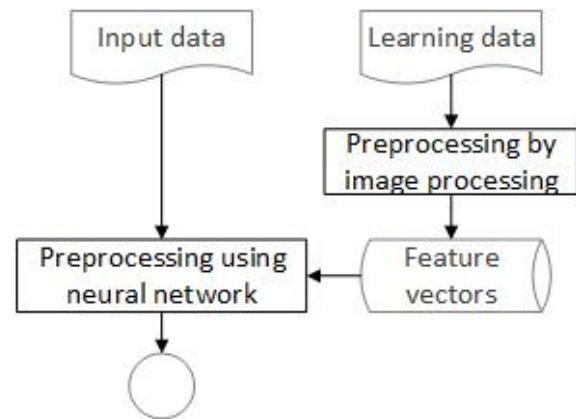


Figure 9: Neural network approach to preprocessing.

A sequential neural network model with the ReLU activation function has been employed for the research purposes. The neural network consists of four layers: an input layer, two hidden layers, and an output layer. The input layer has 363 nodes, the first hidden layer possesses 120 nodes, the second hidden layer has 30 nodes, and the output layer possesses 1 node. As the output, information is received as to whether the pixel under study (located in the centre of the structural object) should be classified as a background or as an object [13]. A comparison of manually processed samples and those processed using neural networks is presented in Figure 10.

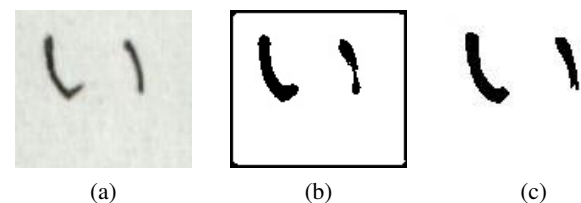


Figure 10: Original sample (a), sample after image processing based preprocessing (b) and sample preprocessed using neural network (c).

#### 4.3 Processing of Single Characters

Upon preprocessing the image containing the object of interest, you obtain the image that requires the further adjustments to extract the features of interest. The next steps of single character processing are shown in Figure 11. The first step is to remove the part of the image that does not carry any information. This is done with the help of the vertical and horizontal histograms. Commencing from the ends of the image, you search for the maximum value of a distance from the edge, where the number of pixels of the object is 0.



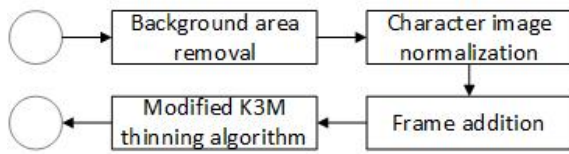


Figure 11: Single character processing

#### 4.4 Development of Feature Vectors and Character Recognition

The next step is the normalisation, which consists in standardising the size of the images [14]. The operations are performed on the images to adjust their size to 62 by 62 pixels. A one-pixel frame is then added to enable to perform further operations on the images (Figure 12). The last operation performed at this stage of the image processing is skeletonization using a modified K3M algorithm. The choice of the skeletonization algorithm has been made using an experimental method.

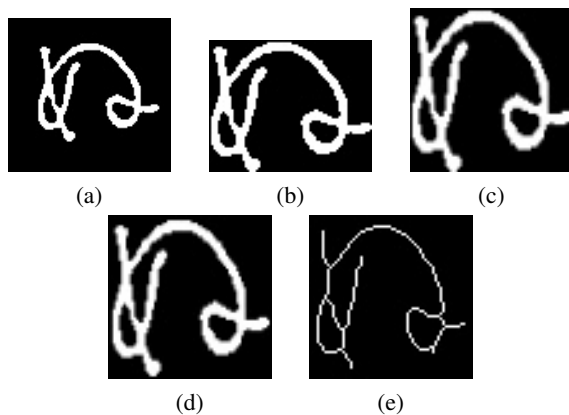


Figure 12: Character after each operation of single character processing step: (a) original image, (b) image after background area removal, (c) image resized, (d) one pixel-wide frame added and (e) image after thinning algorithm application.

While determining the feature vectors, the authors have used the methods applied in the biometric recognition using fingerprints. They use minutiae - characteristic points present in the fingerprints (Figure 13). Among all available types, three types have been selected:

- Ends and beginnings of lines,
- Bifurcations,
- Intersections or Trifurcations.

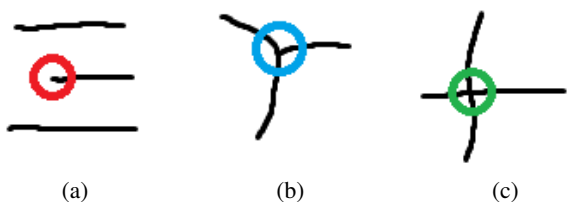


Figure 13: Minutiae types: (a) beginning or ending, (b) bifurcation and (c) trifurcation.

For the purpose of developing a feature vector, the image has been divided into 16 equal areas (16 by 16 pixels) and the sums of each type of minutiae are stored in the appropriate places. Each feature vector contains the feature values and the decision class.

On the basis of an in-house database containing approximately 50 samples for each of the 96 hiragana and katakana characters, experiments have been conducted to test the ability of the SVM (Support Vector Machine) to recognise the characters [15].

## 5 Results

During preprocessing, a prediction match of 86.20% with the intended result has been obtained. It is a relatively high value, but even the smallest deviation from an exemplar one can result in, and has resulted in, the creation of the additional minutiae or a loss of some of them in the image. In the course of the study, the division of samples into a learning set and a test set has been carried out in the ratio 3:2.

Two experiments have been conducted during the character recognition based on the feature vectors of minutiae occurrences. The first experiment has constituted in testing the effect of the choice of preprocessing method on the results of the recognition of the Japanese characters. Three trials have been carried out: data upon a manual segmentation has been used within the first trial, data preprocessed with the use of an algorithm based on the image processing algorithms has been applied in the second trial, and data preprocessed using artificial intelligence has been used in the third one. Table 1 shows the results of the first experiment. Depending on the technique used, the results vary. It can be noted that the more complicated the method processing of a single character is complicated, the lower recognition results were obtained. The highest precision was achieved when using the authors' algorithm is 46.6% when applying image processing approach to preprocessing.

Table 1: First experiment results.

Circumstances	Precision of recognition
Hand preprocessed samples	50,3%
Samples preprocessed using image processing	46,6%
Samples preprocessed using neural network	42,6%

Table 2: Second experiment results.

Circumstances	Precision
Text preprocessed using neural network with added artificial samples	63,3%
Text preprocessed using image processing approach with added artificial samples	65,2%

In the second experiment, a 50 percent chance of development of a descendant from it has occurred. The newly created sample has been subject to twisting and stretching by random parameters. It has been further adapted to the requirements of the algorithm and a new feature vector has been developed on the basis of it. In both

cases the results have been significantly improved by more than 15%, reaching 63.3% when using a neural network and 65.2% when using image processing for preprocessing, as presented in the Table 2.

Table 3: Solution comparison.

Authors	Approach	Database	Declared precision (Best)
Author's Solution	Minutia based	Own database	65,2%
N.T. Ly, C.T. Nguyen, M.Nakagawa [2]	Deep Convolutional Recurrent Network	TUAT Kondate	98,05%
M.Mori, T.Wakahara, K.Ogura [4]	Extended Peripheral Direction Contrubutivity	ETL9B	95%
X. Xia, X. Yu, W. Liu, et al. [7]	Over segmentation strategy	Own database	83.4%
S. Tsuruoka, M. Hattori, M. F. b. A. Kadir, et al. [16]	Modified quadratic discriminate function and users dictionaries	ETL9B	82.5%

## 6 Conclusions

The presented results enable to conclude that the introduced method is not sufficient to allow the unaided character recognition at a high enough level. One of the reasons may be the use of a non-uniform database. The obtained result (62.5%) is the lowest among those compared 3. However, those results allow to suppose that combining this method with other will allow to obtain better solution. The future papers of the authors will focus on the development of the presented solution in order to obtain a system allowing for the off-line text recognition with the use of the hybrid methods.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** This work was supported by grant PROM from the Polish National Agency for Academic Exchange NAWA, supported by the Bialystok University of Technology and partially funded with resources for research by the Ministry of Science and Higher Education in Poland, under Grant WZ/WI-IIT/4/2020.

## References

- [1] P. Szymkowski, K. Saeed, N. Nishiuchi, "SVM Based Hiragana and Katakana Recognition Algorithm with Neural Network Based Segmentation," in Proceedings of the 2020 The 4th International Conference on Graphics and Signal Processing, ICGSP 2020, 51–55, Association for Computing Machinery, New York, NY, USA, 2020, doi:10.1145/3406971.3406978.
- [2] N. T. Ly, C. T. Nguyen, M. Nakagawa, "Training an End-to-End Model for Offline Handwritten Japanese Text Recognition by Generated Synthetic Patterns," in 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), 74–79, 2018, doi:10.1109/ICFHR-2018.2018.00022.
- [3] S. D. Budiwati, J. Haryatno, E. M. Dharma, "Japanese character (Kana) pattern recognition application using neural network," in Proceedings of the 2011 International Conference on Electrical Engineering and Informatics, 1–6, 2011, doi:10.1109/ICEEL.2011.6021648.
- [4] T. Wakahara, Y. Kimura, M. Sano, "Handwritten Japanese character recognition using adaptive normalization by global affine transformation," in Proceedings of Sixth International Conference on Document Analysis and Recognition, 424–428, 2001, doi:10.1109/ICDAR.2001.953825.
- [5] M. Mori, T. Wakahara, K. Ogura, "Measures for structural and global shape description in handwritten Kanji character recognition," 1998, doi:10.1117/12.304621.
- [6] B. Lyu, R. Akama, H. Tomiyama, L. Meng, "The Early Japanese Books Text Line Segmentation base on Image Processing and Deep Learning," in 2019 International Conference on Advanced Mechatronic Systems (ICAMechS), 299–304, 2019, doi:10.1109/ICAMechS.2019.8861597.
- [7] X. Xia, X. Yu, W. Liu, C. Zhang, J. Sun, S. Naoi, "An Efficient off-Line Handwritten Japanese Address Recognition System," in 2019 International Conference on Document Analysis and Recognition (ICDAR), 714–719, 2019, doi:10.1109/ICDAR.2019.00119.
- [8] M. Nixon, A. S. Aguado, Feature Extraction and Image Processing for Computer Vision, Third Edition, Academic Press, Inc., USA, 3rd edition, 2012.
- [9] M. Buczkowski, P. Szymkowski, K. Saeed, "Segmentation of Microscope Erythrocyte Images by CNN-Enhanced Algorithms," Sensors, **21**(5), 2021, doi:10.3390/s21051720.
- [10] M. Tabedzki, K. Saeed, A. Szczepański, "A modified K3M thinning algorithm," International Journal of Applied Mathematics and Computer Science, **26**(2), 439–450, 2016, doi:10.1515/amcs-2016-0031.
- [11] C. Solomon, T. Breckon, Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab, Wiley Publishing, 1st edition, 2011.
- [12] Y. Seki, "Collection of Online and Offline Handwritten Japanese Characters and Handwriting Classification Using the Data," in 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), 264–269, 2020, doi:10.1109/ICFHR2020.2020.00056.
- [13] N.-T. Ly, C.-T. Nguyen, K.-C. Nguyen, M. Nakagawa, "Deep Convolutional Recurrent Network for Segmentation-Free Offline Handwritten Japanese Text Recognition," in 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), volume 07, 5–9, 2017, doi:10.1109/ICDAR.2017.357.
- [14] I. Uddin, D. A. Ramli, A. Khan, J. I. Bangash, N. Fayyaz, A. Khan, M. Kundi, "Benchmark Pashto Handwritten Character Dataset and Pashto Object Character Recognition (OCR) Using Deep Neural Network with Rule Activation Function," Complexity, **2021**, 6669672, 2021, doi:10.1155/2021/6669672.
- [15] K. Nguyen, C. Nguyen, M. Nakagawa, "A Segmentation Method of Single- and Multiple-Touching Characters in Offline Handwritten Japanese Text Recognition," IEICE Transactions on Information and Systems, **E100.D**, 2962–2972, 2017, doi:10.1587/transinf.2017EDP7225.
- [16] S. Tsuruoka, M. Hattori, M. F. b. A. Kadir, T. Takano, H. Kawanaka, H. Takase, Y. Miyake, "Personal Dictionaries for Handwritten Character Recognition Using Characters Written by a Similar Writer," in 2010 12th International Conference on Frontiers in Handwriting Recognition, 599–604, 2010, doi:10.1109/ICFHR.2010.98.