Comparative Study between Texture Feature and Local Feature Descriptors for Silk Fabric Pattern Image Recognition

Thananchai Khamket
Applied Informatics Group
Department of Information Technology
Faculty of Informatics, Mahasarakham University
Maha Sarakham, Thailand
thananchai.k@msu.ac.th

Olarik Surinta
Multi-agent Intelligent Simulation Laboratory (MISL)
Department of Information Technology
Faculty of Informatics, Mahasarakham University
Maha Sarakham, Thailand
Olarik.s@msu.ac.th

ABSTRACT

Thai silk fabrics have unique patterns in different regions of Thailand. The designers may have been inspired and took ideas from the natural environment to create new silk patterns. Hence, many new silk patterns are modified from the original silk pattern. It is challenging for people to recognize a pattern without any prior knowledge and expertise. This paper aims to present a comparative study between texture feature and local feature descriptor for silk pattern image recognition. First, two feature extraction techniques: texture feature and local feature descriptors are proposed to create robustness features from sub-regions that are divided by the grid-based method. Second, the robust features are then classified using the well-known and effective classifier algorithms: K-nearest neighbor (KNN) and support vector machine (SVM) with the radial basis function. We experimented with silk pattern image recognition on two silk fabric pattern image datasets: the Silk-Pattern and Silk-Diff-Pattern. The evaluation results show that the texture feature called the local binary pattern (LBP) when combined with the KNN and SVM algorithms outperforms other feature extraction methods, even deep learning architectures.

CCS Concepts

• Computing methodologies→Image representation • Information systems→Image search • Computing methodologies→Support vector machines.

Keywords

Texture feature; Local feature descriptor; Silk fabric pattern image recognition; Support vector making; K-nearest neighbor.

ICISS 2020, March 19–22, 2020, Cambridge, United Kingdom

ACM ISBN 978-1-4503-7725-6/20/03…$15.00
https://doi.org/10.1145/3388176.3388201

1. INTRODUCTION

In Thailand, people wear silk fabric because silk fibers are tough, lightweight, and can be worn in all weather conditions. People tend to wear a dress that sewn with silk fibers only during celebrations or luxury activities because silk fabric is difficult to maintain and expensive. The pattern design of Thai silk is an identity, so each region has its pattern and style. However, new silk patterns are not much different from the originals due to them being designed and modified from the previous patterns. They are difficult to identify without any knowledge. Maybe only experts of the silk patterns can identify them.

The researchers proposed methods of silk image retrieval and classification using local descriptor and texture feature techniques [7, 19]. Raksaard and Surinta [19] provided the Silk-Pattern dataset to evaluate the image retrieval algorithms. The Silk-Pattern dataset used in the experiments, including 300 silk images in 10 categories, with 30 images per category. The challenge of the Silk-Pattern dataset is to classify the silk pattern in which the test set is randomly cropped 30 and 40% from the whole image. Traditional machine learning techniques and deep learning techniques are proposed to evaluate the performance of silk image retrieval algorithms. As a result, the combination of the histogram of oriented gradients (HOG) and one-nearest neighbor (1NN) performed better than the deep learning techniques (LeNet and AlexNet).

Dittakan and Theera-Ampornpunt [7] proposed texture analysis based on a local binary pattern method: rotated local binary pattern (RLBP) and complete local binary pattern (CLBP), to create the feature vector. Then, three feature selection techniques: chi-squared, information gain, and gain ratio were proposed to select the optimal features. Furthermore, seven classification techniques were used to evaluate the performance of each technique.

Contribution: To address the silk pattern classification problems, as our main contribution, a grid-based method is proposed that uses feature vectors extracted from local binary patterns (LBP) to classify silk fabric pattern images. We evaluate the accuracy result of different techniques on two silk fabric pattern image datasets called a Silk-Pattern dataset [19] and Silk-Diff-Pattern dataset (see Figure 3). In the experiments, we compare the accuracy results between texture feature called local binary patterns (LBP) method and local feature descriptors: Histogram of oriented gradients (HOG) and scale invariant feature transform (SIFT). Additionally, the results also show that the LBP method, when combined with the support vector machine and K-nearest neighbor methods, outperform the convolutional neural network architectures: LeNet and AlexNet architectures.

Paper outline: The remainder of this paper is organized as follows. In Section 2, the silk fabric pattern image recognition approaches are described. In Section 3, we explain a detailed of the silk fabric pattern image dataset. Experimental results are presented in Section 4. Conclusions and suggestions for future work are considered in Section 5.

2. SILK FABRIC PATTERN IMAGE RECOGNITION METHODS

In this research, first, we propose a grid-based method, which has been successful in many areas, such as face recognition [12], people tracking [4], and fast image retrieval [5], for diving the fabric silk image into small regions. Second, we compare two well-known feature extraction techniques: the texture feature and local feature descriptor due to the study of the robustness and effectiveness. Each feature extraction technique is proposed to calculate a feature from a small region. Third, we concatenate the features from each region and use them as the robust feature vector. Finally, two classifiers called the K-nearest neighbor (KNN) and the support vector machine (SVM) [22] methods are used to create a model and classification process. The method of computing local binary pattern (LBP) using the grid-based method is shown in Figure 1.

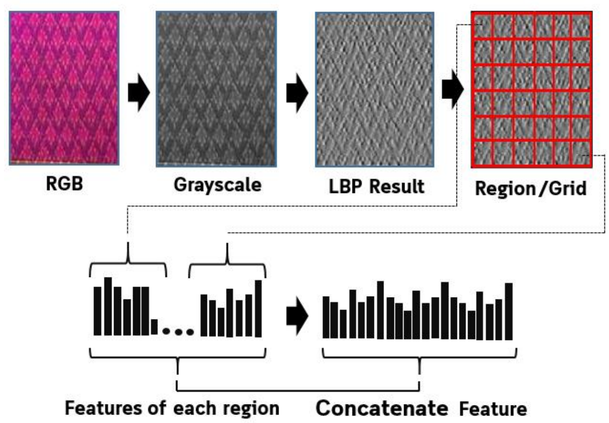


Figure 1. Method of computing the local binary pattern feature from the silk pattern image.

2.1 Texture Feature

Texture features are widely used in many pattern recognition applications such as Melanoma classification [13], hand detection system [16], remote sensing, biomedical imaging [9], character recognition [14], signature recognition [2], and face recognition [11]. The texture always contains information such as colors, intensities, and structure that are represented in the spatial domain. In this research, we focused on the analysis of local structures of silk fabric image using a local binary pattern method (LBP) [23].

Local Binary Pattern (LBP): Wang and He [23] originally proposed the local binary pattern (LBP) for texture spectrum classification. Firstly, the spectrum image was converted to a gray image and resized to 256x256 pixels to calculate the texture spectrum. Secondly, the sub-images of 30x30 pixels were randomly selected. Then, the matrix of 3x3 pixels slides with overlap across the sub-image. Thirdly, the neighborhood pixels around the center pixel of the matrix are transformed into a texture unit with a value of 0, 1, and 2. Finally, the absolute difference values were calculated.

Ojala et al. [18] proposed a robust two-level version of LBP, in which the number of possible texture units was reduced from 6,561 to only 256 (2^8 when 8 is a neighbor pixels). With this method, the matrix size of 3x3 pixels was used. Then, the pixel values around the center pixel of the matrix were transformed into the binary value with a value of 0 and 1 and then considered only the value of 1. From the first to the last neighborhood pixel, the values were calculated as 2^n, where $n = 0,1,..,7$. Finally, the values of the neighborhood pixels were summed and used as a texture unit.

2.2 Local Feature Descriptors

Initially, the local feature descriptors such as the scale-invariant feature transform (SIFT) [15] are proposed for image matching and also applied to image stitching. The local features are used to represent neighborhood pixels of interest points in an image such as edge and corner, called keypoints. In this method, the gradient orientations and magnitudes are computed from each keypoint and count to the orientation histogram. Consequently, the orientation histogram from keypoints is then compared between the two images by the distance function. Hence, the closest distance is the most matching.

In this paper, we use local feature descriptors SIFT and the histogram of oriented gradients (HOG) [6] methods as the representation of the silk fabric image.

2.2.1 Histogram of Oriented Gradients (HOG)

Dalal and Triggs [6] proposed the histogram of oriented gradients (HOG) method for detecting humans in images. Nowadays, it has become a well-known and most successful method in object detection, such as human detection, object tracking and motion detection [1, 8]. In this method, first, the image is transformed into the edge image using a simple convolution kernel, such as Sobel. Second, we divide the image into small blocks. Third, the gradient orientation and magnitudes are computed from each block and stored into orientation bins. Finally, the orientation bins are used for the local feature to create a model using the linear support vector machine (SVM) technique. In the detection process, the detector windows in various sizes are scanned through to the image and then fed the HOG feature to the linear SVM for human classification. In this paper, we extracted HOG features from all grids on the silk fabric image.

The equation of the HOG method [3] can be written as follows:

$$\Phi_f(\mathrm{X}) = \mathbf{Db} * \left[(g_x * \mathbf{X}) \odot (g_y * \mathbf{X})\right] \tag{1}$$

where

X is an image that convolved with the simple kernel g in the horizontal g_x and vertical g_y directions.

Db is orientation bins which are the weighted vote of the gradient orientation and magnitude and normalized using L2-Norm [6].

The g_x and g_y are calculated as follows:

$$g_x = f(x+1, y) - f(x-1, y) \tag{2}$$
$$g_y = f(x, y+1) - f(x, y-1)$$

The gradient magnitude (m) and orientation (θ) are computed as follows:

$$m(x, y) = \sqrt{G_x^2 + G_y^2} \tag{3}$$

$$\theta(x, y) = \tan^{-1} \frac{G_y}{G_x} \tag{4}$$

2.2.2 Scale-Invariant Feature Transform (SIFT)

Lowe [15] invented a method that extracted the distinctive features from scale-invariant keypoints, called the scale-invariant feature transform (SIFT). The features are calculated on the candidate invariant keypoints. Hence, the SIFT features are invariant to different image sizes and orientations.

The SIFT features are computed from each grid due to the grid-based method. The input image is computed using the Gaussian kernel:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{5}$$

where

$I(x, y)$ is the pixel value at location x, y of image I

$G(x, y, \sigma)$ is the Gaussian kernel and σ is the width of the Gaussian kernel

$*$ is the convolution operation

The G_x and G_y are calculated as follows:

$$G_x = L(x + 1, y, \sigma) - L(x - 1, y, \sigma) \tag{6}$$
$$G_y = L(x, y + 1, \sigma) - L(x, y - 1, \sigma)$$

where G_x and G_y are the horizontal and vertical components of the gradients.

The gradient orientation $\theta(x, y)$ and magnitude $m(x, y)$ are computed from the image $L(x, y, \sigma)$ according to Equation (3) and (4).

Then, the region of each keypoint is divided into 4x4 blocks, and each block contains eight orientations. For one key point, the SIFT feature vector contains 128 dimensions.

2.3 Classification Algorithms

In this section, we briefly explain the basic concepts of K-nearest neighbors and the support vector machine algorithms.

2.3.1 K-Nearest Neighbors Algorithm

The K-Nearest neighbors (KNN) algorithm is a supervised learning technique in machine learning. In this method, we compute the distance value between the unknown data and all training data to find the nearest member using a distance function such as the Euclidean function [10, 20]. Euclidean distance function is computed as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2} \tag{7}$$

where N is the number of features. x, y is the data in the training data, and y is the unknown data.

The K nearest members are considered as the candidate members, where K is an odd number. Furthermore, the majority vote is applied to collect the category of candidate members. Finally, the unknown data is assigned to be categorized according to the most vote category. Given an unknown data x_q to be classified. Let $x_1, .., x_k$ denote the k nearest members from training data $(x, f(x))$ that are nearest to x_q. The KNN algorithm [17] is computed as follows:

$$\hat{f}(x_q) = argmax \sum_{i=1}^{k} \delta(v, f(x_i)) \tag{8}$$

where $\delta(a, b) = 1$ if $a = b$, otherwise $\delta(a, b) = 0$.

2.3.2 Support Vector Machine

Vapnik [22] invented the support vector machine (SVM) algorithm for linear binary classification problem. The SVM algorithm finds out the optimal hyperplane, which the largest decision boundary between the two classes, that separates all data points. The training set is $(\mathbf{x}_i, y_i), i = 1, \dots, l$ where $\mathbf{x}_i \in \mathbf{R}^n$ with labels $y_i \in \{+1, -1\}$ [21]. The optimal hyperplane is defined as follows:

$$\mathbf{W}^T\mathbf{X} + w_0 = 0 \tag{9}$$

where $\mathbf{W}$ is the weight vector. $\mathbf{W}^T\mathbf{X} + w_0 > 0$ for $y = +1$ and $\mathbf{W}^T\mathbf{X} + w_0 < 0$ for $y = -1$.

To deal with the non-linear problem, many kernel functions such as radial basis function (RBF) and polynomial kernel are proposed. In this paper, we choose the RBF kernel to handle the non-linear problem. The RBF kernel is defined as follows:

$$K(x_i, x_j) = exp\left[-\gamma \left\|x_i - x_j\right\|^2\right] \tag{10}$$

where γ is the RBF kernel parameter.

$\left\|x_i - x_j\right\|^2$ is the Euclidean distance from the set of feature points x_j

3. THE SILK FABRIC PATTERN IMAGE DATASET

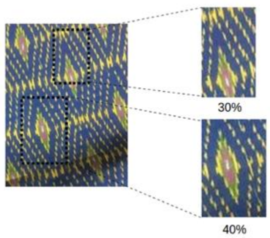
Raksaard [19] introduced a new dataset of Thai silk patterns. The objective of this dataset was to evaluate the retrieval systems including feature extraction method, image classification, and retrieval. The Silk-Pattern dataset was collected from a silk shop in Maha Sarakham, located in the northeast of Thailand. Thai silk fabric includes two parts; the main pattern and the bottom of the fabric, called fabric feet, as shown in Figure 2. In this dataset, the researcher decided to use only the main pattern of the silk fabric, as shown in Figure 2(b).

Figure 2. Illustration of the a) Thai silk fabric, b) main pattern and c) fabric feet.

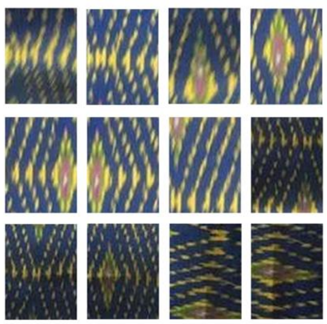
The Silk-pattern dataset consists of ten Thai silk pattern classes and contains 300 images captured using a smartphone to illustrate the silk fabric from different orientations. The Silk-Pattern images are stored in the RGB color space with a size of 450x650 pixels. Sample images of the Silk-Pattern dataset are illustrated in Figure 3.

In the Silk-Pattern dataset, the test set is randomly cropped only 30% and 40% from the whole silk fabric pattern image, as shown in Figure 4(a) and cropped three times from one silk image. The Silk-Pattern images consist of 300 training samples, 900 test samples of cropping 30% (Crop-30), and 900 test samples of cropping 40% (Crop-40). Figure 4(b) illustrates the sample images of the Crop-30.

Figure 3. Illustration of the training set of the Silk-Pattern dataset.



(a)



(b)

Figure 4. Illustration of the test set of the Silk-Pattern dataset. (a) randomly cropped 30% and 40%. (b) The sample images with cropped 30%.

In addition, we introduce a new dataset of Thai silk pattern images, which is more complex and challenging, called the Silk-Diff-Pattern dataset. The advantage of our new dataset is that the silk pattern images include all components of the silk pattern; pattern and fabric feet. Examples of the Silk-Diff-Pattern dataset are shown in Figure 5. Our dataset also has ten classes and contains 300 images. We used the same process as with the Silk-Pattern dataset to collect the test data. The test data consists of two sets; 900 samples of Crop-30 and 900 samples of Crop-40. Furthermore, we carefully checked all the test images. Also, the test image contained only the pattern of the silk fabric.

Figure 5. Illustration of the training set of the Silk-Diff-Pattern dataset.

4. EXPERIMENTAL RESULTS

We briefly explain the experimental setups used for the dataset. After that, the results are presented and discussed.

To provide data for the silk fabric image recognition, we used two silk datasets; the Silk-Pattern and the Silk-Diff-Pattern datasets. Each dataset contains 300 training images. We also divided the image into 2x2 blocks according to the grid-based method. Then, we calculated the robust features by sending each block to the texture feature and local feature descriptors. To test the performance of image recognition, we used two test sets, including Crop-30 and Crop-40 sets. These test sets were randomly cropped three times from the silk fabric image. Therefore each test set included a total of 900 images.

The experimental results are based on 5-fold cross-validation. We compute average recognition accuracies and standard deviations for all experiments. Also, the grid-based method is performed.

Table 1. Evaluation of the classification results on the Silk-Pattern dataset

Method	Accuracy (%)	
	Crop-30	Crop-40
LBP+KNN	**93.48±0.64**	**99.26±0.12**
HOG+KNN [19]	92.05±0.31	89.73±0.79
SIFT+KNN	23.03±0.36	57.99±0.33
LBP+SVM	92.61±0.16	98.46±0.08
HOG+SVM	74.92±1.94	82.68±4.67
SIFT+SVM	42.80±6.98	40.24±1.08
LeNet [19]	64.06±2.25	76.98±2.29
AlexNet	44.70±0.94	55.58±1.04

Table 2. Evaluation of the classification results on the Silk-Diff-Pattern dataset

Method	Accuracy (%)	
	Crop-30	Crop-40
LBP+KNN	65.70±0.84	79.20±0.45
HOG+KNN	76.58±0.98	88.45±0.23
SIFT+KNN	49.59±0.99	59.44±1.14
LBP+SVM	**90.97±0.74**	**92.81±0.74**
HOG+SVM	48.18±1.55	73.22±0.95
SIFT+SVM	41.40±0.94	34.58±1.04

Table 1 shows the recognition accuracy results of the eight different techniques. According to these experimental results, we combined both the texture feature and local feature descriptors with the classifiers, including the K-nearest neighbor (KNN) and support vector machine (SVM) algorithms. The combination of the local binary pattern (LBP), which is the texture feature, and the K-nearest neighbor algorithm (KNN), called LBP+KNN perform the best for the Silk-Pattern dataset on both test sets; Crop-30 and Crop-40. The accuracy result of the LBP+KNN method increased by 1% more than the HOG+KNN [19] method. Moreover, this method shows a significant performance gain of 20-30% compared to the LeNet and AlexNet architectures [19], which are the deep learning techniques.

In Table 2, we report the results of the average accuracy and standard deviation on the Silk-Diff-Pattern dataset. The results show that the LBP with the support vector machine (SVM), called the LBP+SVM, achieves the best accuracy performance on both test sets. The HOG+KNN method is around 4-14% more accurate than the HOG+KNN method, which is the second-best method. Surprisingly, the scale-invariant feature transform (SIFT) method performed very low performance while combined with KNN and SVM algorithms on two datasets.

5. CONCLUSION

In this paper, we compared methods of silk pattern image recognition. First, we proposed a grid-based method for dividing the image into sub-areas. Second, the two well-known feature extraction techniques: texture feature and local feature descriptor, were proposed to extract robust features from each sub-area. The grid-based method allowed the feature extraction technique to extract more useful features. Finally, we concatenated the robust features and fed them to the classifier algorithms: K-nearest neighbor and the support vector machine algorithms.

The results showed that the LBP algorithm outperforms other methods when combined with both the KNN and the SVM algorithms. On the Silk-Pattern dataset, the LBP+KNN method performs much better than the other methods for all test sets. Subsequently, the LBP+SVM method shows the best performance on the Silk-Diff-Pattern dataset. We also compared our results with two basic deep learning architectures: LeNet and AlexNet architectures. We found that the deep learning architectures showed low accuracies when the training set is inadequate. To the best of our knowledge, the scale-invariant feature transform (SIFT) algorithm always showed the best performance. On the other hand, surprisingly, the SIFT algorithm had very low performance when combined with KNN and SVM algorithms on the silk image dataset.

We conclude that deep learning architecture obtain high accuracy on many pattern recognition problems. In future work, we want to study deep learning architecture and apply it to the silk image dataset. We are also interested to improve the performance by using transfer learning and data augmentation.

6. ACKNOWLEDGMENTS

This research was supported by the Faculty of Informatics at Mahasarakham University, Thailand (IT2-03/2561).

7. REFERENCES

[1] A. H. Ahmed et al. 2017. Human Detection Using HOG-SVM, Mixture of Gaussian and Background Contours Subtraction. *13th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (2017), 334–338.

[2] Bharadi, V.A. et al. 2015. Performance Analysis of Grid & Texture Based Feature Vector for Dynamic Signature Recognition. *International Conference on Pervasive Computing (ICPC)* (2015), 1–6.

[3] Bristow, H. and Lucey, S. 2014. Why do linear SVMs trained on HOG features perform so well? (2014).

[4] Chen, L. et al. 2015. Hierarchical Grid-based Multi-People Tracking-by-Detection With Global Optimization. *IEEE Transactions on Image Processing*. 24, 11 (2015), 4197–4212. DOI:https://doi.org/10.1109/TIP.2015.2451013.

[5] Choi, S. and Han, S. 2014. Fast image retrieval with grid-based keypoint detector and binary descriptor. *International Conference on Information and Communication Technology Convergence (ICTC)* (2014), 679–680.

[6] Dalal, N. and Triggs, W. 2005. Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (2005), 886–893.

[7] Dittakan, K. and Author, N.T.-A. 2018. Pum-Riang Thai Silk Pattern Classification Using Texture Analysis. *Pacific Rim International Conference on Artificial Intelligence* (2018), 83–90.

[8] Drozdz, M. and Kryjak, T. 2016. FPGA Implementation of Multi-scale Face Detection Using HOG Features and SVM Classifier. *Image Processing & Communications*. 21, 3 (2016), 27–44. DOI:https://doi.org/10.1515/ipc-2016-0014.

[9] Humeau-HeurTier, A. 2019. Texture Feature Extraction Methods: A Survey. *IEEE Access*. 7, (2019), 8975–9000. DOI:https://doi.org/10.1109/ACCESS.2018.2890743.

[10] Jie, H. et al. 2018. An Improved kNN Based on Class Contribution and Feature Weighting. *Measuring Technology and Mechatronics Automation, International Conference on* (2018), 313–316.

[11] Kambi Beli, I.L. and Guo, C. 2017. Enhancing Face Identification Using Local Binary Patterns and K-Nearest Neighbors. *Journal of Imaging*. 3, 3 (2017). DOI:https://doi.org/10.3390/jimaging3030037.

[12] Karaaba, M. et al. 2015. Robust face recognition by computing distances from multiple histograms of oriented gradients. *Computational Intelligence (SSCI), IEEE Symposium Series on* (2015), 203–209.

[13] Kavitha, J.C. and Suruliandi, A. 2016. Texture and Color Feature Extraction for Classification of Melanoma using SVM. *International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE)* (2016), 1–6.

[14] Liu, L. et al. 2010. Simplified Local Binary Pattern Descriptor for Character Recognition of Vehicle License Plate. *Computer Graphics, Imaging and Visualization, Seventh International Conference on* (Aug. 2010), 157–161.

[15] Lowe, D.G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*. 60, 2 (2004), 91–110. DOI:https://doi.org/10.1023/B:VISI.0000029664.99615.94.

[16] Misra, S. and Laskar, R.H. 2017. Taxonomy of Texture and Color-Texture Features for Developing Hand Detection System under Non-Ideal Conditions. *14th IEEE India Council International Conference (INDICON)* (2017), 1–6.

[17] Mitchell, T. 1997. *Machine Learning*. McGraw Hill.

[18] Ojala, T. et al. 1996. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*. 29, 1 (1996), 51–59. DOI:https://doi.org/https://doi.org/10.1016/0031-3203(95)00067-4.

[19] Raksaard, N. and Surinta, O. 2018. Comparative Study Between Local Descriptors and Deep Learning for Silk Pattern Image Retrieval. *Journal of Science and Technology Mahasarakham University*. 37, 6 (2018), 736–746.

[20] Roempluk, T. and Surinta, O. 2019. A Machine Learning Approach for Detecting Distributed Denial of Service

Attacks. *Digital Arts, Media and Technology, International Conference on* (2019), 146–149.

[21] Surinta, O. et al. 2015. Recognizing Handwritten Characters with Local Descriptors and Bags of Visual Words. *6th International Conference on Engineering Applications of Neural Networks (EANN)* (2015), 255–264.

[22] Vapnik, V.N. 1998. *Statistical Learning Theory*. Wiley.

[23] Wang, L. and He, D.-C. 1990. Texture classification using texture spectrum. *Pattern Recognition*. 23, 8 (1990), 905–910. DOI:https://doi.org/https://doi.org/10.1016/0031-3203(90)90135-8.