Plant Leaf Image Recognition using Multiple-grid Based Local Descriptor and Dimensionality Reduction Approach

Thipwimon Chompookham
PhD Student, Department of Information Technology
Faculty of Informatics,
Mahasarakham University, Thailand
61011261002@msu.ac.th

Sarayuth Gonwirat
PhD Student, Department of Information Technology
Faculty of Informatics,
Mahasarakham University, Thailand
61011261003@msu.ac.th

Siriwiwat Lata
PhD Student, Department of Information Technology
Faculty of Informatics,
Mahasarakham University, Thailand
61011261004@msu.ac.th

Sirawan Phiphiphatphaisit
PhD Student, Department of Information Technology,
Faculty of Informatics, Mahasarakham University,
Thailand
61011261005@msu.ac.th

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

ABSTRACT

The identification process of plant species is one of the significant and challenging problems. In this research area, many researchers have focused on identifying the plant leaf images because the leaves of a plant are found almost all year round. The achieve method of the plant leaf image recognition is based on unique extraction features from the plant leaf and using the well-known machine learnings as a classification method. As a result, recognition accuracy was often not very high. In order to improve recognition accuracy, we proposed a multiple grids technique based on the local descriptors and dimensionality reduction. Firstly, we divided the plant leaf image according to grid size and calculated the local descriptors from each grid. Secondly, the dimensionality reduction is proposed to transform and decrease the correlated variables of the feature vector. Finally, the feature vector with a relatively low-dimensional is transferred to the machine learning techniques, which are the support vector machine and multi-layer perceptron algorithms. We have evaluated and compared the proposed algorithm with the bag of visual words method and the deep convolutional neural network (including AlexNet and GoogLeNet architectures) on the Folio leaf image dataset. The experiments show that the proposed algorithm has improved and obtained very high accuracy on plant leaf image recognition.

CCS Concepts

• Computing methodologies→Object recognition • Computing methodologies→Support vector machines • Computing methodologies→Neural networks.

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

ACM ISBN 978-1-4503-7725-6/20/03…$15.00
https://doi.org/10.1145/3388176.3388180

Keywords

Plant leaf recognition; Multiple grids approach; Local descriptor; Dimensionality reduction; Support vector machine; Multi-layer perceptron.

1. INTRODUCTION

Plants are living things that relate directly to humans in that they are used as a food and medicine. Botanists have collected and studied various plant species which can be of some benefit for humans. However, while the physical characteristics of some plants are similar, they have different benefits and toxins. As such, the ability to distinguish the types of plants requires an advanced knowledge of botany. A typical plant classification problem is the diversity of plants and their botanical characteristics. Researchers find that classification of plant species is a challenging problem. Nowadays, computer vision and machine learning are used as instruments for recognition and classification.

This research aims to use image processing and machine learning for plant classification by classifying plant leaf photos taken from the laboratory.

Wäldchen and Mäder [1] said that over the past 10 years, researchers have tried to bring various parts of the plant including leaves, plant blossom and fruits [2, 3, 4] to study plant classification. Most researchers are interested in the leaves because the plant leaves have specific shape, surface shape, color, and leaf structure [5, 6]. The images of plant leaves used in this research are divided into two forms including 1) Plant leaf taken in an outside environment [7] and 2) Plant leaf taken in a laboratory on a white background [8-10].

In [11], used curvature-scale space for recognizing margin shape (Margin shape recognition) and Leaf identification from the characteristics of plant leaves by Semi-supervised fuzzy C means (FCM) for training the margin shape with 12 terms. Then, it learns with the Pl@ntLeaves database, which is divided into three subsets including Scan, Pseudoscan, and Photograph by using Top-K in the test. The result found that the given K=10 in dataset Scan, Pseudosca, and Photograph, accuracy rates were estimated as 95%, 92%, and 80%, respectively.

Image data of plant leaves taken in the laboratory is presented by Munisami et al. [8] The Folio dataset is a dataset which contains 32 species of plant images. The research suggested the methods to find feature extraction technique including plant shape and color histogram, then used it to classify the plant leaves. The result was an accuracy rate of 87.3%

In [10], have tested the Folio dataset by using deep convolutional neural networks (CNNs) which includes architecture types AlexNet architecture and GoogleNet architecture. Another method was classical local descriptors which include a histogram of oriented gradients (HOG) and bags of visual words (BOW). The support vector machine (SVM), multi-layer perceptron (MLP), and K-nearest neighbor (KNN) were used as instruments for classification of plant leaves images. In the experiments, databases were divided into two parts: 80% dataset for training and 20% of dataset for testing. The result showed that AlexNet Architecture type fine-tuned was the most accurate method, with a 97.67% accuracy rate. Moreover, in the research [9] they used 6 methods of data augmentation. The methods were rotation, blur, contrast, scaling, illumination, and projective transformation. These methods can add up to 25 times the number of datasets for training. Researchers increased the number of images to 11,125 images and tested by using AlexNet architecture. The result could be summarized as increasing dataset contrast methods, which can increase the accuracy rate to 99.04%. When tested with the GoogleNet architecture, it was found that the illumination method had the highest accuracy rate at 99.42%.

Another set of plant leaves images taken in the laboratory was the Flavia dataset presented by [12]. There are 32 species of plant images. The characteristics of the shape feature of the plant leaves were studied before being classified by the SVM method. The accuracy rate was 85%. At the same time, the research [13] developed an automatic leaf classification system by using feature extraction types colored SIFT in cooperation with SVM. In [14] used geometrical and shape feature in cooperation with SVM. The accuracy rates from the test were 98% and 97.69% respectively. In the case of leaf classification by MLP, the research [15] used feature extraction types texture-based with constraint. MLP method must have an input layer, hidden layer, and output layer as 44, 30, and 31 nodes respectively. The accuracy rate of the test was 87.1%.

Contributions: The research focuses on the importance of plant leaf recognition by experiment with (Folio dataset) which collects 32 different species of plants. This research presents multiple grids and dimensionality reduction based descriptors approach, which is simple but effective. The multiple grids divide plant leaves into sub-regions, then it brings the sub-region to calculate the special features using various feature extraction techniques that pull out the distinctive characteristics of the plant leaves. The methods are a histogram of oriented gradients (HOG), local binary pattern (LBP), and color histogram. Finally, the feature will be fed to the dimensionality reduction method by using principal component analysis (PCA) in order to reduce the feature vector size of each method. The size reductions have direct effect on training time and increase the recognition efficiency as well. In this paper, the feature vector was used in training and recognition by a support vector machine (SVM) and Multi-layer perceptron (MLP). This method obtained a very high recognition rate when compared to the deep learning method.

Paper Outline: This paper has been organized as follows. In Section 2, the method for plant leaf recognition is explained. Section 3, the dataset and pre-processing with plant leaf images, which are used in our experiments are described. Section 4, experimental results is presented. The last section discusses the significant findings from this study and describes future work.

2. PROPOSED PLANT LEAF RECOGNITION METHOD

In this study, we use multiple grids and dimensionality reduction based on three feature extraction techniques. Figure 1 shows the process of this research. The input images were forwarded to the multi-grid based process to divide the images into (Sub-regions), then a sub-region was calculated by using three techniques of feature extraction. Each technique was calculated by principal component analysis (PCA) method in order to decrease the amount of feature vector. Finally, researchers put all FE+PCA in concatenate to use it as a feature vector ($f_1, f_2, \ldots, f_n$) then forwarded it to the classification process.

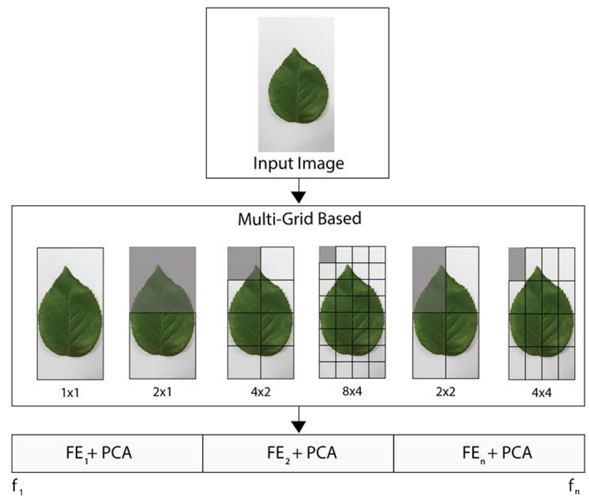

Figure 1. Proposed plant leaf recognition method.

2.1 Multiple Grid-based Technique

The working process of multiple grid-based technique is to divide the picture of the leaves (Input image) into sub-regions by using a grid in the determination of the sub-regions. In these experiments, the Grids were determined at 6 different types, including Grid size of 1x1, 2x1, 4x2, 8x4, 2x2, and 4x4. After that, each sub-area was calculated to find the feature vector by using HOG, LBP, and color histogram.

2.2 Feature Extraction Techniques

2.2.1 Histogram of Oriented Gradients (HOG)

HOG introduced by Dalal and Triggs [16], a method that extracts the characteristics of the image by calculating the oriented gradients from gradient Image by finding gradient in (Horizontal) (G_x) and (Vertical) (G_y) which is calculated from pixel intensities ($I(x, y)$) at (x, y) as the following equation:

$$G_x = I(x+1, y) - I(x-1, y) \quad (1)$$

$$G_y = I(x, y+1) - I(x, y-1) \quad (2)$$

After that, the magnitude (M) and gradient orientation (θ) are calculated as the following equation:

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

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

where $M(x, y)$ is magnitude of gradients, $\theta_{x,y}$ is gradient orientation at x, y. Then, gradient orientation values will be taken to the weighted vote process and will be kept in the orientation bins (β) [17].

Finally, gradient orientation values, which are kept in each orientation bin will be taken to do the Normalization by L2-norm method.

2.2.2 Local Binary Patterns (LBP)

LBP was proposed in [18] for invariant texture classification. LBP is designed for extracting characteristics of pixel points from Neighborhood pixels which are calculated from gray values as the following equation:

$$\mathrm{LBP}_{\mathrm{P,R}} = \sum_{\mathrm{p}=0}^{\mathrm{P}-1} \mathrm{s}(\mathrm{g_p} - \mathrm{g_c})2^{\mathrm{p}} \quad (5)$$

where

$\mathrm{g_c}$ is the gray value of the central pixel.
$\mathrm{g_p}$ is the gray value of its neighbor pixels.
P is the total number of involved neighbors.
R is the radius of the neighborhood.

Then, the central pixel will be used as Threshold value (T) to compare with Neighborhood pixels values, $\mathrm{s(x)} = \begin{cases} 1, \mathrm{x} \geq \mathrm{T} \\ 0, \mathrm{x} < \mathrm{T} \end{cases}$ The next step is to bring the value 1 and 0 from Neighborhood pixels to come together as concatenate. Then, it was converted to decimal. Finally, researchers bring the values into the specified bins.

2.2.3 Color Histogram

This research used two types of color models. There are RGB and HSV color models, while HSV used only hue (H) values because hue values show the true color. Therefore, colors values used for histogram creation consist of red (R), green (G), blue (B), and hue. While, histogram of color RGB values consist of 256 color values, H consist of 360 color values.

2.3 Dimensionality Reduction

From the Multiple-grid based method, a lot of sub-region will be created, which is used for calculation of unique features. This causes high dimensionality of the feature vector and results in computational complexity. Therefore, dimensionality reduction is one of the best ways to minimize the feature vector. This research uses PCA [19] in feature vector reduction. Feature vector from each technique has been reduced to only 80 Features. These techniques improved the accuracy rate as well.

2.4 Classification Algorithms

This research used two types of classification algorithm, including support vector machine (SVM) [20] and multi-layer perceptron (MLP) [21]. The SVM used RBF kernel and MLP by determining the hidden layer as two layers. The dropout method was selected for prevention of an overfitting situation.

3. PLANT LEAF DATASET

The plant leaves images used in the experiment were taken in the laboratory. Thus, most images have a white background. The background makes the leaves prominent and clearly separates them from the background.

3.1 Folio Dataset

The leaves data used in the experiment was the Folio dataset, presented in 2015 [8]. The data represents 32 species of leaves plant images (see Figure 2). All images were taken in the laboratory with a white background. All images were saved in the JPEG format. Size of images is 2322x4128 and 2448x3264 pixel resolution. The plants were in the University of Mauritius farm. Twenty images of each plant species were collected except for mulberry with 19 images and eggplant with 18 images. The dataset contains 637 images.

Some plant leaves are shown twice; they are the same type, but they have different shapes (For example, papaya, chrysanthemum, and ketembilla). Image differentiation of each species are shown in Figure 3. Some plant leaves still have similar shape, e.g., star apple and pomme jacquot (See Figure 4). The factors mentioned above have directly affected the accuracy of recognition.

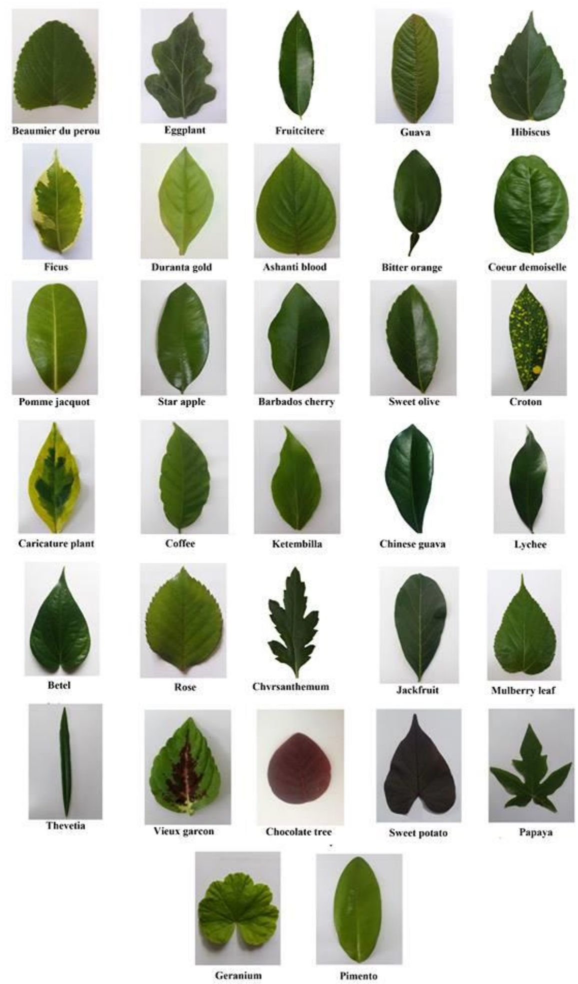

Figure 2. Examples of 32 plant leaves of the Folio dataset.

a)

b)

c)

Figure 3. Some variety examples of plant leaves, a) papaya, b) chrysanthemum, and c) ketembilla leaf images of the Folio dataset.

a) b)

Figure 4. **Similarities shape between different plant leaves. a) The images of star apple and b) pomme jacquot leaves.**

3.2 Dataset Pre-processing

The process of preparing the image of the plant leaves from the Folio dataset is very simple. The process starts by converting all the images to black and white in order to find the plant leaves area (Region of interest: ROI), Then, crop to get ROI. The next step is to check the image of the leaf in the horizontal position and then rotate the image to vertical shape (See Figure 2). After that, the image resizes are resize to 400 pixels. The width of each picture will have different sizes because some plant leaves, e.g., thevetia, lychee, and fruit citere are slender. Therefore, if we assign the size as 400x200 pixel, the plant leaves images will be distorted. Finally, when ROI was identified and resized, the color image for feature extraction process was used.

4. EXPERIMENTAL RESULTS

We compared the feature extraction techniques (i.e., color histogram, local binary pattern (LBP), a histogram of oriented gradients (HOG), and principal component analysis (PCA)) and HOG-bag of word (HOG-BOW) to deep learning techniques (AlexNet and GoogleNet).

In these experiments, we used 5-fold cross-validation to evaluate the results of the plant leaf recognition methods. We used the recognition rate (accuracy) and standard deviation to measure the performance of each feature extraction technique. For the experiments using the support vector machine (SVM) algorithm, the grid-search technique was used to search the best parameters. The best C and gamma (γ) parameters of the SVM with the RBF kernel are 100 and 0.1, respectively. For the multi-layer perceptron (MLP), two hidden layers are used where the size of each layer is 512 and 512 hidden units, respectively. The dropout regularization is used to prevent neural networks from overfitting. The dropout rates of 0.5 for all hidden units are selected. As for the output layer, the softmax function is used. Table 1 and Table 2 show the results (average test accuracy and standard deviation).

The results in Table 1 show the recognition performances obtained from the combination of multiple grid approaches with feature extraction techniques, the result of the HOG-BOW method, and the training time on the Folio dataset. We can see 15 different results. Here, the HOG-BOW method obtains an inferior performance compared to the other feature extraction techniques. On the other hand, the Color-Histogram-LBP-HOG-PCA, when combined with the SVM with the RBF kernel algorithm, significantly outperforms the other techniques and provides a high accuracy of 99.06%. Subsequently, the plant leaf recognition obtains a high accuracy of 98.75% when combined with the Color-Histogram-LBP-HOG-PCA and MLP algorithm.

Table 1. Plant leaf recognition results of the 15 different techniques on the Folio dataset

Multiple Grid Methods	Training Time (Sec)		Accuracy (%)	
	SVM	MLP	SVM	MLP
Color-Histogram	221.86	232.42	96.25±1.87	95.94±1.94
LBP	278.80	284.80	94.45±1.06	91.87±2.22
HOG	201.27	206.83	94.14±2.45	94.14±2.34
Color-Histogram-PCA	182.88	189.49	97.73±1.30	97.11±1.28
LBP-PCA	278.15	285.29	94.14±1.06	94.14±1.74
HOG-PCA	202.12	209.53	93.83±2.62	93.91±1.83
Color-Histogram-LBP	496.61	511.65	97.81±1.15	96.09±1.65
Color-Histogram-HOG	419.10	435.47	98.13±1.39	96.64±1.38
LBP-HOG	481.14	489.10	97.50±1.46	96.87±1.98
Color-Histogram-LBP-HOG	697.46	716.77	98.67±0.91	97.42±1.48
Color-Histogram-LBP-PCA	460.96	469.78	98.67±1.11	98.28±1.51
Color-Histogram-HOG-PCA	384.91	393.20	98.59±1.46	98.28±1.32
LBP-HOG-PCA	480.19	488.94	97.50±1.46	97.58±1.01
Color-Histogram-LBP-HOG-PCA	663.01	672.19	**99.06±0.89**	**98.75±0.92**
HOG-BOW [9]	-	-	92.78±2.17	92.37±1.78

Table 2. Comparing results between proposed method and fine-tuned deep learning methods on the Folio dataset

Method	Accuracy (%)
AlexNet [10]	97.67±1.60
GoogleNet [10]	97.63±1.84
AlexNet data augmentation (Contrast) [9]	99.04±0.38
GoogleNet data augmentation (Illumination) [9]	**99.42±0.38**
Proposed Method (Color-Histogram-LBP-HOG-PCA)	99.06±0.89

We also compared our proposed method with the find-tuned deep convolutional neural networks (CNNs), which are AlexNet and GoogleNet architectures [9]. Furthermore, the data augmentation techniques consisting of contrast and illumination [10] techniques were compared as well. The accuracy results between our proposed method and fine-tuned deep CNNs are shown in Table 2.

The performance of our proposed multiple grids and dimensionality reduction based descriptors approach reaches 99%. Our proposed method performs better than the deep CNN architectures. However, the fine-tuned deep CNNs with the combined data augmentation technique, (contrast and illumination), slightly outperform our proposed method. This is because, the fine-tune deep CNNs were trained from millions of images, and the training data increased 4,005 images from the data augmentation technique our proposed method train and create the plant leaf recognition model from only 510 plant leaf images.

5. CONCLUSION

In this paper, we have investigated many different plant leaf recognition techniques on a Folio dataset. From the experimental results, we conclude that the performance of multiple grids and dimensionality reduction based descriptors, which is our proposed method, is much better than the histogram of oriented gradients combined with bag-of-words technique and fine-tuned deep CNN architectures which are AlexNet and GoogleNet architectures as well. We also have shown that the principal component analysis (PCA), which is the dimensionality reduction technique, increased the accuracy performance and decreased the number of the feature vector of the plant leaf recognition system. Nevertheless, the data augmentation technique can increase the accuracy performance of the plant leaf recognition system. This technique added more than 4,000 illumination images to the training set. Subsequently, we used only 510 images to train the plant leaf recognition system. As a result, the accuracy result of our proposed method is slightly decreased than the fine-tuned deep CNNs with the combined data augmentation technique.

According to the high accuracy of the deep CNNs, in future work, we would like to study the effect of parallel CNN architecture and use this architecture to train the plant leaf images. This technique maybe necessary to improve training times and accuracy performance.

6. REFERENCES

[1] Wäldchen, J. and Mäder, P. 2018. Plant species identification using computer vision techniques: a systematic literature review. In *Archives of Computational Methods in Engineering.* 25, 2 (2018), 507–543.

[2] Caballero, C. and Aranda, M. C. 2010. Plant species identification using leaf image retrieval. In *ACM International Conference on Image and Video Retrieval (CIVR)*, 327–334.

[3] Cerutti, G., Tougne, L., Mille, J., Vacavant, A. and Coquin, D. 2013. Understanding leaves in natural images - A model-based approach for tree species identification. In *Computer Vision and Image Understanding.* 117, 10 (2013), 1482–1501.

[4] Cho, S. Y. 2012. Content-based structural recognition for flower image classification. In *7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 541–546.

[5] Caglayan, A., Guclu, O. and Can, A.B. 2013. A plant recognition approach using shape and color features in leaf images.In *Image Analysis and Processing (ICIA)*, 161–170.

[6] Hossain, J. and Amin, M. A. 2010. Leaf shape identification based plant biometrics. In *13th International Conference on Computer and Information Technology (ICCIT)*, 458–463.

[7] Wang, X. F., Huang, D. S., Du, J. X., Xu, H. and Heutte, L. 2008. Classification of plant leaf images with complicated background. In *Applied Mathematics and Computation.* 205, 2 (2008), 916–926.

[8] Munisami, T., Ramsurn, M., Kishnah, S. and Pudaruth, S. 2015. Plant leaf recognition using shape features and colour histogram with k-nearest neighbour classifiers.In *Procedia Computer Science*. 58, (2015), 740–747.

[9] Pawara, P., Okafor, E. and Schomaker, L. 2017. Data augmentation for plant classification. In *Advanced Concepts for Intelligent Vision Systems. (ACIVS)*, 615-626.

[10] Pawara, P., Okafor, E., Surinta, O., Schomaker, L. and Wiering, M. 2017. Comparing local descriptors and bags of visual words to deep convolutional neural networks for plant recognition. In *the 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, 479–486.

[11] Cerutti, G., Tougne, L., Coquin, D. and Vacavant, A. 2013. Curvature-scale-based contour understanding for leaf margin shape recognition and species identification. In *the International Conference on Computer Vision Theory and Applications (VISAPP)*, 227-284.

[12] Salman, A., Semwal, A., Bhatt, U. and Thakkar, V. M. 2017. Leaf classification and identification using Canny edge detector and SVM classifier. In *International Conference on Inventive Systems and Control (ICISC)*, 1–4.

[13] Arafat, S. Y., Saghir, M. I., Ishtiaq, M. and Bashir, U. 2016. Comparison of techniques for leaf classification. In *Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, 136–141.

[14] Khmag, A., Al-Haddad, S. A. R. and Kamarudin , N. 2017. Recognition system for leaf images based on its leaf contour and centroid. In *IEEE 15th Student Conference on Research and Development (SCOReD)*,467-472.

[15] Chaki, J., Parekh, R. and Bhattacharya, S. 2015. Plant leaf recognition using texture and shape features with neural classifiers. In *Pattern Recognition Letters.* 58, (2015), 61–68.

[16] Dalal, N. and Triggs, B. 2005. Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 886–893.

[17] Karaaba, M., Surinta, O., Schomaker, L. and Wiering, M. 2015. Robust face recognition by computing distances from multiple histograms of oriented gradients.In *IEEE Symposium Series on Computational Intelligence (SSCI)*, 203-209.

[18] Ojala, T., Pietikainen, M. and Maenpaa, T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. In *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 24, 7 (2002), 971–987.

[19] Cootes, T. F., Taylor, C. J., Cooper, D. H. and Graham, J. 1995. Active shape models – their training and application. In *Computer Vision and Image Understanding*. 61, 1 (1995), 38–59.

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

[21] Haykin, S. 2008. *Neural Networks and Learning Machines: A Comprehensive Foundation*.