Optimizing Gabor Filter and Local Binary Patterns for multi-texture classification

Alsadegh Mohamed[#1], Joan Lu[*2], Qiang Xu[#3], Jinwen Ma[*4]

[#] *School of Computing and Engineering, University of Huddersfield, Huddersfield, UK*
[1] alsadegh.mohamed@hud.ac.uk
[3] Q.Xu2@hud.ac.uk

[*] *School of Computing and Engineering-School of Mathematical Sciences, University of Huddersfield-University of Peking, Huddersfield UK- Beijing China*
[2] j.lu@hud.ac.uk
[4] jwma@math.pku.edu.cn

Abstract

In image classification by texture, it is important to maximize the discrimination between different classes by using an effective descriptor. The objective of this research is a new hybrid approach using state of the art feature extraction methods and improving the classification percentage of optimum filter by combining it with optimized LBP and find low dimensional size of features. The Gabor filter (GF) parameters are processed by the Artificial Bee Colony (ABC) algorithm to select the optimum filter, whereas pertinent features from LBP histogram are obtained using Rough Set Theory (RST) without impacting its classification rate. The classification implemented on texture classes is obtained from the Brodatz database. The results from the proposed approach show an improvement in the classification accuracy and processing time of k-folded cross- validation Neural Network classifier over the method of LBP with single filter and a reduced processing time of the classifier.

Keywords : *Feature selection, ABC algorithm optimization, Rough Set Method, GF feature extraction, LBP, Brodatz*

1. Introduction

Most natural sights show a different arrangement of repeated patterns which is called the texture [1]. Feature extraction methods of texture classification aim to obtain informative features from the images in order to classify them into number of classes [2]. Image classification by texture is exploited in different areas topics such as defect detection in the images, Remote Sensing Images, medical images, aerial and satellite images [2].

The information in the feature vector can be improved by integrating features from different methods [3]. In image processing and texture analysis, the GF and LBP are important complementary methods. LBP extracts the details from fine textures, whereas GF pays attention to features on a coarser scale [4]. The GF has localization properties in the spatial and frequency domain with selectivity of orientation and frequency [5]. The main problem of GF is, when constructing sub filters using dissimilar values of orientation and frequency, these filters increase the computation costs of GF [6]. There have been many attempts to reduce its complexities by optimization methods such as [4] [7]. On the other hand, LBP is a simple and fast method of extracting the features from the object [8]. It is effective to deal with local features of the texture by represent these features in small area of pixels. The long size of the LBP histogram may effect the classifier negatively, especially if integrated with other methods [9].

Classification of textures can be improved by integrating different feature extraction methods. However, fusing different features may cause a high dimensionality problem [10]. Feature selection methods are used to remove redundant information from features' size, which has been a vital field of study since 1970 [11]. These methods are divided into filter and wrapper techniques. The first is independent on classifier, with no additional information required. RST is one of the filter methods which was introduced by Pawlak in

1982, based on the concept of set theory [12]. Wrapped methods require a classification algorithm to evaluate the selected features. A genetic algorithm, simulating annealing and swarm colonies like (ABC) are wrapper methods that need a classifier for evaluation [13].

In this paper, RS will be used as a reduction method of the LBP histogram before it is integrated with the nominated optimal filter which is extracted by the ABC algorithm for multi texture classification. Two sides will be considered in the results which are: the size of features which effect the time period of data in the classifier, and the classification percentage. The state of the art feature extraction methods using textures in the proposed hybrid method can be used with many applications such as LBP for analysis of Mammographic of medical images [14] and Remote Sensing Imagery [15], whereas GF with Gastroenterology images [16], defect detection in textured materials [17], and documents processing [18]. The main purpose of optimization is to improve their efficiency in different applications.

Now that the paper's topic has been introduced, the rest of the paper is structured as follows: Section 2 introduces the related work; Section 3 illustrates the theories for the classification method of multi classes; Section 4 explains the applied classifier; Section 5 proposes the used methodology; Section 6 discusses the results; finally, Section 7 is the conclusion, which includes the achievement of this study and proposes future work.

2. Related Work

Dealing with numerous textures for classification or segmentation by optimal GF was the goal of many researchers. Tsai (2001) used a stochastic search technique for two, three and four part textured images. The single optimized filter by simulated annealing (SA) was used to segment the image containing different textures. The procedure depends on maximizing the minimum energy ratio of any two different textures in the image, where the separation between dissimilar textures will be based on the responses of this energy. A defined threshold has been used to separate different regions of textures in the image [7]. The optimal GF is sometimes unable to process multiple textures in the image. In order to cope with this problem, Ma Li (2008) exploited the uncomplicated LBP and combined it with an optimal filter in feature vector for multi-texture segmentation. An immune genetic algorithm was used to select the optimal parameters of GF before it was integrated with LBP in the features level [4].

The GF and LBP are complementary and effective feature extraction methods. In dimensional reduction, LBP and Local Gabor Binary Pattern (LGBP) have been integrated for face classification. In order to deal with high dimensional problem, the Fisher Discriminant Analysis (FDA) is used to reduce their spatial histogram based on feature similarity [19]. Tan in [20] combined GF and LBP representation to give diversity of features and improve the face recognition. Where both methods have high dimensional features, the Principle Component Analysis (PCA) is used to escape from the dimensionality of features methods before integration. Kernel Discriminative Common Vectors (KDCV) are used for fusing the feature vector of previous methods based only on discriminant features for recognition. In [21] Hussain improved the features extracted from the texture, the Oriented Gradient, Local Binary Pattern and Local Ternary Pattern have been used as complementary methods of texture description. Partial Least Squares are used to reduce the dimensionality of the resulting feature vector to decrease the processing time of the classifier without impacting the accuracy.

3. Underpinning Theory

A. Review of Involved Feature Extraction Methods

1) Local Binary Patterns:

Ojala et al [3] in 1996 introduced a new feature extraction method for texture classification which is simple in computation and invariant of the gray scale. The features extracted by LBP reflect micro-texton such as spots, edge ends and flat areas and can be used as a complementary features with other tecniques for texture classification [3].

In its strategy, from a window size of 3 x 3, LBP takes the binary patterns 0 or 1, which consists of a center pixel that is used as the threshold value and compare with the eight neighboring pixels. The values of the threshold's neighbors are multiplied by weights, then the result from neighboring values is summed to find the code of the center pixel, and this repeated on all pixels of the image, as seen in Fig. 1 and (1) of used equation.

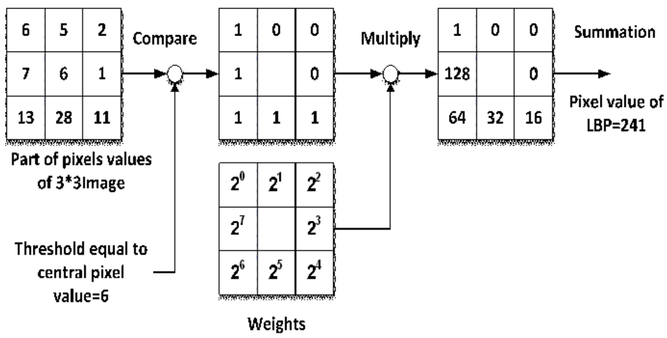

Fig. 1 Computing LPB code of window size 3 x 3 from image pixels

$$LBP_{n,R} = \sum_{n=0}^{n-1} v(p_n - p_c)2^n \qquad (1)$$

$$v(t) = \begin{cases} 1, & if\ t \geq 0 \\ 0, & if\ t < 0 \end{cases}$$

where n is the number of neighborhood pixels, is the value of the central pixel and P_n are the values of neighborhood pixels. For features vector, the histogram (h) collects the different occurrences of the LBP matrix (2) [22].

$$h = \sum_{x,y} B\left(LBP_{n,R}(x,y) = i\right),\ i = 0, \dots, 2^n - 1 \qquad (2)$$

$$B(t) = 1\ when\ (t)\ is\ true\ and\ 0\ otherwise$$

2) Gabor Filter

The Gabor Filter is one of the most effective spectrum methods for texture analysis. It considers the edges information of an image's content, where it extracts orientation information about the texture by using different scales and it is invariant of illumination [23].

The Gabor Filter is a sinusoid wave in the spatial domain which is modulated by a Gaussian envelope (3). The standard deviation of the Gaussian determines the bandwidth envelop of the filter, whereas the sinusoid signal direction and frequency are the direction and frequency of the band pass filter [24].

$$\psi(x, y; f_0, \theta\ ) = \frac{f_0^2}{\pi\gamma\eta}\ e^{-\frac{f_0^2}{\gamma^2}x'^2 + \frac{f_0^2}{\eta^2}y'^2}\ e^{j2\pi\ f_0 x'} \qquad (3)$$

$$x' = x cos\theta + y sin\theta$$

$$y' = -x sin\theta + y cos\theta$$

Where f_0 refers to the central frequency of the filter, and θ is the rotation angle,γ and η Gaussian envelope. The optimal filter depends on proper selection of the last parameters. Fig. 2 shows multi-channel Gabor filters of different frequencies or scales and orientations. The central frequency of the individual filters is calculated by (4), whereas the angles of the filters are calculated with (5).

$$f_n = c^{-n} f_m, n = 0,1 \dots \dots .., n_s - 1 \qquad (4)$$

Where f_m defines the maximum frequency range of the filter, c is the scaling of the frequency factor, n_s is the total scales number, and n is the number of the calculated frequency.

$$\theta_n = \mathrm{n}\frac{2\pi}{n_0}, n = 0,1 \dots \dots \dots, n_0 - 1 \qquad (5)$$

Where n0, is the total number of orientations, and n is the number of the calculated orientation.

The feature vector of GF after convolution with the image consists of mean and standard deviation moments (6) [25].

$$f = [\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots \dots . \mu_{M-1\,N-1}, \sigma_{M-1\,N-1}] \qquad (6)$$

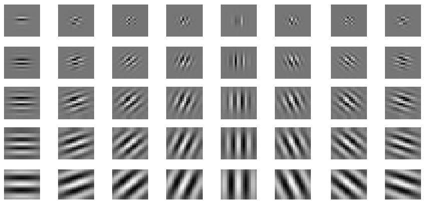

Fig. 2 Multi-channel Gabor filters consist of 40 filters with a number of different scales equal to 5 and a number of different orientations equal to 8.

B. Basic Feature Selection Methods

1) ABC Algorithm

In 2005 Karaboga suggested ABC as an optimization method, which was inspired from the working day of bees [26]. The participating bees included the following types; the employed bees, the onlooker bees and the scout bees. The employed and onlooker bees had an equal number within the colony, and the food sources number were equal to the half colony. The employed bees share obtained information about nectar amount in food sources with onlookers by a special movement in the air called waggling. The food source with the richest nectar will be visited by the largest number of onlookers after receiving the information from employed bees. The food source with the richest nectar is saved and the employed bees become scout bees. The difference between the employed and scout bees is that the former have information about the food sources, whereas the latter are looking for food sources randomly [27] [28].

2) RS Method

The RS method processes the uncertainty and vagueness of certain data mathematically without additional knowledge. The data in information systems is distributed as a table which is categorized into universe and attributes as $S = (U, A)$ In the rows is the universe (U)which consists of a set of objects, whereas in the columns are attributes (A). The decision table is generated by dividing the attributes in the information table into condition attributes and decision attributes to become $S = (U, C \cup D)$ [29].

Attribute reduction is one of the most important tasks of rough set theory [30]. It aims to remove redundant attributes while maintaining the same level of quality of classification as before the reduction. The principle of reduction in the rough set method is based on the granularity structure of the data as result of indiscernibility. Indiscernible attributes occur between objects if a number of them have the same information which is called an elementary set or equivalence.

If S subset is an attribute of A that means $S \subseteq A$ and the equivalence relation is as following:

$$IND(S) = \{(i,j) \epsilon\, U^2, \forall_a \epsilon S, a(i) = a(j)\} \qquad (7)$$

where a(i) is the attribute value of object i.

The attribute is divided into sets based on their relevance by lower and upper approximation, which in turn is deduced as a result of an equivalence or indiscernibility relation between the attributes.

The lower approximation $S_-(t)$ and upper approximation $S^-(t)$ for $T \subseteq \hat{U}$ are:

$$S_-(t) = \{t \epsilon U: S(t) \subseteq T\} \qquad (8)$$

$$S^-(t) = \{t \,\epsilon\, U: S(t) \cap T \neq \emptyset\} \qquad (9)$$

Here, the method defines three regions which are positive, boundary, and negative region. The first includes the objects which definitely belong into relevant set of features, whereas the boundary region holds the objects which are probably relevant features.

$$POS_p(Q) = \cup_{x\epsilon \mathrm{UIQ}} S_-(t) \qquad (10)$$

$$NEG_p(Q) = U - \cup_{x\epsilon \mathrm{UIQ}} S^-(t) \qquad (11)$$

$$SN_s\,(t) = S_-(t) - S_-(t) \qquad (12)$$

where, $P, Q \subseteq A$ be equivalence relations over U.

Based on the boundary region, the crisp state is if this region is empty, whereas the opposite case is a rough state if the region is not empty [29].

4. K-fold Neural Network Classifier

The experiments of multi classes image classification by texture are based on a K-fold cross validation Artificial Neural Network (ANN) Classifier. The ANN algorithm is simple and can be adapted to any problems [31]. ANN is a multi-layer preceptor (MLP) and its weights are updated by back-propagation in the training stage where the involved data is divided between sets for training and sets for testing. Back propagation uses the sum-of-squares error functions to adjust the weights in the layers [32].

The images in the training and testing stages are distributed randomly and K-fold has been used as a cross validation technique [33]. K-fold is a simple cross validation method compared to other techniques [34]. It can be applied with either 5 fold or 10 fold [34] [35]. Due to time, this method used the 5 fold, meaning the features are distributed randomly into five equal parts. One of five parts is used for testing and the remainders for training. The accuracy is calculated five times and their average is the final accuracy.

5. The Reserch Problem and Proposed Methodology

A. Research Problem

Based on previous researches, the main problem to deal with multiple texture classes for classification is to find diverse and effective features that improve the classification accuracy with efficient time in the computation. Where one method cannot give an effective result when used individually, some researchers proposed the solution of integrating the complementary features, however, integrating many features from different methods causes the curse of dimensionality. In this paper, searching for the optimum Gabor filter using an effective algorithm and combined with the output of relevant features from an LBP histogram is the proposed method to deal with this problem.

B. The methodology of Hybrid Method

In this method, the Gabor Bank contributes with the optimal filter using the ABC algorithm as a wrapper method, whereas the RS method is used as a filter method to reduce the histogram size of LBP.

This procedure is to reduce dimensionality and include only the relevant features from feature extraction methods. The main purpose is to improve the accuracy and processing time of the classifier for multi-class classification.

The database from Brodatz Textures as explained in section (VI-A) has been divided into groups and every group consists of a number of image classes. The proposed approach consists of the following stages as shown in Fig. 3:

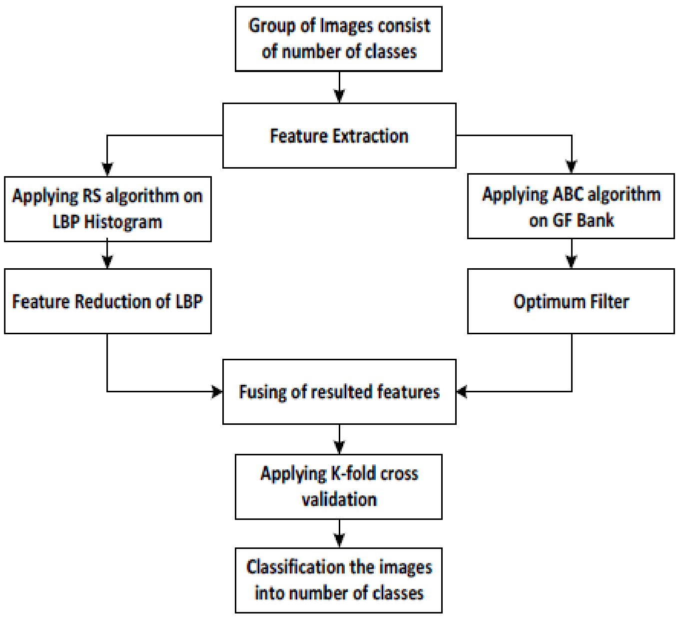

Fig. 3 Flowchart of hybrid multi classification method

In first stage, the exploration and exploitation characteristics of ABC will be implemented for selecting the optimal or suboptimal parameter values of GF. In ABC algorithm, the food sources are GF parameters which are considered as a possible solution for an optimization problem. The initial solution is generated and the exploitation of the solution takes place in the employed and onlooker bees stages. Both of them try to modify the parameter values of GF for better accuracy based on classifiers. This is executed by evaluating the new positions and applying greedy selection to select better solutions. Scout bees explore globally and randomly for expected better parameter values when desperate to improve the solutions locally.

In the second stage, the LBP histogram will be reduced by generating decision rules of RST. The indiscernibility relations apply to the features which here are the LBP histograms to obtain the important features from the texture.

In the final stage, the optimal feature vector of GF will be integrated with the relevant features of LBP. The performance of the proposed classification algorithm is evaluated based on the classification accuracy and time.

6. Expermental Results and Discussion

The new approach is applied to images containing different textures from the Brodatz database. The main objective is to investigate the impact of the concatenate features of optimal GF with relevant features from the LBP histogram. The results of image classification which include the accuracy and spending classifier time have been recorded using Matlab 2016a which was installed on a Windows 7 Core Processor.

A. The database of Brodatz

To evaluate the new classification method with multi classes of textures, the experiment was carried out on textures from the Brodatz database [36]. This database is one of most popular for image classification by texture. There are more than a hundred images in the book which contain a wide variety of natural surfaces of texture, and Fig. 4 exhibits a collection of them. The selected texture images have been divided into eight groups with different classes numbers: 5, 10, 15, 20, 25 and 30 classes.

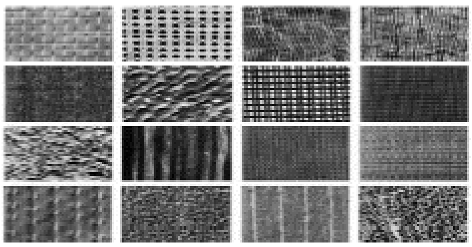

Fig. 4 Part of Brodatz Textures samples

B. Results

Table. 1 explains the number of classes of every image group along with the results of classification accuracy and processing time of data in the classifier. In classification accuracy, fusing optimal GF with a reduced LBP histogram using RS produced more accurate results than combining the optimal filter directly with the LBP histogram. The accuracy exceeded 3% in the groups having 10 and 15 classes whereas with 5 classes the accuracy was 9 %.

In addition, the histogram of LBP was reduced before it was connected with the optimal GF. This method reduced the training and testing data size therefore reducing the time of the classifier algorithm. This is more effective than using the complete LBP histogram.

C. Discussion of Results

The optimum Gabor filter was applied to multiple texture segmentation, where the image contained four parts of different textures [4] [7]. In another study, to reduce the computation cost of using all filters in Gabor Filter Bank, the optimum filter was extracted using a new optimization method (ABC) and executed on 16 groups of images from the University of Maryland, College Park Database (UMD) for binary classification by texture [37]. In multi-texture classification, the accuracy of using a single filter decreases as the number of texture classes increase as shown in Fig. 5.

TABLE I

THE ACCURACY AND TIME OF TEXTURE CLASSIFICATION

Images groups	Classification Accuracy			Classifier Time	
	Optimal GF	Optimal GF with LBP	Optimal GF with LBP & RS	Optimal GF with LBP	Optimal GF with LBP & RS
05 Classes	93.190	86.192	95.451	78.169	18.956
10 Classes	68.261	83.891	86.457	76.782	49.623
15 Classes	59.406	80.623	83.232	112.640	97.032
20 Classes	50.402	82.467	83.457	177.707	139.333
25 Classes	56.761	81.657	81.309	227.773	195.435
30 Classes	53.833	81.869	82.841	340.154	270.665
Average	73.791	88.222	89.748	122.561	86.971

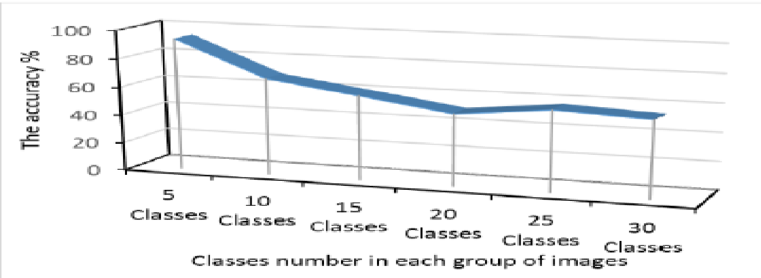

Fig. 5 The accuracy of images classification by optimum Filter on number of image classes.

To overcome the disadvantages of using one filter, the optimal filter has been complemented by LBP [4]. Although LBP has low computational complexity, the results histogram from the LBP matrix which is used for feature analysis is high-dimensional. This may cause the curse of dimensionality if combined with other methods [38]. Here we proposed another solution to improve the accuracy of using the optimum filter. The LBP histogram has been reduced and only the relevant features were selected by the rough set method before it was integrated with the optimal filter. The main purpose is to improve the overall accuracy of a hybrid classification method. However, concatenating the feature vector of the filter with the LBP histogram increases the dimensionality of features of the classifier. The high size of features has a negative impact on the speed of the classifier [9]. RS methods can select relevant features and reduce the dimensionality of the LBP histogram before integrating with the filter.

Fig. 6 shows the impact of applying RS on LBP features for removing the redundant features for texture classification without effecting its accuracy. From the figure, the RS can be used with LBP without effecting its ability in classification, whereas the time taken by the classifier with training and testingdata has clearly been reduced on average from 172.58 to 124.02 for LBP with RS reduction method.

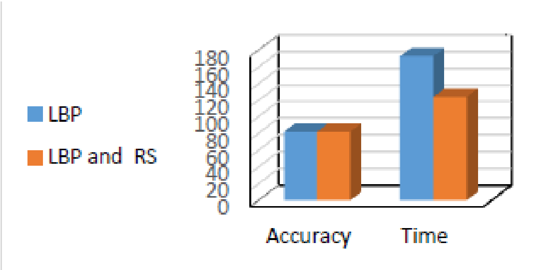

Fig. 6 Comparison the accuracy and time of images classification between original LBP descriptor and LBP descriptor after feature reduction using RS.

Fig. 7 compares the accuracy and time of the classifier between fusing the entire feature size of LBP and optimal GF with integrated relevant features of optimal GF and reduced LBP vector by RS method. The classification rate was from 88.222% to 89.748%, whereas the time was reduced from 122.561 to 86.971 in the new approach method for image classification. These results show the effectiveness of RS on LBP features before being integrated with the optimal filter.

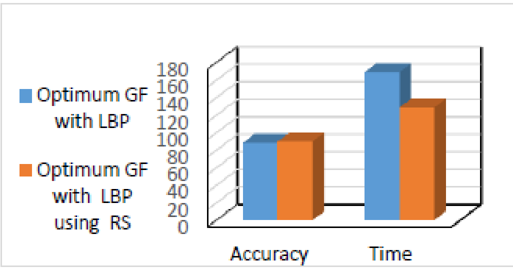

Fig. 7 Comparison classification accuracy and time between Optimum Filter with LBP before and after using RS with LBP histogram for features reduction.

7. Conclusion

This paper developed a hybrid method for multi-classes of texture classification. It extracts complementary improved texture features from GF and LBP. Processing the LBP histogram was executed using RS method and combined with the optimal filter which was extracted using ABC optimization algorithm. It provides a new method for optimized feature extraction. The testing of this method was based on the Brodatz databases.

A. Achievements

The results were effective in extracting the important features and avoiding the curse of dimensionality with improvements in accuracy and classifier time. On average the accuracy improved from using the optimum GF with 73.791 to 89.748 after it was integrated with the optimized features of LBP. In addition, the LBP is simple method, but it includes irrelevant features which were removed successfully by RST.

B. Future Work

While the accuracy and overall time of image classification with a number of classes less than 10 has been improved, our plan is to further improve the classification rate in the groups which contain a high number of classes (more than 10). In addition, this method will be applied to more available datasets in the literature and using different classifiers. Furthermore, the hybrid algorithm, based on widely used descriptors of texture, and these descriptors proved their effectiveness with many applications such as medical image analysis. Therefore, the recommendation is for further research to study the improvement using the optimized algorithm on applications such as medical, remote sensing, documents processing and other imagery.

References

[1] N. Baaziz, O. Abahmane, and R. Missaoui, "Texture feature extraction in the spatial-frequency domain for content-based image retrieval," pp.1-19, 2010.

[2] M. Tuceryan, and A. K. Jain, "Texture analysis". Handbook of pattern recognition and computer vision, 2, pp.207-248, 1993.

[3] T.,Ojala, M. Pietikäinen, and D. Harwood., "A comparative study of texture measures with classification based on featured distributions," Pattern recognition, 29(1): pp. 51-59, 1996.

[4] M. Li, and R. C. Staunton, "Optimum Gabor filter design and local binary patterns for texture segmentation," Pattern Recognition Letters, 29(5), pp. 664-672, 2008.

[5] J. Yang, L. Liu, T. Jiang, and Y. Fan, "A modified Gabor filter design method for fingerprint image enhancement," Pattern Recognition Letters, 24(12), pp. 1805-1817, 2003.

[6] J. Ilonen, J. K. Kamarainen, and H. Kälviäinen, "Fast extraction of multi- resolution gabor features," In Image Analysis and Processing, 14th International Conference on pp. 481-486, September 2007.

[7] D. M. Tsai, S. K. Wu, and M. C. Chen, "Optimal Gabor filter design for texture segmentation using stochastic optimization," Image and Vision Computing, 19(5), pp. 299-316, 2001.

[8] T.Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on pattern analysis and machine intelligence, 24(7), pp. 971-987, 2002.

[9] X. Wang, Z. Hua, and R. Bai, "A hybrid text classification model based on rough sets and genetic algorithms," In Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, SNPD'08. pp. 971-977, August 2008.

[10] A. Jain, and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19(2), pp. 153-158, 1997

[11] H. Liu, and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," IEEE Transactions on knowledge and data engineering, 17(4), pp. 491-502, 2005.

[12] R. Li, and Z. O. Wang, "Mining classification rules using rough sets and neural networks," European Journal of Operational Research, 157(2), pp. 439-448, 2004.

[13] R. Kohavi, and G. H. John, "Wrappers for feature subset selection. Artificial intelligence, 97(1-2), pp. 273-324, 1997.

[14] L. Nanni, A. Lumini, and S. Brahnam, "Local binary patterns variants as texture descriptors for medical image analysis, " Artificial intelligence in medicine, 49(2), pp. 117-125, 2010.

[15] M. Musci, R. Q. Feitosa, G. A. Costa, and M. L. F. Velloso, "Assessment of binary coding techniques for texture characterization in remote sensing imagery," IEEE Geoscience and Remote Sensing Letters, 10(6), pp. 1607-1611, 2013.

[16] F. Riaz, F. B. Silva, M. D. Ribeiro, and M. T. Coimbra, "Invariant gabor texture descriptors for classification of gastroenterology images," IEEE Transactions on Biomedical Engineering, 59(10), pp. 2893-2904, 2012.

[17] A.Kumar, and G. K. Pang, "Defect detection in textured materials using Gabor filters," IEEE Transactions on industry applications, 38(2), pp. 425-440, 2002.

[18] H. Hassen, "A Comparative study of Arabic handwritten characters invariant feature," International Journal of Advanced Computer Science and Applications (IJACSA), pp. 62-68, 2012.

[19] S.Shan, W. Zhang, Y.Su, X. Chen, and W. Gao, "Ensemble of piecewise FDA based on spatial histograms of local (Gabor) binary patterns for face recognition," In: Proc. International Conference on Pattern Recognition, vol. 4, pp. 606–609, 2006.

[20] X. Tan, and B. Triggs, "Fusing Gabor and LBP feature sets for kernel-based face recognition," In: Analysis and Modeling of Faces and Gestures. Lecture Notes in Computer Science, vol. 4778, pp. 235–249, 2007.

[21] S. Hussain, and B. Triggs, "Feature sets and dimensionality reduction for visual object detection," In: Proc. British Machine Vision Conference, pp. 112.1–112.10, 2010.

[22] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on pattern analysis and machine intelligence, 24(7), pp. 971-987, 2002.

[23] D. Zhang, , M. Islam, M. Lu, G., and I. J. Sumana, "Rotation invariant curvelet features for region based image retrieval," International journal of computer vision, 98(2), pp. 187-201, 2012.

[24] A. C. Bovik, M.Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," IEEE transactions on pattern analysis and machine intelligence, 12(1), pp. 55-73, 1990.

[25] B. S. Manjunath, and W. Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Transactions on pattern analysis and machine intelligence, 18(8), pp. 837-842, 1996.

[26] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, 2005.

[27] D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, "A comprehensive survey: artificial bee colony (ABC) algorithm and applications,". Artificial Intelligence Review, 42(1), pp. 21-57, 2014.

[28] Le Dinh, L., D. Vo Ngoc, and P. Vasant., "Artificial bee colony algorithm for solving optimal power flow problem," The Scientific World Journal, pp. 1-9, 2013.

[29] M. Zhang, and J. T. Yao, "A rough sets based approch to feature selection," In Fuzzy Information, 2004. Processing NAFIPS'04. IEEE Annual Meeting of the Vol. 1, pp. 434-439, June 2004.

[30] T. S. Li, "Applying wavelets transform, rough set theory and support vector machine for copper clad laminate defects classification," Expert systems with Applications, 36(3), pp. 5822-5829, 2009.

[31] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," Chemometrics and intelligent laboratory systems, 39(1), pp. 43-62, 1997.

[32] J. Zupan, "Introduction to artificial neural network (ANN) methods: what they are and how to use them," Acta Chimica Slovenica, 41, pp. 327-327, 1994.

[33] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," In Ijcai Vol. 14, No. 2, pp. 1137-1145, August 1995