Ensemble Feature Selection Approach Based on Feature Ranking for Rice Seed Images Classification

Dzi Lam Tran TUAN [1], Thongchai SURINWARANGKOON [2], Kittikhun MEETHONGJAN [3], Vinh Truong HOANG [1]

[1]Department of Image Processing and Computer Graphics, Faculty of Computer Science,
Ho Chi Minh City Open University, 97 Vo Van Tan Street, 700000 Ho Chi Minh City, Vietnam
[2]Department of Business Computer, Faculty of Management Science,
Suan Sunandha Rajabhat University, 1 U Thong Nok Rd, 10300 Bangkok, Thailand
[3]Department of Applied Science, Faculty of Science and Technology,
Suan Sunandha Rajabhat University, 1 U Thong Nok Rd, 10300 Bangkok, Thailand

dziltt.178i@ou.edu.vn, thongchai.su@ssru.ac.th, kittikhun.me@ssru.ac.th, vinh.th@ou.edu.vn

Abstract. *In smart agriculture, rice variety inspection systems based on computer vision need to be used for recognizing rice seeds instead of using technical experts. In this paper, we have investigated three types of local descriptors, such as Local Binary Pattern (LBP), Histogram of Oriented Gradients (HOG) and GIST to characterize rice seed images. However, this approach raises the curse of dimensionality phenomenon and needs to select the relevant features for a compact and better representation model. A new ensemble feature selection is proposed to represent all useful information collected from different single feature selection methods. The experimental results have shown the efficiency of our proposed method in terms of accuracy.*

Keywords

Ensemble feature selection, feature ranking, feature selection, GIST, HOG, LBP, rice seed image.

1. Introduction

Rice is the most important food source of people in many countries including Asia, Africa, Latin America, and the Middle East. Products made from rice, including rice products and indirect products, are indispensable in the daily meals of billions of people around the world. Nowadays, more rice varieties are created with diversified quality and productivity. Different varieties of rice can be mixed during cultivation and trading. We practically need to develop a system to automatically identify rice seeds based on machine vision. Various works have been proposed for automatic inspection and quality control in agriculture [10]. In the past decade, a great number of local image descriptors [13] have been proposed for characterizing images. Each kind of attribute represents the data in a specific space and has precise spatial meaning and statistical properties.

Different local descriptors are extracted to create a multi-view image representation, like Local Binary Pattern (LBP), Histogram of Oriented Gradients (HOG), and GIST. Nhat and Hoang [20] present a method to fuse the features extracted from three descriptors (Local Binary Pattern, Histogram of Oriented Gradient and GIST) for facial images classification. The concatenated features are then applied by canonical correlation analysis to have a compact representation before feeding into the classifier. Van and Hoang [24] propose to reduce noisy and irrelevant Local Ternary Pattern (LTP) features and HOG coding on different color spaces for face analysis. Hoai et al. [12] introduce a comparative study of hand-crafted descriptors and Convolutional Neural Networks (CNN) for rice seed images classification. Mebatsion et al. [17] fuse Fourier descriptors and three geometrical features for cereal grains recognition. Duong and Hoang [9] apply to extract rice seed images based on features coded in multiple color spaces using HOG descriptor. Multi-view

learning was introduced to complement information between different views. While concatenating different feature sets, it is evident that all the features do not give the same contribution for the learning task and some features might decrease the performance. Thus, feature selection methods are applied as a pre-processing stage to high-dimensional feature space. It involves selecting pertinent and useful features, while avoiding and ignoring redundant and irrelevant information [26]. A novel teacher-student feature selection approach [19] is proposed to find the best representation of data in low dimension.

Recently, ensemble feature selection has emerged as a new approach that promises to enhance the robustness and performance. It is the process of performing different feature selection in order to find an optimum subset of features. Instead of using a single selection approach, an ensemble method combines the results of different approaches into a final single subset of features. Seijo-pardo et al. [23] propose to combine different feature selection approaches on heterogeneous data based on a predefined threshold value. Chiew et al. [5] introduce a hybrid ensemble feature selection based on Cumulative Distribution Function gradient. This method can determine an estimation of feature cut-off automatically. Drotar et al. [8] propose a new ensemble feature selection approach methods based on different voting techniques such as plurality, and Borda count. A complete and detailed review of ensemble feature selection methods is introduced in [3].

In this paper, we propose a new ensemble feature selection approach based on multi-view descriptors (LBP, HOG and GIST) extracted from rice seed images. Several feature selection approaches are further investigated and combined to find an optimum subset of features with the purpose to enhance the classification performance. This paper is organized and structured as follows. Section 2. , introduces the feature extracting methods based on three local image descriptors. Section 3. presents a proposed ensemble feature selection framework. Section 4. shows experimental results. Finally, the conclusion is then provided in Sec. 5. .

2. The Feature Extracting Methods

This section briefly reviews three local image descriptors used in experiments for feature extraction.

2.1. Local Binary Pattern

The $LBP_{P,R}(x_c, y_c)$ code of each pixel (x_c, y_c) is calculated by comparing the gray value g_c of the central pixel with the gray values $\{g_i\}_{i=0}^{P-1}$ of its P neighbors, as follows [21]:

$$\text{LBP}_{P,R} = \sum_{p=0}^{P-1} \omega(g_p - g_c)2^p, \qquad (1)$$

where g_c is the gray value of central, g_p is the gray value of P, R is the radius of the circle, and $\omega(g_p - g_c)$ is defined as:

$$\omega(g_p - g_c) = \begin{cases} 1 & \text{if} \quad (g_p - g_c) \geq 0, \\ 0 & \text{otherwise.} \end{cases} \qquad (2)$$

2.2. GIST

GIST is firstly proposed by Oliva and Torralba [22] in order to classify objects which represent the shape of the object. The primary idea of this approach is based on the Gabor filter:

$$h(x,y) = e^{-\frac{1}{2}\left(\frac{x^2}{\delta_x^2} + \frac{y^2}{\delta_y^2}\right)} e^{-j2\pi(u_0 x + v_0 y)}. \qquad (3)$$

For each (δ_x, δ_y) of the image via the Gabor filter, we obtain all the image elements that are close to the point color $(u_0 x + v_0 y)$. The result of the calculated Vector GIST will have many dimensions. To reduce the size of the vector, we averaged each 4×4 grid of the above results. Each image also configures a Gabor filter with 4 scales and 8 directions (orientations), creating 32 characteristic maps of the same size.

2.3. Histograms of Oriented Gradient

HOG descriptors are applied for different tasks in machine vision [7] such as human detection [6]. HOG feature is extracted by counting the occurrences of gradient orientation based on the gradient angle and the gradient magnitude of local patches of an image. The gradient angle and magnitude at each pixel is computed in an 8×8 pixels patch. Next, 64 gradient feature vectors are divided into 9 angular bins 0–180° (20° each). The gradient magnitude T and angle K at each position (k, h) from an image J are computed as follows:

$$\Delta_k = |J(k-1, h) - J(k+1, h)|. \qquad (4)$$

$$\Delta_h = |J(k, h-1) - J(k, h+1)|. \qquad (5)$$

$$T(k, h) = \sqrt{\Delta_i^2 + \Delta_j^2}. \qquad (6)$$

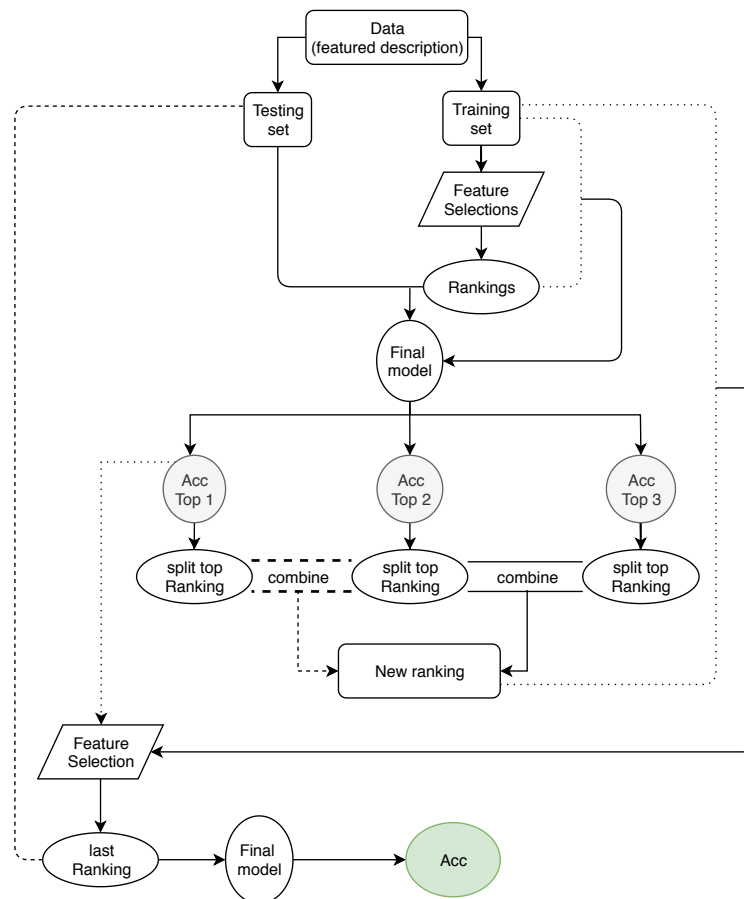$$K(k, h) = tan^{-1}\left(\frac{\Delta_k}{\Delta_j}\right). \qquad (7)$$

Fig. 1: The proposed ensemble feature selection approach.

3. Ensemble Feature Selection

The dimension reduction has several advantages and impacts on data storage, generalization capability and computing time. Based on the availability of supervised information (i.e, class labels), feature selection techniques can be grouped into two large categories: supervised and unsupervised context [1]. Additionally, different strategies of feature selection are proposed based on evaluation processes such as filter, wrapper and hybrid methods [11]. Hybrid approaches incorporate both filter and wrapper into a single structure, in order to give an effective solution for dimensionality reduction [4]. In order to study the contribution of feature selection approaches for rice seed images classification, we propose to apply several selection approaches based on images represented by multi-view descriptors. In the following subsection, we will shortly present the common feature selection methods applied in supervised learning context.

- LASSO (Least Absolute Shrinkage and Selection Operator) allows to compute feature selection based on the assumption of linear dependency between input features and output values. Lasso

minimizes the sum of squares of residuals when the sum of the absolute values of the regression coefficients is less than a constant, which yields certain strict regression coefficients equal to 0 [4] and [25].

- mRMR (Maximum Relevance and Minimum Redundancy) is a mutual information based feature selection criterion, or distance/similarity scores to select features. The aim is to penalize a feature's relevance by its redundancy in the presence of the other selected features [27].

- ReliefF [15] is extended from Relief [14] to support multiclass problems. ReliefF seems to be a promising heuristic function that may overcome the myopia of current inductive learning algorithms. Kira and Rendell used ReliefF as a preprocessor to eliminate irrelevant attributes from data description before learning. ReliefF is general, relatively efficient, and reliable enough to guide the search in the learning process [16].

- CFS (Correlation Feature Selection) mainly applies heuristic methods to evaluate the effect of a single feature corresponding to each group in order to obtain the optimal subset of attributes.

- Fisher [2] identifies a subset of features so that the distances between samples in different classes are as large as possible, while the distances between samples in the same class are as small as possible. Fisher selects the top ranked features according to its scores.

- ILFS (Infinite Latent Feature Selection) is a technique consists of three steps such as preprocessing, feature weighting based on a fully connected graph in each node that connect all features. Finally, energy scores of the path length are calculated, then rank its correspondence with the feature [18].

Figure 1 present the proposed ensemble feature selection framework. Each individual feature selection approach has its pros and cons, the aim of this proposition is to combine the pros of different methods to boost the performance in terms of accuracy. We propose to apply three independent feature selection methods to select the "best" subset of features. Then, a new ranking method is applied for combined feature space. This can increase the dimension space but it allows to collect relevant features determined by different selection methods. The meaning behind is to select the most relevant features so that we have to apply a final ranking to eliminate the redundant and noisy features.
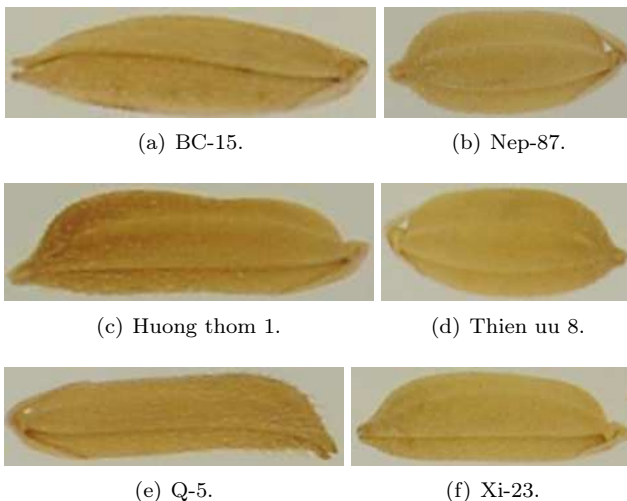
4. Experimental Results

4.1. Experimental Setup

(a) BC-15.

(b) Nep-87.

(c) Huong thom 1.

(d) Thien uu 8.

(e) Q-5.

(f) Xi-23.

Fig. 2: The proposed of ensemble feature selection.

The rice seed images database comprises six rice seed varieties in the northern Vietnam (illustrated in Fig. 2) [9]. We apply the 1-NN and SVM classifiers

to evaluate the classification performance via accuracy rate. A half of the database is selected for the training set and the rest is used for the testing set. We use Hold-out method with ratio (1/2 and 1/2) and split the training and testing set by chessboard decomposition. All experiments are implemented and simulated by Matlab 2019a and conducted on a PC with a configuration of a CPU Xeon 3.08 GHz, 64 GBs of RAM.

4.2. Results

Table 1 shows the accuracy obtained by 1-NN and SVM classifier when no feature selection approach is applied. The first column indicates the features used for representing images. We use three individual local descriptors namely LBP, GIST, and HOG and the concatenation of "LBP + GIST" features. The second column indicates the number of features (or dimension) corresponding to features type. The third and fourth columns show the accuracy obtained by 1-NN and SVM classifier. We observe that the multi-view by concatenating multiple features gives better performance, however it increases the dimension. Hence, the performance of SVM classifier is better than 1-NN classifier with 94.7 % of accuracy.

**Tab. 1:** Classification performance without selection approach for different types of features.

Features	Dimension	1-NN	SVM
LBP	768	53.0	**77.0**
GIST	512	69.4	**88.3**
HOG	21.384	71.5	**94.7**
LBP + GIST	1.280	70.5	**91.7**

The following tables and figures illustrate in detailed the classification in single or multi-view based on three descriptors:

- LBP: Table 2, Fig. 3(a) and Fig. 3(b),

- GIST: Table 4, Fig. 4(a) and Fig. 4(b),

- HOG: Table 5, Fig. 5(a) and Fig. 5(b),

- LBP + GIST: Table 3, Fig. 6(a) and Fig. 6(b).

Table 2 and Fig. 3 show that the classification performance reach 53.0 % by 1-NN classifier on LBP descriptor. After using 6 different feature selection approaches, we obtain three best candidates with descendant accuracy such as mRMR (59.0 %), ILFS (58.4 %) and ReliefF (54.2 %). Based on the proposed method illustrated in Fig. 1, the 85 % percentage of selected features by ReliefF is combined with 43 % of selected feature determined by ILFS method. We obtain the new subset of features which is calculated as follows:

$$(768 \cdot 0.85) + (768 \cdot 0.43) = 983 \text{ dim.} \qquad (8)$$
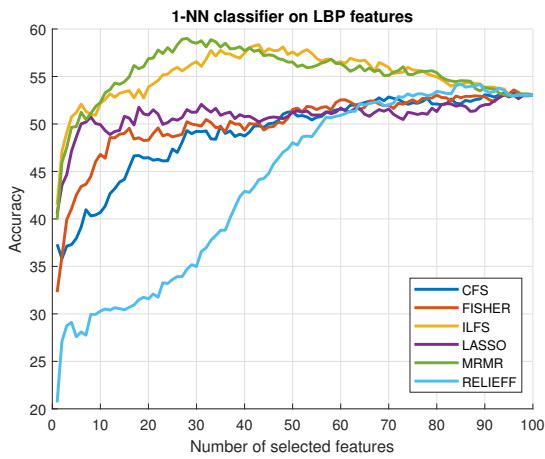
(a) 1-NN.
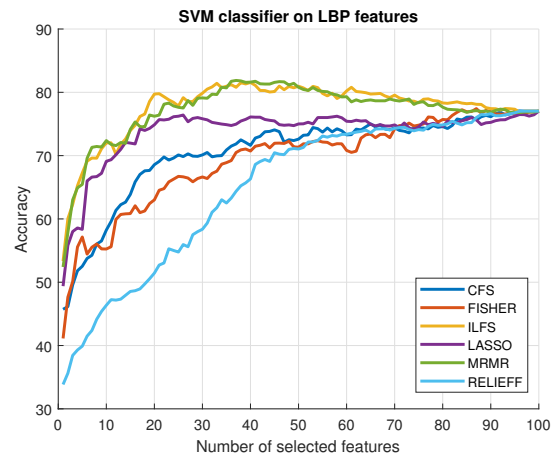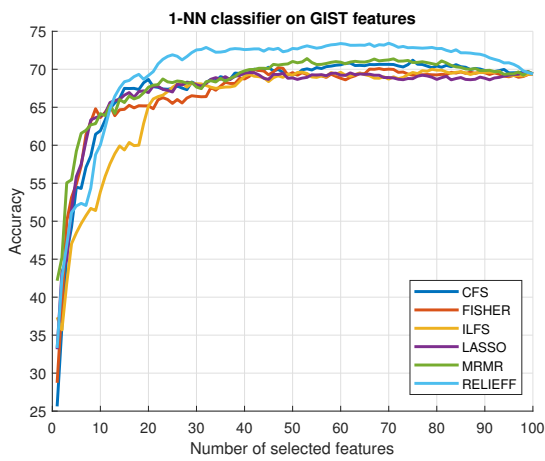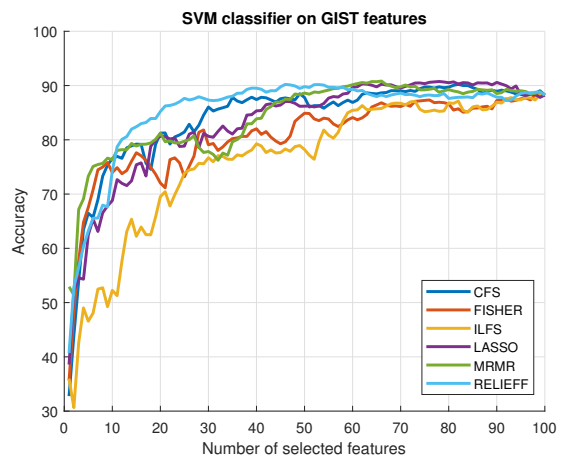
(b) SVM.

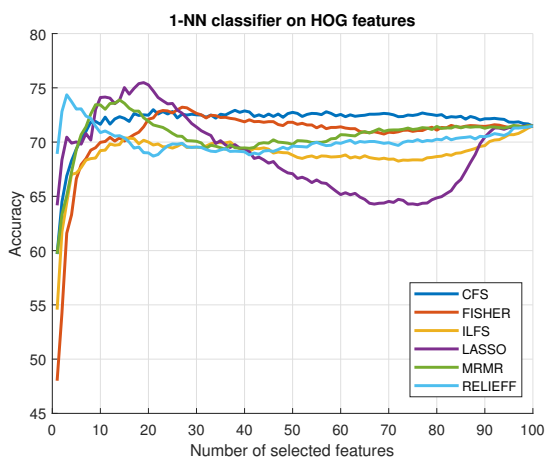**Fig. 3:** 1-NN (a) and SVM (b) classifier on LBP features.
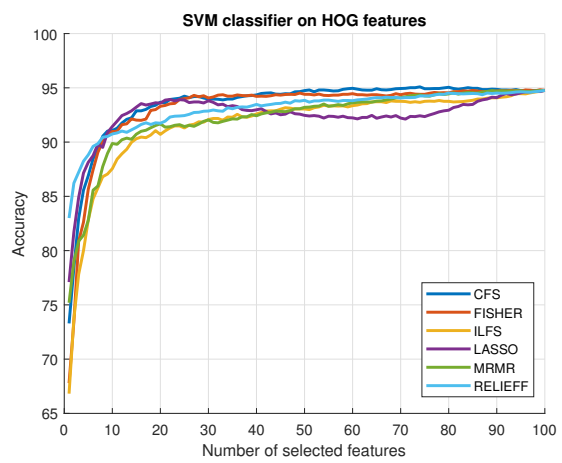

(a) 1-NN.

(b) SVM.

Fig. 4: 1-NN (a) and SVM (b) classifier on GIST features.

(a) 1-NN.

(b) SVM.

Fig. 5: 1-NN (a) and SVM (b) classifier on HOG features.
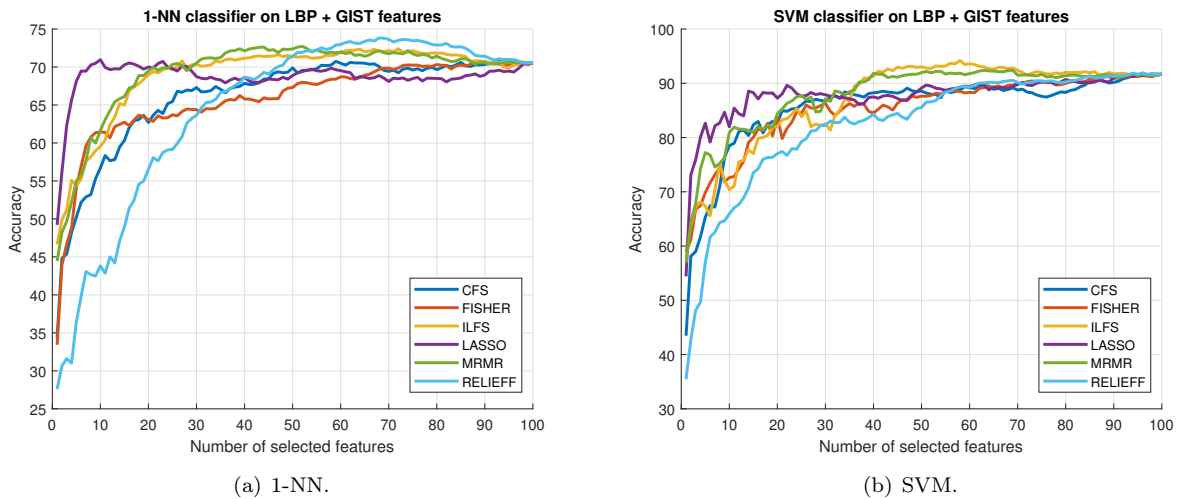
(a) 1-NN.

(b) SVM.

Fig. 6: 1-NN (a) and SVM (b) classifier on LBP + GIST features.

Tab. 2: LBP features - classification performance based on different feature selection methods with 1-NN and SVM classifier. **ACC**: accuracy, **Dim**: dimension, **id %**: percentage of selected features, ⩾ **id %**: percentage of selected features with accuracy equal to all features are used.

LBP	Dim	1-NN						SVM					
		ACC			Max ACC			ACC			Max ACC		
		100 %	⩾ id %	Dim	Max	id %	Dim	100 %	⩾ id %	Dim	Max	id %	Dim
Fisher	768	53.0	80	614	**53.6**	96	**737**	77.0	84	645	**77.4**	87	**668**
mRMR	768	53.0	11	84	**59.0**	28	**215**	77.0	22	169	**81.8**	37	**284**
ReliefF	768	53.0	74	568	**54.2**	85	**653**	77.0	97	745	**77.0**	97	**745**
Ilfs	768	53.0	12	92	**58.4**	43	**330**	77.0	19	146	**81.6**	40	**307**
Cfs	768	53.0	90	691	**52.3**	96	**737**	77.0	96	737	**77.1**	96	**737**
Lasso	768	53.0	94	722	**53.1**	94	**722**	77.0	100	768	77.0	100	768

Tab. 3: LBP + GIST features - classification performance based on different feature selection methods with 1-NN and SVM classifier. **ACC**: accuracy, **Dim**: dimension, **id %**: percentage of selected features, ⩾ **id %**: percentage of selected features with accuracy equal to all features are used.

LBP + GIST	Dim	1-NN						SVM					
		ACC			Max ACC			ACC			Max ACC		
		100 %	⩾ id %	Dim	Max	id %	Dim	100 %	⩾ id %	Dim	Max	id %	Dim
Fisher	1280	70.5	88	1126	**70.7**	88	**1126**	91.7	100	1280	91.7	100	**1280**
mRMR	1280	70.5	31	397	**72.7**	52	**666**	91.7	40	512	**92.4**	69	**883**
ReliefF	1280	70.5	49	627	**73.8**	68	**870**	91.7	94	1203	**91.9**	96	**1229**
Ilfs	1280	70.5	27	346	**72.4**	72	**922**	91.7	41	525	**94.2**	58	**742**
Cfs	1280	70.5	59	755	**70.9**	94	**1203**	91.7	98	1254	91.7	98	**1254**
Lasso	1280	70.5	10	128	**70.9**	10	**128**	91.7	98	1254	91.7	98	**1254**

Tab. 4: LGIST features - classification performance based on different feature selection method with 1-NN and SVM classifier. **ACC**: accuracy, **Dim**: dimension, **id %**: percentage of selected features, ⩾ **id %**: percentage of selected features with accuracy equal to all features are used.

Dim	GIST	1-NN						SVM					
		ACC			Max ACC			ACC			Max ACC		
		100 %	⩾ id %	Dim	Max	id %	Dim	100 %	⩾ id %	Dim	Max	id %	Dim
Fisher	512	69.4	42	215	**70.2**	47	**241**	88.3	98	502	**88.3**	98	**502**
mRMR	512	69.4	39	200	**71.4**	53	**271**	88.3	48	246	**90.8**	66	**338**
ReliefF	512	69.4	21	108	**73.4**	70	**358**	88.3	36	184	**90.2**	46	**236**
Ilfs	512	69.4	49	251	**70.0**	79	**404**	88.3	99	507	**88.4**	99	**507**
Cfs	512	69.4	38	195	**71.2**	75	**384**	88.3	49	251	**90.2**	82	**420**
Lasso	512	69.4	40	205	**69.7**	99	**507**	88.3	58	297	**90.6**	78	**399**

Tab. 5: HOG features - classification performance based on different feature selection method with 1-NN and SVM classifier. **ACC**: accuracy, **Dim**: dimension, **id %**: percentage of selected features, ⩾ **id %**: percentage of selected features with accuracy equal to all features are used.

HOG	Dim	1-NN						SVM					
		ACC			Max ACC			ACC			Max ACC		
		100 %	⩾ id %	Dim	Max	id %	Dim	100 %	⩾ id %	Dim	Max	id %	Dim
Fisher	21384	71.5	20	4277	**73.2**	27	**5774**	94.8	85	18176	**94.8**	99	**21170**
mRMR	21384	71.5	8	1711	**73.9**	14	**2994**	94.8	100	21384	**94.8**	100	**21384**
ReliefF	21384	71.5	2	428	**74.4**	3	**642**	94.8	100	21384	**94.8**	100	**21384**
Ilfs	21384	71.5	100	21384	**71.5**	100	**21384**	94.8	100	21384	**94.8**	100	**21384**
Cfs	21384	71.5	8	1711	**72.9**	21	**4491**	94.8	51	10906	**95.1**	74	**15824**
Lasso	21384	71.5	9	1925	**75.5**	19	**4063**	94.8	100	21384	**94.8**	100	**21384**

Tab. 6: The classification results obtained by single and ensemble feature selection.

Classifier	Dataset			Single FS		Multi FS				
	Description	Dim full		ACC without FS (%)	ACC max of FSs (%)	Acc (%)	Dim	Pair	Dim full	Ranker
1-NN	LBP	768		53.0	59.0	**60.0**	432	Ilfs, ReliefF	983	mRMR
	GIST	512		69.4	73.0	**74.6**	261	mRMR Cfs	655	ReliefF
	HOG	21384		71.5	75.5	**79.3**	3416	mRMR ReliefF	3635	ReliefF
	LBP + GIST	1280		70.5	73.8	**77.1**	698	mRMR Ilfs	1587	ReliefF
SVM	LBP	768		77.0	81.8	**82.4**	544	Ilfs mRMR	591	mRMR
	GIST	512		88.3	90.8	**91.4**	1076	Ilfs mRMR Fisher	1346	mRMR
	LBP + GIST	1280		91.7	94.2	**94.0**	1246	mRMR ReliefF	2112	Ilfs

So, we combine two best subset of features determined by ReliefF and ILFS with a feature space equal to 983. Next, this vector is applied again by mRMR method and 1-NN classifier to remove irrelevant features. Table 6 presents the comparison of a single and ensemble feature selection framework. We observe that the ensemble method outperforms single feature selection method for all kinds of features with 1-NN classifier. For example, we increase 1 % of accuracy compared to a single feature selection method and increase 7 % compared with the classification when no selection method is applied. Similar experimental results are obtained by using SVM classifier on single view descriptor. In terms of dimension, we increase the feature space by combining and selecting useful information of different single feature selection methods. Compared with the aims based on accuracy or time computing, an appropriate approach for such demand has to be chosen.

5. Conclusion

In this paper, we introduced a new ensemble feature selection approach by combining multiple single feature selection methods. A pre-selected subset of features is first determined by considering feature selection and associated classifier. Multiple subsets are then combined to form a final feature space and then applied feature selection method again to eliminate noisy and redundant features. The experimental results on the VNRICE dataset for rice seed images classification have shown the efficiency of the proposed approach.

The future of this work is to determine an appropriate selection method based on each attribute and using different strategies to combine the final feature vector resulting from a single feature selection method.

Acknowledgment

References

[1] BENABDESLEM, K. and M. HINDAWI. Constrained Laplacian Score for Semi-supervised Feature Selection. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Berlin: Springer, 2011, pp. 204–218.

ISBN 978-3-642-23780-5. DOI: 10.1007/978-3-642-23780-5_23.

[2] BISHOP, C. M. *Neural Networks for Pattern Recognition.* 1st ed. Oxford: Oxford university press, 1996. ISBN 978-0-198-53864-6.

[3] BOLON-CANEDO, V. and A. ALONSO-BETANZOS. Ensembles for feature selection: A review and future trends. *Information Fusion.* 2019, vol. 52, iss. 1, pp. 1–12. ISSN 1566-2535. DOI: 10.1016/j.inffus.2018.11.008.

[4] CAI, J., J. LUO, S. WANG and S. YANG. Feature selection in machine learning: A new perspective. *Neurocomputing.* 2018, vol. 300, iss. 1, pp. 70–79. ISSN 0925-2312. DOI: 10.1016/j.neucom.2017.11.077.

[5] CHIEW, K. L., C. L. TAN, K. WONG, K. S. C. YONG and W. K. TIONG. A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. *Information Sciences.* 2019, vol. 484, iss. 1, pp. 153–166. ISSN 0020-0255. DOI: 10.1016/j.ins.2019.01.064.

[6] DALAL, N. and B. TRIGGS. Histograms of oriented gradients for human detection. In: *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05).* San Diego: IEEE, 2005, pp. 886–893. IEEE. ISBN 0-7695-2372-2. DOI: 10.1109/CVPR.2005.177.

[7] DENIZ, O., G. BUENO, J. SALIDO and F. DE LA TORRE. Face recognition using Histograms of Oriented Gradients. *Pattern Recognition Letters.* 2011, vol. 32, iss. 12, pp. 1598–1603. ISSN 1598-1603. DOI: 10.1016/j.patrec.2011.01.004.

[8] DROTAR, P., M. GAZDA and L. VOKOROKOS. Ensemble feature selection using election methods and ranker clustering. *Information Sciences.* 2019, vol. 480, iss. 1, pp. 365–380. ISSN 0020-0255. DOI: 10.1016/j.ins.2018.12.033.

[9] DUONG, H.-T. and V. T. HOANG. Dimensionality Reduction Based on Feature Selection for Rice Varieties Recognition. In: *2019 4th International Conference on Information Technology (InCIT).* Bangkok: IEEE, 2019, pp. 199–202. ISBN 978-1-7281-1019-6. DOI: 10.1109/INCIT.2019.8912121.

[10] GOMES, J. F. S. and F. R. LETA. Applications of computer vision techniques in the agriculture and food industry: a review. *European Food Research and Technology.* 2012, vol. 235, iss. 6, pp. 989–1000. ISSN 1438-2385. DOI: 10.1007/s00217-012-1844-2.

[11] GUYON, I. and A. ELISSEEFF. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research.* 2003, vol. 3, iss. 7, pp. 1157–1182. ISSN 1533-7928. DOI: 10.5555/944919.944968.

[12] D. P. VAN HOAI, T. SURINWARANGKOON, V. T. HOANG, H.-T. DUONG and K. MEETHONGJAN. A Comparative Study of Rice Variety Classification based on Deep Learning and Hand-crafted Features. *ECTI Transactions on Computer and Information Technology (ECTI-CIT).* 2020, vol. 14, iss. 1, pp. 1–10. ISSN 2286-9131. DOI: 10.37936/ecti-cit.2020141.204170.

[13] HUMEAU-HEURTIER, A. Texture Feature Extraction Methods: A Survey. *IEEE Access.* 2019, vol. 7, iss. 1, pp. 8975–9000. ISSN 2169-3536. DOI: 10.1109/ACCESS.2018.2890743.

[14] KIRA, K. and L. A. RENDELL. A Practical Approach to Feature Selection. In: *Machine Learning Proceedings 1992.* Aberdeen: Elsevier, 1992, pp. 249–256. ISBN 978-1-558-60247-2. DOI: 10.1016/B978-1-55860-247-2.50037-1.

[15] KONONENKO, I. Estimating attributes: Analysis and extensions of RELIEF. In: *European Conference on Machine Learning.* Springer, 1994, pp. 171–182. ISBN 978-3-540-48365-6. DOI: 10.1007/3-540-57868-4_57.

[16] KONONENKO, I., E. SIMEC and M. ROBNIK-SIKONJA. Overcoming the Myopia of Inductive Learning Algorithms with RELIEFF. *Applied Intelligence.* 1997, vol. 7, iss. 1, pp. 39–55. ISSN 1573-7497. DOI: 10.1023/A:1008280620621.

[17] MEBATSION, H. K., J. PALIWAL and D. S. JAYAS. Automatic classification of non-touching cereal grains in digital images using limited morphological and color features. *Computers and Electronics in Agriculture.* 2013, vol. 90, iss. 1, pp. 99–105. ISSN 0168-1699. DOI: 10.1016/j.compag.2012.09.007.

[18] MIFTAHUSHUDUR, T., C. B. A. WAEL and T. PRALUDI. Infinite Latent Feature Selection Technique for Hyperspectral Image Classification. *Jurnal Elektronika dan Telekomunikasi.* 2019, vol. 19, iss. 1, pp. 32–37. ISSN 2527-9955. DOI: 10.14203/jet.v19.32-37.

[19] MIRZAEI, A., V. POURAHMADI, M. SOLTANI and H. SHEIKHZADEH. Deep feature selection using a teacher-student network. *Neurocomputing.* 2020, vol. 383, iss. 1, pp. 396–408. ISSN 0925-2312. DOI: 10.1016/j.neucom.2019.12.017.

[20] NHAT, H. T. M. and V. T. HOANG. Feature fusion by using LBP, HOG, GIST descriptors and Canonical Correlation Analysis for face recognition. In: *2019 26th International Conference on Telecommunications (ICT)*. Hanoi: IEEE, 2019, pp. 371–375. ISBN 978-1-7281-0273-3. DOI: 10.1109/ICT.2019.8798816.

[21] OJALA, T., M. PIETIKAINEN and T. MAEN-PAA. A Generalized Local Binary Pattern Operator for Multiresolution Gray Scale and Rotation Invariant Texture Classification. In: *International Conference on Advances in Pattern Recognition*. Rio de Janeiro: Springer, 2001, pp. 399–408. ISBN 978-3-540-41767-5. DOI: 10.1007/3-540-44732-6_41.

[22] OLIVA, A. and A. TORRALBA. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision*. 2001, vol. 42, iss. 3, pp. 145–175. ISSN 0920-5691. DOI: 10.1023/A:1011139631724.

[23] SEIJO-PARDO, B., I. PORTO-DIAZ, V. BOLON-CANEDO and A. ALONSO-BETANZOS. Ensemble feature selection: Homogeneous and heterogeneous approaches. *Knowledge-Based Systems*. 2017, vol. 118, iss. 1, pp. 124–139. ISSN 0950-7051. DOI: 10.1016/j.knosys.2016.11.017.

[24] VAN, T. N. and V. T. HOANG. Kinship Verification based on Local Binary Pattern features coding in different color space. In: *2019 26th International Conference on Telecommunications (ICT)*. Hanoi: IEEE, 2019, pp. 376–380. ISBN 978-1-7281-0273-3. DOI: 10.1109/ICT.2019.8798781.

[25] YAMADA, M., W. JITKRITTUM, L. SIGAL, E. P. XING and M. SUGIYAMA. High-Dimensional Feature Selection by Feature-Wise Kernelized Lasso. *Neural Computation*. 2014, vol. 26, iss. 1, pp. 185–207. ISSN 1530-888X. DOI: 10.1162/NECO_a_00537.

[26] ZHANG, R., F. NIE, X. LI and X. WEI. Feature selection with multi-view data: A survey. *Information Fusion*. 2019, vol. 50, iss. 1, pp. 158–167. ISSN 1566-2535. DOI: 10.1016/j.inffus.2018.11.019.

[27] ZHAO, Z., R. ANAND and M. WANG. Maximum Relevance and Minimum Redundancy Feature Selection Methods for a Marketing Machine Learning Platform. In: *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. Washington: IEEE, 2019, pp. 442–452. ISBN 978-1-7281-4493-1. DOI: 10.1109/DSAA.2019.00059.

About Authors

Dzi Lam Tran TUAN is a graduate student of Ho Chi Minh City Open University, Vietnam. He obtained his M.Sc. degree in Computer Science in 2019. His research interests include feature selection and image analysis.

Thongchai SURINWARANGKOON was born in Suratthani, Thailand in 1972. He obtained a B.Sc. degree in mathematics from Chiang Mai University, Thailand in 1995. He received M.Sc. degree in management of information technology from Walailak University, Thailand in 2005 and Ph.D. in information technology from King Mongkut's University of Technology North Bangkok, Thailand in 2014. He is now a Lecturer in the Department of Business Computer, Suan Sunandha Rajabhat University, Bangkok, Thailand. His research interests include digital image processing, machine learning, artificial intelligence, business intelligence, and mobile application development for business.

Kittikhun MEETHONGJAN is a full lecturer in the Computer Science Program, Head of Apply Science Department, Faculty of Science and Technology, Suan Sunandha Rajabhat University (SSRU), Thailand. He received his B.Sc. Degree in Computer Science from SSRU in 1990, a M.Sc. degree in computer science from King Mongkut's University of Technology Thonburi (KMUTT) in 2000 and a Ph.D. degree in Computer Graphic from the University of Technology, Malaysia, in 2013. His research interest is Computer Graphic, Image Processing, Artificial Intelligent, Biometrics, Pattern Recognition, Soft Computing techniques and, application. Currently, he is an advisor of Master and Ph.D. student in Forensic Science of SSRU.

Vinh Truong HOANG received his M.Sc. degree from the University of Montpellier in 2009 and his Ph.D. degree in computer science from the University of the Littoral Opal Coast, France. He is currently an assistant professor and Head of Image Processing and Computer Graphics Department at the Ho Chi Minh City Open University, Vietnam. His research interests include image analysis and feature selection.