Pertanika J. Sci. & Technol. 25 (S): 315 - 324 (2017)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Binary Classification using SVM for Sick and Healthy Chicken based on Chicken's Excrement Image

Nurhanna Abdul Aziz[1] and Mohd Fauzi Bin Othman[1]*

[1]*Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Jln Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia*

ABSTRACT

The purpose of this paper is to classify between healthy and sick chicken based on their dropping. Most chicken farm management system in Malaysia is highly dependent on human surveillance method. This method, however, does not focus on early disease detection hence, unable to and alert chicken farmers to take necessary action.. Therefore, the need to improve the biosecurity of chicken poultry production is essential to prevent infectious disease such as avian influenza. The classification of sick and healthy chicken based solely on chicken's excrement using the support vector machine is proposed. First, the texture is examined using grey-level co-occurrence matrix (GLCM) approach. A GLCM based texture feature set is derived and used as input for the SVM classifier. Comparison are made using more and then less extracted features, less extracted features and also applying Gabor filter to these features to see the effect it has on classification accuracy. Results show that having more features extracted using GLCM techniques allows for greater classification accuracy.

Keywords: Support vector machine, feature extraction, GLCM, Gabor filter

INTRODUCTION

Poultry, especially chicken, is the primary source of protein in Malaysia. According to the recent USDA statistic chicken meat consumption in Malaysia is the highest in the world. It has increased from 1.4 million metric tons in 2013 to 1.43 million metric tons in 2014. To meet this huge demand imports from China are needed. In order for Malaysia to become a trusted producer in the halal chicken meat industry, it should have a good poultry management system. The significant issues are diseases such as avian influenza. In this field, observation is highly consequential to discover diseases at an early stage because when the disease is in one of the last stages, the chicken is possibly not

ARTICLE INFO

Article history:
Received: 24 August 2016
Accepted: 02 December 2016

E-mail addresses:
nurhanna2@live.utm.my (Nurhanna Abdul Aziz),
mdfauzi@utm.my (Mohd Fauzi Bin Othman)
*Corresponding Author

ISSN: 0128-7680

treatable anymore (den Boer, van den Hout, & Vervloed, 2014). Some of the signs are sudden diarrhoea, decreased egg production, sneezing, nasal discharge, coughing, gasping for air, lack of energy and appetite, swelling of tissues around the eyes and neck, purple discoloration of the wattles, combs and legs and depression, muscular tremors, drooping wings, twisting of head and neck, incoordination and complete paralysis. Chicken disease can also be recognized by its dropping's, colour and density. Early detection of disease and chicken health can facilitate the control of diseases through vector control of vaccination applications, disease-specific approach; and improved productivity. Previous work carried out by (Zhu, Peng, & Ji, 2009) mainly focused on detecting chicken that died. One way to improve the system is to examine chicken excrement images using Gabor- GLCM approach with SVM classification. This paper is divided into three sections. Section One discusses the importance of having early disease detection system in chicken poultry. Section Two emphasizes on the applied methodology, framework used and the experiment's result. We conclude our hypothesis and future work that need to be done in Section Three.

METHODOLOGY

In the methodology section, this paper will describe briefly the designated framework, how the data is collected and pre-processed and also the feature extraction and the classifications methods.

Framework

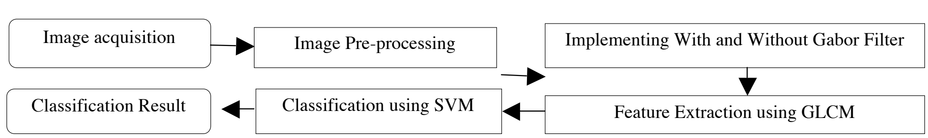

Figure 1. The chicken's disease detection framework

Image Acquisition

Image acquisition of chicken's dropping was obtained by using a static digital camera. The distance of the camera has been set to be at least 90 degrees off the camera. All the photos were decoded in JPEG standard format in 4000x3000 pixels (dot). The manual identification of sick and normal chicken's excrement images was carried out based on their characteristic by an authorized veterinary. The data set contains 20 images of the same background. Figure 2 depicts some of the eight images that are classified as 'sick' or 'stressed' and Figure 3 is a sample of eight images classified as 'healthy' or 'normal'.

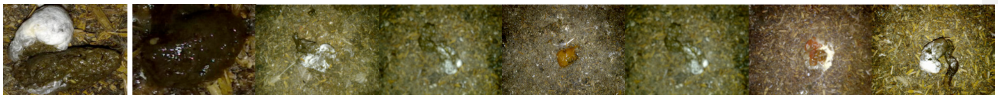

Figure 2. Sick chicken's excrement's images

Figure 3. Healthy chicken's excrement's images

Image Pre-processing

Image pre-processing is used to improve and enhance the quality of the images. Some of the images suffered from noise, blurry and low contrast of quality. Thus, the basic steps of image pre-processing were used and listed as follows:

1. Set the resolution of the images to 300x300 dpi.
2. Resize the image to 200x200 pixels
3. Convert the image from binary to the grayscale.

Feature Extraction Using GLCM and Gabor Filter

One of the major tasks in image processing is feature extraction. It is believed that the nature of the surface can be characterized by the property of the texture (Hammouda & Jernigan, 2000). The texture itself contains much information about the structural arrangement and its relationship to its surrounding (Raju & Durai, 2013). Therefore, it is crucial to have feature extraction method so that the image can be easily classified. It is also a process that represents the raw image that can ease the decision-making process. As described by (Min et al, 2006), texture features are extracted by simulating the perceptual properties such as orientation, coarseness, fineness, and regularity. There are numerous ways to extract and classify the features, but this paper only focuses on Grey-level co-occurrence matrix (GLCM). GLCM is proven to be a very powerful tool for quantifying the intensity variation (Ahmed, Bayraktar, Bhunia, Hirleman, Robinson, & Rajwa, 2013). The use of the GLCM concept for texture can be seen in works done in (Siraj, Salahuddin, & Yusof, 2010; Arebey, Hannan, Begum, & Basri, 2012). As it is defined as the frequencies of grey-level values that occur in an image (thus some of the brief computation of texture feature extraction summarized (Level, Pramunendar, Supriyanto, Novianto, & Yuwono, 2013) is shown below.

$$entropy = -\sum_{i=0}^{N-1} p_{ij} \log p_{ij} \tag{1}$$

$$contrast = \sum_{i=0}^{N-1} n^2 \left\{ \sum_{i=0}^{N} \sum_{i\backslash j=0}^{N} p_{ij} \right\}, |i-j| = n \tag{2}$$

$$correlation = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{(i-\mu_x)(j-\mu_y)}{\sqrt{\sigma_x\sigma_y}}p_{ij} \tag{3}$$

$$autocorr = \sum_{i=0}^{N-1}\sum_{i\setminus j=0}^{N-1}(ij)\mathrm{p}_{ij} \tag{4}$$

$$inertia = \sum_{i=0}^{N-1}\sum_{i\setminus j=0}^{N-1}(i-j)^2\mathrm{p}_{ij} \tag{5}$$

Classification Using SVM

Support vector machine (SVM) is a novel type of learning machine, which is based on statistical learning theory (SLT). That is, an SVM is an approximate implementation of the method of structural risk minimization. SVM has shown to provide a better generalization performance than traditional techniques, including neural networks (Chih-Min et al, 2006). SVM demonstrates good classification performance in (Khedher, Ramírez, Górriz, Brahim, & Segovia, 2015; Dubey, 2014; Ahmed, Kang, Kang, Ko, Cho, Rhee, S. & Yu, 2015). SVM also has been applied in most of the fields such as face recognition, fingerprint, bioinformatics, and it has also been tested and applied to agriculture fields such as research works done by (Khedher et al., 2015; Sharaf-Eldeen, Moawad, El Bahnasy, & Khalifa, 2012). The basic idea of SVM is that it seeks to maximize the distance between two classes, and the distance between classes is traditionally defined by the closest points (Hammouda, K. & Jernigan, E., 2000). It is a very effective method for general purpose pattern recognition. (Chih-Min et al, 2006). SVM is popular for its capability in generalising in and predicting s with a good degree of accuracy. (Siraj et al., 2010). Optimal hyper- plane is derived in a high dimensional feature space that defines a maximum boundary margin between data samples in two classes which provides a better generalization property. With its latest extensions enabled the SVM to learn and classify multiple categories of data, overlapping classes and noisy data by the introduction of slack variables that enable the soft margin classifier (Elhariri et al, 2014). Basically, the SVM is modelled as in (Kazemian & Ahmed, 2015). This paper only classifies and compares sick and healthy chicken's excrement image only.

Experiments and Classification Result

The experiments have been tested on a 4GB RAM, Intel Core i7 CPU 1.6 GHz using Matlab 2012b release. We had 20 samples of chicken's excrement images in which 20% of them is used as our test data, and the balance remains as our training data. GLCM statistical calculation is applied to extract the features which include autocorrelation, contrast, energy, entropy, homogeneity, cluster prominence, cluster shade, correlation, difference entropy, difference variance, dissimilarity, inverse difference moment, information measure of correlation 1, information measure of correlation 2, inverse difference, maximum probability, sum average, sum entropy, sum of squares and sum of variance. Thus, each image will have 19 features subsequently. Based on these extracted features, we design four types of experiments. The first and second sets of experiment are to determine the accuracy rate of the classification by applying Gabor filters with five different orientations with only four features extracted out of 19. Experiments three and four focus on finding accuracy rate SVM classification without applying any Gabor filter with four extracted features and 19 extracted features consequently. Figures 5(a) and 5(b) show some samples of healthy and sick chicken dropping images that applied Gabor filter with different orientation approach while Figures 6(a) and 6(b) show some sample images with no Gabor filter applied.

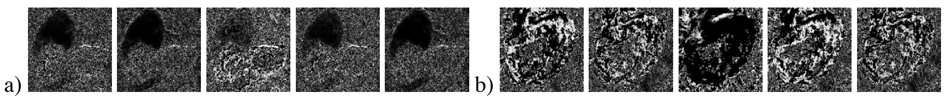

Figure 5. (a) A healthy chicken's excrement image with different orientations (0°, 45°, 90°,135° and 180°) of Gabor filter; (b) Sick chicken's excrement image

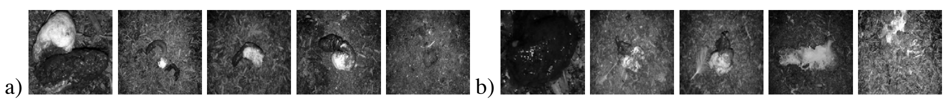

Figure 6. Sample of healthy chicken's excrement with no Gabor filter applied; (b) Sick chicken's excrement

Based on both Gabor and non-Gabor filter images, we applied the statistical calculation of GLCM to extract the features. Table 1 shows some of the GLCM calculation applied on both healthy and sick chicken's excrement images. The reason we choose autocorrelation, contrast, energy and entropy is based on the GLCM calculation in which the result outperform others features. Nevertheless, we still need to do the comparison with other 19 features to determine either this supporting features may affect the result of the classification or not.

Table 1
GLCM statistical calculation applied on both sample healthy and sick images with 0° orientation of GABOR filter

Image#	Autocorrelation	Contrast	Energy	Entropy
Image Healthy 1	6.4733	9.9763	0.0723	3.0739
Image Healthy 2	6.8131	6.1299	0.0558	3.1496
Image Healthy 3	5.7578	4.7044	0.0682	2.952
Image Sick 4	4.9985	4.8265	0.0833	2.7861
Image Sick 5	11.194	53.899	0.0756	3.3975

After the GLCM implementation, all the dataset is scaled and normalised based on the minimum and maximum value of each row before running the classification method. Classification in SVM requires the dataset to be divided into two sets of data where one set of data is used as a training set and another one as a test set. Each row of data contains its feature label. To implement the SVM approach, LibSVM 3.20 was used to fit this experiment purposes being a SVM library tool used widely in classifications (Zhang, 2014; Moraes, Valiati, Gavião, & Neto, 2013; Ahmed et al., 2012; Wang & Chung, 2013; Tang & Sazonov, 2014; Tong, et al., 2014). The toolbox was developed by Chang and Lin (2011) which solves the quadratic programming problem by using a sequential minimal optimization-type algorithm (Kang, S. et al ,2015). We used 80% of the total number of feature vectors as our training sample while the remaining 20% of a test sample. All the data must be converted into LibSVM format before classification can take place. We had labelled sick images equal to -1 and healthy images as one respectively. We used the radial basis kernel function as it is more significant to our case as compared to other kernel function as a polynomial and applied 10-cross validation in our dataset. We also set the parameter C equal to 1 for all the cases.

$$f(x) = \sum_{i=0}^{N-1} \alpha_i y_i \exp(-\| x - x_i \|^2 / 2\sigma^2) + b \tag{6}$$

Tables (2-3) compares the classification accuracy results between four extracted features (correlation, contrast, homogeneity, energy) and 19 extracted features (autocorrelation, contrast, energy, entropy, homogeneity, cluster prominence, cluster shade, correlation, difference entropy, difference variance, dissimilarity, inverse difference moment, information measure of correlation 1, information measure of correlation 2, inverse difference, maximum probability, sum average, sum entropy, sum of squares and sum of variance) by using GLCM techniques with Gabor filter applied. The value is set to be 3.05175 as it is the best value we had tested in our case.

Table 2
Classification result- with GABOR and 4 extracted features within same dataset.Ơ (= 3.05175)

Kernel Type	Orientation	Classification Accuracy	#Iteration	#SV
Radial	0	81.25%	19	16
	45	81.25%	15	16
	90	75%	21	16
	135	68.75%	14	16
	180	81.25	13	16

Table 3
SVM Classification result- with GABOR and 19 extracted features within the same dataset. Ơ (= 3.05175)

Kernel Type	Orientation	Classification Accuracy	#Iteration	#SV
Radial	0	81.25%	18	16
	45	81.25%	31	16
	90	81.25%	21	16
	135	75%	16	16
	180	81.25%	21	16

Both experiments use the same dataset for training and testing. It can be seen that by having more features extracted with Gabor filter applied to the images increase the accuracy rate compares with less feature extracted. Table (4-5) presents the accuracy rate for both four features and 19 features with 80% of the sample data is used for training and the remaining 20% for testing. Using the same training model, the accuracy rate for an image that has 19 extracted features yields much more than the image that only has four features extracted.

Table 4
SVM Classification result- with GABOR and 19 extracted features by using the test sample Ơ = 3.05175; # TEST SAMPLE=4; # TRAINING SAMPLE=16

Kernel Type	Orientation	Classification Accuracy	Iteration	#SV
Radial	0	75%	18	16
	45	50%	31	16
	90	75%	21	16
	135	75%	16	16
	180	75%	21	16

Table 5
SVM Classification result- with GABOR and 4 extracted features by using the test sample Ơ = 3.05175; # TEST SAMPLE=4; # TRAINING SAMPLE=16

Kernel Type	Orientation	Classification Accuracy	#Iteration	#SV
Radial	0	50%	19	16
	45	50%	15	16
	90	50%	21	16
	135	50%	14	16
	180	50%	13	16

Table (6-7) yields the accuracy rate comparison between fewer and more features but without any Gabor filter applied. It proves that even with no Gabor filter applied, the image which has 19 features extracted gives higher accuracy rate than four features extracted.

Table 6
SVM Classification result- with no GABOR and 4 extracted features (Ơ = 3.05175

Kernel Type	Dataset	Accuracy Rate	#Iteration	#SV
Radial	Training	81.25%	13	13
	180	50%	13	16

Table 7
SVM Classification result- with no Gabor and 19 extracted features (Ơ = 3.05175)

Kernel Type	Dataset	Accuracy Rate	#Iteration	#SV
Radial	Training	93.75%	19	16

The performance measures are described regarding true and false positive and true and false negative.

- True Positive (TP): Sick chicken correctly identified as sick chicken.
- True Negative (TN): Healthy chicken correctly identified as healthy chicken.
- False Negative (FN): Sick chicken incorrectly identified as healthy chicken.
- False Positive (FP): Healthy chicken incorrectly identified as sick chicken.

As a conclusion, from Table (2- 7), it indicates that the accuracy rate improves if more features extracted as compared by extracting only four characteristics of the images with Gabor filter applied. It also shows that the implementation of Gabor filter bank does not give the major impact on the accuracy rate. The present finding support (Kazemian & Ahmed, 2015) research work which concluded that as the number of the dataset (in this case is the number of extracted features) increased, the accuracy also increases. Choosing the best optimum kernel for each of the cases also plays an important task in increasing the classification accuracy.

CONCLUSION

This paper presents classification of texture images specifically of chicken's excrement using SVM classification approach and the GLCM method to extract the texture feature of images. It also applies the Gabor filter technique as part of feature extraction method. The paper draws a comparison between quantities of features that are extracted and conducted on g 20 sample images of chicken's dropping. We did four experimental studies. The first is between four features and 19 features by applying the Gabor filter. We found out that having more features extracted thru GLCM techniques yields results with better accuracy as compared with fewer features. The second study is between 4 features and 19 features but without Gabor filter applied which also proves that having more extracted features have more advantage than fewer features even though their result is much better than having Gabor filter applied on. Analysis of the results reveals that choosing the best optimization value for gamma in SVM modeler gives better accuracy. We also prove that having more features extraction for SVM classification give more accurate result specifically for chicken's excrement images in which there is a similarity in color between the features and background color. Our future work will be on texture feature extraction method as we believe that having a good extraction framework would provide better results.There is also the need to focus on identifying nearly sick chicken by analyzing the 'stressed' chicken images using the same framework with expansion on SVM algorithm.

ACKNOWLEDGEMENT

Special thanks to Malaysia- Japan International Institute, Universiti Teknologi Malaysia and also Centre of Artificial Intelligence and Robotic (CAIRO) department.

REFERENCES

Ahmed, F., Al-Mamun, H. A., Bari, A. H., Hossain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection, 40*, 98-104.

Ahmed, W. M., Bayraktar, B., Bhunia, A. K., Hirleman, E. D., Robinson, J. P., & Rajwa, B. (2013). Classification of bacterial contamination using image processing and distributed computing. *IEEE Journal of Biomedical and Health Informatics, 17*(1), 232-239.

Arebey, M., Hannan, M. A., Begum, R. A., & Basri, H. (2012). Solid waste bin level detection using gray level co-occurrence matrix feature extraction approach. *Journal of Environmental Management, 10*4, 9-18.

Chen, C. M., Chen, C. C., & Chen, C. C. (2006, August). A Comparison of Texture Features Based on SVM and SOM. In I*CPR (2)* (pp. 630-633).

den Boer, C. A., van den Hout, N., & Vervloed, L. (2014). Poultry farming

Dubey, S. R., & Jalal, A. S. (2014). Adapted approach for fruit disease identification using images. arXiv preprint *arXiv:1405.4930*.

Elhariri, E., El-Bendary, N., Fouad, M. M. M., Platoš, J., Hassanien, A. E., & Hussein, A. M. (2014). Multi-class SVM based classification approach for tomato ripeness. In *Innovations in Bio-inspired Computing and Applications* (pp. 175-186). Springer International Publishing.

Hammouda, K., & Jernigan, E. (2000). *Texture segmentation using gabor filters*. Canada: Center for Intelligent Machines, McGill University.

Kang, S., Kang, P., Ko, T., Cho, S., Rhee, S. J., & Yu, K. S. (2015). An efficient and effective ensemble of support vector machines for anti-diabetic drug failure prediction. *Expert Systems with Applications, 4*2(9), 4265-4273.

Kazemian, H. B., & Ahmed, S. (2015). Comparisons of machine learning techniques for detecting malicious webpages. *Expert Systems with Applications, 42*(3), 1166-1177.

Khedher, L., Ramírez, J., Górriz, J. M., Brahim, A., Segovia, F., & Alzheimer′ s Disease Neuroimaging Initiative. (2015). Early diagnosis of Alzheimer′ s disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing, 151*, 139-150.

Moraes, R., Valiati, J. F., & Neto, W. P. G. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications, 40*(2), 621-633.

Pramunendar, R. A., Supriyanto, C., Novianto, D. H., Yuwono, I. N., Shidik, G. F., & Andono, P. N. (2013, November). A classification method of coconut wood quality based on Gray Level Co-occurrence matrices. In *Robotics, Biomimetics, and Intelligent Computational Systems (ROBIONETICS), 2013 IEEE International Conference on* (pp. 254-257). IEEE

Raju, J., & Durai, C. A. D. (2013, February). A survey on texture classification techniques. In *Information Communication and Embedded Systems (ICICES), 2013 International Conference on* (pp. 180-184). IEEE.

Sharaf-Eldeen, D. A., Moawad, I. F., El Bahnasy, K., & Khalifa, M. E. (2012, December). Learning and applying range adaptation rules in case-based reasoning systems. In *International Conference on Advanced Machine Learning Technologies and Applications* (pp. 487-495). Springer Berlin Heidelberg.

Siraj, F., Salahuddin, M. A., & Yusof, S. A. M. (2010, September). Digital image classification for malaysian blooming flower. In *2010 Second International Conference on Computational Intelligence, Modelling and Simulation* (pp. 33-38). IEEE.

Song, W., Xiong, B., Sun, Z., Mo, H., & Su, Q. (2010). Development and application of assistant system for diagnosing chicken diseases based on smart mobile phones. *Transactions of the Chinese Society of Agricultural Engineering, 26*(4), 220-226.

Tang, W., & Sazonov, E. S. (2014). Highly accurate recognition of human postures and activities through classification with rejection. *IEEE Journal of Biomedical and Health Informatics, 18*(1), 309-315.

Tong, T., Wolz, R., Gao, Q., Guerrero, R., Hajnal, J. V., Rueckert, D., & Alzheimer's Disease Neuroimaging Initiative. (2014). Multiple instance learning for classification of dementia in brain MRI. *Medical Image Analysis, 18*(5), 808-818.

Wang, S., Wang, J., & Chung, F. L. (2014). Kernel density estimation, kernel methods, and fast learning in large data sets. *IEEE Transactions on Cybernetics, 44*(1), 1-20.

Zhang, C. (2015). Applying data fusion techniques for benthic habitat mapping and monitoring in a coral reef ecosystem. *ISPRS Journal of Photogrammetry and Remote Sensing, 104*, 213-223

Zhu, W., Peng, Y., & Ji, B. (2009, December). An automatic dead chicken detection algorithm based on SVM in modern chicken farm. In *2009 Second International Symposium on Information Science and Engineering* (pp. 323-326). IEEE.