Is color an intrinsic property of construction object's representation? Evaluating color-based models to detect objects by using data mining techniques

Nahyae Hwang [1], Hyojoo Son [2], and Changwan Kim [3]*

[1] Department of Architectural Engineering, Chung-Ang University, Seoul, Korea
[2] Research Institute for Energy and Environment, Chung-Ang University, Seoul, Korea
[3] Department of Architectural Engineering, Chung-Ang University, Seoul, Korea
* Corresponding author (changwan@cau.ac.kr)

Purpose Structural component detection is a prerequisite for various applications, including construction progress measurement and quality inspection. However, it is still a challenge to detect structural components reliably in construction site images taken from a complex and unstructured construction environment. Because construction site images contain numerous unexpected objects, structural components in the images are observed under different poses and varying lighting conditions. The aim of this study is to discover how color information effectively works on structural component detection in construction site images by incorporating hybrid data mining techniques. **Method** To verify the effectiveness of the color-based models for structural components detection, this study involves data collection, feature selection, and color-based model building. First, this study tried to collect the most comprehensive data set on structural components detection before assessment. Second, it attempted to extract the best set of effective color features among all the available color features through feature selection. Third, this study evaluated and compared the performance of the constructed color-based models (defined in terms of accuracy rate) using hybrid data mining techniques. This study then identified the most effective configuration of color features and data mining techniques to detect structural components. **Results & Discussion** The experimental results suggest that color can be a powerful cue for reliable detection of structural components in construction site images. The use of the set of color features in combination with a hybrid data mining technique in structural component detection is highly accurate (accuracy rate above 95%) in detecting structural components composed of major construction materials (e.g. concrete, steel, and wood). The results from structural components detection that are obtained by the proposed combination are reliable for use as an essential input for various applications, including construction progress measurement and quality inspection.

Keywords: automation, color, construction object detection, data mining technique, ensemble model

INTRODUCTION

The automatic detection of objects in photographs acquired from construction sites is essential for a wide range of applications, including, but not limited to, materials tracking, automated control of heavy equipment, progress monitoring, quality control, and the generation of 3D as-built models. As the first step of processing construction photographs, successful and efficient object detection is critical to the success of these applications and greatly affects the subsequent steps. However, it is challenging to automatically detect objects in construction photographs due to the possible appearance and shape variations of the objects. Unlike a well-structured manufacturing environment, a construction site contains objects of unpredictable shape and position due to the uncertain and unstructured nature of the environment. Moreover, a construction photograph contains a variety of objects that are cluttered, intertwined, occluded, and articulated. This complexity, which is inherent to most construction sites, makes the automatic detection of particular objects non-trivial and

difficult. To handle the problem of object detection in such an unstructured and complex environment, an object detection approach based on the colors of the objects has been proposed as an effective way to distinguish an object of interest from the other objects in its vicinity. The color-based object detection approach has obvious and powerful advantages, especially in unstructured and complex environments, because it is independent on the appearance, shape, and position of the objects. In addition, this approach is simple and computationally efficient to implement because it only requires the color values of each pixel in construction photographs. By taking advantage of this approach, color-based object detection has been successfully validated for use in a number of areas (e.g., detection of skin, roads, traffic signs, and so forth). The results of previous studies indicate that color is a useful and robust feature for object detection. This approach seems to be useful for detection of particular objects in construction photographs.

Despite the fact that automatic detection of objects in construction photographs has been explored in

some recent studies within the construction industry, whether the color of an object is actually useful in the automatic detection of that object in such photographs has not been investigated. In addition, a method to test the usefulness of the color of an object in its detection is rarely considered. Consensus about whether the color of an object is genuinely useful in the detection of that object in construction paragraphs has not been reached among researchers in the construction industry. Most researchers are concerned about the perceived difficulty of detecting particular objects in construction photographs using only their inherent color properties, particularly when the color of the object of interest is similar to the others around it or even when the inherent color property of the object of interest is altered due to the effect of varying illumination that occurs in outdoor environments. Based on these concerns, researchers have regarded the detection of specific objects in construction photographs solely by colors as insufficient and difficult. Hence, there have been several attempts to detect particular objects in construction photographs using the color of an object together with its shape[1] or with its texture.[2] However, the methods proposed by these researchers suffer from limitations resulting from a lack of consideration of the variations in object surface properties such as texture, changes in illumination, viewing direction, partial occlusion of the object of interest, and the presence background objects.

The aim of this study is to clearly verify whether an object's color information is useful for detecting that object in construction site photographs, with the questioning centering on manmade objects of concrete and steel and natural objects of wood. This paper collects a comprehensive data set for analysis and describes how we can consider the effects of varying illumination in an outdoor environment to take full advantage of color for object detection. Then, the usefulness of color information for detecting an object is verified through incorporating data mining techniques as the main focus of an investigation. The paper is organized as follows: in the next section, the data set used in the study is described. Section 3 briefly characterizes the eight data mining techniques selected for the study. Section 4 presents experimental results, and the last section concludes with a discussion.

MATERIALS

Data Set

Without a comprehensive data set, it is meaningless to declare the detection of a color-based object "useful" or "not useful," let alone quantify its utility. Since the comprehensive data sets for construction object detection are not readily available in the construction

industry, a total of three data sets for concrete, steel, and wood detection, respectively, were generated.

The appearance of a construction object's surface colors can be affected by environmental factors such as changes in the direction and intensity of illumination. Because most construction sites are outdoors, illumination intensity varies unpredictably and uncontrollably, depending on the time of day, seasonal variations, and weather conditions (sunny, cloudy, or foggy), thereby resulting in large variations in the appearance of a construction object's surface colors. To cover such variations, 108 photographs were taken at a total of 50 construction sites for concrete detection, 91 photographs were taken at a total of 80 construction sites for steel detection, and 50 photographs were taken at a total of 14 construction sites for wood detection.

Each photograph was then divided into 25 × 25 or 50 × 50 pixel subregions. The subregions were categorized and labeled as an object of interest, background, or indeterminate. The subregions categorized as indeterminate were excluded from the dataset. As a result, each component of the data set is composed of object and non-object pixels. The former refers to pixels related to objects such as concrete, steel, and wood, while the latter refers to pixels related to background. In order to verify whether only the color of objects made of concrete, steel, and wood is sufficient for distinguishing these objects from others, this study made a particular effort to collect and include as many objects as possible with similar color properties as the objects of interest. The background objects include all kinds of scenery—bricks, construction equipment, fences, forms, pipes, safety nets, the sky, soil, traffic signs, trees, windows, and other construction-related materials.

In total, the data collected from the concrete, steel, and wood subregions and their background subregions amounted to over 113 million pixels, 95 million pixels, and 35 million pixels, respectively, for concrete, steel, and wood detection. The first concrete data set contains approximately 44 million pixels from the concrete and approximately 69 million pixels from the background pixels. The second steel data set consists of 9 million pixels from the steel and approximately 85 million pixels from the background. The third wood data set has 10 million pixels from the wood and approximately 25 million pixels from the background. The characteristics of the three data sets are provided in Table 1. In summary, the data sets were well balanced in terms of time of day, season, and weather. The percentages of data collected during a.m. and p.m. hours and PM hours were roughly comparable. There was greater variation in the percentages of data collected in different seasons, as well as in the percentages of data collected under different weather conditions, as shown in Table 1.

Table 1. Distribution of data collection by time of day, season, and weather

Category	Data Set	Concrete (%)	Steel (%)	Wood (%)
Time of Day	a.m.	56.9	33.0	30.0
	p.m.	43.1	67.0	70.0
Season	Spring	23.2	9.9	24.0
	Summer	16.7	16.5	28.0
	Fall	46.8	32.9	24.0
	Winter	13.2	40.7	24.0
Weather	Cloudy	13.1	18.7	44.0
	Foggy	11.5	8.8	4.0
	Sunny	75.4	72.5	52.0

Data Representation

Another issue that must be carefully considered to take full advantage of color-based object detection is how to effectively reduce the effect of varying illumination on object color changes that occur in outdoor environments. The values of colors in the RGB (red, green, and blue) colorspace, the most prevalent choice for computer graphics, are particularly subject to deterioration as a result of changes in illumination.[3] Such variations caused by factors in outdoor environments may dramatically affect color properties, potentially impacting detection performance.[4] To deal with this potential artifact, it is important to newly represent the color to minimize the effect of illumination changes. In this study, in order to represent the color in an appropriate form, the color values in RGB colorspace were converted to HSI color space, which is impervious to illumination changes. The HSI colorspace consists of three components: hue (H), saturation (S), and intensity (I). Hue and saturation are related to color, or chromaticity, and the illumination-independent components. The HSI colorspace is defined[5] as:

$$H = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right)$$

$$S = 1 - \frac{\min(R,G,B)}{I}$$

$$I = \frac{(R+G+B)}{3}$$

METHODS

Classification Models

Seven classification models and an ensemble model were applied and compared: 1) artificial neural networks (ANN), 2) support vector machines (SVM), 3) classification and regression tree (CART), 4) quick unbiased efficient statistical tree (QUEST), 5) commercial version 5.0 (C5.0), 6) exhaustive chi-squared automatic interaction detector (CHAID), 7) logistic regression (LR), and 8) ensemble model. The first two models are derived from machine learning tech-

niques, the latter four models from classification and regression-based techniques, and the seventh model from multivariate statistical techniques.

Artificial Neural Networks (ANN)

In this study, a back-propagation (BP) neural network is used for classification. This is a feed forward network that can have 1 or more hidden layers. It utilizes an iterative gradient search technique designed to minimize the mean square error between the actual and desired net outputs. The units in the hidden layer sum their inputs, add a constant, and take a fixed function of the result.[6] The output units are of the same form, but with output function. A three-layer network with 1 hidden layer was proven to be capable of computing any continuous likelihood function required in a classifier and solving complex binary classification problems.[7, 8] The logistic function was selected as the activation function in this study.

Support Vector Machine (SVM)

The SVM, inspired by a statistical learning theory, is one of the most powerful machine learning techniques for solving a large number of complex binary classification problems.[9] It acts as a linear classifier in a high dimensional feature space transformed through a projection from the original input feature space by taking non-linear functions (kernel) of the original data set.[10] Hence, in general, the resulting classifier is non-linear in the input space. The SVM achieves good generalization performances by finding a hyperplane that maximizes the margin between the classes. The radial basis function (RBF) was selected as the kernel function in this study. For further details, see Vapnik[11] and Vladimir and Vapnik.[12]

Classification and Regression Tree (CART)

CART is one of the decision tree algorithms that induces a binary tree on a given set of training data, resulting in a set of "if-then" rules. These rules can then be used to solve classification or regression problems. CART is a robust, easy-to-use decision tree tool that automatically sifts large, complex databases, searching for and isolating significant patterns and relationships.[13] It uses a recursive partitioning, a combination of exhaustive searches and intensive testing techniques, to identify useful tree structures in the data. The knowledge thus discovered is used to generate a decision tree, resulting in reliable, easy-to-grasp predictive models. It constructs the decision tree by splitting subsets of the data set using all predictor variables to create two child nodes repeatedly, beginning with the entire data set. The best predictor is chosen using a variety of impurity or diversity measures. The goal is to produce subsets of the data that are as homogeneous as possible with respect to the target variable.[14]

Quick Unbiased Efficient Statistical Tree (QUEST)
QUEST is a binary-split decision tree algorithm for classification problem. It can be used with univariate or linear combination splits. One of its features is that its method for attribute selection has negligible bias. If all attributes are uninformative with respect to the class attribute, then each attribute has approximately the same chance of being selected to split a node.[15]

Commercial Version 5.0 (C5.0)
The supervised learning algorithm (C5.0) is used to generate the set of rules from the data. As such, it employs a divide-and-conquer approach, rather than a separate-and-conquer one. After inducing the trees, a post-processing step produces the rules. It uses a pruning strategy by which a branch is pruned when the error is one standard error of the existing errors adjusted for the correction of continuity.[16, 17]

Exhaustive Chi-Squared Automatic Interaction Detector (CHAID)
The CHAID is based on the chi-square test of association. It constructs a decision tree by repeatedly splitting subsets of the space into two or more child nodes, beginning with the entire data set.[18] In order to determine the best split at any node, any pair of categories of the predictor variables is merged until there is no statistically significant difference within the pair with respect to the target variable. It naturally deals with interactions between the independent variables that are directly available from an examination of the tree. The final nodes identify subgroups defined by different sets of independent variables.[19]

Logistic Regression (LR)
Logistic regression is a model for binomially distributed dependent variables. It is a generalized linear model that uses the logit as its link function. Binomial (or binary) logistic regression is used when the dependent variable has a dichotomy and the independent variables are of any type. Logistic regression applies a maximum estimation of likelihood after transforming the dependent into a logit variable (the natural log of the odds of the dependent occurring). In this way, logistic regression estimates the probability of occurrence of a certain event. Using logistic regression has advantages over linear discriminant analysis, including coefficients interpretable (odd-ratios), standard errors calculable (but only asymptotic), and maximum likelihood instead of least squares, but at the risk of non-convergence of the algorithm.[20]

Ensemble Model
This study also takes into account the ensemble model to evaluate the performance of the color-based models with regard to construction object detection. An ensemble model involves training multiple single classifiers (so-called member classifiers) and subsequently combining them to form a single classification model. Hence, the classification performance of an ensemble model strongly depends on the performance of the base classification models.[21] Ranking the performance of the above seven classification models enables selection of the best performing models, which can be combined into a single ensemble model. This approach often yields more accurate classification results compared to a single classification model because it aggregates the benefits of multiple models.[22] The ensemble model combined in this manner generally performs at least as well as the best individual models, and often better. In this study, the three best-performing single models were combined to produce an ensemble model. As a way of combining different single models' classification results to obtain the overall classification result, the highest confidence wins was selected in this study.

Model Validation Methods
For evaluation of each of these eight classification models' classification performance, *k*-fold cross validation was used. This method is known for its tendency to minimize bias and variance among all validation methods, including the leave-one-out method[23]. Extensive studies on numerous data sets with different classification models have demonstrated that 10 folds are optimal in terms of computation and estimation of error, and there is theoretical evidence backing this up[24]. Thus, a 10-fold cross validation approach was used to assess the performance of eight classification models.

In each 10-fold cross validation, the data was split into 10 approximately equal folds, with each in turn being used for testing and the remainder being used for training. That is, 9 of the 10 folds are used for training and the *k*th holds out folds for testing, and repeats the procedure for *k* = 1, 2, ..., 10, so that by the end, every instance has been used exactly once for testing. The cross validation estimate of overall performance is then calculated by simply averaging the 10 individual performance evaluation measures for 10-fold cross validation performance.

Performance Evaluation Measures
Various approaches have been suggested for evaluating the performance of classification models. To determine whether a method of color-based object detection is "useful" or "not useful," a partial measure of the performance is inadequate, because evaluation on performance measures needs to be comprehensive. For this reason, performance evaluation measures employed in this investigation consist of six measures: accuracy, precision, sensitivity, specificity, area under the receiver operating characteristics curve (AUC), and overall

average performance score (S). The first five measures are extensively used to evaluate the performance of the classification models.[25, 26, 27] They can be calculated by computing the number of correctly predicted pixels of the object of interest (true positives, TP), the number of correctly predicted pixels that belong to the background objects (true negatives, TN), the number of pixels that were incorrectly assigned to the object of interest (false positives, FP), and the number of pixels of the object of interest that were incorrectly assigned to the background objects (false negatives, FN).[28] The last measure is derived to evaluate the overall performance of the classification models by compounding the effect of the first five measures. It is calculated by averaging the values of the first five measures, with the value of S being positively related to the effectiveness of the overall evaluation measures. The first five measures are determined by the following equations:

$$Accuracy = \left(\frac{TP + TN}{TP + TN + FP + FN} \right)$$

$$Precision = \left(\frac{TP}{TP + FP} \right)$$

$$Sensitivity = \left(\frac{TP}{TP + FN} \right)$$

$$Specificity = \left(\frac{TN}{FP + TN} \right)$$

$$AUC = \frac{1}{2} \left[\left(\frac{TP}{TP + FN} \right) + \left(\frac{TN}{FP + TN} \right) \right]$$

EXPERIMENTAL RESULTS AND DISCUSSION

For the color of an object to be a useful feature in its detection, it should be verified that the color of the object alone is sufficient for distinguishing it from others. Table 2 lists the statistical results regarding the effectiveness and usefulness of color for the detection of concrete, steel, and wood. The results are generated using 10-fold cross validation by seven single classification models and one ensemble model for three data sets for concrete, steel, and wood detection, respectively.

Regardless of the type of classification model, the results imply that concrete, steel, and wood detection can be achieved using color information alone, ranging from 90.40% to 96.24% in accuracy, 95.66% to 98.16% in precision, 83.86% to 94.24% in sensitivity, 96.08% to 98.27% in specificity, and 90.40% to 96.24% in AUC. These results strongly indicate that color information is very effective and useful in construction object detection for objects made of concrete, steel, and wood. To further demonstrate the classification capabilities of different classification models, this study investigated and compared the effectiveness of seven single classification models and an ensemble model in concrete, steel, and wood detection, respectively. The ensemble model demonstrated the best performance for all three data sets in accuracy, precision, sensitivity, specificity, AUC, and S, except for precision in wood data set and specificity in concrete, steel, and wood data sets (see Table 2). SVM followed the ensemble model.

Table 2 draws the following conclusions about the performance of the ensemble model. On the one hand, although the differences in performance be-

Table 2. Classification performance for each data set

Data Set	Methods	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	AUC (%)	S
Concrete	ANN	90.20	93.48	86.45	93.94	90.20	90.85
	SVM	92.24	95.58	88.58	95.90	92.24	92.91
	CART	90.58	95.59	85.09	**96.08**	90.58	91.58
	QUEST	89.15	92.67	85.04	93.27	89.15	89.85
	C5.0	90.04	91.59	88.17	91.91	90.04	90.35
	CHAID	89.44	90.42	88.23	90.65	89.44	89.64
	LR	89.50	91.12	87.52	91.47	89.50	89.82
	Ensemble	**92.79**	**95.66**	**89.66**	95.93	**92.79**	**93.37**
Steel	ANN	94.03	97.64	90.25	97.82	94.03	94.76
	SVM	94.54	98.00	90.93	98.15	94.54	95.23
	CART	92.16	95.45	88.69	95.63	92.16	92.82
	QUEST	89.97	93.42	86.00	93.94	89.97	90.66
	C5.0	92.57	95.95	88.88	96.25	92.57	93.24
	CHAID	92.52	98.04	86.77	**98.27**	92.52	93.62
	LR	83.88	85.93	81.04	86.73	83.88	84.29
	Ensemble	**96.24**	**98.16**	**94.24**	98.24	**96.24**	**96.62**
Wood	ANN	89.26	95.69	82.23	96.29	89.26	90.55
	SVM	89.64	96.62	82.15	97.13	89.64	91.04
	CART	86.48	97.44	74.93	98.03	86.48	88.67
	QUEST	82.86	97.03	67.79	97.93	82.86	85.69
	C5.0	89.61	95.26	83.38	95.85	89.61	90.74
	CHAID	87.66	92.19	82.29	93.02	87.66	88.56
	LR	88.05	**97.65**	77.97	**98.12**	88.05	89.97
	Ensemble	**90.40**	96.49	**83.86**	96.95	**90.40**	91.62

tween the ensemble models and single classification models are not large, the ensemble models for all three data sets generate more successful results than single classification models do. It can be seen that the performances of the ensemble models are better than any other based single classification models, especially the accuracy, sensitivity, and AUC values, which increase slightly. The comparison results confirm that the ensemble model can effectively deal with construction object detection using color information alone and improve classification performance better than single classification models.

CONCLUSIONS

In this study, we attempted to discover whether color information of various objects alone is useful for their detection in photographs acquired from construction sites. For this reason, the effectiveness and usefulness of the color information in construction object detection, such as concrete, steel, and wood, were evaluated using the seven single classification models and one ensemble model for collected comprehensive data sets. Relevant to this debate, our results can be interpreted as evidence for the idea that color is an intrinsic component of the representation of construction objects. This conclusion is based on the results obtained in comparison with single classification models. Color information in handling concrete, steel, and wood detection provides accuracy between 90.40% to 96.24%, precision between 95.66% to 98.16%, sensitivity between 83.86% to 94.24%, specificity between 96.08% to 98.27%, and AUC between 90.40% to 96.24%. Among eight classification models, the overall performance of the ensemble model that combines the three best performing single models was compared to seven single classification models, which were slightly better than those of the single classification models. In summary, the results clearly indicate that construction object detection in construction photographs using color information alone is quite effective.

The results by the proposed method would be reliable for use as an essential source of input for various applications. These include, but are not limited to, materials tracking, automated control of heavy equipment, progress monitoring, quality control, and the generation of 3D as-built models. For future work, the evaluation will be expanded to explore the ability of the color information with regard to a number of types of construction objects, including curtain walls and masonry, leading to a practical implementation. In addition, we believe that further improvements are possible regarding the efficiency and performance of construction object detection using color information by investigating other ensemble models, such as bagging, boosting, and random forest.

References
1. Neto, J.A., Arditi, D., and Evens, M.W., "Using Colors to Detect Structural Components in Digital Pictures", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 17(1), pp. 61–67, 2002.
2. Zhu, Z., and Brilakis, I., "Parameter Optimization for Automated Concrete Detection in Image Data", *Automation in Construction*, Vol. 19(7), pp. 944–953, 2010.
3. Bascle, B., Bernier, O., and Lemaire, V., "Illumination-Invariant Color Image Correction", *Proceedings of the International Workshop on Intelligent Computing in Pattern Analysis/Synthesis*, Xi'an, China, 2006.
4. Doyle, S., Feldman, M., Tomaszewski, J., and Madabhushi, A., "A Boosted Bayesian Multi-Resolution Classifier for Prostate Cancer Detection from Digitized Needle Biopsies", *Transactions on Biomedical Engineering*, Vol. PP(99), pp. 1–14, 2010.
5. Cheng, H.D., Jiang, X.H., Sun, Y., and Wang, J., "Color Image Segmentation: Advances and Prospects", *Pattern Recognition*, Vol. 34(12), pp. 2259–2281, 2001.
6. Ripley. B.D., *Pattern Recognition and Neural Networks*, Cambridge: Cambridge University Press, 1996.
7. Lippmann, R.P., "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine*, Vol. 4(2), pp. 4–22, 1987.
8. Venables, W.N., and Ripley, B.D., *Modern Applied Statistics with S.*, 4th ed., New York: Springer, 2002.
9. Cortes, C., and Vapnik, V., "Support-Vector Network", *Mach Learn*, Vol. 20, pp. 1–25, 1995.
10. Mayer, D., "Support Vector Machines", *R News*, Vol. 1(3), pp. 23–26, 2001.
11. Vapnik, V., *The Support Vector Method of Function Estimation*, In: Nonlinear Modeling: Advanced Black-Box Techniques, 1998.
12. Vladimir, V.N., and Vapnik, V., *The Nature of Statistical Learning Theory*, New York: Springer, 1995.
13. Chandra, D.K., Ravi, V., and Bose, I., "Failure Prediction of Dotcom Companies Using Hybrid Intelligent Techniques", *Expert Systems with Applications*, Vol. 36(3), 48304837, 2009.
14. Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J., *Classification and Regression Trees*, Montery: Wadsworth and Brooks/Cole, 1984.
15. Loh, W.-Y., and Shih, Y.-S., "Split Selection Methods for Classification Trees", *Statistica Sinica*, Vol. 7, pp. 815–840, 1997.
16. Chae, Y.M., Ho, S.H., Cho, K.W., Lee, D.H., and Ji, S.H., "Data Mining Approach to Policy Analysis in a Health Insurance Domain", *International Journal of*

Medical Informatics, Vol. 62(2–3), pp. 103–111, 2001.

17. Quinlan, J.R., *C5.0 and See 5: Illustrative Examples*, RuleQuest Research: <http://www.rulequest.com>, 1997.

18. Michael, J.A., and Gordon, S.L., *Data Mining Technique: For Marketing*, Sales and Customer Support, New York: Wiley, 1997.

19. Magidson, J., and SPSS Inc., *SPSS for Windows CHAID Release 6.0*, Chicago: SPSS Inc., 1993.

20. Lehmann, C., Koenig, T., Jelic, V., Prichep, L., John, R.E., Wahlund, L.-O., Dodge, Y., and Dierks, T., "Application and Comparison of Classification Algorithms for Recognition of Alzheimer's Disease in Electrical Brain Activity (EEG)", *Journal of Neuroscience Methods*, Vol. 161(2), pp. 342–350, 2007.

21. Zhang, Y., Zhang, D., Mi, G., Ma, D., Li, G., Guo, Y., Li, M., and Zhu, M., "Using Ensemble Methods to Deal with Imbalanced Data in Predicting Protein-Protein Interactions", *Computational Biology and Chemistry*, Vol. 36, pp. 36–41, 2012.

22. Alexandre, L.A., Campilho, A.C., and Kamel, M., "On Combining Classifiers Using Sum and Product Rules", *Pattern Recognition Letters*, Vol. 22, pp. 1283–1289, 2001.

23. Kohavi, R., "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", *Proceedings of the International Joint Conference on Artificial Intelligence*, Montréal, Québec, Canada, 1995.

24. Hastie, T., Tibshirani, R., and Friedman, J., *The Elements of Statistical Learning*, 1st ed., New York: Springer, 2001.

25. Ferri, C., Hernández-Orallo, J., and Modroiu, R., "An Experimental Comparison of Performance Measures for Classification", *Pattern Recognition Letters*, Vol. 30(1), pp. 27–38, 2009.

26. Horng, M.-H., "Performance Evaluation of Multiple Classification of the Ultrasonic Supraspinatus Images by Using ML, RBFNN and SVM Classifiers", *Expert Systems with Applications*, Vol. 37(6), pp. 4146–4155, 2010.

27. Kim, Y.S., "Performance Evaluation for Classification Methods: A Comparative Simulation Study", *Expert Systems with Applications*, Vol. 37(3), 2292–2306, 2010.

28. Sokolova, M., and Lapalme, G., "A Systematic Analysis of Performance Measures for Classification Tasks", *Information Processing & Management*, Vol. 45(4), pp. 427–437, 2009.