Automatic classification of melanocytic skin tumors based on hyperparameters optimized by cross-validation using support vector machines

Gokkan, Ozan, Tozburun, Serhat

Ozan Gokkan, Serhat Tozburun, "Automatic classification of melanocytic skin tumors based on hyperparameters optimized by cross-validation using support vector machines," Proc. SPIE 11211, Photonics in Dermatology and Plastic Surgery 2020, 112110B (19 February 2020); doi: 10.1117/12.2542161

SPIE.

Event: SPIE BiOS, 2020, San Francisco, California, United States

Automatic classification of melanocytic skin tumors based on hyperparameters optimized by cross-validation using support vector machines

Ozan Gokkan*[a,b] and Serhat Tozburun[a,c,d]
[a]Izmir Biomedicine and Genome Center, Balcova, 35330, Izmir, TURKEY
[b]Dept. of Biomedical Technologies, Ege University, Bornova, 35100, Izmir, TURKEY
[c]Izmir International Biomedicine and Genome Institute, Dokuz Eylul Univ./Izmir, TURKEY
[d]Dept. of Biophysics, Faculty of Medicine, Dokuz Eylul Univ./Izmir, TURKEY

ABSTRACT

Melanocytic lesions may occur in various areas of the skin and may eventually develop into malignant tissue types as a result of abnormal tissue growth. Although the gold standard for the diagnosis of melanoma is still a histopathological examination, dermatologists often use dermoscopic examination in their routine practice to reduce unnecessary excisions or to prevent misdiagnosis of clinically suspected melanocytic lesions. However, dermoscopic examinations may require special training and experience. Furthermore, even among experts, different evaluation results may occur. For these reasons, image processing and artificial intelligence application studies are performed on dermoscopic images based on information technologies developed in recent years. This study investigated the automatic classification of superficial spreading melanoma and nevocellular nevus using support vector machines. A publicly available and histopathologically verified MED-NODE data set (70 superficial spreading melanomas and 100 nevocellular naevi) was used. For the classification task, first, the energy distributions (power spectral densities) of each image in the spectral domain were obtained. Second, gray-level co-occurrence matrices were created, and the textural features of the matrices were extracted. Finally, the learning model was developed with these features as input for classification. Support vector machines were trained using validation methods, including holdout validation and stratified cross-validation. The hyperparameters were optimized using the regularization factor of 10, the radial basis kernel function, and the gamma factor of 0.0098. Using 10-fold cross-validation, we achieved a mean accuracy of 98.9% (+/- 0.01 standard deviation), 99.4% sensitivity, and 97.5% specificity.

Keywords: Melanocytic skin, superficial spreading melanoma, nevocellular nevus, support vector machines, machine learning, binary classification, feature extraction, hyperparameter optimization

1. INTRODUCTION

The skin is the largest organ in the human body and has many tasks, such as protecting the body from injuries, microorganisms, harmful ultraviolet rays, as well as regulating body temperature, allowing heat and cold sensation. Melanocytic lesions that occur in various regions of the skin, which may undergo change and development over time, have the potential to develop into malignant tissue types as a result of abnormal tissue growth [1]. Although the gold standard for the diagnosis of melanoma is still histopathologic examination [1], dermatologists often use dermoscopy devices in their routine practice to reduce unnecessary excisions or to clinically avoid possible melanoma in lesions [2]. For example, in clinical practice and studies, a risk factor occurs on skin moles based on ABCDE criteria (A: Asymmetry, B: Border irregularity, C: Color, D: Diameter, E: Evolving), 3-point check-list [3], and 7-point check-list. To examine such a risk factor, it is essential to consider that dermoscopic examinations require specialized training and experience, because, in some critical cases, incorrect or inadequate results may occur. Image processing [2] and customized artificial intelligence application [2] development studies on dermoscopic images have recently attracted interest. However, the generalization of the algorithmic processes, including outer edge segmentation, has been challenging [2].

*ozangokkan@gmail.com; phone 90 232 299-5101; fax 90 232 277-6353 tozburunlab.com

Photonics in Dermatology and Plastic Surgery 2020, edited by Bernard Choi, Haishan Zeng, Proc. of SPIE Vol. 11211, 112110B · © 2020 SPIE
CCC code: 1605-7422/20/$21 · doi: 10.1117/12.2542161

Besides, although deep learning-based convolutional neural network architectures can be performed by convolutional filtering-based feature extraction of the skin lesion, large dataset size are required for the development of the artificial neural network [4].

The objective of this study is to investigate the technical performance of the optimized support vector machines on reducing computation time, algorithm complexity, and number of required data set. In this study, the tissue model was superficial spreading melanoma and nevocellular nevus. Validation methods, including holdout and 10-fold stratified cross-validation, were utilized in the training of the support vector machines [5].

2. MATERIALS AND METHODS

In this study, an open-access database, MED-NODE [6], was used to train and test our model, including histopathologically verified images of superficial spreading melanoma (SSM) and nevocellular nevus (NN), respectively. The data set comprised a total of 170 images, including 70 SSM and 100 NN. A non-dermoscopic camera acquired images (Nikon D3 with a Nikkor 2.8/105 mm micro-lens). The distance between the micro-lens and the lesion was ~33 cm, as described in Ref. [6]. The model was developed on HP Z4 workstation with W-2145 Intel Xeon processor.

Figure 1 demonstrates a flowchart of the study. Feature extraction from images was performed in two main steps to obtain a binary classification between melanoma and nevus: First, the power spectral density (PSD) (i.e., the squared magnitude of the two-dimensional Fourier transform) of each image was estimated to analyze the pixel-based energy distribution in the spectral domain. Second, a 2D gray-level co-occurrence matrix (GLCM) of calculated power spectral densities was determined to establish a relationship between patterns of neighboring pixels in an image at a distance.

Figures 2 and 3 show representative images and corresponding power spectral density calculations for superficial spreading melanoma and nevocellular nevus melanoma, respectively.

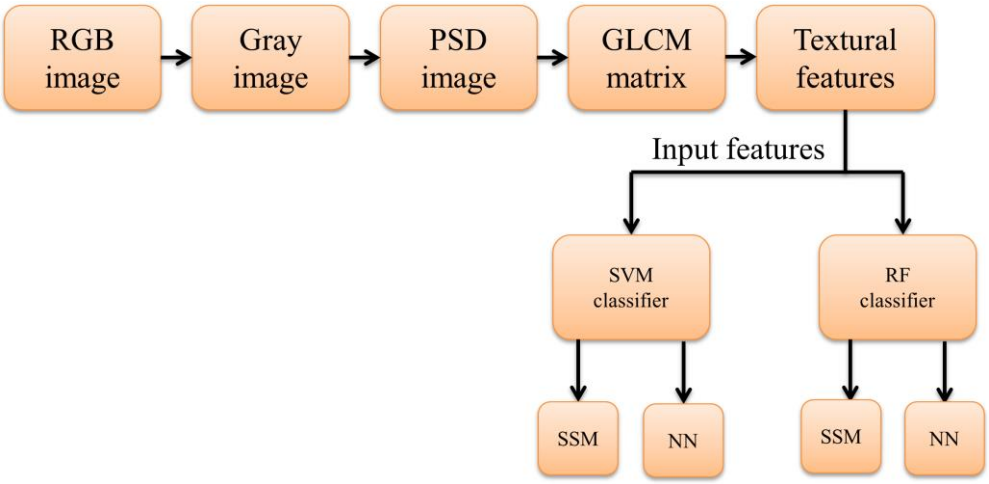

Figure 1. A flowchart of the classification model used in the study.

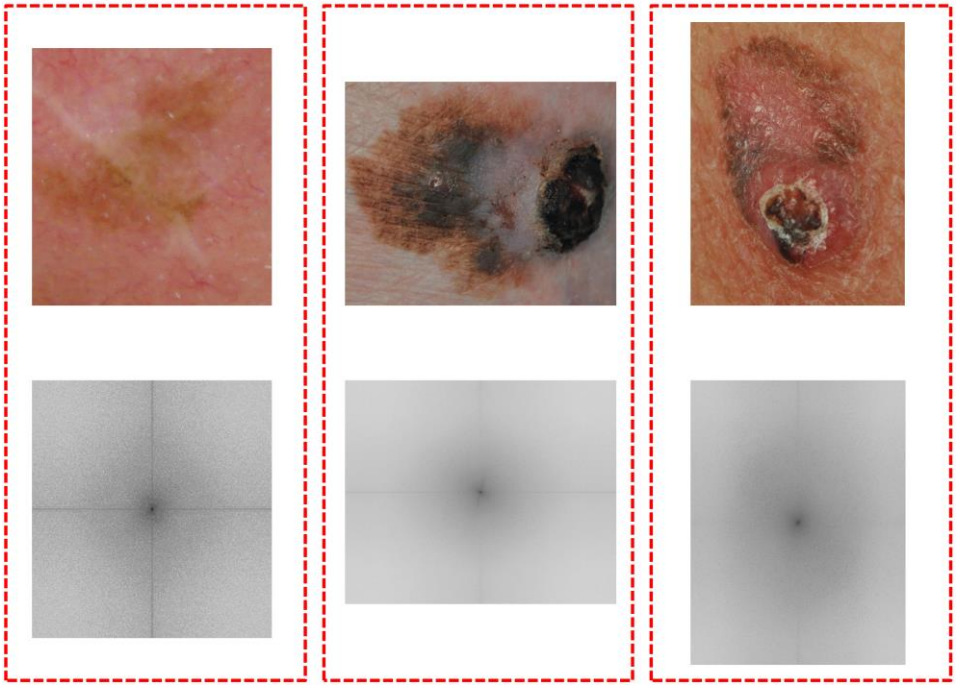

Figure 2. Top row: Superficial spreading melanoma image samples. Bottom row: Corresponding power spectral density calculations of the images.

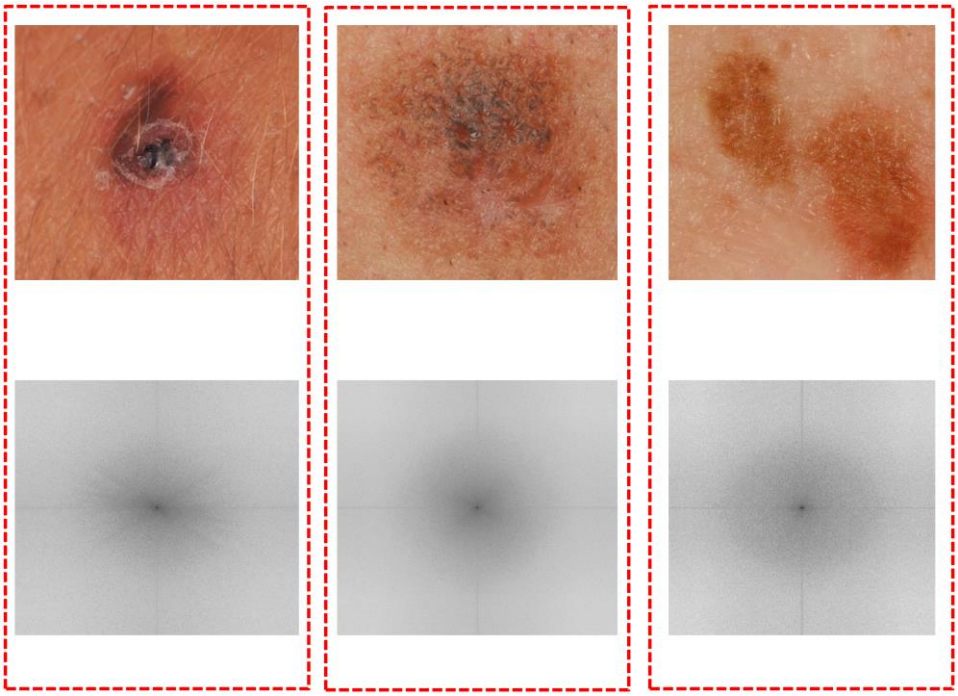

Figure 3. Top row: Nevocellular nevus melanoma image samples. Bottom row: Corresponding power spectral density calculations of the images.

As shown in Figure 4, 2D spectral image-based pattern analysis included both the homogeneous and heterogeneous propagation effects of the lesion, the neovascularization effect [7], arborizing vessel effect [8,10], color tone distribution, and lesion outer (border) irregularity. From a PSD image, the GLCM matrix [9] was determined by evaluating the distances (1, 2, 3, 4) between the reference pixel and neighboring pixels at 0, 45, 90, and 135 degrees for each distance. In the GLCM matrix, 14 different features were extracted for each distance, so that obtaining a 4x14 feature group for four degrees of orientation. These features included angular second moment, contrast, correlation, variance, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation, and maximal correlation coefficient [9]. During the training, two different types of cross-validation methods, (i.e., stratified method and hold-out method) applied, and the hyperparameters, including regularization factor (C = 10), radial basis kernel function (RBF) and gamma factor (gamma = 0.0098) were optimized.

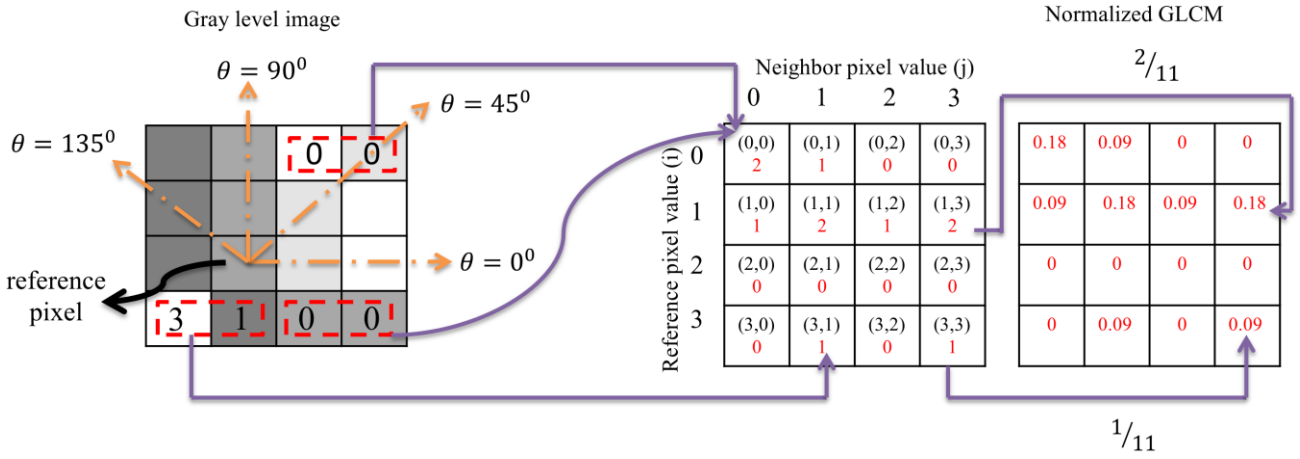

Figure 4. A representative gray-level co-occurrence matrix (GLCM) calculation for distance of 1 at 0°.

The following equations were used to extract the features of each grayscale spatial-dependence matrix [8]. As a notation, *p(i,j)* is a normalized GLCM, shown in Figure 4. $p_x(i)$ was the value of i^{th} coordinate in the probability matrix (normalized GLCM), and calculated by summation of the row of *p(i,j)*. $p_y(j)$ was the value of j^{th} coordinate in the probability matrix (normalized GLCM), and calculated by summation of the column of *p(i,j)*. *N* was the distinct number of gray tone values.

$$p_x(i) = \sum_{j=0}^{N} p(i,j), p_y(j) = \sum_{i=0}^{N} p(i,j) \tag{1}$$

$$p_{x+y}(k) = \sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j), i+j=k, k=2,3,4,...2N \tag{2}$$

$$p_{x-y}(k) = \sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j), |i-j|=k, k=0,1,2,...,N-1 \tag{3}$$

Angular second moment : $$f_1 = \sum_{i}^{N}\sum_{j}^{N}\{p(i,j)\}^2 \tag{4}$$

Contrast : $$f_2 = \sum_{n=0}^{N-1} n^2 \left\{ \sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j), |i-j|=n \right\} \tag{5}$$

Correlation : $f_3 = \dfrac{\sum_i \sum_j (ij)\, p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$ (6)

where $\mu_x, \mu_y, \sigma_x, \sigma_y$ are the means and standard deviations of p_x, p_y.

Sum of squares (variance) : $f_4 = \sum_i^N \sum_j^N (i-\mu)^2 p(i,j)$ (7)

Inverse difference moment : $f_5 = \sum_i \sum_j \dfrac{1}{1+(i-j)^2} p(i,j)$ (8)

Sum average : $f_6 = \sum_{i=2}^{2N} i(p_{x+y}(i))$ (9)

Sum variance : $f_7 = \sum_{i=2}^{2N} (i - f_8)^2 p_{x+y}(i)$ (10)

Sum entropy : $f_8 = -\sum_{i=2}^{2N} p_{x+y}(i) \log(p_{x+y}(i))$ (11)

Entropy : $f_9 = -\sum_i^N \sum_j^N p(i,j) \log(p(i,j))$ (12)

Difference variance : $f_{10} = \text{variance}(p_{x+y})$ (13)

Difference entropy : $f_{11} = -\sum_{i=2}^{2N} p_{x-y}(i) \log(p_{x-y}(i))$ (14)

Information measures of correlation : $f_{12} = \dfrac{HXY - HXY1}{\max(HX, HY)}$ (15)

$$f_{13} = (1 - \exp(-2(HXY2 - HXY)))^{1/2}, HXY = -\sum_i^N \sum_j^N p(i,j) \log(p(i,j)) \quad (16)$$

where HX and HY are entropies of p_x and p_y, respectively.

$$HXY1 = -\sum_i^N \sum_j^N p(i,j) \log(p_x(i) p_y(j)) = -\sum_i^N p_x(i) p_y(j) \log(p_x(i) p_y(j)) \quad (17)$$

$f_{14} = (\text{second largest eigenvalue of } Q)^{1/2}$, where $Q(i,j) = \sum_k \dfrac{p(i,k)\, p(j,k)}{p_x(i)\, p_y(k)}$ (18)

3. RESULTS

For the binary classification of skin tumors, we used two different classifiers, the support vector machines classifier and the random forest classifier [11]. Figure 5 shows the receiver operating characteristic (ROC) performance of the support vector machines classifier. In the 10-fold cross-correlation analysis, the 2nd fold provided the largest area under the curve, as shown in Figure 6.

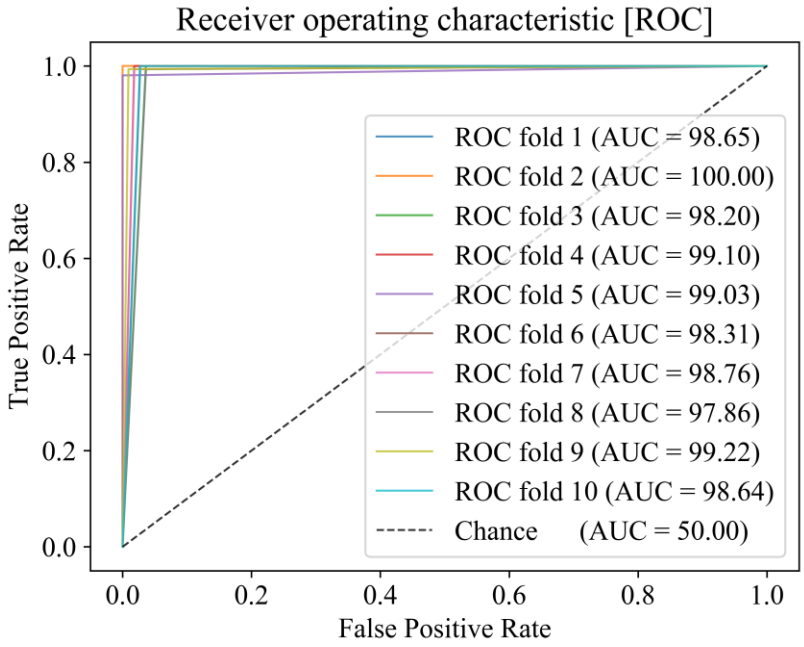

Figure 5. Receiver operating characteristic (ROC) performance of the support vector machines classifier.

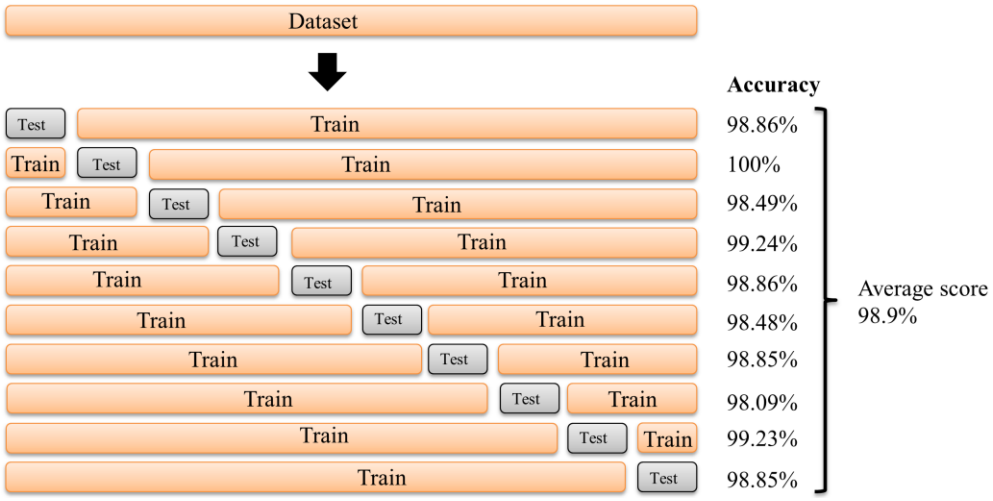

Figure 6. 10-fold stratified cross-validation results for support vector machines classifier.

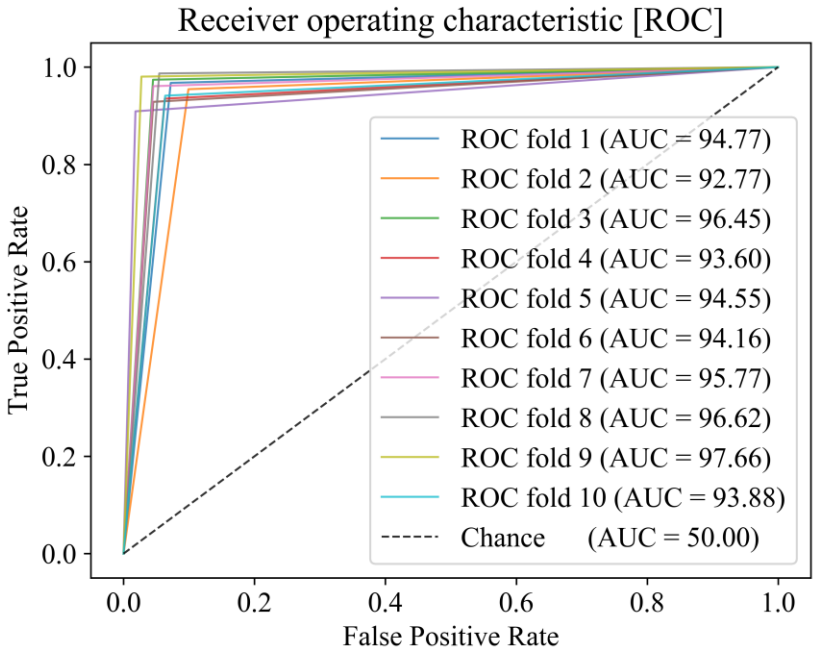

Figure 5. Receiver operating characteristic (ROC) performance of the random forest classifier.

Figure 7 shows the receiver operating characteristic (ROC) of the random forest classifier. In the 10-fold cross-correlation analysis, the 9th fold provided the largest area under the curve.

Table 1 summarize the calculated results including mean accuracy, mean sensitivity, and mean specificity for both classifiers.

Table 1. A performance summary of the support vector machines classifier and the random forest classifier.

Classifiers	meanAccuracy	meanSensitivity	meanSpecificity
Support vector machines	**98.9%**	**99.4%**	**97.5%**
Random forest classifier	94.9%	94.4%	95.4%

4. DISCUSSION AND CONCLUSIONS

Melanocytic lesions may occur in various areas of the skin and may change over time. They also have the potential to develop into malignant tissue types as a result of abnormal tissue growth. Dermatologists often use dermoscopic examinations in their routine practice to reduce unnecessary excisions or to prevent misdiagnosis of clinically suspected melanocytic lesions. However, dermoscopic examinations are considerably dependent on specific training and experience. For this reason, different evaluation results may arise among experts in some cases.

This study investigated the automatic classification of superficial spreading melanoma and nevocellular nevus using support vector machines. Support vector machines were trained using validation methods including holdout and 10-fold stratified cross-validation. Selecting a small value of the regulation factor would result in a large margin or vice versa. So, defining a hyperplane with optimum margin was strongly related to the regularization factor and gamma factor that used in radial basis kernel function. As the gamma in gaussian radial basis kernel function increased, hyperplane or decision boundary would be more wiggling, which might result in overfitting. Therefore, we chose to use the radial basis kernel function because of nonlinear property entries.

Overall, the hyperparameters were optimized using the regularization factor of 10, the radial basis kernel function and the gamma factor of 0.0098. 98.9% mean accuracy, 99.4% sensitivity and 97.5% specificity were obtained.

ACKNOWLEDGMENTS

This work was partially supported by TUBITAK under the project number 116E815. Serhat Tozburun was supported by the Turkish Academy of Sciences Young Scientists Award Programme (2018 TUBA GEBIP).

REFERENCES

[1] D. Ginat, N. Cipriani, C. Barnes *et al.*, "Malignant Skin Neoplasms and Associated Conditions", Springer, (2019).
[2] U. Jamil, S. Khalid, M. Akram *et al.*, "Melanocytic and nevus lesion detection from diseased dermoscopic images using fuzzy and wavelet techniques," Soft Computing, 22, (2018).
[3] P. Soyer, G. Argenziano, I. Zalaudek *et al.*, "Three-Point Checklist of Dermoscopy," Dermatology (Basel, Switzerland), 208, 27-31 (2004).
[4] Z. Al Nazi, and T. A. Abir, "Automatic Skin Lesion Segmentation and Melanoma Detection: Transfer Learning approach with U-Net and DCNN-SVM", IJCCI, (2018).
[5] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, 2, 121-167 (1998).
[6] I. Giotis, N. Molders, S. Land *et al.*, "MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images," Expert Systems with Applications, 42(19), 6578-6585 (2015).
[7] Y. Togawa, "Review of vasculature visualized on dermoscopy," The Journal of Dermatology, 44, 525-532 (2017).
[8] H. Jin, M.-Y. Yang, J.-M. Kim *et al.*, "Arborizing Vessels on Dermoscopy in Various Skin Diseases Other Than Basal Cell Carcinoma," Annals of Dermatology, 29, 288 (2017).
[9] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," IEEE Transactions on Systems, Man, and Cybernetics, SMC-3(6), 610-621 (1973).
[10] N. Marghoob, K. Liopyris, and N. Jaimes, "Dermoscopy: A Review of the Structures That Facilitate Melanoma Detection," The Journal of the American Osteopathic Association, 119, 380-390 (2019).
[11] L. Breiman, "Random Forests," Machine Learning, 45(1), 5-32 (2001).