

Melanoma Classification Using Feature Extraction Methods and Machine Learning Approaches

Sihui He^{1+*}, Zheyang Huang²⁺, and Xinjie Zhong³

¹Faculty of Science, Western University, London, N6A 3K7, Canada

²Faculty of Engineering, The University of Sydney, Sydney, NSW 2006, Australia

³School of Electrical Engineering, Zhejiang University, Hangzhou, 30058, China

⁺They are both first authors

^{*}Corresponding author: she334@uwo.ca

Abstract

Due to the unbalanced melanoma data and the complexity and resolution of the melanoma image backgrounds, classification of the melanoma regions is very challenging. In this paper, EffNet B5 models with different augmentation methods and machine learning models that use feature vectors from feature extraction methods are proposed to solve the classification task. GLCM, LBP, SFTA, and ResNet18 are chosen as the feature extraction methods in this study, while SVM, Random Forest, and XGBoost are chosen to be the machine learning classification models. In this work, the most effective feature extraction algorithm and the classification algorithm with the best AUC scores are determined.

Keywords

Melanoma, Feature Extraction, Classification, Machine Learning

1. Introduction

Melanoma occurs when DNA damage from burning or tanning due to UV radiation triggers changes (or mutations) in the melanocytes, resulting in uncontrolled cellular growth [1]. It is a widespread cancer with 6% of people who are estimated to get melanoma in 2021; Also, an estimated 7,180 people (4,600 men and 2,580 women) will die of melanoma in the U.S. in 2021 [2]. However, it can be cured if we can find it early, with a five-year survival rate estimated to 99%. Therefore, using AI techniques to help people detect melanoma earlier is essential for saving millions of people's lives.

In this paper, our goal is to conduct classification tasks on the melanoma data set 2020 SIIM-ISIC Melanoma Classification [3] and 2019 SIIM-ISIC Melanoma Classification [4-6]. We aim to find the augmentation methods which could improve the performance of EffNets for the binary classification task by training EffNets along with both position and color augmentation methods, such as flipping, hue, saturation.

Also, the machine learning algorithms SVM, Random Forest, and XGBoost along with four different feature extraction algorithms are proposed to implement. The four feature extraction algorithms Gray-Level Co-Occurrence Matrix (GLCM), Local Binary Patterns (LBP), Segmentation Based Fractal Texture Analysis (SFTA), and Residual Neural Network (ResNet18) are also to be used in this paper. By combining the machine learning algorithms with the featured extraction algorithms that are proposed previously, the combination with the best performance on the classification of the melanoma data set and the machine learning algorithm with the highest positive detection rate on the positive cases are both successfully determined.

2. Related Works

Before feature extraction and classification, the image dataset should be pre-treated. Mikołajczyk [7] reviewed the classical data augmentation methods, such as picture rotating, cropping, zooming, histogram-based methods, as well as the deep learning methods, such as style transfer using Generative Adversarial Networks (GANs). The classical methods are still popular and powerful. However, they found that combining both classical methods and style transfer would have a better performance. Style Transfer can generate new images with high-level image synthesis and manipulation. However, it also

has some disadvantages. The style transfer is based on texture and color transfer, so it may limit those images where the structures are essential.

Texture features are thought to be important to the classification problems, as well as the color features (RGB, HSV). Scientists also use dermoscopy features such as ABCD rules (Atypical, Border, Color, and Diameter). Gray level co-occurrence Matrix (GLCM) is a popular method to extract texture features [8-11].

Wu Xuelian, et al use 14 features extracted by GLCM and 16 features extracted by CS-LBP, and they have 84.7615% accuracy using the SVM classifier [10]. The paper “Texture and color feature extraction for classification of melanoma using SVM” shows that only using texture features may have bad performance on classification. The classification results using the same classifier (SVM) on texture features, RGB and HSV color space features, OPP features, and texture + RGB features are 76%, 92%, 89%, 93%. Also, Agnieszka, et al [9] use SVM as classifier and GLCM + ABCDs as feature extraction, and they have a score of 92%.

Researchers also found that using less but important features perform as well as the original features [12, 13]. Wieslaw Paja [12] uses original features (13) with 19.07% error rate. They remove four features and find that the accuracy value is almost the same. They remove half of the features, and the error rate is 22.40%. It shows that the performance will not be worse too much if we remove those less important features. Besides, Suleiman Mustafa [13] extracts shape, color, and geometry features and uses the SVM RBF classifier. They get 86.87% accuracy. After that, they use sequential backward selection (SBS) to select features. Finally, they can reduce features number to six with a performance as good as the original features [9, 11]. Using Principal Component Analysis (PCA) to reduce dimension also has a good performance.

A deep learning way on feature extraction shows a better result on traditional methods [14]. Aimin Yang compares the classical feature extraction algorithms and 2 CNN structures. The results of LBP, LBP_C, LBP_M, LBP_S, Xception and DenseNet are separately 97.12%, 98.20%, 93.21%, 92.01%, 99.01%, 99.16%. Xception and DenseNet are CNN structures. Moreover, they perform better than other algorithms [15].

Besides the different methods of feature extraction, different classifiers also have different performances on the same dataset [11, 15, 16]. Arslan Javaid uses the same features on SVM (Medium Gaussian), Random Forest, and Quadratic Discriminant and gets 88.17%, 93.89% and 90.84%. Şaban Öztürk [16] compares different feature extraction methods (GLCM, LBM, LBGLCM, GLRLM, SFTA) and different classifiers (SVM, KNN, LDA, Boosted Tree). They conclude that SFTA is better (> 92%) than other feature extraction methods, and LBP is worse (< 90%). Furthermore, SFTA with Boosted Tree gets the best performance of 94.3%. What is more, Fábio Perez [17] evaluates 9 CNNs architectures in 5 sets of splits created on the ISIC Challenge 2017 dataset, with 3 repeated measures. The 135 models show that the correlation between the performance of CNN architectures on ImageNet and their performance on target tasks seems smaller than other researchers thought. Also, using multiple models is better than a single model. Although using the validation set on choosing models is better, picking high-performance models at random also has competitiveness.

3. Methodology

3.1. Image Data Set and Data Pre-Processing

The data sets 2020 SIIM-ISIC Melanoma Classification [3] and 2019 SIIM-ISIC Melanoma Classification [4-6] are available for general audiences on the ISIC website. The training data sets and the training ground truth tables are downloaded from the ISIC website. The ground truth tables contain the information of standard lesion diagnosis for every image in the data set.

The training sets from the ISIC website contain 58,457 dermoscopic images in total, with 5,106 positive cases and 53,351 negative cases. Each image comes from the dermoscopic image of a patient identified with a distinct patient id. Histopathology has been used to confirm all malignant diagnoses, whereas expert agreement, longitudinal follow-up, and histopathology have all been used to confirm benign diagnoses. The dataset is randomly resampled into dataset into three classes, 18% for the test set, 18% for the validation set, and 64% for the training set.

3.2. EfficientNets and Augmentation

For the augmentation experiments, different augmentations methods are applied on the training set to obtain a new training set particular, then the new training set is used to train the EfficientNets.

3.2.1. EfficientNets (EffNets)

Mingxing Tan et al [18] initially proposed EfficientNets for more efficient computing while also achieving state-of-the-art 84 percent top-1 accuracy on ImageNet by designing a new baseline network EfficientNet-B0 and applying it with compound scaling to obtain a new family of EfficientNets.

In this paper, EfficientNets-B5 with the batch size 16 are used for our augmentation experiments.

3.2.2. Augmentations

Data augmentation is a common strategy to solve an intrinsic data imbalance problem [19]. Data augmentation allows practitioners to dramatically enhance the diversity of data availability for training sets by position and color augmentation techniques without directly collecting additional data.

The color and position augmentation methods we chose for the experiments are random horizontal flipping, random vertical flipping, random rotation with 90 degrees, random hue with random seed 0.01, random saturation with lower bound 0.7 and upper bound 1.3, random contrast with lower bound 0.8 and upper bound 1.2, random brightness with random seed 0.1.

In addition to the traditional augmentation methods, we would perform two non-traditional augmentation methods: randomly dropping out some selected patches on the images and adding faked hairs to the images.

3.3. Feature Extraction Algorithms

3.3.1. Gray Level Co-occurrence Matrix (GLCM)

Gray Level co-occurrence matrix feature extraction algorithm is a texture-based feature extraction method that determines the spatial relationship between pixels with a specific gray level by extracting second-order statistical features from images [20]. The spatial relationship, also known as the offset, is defined as two horizontally adjacent pixels of an image [21]. By defining the spatial relationship along with different directions and distances, one can obtain a multidimensional feature vector that describes the frequencies of occurrence of relative pixel pairs by applying the GLCM feature extraction method [16].

In this study, we converted the images from the training set to gray, then conducted GLCM feature extraction on each image. The GLCM features we chose to use are dissimilarity and correlation. In addition, the distance we chose is four offsets, and the directions we chose are 45 degrees, 60 degrees, 100 degrees, 120 degrees, 135 degrees, 180 degrees, 200 degrees, 225 degrees, 240 degrees, 300 degrees, and 340 degrees.

3.3.2. Local Binary Patterns (LBP)

Local binary patterns feature extraction algorithm is an effective feature extraction method that is robust to light fluctuations. By considering an image as an examined window, the LBP algorithm divided the window into equal-sized cells. The method compares each pixel in a cell to all of its neighbors by going clockwise or counterclockwise through all the neighbors [22]. The method generates a 256-dimensional feature vector after completing the comparisons on the central pixel. The feature vector is a histogram that contains the frequency values for each combination of pixels that are smaller and larger than the cell's center pixel.

Mathematically, the process of labeling the pixels can be described as [16]:

$$LBP_{\{P,R\}} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (1)$$

$$s(x) \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

where g_c is the gray value of a central pixel, g_p is the value of the neighbor of the central pixel.

3.3.3. Segmentation-based Fractal Texture Analysis (SFTA)

The SFTA feature extraction algorithm decomposes the gray-level input image into a series of binary images from which the fractal dimensions of the resultant regions are calculated to represent segmented texture patterns of the image. After applying multi-level threshold processing to a gray-level input image, the Two-Threshold Binary Decomposition (TTBD) algorithm converts the input image into different binary images [16].

The SFTA algorithm then receives the binary images as input, and extracts the features from the binary images. The SFTA algorithm's mathematical definition is interpreted as [16]:

$$\Delta(x, y) = \begin{cases} 1, & \exists (x', y') \in N_g[(x, y)]: I_b(x', y') = 0, I_b(x, y) = 1 \\ 0, & otherwise \end{cases} \quad (3)$$

where $I_b(x, y)$ as the binary images of the original gray-level image after applying the TTBD method.

3.3.4. Residual Neural Networks (ResNets)

A residual neural network (ResNet) is a kind of artificial neural network (ANN) based on pyramidal cell constructions in the cerebral cortex. The most common ResNet models include double-layer or triple-layer skips with some direct connections in between. The connections are called skip connections and are the core of residual blocks, the stacks of layers set in the ResNet model.

In addition to the traditional feature extraction algorithms, a pre-trained ResNet18 provided by the PyTorch Library `torch.utils.model_zoo` [23] is also used to extract the feature vector for each image in the data set.

3.3.5. Dimensionality Reduction of the Feature Vectors

Due to the strong independence of the ResNet18 features, PCA did not perform well in terms of dimensionality reduction. Consequently, the random projection approach is employed to reduce the dimension of the feature extracted from each feature extraction algorithm. The dimensions of the training set are reduced to 15 and 30 for experimental purposes.

3.4. Machine Learning Classifiers

3.4.1. Support Vector Machines (SVMs)

The Support Vector Machine algorithm is a supervised learning model used to solve regression and classification problems. A support vector machine creates a collection of hyperplanes in a high-dimensional space used for classification [24, 25]. It then uses the nearest points to the hyperplanes as support vectors to determine the optimal decision boundary that divides the data points into two classes with the minimum error. In other words, with the labeled training data points as the inputs, the SVM algorithm generates an optimum hyperplane that can categorize new sample points.

3.4.2. Extreme Gradient Boosting (XGBoost)

XGBoost is a distributed gradient boosting library used for supervised learning problems with better computational speeds compared with other boosting methods [26]. In this study, we applied the XGBoost classifier with `gbtree` as the booster.

3.4.3. Random Forest

The random forest algorithm is an ensemble learning algorithm consisting of many decision trees [27]. When developing every single tree, the random forest algorithm employs bagging and feature randomness to produce an uncorrelated forest of trees whose prediction is more accurate than any single decision tree.

4. Results

In this study, the four feature extraction methods GLCM, pre-trained ResNet18, LBP, and SFTA are utilized. After deducing the dimensions of the feature vectors, the new vectors are used as inputs to train three machine learning models: SVM, Random Forest, and XGBoost on test and validation sets.

4.1. Evaluation Metrics

4.1.1. AUC scores (Area Under the ROC Curve)

The Area Under the ROC Curve (AUC) is defined as the area enclosed with the coordinate axis under the ROC Curve [28]. A ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds. This curve plots with two parameters: the True Positive Rate and the False Positive Rate.

The True Positive Rate (Sensitivity Rate) is defined as

$$TP = \frac{TP}{TP+FN} \quad (4)$$

where TP is the number of true positives and \$FN\$ is the number of false negatives [29]. If an algorithm has a high sensibility rate, it has a better performance on predicting positive cases.

The False Positive Rate is defined as

$$FP = \frac{FP}{FP+TN} \quad (5)$$

where FP is the number of false positives and FN is the number of false negatives [29].

The ROC curve varies when the threshold is changed. In addition, the value of this area is not greater than 1. Since the ROC curve is generally above the straight-line $y=x$, the values of AUC scores range from 0.5 to 1. As the AUC score of a detection model is closer to 1.0, the model has a higher authenticity. However, when the AUC score of a model is 0.5, the model is not applicable for classification tasks.

4.1.2. True Negative Rate (Selectivity Rate)

The true negative rate (selectivity rate) is defined as

$$TN = \frac{TN}{TN+FP} \quad (6)$$

where TN is the number of true negatives, and FP is the number of false positives. The true negative rate is used to measure the performance of models in detecting the negative cases. If a detection model has a higher true negative rate, it has a better performance in detecting the negative cases.

4.2. EffNets with Augmentations

The experimental results are shown in Table 1. The augmentation examples are shown in Figure 1. Some conclusions can also be drawn from the experiments of the augmentation methods. Traditional augmentation methods, such as horizontal flipping and Saturation, can slightly improve the model performance. However, changing the hue of the images might degrade the performance of the model as the AUC score of Hue is 0.0024 lower than the AUC score of the model without any augmentation. Therefore, augmentations might not always lead to model improvement.

However, the two non-traditional augmentations, adding hair to the image and selecting patches on the image and then dropping them, were more effective than other traditional augmentation methods in

improving the performance of EffNets. These two non-traditional augmentation methods have better AUC scores on validation sets.

Table 1
The AUC scores of EfficientNet

Augmentation Method	Test Set	Validation Set
Rotation with 90 Degrees	0.9502	0.9177
Horizontal Flipping	0.9411	0.9148
Contrast	0.9450	0.9176
Saturation	0.9439	0.9120
Hue	0.9368	0.9198
Brightness	0.9422	0.9234
Hair Faking	0.9461	0.9241
Drop Out	0.9501	0.9237
Without Augmentation	0.9392	0.9091

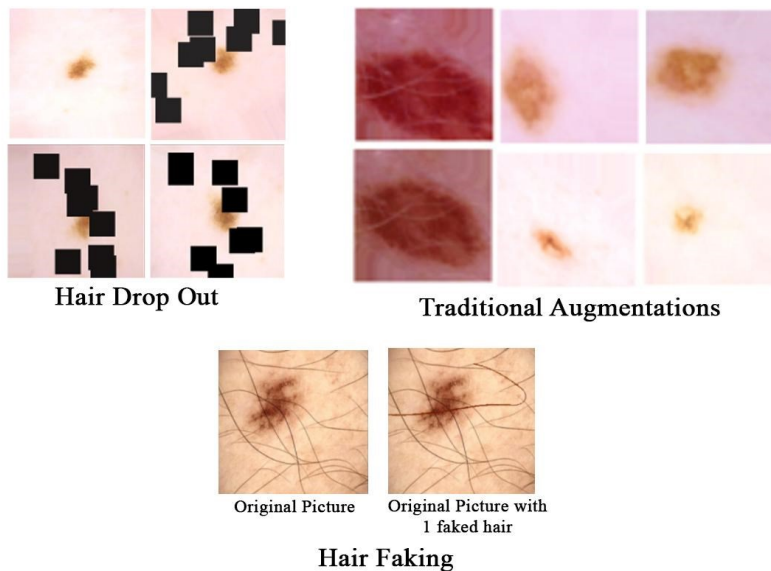


Figure 1: Augmentation Examples

4.3. Machine Learning Classifiers

The AUC scores of the experiments are shown in Table 2; the following findings are made by evaluating the outcomes achieved using various feature extraction methods in different classifiers.

- Increasing the number of features extracted by GLCM and ResNet18 can improve the AUC scores of the models.
- In contrast, increasing the number of features in LBP and SFTA may not improve the model's AUC scores. The new additional features may not have better performance at representing the entire picture from this data set.
- GLCM, on the other hand, has the highest AUC scores among the four feature extraction methods tested on the experimental data.

Table 2
The AUC scores of Different Classifiers with Different Feature Extraction Methods

Classifiers	Feature Extraction	# of Features	Test	Validation
SVM	GLCM	15	0.738	0.7199
		30	0.7326	0.7801

RF	LBP	15	0.7699	0.7808
		30	0.7572	0.7933
	SFTA	15	0.7918	0.8127
		30	0.8042	0.8196
	ResNet18	15	0.7225	0.7073
		30	0.7313	0.712
	GLCM	15	0.6122	0.6842
		30	0.6425	0.7506
	LBP	15	0.6521	0.7463
		30	0.649	0.7704
	SFTA	15	0.6503	0.6875
		30	0.5283	0.6034
ResNet18	15	0.6308	0.6764	
	30	0.6839	0.6758	
XGBoost	GLCM	15	0.6755	0.671
		30	0.7206	0.803
	LBP	15	0.7347	0.7466
		30	0.7112	0.7868
	SFTA	15	0.6734	0.7122
		30	0.6975	0.7429
	ResNet18	15	0.678	0.696
		30	0.6845	0.6979

4.4. Failure Modes Analysis

Since the models with 30 features have better AUC scores, the models with thirty features from the feature extraction methods on the validation set is chosen to be examined. The validation set contains 10,525 images, including 908 positive cases and 9617 negative cases. The predicted results of the algorithms range between 0 and 1. The sensibility rates are shown in Table 3 and the selectivity rates are shown in Table 4. We carefully examined the distributions, choosing 0.25 as the threshold value to determine if a predicted result successfully detects the positive cases. If a predicted value is lower or equal to 0.25, this predicted result is considered failing to detect a positive case.

Table 3
The Sensibility Rates of the Models on the Validation Set

	GLCM	LBP	SFTA	ResNet-18
RF	0.2192	0.9317	0.6101	0.5903
SVM	0.0518	0.4020	0.3425	0.0947
XGBoost	0.2037	0.4725	0.3139	0.1938

Table 4
The Selectivity Rates of the Models on the Validation Set

	GLCM	LBP	SFTA	ResNet-18
RF	0.2655	0.7567	0.7682	0.9337
SVM	0.9071	0.8978	0.9625	0.9799
XGBoost	0.9540	0.9150	0.9546	0.9540

Among the twelve combinations of feature extraction algorithm and machine learning model, the combination with the highest sensibility rate 0.9317 is the random forest model with the GLCM method. However, this combination does not have the highest AUC score. In addition, the combination with the

highest selectivity rate of 0.9799 is the SVM model with ResNet-18. Significantly, the combination with the highest sensibility rate maintains a 0.2655 selectivity rate; meanwhile, the combination with the highest selectivity rate maintains a 0.0518 sensibility rate.

By comparing the AUC scores of the combinations, we found out that even though the random forest algorithm has the lowest AUC scores, it has the best performance on correctly detecting the positive cases. Also, even though a model has a good performance in detecting positive cases, it might fail to detect negative cases.

5. Conclusion

In this paper, four feature extraction methods are used: GLCM, pre-trained ResNet18, LBP, and SFTA. SVM, Random Forest, and XGBoost are trained to perform classification tasks on the test and validation sets using the feature vectors from the prior feature extraction techniques.

Also, Effnets B5 along with different augmentation methods are implemented. From the experimental results, augmentation methods generally improve the performance of EffNets while the hair faking method has the largest AUC score on the validation set and the rotation method has the largest AUC score on the test set. However, random changing the hue of images fails to improve the models as it might blur the complexities of the images.

For the machine learning models with different feature extraction methods, the SVM algorithms with SFTA features have the best AUC scores. However, the random forest algorithm has better performance in terms of detecting positive cases.

However, since the data set is significantly unbalanced, the data set needs to be rebalanced for higher sensibility rates in the future work. The SMOTE approach can be made to balance the positive and negative cases. Meanwhile, since EffNets have higher AUC scores, other deep learning models with transfer learning approaches can also be implemented to improve the accuracy of predictions.

6. References

- [1] Curtin, J. A., Fridlyand, J., Kageshita, T., Patel, H. N., Busam, K. J., Kutzner, H., Cho, K. H., Aiba, S., Bröcker, E. B., LeBoit, P. E., Pinkel, D., & Bastian, B. C. (2005) Distinct sets of genetic alterations in melanoma. *The New England journal of medicine*, 353(20), 2135–2147. <https://doi.org/10.1056/NEJMoa050092>
- [2] American Cancer Society. (2021) Melanoma skin cancer: Understanding melanoma. <https://www.cancer.org/cancer/melanomaskin-cancer.html>.
- [3] Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., Halpern, A., Helba, B., Kittler, H., Kose, K., Langer, S., Lioprys, K., Malvey, J., Musthaq, S., Nanda, J., Reiter, O., Shih, G., Stratigos, A., Tschandl, P., Weber, J. Soyer, P. (2021) A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Scientific data*, 8(1): 1-8.
- [4] Tschandl, P., Rosendahl, C., & Kittler, H. (2018) The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5(1): 1-9.
- [5] Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... & Halpern, A. (2018) Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. pp. 168-172.
- [6] Marc Combalia, Noel C. F. Codella, Veronica Rotemberg, Brian Helba, Veronica Vilaplana, Ofer Reiter, Allan C. Halpern, Susana Puig, Josep Malvey. (2019) BCN20000: Dermoscopic Lesions in the Wild. <https://arxiv.org/pdf/1908.02288.pdf>
- [7] Agnieszka, M., Michał, G. (2018) Data augmentation for improving deep learning in image classification problem. In: International Interdisciplinary PhD Workshop (IIPhDW). Poland. pp. 117-122.
- [8] Wu, X., et al. (2018) Hyphae Detection in Fungal Keratitis Images with Adaptive Robust Binary Pattern. *IEEE Access*, 6: 13449-13460.

- [9] Alquran, H., et al. (2017) The melanoma skin cancer detection and classification using support vector machine. In: IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT). Aqaba, Jordan. pp. 1-5.
- [10] Kavitha, J. C., Suruliandi, A. (2016) Texture and color feature extraction for classification of melanoma using SVM. In: International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16). Kovilpatti, India. pp. 1-6.
- [11] Javaid, A., Sadiq, M., Akram, F. (2021) Skin Cancer Classification Using Image Processing and Machine Learning. In: International Bhurban Conference on Applied Sciences and Technologies (IBCAST). Islamabad, Pakistan. pp. 439-444.
- [12] Paja, W., Wrzesie'n, M. (2013) Melanoma important features selection using random forest approach. In: 6th International Conference on Human System Interactions (HSI). Sopot, Poland. pp. 415-418.
- [13] Mustafa, S., Kimura, A. (2018) A SVM-based diagnosis of melanoma using only useful image features. In: International Workshop on Advanced Image Technology (IWAIT). Chiang Mai, Thailand. pp. 1-4.
- [14] Yang, A., Yang, X., Wu, W., Liu, H., Zhuansun, Y. (2019) Research on Feature Extraction of Tumor Image Based on Convolutional Neural Network. IEEE Access, 7: 24204-24213.
- [15] He, K., Zhang, X., Ren, S., & Sun, J. (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 770-778.
- [16] Öztürk, Ş., & Akdemir, B. (2018) Application of feature extraction and classification methods for histopathological image using GLCM, LBP, LBGLCM, GLRLM and SFTA. *Procedia computer science*, 132: 40-46.
- [17] Perez, F., Avila, S., Valle, E. (2019) Solo or ensemble? choosing a cnn architecture for melanoma classification. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Long Beach, CA, USA. pp. 2775-2783.
- [18] Tan, M. and Le, Q. V. (2020). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks 2020. <https://arxiv.org/pdf/1905.11946.pdf>
- [19] Agnieszka, M., Michał, G. (2018) Data augmentation for improving deep learning in image classification problem. In: *International Interdisciplinary PhD Workshop (IIPhDW)*. Poland. pp. 117-122.
- [20] Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973) Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6): 610-621.
- [21] Mohanaiah, P., Sathyanarayana, P., & GuruKumar, L. (2013) Image texture feature extraction using GLCM approach. *International journal of scientific and research publications*, 3(5): 1-5.
- [22] Ojala, T., Pietikainen, M., & Harwood, D. (1994, October) Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In *Proceedings of 12th international conference on pattern recognition*. Jerusalem. Vol. 1, pp. 582-585.
- [23] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S. (2019) Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32: 8026-8037.
- [24] Vapnik, V. (1998) The support vector method of function estimation. In *Nonlinear modeling*. Springer, Boston, MA. pp. 55-85.
- [25] Vapnik, V. (1995) *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
- [26] Chen, T., Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM. <https://doi.org/10.1145/2939672.2939785>
- [27] Ho, T. K. (1995) Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (Vol. 1, pp. 278–282).
- [28] Hand, D. J. (2009) Measuring classifier performance: a coherent alternative to the area under the ROC curve. *Machine learning*, 77(1), 103-123.
- [29] Sultana, N., Puhan, N., Mandal, B. (2018) DeepPCA Based Objective Function for Melanoma Detection. In: *International Conference on Information Technology (ICIT)*. Bhubaneswar, India. pp. 68-72.