Skin Lession Classification Using Image Processing *Note: Sub-titles are not captured in Xplore and should not be used

[1]Mr Sachin Mahadeo Bagade, [2]Dr P M Mahajan, [3]Dr A M Patil, [4]Prof O K Firke

[1]M Tech Student , *Department of E&TC* , [2]Dean of Academic, Department of E&TC
[3]HOD, Department of E&TC Engineering, [4]Associate Professor, Department of E&TC
[1,2,3,4] J T Mahajan College of Engineering, Faizpur

Abstract: Melanoma is a type of skin cancer with a high mortality rate. The different types of skin lesions result in an inaccurate diagnosis due to their high similarity. Accurate classification of the skin lesions in their early stages enables dermatologists to treat the patients and save their lives. This paper proposes a model for a highly accurate classification of skin lesions. The proposed model utilized the GLCM and Gabor features. Performance of SVM, KNN and Naïve Bayes Classifiers is evaluated. The latest well-known public challenge dataset, ISIC 2019, is used to test the ability of the proposed model to classify different kinds of skin lesions. The proposed model successfully classified the nine different classes of skin lesions, namely, melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, benign keratosis, dermatofibroma, vascular lesion, and Squamous cell carcinoma. The achieved classification accuracy, using KNN and Gabor features are 96.07. The proposed model can detect images that do not belong to any one of the nine classes where these images are classified as unknown images.

INTRODUCTION

In 2018 [1], the WHO reported that there are more than 14 million new cancer patients and more than 9.6 million deaths over the world because of cancer. These statistics show that cancer is the leading cause of human death [2], [3]. Skin Cancer initially occurs on the upper layer of the skin, the epidermis, where it is noticeable and can be seen by human eyes [4]. Skin cancer is one of the significant contributors to the cause of death over the world [5].

Different types of skin cancers have been discovered. Melanoma is a well-known kind of skin cancer, which usually is the most malignant lesion compared to other skin lesions types [6], [7]. Melanoma is one of the fastest spreading skin cancers where recent studies show that the number of skin cancer patients increased year by year [8], [9]. Automatic computer-aided systems for accurate classification of skin lesions are beneficial to saving human life. In the present era, Computer-Aided Diagnosis (CAD) systems become a necessity for preliminary diagnosis of different diseases. Image-based CAD systems utilize skin lesions images without any other medical information to provide dermatologists with an accurate diagnosis [10]. Furthermore, an image-based CAD system could classify different skin lesions based on features extracted from the colors in images of the skin [11]. Based on its accuracy, a CAD system could lead to early diagnosis of skin cancer and then open the door for treatments to save human life [12].

Dermoscopy is the most well-known method of skin imaging, which showed an improvement in the diagnosis of melanoma compared to that of the naked eye [13]. Nevertheless, the benchmark datasets of skin cancer are limited and contain a few numbers of classes. Besides, the available number of images in these classes is limited. Three different issues make the automatic identification of melanoma from dermoscopy images a challenging task. First, although the skin lesions belong to different classes, the characteristics of these lesions, such as size, texture, color, and shape, are very similar, which makes classification a challenging task. Second, there is a high correlation between melanoma and non-melanoma lesions. Third, environmental conditions, such as hair, veins, and illumination [14].

Several attempts made to overcome these challenging issues. Oliveira *et al.* [15] used low-level hand-crafted features to differentiate between melanoma and non-melanoma lesions. Unfortunately, this trial led to wrong results due to the high visual similarity, significant intra-class variations, and dermoscopic artifacts [16]. Pathanbet *et al.* [17] and Shimizu *et al.* [18] segmented the input images to remove the background and unnecessary contents to improve the classification of the skin lesions. Nonetheless, poor results were obtained when segmentation and classification procedures focused on features at low levels, which results in low discrimination rates [19].

Many techniques, such as ABCDE rule, genetic algorithms, support vector machines (SVMs), and artificial neural networks (ANNs), proposed to assess the skin lesion and classify it as either melanoma or benign [20]–[24]. All of these procedures were efficient, cost-effective, and less painful than conventional medical techniques.

The contributions in this work can be summarized through the following points:

1. We proposed the architecture of the machine learning using GLCM and Gabor features extraction.
2. In the proposed model, during pre-processing we used median filtering for noise removal. We have evaluated the performance using SVM, KNN and Naïve Bayes algorithms.
3. In this paper, we classified this ISIC 2019 challenging dataset with high-performance measures in two different ways.
4. The Gabor features extraction with KNN classifier achieved higher classification rates than other machine learning architectures.

The rest of this work is organized as follows: The utilized methods are described in section 2; the datasets, the performed experiments, and the discussion are described and discussed in section 3, and finally, the conclusion is presented in section 4.

DATASET

The goal for ISIC 2019 is classify dermoscopic images among nine different diagnostic categories:

1. Melanoma
2. Melanocytic nevus
3. Basal cell carcinoma
4. Actinic keratosis
5. Benign keratosis (solar lentigo / seborrheic keratosis / lichen planus-like keratosis)
6. Dermatofibroma
7. Vascular lesion
8. Squamous cell carcinoma
9. None of the others

The dataset for workshop ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection 1 is used [21], [22]. In the training set, there are a total of 10015 skin lesion images from seven skin diseases- Melanoma (1113), Melanocytic nevus (6705), Basal cell carcinoma (514), Actinic keratosis (327), Benign keratosis (1099), Dermatofibroma (115) and Vascular (142). The validation dataset consists of 193 images. Figure 1 shows some example of these 7 types of the skin lesion. Task- 3: the goal of task-3 is to find improved automated predictions of disease classification within dermoscopic images. Possible disease categories are in below figure 1.

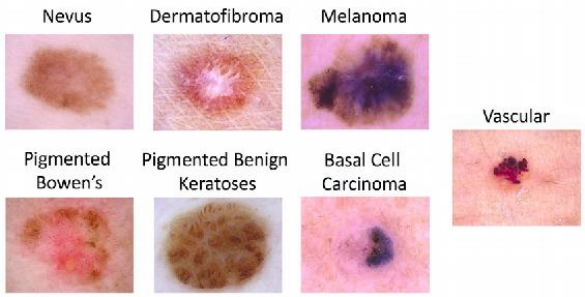

Figure 1: Sample database images from ISIC 2019

I. GENERAL ARCHITECTURE FOR SKIN CANCER CLASSIFICATION

The system contains five major modules: skin disease input database, preprocessing/ Noise removal, feature extraction, classifier & recognized output as illustrated in fig.2. Overall, the system is based on preprocessing/ noise removing mechanism of skin disease images, extracting some of features which contain information about textural features of the image & taking appropriate pattern recognition model to identify the type of skin disease image disease

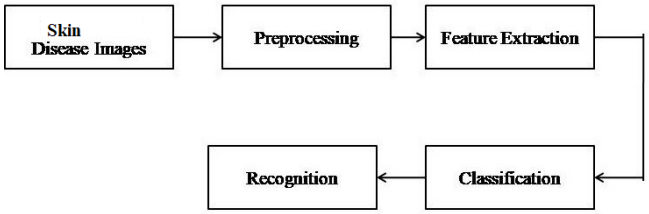

Fig.2. Structure of Skin Disease Classification

Various steps for classification of skin disease from disease database images as follows:

1. Database Gathering and Preprocessing:Obtaining the skin disease images Data from ISIC 2019 Dataset
2. Preprocessing and noise removal of skin disease images using Median Filter.
3. Feature Extraction : GLCM, Gabor features are extracted from skin disease images
4. Design and development of the system for classification of skin disease images as normal or diseased containing 9 types of skin diseases.
5. Evaluate the performance of different classifiers

- SVM
- KNN
- Naïve Bayes

GUI is developed in MATLAB as shown in figure 3

a. Gabor Features Extraction

In image processing, a Gabor filter, named after Dennis Gabor, is a linear filter used for texture analysis, which means that it basically analyzes whether there are any specific frequency content in the image in specific directions in a localized region around the point or region of analysis. Frequency and orientation representations of Gabor filters are claimed by many contemporary vision scientists to be similar to those of the human visual system, though there is no empirical evidence and no functional rational to support the idea. They have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave.

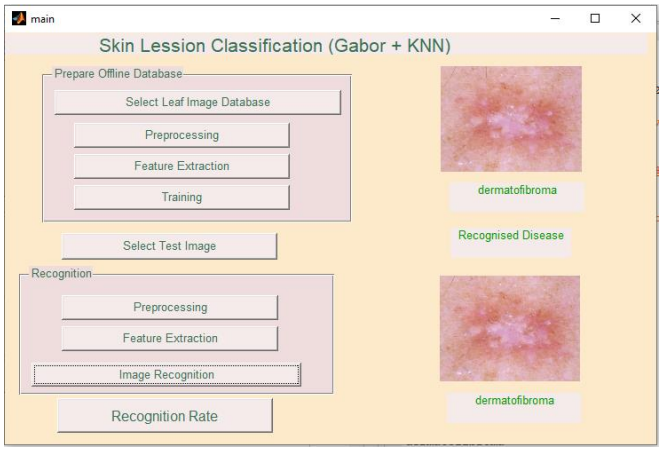

Fig.3. Matlab Based GUI for Skin disease Disease Classification

b. Statistical Features

stats = graycoprops(glcm, properties) calculates the statistics specified in properties from the gray-level co-occurence matrix glcm. glcm is an m-by-n-by-p array of valid gray-level co-occurrence matrices. If glcm is an array of GLCMs, stats is an array of statistics for each glcm.

graycoprops normalizes the gray-level co-occurrence matrix (GLCM) so that the sum of its elements is equal to 1. Each element (r,c) in the normalized GLCM is the joint probability occurrence of pixel pairs with a defined spatial relationship having gray level values r and c in the image. graycoprops uses the normalized GLCM to calculate properties

Different statistical features considered here are contrast, correlation, Energy and Homogeneity

c. Support Vector Machines

In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on the side of the gap on which they fall.

In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. When data are unlabelled, supervised learning is not possible, and an unsupervised learning approach is required, which attempts to find natural clustering of the data to groups, and then map new data to these formed groups.

d. KNN Classification

In pattern recognition, the k-nearest neighbors algorithm (k-NN) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression:

In k-NN classification, the output is a class membership. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighbor.

In k-NN regression, the output is the property value for the object. This value is the average of the values of k nearest neighbors. k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The k-NN algorithm is among the simplest of all machine learning algorithms. Both for classification and regression, a useful technique can be used to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. For example, a common weighting scheme consists in giving each neighbor a weight of 1/d, where d is the distance to the neighbor. The neighbors are taken from a set of objects for which the class (for k-NN classification) or the object property value (for k-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. A peculiarity of the k-NN algorithm is that it is sensitive to the local structure of the data.

e. Naïve Bayes Classification

Naive Bayes classifiers are a collection of classification algorithms based on **Bayes' Theorem**. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other. Naïve Bayes is a classification algorithm that relies on strong assumptions of the independence of covariates in applying Bayes Theorem. The Naïve Bayes classifier assumes independence between predictor variables conditional on the response, and a Gaussian distribution of numeric predictors with mean and standard deviation computed from the training dataset.

Naïve Bayes models are commonly used as an alternative to decision trees for classification problems. When building a Naïve Bayes classifier, every row in the training dataset that contains at least one NA will be skipped completely. If the test dataset has missing values, then those predictors are omitted in the probability calculation during prediction.

II. DIFFICULTIES IN SKIN CANCER DETECTION USING MACHINE LEARNING

Difficulties of S.C.D. Identifying skin lesions is difficult because of various image types and sources, as depicted in Fig. 2. The look of human skin color varies greatly, making skin identification challenging and time-consuming. The following are some of the issues posed by the complex visual qualities of skin lesions images (Naji et al., 2019):

- The wide range of skin lesions adds to the intricacy of these images and makes precise skin lesion diagnosis extremely challenging. Lesion location, size, and shape are all extremely variable. Thus, most image analysis approaches require image preprocessing to analyze skin lesion images accurately
- During the acquisition of an image, noise is introduced. In the presence of noise and artifacts, skin lesion photos may be difficult to identify. Human and even computer-assisted approaches to skin lesion segmentation can influence an image's interpretation if these compromising signals are present in the picture. Hair, bubbles, and blood vessels are just a few examples
- The presence of fuzzy and uneven boundaries in some skin lesion images complicates several contour refinement and boundary localization procedures. Preprocessing skin lesion images for simple asymmetry prediction can be difficult
- Low contrast from adjacent tissues can also be observed, adding to the problem. The low difference between the lesion and the surrounding skin makes segmentation challenging
- The lesion's color texture, light beams, and reflections influence an image's lighting

III. RESULT ANALYSIS

Matlab-based GUI-driven tool is developed for effective classification of skin diseases. Fig.3 shows graphical user interface(GUI) developed for proposed algorithm before execution. GUI for this software is divided into number of subgroups according to their functionality.

a. Database Selection and Preprocessing:

Skin disease images training database is selected. Then for preprocessing, median filter is used for noise removal.

b. Features Extraction

From the preprocessed training images Gabor, and GLCM statistical features are extracted. Features matrix is constructed.

c. Classification

Different classifiers including SVM, KNN, Naïve Bayes are trained with various features for skin disease classification. This module deals with skin disease detection and classification. The performances of different classifiers have been evaluated by considering different number of training images. Four parameters are used for evaluating performance of the algorithm. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN)

True Positive: Skin disease detection coincides with actual labelled data

True Negative: both classifier and actually labelled absence of Skin disease

False Positive: system labels a healthy case as an Skin disease one

False Negative: system labels Skin disease image as healthy

Accuracy: Accuracy is the ratio of number of correctly classified cases, and is given by

$$\text{Accuracy} = (TP+TN) / N$$

Total number of cases are N

Precision is the fraction of retrieved images that are relevant to the query. Precision takes all retrieved images into account, but it can also be evaluated at a given cut-off rank, considering only the results returned by the system

Precision : Precision is defined as

$$\text{Precision} = TP/ (TP+FP)$$

Recall: Recall is the fraction of the relevant images that are successfully retrieved. In binary classification, recall is called sensitivity. It can be viewed as the probability that a relevant document is retrieved by the query. It is trivial to achieve recall of 100% by returning all documents in response to any query. Therefore, recall alone is not enough but one needs to measure the number of non-relevant documents also, for example by also computing the precision.

Recall is defined as

$$\text{Recall} = TP / (TP+FN)$$

F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but F1 is usually more useful than accuracy, especially if you have an uneven class distribution. Accuracy works best if false positives and false negatives have similar cost. If the cost of false positives and false negatives are very different, it's better to look at both Precision and Recall. In our case, F1 score is 0.701.

$$\text{F1 Score} = 2*(\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

Table 1 shows the performance of the porposed system in terms of precision

TABLE I. PERFORMANCE MEASURE USING PRECISION

Precision		
	Gabor	GLCM
SVM	0.4872	0.2794
KNN	0.5032	0.4872
NB	0.0777	0.0491

Table 2 shows the performance of the porposed system in terms of recall

TABELE II PERFORMANCE MEASURE USING RECALL

Recall

	Gabor	GLCM
SVM	1	0.7922
KNN	17.8596	1
NB	0.7895	0.7609

Table 3 shows the performance of the porposed system in terms of F measure

TABLE III PERFORMANCE MEASURE USING F MEASURE

Recall		
	Gabor	GLCM
SVM	0.6552	0.4474
KNN	0.9788	0.6552
NB	0.1415	0.1129

Table 4 shows the performance of the porposed system in terms of Accuracy

TABLE IV PERFORMANCE MEASURE USING ACCURACY

Recall		
	Gabor	GLCM
SVM	0.9465	0.7387
KNN	0.9607	0.9465
NB	0.5134	0.4848

As depicted in tables I to IV, KNN classification using Gabor features gives better accuracy of 96.07%

As depicted in tables I to IV, Gabor features with all classifiers give promising result for skin disease classification. Table V shows the performance evaluation of the algorithms in terms of execution time. The KNN classifier requires more execution time for training and evaluation.

TABLE V PERFORMANCE MEASURE USING ACCURACY

Execution Time		
	Gabor	GLCM
SVM	2.0632	1.8554
KNN	4.8612	0.0643
NB	0.2112	0.1826

Fig.4 shows the performance measure of different classifiers using Gabor and GLCM features for skin disease classification.

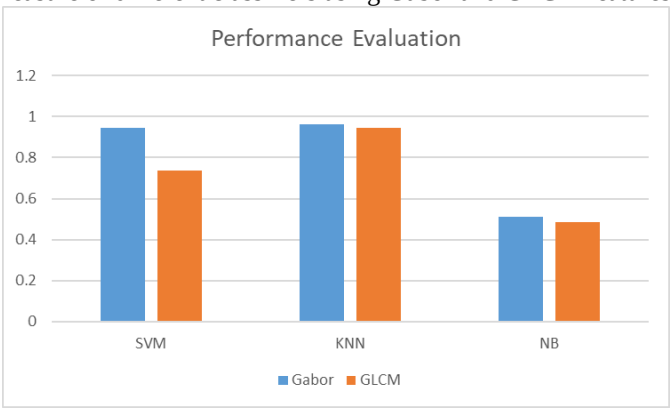

Fig.4. Performance Measure Using Gabor and GLCM Features

As depicted in Fig.4 the accuracy of KNN classification is better but the execution time is more. So for skin diseases classification, KNN classification is better with better performance.

IV. CONCLUSION

A method for ISIC 2019 challenge dataset has been developed here by using machine learning and GLCM and Gabor features extraction. The proposed method can classify nine different types of lesions accurately, even with the imbalance of images between classes. The performance measures of the proposed methods were accuracy, precision, recall, f measure and execution time. Using Gabor features with KNN classification we achieve the better performance of 96.07%. But the execution time for KNN using Gabor features is more. In future we will try to reduce the execution time of training KNN along with increasing performance. We will also try to generalize the system for classification of any type of disease images.

VII .REFERENCES

1. *Cancer*, Mar. 2018, [online] Available: https://www.who.int/news-room/fact-sheets/detail/cancer.
2. Y.-E. Choi, J.-W. Kwak and J. W. Park, "Nanotechnology for early cancer detection", *Sensors*, vol. 10, pp. 428-455, Jan. 2010.
3. *National Cancer Institute. Cancer Statistics*, Oct. 2017, [online] Available: https://www.cancer.gov/aboutcancer/understanding/statistic.
4. *National Cancer Institute. Skin Cancer (Including Melanoma)-Patient Version*, Oct. 2018, [online] Available: https://www.cancer.gov/types/skin.
5. H. E. Kanavy and M. R. Gerstenblith, "Ultraviolet radiation and melanoma", *Seminars Cutaneous Med. Surg.*, vol. 30, no. 4, pp. 222-228, Dec. 2011.
6. C. K. Bichakjian, A. C. Halpern, T. M. Johnson, A. F. Hood, J. M. Grichnik, S. M. Swetter, et al., "Guidelines of care for the management of primary cutaneous melanoma", *J. Amer. Acad. Dermatol.*, vol. 80, no. 1, pp. 208-250, 2019.
7. American Cancer Society. Facts & Figures, Atlanta, Ga, USA, 2020.
8. R. J. Friedman, D. S. Rigel and A. W. Kopf, "Early detection of malignant melanoma: The role of physician examination and self-examination of the skin", *CA A Cancer J. Clinicians*, vol. 35, no. 3, pp. 130-151, 1985.
9. Z. Apalla, D. Nashan, R. B. Weller and X. Castellsagué, "Skin cancer: Epidemiology disease burden pathophysiology diagnosis and therapeutic approaches", *Dermatol. Therapy*, vol. 7, no. S1, pp. 5-19, Jan. 2017.
10. M. Q. Khan, A. Hussain, S. U. Rehman, U. Khan, M. Maqsood, K. Mehmood, et al., "Classification of melanoma and nevus in digital images for diagnosis of skin cancer", *IEEE Access*, vol. 7, pp. 90132-90144, 2019.
11. K. Doi, "Computer-aided diagnosis in medical imaging: Historical review current status and future potential", *Computerized Med. Imag. Graph.*, vol. 31, no. 4, pp. 198-211, Jun. 2007.

A. M. Abdel-Zaher and A. M. Eldeib, "Breast cancer classification using deep belief networks", Expert Syst. Appl., *vol. 46, pp. 139-144, Mar. 2016.*

12. M. Binder, "Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists", *Arch. Dermatol.*, vol. 131, no. 3, pp. 286-291, Mar. 1995.

A. Barata, M. E. Celebi and J. S. Marques, "A survey of feature extraction in dermoscopy image analysis of skin cancer", IEEE J. Biomed. Health Informat., *vol. 23, no. 3, pp. 1096-1109, May 2019.*

13. R. B. Oliveira, A. S. Pereira and J. M. R. S. Tavares, "Computational diagnosis of skin lesions from dermoscopic images using combined features", *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6091-6111, Oct. 2019.
14. M. E. Celebi, Q. Wen, H. Iyatomi and K. Shimizu, "A state-of-the-art survey on lesion border detection in dermoscopy images", *Dermoscopy Image Analysis*, vol. 10, pp. 97-129, Sep. 2015.
15. S. Pathan, K. G. Prabhu and P. C. Siddalingaswamy, "Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review", *Biomed. Signal Process. Control*, vol. 39, pp. 237-262, Jan. 2018.
16. K. Shimizu, H. Iyatomi, M. E. Celebi, K.-A. Norton and M. Tanaka, "Four-class classification of skin lesions with task decomposition strategy", *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 274-283, Jan. 2015.
17. F. Peñaranda, V. Naranjo, G. Lloyd, L. Kastl, B. Kemper, J. Schnekenburger, et al., "Discrimination of skin cancer cells using Fourier transform infrared spectroscopy", *Comput. Biol. Med.*, vol. 100, pp. 50-61, Sep. 2018.
18. N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, et al., "Convolutional neural networks for medical image analysis: Full training or fine tuning?", *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299-1312, May 2016.

A. Ruiz, V. Berenguer, A. Soriano and B. Sánchez, "A decision support system for the diagnosis of melanoma: A comparative approach", Expert Syst. Appl., *vol. 38, no. 12, pp. 15217-15223, Nov. 2011.*

B. Murugan, S. A. H. Nair and K. P. S. Kumar, "Detection of skin cancer using SVM random forest and kNN classifiers", *J. Med. Syst.*, vol. 43, no. 8, pp. 269, Aug. 2019.

19. Z. Ma and J. M. R. S. Tavares, "A novel approach to segment skin lesions in dermoscopic images based on a deformable model", *IEEE J. Biomed. Health Informat.*, vol. 20, no. 2, pp. 615-623, Mar. 2016.
20. S. B. Sulistyo, W. L. Woo and S. S. Dlay, "Regularized neural networks fusion and genetic algorithm based on-field nitrogen status estimation of wheat plants", *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 103-114, Feb. 2017.
21. L. Yu, H. Chen, Q. Dou, J. Qin and P.-A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks", *IEEE Trans. Med. Imag.*, vol. 36, no. 4, pp. 994-1004, Apr. 2017.