Towards Skin Cancer Classification Using Machine Learning and Deep Learning Algorithms: A Comparison

Iqra Kiran[1], Muhammad Zohaib Siddique [2], Ateeq Ur Rehman Butt[1], Dr. Azhar Imran Mudassir[3], Muhammad Azeem Qadir[1], Sundus Munir[4]

[1] Department of Computer Science (National Textile University).

[2] Department of Computer Science (Riphah International University).

[3] Department of Creative Technologies (Air University Islamabad).

[4] Department of Computer Science (Lahore Garrison University).

S kin cancer is an uncontrolled development of abnormal skin cells potentially due to excessive exposure to sun, history of sunburns, less melanin, Precancerous skin lesions, moles, etc. This occur when unrepaired DNA damages the cells of the skin. It is one of the diseases that are viewed on its quick evolution and the most common type of cancer that endangers life. Researchers have implemented several machine learning and deep learning techniques for classification of skin cancer. In this research paper, different cancer categories are classified using significant attributes. We have used International Skin Imaging Collaboration (ISIC) dataset for classification purposes. This dermoscopic attributes dataset includes 1000 images and 10016 instances, seven categories, 5 features and 2 Meta attributes. We implemented K-Nearest Neighbor, Logistic Regression, Convolutional Neural Network, Naïve Bayes, and Decision Tree for classification and compared their performance. In order to implement classification algorithm, we used Orange which is an open-source machine learning, data mining, and data visualization toolkit. The models are evaluated based on matrices that include Accuracy, C. Automation, F1 score, Precision, Recall, and AUC. Furthermore, frequency of features is visualized using graphical method and the ROC analysis is also performed for the classifiers. It is observed that CNN technique provided the highest accuracy of 89% and the mentioned results are the highest results of classification with the state of the art techniques. For future, the improved and recent dataset and ensemble modelling techniques based on deep learning can used to enhance classification results. The research can also be extended for other cancer types using CNN.

Keywords: skin cancer; classification;

Introduction

Skin is the principal and greatest organ of the human body. Skin covers the bones, muscles, and all body parts. Even a small problem can lead to major damage. Various types of skin cancers may occur due to some infections. Usually, people visit the consultant when cancer reaches its critical stage and the patient may face trouble in recovering [1]. Skin cancer is an aggressive form of cancer amongst various types of cancers. The number of patients affected by skin cancer has risen to 53% over the last decade. In the United States, 01 in 52 women, and 01 in 32 men were infected with melanoma, and approximately 10 million people have died from melanoma. Latest studies revealed that 98% of the patients who survive early identified with melanoma and only 17% survive when melanoma was left incurable at its initial level [2][3]. Approximately 178,560 new melanoma cases have been reported in the US since 2018, including non-invasive 87,290 and invasive 91,270 cases. Furthermore, melanoma-related deaths have reached up to 9,320, which involves 3330 women and 5990 men [4][5].
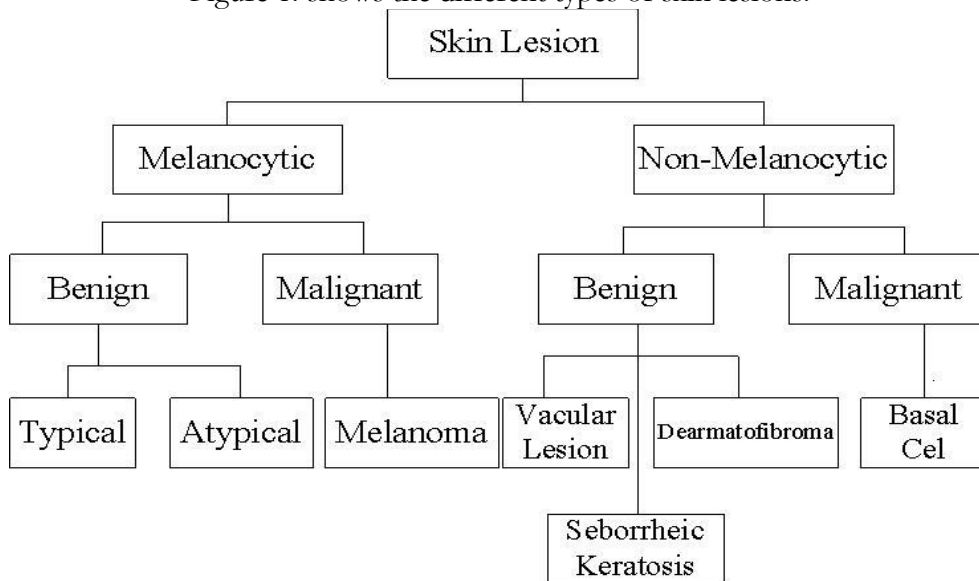
Figure 1. shows the different types of skin lesions.

Figure 1. Skin lesion classification tree

Skin cancer is categorized into seven forms. Actinic keratoses (akiec) Is a rough, scaly patch, which is commonly found at face, lips, and back. Basal Cell Carcinoma (BCC) is a category of a non-melanocytic malignant lesion. It is a public type, but the least dangerous form of a tumor. It grows slowly and is most common in areas of skin that is exposed to the sun more often like face. Mostly on the neck, skull, and upper torso. Squamous Cell Carcinoma (SCC) is mostly on dark skin color. It is mostly on the legs and feet. Benign keratosis (bkl) natural, harmless, non-cancerous growth on the skin is seborrheic keratosis. It occurs generally as a red, black, or brown growth on the back, arms, chest, or face, dermatofibroma (df) is a common type of benign skin tumor seen most often on the legs, small, slow-growing, typically firm, red-to-brown bump, Pigmented nevi (moles) are skin lesions that are generally black, brown or skin-colored, melanoma (melanoma), and vascular skin (vascular), is the most commonly diagnosed lesions in skin cancer. The physicians developed various strategies for skin lesion assessment comprising CASH, 7-point checklist, and ABCD rule [6].

ISIC database is the world's biggest publicly available repository for dermoscopic images of skin lesions. In 2018 ISIC conducted an image analysis competition in which participants were to practice a dermoscopic image using the HAM 10,000 dataset to recognize one of seven classes: melanoma, melanocytic NV, bcc, actinic keratosis, bkl, df, and vascular lesion [7]. This initiative is a follow-up to a similar approach last year, which was held in collaboration with the 2017 International Biomedical Imaging Symposium [8].

Dubal et al. [9] used a Neural Network and ABCD rule has been used to classify the image to a high degree of accuracy. Their techniques were applied as an expert software program, where users provide input in terms of cancer images and identified skin lesions are beginning or malignant. Farooq et al. [10] applied different algorithms on the different datasets and provided comparative analysis. They used SVM and NN classifiers to classify the segmented moles. R. Ashraf et al. presented the use of deep learning in their research work for the classification of skin cancer images in an effective way [11]. Mhaske et al. [12] compared different ML algorithms by using the ph2 dataset. They classified melanoma skin cancer on supervised and unsupervised Machine learning. Murthi et al. suggested a computer-aided melanoma skin cancer identification using ANN. They calculate accuracy on MATLAB and achieved the highest accuracy of 96%. Ramlakhan et al. [13] used the classical machine learning techniques to design a technique to classify benign and malignant lesions. When tests were conducted on 83 images, with an accuracy of 66.7 percent. Some scientists also focus on skin diseases other than skin cancer. AUR Butt et al. presented a computer-aided diagnosis for segmentation and classification of burnt human skin by using different machine learning algorithms by incorporating the comparison between all the used algorithms [14].

Islam et al. [15] concentrated on the development of a portable classification system for pigmented skin lesions. The classification system proposed for skin lesions uses image processing and artificial intelligence to evaluate the texture-based characteristics derived from the image of the disease. The arsenic detection accuracy and recall rates were 88% and 84%. A mobile-based optimization approach for the classification of benign and malignant lesions were implemented by Aleem et al. [16]. Trained and tested on the dataset, the smartphone app contained only 84 images. Research and testing on small dataset, results in 80% sensitivity and 75% specificity. Hekler et al. [17] integrated human intelligence and artificial intelligence to identify skin cancer. Using 11,444 dermoscopic images a specific CNN was trained to classify skin lesion images into five groups. These images have also been classified by dermatologists, and it has been discovered that human and artificial intelligence combines to accomplish superior results. Their proposed methodology achieved 82.95% accuracy. Mobile phones are also utilized in the field to classify skin lesions.

F. A. Khan et al. presented the use of deep convolutional neural networks in their research work for the segmentation and classification of burnt skin images of human beings [18]. One such attempt Ahmed et al. [19] did where 48,373 dermoscopic images were trained for binary classification of skin lesions using the Convolutional Neural Network model MobileNetV2. Using the qualified model, a skin lesion image with an accuracy rate of 91.33% was graded as benign or malignant. After the classification model was trained, an iOS-based mobile app was designed to assess its efficiency on unseen images. Abbas et al. [20] used the standard approach to machine-learning to identify benign and malignant skin lesions. 900 images were included in their proposed approach. Second, segmentation of the image using an edge detection technique to remove the ROI. Texture-based characteristics were extracted and SVM was applied to them to achieve the overall classification mark and 99.02% accuracy was obtained. F. A. Khan et al. presented the use of DCNN for segmentation and depth classification of burnt human skin by claiming that their obtained results are the best and highest results among the previous results of the state-of-the-art techniques of related works [21].

Deep learning methods are also being used for the classification of skin lesions by using pre-trained learning models. Mahbod et al. [22] have suggested a fully automated classification scheme. Features were produced using AlexNet [23], VGG16 [24], and ResNet-18 [25] in the implemented classification scheme, and passed for final prediction to the SVM classifier. The experimental classification scheme was tested on 150 images and the melanoma and seborrheic keratosis yielded a region under the curve of 83.83 % and 97.55% respectively.

In this article, the flow is the programmer's design section in which the system's mathematical model that defines the console's input and the output state is clarified. The classification results

obtained by these algorithms and their precision are also discussed. Finally, the results are discussed in terms of accuracy for various kinds of classifiers.

In this work, we experimented with the ISIC 2018 dataset. The classification is performed using Convolutional Neural Network, Logistic Regression, K-Nearest Neighbor, Naïve Bayes, and Decision Tree.

Material and Methods
Dataset site

HAM 1000 [8] (Human against a machine with 10000 images) released by International Skin Imaging Collaboration (ISIC) including 10016 instances. The dataset is a multi-classification with seven different labels. The main objective of using a dataset is to classify the different categories of skin cancer. It is publicly available for academics used to perform machine learning processes. The dataset consists of dermoscopic attributes. 40% dataset are considered as training and 60% are considered as testing. The dataset consists of 5 features and 2 Meta attributes. The description of data is defined in Table 2.

Table 2. Dataset description.

Dx	Akies, bcc, bkl, df, mel, nv, vasc
dx-type	Confocal, consensus, follow-up, Histo
Sex	Female, male, unknown
Age	25-80
Localization	Abdomen, acral, back, chest, ear, face, foot, genital, hand, lower extremity, neck, scalp, trunk, unknown, upper extremity

Decision Tree

Decision Tree (DT) is supervised learning among which one of the dissimilar methods is built for classification. It uses inductive reasoning to generate a tree structure in which each node indicates an attribute while the node's each descending branch represents one of the possible outcomes for that attribute. Each node of the tree is the distinguishing equation when classifying all data. It is a common method that provides both classification and predictive function at the same time [15]. We applied DT for the classification of skin cancer by setting criterion gain ratio, highest depth of the tree measured as 100. The minimal instance in leaf size = 2, and did not split the minimal size that is less than 5. DT showed an overall accuracy of 86%, precision 67%, and recall 70%.

Naive Bayes

Naive Bayes (NB) is essentially developed on the basis of base theorem that is used to test the theory of probability. It utilizes the theory of probability for data classification. The algorithm provides accuracy, precision, and recall rate were 87%, 66%, and 69% respectively.

Logistic Regression

Logistic Regression (LR) is a classification algorithm for Machine Learning employed to predict the probability of categorical dependent variables. We apply LR for skin cancer classification. We used ridge regularization (L2) and set the probability as 1. We achieved 88% overall accuracy by using logistic regression and 62% precision and 70% recall.

Convolutional Neural Network

A convolutional neural network is a kind of machine learning which consists of many neural network layers. Two different types of convolutional and pooling layer. The last stage is typically made by a fully connected layer. In our proposed method ReLU is an activation function and the learning rate is set 0.001. The maximum number of iterations was 200 and the number of hidden layer neurons was 100. CNN achieved the highest accuracy which is 89% and is best among all.

K-Nearest Neighbor

K-Nearest Neighbor is also one of the supervised learning methods being used for classification problems. KNN selects data on the base of the k value of the nearest neighbor then

decides the relevance with the given points. We apply K-Nearest Neighbor with k value 5 and 40% data split for training and 60% for testing then made a 2-fold of cross-validation. By using KNN we achieved a precision of 67%, and recall 66%, whereas we achieved an accuracy of 81% by using the KNN algorithm.

Result

The proposed work has been evaluated using core i3 with 4 GB of RAM and developed using an open-source machine learning, data visualization, and data mining toolkit known as Orange (Version: 3-3.26.0). ISIC dataset was used for classification which contains 10016 instances. Seven categories of skin cancer are used for classification. We have used multiple algorithms in our work for determining the cancer type classification, which is represented in Table 3.

Table 3. An average accuracy of all model

Model	Accuracy	C.Automation	F1	Precision	Recall
KNN	81%	0.68	0.67	0.67	0.68
Decision Tree	86%	0.70	0.68	0.67	0.70
Neural Network	89%	0.70	0.67	0.65	0.70
Naïve Bayes	87%	0.69	0.67	0.66	0.69
Logistic Regression	88%	0.70	0.65	0.62	0.70

Discussion

We have observed that CNN offers the highest accuracy of 89% and has recall rate of 0.70 which is similar decision tree and logistic regression. Meanwhile, all the other algorithms offer comparable accuracy. Logistic regression has the second highest accuracy of 88% while KNN provides the lowest accuracy for the classification, 81%. Moreover, CNN offers precession of 0.65 which is as average value when compared to the precision value of other algorithms. Decision tree and KNN provides the highest precision of 0.67. Furthermore, Decision tree, neural network and logistic regression has highest recall of 0.70. The F1 score is highest for decision tree which is perhaps due to the highest values attained by this classifier for recall and precision among other classifiers.

The specificity and sensitivity of all classifier based on seven categories of skin cancer can be noted from the graphs in Figure 3.2. The X-axis shows that FP (specificity) and Y-axis shows TP (sensitivity). The visualizing frequencies of sex vs dx (a) and age vs dx (b) can be observed from Figure 3.3. In (b) age contain < 42 to > = 67. Nv affects a patient of male and female as compared to other categories.

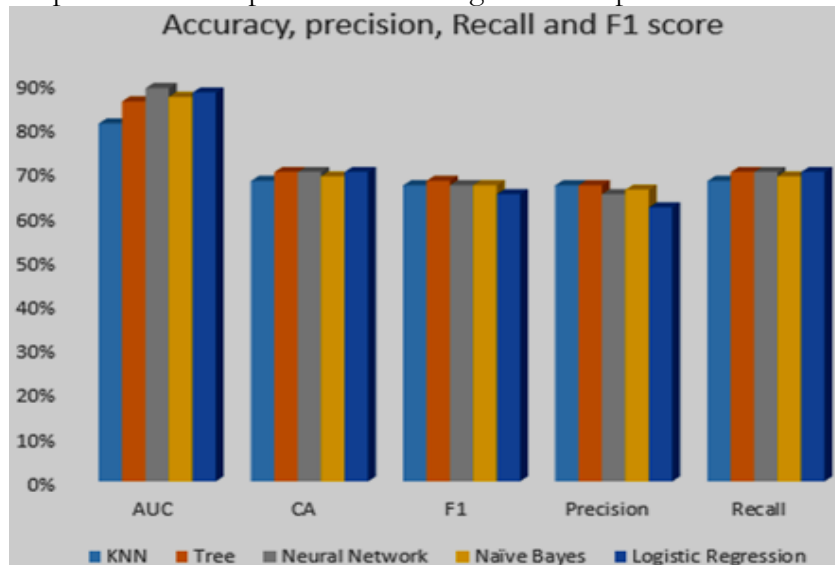The graphical representation of performance of algorithms is presented below in Figure 2.

Figure 2. An average accuracy of all models

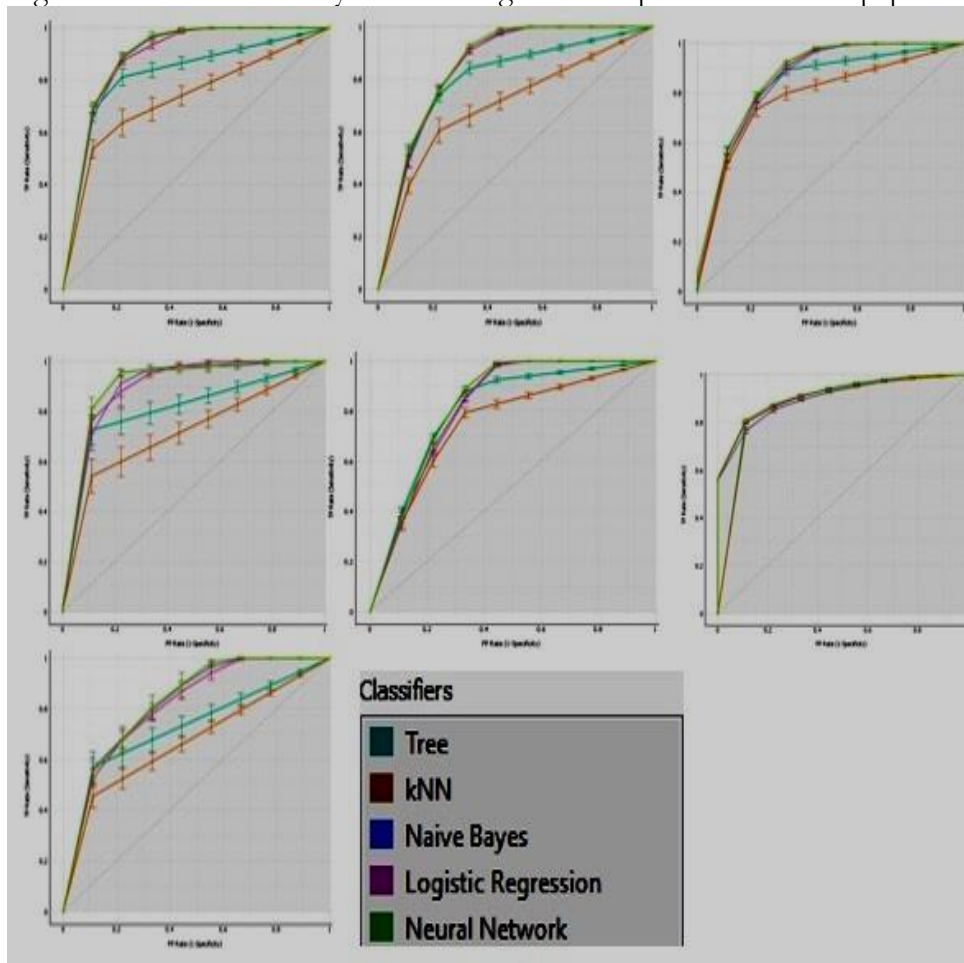Figure 3. shows ROC analysis for the algorithm implemented in this paper.

Figure 3. ROC Analysis

Figure 4. shows the visualizing frequencies of sex vs dx (a) and age vs dx (b).
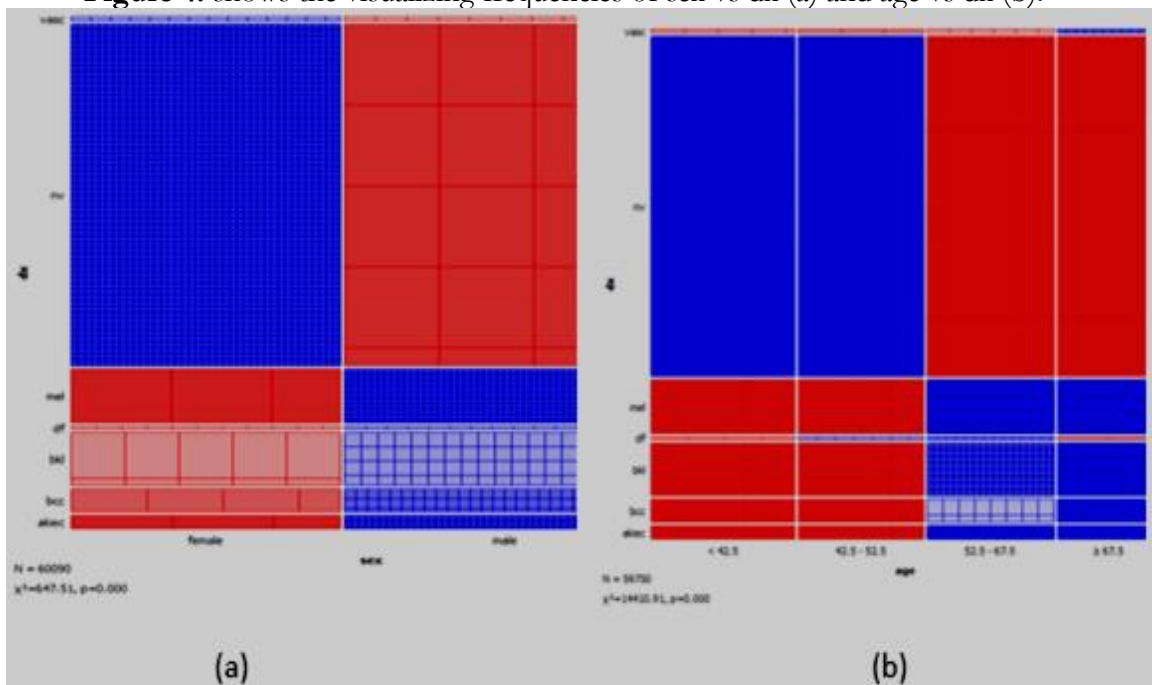

Figure 4. Graphical method for visualizing frequencies.

Conclusion

This research article discusses the classification technique for skin cancer. The investigation was conducted on the Intel Core i3 CPU having 4GB of RAM. Orange v3-3.26.0 has been used to examine & train the classification model. Few existing classification methods for the medical diagnosis of cancer patients have been discussed on basis of accuracy. Five machine learning technique was applied to the ISIC dataset. The results showed that CNN outperforms other models. We achieved 89% accuracy to classify 7 categories of skin cancer. In future, recent and improved dataset can be used for achieving even better accuracy. Furthermore, ensemble models of deep leaning algorithms can also be used to enhance the performance of classification model. In addition to this, the proven classification algorithms can be used for detection of less common skin cancers like Kaposi sarcoma, Merkel cell carcinoma, Sebaceous gland carcinoma.

References

[1] U. Jamil, A. Sajid, M. Hussain, O. Aldabbas, A. Alam, and M. U. Shafiq, "Melanoma segmentation using bio-medical image analysis for smarter mobile healthcare," J. Ambient Intell. Humaniz. Comput., 2019, vol. 10, no. 10, pp. 4099–4120.

[2] J. Burdick, O. Marques, J. Weinthal, and B. Furht, "Rethinking Skin Lesion Segmentation in a Convolutional Classifier," J. Digit. Imaging, vol. 31, 2018, no. 4, pp. 435–440.

[3] K. D. Miller et al., "Cancer treatment and survivorship statistics, 2016," CA. Cancer J. Clin.,2016, vol. 66, no. 4, pp. 271–289.

[4] T. Akram, M. A. Khan, M. Sharif, and M. Yasmin, "Skin lesion segmentation and recognition using multichannel saliency estimation and M-SVM on selected serially fused features," J. Ambient Intell. Humaniz. Comput., 2018.

[5] E. Nasr-Esfahani et al., "Melanoma detection by analysis of clinical images using convolutional neural network," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, 2016, vol. 2016-Octob, pp. 1373–1376.

[6] M. A. Khan, M. Y. Javed, M. Sharif, T. Saba, and A. Rehman, "Multi-model deep neural network-based features extraction and optimal selection approach for skin lesion classification," 2019 Int. Conf. Comput. Inf. Sci. ICCIS 2019, 2019, pp. 1–7.

[7] M. Properties and H. O. Mucosa, "Technical Report", 2015, pp. 2–3.

[8] P. Tschandl, C. Rosendahl, and H. Kittler, "Data Descriptor: The HAM 10000 dataset, a large collection of multi-sources dermatoscopic images of common pigmented skin lesions," Nat. Publ. Gr., 2018, vol. 5, pp. 1–9.

[9] P. Dubai, S. Bhatt, C. Joglekar, and S. Patii, "Skin cancer detection and classification," Proc. 2017 6th Int. Conf. Electr. Eng. Informatics Sustain. Soc. Through Digit. Innov. ICEEI 2017, 2018, vol. 2017-Novem, pp. 1–6.

[10] M. A. Farooq, M. A. M. Azhar, and R. H. Raza, "Automatic Lesion Detection System (ALDS) for Skin Cancer Classification Using SVM and Neural Classifiers," Proc. - 2016 IEEE 16th Int. Conf. Bioinforma. Bioeng. BIBE 2016, 2016, pp. 301–308.

[11] R. Ashraf, I. Kiran, T. Mahmood, A. Ur Rehman Butt, N. Razzaq, and Z. Farooq, "An efficient technique for skin cancer classification using deep learning," Proc. - 2020 23rd IEEE Int. Multi-Topic Conf. INMIC 2020, 2020.

[12] H. R. Mhaske and D. A. Phalke, "Melanoma skin cancer detection and classification based on supervised and unsupervised learning," 2013 Int. Conf. Circuits, Control. Commun. CCUBE 2013, pp. 1–5, 2013.

[13] K. Ramlakhan and Y. Shang, "A mobile automated skin lesion classification system," Proc. - Int. Conf. Tools with Artif. Intell. ICTAI, 2011, pp. 138–141.

[14] A. U. Rehman Butt, W. Ahmad, R. Ashraf, M. Asif, and S. A. Cheema, "Computer Aided Diagnosis (CAD) for Segmentation and Classification of Burnt Human skin," 1st Int. Conf. Electr. Commun. Comput. Eng. ICECCE 2019, 2019, no. July, pp. 24–25.

[15] C. L. Chang and C. H. Chen, "Applying decision tree and neural network to increase quality of dermatologic diagnosis," Expert Syst. Appl., 2009, vol. 36, no. 2 PART 2, pp. 4035–4041.

[16] R. Z. B and E. Conchon, "eHealth 360°," , 2017, vol. 181, pp. 407–418.

[17] A. Hekler et al., "Superior skin cancer classification by the combination of human and artificial intelligence," Eur. J. Cancer, 2019, vol. 120, pp. 114–121.

[18] F. A. Khan et al., "Computer-aided diagnosis for burnt skin images using deep convolutional neural network," Multimed. Tools Appl., 2020, vol. 79, no. 45–46, pp. 34545–34568.

[19] A. Ech-Cherif, M. Misbhauddin, and M. Ech-Cherif, "Deep Neural Network Based Mobile Dermoscopy Application for Triaging Skin Cancer Detection," 2nd Int. Conf. Comput. Appl. Inf. Secur. ICCAIS 2019, 2019, pp. 1–6.

[20] Z. Abbas, M. U. Rehman, S. Najam, and S. M. Danish Rizvi, "An Efficient Gray-Level Co-Occurrence Matrix (GLCM) based Approach Towards Classification of Skin Lesion," Proc. - 2019 Amity Int. Conf. Artif. Intell. AICAI 2019, 2019, pp. 317–320.

[21] F. A. Khan, A. U. Rehman Butt, M. Asif, H. Aljuaid, A. Adnan, S. Shaheen, "Burnt Human Skin Segmentation and Depth Classification Using Deep Convolutional Neural Network (DCNN)," J. of Medical Imaging and Health Informatics, 2020, vol. 10, no. 10, pp. 2421-2429.

[22] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinger, "Skin Lesion Classification Using Hybrid Deep Neural Networks", Institute for Pathophysiology and Allergy Research, Medical University of Vienna, Austria Department of Research and Development, TissueGnostics GmbH , Austria Department of Computer Science , Loughborough University , U . K . Department of Biomedical, 2019, pp. 1229–1233.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, 2017, vol. 60, no. 6, pp. 84–90.

[24] F. O. R. L. Arge and C. I. Mage, "V d c n l -s i r," , 2015, pp. 1–14.

[25] V. Sangeetha and K. J. R. Prasad, "Syntheses of novel derivatives of 2-acetylfuro[2,3-a]carbazoles, benzo[1,2-b]-1,4-thiazepino[2,3-a]carbazoles and 1-acetyloxycarbazole-2-carbaldehydes," Indian J. Chem. - Sect. B Org. Med. Chem., 2006, 1954, vol. 45, no. 8, pp. 1951.