iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 10 **Issue:** VII **Month of publication:** July 2022

**DOI:** https://doi.org/10.22214/ijraset.2022.45182

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 10 Issue VII July 2022- Available at www.ijraset.com*

# A Review of the Use of GUI and Deep Neural Networks for Skin Cancer Classification

Mohammad Altaf [1], Dr. Gurinder Kaur Sodhi[2]
[1]*M. Tech Scholar,* [2]*Assistant Professor, Department of ECE Engineering, Desh Bhagat University, Mandi Gobingarh, Punjab, India*

***Abstract:*** ***Modern classifiers based on convolutional neural networks (CNNs) have been shown to categorise photos of skin cancer on par with dermatologists, potentially enabling lifesaving and rapid diagnosis even outside of the hospital via mobile app installation.***
***There is currently no overview of current research in this field that we are aware of. We searched the Google Scholar, PubMed, Medline, Science Direct, and Web of Science databases for systematic reviews and original research publications published in English.***
***In this evaluation, only publications with appropriate scientific proceedings were considered. CNNs perform admirably as cutting-edge skin lesion classifiers.***
***Unfortunately, comparing different categorization algorithms is difficult because some approaches use non-public datasets for training and/or testing, which makes repeatability problematic. Future papers should use publicly available benchmarks and fully describe the training methodologies used to allow for comparison.***
***Keywords: skin cancer; convolutional neural networks; lesion classification; deep learning***

## I. INTRODUCTION

From 2008 to 2018, the yearly incidence of melanoma cases increased by 53%, owing in part to increased UV exposure [1,2]. Despite the fact that melanoma is one of the deadliest types of skin cancer, early detection can result in a high chance of survival. The first step in diagnosing a malignant lesion is a visual examination of the suspect skin region by a dermatologist. Because several lesion types are identical, accurate diagnosis is critical; additionally, diagnostic accuracy is strongly related to the physician's professional experience [3].

Dermatologists have a melanoma diagnosis accuracy rate of 65 to 80 percent without the use of additional technology [4]. Dermatoscopic images captured with a special high-resolution and magnifying camera are added to the visual assessment in questionable situations.

Throughout the filming, the lighting is controlled, and a filter is used to reduce reflections on the skin, allowing deeper skin layers to be seen. With this technological assistance, the accuracy of skin lesion diagnosis can be increased by 49 percent. [5]. Dermatologists can identify melanoma with an absolute accuracy of 75 percent to 84 percent using a combination of ocular inspection and dermatoscopic imaging [6,7].

The challenge of categorizing skin lesions has been a focus of machine learning community for some time. Automated lesion categorization, through the installation of applications on mobile devices, can both assist physicians in their daily clinical practise and provide rapid and affordable access to vital diagnoses even outside the hospital [8,9].

Prior to 2016, the majority of machine learning research followed the standard workflow of preprocessing, segmentation, feature extraction, and classification [9-11]. However, feature extraction, in particular, requires a high level of application-specific expertise, and selecting appropriate features takes time.Furthermore, mistakes, information loss in earliest steps of processing have significant impact on categorization quality. Poor segmentation, , frequently results in poor feature extraction and, as a result, low classification accuracy. There was a shift in the study of lesion classification techniques in 2016. This shift is evidenced by the approaches presented at the 2016 International Symposium on Biomedical Imaging (ISBI) [12].Traditional machine learning methods were not used by any of the 25 competing teams; Instead, they all employed a deep learning technique known as convolutional neural networks (CNNs) [13]. This is the first comprehensive review of cutting-edge research on using CNNs to classify skin lesions.

©IJRASET: All Rights are Reserved 

The approaches given are divided into two categories: those that employ a CNN just as a feature extractor and those that use it for end-to-end learning. The paper's closure highlights why comparing the methodologies offered is challenging and what difficulties must be solved in the future.

## II. METHODOLOGY

Techniques for Searching Google Scholar, PubMed, Medline, ScienceDirect, and Web of Science databases were searched for systematic reviews and original research publications published in English. Search terms included convolutional neural networks, deep learning, skin cancer, lesions, melanoma, and carcinoma.

In this evaluation, only publications with appropriate scientific proceedings were considered. Choosing a Study We only looked at approaches for classifying skin lesions. Methods such as Demyanov et al [14] that use CNN exclusively for lesion segmentation or dermatoscopic pattern categorization are not examined in this work. Furthermore, this evaluation includes only studies that demonstrate a suitable scientific procedure. This last requirement is providing the techniques in an intelligible way and sufficiently addressing the outcomes.

Carcagn et al [15], for example, and Dorj et al [6] do not evaluate works in which the genesis of the performance is not credible. Convolutional Neural Networks (CNNs) are a type of neural network that makes use of convolutions. CNNs are specialised neural networks with a one-of-a-kind design that have been shown to be extremely effective in image recognition and classification [7]. CNNs have been demonstrated to be superior to humans in recognising faces, objects, and traffic signals, and can thus be found in robots and self-driving cars. CNNs are supervised learning methods that teach themselves using labelled data. In essence, CNNs learn relationship between the input objects and the class labels, and they are composed of two parts: hidden layers where the features are extracted, and fully connected layers at the end of the processing where the actual classification task is performed. Unlike traditional neural networks, the hidden layers of a CNN have a distinct design. Each layer of a normal neural network is made up of a group of neurons, and the neuron in each layer is connected to the neuron in the layer before it.

The design of hidden layers in a CNN varies slightly. The neurons in a layer are not linked to all of the neurons in the preceding layer; rather, they are linked to a subset of them. Translation-invariant properties are obtained by restricting connections to local connections and adding additional pooling layers that aggregate local neuron outputs into a single value. As a result, training procedure is streamlined, and model's complexity is reduced.Convolutional Neural Networks-based Classifiers for Skin Lesions Currently Available This section describes the individual CNN approaches used to classify skin lesions. CNNs can categorise cutaneous lesions in two ways. On the one hand, a CNN pretrained on another large dataset, such as ImageNet [8,] can be used as a feature extractor. A different classifier, such as k-nearest neighbours, support vector machines, or artificial neural networks,does the classification in this situation. A CNN, on the other hand, may learn the association between raw pixel data and class labels directly via complete learning. . In opposite to traditional machine learning workflow, feature extraction is now regarded an inherent aspect of classification rather than a distinct, self sufficient processing step. I If the CNN is trained using end-to-end learning, the study can be divided into two parts: learning the model from scratch and learning the model through transfer learning.

Figure 1 depicts a summary of the offered CNN approaches. The availability of adequate training data labeled with classes is basic condition for the efficient training of deep CNN models. Otherwise, there is a risk of overfitting the neural network, which could lead to failure. the network's generalization characteristic for unknown input data being inadequate. The categorization of skin lesions has a relatively limited quantity of publically available data. Almost every published approach makes use of datasets with fewer per training class more than 1000 training data points In contrast, well-known CNN models for image classification, such as AlexNet [8, VGG [12], GoogLeNet [20], or ResNet [11], are trained on the massive image database ImageNet, with each training class consisting of over 1000 images. Even if only a small amount of data is available for training, sophisticated CNN models with several million free parameters can be used for classification using a training method known as transfer learning. In this example, the CNN is pretrained on a large dataset such as ImageNet before being used to initialise the CNN for the task at hand.

The number of training classes in the real classification challenge is used to modify the pretrained CNN model's last fully connected layer. The pretrained CNN's weights can then be fine-tuned in one of two ways: fine-tune all layers of the CNN or freeze some of the front layers due to overfitting issues and fine-tune only selected rear layers of the network.

The logic behind this method is that the top layers of a CNN contain more generic features (e.g., edge or color-blob detectors) that can be used for a variety of tasks, whereas the rear layers of the CNN become increasingly specific to the characteristics of the classes in the original dataset.Statistical quantities for evaluating different classifiers are introduced in the following discussion. Following that, strategies for using the CNN as a feature extractor are discussed. The last paragraph gives an overview of strategies used in end-to-end learning using CNN.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

## III. CLASSIFIERS' PERFORMANCE

Each object is assigned to a class by a classifier. In general, this assignment isn't flawless, and objects may be allocated to the incorrect class.

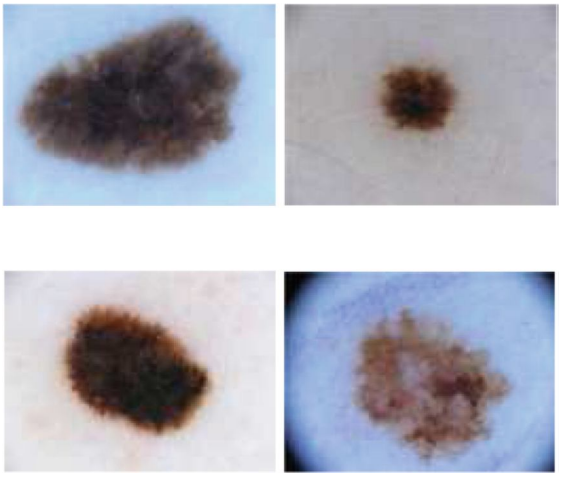


Figure 1 Illustration of Benign v/s Malignant melanoma

The true class of objects must be known in order to assess a classifier. The classifier's assigned class is compared to the real class to determine the classification quality. This allows the items to be separated into the four subgroups shown below.:

*1)* True positive (TP): the classifier predicts the positive class correctly.
*2)* True negative (TN): the classifier predicts the negative class correctly.
*3)* False positive (FP): the classifier predicts the positive class incorrectly.

### *A. FN (False Negative)*

The classifier guesses negative class erroneously. Statistical values for the classifier may now be determined based on the cardinality of these subgroups. Accuracy is a common and commonly used metric, However, it is only useful if the different classes in the dataset are distributed fairly evenly..

(TP + TN)/(TP + TN + FP + FN) is formula for calculating accuracy. It indicates the percentage of items that have been categorised properly. Sensitivity and specificity are two more key measures that may be used even if different classes are not evenly distributed. TP/(TP + FN) is used to compute sensitivity, which is the ratio of properly identified positive items to the total number of positive objects in the dataset. Specificity is determined as TN/( the number of positive objects in the dataset as a whole The percentage of negative items correctly classified as negative out of the total number of negative objects in the provided dataset is calculated as TN/(TN + FP).TN + FP) as the percentage of negative items properly categorized as negative out of the total number of negative objects in the provided dataset.

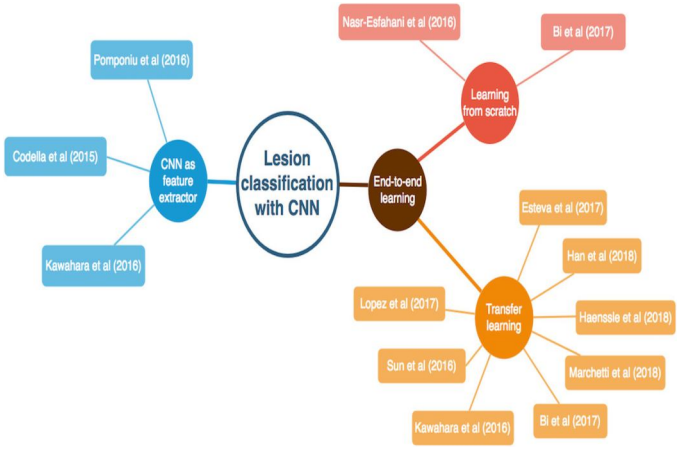



Figure 2. CNN & Associated Classification

A binary classifier's output is interpreted as a probability distribution over the classes. Normally, items with an output value greater than.5 are assigned to the positive class, while objects with an output value less than.5 are assigned to the negative class in a binary classifier. A different method based on the receiver operating characteristic is used (ROC). The categorization threshold is varied between 0 and 1, and the sensitivity and specificity are calculated for each threshold. The ROC curve is created by plotting sensitivity versus 1-specificity and can be used to evaluate the classifier. The further the ROC curve deviates from the diagonal, the better the classifier. The curve's area under the curve is a good overall metric (AUC)

## IV. SIMULATION AND RESULTS

In this classifier, the Convolutional Neural Network is used as a feature extractor. It may be used in classification by deleting the fully connected layers of a CNN that was pretrained with a large dataset. ImageNet is used for skin lesion categorization pretraining. Despite having been learned in a nonmedical image domain, the learned features are of sufficient quality to classify lesions [2]. Pomponiu et al [3] used only 399 photos from a regular camera to differentiate between melanomas and benign nevi. The first step was to complete data preparation and augmentation. Then, to extract representational characteristics, a pretrained AlexNet was used. After that, the lesions were classified using cosine distance metrics and a k-nearest-neighbor classifier. The technique was tested using only cross-validation; no independent test dataset was used.

Aside from the lack of an independent test dataset, the region of interest for each skin lesion must be manually marked. For feature extraction, Codella et al [4] also used an AlexNet model. Unlike Gutman et al [12], a total of 2624 dermatoscopic images from the International Skin Imaging Collaboration (ISIC) database were used to differentiate melanoma from nonmelanoma lesions or melanoma from atypical nevi. In addition to the modified AlexNet outputs, the authors used low-level handmade features, sparse coding features, a deep residual network, and a convolutional U-network. To classify the data based on all of these attributes, a support vector machine was used.

The authors obtained an accuracy of 93.1 percent, a sensitivity of 94.9 percent, and a specificity of 92.8 percent for distinguishing melanoma from nonmelanoma. The more difficult distinction between melanomas and atypical nevi was found to have an accuracy of 73.9 percent, a sensitivity of 73.8 percent, and a specificity of 74.3 percent. The authors also demonstrated that deep features outperform low-level handmade features in classifiers.

A linear classifier was employed by Kawahara et al [25] to categorize 10 distinct skin lesions. For feature extraction, an AlexNet with a convolutional layer in place of the last fully connected layer was also used. This significantly modified AlexNet was tested using the public Dermofit Image Library, which contains 1300 clinical photographs of 10 skin lesions. An accuracy of 81.8 percent was obtained based on the entire dataset of ten distinct types of skin lesions. Transfer Learning and Convolutional Neural Network Model Training Transfer learning is a popular approach for skin lesion categorization because publicly accessible datasets are limited. As a result, all of these studies use the ImageNet dataset to pretrain a CNN, which is then fine-tuned to the classification task at hand. Esteva and colleagues [6] published a seminal paper. For the first time, a CNN model was trained using a large amount of data: 129,450 images, 3374 of which were taken with dermatoscopic instruments and represented 2032 different skin lesions. Two binary classification difficulties were investigated: keratinocyte carcinomas versus benign seborrheic keratosis and malignant melanomas versus benign nevi.Both clinical and dermatoscopic pictures were subjected to the final classification distinction. For the categorization, the scientists utilized a GoogLeNet Inception v3 model that was pre-trained with ImageNet, a big picture database. . The CNN model was fine-tuned to categorise skin lesions using transfer learning. A distinguishing feature of this technique is the use of a novel tree-structured disease taxonomy, in which individual illnesses serve as the tree's leaves.Individual disorders that are aesthetically and clinically similar are grouped together in the inner nodes. The output of the CNN is probability distribution with over 757 training classes, rather than a two-dimensional vector. The probabilities of a coarser lesion class (ie, an inner node at a higher level in the tree) are established by adding the probabilities of the child nodes of a coarser lesion class (ie, an inner node at a higher level in the tree). The authors demonstrate that a CNN that has been trained for finer classes outperforms a CNN that has been trained for the various classes that are relevant to the assessment problem. The trained CNN achieved a ROC AUC of.96 for carcinomas, a ROC AUC of.96 for melanomas, and a ROC AUC of.94 for melanomas categorised solely by dermatoscopic images using completely biopsy-proofed test data.

Esteva et al [26] used a strategy similar to Haenssle et al [3]. A GoogLeNet Inception v3 model for skin lesion classification was created using transfer learning, with weights fine-tuned in all layers. The study was limited to dermatoscopic images of melanoma vs benign nevi, and the AUC ROC for this job was.86 (Esteva et al [6]: .94). The number of training data points was not specified, and not all of the data was biopsied. However, the study included the most dermatologists (n=58) to date and was the first to demonstrate that having more clinical data increases dermatologists' sensitivity and specificity.

Han et al. [7] are notable for their scientific transparency, having made their computer algorithm openly available for external testing. Based on clinical photos, the researchers demonstrated a classifier for 12 distinct skin conditions. They created a ResNet model using 19,398 training photos to fine-tune it. The CNN model attained ROC AUCs of.96,.83,.82, and.96 to detect basal cell carcinoma, squamous cell carcinoma, intraepithelial carcinoma, and melanoma,respectively, using the publically available Asan dataset..

Marchetti et al [13] present a CNN ensemble for distinguishing melanomas from nevi or lentigines. They combined all automated predictions from the ISBI 2016 Challenge's 25 teams into a single classification result using five methods. To accomplish this, they used two nonlearning methodologies and three machine learning methods. Following training with 279 dermatoscopic images from the ISBI 2016 Challenge dataset, the fusion algorithms were evaluated with 100 additional dermatoscopic images from the ISBI 2016 Challenge dataset. Greedy fusion was the best-performing ensemble approach in terms of average accuracy, with a sensitivity of 58% and a specificity of 88%.

Bi et al [8] proposed a different sort of CNN ensemble. Using dermatoscopic pictures, they looked at how to classify melanomas, seborrheic keratosis, and nevi. Instead of training multiple CNNs for the same classification problem, they fine-tuned a pre-trained CNN to train three ResNets for different problems: one for the original three-class problem and two binary classifiers (melanoma versus both other lesion classes or seborrheic carcinoma versus both other lesion classes). The test used 150 dermatoscopic images and produced ROC AUCs of.854 for melanomas,.976 for seborrheic carcinomas, and.915 for all classes.

. Kawahara et al [9] provide unique design for a CNN ensemble. The CNN was made up of many components, each of which looked at same image at a different resolution. The outputs from numerous resolutions are then combined into a single layer by an end portion. The CNN detects interactions across a wide range of image resolutions, and end-to-end learning optimises the weighting parameters. The algorithm had an average classification accuracy of 79.5 percent in the public Dermofit Image Library.

Sun et al [10], like Esteva et al [26], proposed a classifier with 198 highly precisely specified training classes. For this classification challenge, 6584 clinical images from the publicly available image collection DermQuest were used for training and testing, and the performance of the CNN models CaffeNet and VGGNet was assessed. The best average accuracy across all 198 classes was 50.27 percent for a pretrained VGGNet that was refined by fine-tuning the weighting parameters.

Lopez et al. [11] used modified VGGNet to rank melanoma against nevi or lentigines using dermatoscopic pictures and a modified VGGNet. The accuracy of CNN trained from scratch, a pretrained CNN with transfer learning and frozen layers, and pretrained CNN with transfer learning and fine-tuning of weighting parameters was compared by the authors. With 379 photos from the ISBI 2016 Challenge dataset, all three configurations were examined, with the last one achieving the maximum accuracy of 81.33 percent.

## V. DISCUSSION

One challenge with comparing skin lesion categorization systems is that the separate works' considered problem formulations differ, sometimes only slightly. This is true not just for training classes that are being examined and the data that is being used, but also for the statistical numbers that are being provided. In addition to publicly accessible data archives, several studies employ nonpublic skin clinic archives [3,6]. This makes replicating the results considerably more challenging. Since 2016, ISIC Melanoma Project has worked to rectify this by creating a publicly available collection of dermatoscopic skin lesion pictures that can be used as a baseline for education and research [12].. They also introduced an annual challenge in which participants must solve a clearly stated problem. More wsork comparing itself to this standard would be useful in order to improve the ranking of processes in the state of research. Another major research topic is the creation of massive public picture archives containing photographs that are as representative of the global population as feasible [13]. Skin lesions from light-skinned persons are mostly found in existing photographic collections. For example, the photographs in the ISIC database are mostly from the United States, Europe, and Australia. The CNN should learn to abstract from the color of skin in order to accomplish good categorization for dark-skinned persons as well. This, however, can only happen if it sees enough dark-skinned humans throughout training. Clinical data (e.g., age, gender, ethnicity, skin type, and anatomic location) might be used as inputs for the classifiers to increase classification quality. As Haenssle et al [3] demonstrate, this additional knowledge helps dermatologists make better decisions. These considerations should be taken into account in future development.

## VI. CONCLUSION

Comparing the efficacy of published categorization findings appears difficult, if not impossible, given that the majority of authors use private datasets for training or testing. Future research must contrast readily available data. Three CNNs—one trained from start, one pretrained with transfer learning and frozen layers, and one pretrained with transfer learning and adjusting the weighting parameters—were each subjected to a precision test. benchmarks and fully describe the training methods used.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

## REFERENCES

[1] Nami N, Giannini E, Burroni M, Fimiani M, Rubegni P. Teledermatology: State-of-the-art and future perspectives. Expert Rev Dermatol 2014 Jan 10;7(1):1-3. [doi: 10.1586/edm.11.79]

[2] Fabbrocini G, Triassi M, Mauriello MC, Torre G, Annunziata MC, De Vita V, et al. Epidemiology of skin cancer: Role of some environmental factors. Cancers (Basel) 2010 Nov 24;2(4):1980-1989 [FREE Full text] [doi: 10.3390/cancers2041980] [Medline: 24281212]

[3] Haenssle H, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, Reader Study Level-I and Level-II Groups. Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 2018 Aug 01;29(8):1836-1842. [doi: 10.1093/annonc/mdy166] [Medline: 29846502]

[4] Argenziano G, Soyer HP. Dermoscopy of pigmented skin lesions: A valuable tool for early diagnosis of melanoma. Lancet Oncol 2001 Jul;2(7):443-449. [Medline: 11905739]

[5] Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. Lancet Oncol 2002 Mar;3(3):159-165. [Medline: 11902502]

[6] Ali ARA, Deserno TM. A systematic review of automated melanoma detection in dermatoscopic images and its ground truth data. Proc SPIE Int Soc Opt Eng 2012 Feb 28;8318:1-6. [doi: 10.1117/12.912389]

[7] Fabbrocini G, De Vita V, Pastore F, D'Arco V, Mazzella C, Annunziata MC, et al. Teledermatology: From prevention to diagnosis of nonmelanoma and melanoma skin cancer. Int J Telemed Appl 2011 Sep 01;2011(17):125762 [FREE Full text] [doi: 10.1155/2011/125762] [Medline: 21776252]

[8] Foraker R, Kite B, Kelley MM, Lai AM, Roth C, Lopetegui MA, et al. EHR-based visualization tool: Adoption rates, satisfaction, and patient outcomes. EGEMS (Wash DC) 2015;3(2):1159 [FREE Full text] [doi: 10.13063/2327-9214.1159] [Medline: 26290891]

[9] Fabbrocini G, Betta G, Di Leo G, Liguori C, Paolillo A, Pietrosanto A, et al. Epiluminescence image processing for melanocytic skin lesion diagnosis based on 7-point check-list: A preliminary discussion on three parameters. Open Dermatol J 2010 Jan 01;4(1):110-115. [doi: 10.2174/1874372201004010110]

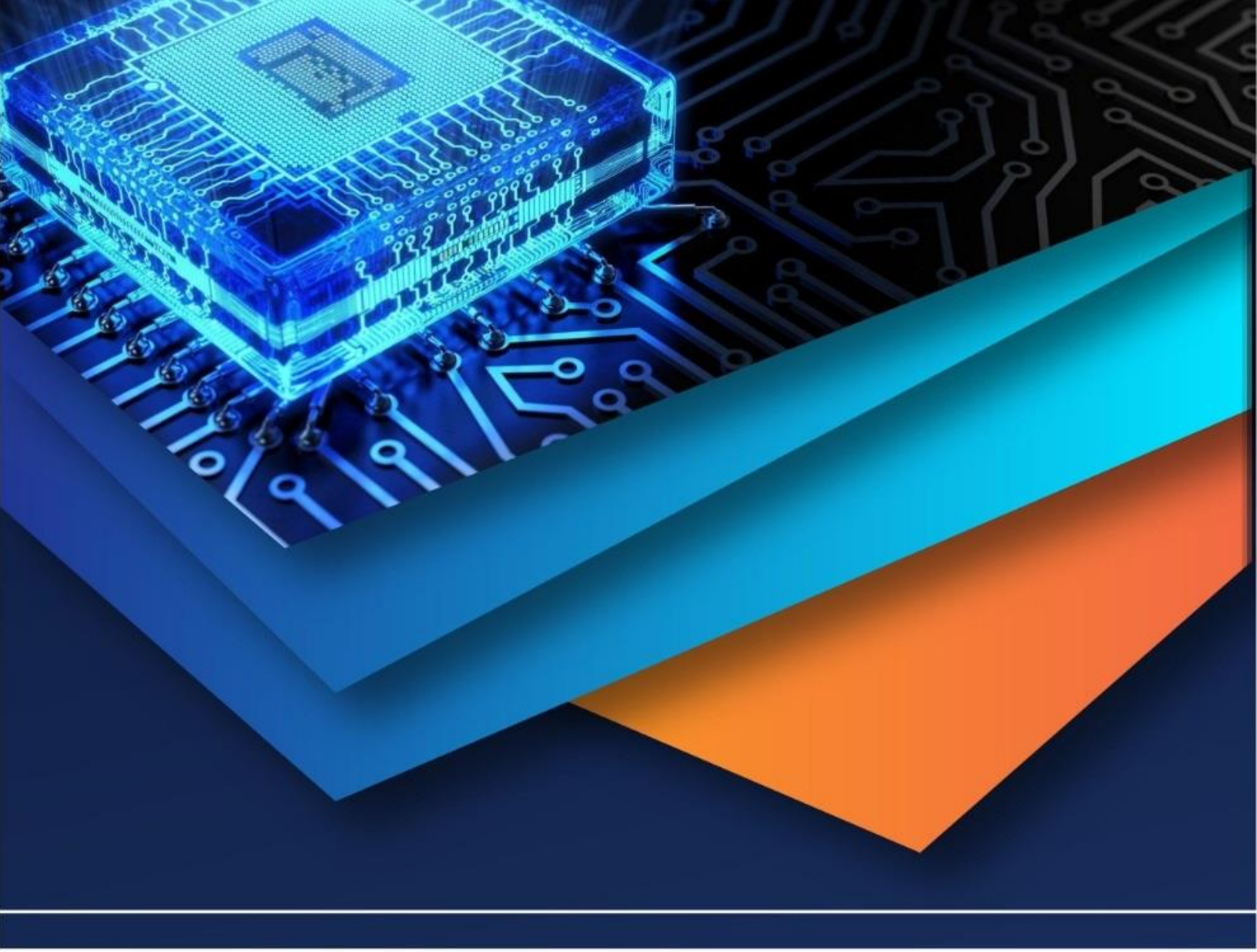
[10] Hart PE, Stork DG, Duda RO. Pattern Classification. 2nd edition. Hoboken, NJ: John Wiley & Sons; 2000.

[11] Oliveira RB, Papa JP, Pereira AS, Tavares JMRS. Computational methods for pigmented skin lesion classification in images: Review and future trends. Neural Comput Appl 2016 Jul 15;29(3):613-636. [doi: 10.1007/s00521-016-2482-6]

[12] Gutman D, Codella NCF, Celebi E, Helba B, Marchetti M, Mishra N, et al. arXiv. 2016 May 04. Skin lesion analysis toward melanoma detection: A challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC) URL: https://arxiv.org/pdf/1605.01397 [accessed 2018-10-06] [WebCite Cache ID 72yoIScsz]

[13] Marchetti MA, Codella NCF, Dusza SW, Gutman DA, Helba B, Kalloo A, International Skin Imaging Collaboration. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. J Am Acad Dermatol 2018 Dec;78(2):270-277.e1. [doi: 10.1016/j.jaad.2017.08.016] [Medline: 28969863]

[14] Demyanov S, Chakravorty R, Abedini M, Halpern A, Garnavi R. Classification of dermoscopy patterns using deep convolutional neural networks. In: Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI).: IEEE; 2016 Presented at: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI); April 13-16, 2016; Prague, Czech Republic p. 2-12.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)