IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Malignant Melanoma Classification Using Deep Learning: Datasets, Performance Measurements, Challenges and Opportunities

Ahmad Naeem[1], Muhammad Shoaib Farooq[1], (MEMBER, IEEE), Adel Khelifi[2], Adnan Abid[1], (Member, IEEE)

[1]Department of Computer Science, University of Management and Technology, Lahore 54000, Pakistan
[2]Abu Dhabi University, Abu Dhabi, United Arab Emirates.
Corresponding author: M. S. Farooq (shoaib.farooq@umt.edu.pk).

ABSTRACT Melanoma remains the most harmful form of skin cancer. Convolutional neural network (CNN) based classifiers have become the best choice for melanoma detection in the recent era. The research has indicated that classifiers based on CNN classify skin cancer images equivalent to dermatologists, which has allowed a quick and life-saving diagnosis. This study provides a systematic literature review of the latest research on melanoma classification using CNN. We restrict our study to the binary classification of melanoma. In particular, this research discusses the CNN classifiers and compares the accuracies of these classifiers when tested on non-published datasets. We conducted a systematic review of existing literature, identifying the literature through a systematic search of the IEEE, Medline, ACM, Springer, Elsevier, and Wiley databases. A total of 5112 studies were identified out of which 55 well-reputed studies were selected. The main objective of this study is to collect state of the art research which identify the recent research trends, challenges and opportunities for melanoma diagnosis and investigate the existing solutions for the diagnosis of melanoma detection using deep learning. Moreover, proposed taxonomy for melanoma detection has been presented that summarizes the broad variety of existing melanoma detection solutions. Lastly, proposed model, challenges and opportunities have been presented which helps the researchers in the domain of melanoma detection.

INDEX TERMS Deep learning, CNN, Skin Cancer, Melanoma, Detection, Diagnosis

I. INTRODUCTION

The report of the World Health Organization (WHO) shows that cancer is one of the world's leading causes of death [2]. It predicts that in the next two decades, the number of people diagnosed with cancer will be double [7][85]. Death rates caused by cancer can be reduced if the cancer is detected and treated in the early stages [12]. Investing research effort in the development of early cancer detection strategies is the primary concern of researchers.

The most harmful form of skin cancer is melanoma. It has been ranked at the ninth position among the most common cancer. More than 132,000 cases have been diagnosed every year [3]. A report published in 2019 by The American Cancer Organization estimates that 192,310 people were diagnosed with melanoma in U.S [12]. Over the past 30 years, melanoma cases have been gradually increasing like other cancer cases. A minor surgery can increase the chances of recovery if the melanoma is diagnosed in the early stages [8]. Dermoscopy is one of the dermatologists' most popular imaging techniques. It magnifies the skin lesion surface and its structure became more visible to the dermatologist for an examination. However, this technique can only be used effectively by trained physicians, because it is totally based on the practitioner's visual acuity and experience [8]. These challenges motivate the research community to develop new techniques for visualization and diagnosing of melanoma. Computer-aided diagnosis (CAD) system assists in the diagnosing of melanoma cancer. The CAD tool provides a user-friendly environment for non-experienced dermatologists [1]. Evidence of CAD diagnostic tool can be used as a second opinion in diagnosing melanoma cancer.

An Expert dermatologist can achieve an average accuracy of 65% to 75% by using Dermoscopy [1]. Moreover, accuracies can be further improved for suspicious cases, by using a camera with a special high-resolution and a magnifying lens to capture dermoscopic images for visual inspection. The

light is regulated during the image capturing process which improves the visibility of deep layers of skin. The accuracy of a skin cancer diagnosis can be improved by an estimated 50 % with this technological support [2]. Dermatologist precision improved by combining visual perception and dermoscopic images [6]. Automated identification of melanoma can assist doctors in their day-to-day clinical routine by allowing quick and economical access to life-saving diagnoses [10]. The aforementioned issues and challenges emphasize the Machine learning community to put a primary focus on the classification of melanoma [4]. Machine learning implements statistical algorithms for learning purposes, which first train the data, then test that data, by adopting their parameters [11]. Prior to 2016, the study mainly focused on the classical machine learning workflow which consists of pre-processing, segmentation, extraction of features, and classification [5]. Moreover, a decent level of expertise necessarily required for the extraction of features from cancerous images. A poor segmentation can lead to bad feature selection which decreases the accuracy of classification [6]. In 2016 a transition occurs in the field of skin lesion classification techniques. The approaches presented to the International Symposium on Biomedical Imaging (ISBI) 2016 indicate this transition. The research contributors didn't apply conventional machine learning algorithms rather they all used a technique of deep learning: convolution neural networks (CNNs)[6][86].

Zilong *et al.* [9] review the few techniques for skin cancer detection using images. This study was not limited to melanoma detection, it provides an over about different types of cancers that use images for their diagnosis. Few studies for melanoma diagnosis were included in this paper. In addition to this, a review presented by Sultana *et al.* [15] provides an overview of a few deep learning methods and discusses some benchmark datasets. However, it did not give information about non published, internet collected and combine datasets. Another systematic literature on classifying a skin lesion using CNNs by Brinker *et al.* [36] has been done, it provides a brief overview of the different methods of deep learning. However, the major limitation of this study was that it did not provide any information about the datasets.

We have made a search string to collect relevant research available in the field of deep learning for melanoma detection. For this purpose, papers from renowned journals, conferences, and symposiums were searched. We select 55 research papers out of 5112 papers for further review and analysis. The finalized papers were tested empirically and qualitatively across multiple aspects.

In this paper, we have discussed different CNNs based models and benchmark, as well as unpublished internet collected and combined benchmark datasets for melanoma detection. The key emphasis of this work was to provide a complete review of deep learning literature in the detection of skin cancer. A lot of research has been done over the past few years on automatic melanoma detection using deep learning. Collecting, evaluating, classifying and summarizing the state of the art work remains critical. This SLR gives an overview of the proposed taxonomy and model for melanoma detection by exploring the different types of deep learning techniques. Moreover, this review identifies the recent research trends, open issues and challenges in the field of melanoma diagnosis. In addition to this, it examines the various public, non-public, internet collected and combine datasets. Furthermore, this study categorizes the main shortcomings of existing solutions and point out the research areas where further improvement should be done in the future.

This paper has been divided into six sections. We present an Introduction and Objective in section I. Section II introduces the research method for performing systematic literature review by unfolding research questions, literature review scope, information source, search criteria, search string, information extraction process and selection/evaluation criteria of the study. In Section III, the results of every selected paper have been discussed. Moreover, deep learning-based classifiers, their performances and datasets were discussed in this section. In section IV, discuss the results taxonomy for melanoma detection and in addition to this, a model has been proposed by analyzing selected papers, which summarizes our findings. In section V, challenges and opportunities have been discussed. Section VI provides the conclusion of this study.

II. Research Methodology

The purpose of this systematic review was to identify the finest accessible classifiers, methods and datasets rely on deep learning for the detection of melanoma. The systematic literature review process helps to identify and analyze the available research in the relevant study domain [14]. The findings of the study provide scientific evidence through the classification of relevant studies. This systematical research methodology has recommended by Kitchenham et al. [13] and the choice of method in the prime study was followed by [14]. This study includes publications that were found in particular sources and use techniques relevant to CNN or pre-trained models based on CNN for the detection of melanoma.

A. REVIEW PROCEDURE

The initial step towards a systematic review was to establish a review procedure. The Systematic Literature Review (SLR) procedure helps to identify a study strategy as a search method for extracting the relevant literature. This procedure provides criteria for review of literature which includes research questions, search string, information sources such as conferences, journals and symposiums, requirements for inclusion and exclusion. The systematic literature review process follows the steps illustrated in Fig. 1.

TABLE I

Research Questions

	Research Question Statement	Objective of questions
Q1	What type of deep learning-based classifiers are used for diagnosing melanoma?	The motivation of this research question is to classify the state of art research for the detection of melanoma.
Q1.1	What type of challenges and opportunities exist in deep learning algorithms for the diagnosis of melanoma?	This question aims at identifying the strengths and limitation of existing deep learning methods
Q2	What performance metrics are used by classification methods in diagnosing melanoma?	The accuracy of different studies is calculated by the evaluation metrics like True Positive (TP) also known as Sensitivity or Recall, Misclassification Rate (Error Rate), False Positive (FP), True Negative (TN) also called Specificity, False Negative (FN), Precision, ROC.
Q2.1	What are the accuracies of classifiers? How taxonomy is proposed by considering these accuracies?	A taxonomy was proposed to emphasize effective methods for deep learning systems for the diagnosis of melanoma
Q3	What type of datasets are available for diagnosing melanoma?	This research addresses the availability of benchmark datasets as well as non-publish, non-listed and internet collected dataset. The number of images available for training and testing in each dataset is also investigated
Q3.1	What is the reliability of datasets?	

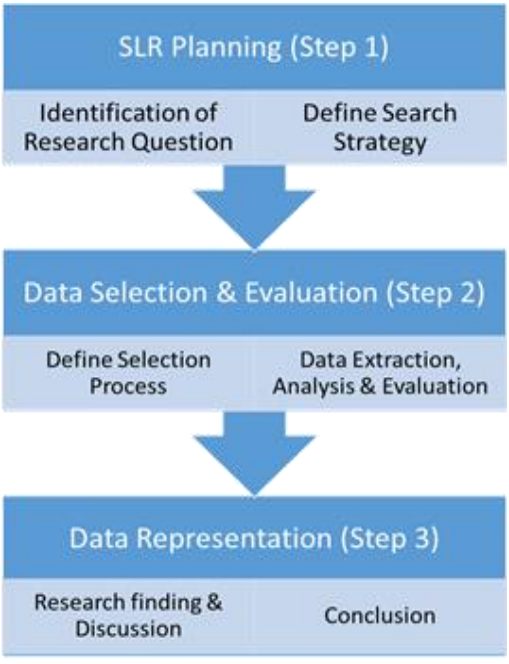

Fig. 1. Review procedure

1) RESEARCH OBJECTIVES

The primary objectives of this paper were following

- Focused on state-of-the-art research on deep learning techniques which were used for melanoma diagnosis
- To identify the recent research trends, challenges and opportunities in the field of melanoma diagnosis
- Investigate the existing solutions for the diagnosis of melanoma and provides a systematic review of these solutions on the basis of similarities and differences
- Proposed a taxonomy to emphasize on effective methods of deep learning systems for melanoma diagnosis

2) RESEARCH QUESTIONS

Analyzing the primary research questions remains important for a systematic review. The analysis procedure involves designing the search strategies to find and extract relevant studies after defining the research questions [17]. The answers to these questions fetched through the published literature, according to the methodology suggested by Kitchenham et al [13]. The fundamental purpose of this study was to summarize the current, state of the art techniques for melanoma detection in the context of CNN-based models. The research questions were formulated to evaluate the importance of the study, as outlined in Table I.

3) SEARCH STRATEGY

Well-organized research remains a prerequisite to extract suitable information and eliminate unrelated studies from focused research areas. We shortlisted specific articles in this systematic study that developed or validated methods for melanoma detection using CNN. We have searched the databases of IEEE, Springer, Elsevier, Wiley online, ScienceDirect, ACM, Scopus, Medline, MDPI from Jan 1, 2014, to March 31, 2020, this study contains only those papers that show ample scientific proceedings. Search string has been made with the combination of primary, secondary & additional keywords as shown in Fig. 2.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

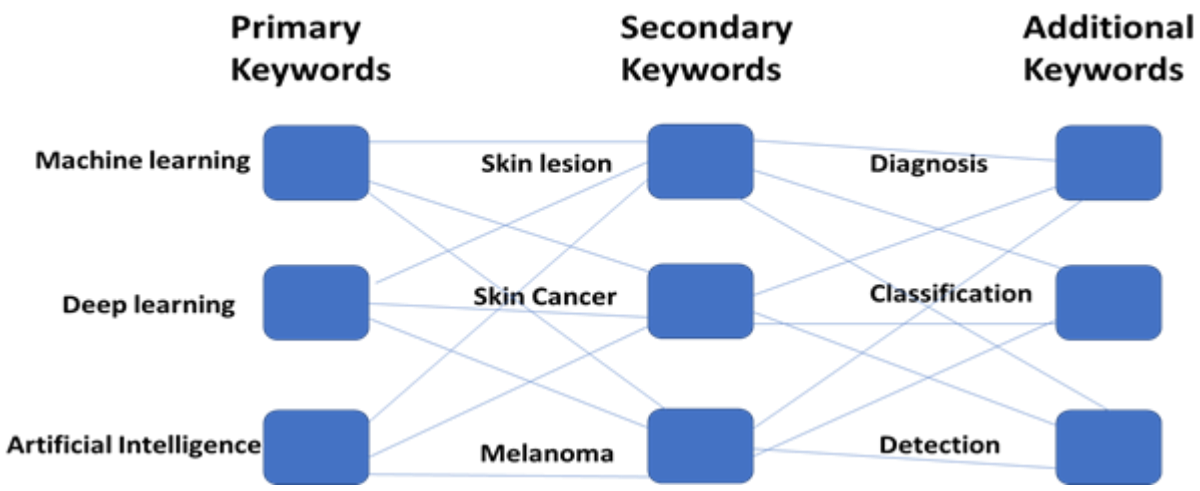

Fig. 2. **Search string keywords for selecting studies**

TABLE II

Search strategies for databases

Database	Search Strategy
IEEE Xplore	((“Document Title”: “machine learning” OR “deep learning” OR “artificial intelligence” OR “neural network” OR “vector machine” OR “Bayesian” OR “Supervised Learning”) AND (“Abstract”: “skin lesion,” OR “skin cancer,” OR “Melanoma”) AND (“diagnosis” OR “classification” OR “detection”)) Publication Year: Year: 2014-2020
ACM Digital library	((“machine learning” OR “deep learning” OR “artificial intelligence” OR “neural network”) AND (“skin lesion,” OR “skin cancer,” OR “Melanoma”) AND (“diagnosis” OR “classification” OR “detection”)) Publication Year: Year: 2014-2020
Medline	("machine learning"[All Fields] OR "deep learning"[All Fields] OR "artificial intelligence"[All Fields] OR "neural network"[All Fields]) AND ("skin lesion,"[All Fields] OR ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "cancer"[All Fields]) OR "skin cancer"[All Fields]) OR "Melanoma"[All Fields]) AND ("diagnosis"[All Fields] OR "classification"[All Fields] OR "detection"[All Fields]) Publication Year: 2014-2020
Science Direct	(“machine learning” OR “deep learning” OR “artificial intelligence” OR “neural network”) AND (”skin lesion,” OR “skin cancer,” OR “Melanoma”) AND (“diagnosis” OR “classification” OR “detection”)) Publication Year: 2014-2020
Wiley online	machine learning OR deep learning OR artificial intelligence OR neural network AND skin lesion, OR skin cancer, OR Melanoma AND diagnosis OR classification OR detection Year: 2014-2020
Springer Link	((“machine learning” OR “deep learning” OR “artificial intelligence” OR “neural network”) AND (“skin lesion” OR “skin cancer” OR “Melanoma”) AND (“diagnosis” OR “classification” OR “detection”)) Year: 2014-2020
Scopus	TITLE-ABS-KEY (“machine learning” OR “deep learning” OR “artificial intelligence” OR “neural network”) AND (“skin lesion” OR “skin cancer” OR “Melanoma”) AND (“diagnosis” OR “classification” OR “detection”)) Year: 2014-2020

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

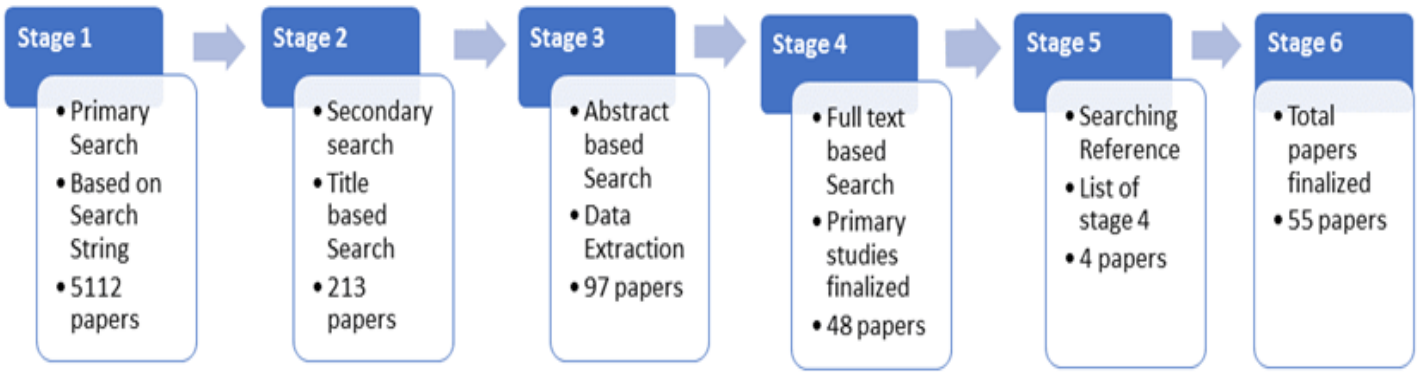

Fig.3.Selectionand screening process

Furthermore, in this article primary keywords (PK) were machine learning, deep learning and artificial intelligence. Whereas secondary keywords (SK) were skin lesion, melanoma and skin cancer. However, additional keywords (AK) were detection, diagnosis and classification. OR operators used in same level keywords whereas, AND operators used in different level keywords. The combination of PK, SK and AK was used to search the relevant papers from databases. The standard search criteria for each repository has been followed. The criteria for making search strings given in [88], [89], [90], [91] and [92]. The search string for each dataset is given in Table II.

4) STUDY INCLUSION/EXCLUSION CRITERIA

Prisma statement [18] provides the inclusion criteria for research. The articles portrayed recent and emerging research for melanoma detection were included. Moreover, the inclusion criteria have been limited to the search strings mention in Table II which contain search stings for different databases. This criterion was consistent with the research area. Upon evaluating the criteria for inclusion, the next step is to apply exclusion criterion Papers were omitted on the basis of the following points

- Papers which were not focused on the binary classification of disease.
- Melanoma diagnosis without medical images were excluded.
- Works, where the performance's origin was not credible, were omitted.
- Research papers were omitted on the basis of non-human samples

B. SELECTION AND SCREENING CRITERIA

The selection process was carried out by identifying the most relevant articles, which were aligned with this systematic study's objective [16]. Articles that provides a significant research contribution have been selected in this review.

In the first stage, 5112 studies were included by using the search string. Not all the papers in the study were specifically related to research questions; thus, according to the actual relevancy they needed to assessed. Irrelevant and duplicate papers were manually omitted in the second stage on the basis of titles. 213 were listed as appropriate. In the third stage, the abstracts of research papers were analyzed. Papers whose abstracts show a significant working were included. 97 papers remained on the list after this point. A complete text-based analysis was performed in the next stage. 48 papers were finalized at this stage.

Snowball tracking was applied after all these filters by searching through the reference of each selected study and ensuring that no significant study was missing. Seven more studies were selected at this stage and added them to selected papers, as a result a total of 55 primary studies were included in this review. Fig. 3. shows the complete Selection process.

III. Data analysis and results

This section summarizes the findings and includes a tabular format for a brief assessment of each study. Each section addresses the problem description, proposed solution, strengths and weaknesses. Summary and suggestions have been taken on the basis of the findings of the study data.

A. QUALITY ESTIMATION

The quality estimation has become an integral part of a systematic study. In order to enhance the quality of our research, a questionnaire was designed to review the quality of the included articles. The quality assessment was carried out by the two authors who retrieved the studies.

1) Has the study uses deep learning algorithms for melanoma diagnosis? The possible answer was ‘‘Yes (+1)’’ and ‘‘No (+0).’’

2) Has the study provided a clear solution to the problems of disease diagnosis using datasets? The possible answer was ‘‘Yes (+1)’’ and ‘‘No (+0).’’

3) The article has been published in a known and reliable source of publication. This query was ranked by seeing the Journal Citation Reports (Q1, Q2, Q3 & Q4) quartile ranking and the computer science conference rankings (CORE) (A, B, and C).

For conferences and seminars, the possible answers to this question were:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

- (+2) if it is ranked CORE A,
- (+1.5) If it is ranked CORE B,
- (+1) if it is ranked CORE C,
- (+0) If it is in CORE ranking.

The possible answers to this question were for journals, letters and scientific reports:

- (+2) if it is ranked Q1,
- (+1.5) if it is ranked Q2,
- (+1) if it is ranked Q3
- (+0.5) if it is ranked Q4
- (+0) If it has no JCR ranking.

The quality criterion score (c) mentions to the fact that journals were more valuable than conferences and workshops because the authors assume that it may be more difficult to publish work in Q1, Q2, Q3, Q4 journal than in other publication channels. Appendix Table I. provides the publication source while Appendix Table II. presented the Quality Assessment of 55 selected studies.

B. SEARCH RESULTS

This section describes the results related to the systematic study questions. The Fig. 4. shows selected papers distribution, it provides a graphical view of all 55 finalized articles which have different publication channels, and the number of articles per publication source. Two different publication channels were identified in addition to one symposium and a scientific report. About 31 % of the selected papers appeared at the conference. Whereas 65% were published in journals and 2 % each were presented in the symposium and book chapter.

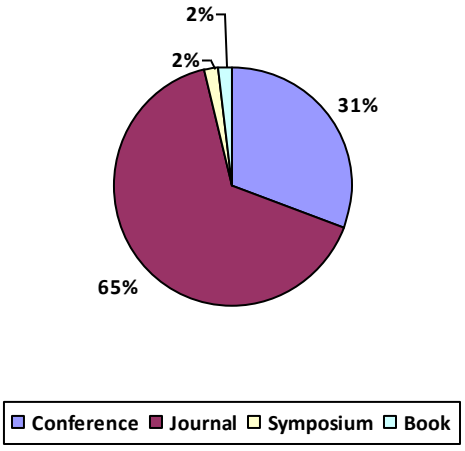

Fig. 4. Selected papers distribution

Furthermore, a list of all sources and the division of overall selected papers from the respective source i.e. IEEE, ACM, Springer, Science Direct, Wiley, Medline has been illustrated using Pie-Chart diagrams. Whereas, Fig. 4. (a) selected journal paper distribution, explains that overall papers from conferences have opted from IEEE, has a share of 29 %, similarly, springer has also share of 29 %, Science Direct has a share of 24%, MDPI has a share of 9% whereas, Wiley and Springer each has a share of 6% each, paper share for our study.

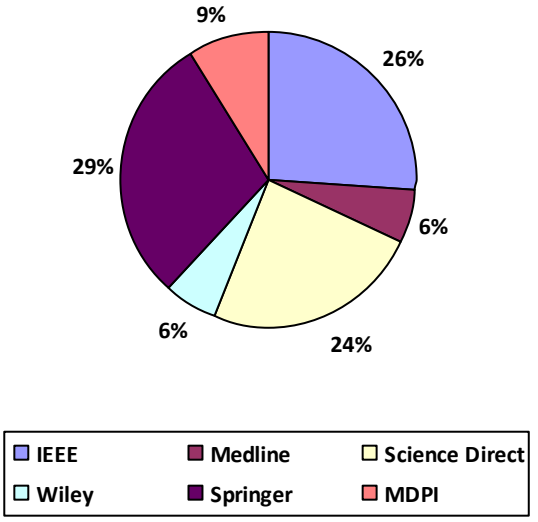

Fig 4. (a) Selected conference papers distribution

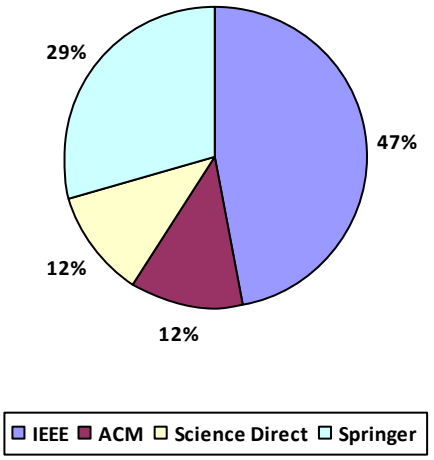

Fig. 4. (b) Selected journal papers distribution

Fig. 4. (b) selected journal paper distribution, the journal papers have opted from Springer, has a 30% share, IEEE has 29%, Science Direct has a 23%, Wiley, MDPI and Medline have a 6% each share in selected publications

C. RESEARCH QUESTIONS ASSESSMENT AND DISCUSSION

Melanoma remains the most fatal type of cancer. Its spreading ratio increases day by day in the world. The effective and pre-diagnosis of this disease stays very important [12]. Based on our research questions, we have

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

examined the 52 finalized papers. After the thorough examination of selected papers, we take out the pieces of evidence from this diverse research domain for assessment.

1) EVALUATION OF RQ1: WHAT TYPE OF DEEP LEARNING BASED CLASSIFIERS ARE USED FOR DIAGNOSING MELANOMA DISEASE?

Melanoma considered as one of the most common type of cancer. A lot of extensive research done in this field, some new techniques and new algorithms have been relying on deep learning for diagnosing melanoma.

a) Pre-trained Models

A pre-trained convolution neural networks CNN technique; named AlexNet, has been proposed by Pomponiu *et al.* [19], which produce high level skin samples for the classification of skin disease. Moreover, the proposed method derives features from the last three entirely linked layers which were utilized to prepare a k nearest neighbor (NN) classifier for skin malignant classification. In comparison to this, Esteva *et al.* [20] suggested a pre-trained CNNs for skin malignant diagnosis using a huge dataset for their analysis (129,450 clinical images). Moreover Mahbod *et al*.[21] used pre-trained CNN to study the classification of skin lesions, a pre-trained AlexNet and VGG-16 architecture were implemented in their algorithm for the classification of skin lesions to extract distinct features from dermoscopic imageries. Whereas Soudani *et al.* [46] used a hybrid technique of amalgamation of pre-trained architectures (VGG16 and ResNet50) to extract characteristics from the convolutionary sections. A classifier with an output layer has been designed which consists of five nodes. These nodes represent the classes of the segmentation methods and predict the most effective skin lesion detection and segmentation technique in any image data. A pre-trained deep learning network and transfer learning were proposed in Khalid *et al.* [31]. The transfer learning was applied to AlexNet to identify skin lesions in addition to fine-tuning and data increase. An automated system for melanoma detection has been developed by Bisla *et al.* [40] which counter the limitation of datasets. Moreover, proposed method relies heavily on the processing unit to eliminate image occlusions and the unit for data generation for skin lesion classification

A classification method proposed by Aldwgeri *et al.* [48] uses CNN and transfer learning to enhance skin classification. Various pre-trained models, including VGG-Net, ResNet50, InceptionV3, Xception, and DenseNet121, have been applied. Georgakopoulos *et al.* [49] studied the effects of unifying transfer learning in CNN architecture training. The outcome of such a hybrid system shows that the effects of classification were substantially improved. Gavrilov *et al.* [51] proposed an early diagnostic algorithm focused on deep convolutional neural networks which efficiently differentiate between benign and malignant skin cancer. Mahbod *et al.* [53] provide a fully automatic computerized system for the classification of skin lesions that uses optimized deep features from a range of well-established pre-trained CNN models such as AlexNet, VGG16 and ResNet-18 and then trains vector machine classifiers.

b) Handcrafted Methods

In addition to using pre-trained CNNs like Mahbod *et al* [21], some research papers have built their own method using CNNs for efficient melanoma detection. In [22], Massod *et al*. introduce a semi-supervised, self-assisted learning method for melanoma diagnosis in dermoscopic images. The proposed system trained a deep belief network and two self-assisted support vector machines (SA-SVMs) with radial basis function (RBF) kernel and polynomial kernel on three different datasets, respectively. The three datasets were generated from both labeled data and unlabeled data. Throughout the training period, a fine-tuning technique with an exponential loss function implemented to improve the marked results. While Majtner *et al*. [23], introduced a melanoma identification system incorporating handcrafted features and comprehensive features. This study uses deep learning and support vector machine for melanoma diagnosis. For grayscale images, a support vector machine was used to extract features and for raw color image convolution, a neural network was implemented to produce likelihood scores. Results were determined based on high scores. However, Sabbaghi *et al*. [24], proposed a method which uses deep neural network for extracting features from images and transmit these features into a bag of features (BoF) space to enhance classification accuracy. In comparison to [24], Demyanov *et al.* [25] suggested a method that utilizes CNNs to perceive two pattern forms in dermoscopic images of the skin (typical network and regular globules). The regular stochastic gradient descent algorithm trained CNN in the proposed method. In addition to the work suggested overhead, Nasr-Esfahani *et al.* [27]suggested a method for the identification of melanoma lesions feeding previously handled experimental images on deep learning model. Whereas, Sabouri *et al.* [28]recommended a border detection system based on CNNs for the recognition of skin lesions. However, Sreelatha *et al.* [41] use the pattern Gradient and Adaptive Contour Function (GFAC) to detect skin cancer melanoma. In this method, techniques of pre-processing image segmentation and techniques of noise reduction used to reduce noise. Mukherjee *et al.* [43] suggested an architecture called CNN malignant lesion detection (CMLD) that was used to calculate image classification accuracy. Majtner *et al.* [50] provide an improved method for melanoma detection based on the combination of Linear Discriminant Analysis (LDA) and features derived from the Deep Learning approach. Using the LDA method improves the accuracy of classification. However, Namozov *et al.* [52] introduce a deep neural network model with adaptive linear piecewise units that can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

achieve excellent melanoma recognition performance. Khan *et al.* [54] proposed an automated system based on transfer learning for features extraction. whereas, optimal features were extracted using kurtosis-controlled part theory (KcPCA) for the effective diagnosis of melanoma.

c) Fully Convolutional Networks (FCN) based Methods

In [26], Yu *et al.* proposed a method which utilize deep residual networks to diagnose melanoma using dermoscopic images. The proposed method utilizes two FCN that substitute conventional convolutional layers with residual blocks as mention in the architecture of FCN. Moreover, proposed method generates a grade map from the dermoscopic images to the segment skin lesion. The area of interest which contains the skin lesion has been resized, cropped and transferred for melanoma classification. Moreover, Jayapriya *et al.* [47] employ a hybrid framework that includes two FCNs (VGG 16 & GoogleNet). The classification was performed using deep residual network and a hand-crafted tool to remove the feature from the segmented lesion.

d) Ensemble deep learning Methods

Many studies use ensemble deep learning techniques for melanoma classification like Milton *et al.* [44] uses a cooperative deep learning model, which was tested on a benchmark dataset of ISIC 2018. To identify skin lesions from dermoscopic images. However, Mahbod *et al.* [45] propose a cooperative deep learning-based technique that was designed by unifying intra-architecture and inter-architecture network fusion for convolutional neural networks (CNNs).This was a completely automatic and instinctive computerized process.

The summary of the classifiers has been given in Table III.

2) EVALUATION OF RQ2: WHAT IS THE PERFORMANCE MATRICS USED BY CLASSIFICATION METHODS TO VALIDATE THE EFFICIENCY IN DIAGNOSING MELANOMA?

The accuracy of different studies was calculated by the evaluation metrics such as sensitivity, specifity, precision, accuracy and area under curve (AUC). The reliability of every classifier judge on these parameters. Table. V. provides a summary of performance metrics.

a) Efficiency calculation on single datasets

Efficiency is the key factor for the model reliability. For this purpose, Khalid *et al.* [31]uses AlexNet for transfer learning and classify three different lesions. The proposed system was trained and tested on the PH2 dataset only, the achieve accuracy rates were high, but the credibility of the method was low due to the use of only one dataset. A different technique was implemented by Gulati *et al.* [57], which uses two pre-trained models VGG16 and Alex Net and experimentation was done on the PH2 dataset, which contains only 200 images for testing and training, although results were impressive but. Large datasets have to be used to check the performance of a given model. Moreover, Jianu *et al.* [62] use deep convolution neural networks. The main limitation of this study was to use fewer numbers of images for rare cancers like actinic keratosis, dermatofibroma and vascular lesion Whereas Warsi *et al.* [66], use the PH2 dataset for its proposed model. For maximum accuracy, this model has to be tested on multiple datasets which contains a large number of images for training. Results come more efficient when datasets contain a limited number of images.

Similarly, to the aforementioned studies, Khan *et al.* [54] implemented transfer learning, while Abbas *et al.* [55] implemented the fusion of feature vector with deep learning. Whereas Wang et al [75] proposed a dual deep CNN, which can mutually learn from each other, in comparison to this Nida *et al.* [78] incorporated deep regional convolutional neural network (RCNN) with Fuzzy C-mean (FCM) clustering, for testing their method, and they use only ISBI 2016 dataset. Lopez *et al.* [68] uses transfer learning and Yunhao *et al.* [73] incorporated fully convolution neural network, they implemented their methods on ISBI 2016 dataset, these methods have to use large datasets for computing actual results. ISBI 2016 has a limited number of images due to which the accuracy of these studies remains ambiguous.

Namozov *et al.* [52] implemented transfer learning whereas, Yu *et al.* [61] incorporated CNN with feature vector while, Yang *et al.* [71] incorporated region average pooling method (RAPooling) with RankOpt to achieve efficacious results, whereas, both these methods use ISBI 2017 datasets to validate their results. However, *Li et al.* [77] implemented two fully convolutional neural networks, use ISBI 2017 datasets for the validation of results.

Adwgeri *et al.* [48], uses ISBI 2018 dataset and incorporate pre-trained model like VGG-Net, ResNet50, InceptionV3, Xception and DenseNet. Shahin *et al.* [60] introduce an ensemble CNN, which can enhance the results when tested on ISIC 2018 dataset. However, Hagerty *et al.* [69] presented a handcrafted ensemble technique that uses ISBI 2018 dataset and achieves optimal accuracy.

Moreover, Albahar *et al.* [67], implemented Skin Lesion Classification using CNN & Novel Regularizer while Hasan *et al.* [72] proposed neural networks with feature selection, and uses ISIC archive to test their results. This archive consists of more than 23000 images from different skin cancers, which increases the efficiency of the method

In addition to the above-mentioned studies, Albert *et al.*[79] implemented Predict-Evaluate-Correct K-fold (PECK), Synthesis and Convergence of Intermediate Decaying Omni gradients (SCIDOG), PECK trains ensembles by utilizing limited data while SCIDOG easily detects lesion even if there is noise in the image. These algorithms were applied on the Mednode dataset that contains 170 images in which 100 images were melanoma and 70 images were benign.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

Almaraz *et al.* [93] proposed a method which fuses deep learning features with handcrafted features via mutual information measures to extract the important information from both type of features. The proposed method uses several methods such as linear regression, support vector machine and relevant vector machines for classification. The efficiency of proposed method has been tested on ISBI challenge 2018.

b) Efficiency calculation on Multiple datasets

Masni *et al.* [63], uses images of PH2 dataset and ISBI 2017 challenge. To achieve maximum performance, a large number of images for training have been used and the segmentation of each class must be improved. Xie *et al.* [64], uses PH2, ISBI 2016 and 2017 image dataset, by using multiple datasets, the proposed method achieves high performance with better segmentation, it also provides robustness against hair fibers. Sarkar *et al.* [70] use three datasets which include PH2, ISIC2016 and MEDNODE datasets which provide the high performance of a proposed method, one of the limitations of the proposed system was that it uses only dermoscopic images for melanoma diagnosis. This method has used non-dermoscopic images for diagnosis. Jayapriya *et al.* [47], uses transfer learning which was based on GoogleNet and VGG16. To validate the proposed method ISBI 2016 and 2017 datasets were used. Khan *et al.* [54] use HAM 1000, ISBI 2016 & 2017, it uses the technique of transfer learning for feature extraction and feature were selected using technique KcPCA, which generates better results by using these techniques. Performance can further be enhanced if the proposed method uses autoencoder for feature selection and classification. Attique *et al.* [59] incorporated optimized color feature segmentation with CNN, and it uses ISBI 2016, ISBI 2017 and ISBI 2018 datasets to check the effectiveness of its proposed method.

El-Khatiab *et al.* [81] use transfer learning which was based on Google Net, ResNet and NasNet. To validate the proposed method PH2 and ISBI 2019 were used. Whereas, Adegun *et al.[*82] proposed an end to end and pixel-wise learning using DCNN and uses ISBI 2018 and PH2 to validate their results.

Dugonik *et al.*[94] proposed an ensemble method consist of ResNet, DenseNet, SE-ResNext,and NasNet. The efficiency of proposed method tested on ISIC Archive and Dermnet. The proposed method shows promising results on both datasets. In addition to this, Khan *et al.*[95] proposed a method which implement DenseNet for feature extraction and for feature selection proposed method used iteration-controlled Newton-Raphson (IcNR) method. ISBI 2016 and 2017 datasets were used to test the performance of proposed method.

c) Efficiency calculation on Combined datasets

Amin *et al.* [65], proposes a system that has implemented a fusion of Alex net and VGG16, this system was tested on a combined dataset of PH2 + ISBI 2016 +ISBI 2017 which consists of 3100 images in total. This system performs efficiently on a diverse and large dataset.

Pham *et al.* [58] have proposed a Deep CNN with Data Augmentation and to test the method, it combines PH2+ ISIC archive, +ISBI 2017 dataset to make a large dataset.

The segmentation process can be made efficient by strengthening it against hair and artifacts. Zhang *et al.* [74], implemented optimized CNN and uses a combined dataset dermquest and dermIS, which provides more than 22,000 images for training and testing. Hosnay *et al.* [76] incorporated transfer learning which first tested on ISIC archive then, it also tested on a combined dataset, which consists of dermIS and dermQuest for checking the accuracy of their proposed system. Nahata *et al.*[80] incorporated ensemble techniques by utilizing Resnet 50 and inception v3 and this study combine ISBI 2018 and ISBI 2019 to make a large dataset for checking the efficiency of their proposed technique.

d) Efficiency calculation on Internet Collected images

Dori *et al.* [56] use unpublished image data set. The performance of the method has to be tested on publicly available benchmark datasets for actual performance analysis. The summary of these publications has been given in Table V.

TABLE III
Deep learning classifiers

Contribution of Publication	Image Type	Architecture	Training Technique	Dataset	Reference
Transfer learning is applied on two datasets for malignant melanoma diagnosis	Dermoscopy	DNN	Transfer learning	DermIS [29], DermQuest [30]	[19]
Skin Cancer classification using pretrained model	Dermoscopy	DNN	Transfer learning	Non-published medical dataset	[20]

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

on large dataset (129,450 images)					
Skin Lesion Classification Using pretrained AlexNet, VGG16 & Resnet	Dermoscopy	Hybrid DNN	Transfer Learning (AlexNet, VGG16 and ResNet-18)	ISIC 2016 [5]	[21]
Self-supervised learning model for skin cancer diagnosis using SVM & CNN	Dermoscopy	Deep belief architecture (DBA)	SVM + CNN	Non-published medical dataset	[22]
Skin lesion classification by using SURF, LBP & CNN	Dermoscopy	CNN	DNN	ISIC 2016 [5]	[23]
Classification of melanomas using bag-of-features(BoF) & CNN	Dermoscopy	CNN	DNN+BOF	Non-published medical dataset	[24]
Classification of skin cancer using GDA for CNN training	Dermoscopy	DCNN	Gradient descent Algorithm(GDA)	ISIC 2016 [5]	[25]
Automated Melanoma Recognition using FCN	Dermoscopy	Hybrid (FCN+CNN)	Deep Residual Network	ISIC 2017 [37]	[26]
Melanoma diagnosis using CNN	Dermoscopy	CNN	DNN	MED-NODE [32]	[27]
Melanoma detection using CNN	Dermoscopy	CNN	DNN	DermIS [29], DermQuest [30]	[28]
Skin cancer classification using pretrained model (AlexNet)	Dermoscopy	CNN	Transfer learning (AlexNet)	PH2[38]	[31]
Automated melanoma detection using transfer learning	Dermoscopy	CNN	Transfer learning	PH2[38], ISIC 2017 [37], ISIC 2018 [39]	[40]
skin cancer detection by preforming image segmentation & noise reduction	Dermoscopy	CNN	Gradient and Feature Adaptive Contour (GFAC)	PH2[38]	[41]
Malignant Melanoma Classification using CMLD architecture	Dermoscopy	CNN	CNN malignant lesion detection (CMLD) architecture	Dermofit [42] MED-NODE [32]	[43]
Skin lesion	Dermoscopy	CNN	Ensemble DNN	ISIC 2018 [39]	[44]

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

classification using ensemble deep learning model					
Skin lesion classification using inter & intra architecture network fusion for CNN	Dermoscopy	CNN	Ensemble DNN	ISIC 2017 [37]	[45]
Melanoma Diagnosis using VGG 16 & ResNet50	Dermoscopy	CNN	Transfer learning + Crowd sourcing	ISIC 2017 [37]	[46]
Skin lesion segmentation using two FCN's and hand-crafted tools	Dermoscopy	CNN	FCN's	ISIC 2017 [37], ISBI 2016 [5]	[47]
Ensemble of Deep Convolutional Neural Network for Skin Lesion Classification	Dermoscopy	CNN	Ensemble CNN	PH2 [38]	[48]
Detection of Malignant Melanomas in Dermoscopic Images by using transfer learning	Dermoscopy	CNN	Fine Tuning + Transfer learning	Unpublished dataset	[49]
Melanoma detection using LDA & CNN	Dermoscopy	CNN	Linear discriminant analysis (LDA) + CNN	ISIC Archive [35]	[50]
Use of Neural Network Based Deep Learning Techniques for the Diagnostics of Skin Diseases	Dermoscopy	CNN	Transfer learning	ISIC 2017 [37]	[51]
Parameterized Activation Function used for Melanoma Classification	Dermoscopy	CNN	Parameterized Activation Function	ISIC 2018 [39]	[52]
Skin Lesion Classification Using Hybrid Deep Neural Networks	Dermoscopy	CNN	Hybrid DNN	ISBI 2017[37]	[53]
Multi-Model Deep Neural Network based Feature Extraction and Optimal Selection Approach for Skin Lesion Classification	Dermoscopy	CNN	Transfer learning	ISIC 2016 [5], ISIC 2017 [37]	[54]

3) EVALUATION OF RQ3: WHAT TYPE OF DATASETS ARE AVAILABLE FOR DIAGNOSING MELANOMA?

There were several data sets available for the detection of skin lesions. Some were open to the public and some were not publicly available. Table IV. provides a summary of the benchmark datasets.

Benchmark Datasets: Below mentions, datasets were considered as a benchmark because of its excessive usage in studies for melanoma detection.

a) ISBI Challenge 2016 Dataset [5]: The research articles [47], [48], [54], [55], [59], [61], [64], [68], [73], [75] and [78] have used this dataset to check the accuracy of their proposed methods The Dataset Challenge includes 900 training images (273 melanomas) and 379 evaluation images (115 melanomas).

b) DermIS [29] & DermQuest [30]: In this review, article [70], [74] and [76] use these datasets. DermIS provides images for the diagnosis of different types of skin cancers. On the internet, dermIS provides the largest online image information for skin cancer diagnosis. It contains 146 melanoma images Dermquest provides medical images for dermatologists. The well-known international editorial board approved and reviewed these images. It provides 22,000 clinical images to a dermatologist for analysis purposes.

c) MEDNODE dataset [32]: It consists of 100 melanoma and 70 naevus images which were collected from the University of Medical Center's Department of Dermatology, Groningen. In our review [70], [76] and [79] use this dataset for experimentation.

d) ISIC Archive[35]: A dermoscopic image dataset with 23,906 dermoscopic images that has been publicly available. In this review [67], [70], [72], [76] and [94] use this dataset to check the accuracy of their methods

e) ISBI 2017 Dataset Challenge[37]: This dataset was used by [47], [50], [58], [59], [63], [64], [71] and [77] for testing their proposed methods. It contained 2,000 dermoscopic images in which 374 were melanomas, 254 were seborrheic keratoses, and 1,372 benign nevi images.

f) PH2 Dataset [38]: This is a dermoscopic image database obtained from the Pedro Hispano Clinic, Portugal Dermatology Service. In this Review [31], [57], [62], [63], [64], [66], [70] and [82] uses PH2 for diagnosis of melanoma. PH2 contains a total of 200 dermoscopic images in which 40 were melanoma and 160 were of non-melanoma images.

g) 2018 ISIC challenge [39]: This dataset contained more than 10,000 dermoscopic images of 7 types of diseases (melanoma, nevi, seborrheic keratosis, BCC, Bowen's disease and actinic keratosis, vascular lesions, and dermatofibromes). The studies include [48], [59], [60] and [69] have used this dataset

h) Dermofit Image Library [42]: This dataset has a collection of 1,300 high-quality images of skin lesions which were collected under standardized color conditions. There were 10 different classes of lesions: Actinic Keratosis, Basal Cell Carcinoma, Melanocytic Nevus (mole), Seborrhoeic Keratosis, Squamous Cell Carcinoma, Intraepithelial Carcinoma, Pyogenic Granuloma, Haemangi-oma, and Dermatofibroma. In this review, publication [43] used this dataset for the validation of their proposed method.

i) 2019 ISIC challenge [83]: This dataset contains 25,333 images of 8 different types of skin lesion diseases named as Actinic Keratosis, Melanoma, Vascular Lesion, Squamous, Cell Carcinoma, Basal Cell Carcinoma, Benign Keratosis, Melanocytic Nevus, Dermatofibroma. In this review [80] and [81] uses this dataset.

Non- Public Dataset: The no-public datasets have been mostly used benchmark datasets. Whereas, there exist other datasets which were not publicly used, has been given below.

a) Interactive Dermoscopy Atlas [33]: the dataset consists of 112 images of malignant lesions (containing melanoma and basal cell carcinoma (BCC)) and 298 images of benign lesions (congenital, organic, dermal, Clark, spitz, and blue nevus; dermatofibroma; and seborrheic keratosis). There have two modalities for each image, dermoscopic and clinical.

b) Dermnet [34]: This database contains more than 23,000 images of skin lesions divided into 23 skin disease classes. In this review [94] uses this dataset.

3) Non-Listed Dataset

a) IRMA Skin Lesion Dataset: 747 dermoscopic images (186 melanomas). This dataset only available on request and a licensing agreement has been signed.

b) Dataset MoleMap: Dataset MoleMap NZ Ltd has been obtained between 2003 and 2015 with both dermoscopic and clinical images. The number of images were 32,195 photographs of 8,882 patients from 15 disease groups with 14,754 lesions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

TABLE IV.
Publicly available datasets

Data Set	No. of dermoscopic Images	Melanoma images
2016 ISBI Challenge[5]	900	273
2017 ISBI Challenge[37]	2000	374
2018 ISBI Challenge [39]	10000	1113
2019 ISBI Challenge[83]	25333	4522
PH2 [38]	200	40
ISIC Archive [35]	23906	21659
Dermofit Image library [42]	1300	76
Dermis [29]	397	146
MED-NODE [32]	170	100

IV. Discussion

This study presented a comprehensive review of the state-of-the-art techniques for the detection of melanoma. A taxonomy for melanoma was provided to summarize the results of this research, that has been shown in Fig. 4. Moreover, a model for melanoma detection has also been proposed as shown in Fig. 5.

A. TAXONOMY FOR MELANOMA DETECTION

The designed taxonomy initially perform two operations on the skin lesion and categorized it into melanoma and benign, majority of findings reviewed in this analysis were focused on a deep learning for the binary classifying of disease. Once the cancer was diagnosed as melanoma, it was further studied for identification of a suitable form. Melanoma consists of four primary types which are superficial spreading, Noda melanoma, Acrel lentigious and letigo maligna. Deep learning methods were trained to detect the specific type. All these types have different shapes, locations, structures, size and colors. Superficial spreading melanoma has a dark spot on the skin, it changes its color and it has an irregular border. Furthermore, Acral lentiginous and Lentigo maligna have irregularly shaped, both change its size and color. Moreover, if the disease has been detected as a benign, then it is classified into three different types as Dermal, Melanocytic and Epidermal. These types fall in the category of a non-cancerous disease which has been formed on the skin and has a resemblance to melanoma. By establishing a taxonomy for melanoma identification using deep learning, the results of this research have been summarized in Fig. 5.

B. MODEL FOR MELANOMA DETECTION

The effective model for melanoma detection has been proposed in Fig. 6. that has comprised of five main components, which based on data acquisition, fine-tuning, feature selection, deep learning, and model finalization. The first step is a data acquisition, in which dataset for skin cancer detection is selected from publicly available benchmark datasets, non-listed and non-public datasets such as internet collected images for melanoma detection.

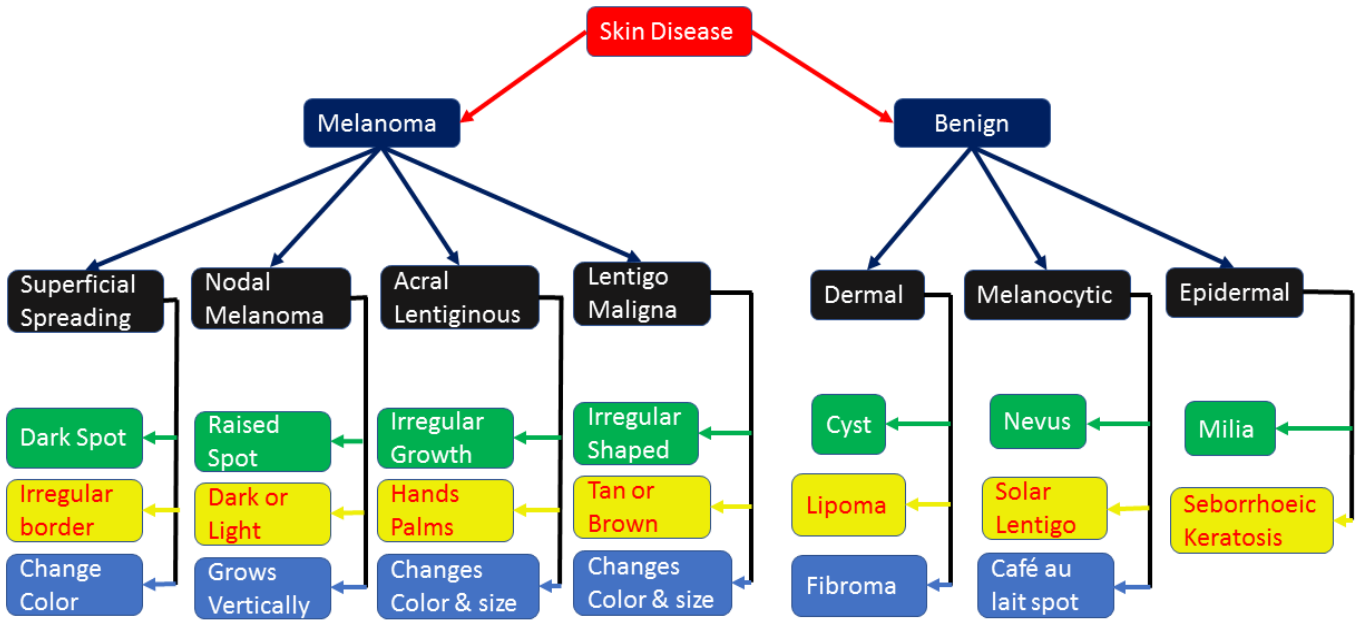

Fig. 5. **Taxonomy for Melanoma Detection**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

In publicly available benchmark datasets, PH2, DermIS, DermQuest, Mednode and ISBI Challenge 2016, 2017, 2018 and 2019 have been available. While in the non-public datasets the available options were Interactive Dermoscopy Atlas and Dermnet. Moreover, IRMA Skin Lesion Dataset and MoleMap datasets were available in the category of Non-listed datasets. Furthermore, the Internet collected image datasets were also used.

Fine-tuning were applied to images of datasets which contained unnecessary information such as wrinkle, dark spots and hair were removed from the skin surface for achieving the best results. Moreover, testing and training sets of images have also been defined at this stage. Whereas, appropriate feature selection has a significance important, after that feature reduction techniques were applied to suppress the bulk of features and fetch the most important features from available data. After applying all the previous steps, deep learning methods were applied to the datasets for checking the reliability of given methods.

Many deep learning-based methods were available, in which some deep learning techniques were based on transfer learning, while some were based on ensemble techniques, whereas some techniques utilize fully convolutional neural network models and hybrid methods [87]. All the results from different techniques have been evaluated to choose the best technique as shown in Fig. 6.

V. Challenges and Opportunities

There exist many challenges and opportunities for the detection of Melanoma using deep learning techniques. This section discusses the challenges which were identified from the literature.

A. Datasets Variation

There were different datasets available for the classification of melanoma. Some datasets were available publicly while others were not. It has been observed that numbers of images varied in different datasets. Moreover, some articles made a self-collected image dataset using the internet.

1) Limited number of images in datasets

Available benchmark datasets have a limited number of images for training and testing and it has also observed that benchmark datasets have a small number of images for testing and training. Proposed methods perform well on a small number of images and there reliably is uncertain when test on a large image set. The PH2 [38] dataset only contains 200 images for testing and training. To overcome this issue ISIC, announce an annual challenge to address the defined issue from 2016[35]. In addition to this, some researcher [58], [65], [74], [76] and [80] combines the different datasets to form one large image dataset and then validate their proposed methods.

2) NON-PUBLIC DATASETS

Some researchers [56] use non-public datasets and internet collected images. Which makes replication of the results more complicated due to non-availability of the dataset.

3) LIGHT SKINNED PEOPLE DATASETS

Since 2016 ISIC [35] organized an annual challenge for melanoma diagnosis but one of the limitations of the ISIC challenge is that it has data of mostly light-skinned people. Images of dark-skinned people should be included in datasets.

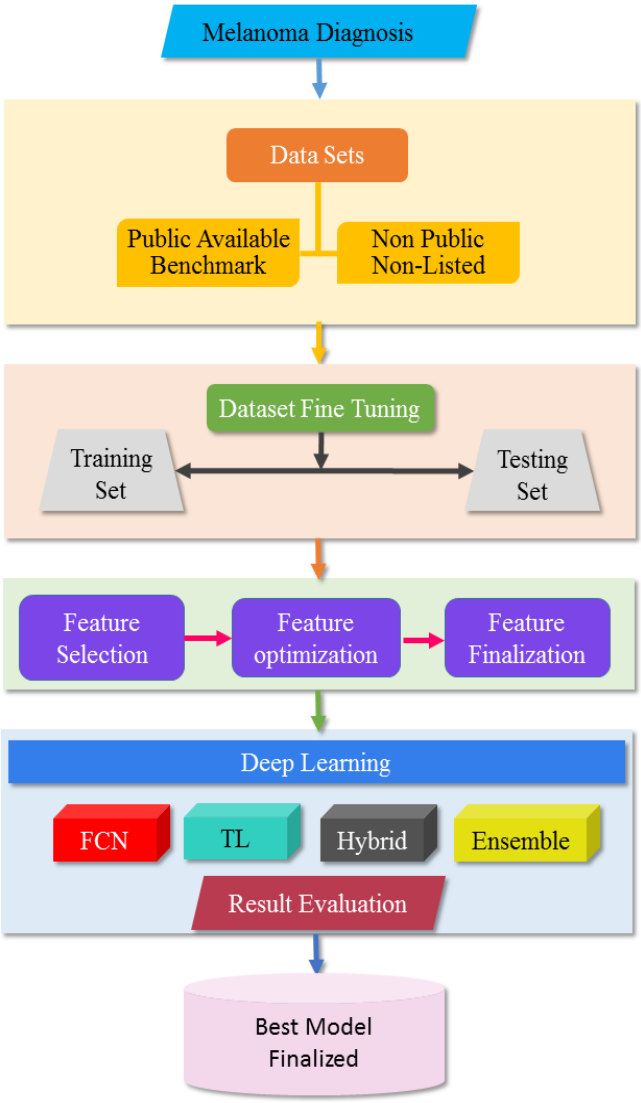

Fig. 6. **Model for Melanoma Detection**

B. SIZE OF LESION

It has been also observed that the size of the lesion has significant importance. If the size of the lesion is less than 6mm then it is difficult to detect melanoma and the accuracy of detection drastically dropped, lesion above 6mm is considered as melanoma [84].

C. POTENTIAL CLASSIFIERS FOR MELANOMA

It has been observed from the literature that pre trained models and handcrafted method based on deep leaning showed promising results for melanoma detection with high accuracies.

D. ACCURACY OF DEEP LEARNING METHOD

After reviewing the selected studies, it has been observed that deep learning methods perform well when 70 % images were used for training and 30% used for testing. Moreover, some studies increase the training ratio for promising results. Furthermore, deep learning methods perform well when the optimal ratio for training and testing is set for proposed methods.

VI. Conclusion

In this paper, state of the art research for melanoma detection has been discussed. Moreover. open issues and challenges have been identified. Furthermore, this study analyzed in-depth several deep learning-based techniques such as fully convolution neural network, pre-trained model, ensemble and handcrafted methods to detect melanoma. It has observed that using deep learning techniques, there has no dire need for complex and composite pre-processing techniques such as image resize, crop and pixel value normalization. Proposed taxonomy and proposed model have been presented by exploring relevant studies. Moreover, this study categories the main shortcomings of the existing methods and point out the areas where further improvements are needed. The handcrafted methods showed better results than the conventional deep learning methods. In some studies, handcrafted features were used with the extraction functionality of preprocessing, and segmentation. Furthermore, labeling of images considered as the most important task in medical demographic image databases. A large number of labeled benchmark datasets like PH2, ISBI (2016, 2017, 2018 challenges), DermIS, Dermquest, Mednode and open-access datasets, have been available for researchers to evaluate their work. Moreover, Non-published / Non-Listed datasets were also available for the detection of melanoma. However, it become difficult to compare the results due to the diversified availability of datasets. In future work, the researcher must use a larger dataset by performing fine-tuning to hyper-parameters which can reduce the chances of overfitting. Moreover, CNN must learn to fetch data from dark-skinned people to achieve high accuracy. Moreover, age, gender, race must be added to achieve better results. However, increasing the accuracy rate remains an open challenge. The aim has to achieve maximum sensitivity while improving the specificity and overall accuracy of the methods.

TABLE V
Performance metrics

Dataset	Publication	Classifier	Sensitivity	Specifity	Precision	Accuracy	AUC
PH2 [38]	[31]	Transfer Learning	98.3%	98.9%	97.7%	98.6%	-----
PH2 [38]	[57]	AlexNet (TL) VGG16(TL)	100%	96.87%	-----	97.5%	-----
PH2 [38]	[62]	DCNN	72%	89%	-----	80.5%	-----
PH2 [38]	[63]	Full resolution convolutional network (FrCN) method for skin lesion segmentation	91.6%	96.5%	-----	94.6%	-----
PH2 [38]	[64]	High-resolution feature blocks (HRFB)	96.3%	94.2%	-----	94.9%	-----
PH2 [38]	[66]	3D color-texture feature (CTF)	98.1%	93.8%	-----	97.5%	-----

PH2 [38]	[70]	Deep depth wise separable residual convolutional algorithm	100%	-----	90%	96.7%	99.49%
PH2 [38]	[81]	Transfer learning	92.31%	94.12%	-----	93.33	-----
PH2 [38]	[82]	End to end and pixelwise learning using DCNN	93%	95%	-----	95%	-----
ISBI 2016[5]	[47]	FCNs based on VGG-16 and GoogLeNet	69.33%	93.75%	-----	88.92%	-----
ISBI 2016[5]	[54]	Transfer learning	90.5%	99.2%	92.1%	90.2%	89%
ISBI 2016[5]	[55]	Fusion of multiple visual features and deep-neural-network	93%	80%	-----	95%	96%
ISBI 2016[5]	[59]	Optimized colour feature(OCF) of lesion segmentation & DCNN	92%	90%	-----	92.1%	97%
ISBI 2016[5]	[61]	DCNN-FV (fusion)	-----	-----	68.49%	86.81%	85.2%
ISBI 2016[5]	[64]	High-resolution feature blocks (HRFB)	87%	96.4%]	-----	93.8%	-----
ISBI 2016[5]	[68]	Transfer learning	78.6%	84%	-----	81.3%	-----
ISBI 2016[5]	[73]	FCN + CNN	94%	93%	-----	92%	-----
ISBI 2016[5]	[75]	Dual deep CNN to mutually learn from each	-----	-----	67.3%	86.5%	82.5%
ISBI 2016[5]	[78]	Deep regional convolutional neural network (RCNN) and Fuzzy C-mean (FCM) clustering.	95%	94%	-----	94.2%	-----
ISBI 2016[5]	[95]	DenseNet with IcNR	94.2%	-----	94.4%	94.20%	98%

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3001507, IEEE Access

ISBI 2017[37]	[47]	FCNs based on VGG-16 and GoogLeNet	81.28%	86.22%	-----	85.3%	50%
ISBI 2017[37]	[50]	Linear discriminant analysis (LDA) + CNN	52%	97.4%	55.5%	85.8%	80.5%
ISBI 2017[37]	[52]	Transfer learning	95.51%	95.5%	95.5%	95.6%	98%
ISBI 2017[37]	[58]	Deep CNN & Data Augmentation	-----	-----	73.9%	89.0%	89.2%
ISBI 2017[37]	[59]	Optimized color feature (OCF) of lesion segmentation & DCNN	96.5%	97.0%	-----	96.5%	99%
ISBI 2017[37]	[63]	Full resolution convolutional network (FrCN) method for skin lesion segmentation	78.9%	96.%	-----	90.7%	----
ISBI 2017[37]	[64]	High-resolution feature blocks (HRFB)	87%	96.4%	-----	93.8%	-----
ISBI 2017[37]	[71]	CNN + RAPooling +RankOpt	60.7%	88.4	-----	83%	84.2%
ISBI 2017[37]	[77]	Two (FCRN) & LICU	49%	96.1%	72.9%	85.7%	91.2%
ISBI 2017[37]	[95]	DenseNet with IcNR	93.0%	-----	93.20%	93.4%	97%
ISBI 2018[39]	[48]	Transfer learning	80%	98.1%	-----	97%	-----
ISBI 2018[39]	[59]	Optimized colour feature (OCF) of lesion segmentation & DCNN	85.0%	84.0%	-----	85.1%	92%
ISBI 2018[39]	[60]	Ensemble (Resnet & inception V3)	79.6%	-----	86.2%	89.(%	-----
ISBI 2018[39]	[69]	Hand crafted deep learning ensemble technique	-----	-----	-----	-----	90%
ISBI 2018[39]	[93]	Hand crated and deep feature fusion	86.41	90	92.08	92.40	-----
ISBI 2019 [83]	[81]	Transfer learning	88.46%	88.24%	-----	88.33%	-----

ISIC Archive [35]	[67]	Skin Lesion Classification using CNN & Novel Regularizer	94.3%	93.6%	-----	97.49%	98%
ISIC Archive [35]	[70]	Deep depth wise separable residual convolutionl algorithm	99.3%	-----	99.6%	99.5%	99.4%
ISIC Archive[35]	[72]	Neural networks; Feature selection	84%	-----	83.2%	93%	-----
ISIC Archive [35]	[76]	Transfer learning and the pre-trained deep neural network	88.4%	93%	92.3%	95.9%	-----
ISIC Archive [35]	[94]	Ensemble (ResNet, DenseNet, SE-ResNext, NASNet)	80.46	96.57	85.02	-----	-----
Internet collected images	[56]	ECOC SVM with deep convolutional neural network	97.83%	90.74%	-----	94.2%	-----
Combined data sets (PH2 + ISBI,2016, + ISBI 2017)	[65]	Fusion of Alex net and VGG16	99.5%	98.4%	-----	99%	-----
Combined dataset (DermIS + Dermquest)	[74]	Optimized CNN	99.4%	94%	-----	93%	-----
Combined data sets (PH2+ ISIC archive, +ISIC challenge 2017)	[58]	Deep CNN & Data Augmentation	-----	-----	73.9%	89.0%	87.4%
Combined dataset(DermIS[29] + DermQuest[30])	[76]	Transfer learning and the pre-trained deep neural network	96.9%	95.6%	94.9%	96.3%	-----
Combined dataset (ISBI 2018[39] + ISBI 2019[83])	[80]	Ensemble (Resnet and Inception)	-----	-----	91%	91%	-----
DermIS[29]	[70]	Deep depth wise separable residual convolutional algorithm	100%	-----	91.66%	94.44%	-----
MED-NODE [32]	[70]	Deep depth wise separable residual convolutional algorithm	92.3%	-----	100%	95.2%	94.4%
MED-NODE [32]	[76]	Transfer learning and the pre-trained deep neural network	97.3%	97.4%	97.9%	97.7%	-----

MED-NODE [32]	[79]	Predict-Evaluate-Correct k-fold (PECK)	89%	93%	92%	91%	-----
Dermnet[34]	[94]	Ensemble (ResNet, DenseNet, SE-ResNext, NASNet)	79.94	98.40	79.82	-----	-----

APPENDIX

TABLE I
Publication source

Publication Source	Channel	Reference	No.	%
2016 IEEE International Conference on Image Processing (ICIP)	Conference	[19]	1	1.6
Nature Medicine 17	Journal	[20]	1	4.7
ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)	Conference	[21]	1	1.6
2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)	Conference	[22]	1	3.1
2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)	Conference	[23]	1	1.6
2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)	Conference	[24]	1	1.6
2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)	Symposium	[25]	1	4.7
IEEE Transactions on Medical Imaging (Volume: 36 , Issue: 4 , April 2017)	Journal	[26]	1	1.6
2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)	Conference	[27]	1	1.6
2016 IEEE Congress on Evolutionary Computation (CEC)	Conference	[28]	1	1.6
2018 9th Cairo International Biomedical Engineering Conference (CIBEC)	Conference	[31]	1	4.7
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2019	Conference	[40]	1	1.6
Journal of Medical Systems July 2019	Journal	[41]	1	1.6
Recent Trends in Signal and Image Processing 19	Conference	[43]	1	1.6
arXiv preprint arXiv:1901.10802	Journal	[44]	1	1.6
Computerized Medical Imaging and Graphics Volume 71, January 2019, Pages 19-29	Journal	[45]	1	1.6
Expert Systems with Applications Volume 118, 15 March 2019	Journal	[46]	1	1.6
International Journal of imaging system and technology 2019	Journal	[47]	1	1.6
International Visual Informatics Conference IVIC 2019: Advances in Visual Informatics	Conference	[48]	1	1.6
International Conference on Engineering Applications of Neural Networks EANN 2017: Engineering Applications of Neural Networks	Conference	[49]	1	1.6
Multimedia Tools and Applications May 2019	Journal	[50][55][56]	3	4.7

Biomedical Engineering Letters	Journal	[51]	1	1.6
2018 International Conference on Information and Communication Technology Convergence	Conference	[52]	1	1.6
IEEE International Conference on Acoustics, Speech and Signal Processing 2019	Conference	[53]	1	1.6
International Conference on Computer and Information Sciences 2019	Conference	[54]	1	1.6
ICACDS 2019: Advances in Computing and Data Sciences	Conference	[57]	1	1.6
Asian Conference on Intelligent Information and Database Systems				
ACIIDS 2018: Intelligent Information and Database Systems	Conference	[58]	1	1.6
Expert system	Journal	[59]	1	1.6
ICMLC 2018: Proceedings of the 2018 10th International Conference on Machine Learning and computing	Conference	[60]	1	1.6
Computerized Medical Imaging and Graphics				
Volume 71, January 2019, Pages 19-29	Journal	[61]	1	1.6
Computer Methods and Programs in Biomedicine	Journal	[62][63]	2	3.1
Pattern Recognition Letters	Journal	[95]	1	1.6
Volume 131, March 2020, Pages 63-70	Journal	[64]	1	1.6
Informatics in Medicine Unlocked				
Volume 17, 2019, 100176	Journal	[65]	1	1.6
IEEE Access	Journal	[66][70][79]	3	4.7
2017 13th IASTED International Conference on Biomedical Engineering (BioMed)	Conference	[67]	1	1.6
IEEE Journal of Biomedical and Health Informatics (Volume: 23 , Issue: 4 , July 2019)	Journal	[68]	1	1.6
IET Image Processing,Volume 13, Issue 12	Journal	[69]	1	1.6
In Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence (pp. 254-258). ACM	Journal	[71]	2	3.1
In Proceedings of the 2018 10th International Conference on Machine Learning and Computing (pp. 252-256). ACM.	Journal	[72]	1	1.6
Artificial Intelligence in Medicine, 102, 101756	Journal	[73]	1	1.6
In International Conference of Pioneering Computer Scientists, Engineers and Educators (pp. 214-222). Springer	Journal	[74]	1	1.6
In Asian Conference on Intelligent Information and Database Systems (pp. 573-582). Springer	Journal	[75]	1	1.6
PLOS one	Journal	[76]	1	1.6
Sensors,MPDI	Journal	[77][81]	2	3.1
International journal of medical informatics	Journal	[78]	1	1.6
Machine Learning with Health Care Perspective, Learning and Analytics in Intelligent Systems	Book	[80]	1	1.6
International Conference on Advanced Concepts for Intelligent Vision Systems	Conference	[82]	1	1.6
Entropy MDPI	Journal	[93]	1	1.6
Applied Sciences MDPI	Journal	[94]	1	1.6

TABLE II

		Quality assessment				
Reference	Classification		Quality Assessment			
	Publication Channel	Year	a	b	c	score
[19]	Conference	2016	1	1	2	4
[20]	Journal	2017	1	1	2	4
[21]	Journal	2017	1	1	0	2
[22]	Conference	2015	1	1	2	4
[23]	Conference	2015	1	1	2	4
[24]	Conference	2016	1	1	2	4
[25]	Symposium	2016	1	1	1.5	3.5
[26]	Journal	2017	1	1	2	4
[27]	Conference	2016	1	1	2	4
[28]	Conference	2016	1	1	2	4
[31]	Conference	2018	1	1	2	4
[40]	Conference	2019	1	1	2	4
[41]	Journal	2019	1	1	1.5	3.5
[43]	Book	2019	1	1	1.5	3.5
[44]	Journal	2019	1	1	0	2
[45]	Journal	2019	1	1	1	2
[46]	Journal	2019	1	1	2	4
[47]	Journal	2019	1	1	1.5	3.5
[48]	Conference	2019	1	1	2	4
[49]	Conference	2017	1	1	2	4
[50]	Journal	2019	1	1	2	4
[51]	Journal	2019	1	1	1	3
[52]	Conference	2018	1	1	2	4
[53]	Conference	2019	1	1	2	4
[54]	Conference	2019	1	1	2	4
[55]	Journal	2019	1	1	2	4
[56]	Journal	2018	1	1	2	4
[57]	Conference	2019	1	1	2	4
[58]	Conference	2018	1	1	2	4
[59]	Journal	2019	1	1	1.5	3.5
[60]	Conference	2018	1	1	2	4
[61]	Journal	2018	1	1	2	4
[62]	Symposium	2019	1	1	1.5	3.5
[63]	Journal	2018	1	1	2	4
[64]	Journal	2019	1	1	2	4
[65]	Journal	2019	1	1	2	4
[66]	Journal	2019	1	1	1	3
[67]	Journal	2019	1	1	2	4
[68]	Conference	2017	1	1	2	4
[69]	Journal	2019	1	1	2	4

[70]	Journal	2019	1	1	2	4
[71]	Journal	2018	1	1	2	4
[72]	Journal	2019	1	1	2	4
[73]	Journal	2018	1	1	2	4
[74]	Journal	2019	1	1	2	4
[75]	Conference	2019	1	1	2	4
[76]	Journal	2019	1	1	2	4
[77]	Journal	2018	1	1	1.5	3.5
[78]	Journal	2019	1	1	2	4
[79]	Journal	2020	1	1	2	4
[80]	book	2020	1	1	1.5	3.5
[81]	Journal	2020	1	1	1.5	3.5
[82]	Conference	2020	1	1	1.5	3.5
[93]	Journal	2020	1	1	2	4
[94]	Journal	2020	1	1	1.5	3.5
[95]	Journal	2020	1	1	2	4

REFERENCES

[1] Nami, N., Giannini, E., Burroni, M., Fimiani, M., & Rubegni, P. (2012). Teledermatology: state-of-the-art and future perspectives. Expert Review of Dermatology, 7(1), 1-3.

[2] Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., ... & Uhlmann, L. (2018). Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Annals of Oncology, 29(8), 1836-1842.

[3] FWHO | *Ultraviolet radiation*. Accessed Dec. 11, 2019. [Online].Available:http://www.who.int/uv/faq/skincancer/en/index1.html

[4] Oliveira, R. B., Papa, J. P., Pereira, A. S., & Tavares, J. M. R. (2018). Computational methods for pigmented skin lesion classification in images: review and future trends. Neural Computing and Applications, 29(3), 613-636.

[5] Gutman, D., Codella, N. C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., & Halpern, A. (2016). 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). arXiv preprint arXiv:1605.01397.

[6] Marchetti, M. A., Codella, N. C., Dusza, S. W., Gutman, D. A., Helba, B., Kalloo, A., ... & Jaimes, N. (2018). 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. Journal of the American Academy of Dermatology, 78(2), 270-277.

[7] Australian Institute of Health and Welfare. (2017). Cancer in Australia 2017. Cancer series no. 101. Cat. no. CAN 100.

[8] Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. CA: A Cancer Journal for Clinicians, 68(1), 7-30.

[9] Hu, Zilong, Jinshan Tang, Ziming Wang, Kai Zhang, Ling Zhang, and Qingling Sun. "Deep learning for image-based cancer detection and diagnosis− a survey." Pattern Recognition 83 (2018): 134-149.

[10] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[11] Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—addressing ethical challenges. The New England journal of medicine, 378(11), 981.

[12] American Skin cancer organization | *Skin Cancer Information* Accessed:Nov.29,2019.[Online].Avalable:https://www.skincancer.org/skin-cancer-information/skin-cancer-facts.

[13] Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. Information and software technology, 51(1), 7-15.

[14] Kitchenham, B. (2004). Procedures for undertaking systematic reviews: Joint technical report. Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd.(0400011T. 1).

[15] Sultana, N. N., & Puhan, N. B. (2018, January). Recent deep learning methods for melanoma detection: a review. In International Conference on Mathematics and Computing (pp. 118-132). Springer, Singapore.

[16] Centre for Reviews and Dissemination | *Home*. Accessed Nov. 29,2019.[Online].Available:http://www.york.ac.uk/inst/crd/faq4.htm

[17] Aryani, A., Peake, I. D., & Hamilton, M. (2010, September). Domain-based change propagation analysis: An enterprise system case study. In 2010 IEEE International Conference on Software Maintenance (pp. 1-9). IEEE.

[18] Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Annals of internal medicine, 151(4), W-65.

[19] Pomponiu, V., Nejati, H., & Cheung, N. M. (2016, September). Deepmole: Deep neural networks for skin mole lesion classification. In 2016 IEEE International Conference on Image Processing (ICIP) (pp. 2623-2627). IEEE.

[20] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[21] Mahbod, A., Schaefer, G., Wang, C., Ecker, R., & Ellinge, I. (2019, May). Skin lesion classification using hybrid deep neural networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1229-1233). IEEE.

[22] Masood, A., Al-Jumaily, A., & Anam, K. (2015, April). Self-supervised learning model for skin cancer diagnosis. In 2015

7th International IEEE/EMBS Conference on Neural Engineering (NER) (pp. 1012-1015). IEEE.

[23] Majtner, T., Yildirim-Yayilgan, S., & Hardeberg, J. Y. (2016, December). Combining deep learning and hand-crafted features for skin lesion classification. In 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1-6). IEEE.

[24] S Sabbaghi, S., Aldeen, M., & Garnavi, R. (2016, August). A deep bag-of-features model for the classification of melanomas in dermoscopy images. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1369-1372). IEEE.

[25] Demyanov, S., Chakravorty, R., Abedini, M., Halpern, A., & Garnavi, R. (2016, April). Classification of dermoscopy patterns using deep convolutional neural networks. In 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI) (pp. 364-368). IEEE.13th International Symposium on, IEEE2016, pp. 364-368.

[26] Yu, L., Chen, H., Dou, Q., Qin, J., & Heng, P. A. (2017). Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks. IEEE transactions on medical imaging, 36(4), 994.

[27] Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S. M. R., Jafari, M. H., Ward, K., & Najarian, K. (2016, August). Melanoma detection by analysis of clinical images using convolutional neural network. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1373-1376). IEEE.

[28] Sabouri, P., & GholamHosseini, H. (2016, July). Lesion border detection using deep learning. In 2016 IEEE Congress on Evolutionary Computation (CEC) (pp. 1416-1421). IEEE.

[29] DermIS Dermatology Information System | *Home* Accessed Nov. 30 2019.[Online]. Available: https://www.dermis.net

[30] DermQuest Online medical resource | *Home*. Accessed Dec. 01, 2019.[Online]. Available: http://www.dermquest.com.

[31] Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2018, December). Skin Cancer Classification using Deep Learning and Transfer Learning. In 2018 9th Cairo International Biomedical Engineering Conference (CIBEC) (pp. 90-93). IEEE.

[32] Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M. F., & Petkov, N. (2015). MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert systems with applications, 42(19), 6578-6585.

[33] Interactive dermatology atlas | *Home*. Accessed Dec. 01, 2019.[Online].Available:http://www.dermatlas.net/atlas/index.cfm.

[34] Dermnetnz online medical resources | *Home* Accessed Dec. 01, 2019. [Online].Available: http://www.dermnetnz.org.

[35] The International Skin Imaging Collaboration | *Challenges* Accessed Dec. 01, 2019. [Online].Available https://www.isicarchive.com/#!/topWithHeader/tightContentTop/challenges

[36] Brinker, T. J., Hekler, A., Utikal, J. S., Grabe, N., Schadendorf, D., Klode, J., ... & von Kalle, C. (2018). Skin cancer classification using convolutional neural networks: systematic review. Journal of medical Internet research, 20(10), e11936.

[37] Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... & Halpern, A. (2018, April). 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). IEEE.

[38] Mendonça, T., Ferreira, P. M., Marques, J. S., Marcal, A. R., & Rozeira, J. (2013, July). PH 2-A dermoscopic image database for research and benchmarking. In 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 5437-5440). IEEE.

[39] Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., ... & Kittler, H. (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368.

[40] Bisla, D., Choromanska, A., Berman, R. S., Stein, J. A., & Polsky, D. (2019). Towards Automated Melanoma Detection with Deep Learning: Data Purification and Augmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 0-0).

[41] Sreelatha, T., Subramanyam, M. V., & Prasad, M. G. (2019). Early Detection of Skin Cancer Using Melanoma Segmentation technique. Journal of medical systems, 43(7), 190.

[42] Ballerini L, Fisher RB, Aldridge RB, Rees J (2013) A color and texture based hierarchical KNN approach to the classification of non-melanoma skin lesions, color medical image analysis.In: Celebi ME, Schaefer G (eds) Lecture notes in computational vision and biomechanics, vol 6, pp 63–86.

[43] Mukherjee, S., Adhikari, A., & Roy, M. (2019). Malignant Melanoma Classification Using Cross-Platform Dataset with Deep Learning CNN Architecture. In Recent Trends in Signal and Image Processing (pp. 31-41). Springer, Singapore.

[44] Milton, M. A. A. (2019). Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Challenge. arXiv preprint arXiv:1901.10802.

[45] Mahbod, A., Schaefer, G., Ellinger, I., Ecker, R., Pitiot, A., & Wang, C. (2019). Fusing fine-tuned deep features for skin lesion classification. Computerized Medical Imaging and Graphics, 71, 19-29.

[46] Soudani, A., & Barhoumi, W. (2019). An image-based segmentation recommender using crowdsourcing and transfer learning for skin lesion extraction. Expert Systems with Applications, 118, 400-410.

[47] Jayapriya, K., & Jacob, I. J. (2019). Hybrid fully convolutional networks-based skin lesion segmentation and melanoma detection using deep feature. International Journal of Imaging Systems and Technology.

[48] Aldwgeri, A., & Abubacker, N. F. (2019, November). Ensemble of Deep Convolutional Neural Network for Skin Lesion Classification in Dermoscopy Images. In International Visual Informatics Conference (pp. 214-226). Springer, Cham.

[49] Georgakopoulos, S. V., Kottari, K., Delibasis, K., Plagianakos, V. P., & Maglogiannis, I. (2017, August). Detection of malignant melanomas in dermoscopic images using convolutional neural network with transfer learning. In International Conference on Engineering Applications of Neural Networks (pp. 404-414). Springer, Cham.

[50] Majtner, T., Yildirim-Yayilgan, S., & Hardeberg, J. Y. (2019). Optimised deep learning features for improved melanoma detection. Multimedia Tools and Applications, 78(9), 11883-11903.

[51] Gavrilov, D. A., Melerzanov, A. V., Shchelkunov, N. N., & Zakirov, E. I. (2019). Use of Neural Network-Based Deep Learning Techniques for the Diagnostics of Skin Diseases. Biomedical Engineering, 52(5), 348-352.

[52] Namozov, A., & Im Cho, Y. (2018, October). Convolutional Neural Network Algorithm with Parameterized Activation Function for Melanoma Classification. In 2018 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 417-419). IEEE.

[53] Mahbod, A., Schaefer, G., Wang, C., Ecker, R., & Ellinge, I. (2019, May). Skin lesion classification using hybrid deep neural networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1229-1233). IEEE.

[54] Khan, M. A., Javed, M. Y., Sharif, M., Saba, T., & Rehman, A. (2019, April). Multi-Model Deep Neural Network based Features Extraction and Optimal Selection Approach for Skin Lesion Classification. In 2019 International Conference on Computer and Information Sciences (ICCIS) (pp. 1-7). IEEE.

[55] Abbas, Q., & Celebi, M. E. (2019). DermoDeep-A classification of melanoma-nevus skin lesions using multi-feature fusion of visual features and deep neural network. Multimedia Tools and Applications, 1-22.

[56] Dorj, U. O., Lee, K. K., Choi, J. Y., & Lee, M. (2018). The skin cancer classification using deep convolutional neural network. Multimedia Tools and Applications, 77(8), 9909-9924.

[57] Gulati, S., & Bhogal, R. K. (2019, April). Detection of Malignant Melanoma Using Deep Learning. In International Conference on Advances in Computing and Data Sciences (pp. 312-325). Springer, Singapore.

[58] Pham, T. C., Luong, C. M., Visani, M., & Hoang, V. D. (2018, March). Deep CNN and data augmentation for skin lesion classification. In Asian Conference on Intelligent Information and Database Systems (pp. 573-582). Springer, Cham.1,2.

[59] Khan, M. A., Sharif, M. I., Raza, M., Anjum, A., Saba, T., & Shad, S. A. Skin lesion segmentation and classification: A unified framework of deep neural network features fusion and selection. Expert Systems.

[60] Shahin, A. H., Kamal, A., & Elattar, M. A. (2018, December). Deep ensemble learning for skin lesion classification from dermoscopic images. In 2018 9th Cairo International Biomedical Engineering Conference (CIBEC) (pp. 150-153). IEEE.

[61] Yu, Z., Jiang, X., Zhou, F., Qin, J., Ni, D., Chen, S., ... & Wang, T. (2018). Melanoma recognition in dermoscopy images via aggregated deep convolutional features. IEEE Transactions on Biomedical Engineering, 66(4), 1006-1016.

[62] Jianu, S. R. S., Ichim, L., & Popescu, D. (2019, March). Automatic Diagnosis of Skin Cancer Using Neural Networks. In 2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE) (pp. 1-4). IEEE.

[63] Al-Masni, M. A., Al-antari, M. A., Choi, M. T., Han, S. M., & Kim, T. S. (2018). Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. Computer methods and programs in biomedicine, 162, 221-231.

[64] Xie, F., Yang, J., Liu, J., Jiang, Z., Zheng, Y., & Wang, Y. (2019). Skin lesion segmentation using high-resolution convolutional neural network. Computer Methods and Programs in Biomedicine, 105241.

[65] Amin, J., Sharif, A., Gul, N., Anjum, M. A., Nisar, M. W., Azam, F., & Bukhari, S. A. C. (2019). Integrated Design of Deep Features Fusion for Localization and Classification of Skin Cancer. Pattern Recognition Letters.

[66] Warsi, F., Khanam, R., Kamya, S., & Suárez-Araujo, C. P. (2019). An efficient 3D color-texture feature and neural network technique for melanoma detection. Informatics in Medicine Unlocked, 100176.

[67] Albahar, M. A. (2019). Skin Lesion Classification Using Convolutional Neural Network With Novel Regularizer. IEEE Access, 7, 38306-38313.

[68] Lopez, A. R., Giro-i-Nieto, X., Burdick, J., & Marques, O. (2017, February). Skin lesion classification from dermoscopic images using deep learning techniques. In 2017 13th IASTED International Conference on Biomedical Engineering (BioMed) (pp. 49-54). IEEE.,

[69] Hagerty, J., Stanley, J., Almubarak, H., Lama, N., Kasmi, R., Guo, P., ... & Stoecker, W. V. (2019). Deep Learning and Handcrafted Method Fusion: Higher Diagnostic Accuracy for Melanoma Dermoscopy Images. IEEE journal of biomedical and health informatics.

[70] Sarkar, R., Chatterjee, C. C., & Hazra, A. (2019). Diagnosis of melanoma from dermoscopic images using a deep depthwise separable residual convolutional network. IET Image Processing, 13(12), 2130-2142.

[71] Yang, J., Xie, F., Fan, H., Jiang, Z., & Liu, J. (2018). Classification for dermoscopy images using convolutional neural networks based on region average pooling. IEEE Access, 6, 65130-65138.assification for Dermoscopy Images using Convolutional Neural Networks based on Region Average Pooling Jiawen Yang.

[72] Hasan, M., Barman, S. D., Islam, S., & Reza, A. W. (2019, April). Skin Cancer Detection Using Convolutional Neural Network. In Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence (pp. 254-258). ACM.

[73] Ge, Y., Li, B., Zhao, Y., Guan, E., & Yan, W. (2018, February). Melanoma segmentation and classification in clinical images using deep learning. In Proceedings of the 2018 10th International Conference on Machine Learning and Computing (pp. 252-256). ACM.

[74] Zhang, N., Cai, Y. X., Wang, Y. Y., Tian, Y. T., Wang, X. L., & Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. Artificial Intelligence in Medicine, 102, 101756.

[75] Wang, Y., Pan, H., Yang, B., Bian, X., & Cui, Q. (2019, September). Mutual Learning Model for Skin Lesion Classification. In International Conference of Pioneering Computer Scientists, Engineers and Educators (pp. 214-222). Springer, Singapore.

[76] Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2019). Classification of skin lesions using transfer learning and augmentation with Alex-net. PloS one, 14(5), e0217293.

[77] Li, Y., & Shen, L. (2018). Skin lesion analysis towards melanoma detection using deep learning network. Sensors, 18(2), 556.

[78] Nida, N., Irtaza, A., Javed, A., Yousaf, M. H., & Mahmood, M. T. (2019). Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. International journal of medical informatics, 124, 37-48.

[79] Albert, B. A. (2020). Deep Learning From Limited Training Data: Novel Segmentation and Ensemble Algorithms Applied to Automatic Melanoma Diagnosis. *IEEE Access*, *8*, 31254-31269.

[80] Nahata H., Singh S.P. (2020) Deep Learning Solutions for Skin Cancer Detection and Diagnosis. In: Jain V., Chatterjee J. (eds) Machine Learning with Health Care Perspective. Learning and Analytics in Intelligent Systems, vol 13. Springer, Cham

[81] El-Khatib, H., Popescu, D., & Ichim, L. (2020). Deep Learning–Based Methods for Automatic Diagnosis of Skin Lesions. *Sensors*, *20*(6), 1753.

[82] Adegun, A., & Viriri, S. (2020, February). Deep Convolutional Network-Based Framework for Melanoma Lesion Detection and Segmentation. In International Conference on Advanced Concepts for Intelligent Vision Systems (pp. 51-62). Springer, Cham.

[83] International Skin Imaging Collaboration, | ISIC 2019. Accessed April. 01,2020 [Online] Available: https://challenge2019.isic-archive.com

[84] Pathan, S., Prabhu, K. G., & Siddalingaswamy, P. C. (2018). Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review. *Biomedical Signal Processing and Control*, *39*, 237-262.

[85] Rathee, G., Sharma, A., Saini, H., Kumar, R., & Iqbal, R. (2019). A hybrid framework for multimedia data processing in IoT-healthcare using blockchain technology. Multimedia Tools and Applications, 1-23.

[86] Khan, M. Z., Harous, S., Hassan, S. U., Khan, M. U. G., Iqbal, R., & Mumtaz, S. (2019). Deep Unified Model For Face Recognition based on Convolution Neural Network and Edge Computing. IEEE Access, 7, 72622-72633.

[87] Bhatti, M. H., Khan, J., Khan, M. U. G., Iqbal, R., Aloqaily, M., Jararweh, Y., & Gupta, B. (2019). Soft computing-based EEG classification by optimal feature selection and neural networks. IEEE Transactions on Industrial Informatics, 15(10), 5747-5754.

[88] IEEE | IEEE Xplore, Accessed May. 04,2020 [Online] Available: https://ieeexplore.ieee.org/search/advanced

[89] Springer Link | Advance search, Accessed May. 04,2020 [Online] Available: https://link.springer.com/advanced-search

[90] ACM DIGITAL LIBRARY | Advance search, Accessed May. 04,2020 [Online] Available: https://dl.acm.org/search/advanced

[91] Wiley Online Library | Advance search, Accessed May. 04,2020 [Online] Available: https://onlinelibrary.wiley.com/search/advanced

[92] NCBI | Advance search, Accessed May. 04,2020 [Online] Available: https://www.ncbi.nlm.nih.gov/pubmed/advanced

[93] Almaraz-Damian, J. A., Ponomaryov, V., Sadovnychiy, S., & Castillejos-Fernandez, H. (2020). Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures. *Entropy*, *22*(4), 484.

[94] Bajwa, M. N., Muta, K., Malik, M. I., Siddiqui, S. A., Braun, S. A., Homey, B., ... & Ahmed, S. (2020). Computer-Aided Diagnosis of Skin Diseases Using Deep Neural Networks. *Applied Sciences*, *10*(7), 2488.

[95] Khan, M. A., Sharif, M., Akram, T., Bukhari, S. A. C., & Nayak, R. S. (2020). Developed Newton-Raphson based deep features selection framework for skin lesion recognition. *Pattern Recognition Letters*, *129*, 293-303.