A Critical Analysis of Computer Aided Approaches for Skin Cancer Screening

Minoj Selvarasa
Informatics Institute of Technology
Colombo 6, Sri Lanka
minojselvaraj@gmail.com

Achala Aponso
Informatics Institute of Technology
Colombo 6, Sri Lanka
ach.chathuranga@gmail.com

2020 International Conference on Image Processing and Robotics (ICIP) | 978-1-7281-6541-7/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICIP48927.2020.9367370

***Abstract*—Skin Cancer is life-threatening when diagnosed at a later stage. Early detection of skin cancers such as melanoma indicates a higher survival rate for the patient. Non-computer aided tools were used in the past such as the visual inspection using tools like the dermoscopy. Commercial tools were later introduced that allowed the examiners to examine the images obtained from the dermoscopy using techniques such as the ABCD rule and 7-point checklist. Deep Learning has proven to be the state-of-the-art for computer vision problems such as image classification. A lot of research has been carried out in the application of deep learning for automating skin cancer screening. This paper presents an analysis of the existing work carried out in the area of automatic skin cancer screening and the different steps involved in building a skin cancer classification tool for skin cancer screening. The limitations of the various existing approaches are explored, and the results of the analysis will be used as part of an ongoing research to design and develop a robust system that will address the identified cons.**

Keywords—Skin Cancer, Screening, Computer Vision, Deep Learning, Augmentation, Segmentation, CNN

I. INTRODUCTION

Cancer is a health hazard and is considered to be one of the deadliest diseases in the world. The mortality rate for cancer is high compared to that of other diseases [1]. Skin Cancer, a form of cancer, can be cured if diagnosed early but becomes lethal if not diagnosed early and allowed to grow [1]. According to World Health Organization (WHO), there is an increase in the number of skin cancer cases over the past decade [2]. In countries such as the United Kingdom and Brazil, the number of skin cancer cases have increased by 2-fold in just the past decade [3][4]. Countries such as the United States of America are experiencing a shortage in the number of experienced dermatologists [5][6]. A study carried out in Sri Lanka at the National Hospital of Sri Lanka (NHSL), shows that out of the 123 respondents that are doctors only 10 had received formal training on how to perform full body examination and out of the 13 doctors that had performed full body examinations, only 2 had carried out more than five examinations in the preceding 12 months [7]. This shortage of well-trained dermatologists calls for an automated skin screening tool that can aid dermatologists in the screening process.

II. PROBLEM & MOTIVATION

Unaided visual inspection was used for the screening of patients before the invention of dermoscopy. This approach only had an accuracy of about 60% [8]. The diagnostic accuracy among dermatologists with training under 5 years is very low but the accuracy is higher for those who have an experience of over 10 years [9]. Several clinicians have failed to adopt the algorithms that were developed to improve scalability in favor of experience. Tools were later developed to assist dermatologists in the screening of patients for skin cancer. Dermoscopy, as mentioned earlier, is one such popular tool that was adopted by dermatologists around the world.

Dermoscopy is a widely used non-invasive medical imaging technique that is used for skin cancer screening and early detection of skin cancer. The tool uses a microscope and incident light to visualize the subsurface features of the skin providing a dermoscopic image of the skin [8][9][10]. The dermoscopy shows a significant improvement in terms of accuracy when compared to that of unaided visual inspection. The accuracy increases from around 60% to 90% when used by a well-experienced dermatologist [8][9][10]. As mentioned earlier, the accuracy is subjective and high depends on the expertise of the dermatologist. The accuracy is no better than that of unaided visual inspection when the tool is used by a non-experienced practitioner [9][11]. The understanding of techniques such as ABCD rule, Menzies scoring method and 7-point checklist is also a requirement to use dermoscopy as such techniques are used to analyse the images obtained from the dermoscopy [11]. A study suggests the involvement of two or more experienced dermatologists when carrying out the skin examination for better accuracy [10]. But with the shortage of well trained and experienced dermatologists, this is a major issue. In such scenarios, an automated skin screening tool can aid dermatologists in the screening process allowing them to screen many patients accurately in a short period of time. The number of misdiagnoses can also be reduced by utilizing such an approach.

III. EXISTING WORK

Several researches have been carried out in the application of Machine Learning for the classification of cancer diseases. This successfully automates the skin screening process. Most of these approaches have utilized dermoscopic images for the training process.

The use of a machine learning approach involves several steps as part of the classification process. This review also analyses the different approaches that have been explored for each of these steps involved in constructing an automated skin screening tool while also identifying the main limitations faced during these researches.

The initial step carried out is data gathering where data related to the problem being solved needs to be gathered. Here, skin cancer related datasets need to be collected in order to carry out the classification. The next step would be to pre-process and prepare the images. In case of dataset imbalances and lack of data, data augmentation can be carried out which can drastically improve the classification accuracy as well.

Authorized licensed use limited to: East Carolina University. Downloaded on June 16,2021 at 16:01:22 UTC from IEEE Xplore. Restrictions apply.

This is because when techniques such as deep learning is used, they tend to benefit from a larger dataset.

A. The Dataset

Unlike other approaches such as the use of other simpler techniques like the ABCD rule [12], the application of machine learning heavily relies on the dataset. Dataset containing images that represent all the different diseases that needs to be classified is required. Lack of datasets related to skin cancer or very little images were an early limitation that was experienced by several researches. But due to the recent surge in interest in skin cancer screening using machine learning has given rise to larger datasets. The most popular dataset that consists of a large collection of images across several diseases is the ISIC 2019 dataset [13][14] which is was prepared as part of the skin image analysis challenge hosted by the International Skin Imaging Collaboration (ISIC). The dataset consists of over 25,000 images across 8 different skin diseases that are both malignant and benign. Other datasets that are much smaller include the SD-198, Derm7pt, and DermNet which all consists of a smaller number of images with lesser variation in the diseases.

B. Preprocessing

Preprocessing is an important step in machine learning. Data needs to be prepared for the consumption of the classification algorithm. The importance of pre-processing steps has been explored by Hoshyar, Al-Jumaily and Hoshyar in a comparison carried out between the different preprocessing steps that skin cancer classification can benefit from [18]. Another review presents the important steps that needs to be carried out in skin cancer classification and specifically highlights on the importance of pre-processing for skin cancer classification [15]. Pre-processing is considered important as it allows to improve the quality of the image. A good selection of pre-processing steps can greatly improve the accuracy of the classification that is to be carried out later [16].

1) Image Scaling

Image Scaling is the first step that is carried out in improving the image. It is important to scale the images as the images are gathered from multiple data sources and can be in different sizes, therefore all the images need to be resized to the same size [17].

2) Contrast Enhancement

Contrast enhancement is a pre-processing technique carried out so that it can help to differentiate the lesion and the surrounding skin [15][18][19]. There are two types of contrast enhancement methods, namely “Linear Contrast Enhancement” and “Non-Linear Contrast Enhancement”. Non-Linear Contrast Enhancement is most suitable for preprocessing skin lesion images according to Hoshyar, Al-Jumaily and Hoshyar.

3) Artefact Removal

Another pre-processing step would be to remove artefacts such as hair and rulers (if present). This is crucial as the presence of hair and details that are not visualized properly can lead to misclassifications, which can be an expensive error to both the doctor and patient [15][18][20]. Mathematical morphological methods, inpainting based approach and an automated software called DullRazor are the most commonly used approaches [21].

C. Image Segmentation

Image segmentation can be considered to be a part of preprocessing or separate. This step helps to differentiate the Region of Interest (ROI) from the skin. The output of the method is fed to the network to perform the classification [15][19]. The application of Deep Learning for skin lesion segmentation has been explored by several researchers [6], [11][20][22]. The use of U-Net architecture has been explored by a few researchers for image segmentation. The architecture is most suitable for image segmentation tasks and displays exceptional performance in most image segmentation tasks [13][14][20][22]. A modified U-Net architecture was explored by Bisla et al. for image segmentation. The network displays performance that is better than the traditional baseline and the previous ISIC challenge winners [20].

D. Image Augmentation

Although the previously explored dataset, ISIC, consists of over 25,000 images, it is still considered to be a smaller dataset as computer vision approaches such as deep learning perform well only in the presence of a much larger dataset. Also, the ISIC dataset presents a huge imbalance in the different diseases present in the dataset. The presence of a larger dataset can help reduce overfitting when deep learning approaches are used. Therefore, this presents a need for data augmentation [23][24]. The data augmentation process allows to synthetically generate data.

Several augmentation techniques exist. The simplest augmentation technique for images being simple random transformations such as geometric and color augmentations [23]. More complex augmentation techniques such as the use of Generative Adversarial Networks (GANs) [20][23] and Variational Auto-Encoders have been explored by researchers. Using complex augmentation techniques such as the use of GANs result in better variation in the images. Another augmentation technique approached in [23] is learning the augmentation using Neural Networks. A comparison carried out in [23] show that image augmentation results in better accuracy when compared with no augmentation. The other two approaches, learning the augmentation and using GANs, seem to perform on par with traditional transformation. The authors do suggest a combination of augmentation techniques.

E. Classification

1) Non-Deep Learning approaches

a) Support Vector Machines

Support Vector Machines (SVM) are a class of algorithms that has been explored in the area of computer vision. SVM aims to determine a hyperplane based on the provided data. The hyperplane divides the data into defined number of classes. The process of learning the optimal decision boundary that divides the data points into its respective classes is called training the SVM [25]. Feature extraction is an important step when SVMs are used. Features need to be extracted from the images. Automated approaches and/or hand-crafted features need to be used with SVMs. Feature extraction algorithms such as RSurf, Local Binary Patterns and CNN are used with SVM. [26][27]. As the dataset sizes are small, the results from these approaches are promising but also requires manual feature extraction.

Authorized licensed use limited to: East Carolina University. Downloaded on June 16,2021 at 16:01:22 UTC from IEEE Xplore. Restrictions apply.

b) K-Nearest Neighbor

KNN is a simple supervised algorithm that is used for both classification and regression. The algorithm assumes that similar things exist in close proximity. Therefore, the algorithm is based on the idea of similarity. KNN is suitable when classifying images using features described by local features [28]. Similarity functions are required to perform the image classification. A research shows that CNN performs better than kNN in terms of overall accuracy [29]. When results from other approaches that have been explored in this paper show that these other approaches outperform a kNN only approach.

2) Deep Learning approaches

a) Training from Scratch

A deep learning algorithm is trained from scratch using a skin lesion dataset. This would require the architecture to be defined manually or existing architectures such as VGG16 can be used. This will require a large dataset. The research by Romero-Lopez et al. show does not produce a good enough performance in comparison to other approaches explored. This is mainly because there are no large datasets available for skin lesions [30]. Another approach explored in [11] explores the use of CNNs for automated image segmentation. An ensemble of Fully Convolutional Networks (FCN) has also been explored in this approach which resulted in better performance over individual networks [11].

b) Ensembling

This approach uses a combination of deep learning and machine learning algorithm. This research shows a comparison between hand crafted features and automated feature extraction using CNN with SVM as the classifier. The results from the research show that the automated feature extraction method outperforms handcrafted features [27]. Another research by Romero-Lopez et al. have explored the use of pretrained deep learning models for feature extraction. Very few pretrained models have been explored in this research: AlexNet, variation of VGG16 and a variation of ResNet-18. SVMs have been used as the classifier. The outputs from the three individual SVMs are then fused together [30]. The usage of CNN for feature extraction and SVM as the classifier has been explored in [27]. Transfer Learning has been used for the feature extractor network. The use of CNN as feature extractor has resulted in better results in comparison to hand tuned features [27].

c) Transfer Learning

Transfer Learning is an approach where a model that has been previously trained on a large dataset is retrained on a more domain specific dataset. A lot of research has been and is being carried out in this area as it allows for better models with less data. The layers that learn more specialized features are retrained to learn more specialized features of the dataset that the model is being retrained on. Research has been carried out where Transfer Learning is used for both the classification and feature extraction [30]. A study by Romero-Lopez et al show a comparison between three different approaches: I. Training an existing architecture from scratch, II. Leveraging from a pretrained VGGNet (trained on ImageNet dataset) and III. Transfer learning and fine tuning the CNN architecture. Results show that the third method performs better. This is mainly because the dataset used for the training is small and therefore when training from scratch, the algorithm fails [30]. A study by Menegola et al. also show a comparison of multiple approaches that involve transfer learning and hybrid approaches. The comparison shows the results obtain by training a network from scratch, transferring knowledge from a network trained on a related domain (diabetic retinopathy) and transferring knowledge from a network trained on the ImageNet dataset [31]. Fine-tuning has also been explored alongside transfer learning in the above approach. The results show that transferring from a related domain doesn't have much of an impact as using the knowledge from the ImageNet dataset has resulted in better accuracy. Also, according to the research, fine tuning has a major positive impact on the accuracy [31]. Another approach has explored the use of Transfer Learning as part of the feature extractor. The results show that the performance from the use of Transfer Learning is better than that of state-of-the-art approaches based on hand-crafted features [24].

IV. Conclusion

All the steps involved in constructing an automated tool for skin cancer screening has been explored. It is evident from the researches explored previously that lack of data is one of the main limitations identified by most of the researchers when carrying out tasks such as classification. This adversely affects the performance of the classifier as most deep learning classifiers tend to work well in the presence of a larger number of images. This limitation can be addressed by either collecting more data, which is cumbersome, or by the use of data augmentation techniques. As one of the researchers have pointed out a combination of augmentation techniques that involve the use of Generative Adversarial Networks (GAN) and traditional transformation can be explored for this purpose. Another limitation would be the absence of parameter tuning. Although model parameter tuning has been explored the tuning of hyperparameters have not been explored. The architecture of a network can be considered a hyperparameter. The fine tuning of the architecture has shown to result in better performance in other areas of research. Common approach used for fine tuning an architecture is Neural Architecture Search (NAS). Fine tuning of the architecture helps build architectures that are more specific to the dataset and therefore perform exceptionally well. But the tuning of architectures is a very GPU intensive task and therefore requires a massive number of GPUs. Further research can be carried out on how such an approach can be used for designing the architecture for the classification network for skin cancer classification and a combination of data augmentation techniques can also be explored to address the lack of data and class imbalances. With the knowledge gained by identifying the limitations of the existing work that has been carried out, a tool can be built that addresses the limitations.

References

[1] Cancer Statistics", National Cancer Institute, 2020. [Online]. Available: https://www.cancer.gov/about-cancer/understanding/statistics. [Accessed: 06- Feb- 2020].

[2] "Skin cancers", World Health Organization, 2020. [Online]. Available: https://www.who.int/uv/faq/skincancer/en/index1.html. [Accessed: 06- Feb- 2020].

[3] F. Robertson and L. Fitzgerald, "Skin cancer in the youth population of the United Kingdom", Journal of Cancer Policy, vol. 12, pp. 67-71,

Authorized licensed use limited to: East Carolina University. Downloaded on June 16,2021 at 16:01:22 UTC from IEEE Xplore. Restrictions apply.

2017. Available: https://www.sciencedirect.com/science/article/abs/pii/S2213538316300479. [Accessed 6 February 2020].

[4] P. Minango, Y. Iano, A. Borges Monteiro, R. Padilha França and G. Gomes de Oliveira, "Classification of Automatic Skin Lesions from Dermoscopic Images Utilizing Deep Learning", SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING, vol. 2019, no. 1, pp. 107-114, 2019. Available: 10.18580/setijbe.2019.14 [Accessed 5 February 2020].

[5] A. Glazer and D. Rigel, "Analysis of Trends in Geographic Distribution of US Dermatology Workforce Density", JAMA Dermatology, vol. 153, no. 5, p. 472, 2017. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5470415/. [Accessed 6 February 2020].

[6] N. Codella et al., "Deep learning ensembles for melanoma recognition in dermoscopy images", IBM Journal of Research and Development, vol. 61, no. 45, pp. 5:1-5:15, 2017. Available: https://ieeexplore.ieee.org/abstract/document/8030303. [Accessed 6 February 2020].

[7] H. Herath et al., "Knowledge, attitudes and skills in melanoma diagnosis among doctors: a cross sectional study from Sri Lanka", BMC Research Notes, vol. 11, no. 1, 2018. Available: https://bmcresnotes.biomedcentral.com/articles/10.1186/s13104-018-3499-y. [Accessed 6 February 2020].

[8] C. Barata, M. Celebi and J. Marques, "A Survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer", IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 3, pp. 1096-1109, 2019. Available: https://ieeexplore.ieee.org/abstract/document/8377976. [Accessed 6 February 2020].

[9] J. Dinnes et al., "Tests to assist in the diagnosis of cutaneous melanoma in adults: a generic protocol", Cochrane Database of Systematic Reviews, 2015. Available: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011902/full. [Accessed 6 February 2020].

[10] H. Kittler, H. Pehamberger, K. Wolff and M. Binder, "Diagnostic accuracy of dermoscopy", The Lancet Oncology, vol. 3, no. 3, pp. 159-165, 2002. Available: https://www.ncbi.nlm.nih.gov/pubmed/11902502. [Accessed 6 February 2020].

[11] Y. Yuan, M. Chao and Y. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance", IEEE Transactions on Medical Imaging, vol. 36, no. 9, pp. 1876-1886, 2017. Available: https://ieeexplore.ieee.org/document/7903636. [Accessed 7 February 2020].

[12] R. Johr, "Dermoscopy: alternative melanocytic algorithms—the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist", Clinics in Dermatology, vol. 20, no. 3, pp. 240-247, 2002. Available: 10.1016/s0738-081x(02)00236-5 [Accessed 5 February 2020].

[13] N. Codella et al., "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)", 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018.

[14] N. Codella et al., "Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)", Arxiv, 2019. Available: https://arxiv.org/abs/1902.03368. [Accessed 6 February 2020].

[15] P. Mehta and B. Shah, "Review on Techniques and Steps of Computer Aided Skin Cancer Diagnosis", Procedia Computer Science, vol. 85, pp. 309-316, 2016. Available: https://www.sciencedirect.com/science/article/pii/S1877050916305865. [Accessed 6 February 2020].

[16] J. D. Narain Ponraj, "A Survey on the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer", Citeseerx.ist.psu.edu, 2011. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.651.592. [Accessed: 06- Feb- 2020].

[17] A. Grigoryan and S. Agaian, "A New Measure of Image Enhancement", Citeseerx.ist.psu.edu, 2000. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.4021. [Accessed: 06- Feb- 2020].

[18] A. Hoshyar, A. Al-Jumaily and A. Hoshyar, "The Beneficial Techniques in Preprocessing Step of Skin Cancer Detection System Comparing", Procedia Computer Science, vol. 42, pp. 25-31, 2014. Available: https://www.sciencedirect.com/science/article/pii/S1877050914014677. [Accessed 6 February 2020].

[19] S. Jain, V. jagtap and N. Pise, "Computer Aided Melanoma Skin Cancer Detection Using Image Processing", Procedia Computer Science, vol. 48, pp. 735-740, 2015. Available: https://www.sciencedirect.com/science/article/pii/S1877050915007188. [Accessed 5 February 2020].

[20] D. Bisla, A. Choromanska, J. Stein, D. Polsky and R. Berman, "Towards Automated Melanoma Detection with Deep Learning: Data Purification and Augmentation", Arxiv, 2019. Available: http://arxiv.org/abs/1902.06061. [Accessed 5 February 2020].

[21] T. Lee, V. Ng, R. Gallagher, A. Coldman and D. McLean, "Dullrazor®: A software approach to hair removal from images", Computers in Biology and Medicine, vol. 27, no. 6, pp. 533-543, 1997. Available: https://www.sciencedirect.com/science/article/abs/pii/S0010482597000206. [Accessed 6 February 2020].

[22] J. Ebenezer and J. Rajapakse, "Automatic segmentation of skin lesions using deep learning", Arxiv, 2018. Available: https://arxiv.org/abs/1807.04893. [Accessed 6 February 2020].

[23] L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning", Arxiv, 2017. Available: https://arxiv.org/abs/1712.04621. [Accessed 7 February 2020].

[24] V. Pomponiu, H. Nejati and N. Cheung, "Deepmole: Deep neural networks for skin mole lesion classification", 2016 IEEE International Conference on Image Processing (ICIP), 2016. Available: https://ieeexplore.ieee.org/document/7532834. [Accessed 7 February 2020].

[25] U. Maulik and D. Chakraborty, "Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques", IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 1, pp. 33-52, 2017. Available: https://ieeexplore.ieee.org/document/7882747. [Accessed 7 February 2020].

[26] N. Das, A. Pal, S. Mazumder, S. Sarkar, D. Gangopadhyay and M. Nasipuri, "An SVM Based Skin Disease Identification Using Local Binary Patterns", 2013 Third International Conference on Advances in Computing and Communications, 2013. Available: https://ieeexplore.ieee.org/document/6686372. [Accessed 6 February 2020].

[27] T. Majtner, S. Yildirim-Yayilgan and J. Hardeberg, "Combining deep learning and hand-crafted features for skin lesion classification", 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), 2016. Available: https://ieeexplore.ieee.org/document/7821017. [Accessed 7 February 2020].

[28] G. Amato and F. Falchi, "kNN based image classification relying on local feature similarity", Proceedings of the Third International Conference on SImilarity Search and APplications - SISAP '10, 2010. Available: https://dl.acm.org/doi/10.1145/1862344.1862360. [Accessed 7 February 2020].

[29] S. Yu, S. Jia and C. Xu, "Convolutional neural networks for hyperspectral image classification", Neurocomputing, vol. 219, pp. 88-98, 2017. Available: https://www.sciencedirect.com/science/article/abs/pii/S0925231216310104. [Accessed 6 February 2020].

[30] A. Romero-Lopez, X. Giro-i-Nieto, J. Burdick and O. Marques, "Skin Lesion Classification from Dermoscopic Images Using Deep Learning Techniques", Biomedical Engineering, 2017. Available: https://ieeexplore.ieee.org/document/7893267. [Accessed 8 February 2020].

[31] A. Menegola, M. Fornaciali, R. Pires, F. Bittencourt, S. Avila and E. Valle, "Knowledge transfer for melanoma screening with deep learning", 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), 2017. Available: https://ieeexplore.ieee.org/document/7950523?section=abstract. [Accessed 11 February 2020].

Authorized licensed use limited to: East Carolina University. Downloaded on June 16,2021 at 16:01:22 UTC from IEEE Xplore. Restrictions apply.