Skin Cancer Detection Based on Deep Learning

Reza Ahmadi Mehr[1], Ali Ameri[2]*

ABSTRACT

Background: The conventional procedure of skin-related disease detection is a visual inspection by a dermatologist or a primary care clinician, using a dermatoscope. The suspected patients with early signs of skin cancer are referred for biopsy and histopathological examination to ensure the correct diagnosis and the best treatment. Recent advancements in deep convolutional neural networks (CNNs) have achieved excellent performance in automated skin cancer classification with accuracy similar to that of dermatologists. However, such improvements are yet to bring about a clinically trusted and popular system for skin cancer detection.

Objective: This study aimed to propose viable deep learning (DL) based method for the detection of skin cancer in lesion images, to help physicians in diagnosis.

Material and Methods: In this analytical study, a novel DL based model was proposed, in which other than the lesion image, the patient's data, including the anatomical site of the lesion, age, and gender were used as the model input to predict the type of the lesion. An Inception-ResNet-v2 CNN pretrained for object recognition was employed in the proposed model.

Results: Based on the results, the proposed method achieved promising performance for various skin conditions, and also using the patient's metadata in addition to the lesion image for classification improved the classification accuracy by at least 5% in all cases investigated. On a dataset of 57536 dermoscopic images, the proposed approach achieved an accuracy of 89.3%±1.1% in the discrimination of 4 major skin conditions and 94.5%±0.9% in the classification of benign vs. malignant lesions.

Conclusion: The promising results highlight the efficacy of the proposed approach and indicate that the inclusion of the patient's metadata with the lesion image can enhance the skin cancer detection performance.

Introduction

The most commonplace neoplasms in humans are skin cancers [1,2], which are mainly divided into two types: melanoma and non-melanoma. However, melanoma (the most threatening type of skin cancer) is not common compared to other malignant types, it is more likely to spread than other skin cancers [3,4]. In 2020, an estimation 324,635 new registered cases of melanoma from 185 countries was diagnosed, of which 57,043 cases led to death [2].

Non-melanoma skin cancer (NMSC) mainly consists of Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC), grouped and called "keratinocyte carcinomas" (KC) due to the formation of a type of skin cell called keratinocyte. Histologically, approximately 70% of

[1]MSc, Department of Biomedical Engineering and Medical Physics, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2]PhD, Department of Biomedical Engineering and Medical Physics, Shahid Beheshti University of Medical Sciences, Tehran, Iran

*Corresponding author:
Ali Ameri
Department of Biomedical Engineering and Medical Physics, Shahid Beheshti University of Medical Sciences, Tehran, Iran
E-mail: ali.ameri@unb.ca

NMSCs are classified as BCC and 20% as SCC [5,6]. It is also estimated that 5.4 million new cases of KCs occur in the US, annually [1,2]. In 2020, among all types of NMSC, 63,731 individuals from 185 countries died worldwide [2].

By 2017, melanoma led to approximately 75% of all skin cancer-related deaths [7]. However, even though the incidence rate of melanoma is on the rise, it no longer accounts for more than 50% of all skin cancer-related deaths [2]. Early detection of skin cancers is effective for their treatment. Most cases of NMSCs can be cured, especially in the beginning stages. Melanoma is also highly curable when detected in its earliest stages.

The detection of melanoma using no special visual aid equipment by expert dermatologists has about 60% of accuracy [7]. The detection performance is even lower for primary care clinicians with only 23-46% accuracy [8]. However, the recent developments of deep-learning-based methods have led to increased performance in medical and non-medical fields. Further, they can assist dermatologists to track skin lesions in images to detect cancer earlier.

Images are applied in three conventional modalities for skin lesions: dermoscopic, histological, and photographic images. Dermoscopy images are obtained by a specialized instrument, leading to high-resolution skin imaging with a decrease in the skin surface reflectance [9]. Histological images are acquired via invasive biopsy and microscopy [7]. Both dermoscopy and histological methods result in great standardized images, while simple photographic images (using smartphones and cameras) show discrepancies in the zoom, angle, and lighting and may contain irrelevant backgrounds, making automated classification significantly more challenging [10]. To overcome this challenge, we require a data-driven approach using millions of pre-training and training images to make classification robust to photographic variability [11].

In the recognition of skin cancer, the main challenge is to accurately diagnose malignant versus benign lesions with the same etiology and similarities in shape, border, and color, which is complicated for dermatologists, since both are derived from melanocyte cells. The same challenge exists for distinguishing between malignant keratinocyte carcinoma (BCC and SCC) versus benign keratosis [7]. Moreover, malignant cutaneous lymphomas versus inflammatory non-neoplastic eczema and dermatitis are not easily differentiated [12]. Malignant dermal lesions (Kaposi sarcoma) are also difficult to be distinguished from benign dermal lesions (e.g., dermatofibroma and vascular lesions) [7]. Since the visual inspection contains errors, biopsy and histopathological examination is the gold standard of diagnosis. A discernment between these eight categories will cover approximately 97% of the incidence rate of all skin-related cancers.

The present analytical study aimed to introduce a new viable DL based approach for the automatic prediction of skin lesion type as a tool, which can help physicians interpret lesion images. Additionally, the proposed model used the patient data, such as the anatomical site of the lesion, age, gender, and the lesion image to predict the lesion type. Recent studies [1,2,13] have shown that the skin cancer incidence rate increases with age and men are 10% more likely to develop melanoma skin cancer than women and are 4% more likely to die from melanoma than women. Also, a correlation between the lesion type and its anatomical site on the body has been found [1,2,13]. These findings motivated us to investigate the impact of including this information (age, gender, and anatomical site) as the input of our automated model for skin cancer detection.

Material and Methods

Methods

In this analytical study, we focus on the most

critical and widespread types of malignant lesions (approximately 97% of the incidence rates of all skin-related cancers). Figure 1 shows various skin lesions demonstrating the difficulty in distinguishing between malignant and benign lesions. The proposed method conducts two tasks: firstly, binary classification of malignant vs. benign lesions and secondly, a more detailed classification between eight skin conditions, including 4 benign and 4 malignant lesion types as seen in Figure 2.

Datasets

Three databases, including the International Skin Imaging Collaboration (ISIC) Dermoscopic Archive released for 2019 [14-16] and 2020 [17,18] melanoma detection challenges,

the PAD-UFES-20 [19,20] and a part of images from the Fitzpatrick17k [21] were used to train and evaluate the proposed system. ISIC 2019 consists of images with nine different diagnostic categories. All images are dermoscopic, biopsy-proven, and annotated as malignant or benign, and ISIC 2020 has images of unique benign and malignant skin lesions from over 2,000 patients. All malignant diagnoses were confirmed via histopathology; benign diagnoses were confirmed using either expert agreement, longitudinal follow-up, or histopathology. Both ISIC datasets include metadata with extra information about the gender and age of the patients and the anatomical site of the lesion on the body. The PAD-UFES-20 consists of photographic samples of six different types
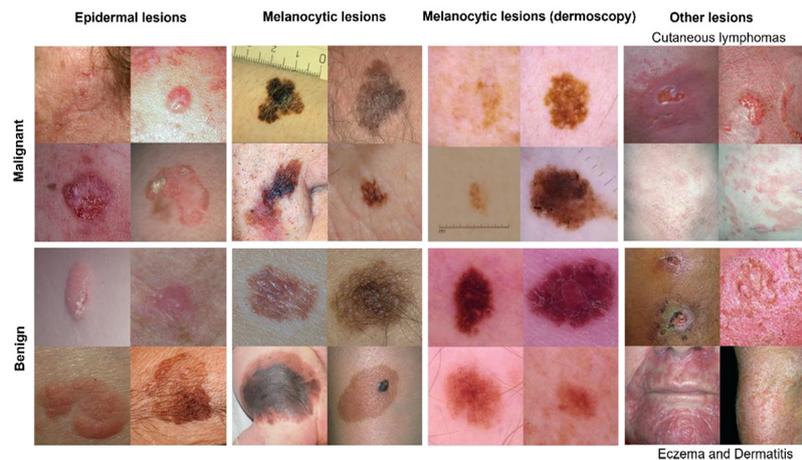

Figure 1: Example images from the dataset; both photographic and dermoscopic images are shown for melanocytic lesions to visualize the difference.
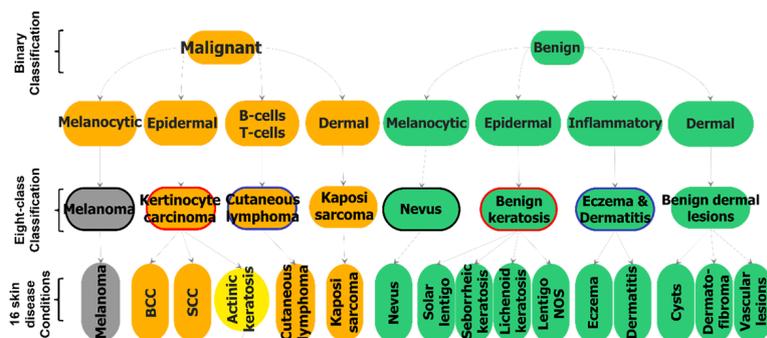

Figure 2: A schematic diagram of important skin lesions (orange: malignant, green: benign, gray: melanoma, yellow: actinic keratosis, a pre-malignant condition that is an early form of squamous cell carcinoma).

of skin lesions so that each image is defined as whether it is biopsied or not, while all malignant lesions are biopsy-proven. This dataset also provides metadata, such as the gender and age of the patient and the anatomical site of the lesion. The Fitzpatrick17k offers a three-partition label for photographic images, including malignant, benign, and non-neoplastic lesions. However, metadata were not provided in Fitzpatrick17k. Our dataset, derived from the three databases described above comprises 66735 clinical images representing 16 different skin-disease conditions, including 58031 and 8704 dermoscopy and photographic images, respectively.

Training algorithm

The Inception-ResNet-v2 convolutional neural network (CNN) was used to classify lesions due to its best performance for object recognition compared to all state-of-the-art CNN models, demonstrating low computational cost and high accuracy [22]. Inception-ResNet-v2 was pre-trained for object recognition on 1.28 million images (1,000 object categories) from the ImageNet Large Scale Visual Recognition Challenge. For transfer learning, the classification layer was replaced to account for the corresponding number of classes. The network was trained on our dataset, while all trainable parameters were fine-tuned across all layers. Inception-ResNet-v2 takes as input, images of 299×299 pixels with normalized pixel values between -1 and 1. Consequently, all images were resized to 299×299 pixels and the values of pixels were scaled between -1 and 1 before their feeding into the network. The CNN was trained with backpropagation using the Adam optimizer with a global learning rate of 0.001 and a decay factor of 16 every 30 epochs. For other arguments of Adam optimizer, beta_1, beta_2, and epsilon were set to 0.9, 0.999, and 0.1, respectively. Training and testing of the network were performed using Keras and Google's TensorFlow deep learning frameworks. For data augmentation, a rotation (with a random angle between $0°$ and $90°$), and also flipping (either vertical or horizontal with a probability of 0.5) of the original, and rotated images were conducted so that the number of images increased by a factor of 4.

The performance of the proposed model was assessed using a 10-fold cross validation approach without any overlap of the training or test set, i.e. no image was included in both training and test sets. For each skin disease, an equal number of images were used across different folds. Note that most images were dermoscopic and biopsy-proven, and only some skin conditions had only-photographic images (from the Fitzpatrick17k dataset), which were expert-annotated. These skin conditions were cutaneous lymphomas, Kaposi sarcoma, eczema, dermatitis, and benign dermal cysts.

Classification approaches

Two approaches were implemented as follows: 1) denoted as direct, a CNN was directly trained and tested for either binary or 8-class classification (Figure 3a, b), 2) denoted as probability-based (PB), a CNN was trained to classify 16 different diseases; for the test, the probabilities of each 16 classes were used to conduct the binary and 8 class classification tasks by summing up the corresponding children nodes' probabilities shown in Figure 2; a diagram of this approach is presented in (Figure 3c). As novel contribution of this work, a CNN was applied to metadata embedded lesion images, using either of the two sub-approaches, namely direct (denoted as metadata direct) or probability-based (denoted as metadata-PB). The patient's metadata, including the gender, age, and anatomical site of the lesion, were embedded in the lesion image. The proposed approach encoded the metadata related to each image using a pseudo-QR-code to replace the image's top-left corner pixels, which were always backgrounded without relevant information.

Both the ISIC dermoscopic archives and the PAD-UFES-20 offer patients' metadata related
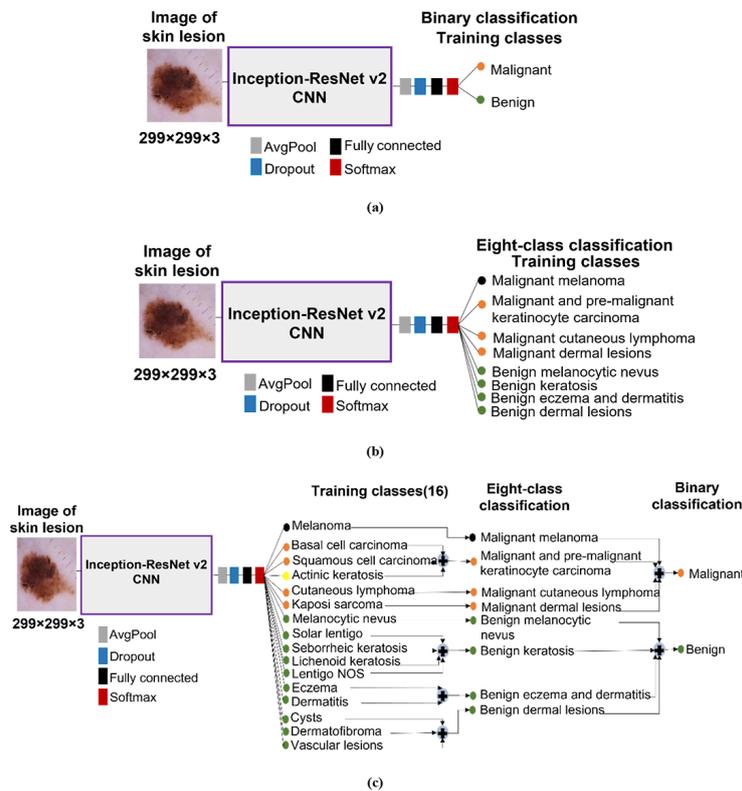
Figure 3: The workflow of (a) direct approach: binary classification, (b) direct approach: 8-class classification, and (c) Probability-based (PB) approach: the Convolutional Neural Network (CNN) trained on 16 skin disease conditions, and the 8-class and binary classification task calculated by summing up the corresponding children nodes' probabilities.

to the age, gender, and anatomical site of the lesion. However, their anatomical site information standard is different. We removed this discrepancy by simply mapping them according to Table 1. The metadata was encoded as a feature vector to produce a simple pseudo-QR code. A one-hot encoding was applied for the categorical features, i.e., the anatomical site and gender (male: 01, female: 10, NaN: 00) [23,24] (Table 2), whereas for the age, which is a real/integer feature, a thermometer encoding was conducted [24] (Table 3).

An additional bit in the feature vector was considered to encode the image type; (dermoscopic: 1, photographic: 0). Dermoscopic images have much higher quality, and the knowledge of the image type may help the CNN in the classification

The resulting 14-bit feature vector was constructed as shown in Figure 4a and rearranged

as a 4×4 pseudo-QR-code, in which zeros and ones were shown as white and black pixels, respectively (Figure 4b). The 15th and 16th pixels were blank and set to white for all images. Note that for images with no metadata, such as those from the Fitzpatrick17k dataset, all 4×4 pixels were set to white. The top-left corner of the lesion image was replaced with the pseudo-QR-code, as displayed in Figure 4c. The metadata-embedded images were then used as inputs to the CNN.

Results

Table 4 shows the classification accuracies of four approaches on the binary and 8-class tasks, including malignant melanoma, benign melanocytic nevus, malignant and pre-malignant KC, benign keratosis, malignant cutaneous lymphoma, benign eczema and dermatitis, malignant dermal lesions, and benign dermal

Table 1: Mapping anatomical site annotations of different datasets to the final site annotations. The hyphen indicates no images related to that group.

ISIC 2019 annotation	ISIC 2020 annotation	PAD-UFES-20 annotation	Final site annotation
Head/neck	Head/neck	Face-Ear-Nose-Lip-Neck-Scalp	Head/neck
Upper extremity	Upper extremity	Arm-Forearm-Hand	Upper extremity
Lower extremity	Lower extremity	Thigh-Foot	Lower extremity
Oral/genital	Oral/genital	-	Oral/genital
Palms/soles	Palms/soles	Hand-Foot	Palms/soles
Anterior torso	Torso	Chest	Torso
Posterior torso	Torso	Back	Torso
Lateral torso	Torso	-	
*NaN	NaN	-	NaN

*NaN indicates missing anatomical site information of an image, ISIC: International Skin Imaging Collaboration

Table 2: One-hot encoding applied for the anatomical site annotation

Target site annotation	One hot encoding
Head/neck	100000
Upper extremity	010000
Lower extremity	001000
Oral/genital	000100
Palms/soles	000010
Torso	000001
*NaN	000000

*NaN indicates missing anatomical site information of an image

Table 3: Thermometer encoding applied for the age

Age spans	Thermometer encoding
0-19	00001
20-39	00010
40-59	00100
60-78	01000
≥79	10000
*NaN	00000

*NaN indicates missing anatomical site information of an image

lesions. The binary classes include malignant and benign. For the 8 class task, the confusion matrices for the worst- and best-performing approaches, i.e., direct and metadata-PB, were plotted for comparison in Figure 5, showing that the metadata-PB approach was confused less between malignant melanoma vs. benign melanocytic nevus (classes 0 and 1). Metadata-PB performed even better as it did not misclassify melanocytic and keratinocytic lesions as cutaneous lymphoma, eczema, and dermatitis, or Kaposi sarcoma. The performance of detecting malignant cutaneous lymphoma, Kaposi sarcoma, eczema, and dermatitis did not considerably change in metadata-PB. However, they were better differentiated from melanocytic and kertinocytic lesions. Column 7 of the confusion matrices in Figure 5 shows that other lesion types were confused with benign dermal lesions, and row 6 shows that malignant dermal lesions were mislabeled. Together, they indicate the difficulty of discernment between dermal lesions versus other classes.

Despite the success of the proposed approach, it did not considerably enhance the performance of detecting cutaneous lymphoma, Kaposi sarcoma, eczema, and dermatitis,
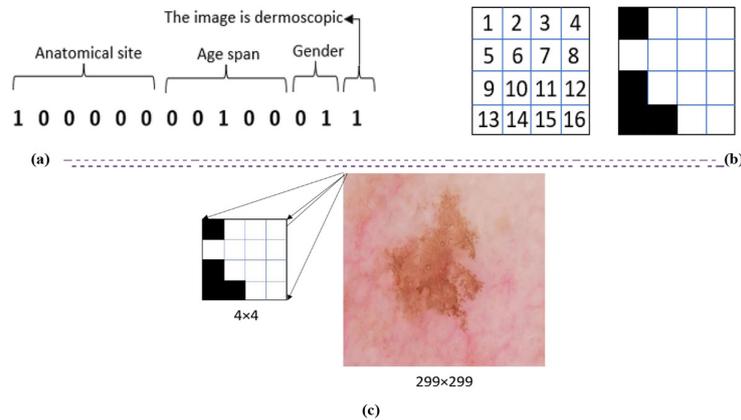
Figure 4: (a) The Feature vector, (b) the corresponding 4x4 pseudo-QR-code representation, and (c) the metadata embedded lesion image for a dermoscopic image from a lesion on the neck of a 45 years old male. The pseudo QR code replaced the top-left corner pixels of the lesion image.

Table 4: The classification accuracies (mean±standard deviation) using 10-fold cross-validation.

Method	Binary accuracy (%)	8-class accuracy (%)
direct	75.3±0.7	62.9±1.5
PB	76.6±0.9	66.1±1.6
metadata-direct	80.9±0.8	68.1±1.9
metadata-PB	84.8±0.9	71.5±1.8

PB: Probability-based

which only contained photographic images from the Fitzpatrick17k dataset. Therefore, the ISIC dataset, including 57536 dermoscopic in 9 skin disease conditions (melanoma, BCC, SCC, actinic keratosis, nevus, solar lentigo, seborrheic keratosis, lichenoid keratosis, and lentigo NOS) were used to evaluate the effectiveness of the proposed method on dermoscopic-only images. A binary (malignant, benign) and a 4-class task (malignant melanoma, benign melanocytic lesions, malignant and pre-malignant keratinocyte carcinoma, and benign keratosis) were considered to assess the proposed method. The classification accuracies (%) for dermoscopic only images for the metadata-PB method were 94.5±0.9 and 89.3±1.1, for binary and 4-class tasks, respectively, demonstrating improvement over the findings in Table 4, due to the higher quality of the images, fewer classes, and a more balanced number of images across different skin conditions.

Table 5 also presents the results of recent studies in comparison to those of our proposed method, demonstrating that the proposed method outperformed previous approaches; however, a direct comparison is not possible due to the difference in datasets and the number of classes.

Discussion

This work introduced a novel approach by embedding the patient's metadata in the lesion images to improve classification accuracy with deep CNNs. The proposed method enhanced the accuracy substantially (at least 5% in all cases, $P<0.001$), highlighting the potential of this approach. These findings also show that the metadata contains valuable information useful for more efficient skin lesion classification.

The direct and PB approaches are similar to those applied in a study by Esteva et al. [7] using a private dataset of 129,450 images with 757 different skin conditions. Despite different
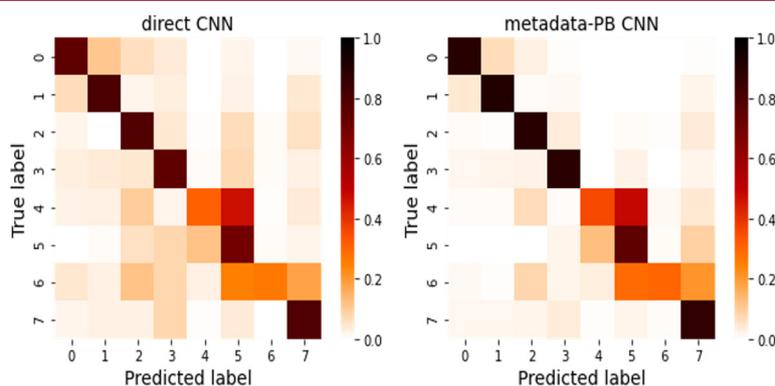
Figure 5: The confusion matrices for the direct and metadata-PB (Probability-based) approaches for the 8-class task.

Table 5: The results of recent studies in comparison to those of our proposed method

Study	Datasets	Number of classes	Best Classification Accuracy
[7]	129,450 clinical images from a combination of open-access dermatology repositories, the ISIC Dermoscopic Archive, the Edinburgh Dermofit Library, and data from the Stanford Hospital.	9 classes & 3 classes	55.4±1.9% & 72.1±0.9%
[25]	3,753 RGB images of skin cancers collected from the Internet	4 groups	94.2%
[26]	11,444 dermoscopic images from the ISIC Dermoscopic Archive	5 classes	82.95%
[27]	5,846 clinical images of pigmented skin lesions	6 classes & Binary classes	86.2% & 91.5%
[28]	The 2019 International Skin Imaging Collaboration Grand Challenge (HAM10000 & BCN20000 datasets)	8 classes	58.5% on the BCN20000 82% on the HAM10000
[29]	HAM10000 dataset	7 classes	87.91%
The proposed method	57,536 dermoscopic images from the ISIC Dermoscopic Archive	4 classes & Binary classes	89.3±1.1% & 94.5±0.9%

ISIC: International Skin Imaging Collaboration, RGB: Red-Green-Blue, HAM: Human Against Machine, BCN20000: Hospital Clinic Barcelona

datasets and the number of skin conditions, the results of the current study were compared to those of their work [7]. In [7], three-partition and nine-partition labels were used, in which the three-partition label divided the skin conditions into malignant, benign, and non-neoplastic lesions. They [7] found that non-neoplastic lesions were better diagnosed by dermatologists rather than CNN's. We included only two critical groups of non-neoplastic inflammatory lesions, namely eczema and dermatitis in our dataset because these two conditions are more challenging for dermatologists to detect due to their similarity to malignant cutaneous lymphoma [12]. In [7], the three-way classification achieved an accuracy of 69.4±0.8% and 72.1±0.9% using the direct and PB methods, respectively, whereas the nine-way classification attained an accuracy of 48.9±1.9% and 55.4±1.7% using the direct and PB methods, respectively. A comparison of the above results with those of our work indicates that the

proposed approach outperformed the method in [7], possibly due to the fewer classes in our work. Moreover, in comparison with [7], the fewer total skin conditions in this study are responsible for the reduced difference in performance between the direct and PB methods.

Conclusion

This study introduced a novel method for skin lesion classification in metadata- embedded images using a deep CNN (Inception-ResNet-v2). The results indicate that the proposed method improved the skin lesion classification performance by at least 5% by including the patient's metadata (i.e. the anatomical site of the lesion, age, and gender) as the model's input data. The proposed method achieved 89.3% in the classification of 4 major skin conditions and 94.5% in distinguishing between malignant and benign lesions. Future work in this field should focus on developing larger public datasets, which provide metadata to facilitate the research community's effort and to enhance the deep learning-based automated classification of skin lesions. In addition, future studies are warranted to investigate the impact of augmentation algorithms such as cutout regularization on CNN performance.

Authors' Contribution

R. Ahmadi Mehr analyses and interpretation of the results, drafting the manuscript. A. Ameri study conception and design, project supervision, and reviewing the manuscript. All authors reviewed and approved the final manuscript.

Ethical Approval

Ethics approval of the datasets has been obtained from the Medical University of Vienna (Protocol-No. 1804/2017), the University of Queensland (Protocol-No. 2017001223), Memorial Sloan Kettering Cancer Center (approval number 16–974), Hospital Clínic Barcelona (approval number HCP/2019/0413), Human Research Ethics Committees of Metro South Health (HREC/17/QPAH/816), the University of Queensland (2018000074), Medical University of Vienna (approval number 1804/2017), Melanoma Institute Australia and the Sydney Melanoma Diagnosis Centre (approval number X20-0241 & 2020/ETH01411), Federal University of Espírito Santo (nº 500002/478), and the Brazilian government through Plataforma Brasil (nº 4.007.097).

Informed consent

Data used in this work have been collected in previous studies in line with the principles of the Declaration of Helsinki. Approvals were granted by local Ethics Committees.

Conflict of Interest

None

References

1. American Cancer Society. Cancer Facts & Figures. 2021. [accessed 2021 September 3]. Available from: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2021.html.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021;**71**(3):209-49. doi: 10.3322/caac.21660. PubMed PMID: 33538338.
3. Stern RS. Prevalence of a history of skin cancer in 2007: results of an incidence-based model. *Arch Dermatol.* 2010;**146**(3):279-82. doi: 10.1001/archdermatol.2010.4. PubMed PMID: 20231498.
4. Miller AJ, Mihm MC Jr. Melanoma. *N Engl J Med.* 2006;**355**(1):51-65. doi: 10.1056/NEJMra052166. PubMed PMID: 16822996.
5. Madan V, Lear JT, Szeimies RM. Non-melanoma skin cancer. *Lancet.* 2010;**375**(9715):673-85. doi: 10.1016/S0140-6736(09)61196-X. PubMed PMID: 20171403.
6. Rudolph C, Schnoor M, Eisemann N, Katalinic A. Incidence trends of nonmelanoma skin cancer in Germany from 1998 to 2010. *J Dtsch Dermatol Ges.* 2015;**13**(8):788-97. doi: 10.1111/ddg.12690. PubMed PMID: 26213814.
7. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;**542**(7639):115-8. doi: 10.1038/nature21056. PubMed PMID: 28117445. PubMed PMCID: PMC8382232.
8. Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. *Lancet Oncol.* 2002;**3**(3):159-65. doi: 10.1016/s1470-2045(02)00679-4. PubMed PMID: 11902502.

9. Codella NC, Nguyen QB, Pankanti S, Gutman DA, Helba B, Halpern AC, Smith JR. Deep learning ensembles for melanoma recognition in dermoscopy images. *IBM Journal of Research and Development*. 2017;**61**(4/5):5-1. doi: 10.1147/JRD.2017.2708299.

10. Ramlakhan K, Shang Y. A mobile automated skin lesion classification system. In 23rd International Conference on Tools with Artificial Intelligence; Boca Raton, FL, USA: IEEE; 2011. p. 138-41.

11. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems 25. NeurIPS Proceedings; 2012.

12. Silvestre Salvador JF, Romero-Pérez D, Encabo-Durán B. Atopic Dermatitis in Adults: A Diagnostic Challenge. *J Investig Allergol Clin Immunol*. 2017;**27**(2):78-88. doi: 10.18176/jiaci.0138. PubMed PMID: 28071589.

13. The Skin Cancer Foundation. Skin Cancer Facts & Statistics. 2021. [accessed 2021 September 4]. Available from: https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/.

14. Combalia M, Codella NC, Rotemberg V, Helba B, Vilaplana V, Reiter O, et al. Bcn20000: Dermoscopic lesions in the wild. *ArXiv*. 2019. doi: 10.48550/arXiv:1908.02288.

15. Codella NCF, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, 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). International Symposium on Biomedical Imaging; Washington, DC, USA: IEEE; 2018.

16. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data*. 2018;**5**:180161. doi: 10.1038/sdata.2018.161. PubMed PMID: 30106392. PubMed PMCID: PMC6091241.

17. International Skin Imaging Collaboration. The ISIC 2020 Challenge Dataset. ISIC; 2020. Available from: https://challenge2020.isic-archive.com/.

18. Rotemberg V, Kurtansky N, Betz-Stablein B, Caffery L, Chousakos E, Codella N, et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Sci Data*. 2021;**8**(1):34. doi: 10.1038/s41597-021-00815-z. PubMed PMID: 33510154. PubMed PMCID: PMC7843971.

19. Pacheco AGC, Lima GR, Salomão AS, Krohling B, Biral IP, De Angelo GG, et al. PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones. *Data Brief*. 2020;**32**:106221. doi: 10.1016/j.dib.2020.106221. PubMed PMID: 32939378. PubMed PMCID: PMC7479321.

20. Pacheco AGC, Krohling RA. The impact of patient clinical information on automated skin cancer detection. *Comput Biol Med*. 2020;**116**:103545. doi: 10.1016/j.compbiomed.2019.103545. PubMed PMID: 31760271.

21. Groh M, Harris C, Soenksen L, Lau F, Han R, Kim A, et al. Evaluating Deep Neural Networks Trained on Clinical Images in Dermatology with the Fitzpatrick 17k Dataset. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); IEEE; 2021. p. 1820-8.

22. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In Thirty-first AAAI conference on artificial intelligence; California: AAAI Press; 2017.

23. Lantz B. Machine learning with R: learn how to use R to apply powerful machine learning methods and gain an insight into real-world applications. UK: Packt Publishing Ltd; 2013.

24. Buckman J, Roy A, Raffel C, Goodfellow I. Thermometer encoding: One hot way to resist adversarial examples. International Conference on Learning Representations; 2018.

25. Dorj UO, Lee KK, Choi JY, Lee M. The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications*. 2018;**77**:9909–24. doi: 10.1007/s11042-018-5714-1.

26. Hekler A, Utikal JS, Enk AH, Hauschild A, Weichenthal M, Maron RC, et al. Superior skin cancer classification by the combination of human and artificial intelligence. *Eur J Cancer*. 2019;**120**:114-21. doi: 10.1016/j.ejca.2019.07.019. PubMed PMID: 31518967.

27. Jinnai S, Yamazaki N, Hirano Y, Sugawara Y, Ohe Y, Hamamoto R. The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning. *Biomolecules*. 2020;**10**(8):1123. doi: 10.3390/biom10081123. PubMed PMID: 32751349. PubMed PMCID: PMC7465007.

28. Combalia M, Codella N, Rotemberg V, Carrera C, Dusza S, Gutman D, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. *Lancet Digit Health*. 2022;**4**(5):e330-9. doi: 10.1016/S2589-7500(22)00021-8. PubMed PMID: 35461690. PubMed PMCID: PMC9295694.

29. Ali K, Shaikh ZA, Khan AA, Laghari AA. Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer. *Neuroscience Informatics*. 2022;**2**(4):100034. doi: 10.1016/j.neuri.2021.100034.