# Skin Lesions Classification: A Radiomics Approach with Deep CNN

Gabriele Piantadosi[1(✉)], Giampaolo Bovenzi[1], Giuseppe Argenziano[2], Elvira Moscarella[2], Domenico Parmeggiani[3], Ludovico Docimo[3], and Carlo Sansone[1]

[1] Dip. di Ingegneria Elettrica e delle Tecnologie dell'Informazione, University of Naples Federico II, Naples, Italy
{gabriele.piantadosi,giampaolo.bovenzi,carlo.sansone}@unina.it
[2] Dip. di Salute Mentale e Fisica e Medicina Preventiva, University of Campania Luigi Vanvitelli, Caserta, Italy
{giuseppe.argenziano,elvira.moscarella}@unicampania.it
[3] Dip. di Scienze Mediche e Chirurgiche Avanzate, University of Campania Luigi Vanvitelli, Caserta, Italy
{domenico.parmeggiani,ludovico.docimo}@unicampania.it

**Abstract.** Supporting the early diagnosis of skin cancer is crucial for the sake of any kind of treatment or surgery. This work proposes to improve the outcome of automatic diagnoses approaches by using an ensemble of pre-trained deep convolutional neural networks and a suitable voting strategy. Moreover, a novel patching approach has been deployed. The proposal has been fairly evaluated with the literature proposals demonstrating good preliminary results.

## 1 Introduction

The skin cancers can be roughly divided into two categories: melanoma and non-melanoma. The most common skin cancer belongs to the category of non-melanoma tumours and is the basal cell carcinoma and squamous cell carcinoma representing the 5th most commonly occurring cancer in men and women, with over 1 million diagnoses worldwide in 2018, although this is likely to be an underestimate [2]. Skin cancer diagnosis is a non-trivial challenge for several reasons: this disease is usually underestimated. Due to these factors, it is likely that the reported global incidence of skin cancer is an underestimate to the extent that some people with cancer do not consult a physician; moreover there are different sub-types of skin cancer, leading to issues in data collection; for example, non-melanoma skin cancer may be not completely tracked by cancer registries.

According to the American Cancer Society, each year, more than 60000 people in the United States were diagnosed with melanoma and more than 8000 Americans died of disease [2]. Melanoma is the most aggressive form of skin cancer and the surgical treatments are the main approach. However, patients


M. Cristani et al. (Eds.): ICIAP 2019 Workshops, LNCS 11808, pp. 252–259, 2019.
https://doi.org/10.1007/978-3-030-30754-7_26

who undergo late treatment may develop metastasise to regional lymph nodes or distant metastases. The median survival date in that conditions is about 6–9 months, and the five-year survival rate is less than 5%. As a consequence, it is important to rely on more effective strategies to decrease mortality rates from melanoma such as early detection. In recent years, Deep learning (DL) approaches, and in particular deep Convolutional Neural Networks (CNNs), obtain wide popularity in many pattern recognition tasks thanks to their ability to autonomous learn compact hierarchical features that best fit the specific input domain. The aim of this work is to support the skin lesion diagnosis by means of deep learning approaches with a suitable ensemble of deep CNN.

The paper is organised as follows: Sect. 2 outlines some traditional approaches and related works; Sect. 3 describes the proposal, introduces the dataset and illustrates the evaluation approach; Sect. 4 reports our results compared with those obtained by using some literature proposals; finally, Sect. 5 discusses the obtained results and provides some conclusions.

## 2 Related Works

Nowadays, detection of melanoma is strongly improved by means of novel approaches to support the visual inspection such as total body photography, dermoscopy, automated diagnostic systems and reflectance confocal microscopy. Dermoscopy is the most spread method for skin cancer detection. It consists of a non-invasive, *in-vivo* technique and it is performed by means of a manual instrument called dermatoscope. The procedure allows for the visualization of subsurface skin structures in the epidermis, at the dermoepidermal junction, and in the upper dermis; these structures are usually not visible to the naked eye [1,9,10]. An interesting aspect of such a technique is the possibility of the images to be digitised for storage, far transmission or sequential analysis. Thanks to this feature dermatologists have begun to incorporate novel imaging techniques into diagnostic algorithms.

The automatic skin cancer diagnosis techniques are traditionally based on computer vision approaches that implement one or more of the following approaches:

**Geometry analysis:** like all the tumour forms, the growth of the lesion has an irregular pattern which, in the case of skin tumour forms, leads to asymmetry in the shape. It is important, therefore, to rely on powerful descriptors such as area, perimeter, convexity, the major/minor-axis' length and angle, compactness, elongation, eccentricity (also known as ellipticity), roundness (or circularity) and sphericity [1,6].

**Color analysis:** the tonality of the skin when a tumour grows may change in a discriminant way. Therefore, automatic approaches may benefit from colour analysis. This descriptors are mainly histogram derived considerations and require the choice of a colour space such as RGB, Hue-Saturation-Value (HSV), YCrCb, and the novel Hue-Min-Max-Difference (HMMD). Among the common colour descriptors there are the Scalable Color Descriptor (SCD),

defined by a fixed colour space quantization, and the Haar transform encoding; the dominant colour (DC) quantify the distribution of the salient colours in the image; the Color Layout Descriptor (CLD) captures the spatial layout of the dominant colours; and so on [8,12,13].

**Texture analysis:** texture, like colour, is a powerful low-level descriptor in skin tumour classification. Computer Vision literature is plenty of textural descriptors that can be roughly divided into local or global approaches operating in space or frequency domain [8,12].

It is, therefore, necessary to rely on robust segmentation and feature extraction approaches that must be developed by domain experts.

While newer hand-crafted features for skin cancer detection are continuously proposed by domain experts, in the last years deep learning approaches have gained popularity in many pattern recognition and computer vision tasks, being able to outperform classical machine learning approaches in different fields [5]. Among these 'Deep Approaches', the most used are the Deep CNN (better discussed in the Sect. 3), composed of different convolutional layers stacked in a deep architecture meant to automatically learn the best data representation as composed by simpler concepts. They usually perform better than classifiers trained on hand-crafted features because are able to learn a compact and hierarchical representation of an image, that well fits the specific task to solve.

Among the deep approaches for skin cancer diagnosis Codella et al. [3] propose to feed a pre-trained network called AlexNet [7]. The pre-trained network was used as a feature-extractor (better discussed in Sect. 3) to map the images in a bigger feature-space (4096 features). The vectors in the new space were used to train a Support-Vector Machine (SVM) model.

Pomponiu et al. [11] propose a similar approach by applying data augmentation to feed a bigger number of images. This feeding approach consists of applying affine transformations to the input images during the training phase with the aim of improving generalization behaviour. The authors also change the classification model with a weaker k-Nearest Neighbors (k-NN) model.

A novel deep approach proposed by Haenssle et al. [4] consists of using a more complex pre-trained network called GoogLeNet Inception-v3 [7]. The difference consists in using the pre-trained network in a fine-tuning flavour (transfer learning approaches are better discussed in Sect. 3) to adapt the whole network for the skin cancer detection task.

## 3 Materials and Methods

In this work we propose to address the skin lesion classification by means of Convolutional Neural Networks (CNNs) using transfer learning as a training strategy.

Convolutional Neural Networks (CNNs) are machine learning models borrowed from traditional Neural Networks (NNs). Such architectures share most of the features: they are both made of neurons, usually organized in layers; layers are stratified to create a feed-forward network in which the output of a

layer is the input of next layer. However, while traditional NNs operate on the features designed and extracted by a domain expert, CNNs use a hierarchy of convolution operations to autonomously design the features that better model the problem under analysis. Although this characteristic gives to CNNs a great representational ability, it also comes with a huge number of parameters to learn and, despite efforts made to design increasingly compact networks, CNNs tend to strongly focus on the input dataset (also know as overfit behaviour) when the training phase is limited to few training instances. The solution would be to provide a large number of input instances but, when these are not easily available, a valid solution is to transfer the knowledge from a previous trained CNN (also on a different domain) to the new input domain by training it to solve the new task with a proper amount of available training data. This approach is also known as *Transfer Learning* and can be obtained by following two approaches:

**Fine-Tuning:** consists in train only the last few layers of a pre-trained CNN (usually called fully connected layers). This fine-tuning of part of the weights not only address the novel input problem but also adapt the last layer of the network to the different output classes.

**Feature-Extraction:** uses the pre-trained CNN as is by relying on the output of the second/third layer; no further training is required to obtain a feature vector to be used as input of different models, such as traditional ML approaches, or a mapping function in a novel and bigger vectorial space. This approach is strongly promoted because the knowledge of the pre-trained network is able to efficiently map the instances for several tasks (also for very different domains).

Both the approaches aim to reduce the computational burden of to train a network from scratch but only the Feature-Extraction approach also performs a space transformation that may help to index a bunch of images efficiently and relying on visual peculiarity not easy to catch up with handcrafted features.

### 3.1 Proposed Approach

The proposed architecture is depicted in the Fig. 1.

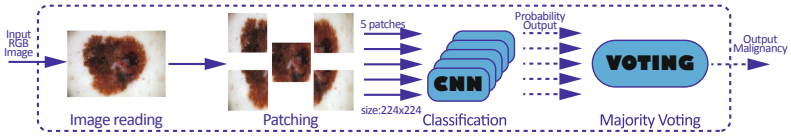


**Fig. 1.** Proposed approach.

The approach consists of four main stages organised as follows:

**Image Reading:** Since there are several tools and cameras for dermoscopy that can provide digital images suitable for computer vision aided analysis, it is required a suitable read, conversion and representation format able to be compliant with the subsequent steps. The formats accepted are JPEG, BMP, PNG, DICOM, NIFTI and the image is, therefore, converted into RGB matrices.

**Patching:** The whole image is divided into five patches extracted as big as half the dimensions. The position of each patch is: four from the corners and one in the centre. Each patch is, then, resized to the squared size of $224 \times 224$ to be compliant with the input of the subsequent stage (the input of the neural network).

**Classification:** Each patch undergoes a neural network to create an ensemble classification of the same lesion with five points of view. The result is obtained with a pre-trained Alexnet [7] fine-tuned on our dataset. The output of this stage is the probability of each patch to belong to malignant melanoma (a five element vector).

**Voting:** The five element vector is, therefore, combined to produce an ensemble decision by applying a weighted majority voting. The most probable class (malignant/non-malignant) is returned.

### 3.2 Dataset

The dataset was collected with the "Synergy-net: Research and Digital Solutions against Cancer" project (funded in the framework of the POR Campania FESR 2014–2020) and labelled by experienced dermatologists. The size consists of 200 patients (77 female and 123 male subjects) with several images per patient as described in Table 1. The table also shows the four different skin lesions included in the dataset and the relative class considered for the final aim of this article. Lesions belong to patients with age from 1 to 100 years old and are from different parts of the body.

**Table 1.** Dataset composition.

| Lesion | Patients | Male | Female | Images | Class |
|---|---|---|---|---|---|
| Nevus (Mole) | 23 | 13 | 10 | 24 | (B) Benignant |
| Angioma | 27 | 14 | 13 | 27 | (B) Benignant |
| Seborrheic Keratosis | 50 | 32 | 18 | 51 | (B) Benignant |
| Melanoma | 100 | 64 | 36 | 107 | (M) Malignant |
| **Total** | **200** | **123** | **77** | **209** | $\frac{B}{M} = \frac{102}{107}$ |

The colour images composing the dataset have different resolutions ranging from a minimum of $1999 \times 1333$ pixel to a maximum of $5184 \times 3456$ and three RGB channels.

### 3.3 Experimental Setup

The proposed CNNs have been evaluated using the high-level neural networks API Keras (Python 3.6) with TensorFlow 1.6 as back-end. Python scripts have been executed on a physical server hosted in our university HPC center[1] equipped with 2 x Intel(R) Xeon(R) Intel(R) 2.13 GHz CPUs (4 cores), 32 GB RAM and a Nvidia Titan Xp GPU (Pascal family) with 12 GB GRAM.

All the competitors have been developed by strictly implementing the design published in their works and evaluated on the same set of data and configuration of cross-validation (fixing the seeds of the pseudo-random generators). To face the size limitation of the training dataset a data augmentation approach has been applied. Spatial affine transformations such as Left/Right flip, Top/Bottom flip, 90° rotations have been applied.

Finally, in order to train the proposed models, a cross-entropy loss has been minimized and the performance has been evaluated in 10-fold cross-validation considering the median values of the Area Under the Receiver Operating Characteristic (ROC) Curve or AUC-ROC.

## 4 Results

This section reports the results of the described approach compared with some literature proposals. Table 2 reports our approach by varying the training stage (with or without the augmentation). The subpatching strategy has been also evaluated by considering the baseline approach (whole image).

**Table 2.** Proposed approach variants comparison. AUC median values (obtained with a 10-folds CV).

| Augmentation | SubPatch | Voting | AUC-ROC |
|---|---|---|---|
| Yes | | | 84.19% |
| | Yes | OR | 55.28% |
| | Yes | Majority | 60.91% |
| Yes | Yes | OR | 83.63% |
| Yes | Yes | Majority | **87.27%** |

To evaluate the effectiveness of the proposed approach, our results were compared with those obtained by applying the algorithms described in Sect. 2, by using a 10-fold Cross Validation (CV) ensuring that, to obtain a reliable and fair evaluation, the lesion from the same subject are always separated across the CV folds. Since Table 2 clearly shows that a form of augmentation improves the model to generalise the problem and thus improves the final result, we evaluated competitors that did not already apply this approach using our own augmentation technique. Results are reported in Table 3.

[1] http://www.scope.unina.it.

**Table 3.** Our approach compared with the literature proposals. AUC median values (obtained with a 10-folds CV).

| Approach | Methodology | AUC-ROC |
|---|---|---|
| Haenssle et al. [4] | Inception-v3 (FineTuning) | 55.68% |
| Pomponiu et al. [11] | Alexnet (FeatureExtraction) + kNN | 76.41% |
| Codella et al. [3] | Alexnet (FeatureExtraction) + SVM | 78.06% |
| Haenssle et al. [4] + Augmentation | Inception-v3 (FineTuning) | 80.91% |
| Codella et al. [11] + Augmentation | Alexnet (FeatureExtraction) + SVM | 84.54% |
| Our approach | Ensamble of Alexnet (FineTuning) | **87.27**% |

## 5 Discussions and Conclusions

The aim of this work was to apply deep convolutional neural networks to support the early diagnosis of skin cancer. Table 2 demonstrates how our novel patching approach and the suitable ensemble of pre-trained networks is able to produce a reliable diagnosis.

Table 3 compares the best CNN-based approach with other approaches presented so far in the literature and described in Sect. 2, showing that, even if we try to train the competitors in the best condition, our proposal achieves the best classification results in terms of AUC. Table 2 also makes it clear that our dataset, due to the size limitation, requires data augmentation approach to successful train the network. For this reason and for the sake of the fairness we also train the competitors with data augmentation (middle part of Table 3).

As part of the "Synergy-net: Research and Digital Solutions against Cancer" project, the data acquisition is currently going on providing novel data with the aim of retraining the model and further validating the insights. Moreover, we are currently investigating the data fusion of geometrical features, textural features and deep features to provide a more robust ML/DL approach in skin lesion diagnosis.

**Acknowledgments.** The authors gratefully acknowledge the support of the Calculation Centre SCoPE of the University of Naples Federico II and his staff. This work is part of the "Synergy-net: Research and Digital Solutions against Cancer" project (funded in the framework of the POR Campania FESR 2014–2020 - CUP B61C17000090007).

## References

1. Argenziano, G., et al.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet. J. Am. Acad. Dermatol. **48**(5), 679–693 (2003)
2. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A.: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. **68**(6), 394–424 (2018)

3. Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., Smith, J.R.: Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images. In: Zhou, L., Wang, L., Wang, Q., Shi, Y. (eds.) MLMI 2015. LNCS, vol. 9352, pp. 118–126. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24888-2_15
4. Haenssle, H.A., et al.: Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann. Oncol. **29**(8), 1836–1842 (2018)
5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
6. Henning, J.S., et al.: The cash (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. J. Am. Acad. Dermatol. **56**(1), 45–52 (2007)
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
8. Manjunath, B.S., Ohm, J.R., Vasudevan, V.V., Yamada, A.: Color and texture descriptors. IEEE Trans. Circuits Syst. Video Technol. **11**(6), 703–715 (2001)
9. Marghoob, A.A., et al.: Instruments and new technologies for the in vivo diagnosis of melanoma. J. Am. Acad. Dermatol. **49**(5), 777–797 (2003)
10. Menzies, S., Ingvar, C., McCarthy, W.: A sensitivity and specificity analysis of the surface microscopy features of invasive melanoma. Melanoma Res. **6**(1), 55–62 (1996)
11. Pomponiu, V., Nejati, H., Cheung, N.M.: Deepmole: deep neural networks for skin mole lesion classification. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 2623–2627. IEEE (2016)
12. Yu, H., Li, M., Zhang, H.J., Feng, J.: Color texture moments for content-based image retrieval. In: Proceedings of International Conference on Image Processing, vol. 3, pp. 929–932. IEEE (2002)
13. Zalaudek, I., et al.: Three-point checklist of dermoscopy: an open internet study. Br. J. Dermatol. **154**(3), 431–437 (2006)