Original paper

# Convolutional neural networks for the detection of malignant melanoma in dermoscopy images

**Dominika Kwiatkowska[1], Piotr Kluska[2], Adam Reich[1]**

[1]Department of Dermatology, Institute of Medical Sciences, Medical College of Rzeszow University, Rzeszow, Poland
[2]Wroclaw University of Science and Technology, Wrocław, Poland

Adv Dermatol Allergol 2021; XXXVIII (3): 412–420
DOI: https://doi.org/10.5114/ada.2021.107927

## Abstract

**Introduction:** Convolutional neural networks gained popularity due to their ability to detect and classify objects in images and videos. It gives also an opportunity to use them for medical tasks in such specialties like dermatology, radiology or ophthalmology. The aim of this study was to investigate the ability of convolutional neural networks to classify malignant melanoma in dermoscopy images.

**Aim:** To examine the usefulness of deep learning models in malignant melanoma detection based on dermoscopy images.

**Material and methods:** Four convolutional neural networks were trained on open source dataset containing dermoscopy images of seven types of skin lesions. To evaluate the performance of artificial neural networks, the precision, sensitivity, F1 score, specificity and area under the receiver operating curve were calculated. In addition, an ensemble of all neural networks' ability of proper malignant melanoma classification was compared with the results achieved by every single network.

**Results:** The best convolutional neural network achieved on average 0.88 precision, 0.83 sensitivity, 0.85 F1 score and 0.99 specificity in the classification of all skin lesion types.

**Conclusions:** Artificial neural networks might be helpful in malignant melanoma detection in dermoscopy images.

**Key words:** dermoscopy, melanoma, neural networks, deep learning.

## Introduction

Early detection of malignant melanoma – one of the most aggressive skin tumours – plays a key role in reducing mortality due to this neoplasm. The increase in melanoma morbidity is observed all over the world, especially in the Caucasian population [1, 2]. One of the most important prognostic factors in patients with malignant melanomas is the tumour thickness, histologically assessed according to the Breslow scoring system. The 5-year survival rate for patients with stage IA (with tumours less than 1 mm thick) is above 90%, whereas for patients with tumours thicker than 4 mm is far less satisfactory [2, 3]. Even though a dynamic development of a number of novel anti-cancer immunotherapies has been observed in recent years leading to a significant improvement of the survival rate of patients with advanced malignant melanoma, still many patients die due to this malignancy [4, 5]. Thus, detection of the tumour in the primary stage remains crucial for patient prognosis. A large number of early malignant melanomas are asymptomatic, and frequently they do not arouse any suspicion among patients, even though they are usually visible to the naked eye. The early detection of malignant melanoma is often possible with the help of dermoscopy, but the accuracy of this diagnostic technique relies on the physician's skills and experience. Implementation of an automatic lesion classification with deep learning methods might help clinicians in faster malignant melanoma detection and proper differentiation from benign pigmented lesions and other malignancies. Convolutional neural networks have increased performance in computer vision tasks, like classification, object localization and detection, being even able to outperform humans [6, 7]. Here, we have tested various convolutional neural networks as an aid for proper classification of pigmented skin lesions and differentiating malignant melanoma from other skin tumours.

**Address for correspondence:** Prof. Adam Reich, Department of Dermatology, Institute of Medical Sciences, Medical College of Rzeszow University, 2 Szopena St, 35-055 Rzeszow, Poland, phone: +48 605 076 722, e-mail: adi_medicalis@go2.pl
**Received:** 5.08.2019**, accepted:** 13.01.2020.

*This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). License (http://creativecommons.org/licenses/by-nc-sa/4.0/)*

## Aim

The aim of this study was to evaluate the accuracy of deep learning models in malignant melanoma detection based on dermoscopic images.

## Material and methods

### Dataset

The dataset used in this study was extracted from "ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection" grand challenge [8, 9]. Authors of HAM10000 dataset also provided supplementary data about the origin of the lesion with a unique identifier. Based on this unique identifier we have split our dataset, to avoid training-test data leakage. The diagnosis of all skin lesions was confirmed by histological examination with reference to the information provided in the dataset [8, 9]. This dataset had a total of 10015 images assigned to one of the following categories with a given quantity:

1. Melanoma – 1113 images,
2. Melanocytic nevus – 6705 images,
3. Basal cell carcinoma – 514 images,
4. Actinic keratosis/Bowen's disease (intraepithelial carcinoma) – 327 images,
5. Benign keratosis (solar lentigo/seborrheic keratosis/lichen planus-like keratosis) – 1099 images,
6. Dermatofibroma – 115 images,
7. Vascular lesion – 142 images.

For training, monitoring and evaluating models we proportionally split the dataset into following parts:

1. Training dataset – 8123 images,
2. Validation dataset – 886 images,
3. Test dataset – 1006 images.

Distribution of skin lesion types across datasets can be found in Table 1.

### *Convolutional neural network (CNN)*

Deep learning is a class of machine learning algorithms which uses multiple layers to progressively extract higher level features from the raw input, e.g. in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits or letters or faces [10, 11]. Those models can learn – without explicit programming – different features at multiple levels of abstraction directly from data [10]. The CNN is a variant of a deep learning model widely used for image processing in which core operation is performed by a convolutional layer. A convolution is the simple application of a filter to an input (e.g. image) which results in an activation. Repeated application of the same filter to an input results in a map of activations called a feature map, indicating the locations and strength of a detected feature in an input (e.g. an image). The convolutional layer consists of learnable filters which are applied over the input data to extract features [10]. In our experiments, a batch normalization layer was inserted after each convolutional layer, followed by rectified linear unit activation function [11]. Batch normalization is a technique for improving the speed, performance, and stability of artificial neural networks [10, 11].

### Models

We have compared the performance of four different architectures of artificial neural networks. As the base structure, we used ResNet-101 [12] and its variations – ResNeXt [13], SE-ResNet, SE-ResNeXt [14]. ResNet-101 is built with 33 residual blocks as shown in Figure 1 A and, in total, consists of 100 convolution operations. Residual blocks introduce a shortcut connection. The shortcut connection adds identity mapping between the input of the residual block and its output. The architecture of the residual block showed a positive impact on training deeper CNNs.

ResNeXt introduces new hyperparameter – cardinality, which can be achieved by a grouped convolution operation as shown in Figure 1 B. This operation divides the input into 32 groups. This is equivalent to performing 32 smaller convolution operations side by side.

Squeeze and excitation (SE) networks (SE-ResNet, SE-ResNeXt) are built on that idea with an additional operation block (Figure 1 C) to increase sensitivity to descriptive features.

**Table 1.** Distribution of skin lesion types across training, validation and test datasets

| Type of skin lesion | Training dataset | Validation dataset | Test dataset | Total number |
|---|---|---|---|---|
| Malignant melanoma | 896 (11.0%) | 103 (11.6%) | 114 (11.3%) | 1113 (11.1%) |
| Melanocytic nevus | 5446 (67.0%) | 589 (66.5%) | 670 (66.6%) | 6705 (66.9%) |
| Basal cell carcinoma | 416 (5.1%) | 46 (5.2%) | 52 (5.2%) | 514 (5.1%) |
| Actinic keratosis/Bowen's disease (intraepithelial carcinoma) | 264 (3.3%) | 33 (3.7%) | 30 (3.0%) | 327 (3.3%) |
| Benign keratosis (solar lentigo/seborrheic keratosis/lichen planus-like keratosis) | 896 (11.0%) | 92 (10.4%) | 111 (11.0%) | 1099 (11.0%) |
| Dermatofibroma | 94 (1.2%) | 10 (1.1%) | 11 (1.1%) | 115 (1.1%) |
| Vascular lesion | 111 (1.4%) | 13 (1.5%) | 18 (1.8%) | 142 (1.4%) |
| Total number | 8123 (100%) | 886 (100%) | 1006 (100%) | 10015 (100%) |

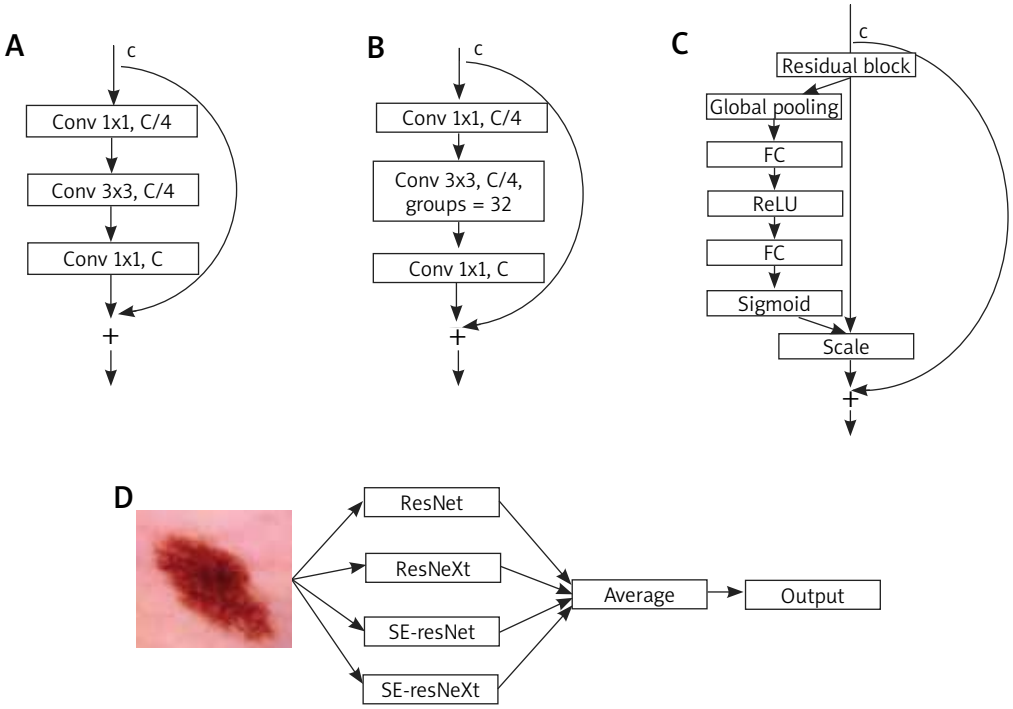



**Figure 1.** Example of a residual block (**A**) of ResNet-101 with input of C channels. The first convolutional layer decreases channels by 4 and the last one restores it to C channels. ResNeXt block (**B**). SE-ResNet block (**C**), where residual block can be a block from (A) or (B). **D** – Scheme of the ensemble of CNNs

### *Ensemble of CNNs*

At the end of our experiments, we have combined all CNN models as a stacked ensemble [15] to test whether the achieved prediction of malignant melanoma could be further improved. Output produced by this method was calculated by averaging all predictions of CNNs as seen in Figure 1 D.

### *Training*

A good practice is to use transfer learning when faced with CNN training on a small dataset. We have used models, which have been already pre-trained on ImageNet dataset [6], which has 1000 object classes. To adapt it for our needs, the last fully connected layer consisting of 1000 output nodes was removed and replaced with a fully connected layer with 7 output neurons. On the product of the last fully connected layer, a softmax activation function was applied. Each of those models was trained on the training dataset with Adam optimizer [16]. To work with the imbalanced dataset, weighted cross-entropy loss function was utilized, in which weights were equal to inverse cardinality in the training dataset. As most of the images in our dataset had lesions located in the centre of the image, during training an input image of 600 × 450 size was augmented by random rotation by 180°, and image centre was cropped to 300 × 400 size in order to cut out black or skin background. Finally, each image was resized to 224 × 224 which is the input size of a CNN. CNNs were trained up to 20 epochs. If the network started to show signs of overfitting on the validation dataset, the training was stopped. Continuation of the training could result in better CNN fitting to the images presented in the training set, but this outcome was not desired, as next the CNN would not generalize well on the test dataset.

### Evaluation

To measure the performance of the algorithm we have computed precision, sensitivity, specificity, F1 score (1) and area under the receiver operating curve (AUC) on the test dataset. F1 score was calculated as a harmonic mean of precision and sensitivity according to the formula:

$$F1 = 2 \times \frac{1}{\frac{1}{Precision} + \frac{1}{Sensitivity}} = \frac{2TP}{2TP + FP + FN}$$

where TP = true positives, FP = false positives, FN = false negatives.

### Visual explanation

In order to explain the classifications of the best CNN model, Grad-CAM was used [17]. This method allowed us to visualize which regions, according to CNN, were the most important for given class prediction. Grad-CAM produces heatmap that can be overlaid over the tested image. The area coloured in red corresponds to the highest score of activation of the CNN, which potentially de-

termines the membership of the appropriate category of skin lesions, whereas the light blue colour indicates the least specific area for a correct prediction to the network. In addition, we present Guided Grad-CAM, which shows in detail the most distinctive features, based on which CNN decided about image classification to a given class.

### Software list

The software with given versions used in the current study is listed below:

- Python 3.6.
- PyTorch 0.4.1.
- NumPy 1.15.
- Pandas 0.23.4.
- Scikit-learn 0.19.2.
- Pillow 5.2.0.

## Results

### Evaluation of CNNs

After training, the results of each CNN was evaluated as described above their average metrics were compared between models.

All of the examined neural networks achieved the best precision, sensitivity and F1 score in the evaluation of melanocytic nevi, which can be explained by the fact that they accounted for the vast majority of lesions in the training dataset. Interestingly, the specificity of melanocytic nevus detection achieved the lowest score in all CNN compared to other calculated parameters (Table 2). Regarding malignant melanoma assessment, the ResNeXt appeared to be superior, achieving the best precision, sensitivity and F1 score, followed by ResNet which precision was lower by 0.01 and sensitivity by 0.02 than in ResNeXt. SE-ResNet and SE-ResNeXt scored 0.05 and 0.07 fewer precision points compared to the best model, respectively (Table 2).

Based on average metrics, ResNeXt again turned out to be the best among all the analysed CNNs (Table 2). However, ResNet matched ResNeXt regarding average sensitivity.

Despite the fact that ResNeXt turned out to be the best tested model, we observed misclassifications of malignant melanoma by this model in reference to other CNNs (Table 3). A total percentage of wrongly classified images of malignant melanoma by ResNeXt was 28%, from which half were correctly classified by other neural networks. The remaining ones were improperly diagnosed not only by ResNeXt, but also by other CNNs. The considerable number of false negative results of malignant melanoma image prediction were melanocytic nevi or benign keratoses.

The ensemble of CNNs' precision, F1 score and specificity surpassed the best result of ResNeXt in malignant melanoma prediction. However, sensitivity remained similar with this model as shown in Figure 2 A. The normalized confusion matrix is presented in Figure 2 B.

### Visual explanation

Based on Grad-CAM activation heatmaps we have shown how ResNeXt learned to distinguish skin lesion from unaltered skin (Figure 3). Figures 3 A and B demonstrate properly classified images of melanocytic nevus (Figure 3 A) and malignant melanoma (Figure 3 B). On the other hand, ResNeXt misclassified malignant melanoma with melanocytic nevus as shown in Figure 3 C. Presumably, the spatial size of skin lesion as well as its colour may have an impact on the accurate assignment of the lesion type. Likewise, zoom, lighting, and angle of dermatoscopic images taken might be important factors that contribute to malignant melanoma prediction, as seen in Figure 3 D. Only the first picture of the disease was correctly classified as malignant melanoma.

## Discussion

Crucial factors of the correct malignant melanoma diagnosis include clinical experience and proper training of physicians. As shown by Haenssle *et al.* [18], the CNN might outperform clinicians in the differentiation of malignant skin lesions from benign ones. For the needs of that research, an international group of 58 dermatologists, which included 30 experts in the field, were involved. In the level-I study, physicians had to make a diagnosis only based on dermatoscopic images. The artificial neural network achieved 0.86 area under the receiver operating characteristic curve (ROC AUC) compared to ROC area of 0.79 ($p < 0.01$) for all dermatologists. As in real life clinicians usually have more information about the patient condition, in the level-II study, dermatologists were provided with an extra close-up dermoscopy image and additional clinical information (age, sex and body site). In the second study, the mean ROC area for all dermatologists increased to 0.82, but was still lower than CNN. Based on these experiments, we do believe that even experienced dermatologists may benefit from an aid of CNN, which can give a supportive opinion while diagnosing a suspicious pigmented skin lesion, although CNN cannot replace a well-skilled physician.

Additionally, Yap *et al.* [19] showed that there is a possibility to combine CNN trained on dermatoscopic images with another CNN trained on macroscopic images. This research focused on the classification of five skin lesion types: melanocytic nevus, malignant melanoma, basal cell carcinoma, squamous cell carcinoma, and pigmented benign keratosis. Based only on one image type, artificial neural network prediction had achieved 0.647 ±0.01 and 0.707 ±0.01 accuracy for macroscopic and dermatoscopic images, respectively. However, when they were combined, their accuracy increased to 0.721 ±0.007.

Despite the fact that the accuracy of macroscopic CNN alone is worse than the one trained with dermatoscopic images, their combination had a positive impact on the diagnosis. However, to implement such systems in real life scenarios, it requires additional macroscopic image dataset for CNN.

Access to medical databases is a limiting factor for CNN development, because they require collection of personal data and, as a result, usually they are not freely available for the public. Preparation of such dataset takes time and the proper labelling of skin lesions on the images requires expert knowledge. The HAM10000 dataset [8, 9], which was

**Table 2.** Results of precision, sensitivity, F1 score and specificity for ResNet, ResNeXt, SE-ResNet and SE-ResNeXt in classification of each disease

| Skin lesion | Precision | Sensitivity | F1 score | Specificity | AUC (ROC) |
|---|---|---|---|---|---|
| ResNet: | | | | | |
| Malignant melanoma | 0.76 | 0.70 | 0.73 | 0.97 | 0.96 |
| Melanocytic nevus | 0.95 | 0.96 | 0.95 | 0.89 | 0.98 |
| Basal cell carcinoma | 0.86 | 0.85 | 0.85 | 0.99 | 1.00 |
| Actinic keratosis/Bowen's disease | 0.72 | 0.77 | 0.74 | 0.99 | 0.99 |
| Benign keratosis | 0.79 | 0.80 | 0.80 | 0.97 | 0.97 |
| Dermatofibroma | 0.9 | 0.82 | 0.86 | 1.00 | 0.98 |
| Vascular lesion | 0.94 | 0.89 | 0.91 | 1.00 | 1.00 |
| Average | 0.85 | 0.83 | 0.84 | 0.97 | 0.98 |
| ResNeXt: | | | | | |
| Malignant melanoma | 0.77 | 0.72 | 0.74 | 0.97 | 0.95 |
| Melanocytic nevus | 0.95 | 0.96 | 0.96 | 0.90 | 0.98 |
| Basal cell carcinoma | 0.85 | 0.9 | 0.88 | 0.99 | 0.99 |
| Actinic keratosis/Bowen's disease | 0.84 | 0.70 | 0.76 | 1.00 | 0.99 |
| Benign keratosis | 0.79 | 0.81 | 0.80 | 0.97 | 0.98 |
| Dermatofibroma | 1.00 | 0.73 | 0.84 | 1.00 | 1.00 |
| Vascular lesion | 0.95 | 1.00 | 0.97 | 1.00 | 1.00 |
| Average | 0.88 | 0.83 | 0.85 | 0.99 | 0.99 |
| SE-ResNet: | | | | | |
| Malignant melanoma | 0.72 | 0.69 | 0.71 | 0.97 | 0.96 |
| Melanocytic nevus | 0.94 | 0.94 | 0.94 | 0.88 | 0.97 |
| Basal cell carcinoma | 0.78 | 0.90 | 0.84 | 0.99 | 0.99 |
| Actinic keratosis/Bowen's disease | 0.85 | 0.57 | 0.68 | 1.00 | 0.98 |
| Benign keratosis | 0.75 | 0.78 | 0.77 | 0.97 | 0.97 |
| Dermatofibroma | 0.89 | 0.73 | 0.8 | 1.00 | 0.98 |
| Vascular lesion | 0.94 | 0.89 | 0.91 | 1.00 | 1.00 |
| Average | 0.84 | 0.79 | 0.81 | 0.97 | 0.98 |
| SE-ResNeXt: | | | | | |
| Malignant melanoma | 0.70 | 0.67 | 0.68 | 0.96 | 0.95 |
| Melanocytic nevus | 0.95 | 0.95 | 0.95 | 0.90 | 0.97 |
| Basal cell carcinoma | 0.82 | 0.87 | 0.84 | 0.99 | 0.99 |
| Actinic keratosis/Bowen's disease | 0.73 | 0.53 | 0.62 | 0.99 | 0.98 |
| Benign keratosis | 0.74 | 0.83 | 0.78 | 0.96 | 0.97 |
| Dermatofibroma | 0.89 | 0.73 | 0.80 | 1.00 | 0.99 |
| Vascular lesion | 0.89 | 0.89 | 0.89 | 1.00 | 1.00 |
| Average | 0.82 | 0.78 | 0.79 | 0.97 | 0.98 |

**Table 3.** Differences in the classification of various images by analysed CNNs. Correctly classified malignant melanomas on images were marked in green, whereas incorrect predictions were highlighted in red

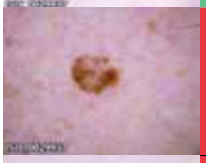
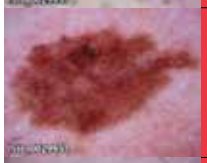
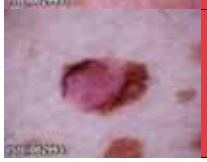
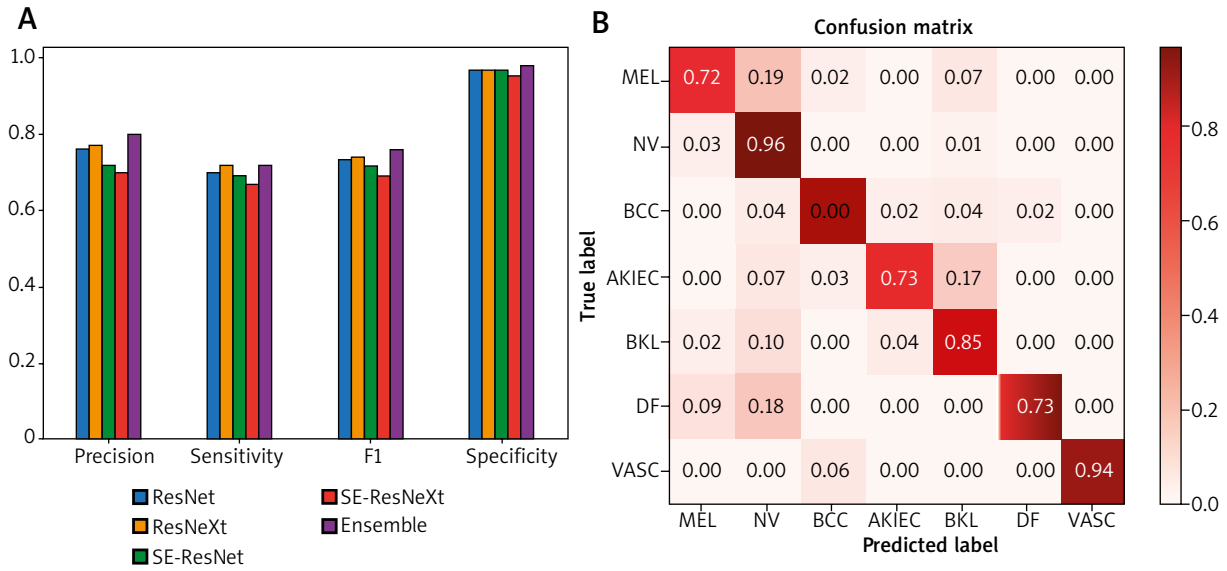
| Image | ResNet | ResNeXt | SE-ResNet | SE-ResNeXt |
|---|---|---|---|---|
| | Malignant melanoma | Melanocytic nevus | Malignant melanoma | Malignant melanoma |
| | Malignant melanoma | Melanocytic nevus | Melanocytic nevus | Melanocytic nevus |
| | Malignant melanoma | Melanocytic nevus | Melanocytic nevus | Malignant melanoma |
| | Malignant melanoma | Benign keratosis | Benign keratosis | Malignant melanoma |
| | Malignant melanoma | Melanocytic nevus | Malignant melanoma | Melanocytic nevus |
| | Malignant melanoma | Benign keratosis | Basal cell carcinoma | Benign keratosis |
|  | Melanocytic nevus | Benign keratosis | Malignant melanoma | Melanocytic nevus |
| | Melanocytic nevus | Melanocytic nevus | Melanocytic nevus | Malignant melanoma |
| | Malignant melanoma | Melanocytic nevus | Malignant melanoma | Melanocytic nevus |
| | Malignant melanoma | Benign keratosis | Malignant melanoma | Malignant melanoma |
| | Malignant melanoma | Benign keratosis | Benign keratosis | Benign keratosis |

**Table 3.** Cont.

| Image | ResNet | ResNeXt | SE-ResNet | SE-ResNeXt |
|---|---|---|---|---|
| | Malignant melanoma | Melanocytic nevus | Melanocytic nevus | Melanocytic nevus |
| | Benign keratosis | Benign keratosis | Benign keratosis | Malignant melanoma |
|  | Melanocytic nevus | Melanocytic nevus | Malignant melanoma | Melanocytic nevus |
|  | Benign keratosis | Melanocytic nevus | Malignant melanoma | Benign keratosis |




MEL – malignant melanoma, NV – melanocytic nevus, BCC – basal cell carcinoma, AKIEC – actinic keratosis/Bowen's disease, BKL – benign keratosis, DF – dermatofibroma, VASC – vascular lesion.

**Figure 2. A –** Comparison of scores at malignant melanoma prediction between ResNet, ResNeXt, SE-ResNet, SE-ResNeXt and ensemble of all convolutional neural networks. **B** – Normalized confusion matrix of the ensemble

used in our research, is relatively new and small compared to other ones commonly used in computer vision challenges (ImageNet [6], MS COCO [7], etc.). This dataset was highly imbalanced between classes, which may have an impact on the final performance of the convolutional neural network, however, despite this limitation, a huge number of images was an attractive and valid dataset for performing our experiments. Gathering more images to this dataset might help to overcome some problems we have faced, such as different classification of images based on the camera angle, zoom and lighting of taken pictures (for details see Figure 3). This could also decrease the bias towards overrepresentation of melanocytic nevi, which had been encountered during training. Overall, having a better image database could aid researchers in developing a more robust algorithm that generalizes well over new skin lesion images.

Last but not least, the question arises whether dermatologists will be willing to use an artificial neural

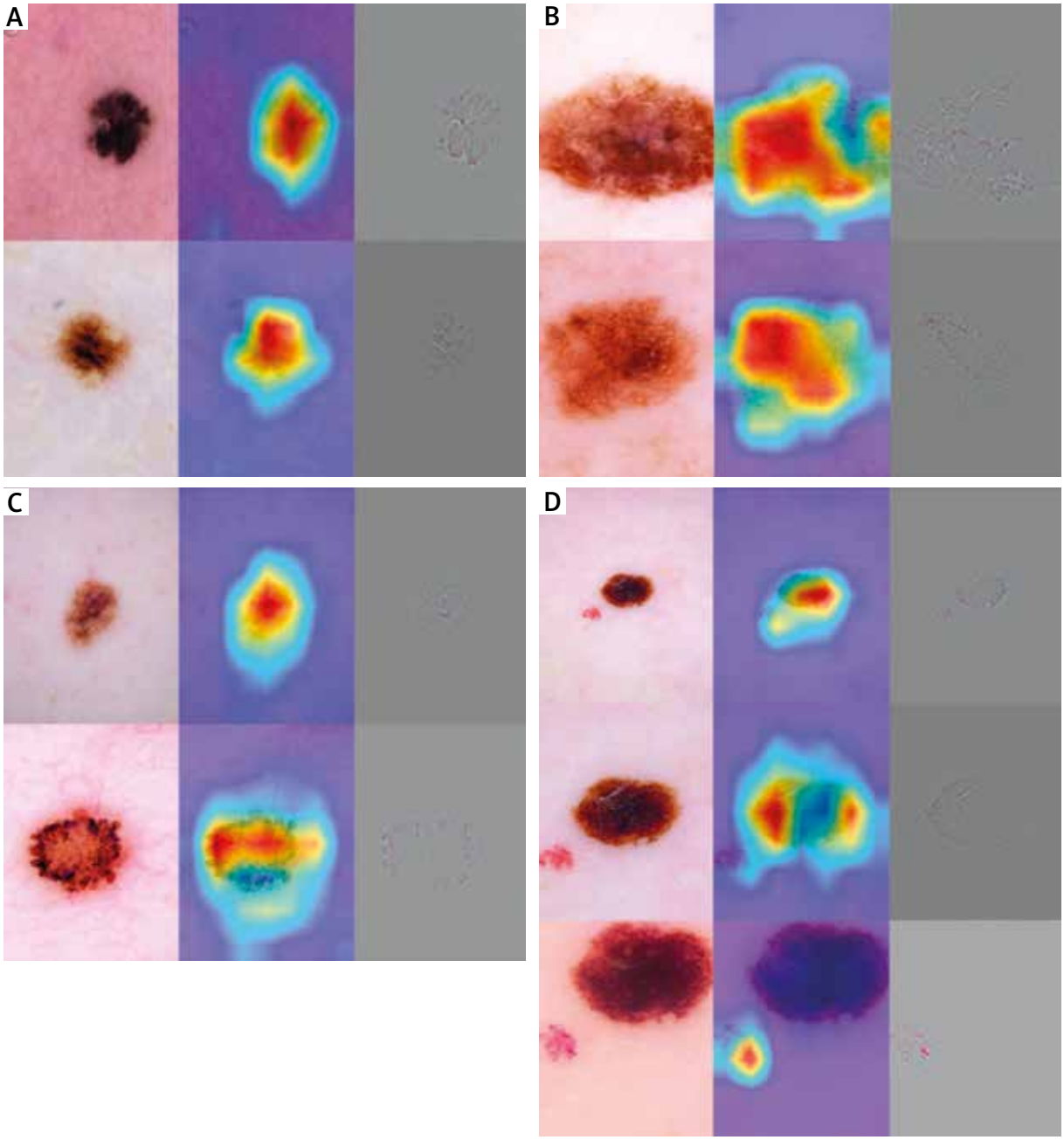



**Figure 3. A** – Examples of a correctly classified melanocytic nevus. **B** – Examples of a correctly classified malignant melanoma. **C** – Wrongly classified examples of malignant melanoma (these images were predicted as melanocytic nevus). **D** – From top to bottom ResNeXt classified those images as malignant melanoma, melanocytic nevus and vascular lesion (from left: original image, Grad-CAM heatmap on original image and Guided Grad-CAM visualization)

network, because of its opacity. Deep learning models have low interpretability, thereby they are often called as a black box algorithm [20]. The grad-CAM method that has been used in the current study allowed us to visually explain CNN prediction for the given image via a saliency map. Utilizing it or another similar method in a computer system might give dermatologists an insight which features of skin lesion were important for the particular model during classification.

## Conclusions

Our research showed that deep learning models had achieved satisfactory accuracy in malignant melanoma

detection in dermoscopy images. In future investigations, one should focus on a better understanding of CNN predictions.

## Acknowledgments

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. We would like also to thank the International Skin Imaging Collaboration: Melanoma Project (https://www.isic-archive.com) for publicly available images of skin lesions that were used in this research.

Dominika Kwiatkowska and Piotr Kluska contributed equally to this work.

## Conflict of interest

The authors declare no conflict of interest.

### References

1. Darmawan CC, Jo G, Montenegro SE, et al. Early detection of acral melanoma: a review of clinical, dermoscopic, histopathologic, and molecular characteristics. J Am Acad Dermatol 2019; 81: 805-12.
2. Stewart BW, Wild CP. World cancer report 2014. World Health Organisation 2014: 1-2. doi:9283204298.
3. Moran B, Silva R, Perry AS, Gallagher WM. Epigenetics of malignant melanoma. Semin Cancer Biol 2018; 51: 80-8.
4. Luke JJ, Flaherty KT, Ribas A, Long GV. Targeted agents and immunotherapies: optimizing outcomes in melanoma. Nat Rev Clin Oncol 2017; 14: 463-82.
5. Kwiatkowska D, Kluska P, Reich A. Beyond PD-1 immunotherapy in malignant melanoma. Dermatol Ther 2019; 9: 243-57.
6. Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition; 2009: 248-55.
7. Lin TY, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2015; 3686-93.
8. Codella NCF, Gutman D, Celebi ME, 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). In: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium On 2018: 168-72.
9. 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.
10. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015; 521: 436-44.
11. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv Prepr arXiv150203167, 2015.
12. 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; 2016: 770-8.
13. Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks. In: Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference On 2017: 5987-95.
14. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. arXiv Prepr arXiv170901507, 2017; 7.
15. Tsoumakas G, Partalas I, Vlahavas I. A taxonomy and short review of ensemble Selection. ECAI 2008, Work Supervised Unsupervised Ensemble Methods Their Appl. 2008.
16. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv Prepr arXiv14126980, 2014.
17. Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. ICCV 2017: 618-26.
18. Haenssle HA, Fink C, Schneiderbauer R, 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 2018; 29: 1836-42.
19. Yap J, Yolland W, Tschandl P. Multimodal skin lesion classification using deep learning. Exp Dermatol 2018; 27: 1261-7.