DEEP LEARNING BASED SKIN LESIONS DIAGNOSIS

D.A. Gavrilov [1]*, N.N. Shchelkunov [1], A.V. Melerzanov [1]

[1] Moscow Institute of Physics and Technology Dolgoprudny, Moscow region, Russia - (gavrilov.da, shchelkunov.nn, melerzanov.av)@mipt.ru

Commission II, WG II/10

KEY WORDS: Deep convolutional neural networks, Melanoma diagnosis, Computer vision, Telemedicine

ABSTRACT:

Melanoma is one of the most virulent lesions of human's skin. The visual diagnosis accuracy of melanoma directly depends on the doctor's qualification and specialization. State-of-the-art solutions in the field of image processing and machine learning allows to create intelligent systems based on artificial convolutional neural network exceeding human's rates in the field of object classification, including the case of malignant skin lesions. This paper presents an algorithm for the early melanoma diagnosis based on artificial deep convolutional neural networks. The algorithm proposed allows to reach the classification accuracy of melanoma at least 91%

1. INTRODUCTION

Melanoma is a malignant tumor predominantly cutaneous localization, which is today one of the most dangerous types of cancer. According to the World Health Organization (WHO), the incidence of this disease is steadily increasing from year to year. Melanoma makes the largest contribution to mortality from all types of skin cancer and is one of the fastest growing types of cancer. Late diagnosis explains the high mortality rate for melanoma. At the same time, surgical treatment gives good prognostic results and can provide almost 100% survival with the detection of neoplasms in the early stages.

The primary diagnostics of skin neoplasms is performed by specialists using macroscopic and dermatoscopic photographs with powerful magnification and uniform illumination of the part of skin being imaged (Binder M. et al. 1995). The most widely distributed symptom complex for the diagnosis of melanoma is the ABCDE test (*The ABCDEs of Melanoma*, без даты). This test was proposed by Friedman et al. in 1985 for general practitioners (Friedman R.J. et al. 1985), and it allows various parameters to be followed: A (asymmetry) − asymmetry of the pigmented spot; B (border irregularity) − unevenness of the margin; C (color) − irregular coloration; D (diameter) − having a diameter of >6 mm. The presence of three or more of these features is evidence pointing towards a malignant tumor. An additional criterion, E, is used for repeated dynamic observations of people in the at_risk group. Parameter E assesses the dynamics of changes in color, shape, and size of the pigmented skin area. In addition, there is a series of other characteristics allowing malignant neoplasms to be discriminated from benign using nothing more than an image.

The Primary diagnosis of melanoma, as a rule, is carried out by a physician visually, and further diagnosis is made using biopsy and morphological studies.

The concept of using computer vision to solve the task of identifying skin cancers arose relatively recently. Computer technology or machine vision consists of a set of methods providing for the identification and classification of different objects. Scientific studies aiming to improve the differential diagnostics of melanoma using computer techniques and expert systems were started in many large centers in Germany, Austria and other countries from 1987, following the suggestion of N. Cascinelli, the president of the WHO Melanoma Program, which has now developed into the World Melanoma Society. For a long time, results were not accurate enough for practical exploitation. Examples include systems for assessing the components of the ABCDE test (Jain et al. 2015) reaching 70% accuracy.

Clinics currently use special dermatology systems providing digital solutions allowing diagnostics using a dermatoscope, camera and special professional software. In general, such programmable systems provide computer analysis of images of pigmented skin formations and provide the opportunity to store the images obtained and to create and archive patient records and perform accurate assessments of the course of disease.

Secondary prevention or preventive medicine comes to the fore in the fight against melanoma death. Timely examinations, therapy and control of the spread of relapses. According to some authors, the timely detection and initiation of treatment for melanoma with regular self-examination or dermatological examination reduces patient mortality by 63% (Berwick M. et al. 1996).

Thus, a responsible approach to self-examination for the emersion of atypical pigmented skin lesions can provide high results of primary diagnosis and, if necessary, allow the patient to seek medical help promptly (De Giorgi V. et al. 012).

It is especially important that the percentage of melanoma detection, when using modern approaches and techniques, is often slightly different in the case of self-examination of the patient (33%) from diagnosis when examined by a professional physician as part of a general examination (36%) (De Giorgi V. et al. 2012). Thus, a responsible approach to self-examination for the appearance of atypical pigmented skin lesions can

---

* Corresponding author

provide high results of primary diagnosis and, if necessary, allow the patient to promptly seek medical help.

Current development in image processing and machine learning techniques have produced systems based on artificial neural convolutional networks which are better than humans in object classification tasks – including the diagnostics of skin neoplasms (Gavrilov, 2018). There are algorithms for automated computer analysis of dermatological images, allowing to determine the border, brightness, diameter and symmetry of pigmentation (Gareau D.S. et al. 2017) and, thus, provide assistance to doctors to improve the accuracy of diagnosis (Fink C. et al. 2017).

Recently, systems based on artificial intelligence technologies, in particular convolutional neural networks, and providing the possibility of independent detection of malignant tumors at an early stage, have been actively being developed (Esteva et al., 2017). Studies show a high potential for the use of neural network technologies for the diagnosis of skin neoplasms (Haenssle H.A. et al. 2018).

Using of computer programs for the skin cancer diagnosis can provide substantial support in diagnosis to both dermatologists and general practitioners. In addition, with the introduction of such diagnostic systems, it becomes possible to use expert decision support systems, as well as tools for remote consulting, the so-called telemedicine.

Presented here is the development of an algorithm for the early diagnosis of melanoma based on artificial deep convolutional neural networks. The algorithm discriminates between benign and malignant skin lesions with accuracy of at least 91% using automatic analysis of images of pigmented skin formations.

2. METHODS

The development of the modern technology in the field of image processing and machine learning allows one to create systems based on artificial convolutional neural networks, which are prevailing over humans in object classification tasks, including the diagnosis of malignant skin tumors (Haenssle H.A. et al. 2018).

The main problem of building deep convolutional neural networks, now, is that there is no enough public selection of images to train the system and to set up models. One of the largest archives of skin lesions is the International Skin Imaging Collaboration (ISIC) (*ISIS Archiv*, без даты). However, the presented data sets contain an insufficient quantity of required images, often the images were obtained in conditions of insufficient light, do not contain linear scale information, and also, do not have a unified disease classification system.

The authors have developed algorithms that allow the analysis of images distorted by interference in a limited sampling conditions, the variability of lighting and scatter in the shooting conditions (Figure 1).
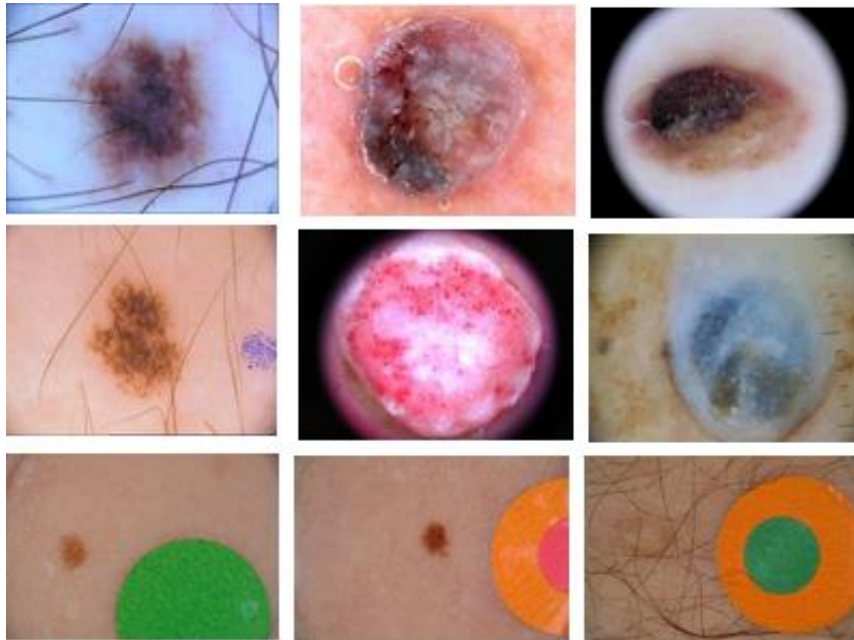

Figure 1. Examples of noise in images

An approach known as "transfer learning" was used when solving the problem. In this case, it was not envisaged teaching the neural network to classify skin diseases from scratch (Torrey et al., 2009).

The neural network with the Inception.v.3 (Szegedy et al., 2015; Shlens, 2016) architecture was chosen as a model (Figure 2). Inception.v.3 demonstrates the high quality classification of

various images in an ILSVRC (ImageNet Large Scale Visual Recognition Challenge) competition. The selected pre-trained neural network has the ability to be retrained for later use as a component of a larger network. The Inception.v.3 network was used to prepare an image classification model for the ImageNet Challenge archive data (*ImageNet Large Scale Visual Recognition Competition (ILSVRC)* (Russakovsky O. et al. 2015).
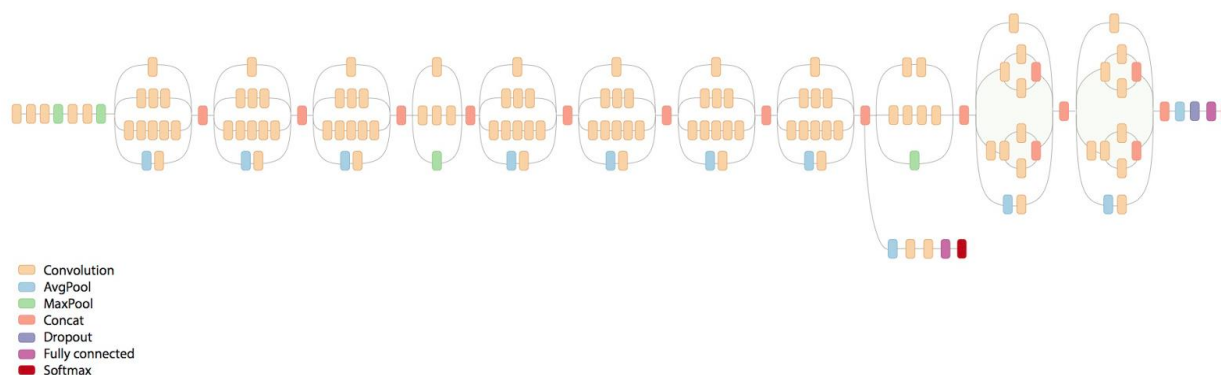
Figure 2. The architecture of the neural network "Inception v.3" (Szegedy et al. 2015; Shlens, 2016)

The pre-trained neural network tuned in to the classification of skin diseases by removing the upper classifying layers and adding new neurons to determine skin diseases. Due to the deficiency of training samples, the network was pre-trained with ImageNet images. Learning from scratch on an insufficiently large sample base could lead to retraining and the impossibility of further qualitative new data classification.

Subsequently, the resulting model was reconfigured in 10,000 photographs of skin formations, to enable the neural network to distinguish the type of skin formations. The number of source images was expanded to 1,000,000 by augmentation of the data - using various distortions to expand the training set. Insertion distortion included: turns, specular reflections, cutting out only a part of the image, stretching and compression, options for changing lighting.

Often in the sample there are images with interference, in addition, the artificial distortions introduced during the augmentation of the original images include, inter alia, interference. Interference factors and other distortions affect the recognition accuracy; however, this effect is not always negative. It should be noted that the accuracy of the network increases with the growth of the training sample. Thus, the network learns to recognize and distorted images.

The unevenness of the brightness of the images, different distances of the camera from the object when shooting, as well as the difference in the size of objects in the images were taken into account in the training set by increasing possible variations.

To control learning, a test sample was used, which images were not used in the learning process.

In total, to improve the quality of recognition and classification, 5 networks were trained. The networks were of the same architecture, but different weights. After that, all the trained neural networks were combined into an ensemble of models that ensure decision-making by majority voting. The sensitivity and specificity of the best of these models were 85% and 92% respectively, which is comparable with the diagnostic accuracy of clinical examination using dermatoscopy

The original selected model Inception.v.3 worked exclusively with images of no more than 300x300 pixels resolution, which is substantially less than the usual size of dermatological images. Taking into account the fact that for accurate diagnosis of diseases, fine details of the image, indistinguishable at this

resolution, are extremely important, convolutional layers have been added to the existing model, which allow higher resolution images classification. The modified model allowed conducting studies of dermatoscopic images with a resolution of up to 700x700 pixels.

The final test model uses the images of skin neoplasms, made with a mobile phone in daylight conditions. At the output, the system gives the probability of matching the resulting image to one of 4 classes: melanoma, nevus, seborrheic keratosis or another disease.

3. RESULTS

The main criteria for the quality of the network performance result were the area under the accuracy-completeness curve and the total probability of detection.

This ensures the accuracy of recognition of melanoma of the skin more than 91% with AUC-ROC 0.96 (ROC-curve - a graph that allows to evaluate the quality of the binary classification), which is comparable with the results of diagnostics of highly qualified doctors. Figure 3 shows the graph of AUC-ROC recognition of melanoma-nevus for an ensemble of models.
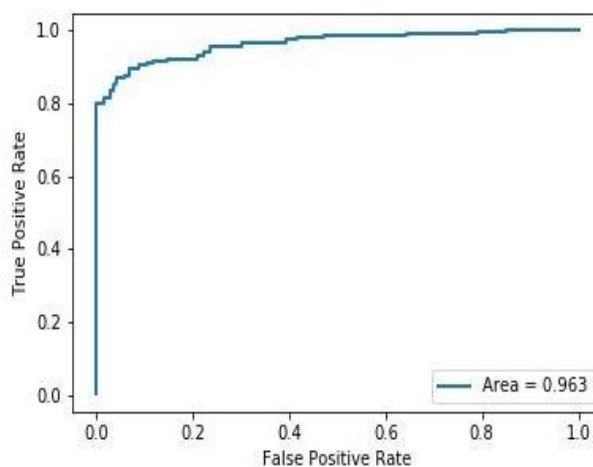

Figure 3. The graph of AUC-ROC recognition of melanoma-nevus for an ensemble of models

The variety of skin neoplasms, the complexity of their structure, and the similarity of the clinical picture of various forms of skin lesions causes difficulties of visual inspection and diagnosis, even among specialists. Taking into account the fact that the accuracy of visual diagnosis of melanoma directly depends on the qualifications and specialization of the doctor, as well as on the frequency of occurrence of the disease in his daily practice, the developed model allows to separate malignant neoplasms from benign ones as an additional task. For example, the accuracy of determining such a benign formation as seborrheic keratosis from a dermatological photograph is 97% and AUC-ROC 0.99 (Figure 4). Thus, the proposed algorithm allows us confident distinguishing melanoma from seborrheic keratosis, as well as melanoma from benign nevus.
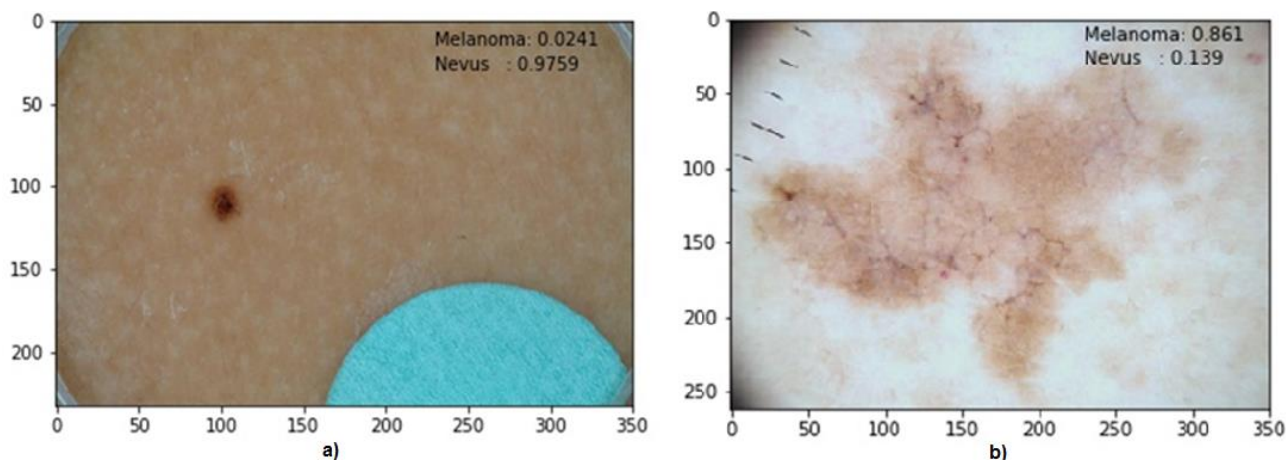

Figure 4. The examples of classification of a) nevus, b) melanoma

4. CONCLUSION

The work has created a high-precision deep learning neural network for automated diagnosis of skin tumors.

During the works, a training sample was formed, comprising real photographs of skin lesions and the images obtained from the source through various distortions. The unique augmentation algorithms, coupled with the loss functions, are the main authoring, allowing to achieve a high quality classification of skin diseases with a limited training set.

The developed model allows quality diagnosis of melanoma with an accuracy of not less than 91%, which is comparable with the diagnostic capabilities of highly qualified dermatologists. The use of intelligent systems of this type to identify skin diseases will provide substantial support in the diagnosis of both dermatologists and general practitioners.

The system is available as a test version at skincheckup.online. The accessibility of the project will allow anyone to carry out preliminary self-diagnostics using personal photographs.

REFERENCES

Berwick M., Begg C.B., Fine J.A., Roush G.C., Barnhill, R. L. (1996) «Screening for cutaneous melanoma by skin self-examination», *J Natl Cancer Inst*, 88(1), cc. 17–23.

Binder M., Schwarz M., Winkler A., Steiner A., Kaider A., Wolff K., Pehamberger, H. (1995) «Epiluminescence Microscopy. A Useful Tool for the Diagnosis of Pigmented Skin Lesions for Formally Trained Dermatologists», *Archives of Dermatology*, 131(3), cc. 286–291.

Esteva, A. *и др.* (2017) «Dermatologist-level classification of skin cancer with deep neural networks», *Nature*. Nature Publishing Group, 542(7639), cc. 115–118. doi: 10.1038/nature21056.

Fink C., Jaeger C., Jaeger K., Haenssle, H. A. (2017) «Diagnostic performance of the MelaFind device in a real-life clinical setting», *J Dtsch Dermatol Ges*, 15(4), cc. 414–419.

Friedman R.J., Rigel D.S., Kopf, A. W. (1985) «Early detection of malignant melenoma:therole of physician exsamination and self-examination o the skin», *CA: Cancer J.Clin.*, (35), cc. 130–151.

Gareau D.S., da Rosa J.C., Yagerman S., Carucci J.A., Gulati N., DeFazio J. L., Suárez- Fariñas M., Marghoob A., Krueger, J. G. (2017) «Digital imaging biomarkers feed machine learning for melanoma screening», *Experimental Dermatology*, 26, cc. 615–618.

Gavrilov, D. A. (2018) *Artifical intelligence-Al image recognition for helthcare*, *16th Anti-Aging and Aesthetic Medicine World Congress*.

De Giorgi V., Grazzini M., Rossari S., Gori A., Papi F., Scarfi, F. (2012) «Is skin self-examination for cutaneous melanoma detection still adequate? A retrospective study», *Dermatology*, 225(1), cc. 1–6.

Haenssle H.A., Fink C., Schneiderbauer R., Toberer F., Buhl T., Blum A., KallooA A., Hassen A. B. H., Thomas L., Enk A., Uhlmann, L. (2018) «Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists», *Annals of Oncology*, 29(8), cc. 1836–1842.

ImageNet Large Scale Visual Recognition Competition (ILSVRC) (без даты). [Electronic resource]. URL: http://www.image-net.org/challenges/LSVRC/ (accessed: 13.07.2018 г.).

ISIS Archiv Kitware, Inc. [Electronic resource]. URL: https://isic-archive.com/ (accessed: 24.01.2018 г.).

Jain, S., Jagtap, V. и Pise, N. (2015) «Computer aided melanoma skin cancer detection using image processing», *Procedia Computer Science*. Elsevier Masson SAS, 48, cc. 736–741. doi: 10.1016/j.procs.2015.04.209.

Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A.C., Fei-Fei, L. (2015) «ImageNet Large Scale Visual Recognition Challenge», *International Journal of Computer Vision*, 115(3), cc. 211–252.

Shlens, J. (2016) *Train your own image classifier with Inception in TensorFlow*. [Electronic resource]. URL: https://research.googleblog.com/2016/03/train-your-own-image-classifier-with.html (accessed: 24.01.2018 г.).

Szegedy, C. *и др.* (2015) *Rethinking the Inception Architecture for Computer Vision*. doi: 10.1109/CVPR.2016.308.

Talha Khan, B. (2018) «Machine learning model in melanoma», *The Lancet Oncology*, 19(7), с. 340.

The ABCDEs of Melanoma [Electronic resource]. URL: https://www.melanoma.org/understand-melanoma/diagnosing-melanoma/detection-screening/abcdes-melanoma (accessed: 24.01.2018 г.).

Torrey, L. и Shavlik, J. (2009) *Transfer Learning*, *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*. IGI Global. doi: 10.1016/j.jbi.2011.04.009.