

Research Article

Artificial Intelligence (Neural Network) in the Diagnosis of Benign Skin Tumors in Pediatric Patients

Rina Khabarova^{1,2,*} , **Svetlana Kulyova^{1,2}** , **Evgeniy Senchurov^{1,2}** ,
Elena Mikhailova¹ , **Kseniya Borokshinova¹** , **Alika Kulyova¹** 

¹Children Oncology Department, N. N. Petrov National Medical Research Center of Oncology, Ministry of Health of the Russian Federation, Saint Petersburg, Russian Federation

²Department of Oncology, Children's Oncology and Radiotherapy of Saint Petersburg State Pediatric Medical University, Ministry of Health of the Russian Federation, Saint Petersburg, Russian Federation

Abstract

Relevance: ten years ago, artificial intelligence (AI), particularly neural networks (NN), as a diagnostic option in practice seemed a distant prospect. Today, the use of AI is becoming an increasingly popular and daily improving approach in all aspects of clinical and fundamental medicine. Purpose: design and learning of a NN to recognize four types of benign melanocytic skin tumors, integration into a mobile app to apply in practice. Material and methods: clinical and dermatoscopic analysis of skin tumors was carried out in 600 children. In 65 cases the tumors were removed. Histological types were dermal nevus – 43% (n=28), complex nevus - 33.8% (n=22), pyogenic granuloma - 10.8% (n=7), Spitz-nevus - 6.2% (n=4), blue nevus - 3.1% (n=2), melanoma - 3.1% (n=2). Seven patients with pyogenic granulomas and two patients with melanoma were excluded. The test set included 56 dermatoscopic images. Due to the small number of images augmentation was performed. The database has been increased from 600 images to 1800. NN is written in the machine language Python. The machine learning framework was TensorFlow 2.0. The network architecture is based on the pre-trained model “EfficientNet B7”. This model uses the “supervised learning” paradigm. Each element of the sample had a class affiliation. Results: an accuracy of 83% was achieved after a period of learning on the test set. Mathematical metrics calculated in the Scikit-learn library. Sensitivity was 100% (blue nevus), 73% (complex nevus), 93% (dermal nevus), 75% (Spitz-nevus), and specificity were 98%; 94%; 82%; 98%, respectively. AI was integrated into the mobile app “KIDS NEVI”. Conclusion: AI as an auxiliary method for the skin tumors diagnosis in children and adolescents has demonstrated high potential and great opportunities. Dermatoscopic analysis of a skin tumor and a mobile app are able to provide “double control”, quick and correct clinical diagnosis and determination treatment tactics.

Keywords

Children, Benign Skin Tumors, Artificial Intelligence, Neural Network, Dermatoscopy

*Corresponding author: izmozherova@yandex.ru (Rina Khabarova)

Received: 17 July 2024; **Accepted:** 15 August 2024; **Published:** 6 September 2024



1. Introduction

Childhood is a dynamic period of melanocytic skin tumors development. The importance of a quick diagnostics and decision to follow-up or remove the tumor is unquestionable [1, 2].

Information about natural evolutionary transformations of nevi is still rare. In some cases, diagnostics is difficult. Clinical or the "naked eye" research of a skin tumor has quite low information [2-4]. Dermatoscopy is a non-invasive, highly informative diagnostic method increasingly used in clinical oncologic and dermatologic practice. This painless technique is optimal for assessing skin tumors in children and adolescents. The high accuracy of dermoscopy was also confirmed in a study by the N. N. Petrov National Medical Research Center of Oncology on pediatric cohort [4].

Today, the need to create auxiliary methods of non-invasive diagnostics is evident. New developments will make it possible to determine patient treatment tactics and avoid unnecessary surgery.

Ten years ago, artificial intelligence (AI), particularly neural networks (NN), as a diagnostic option in practice seemed only a distant prospect. The understanding of the potential machine learning and the technical component was beginning. Today the use of NN in all aspects of clinical and fundamental medicine is an increasingly popular and daily improving approach [5].

No publications have been found on the topic of AI in the diagnosis of skin tumors in children and adolescents. The absence of articles on this theme notes the scientific novelty of the research.

The main aims of study were the development and learning a neural network to recognize four types of benign melanocytic skin tumors, evaluate diagnostic information content, and integrate the AI into a mobile app for practical use. The combined use of dermatoscopic and mobile applications can provide "double control", fast and correct clinical diagnosis, and determination of treatment tactics.

The NN was developed on the basis of the "EfficientNet B7" architecture. Learning was conducted on 1800 (including augmentation) dermatoscopic images of pediatric and adolescent patients observed at the N. N. Petrov National Medical Research Center of Oncology. The accuracy was assessed on the test sample and the diagnostic information value (sensitivity, specificity and other metrics) was determined.

2. Materials and Methods

Clinical and dermatoscopic analysis of skin tumors was performed in 600 pediatric and adolescent patients from 2016 to 2022. The study protocol was approved by the biomedical ethics committee of NMRC of Oncology named after N.N.Petrov of MoH of Russia. Protocol №42 dated 28.10.2020. Removal of lesions with morphologic verification was performed in part of patients (n = 65). The indications for removal were different: from signs of superficial

trauma to atypical dermatoscopic picture requiring pathomorphological and immunohistochemical analysis.

The gender ratio was 1: 1; boys 33 (50.8%) and girls 32 (49.2%).

The age of the patients ranged from 2 to 17 years, with a median of 14 years. More than half of the children were of pubertal age (n = 45 or 69.2%).

Maximum follow-up period was 1569 days, with a median of 16.5 days.

Diameter of skin neoplasms varied from small nevi to the size of giant congenital nevi. The minimum diameter was 0.2 cm and the maximum diameter was 60 cm.

Localization of tumors varied: 16 (24.6%) patients had elements on the limbs, 61.5 (40%) on the trunk, 2 (3.1%) on the scalp, and 7 (10.8%) on the neck (Figure 1).

The dermatoscopic pattern was analyzed according to generally accepted algorithms for the diagnosis of adult skin neoplasms with an age-adapted approach according to the ABCDE rule (A - asymmetry, B - border, C - color, D - diameter, E - evolution) (Stolz W. et al., 1994), 11 - positional test (Menzies S. W. et al., 1996), 7 - positional test (Argenziano G. et al., 1998). Dermatoscopic criteria (abs., %) are presented in Table 1.

Skin lesions were removed by excision, ablative laser surgery, and total punch biopsy. Staged morphological diagnosis verifies the diagnosis in all patients, allowing to choose the adequate extent of surgical procedures. Dermal nevus was detected in 43% of cases (n = 28), complex nevus was detected in 33.8% (n = 22), pyogenic granuloma was detected in 10.8% (n = 7), Spitz-nevus was detected in 6.2% (n = 4), blue nevus was detected in 3.1% (n = 2), and melanoma was detected in 3.1% (n = 2) (Figure 2).

Benign skin melanocytic tumors of four histological types - blue nevus, complex nevus, dermal nevus and Spitz - nevus - were selected for further analysis and NN development.

Seven patients with vascular neoplasms - pyogenic granulomas (capillary-lobular granulomas) and 2 patients with skin melanoma were excluded from the test sample. Exclusion of patients with pyogenic granulomas is justified by another tumorigenesis. Children with verified melanoma were also removed from study due to poor photographic material. Melanoma is an acute problem in the structure of oncologic pathology. This tumor is extremely rare in young patients [6, 7]. Pediatric melanoma is detected in less than 2 cases per 1 million [8]. The rarity of disease limits the dermatoscopic image database set and makes it impossible to attempt correct machine learning (ML) and NN development.

Augmentation was performed due to the small number of images in the training sample. The database was enlarged from 600 to 1800 by reflecting, enlarging, lighting, changing the lighting and contrast of the images (Figure 3). Image augmentation generates random images based on existing training data to improve the generalization ability of models.

The accuracy of the model did not exceed 70 percent without the use of augmentation. The test sample included 56 der-

matoscopic images.

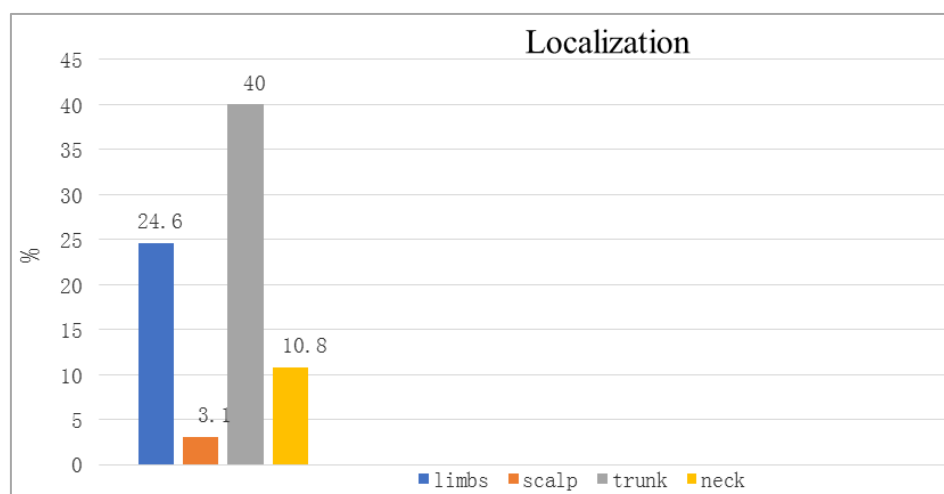


Figure 1. Distribution of tumor localization (16 (24.6%) on the limbs, 61.5 (40%) on the trunk, 2 (3.1%) on the scalp, and 7 (10.8%) on the neck).

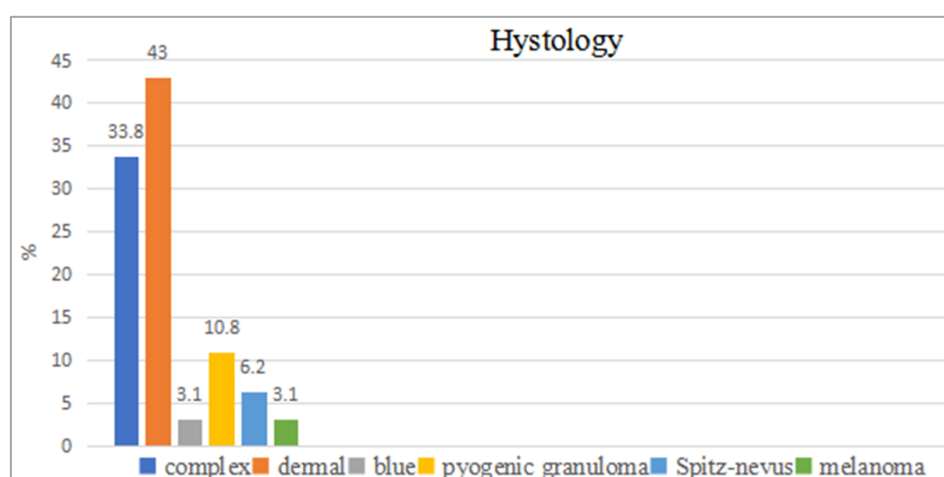


Figure 2. Distribution of histological type (dermal nevus was detected in 43% of cases ($n = 28$), complex nevus in 33.8% ($n = 22$), pyogenic granuloma in 10.8% ($n = 7$), Spitz-nevus in 6.2% ($n = 4$), blue nevus in 3.1% ($n = 2$), melanoma in 3.1% ($n = 2$)).

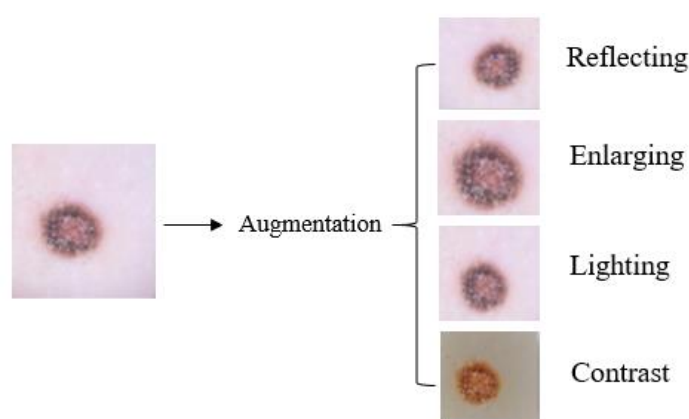


Figure 3. Augmentation.

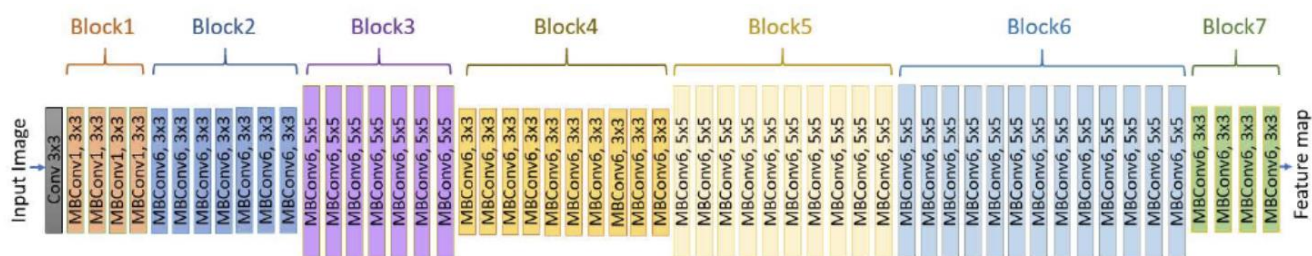


Figure 4. The architecture model "EfficientNet B7".

```

Epoch 8/15
57/57 [=====] - 93s 2s/step - loss: 0.1335 - accuracy: 0.9610 - val_loss: 1.5669 - val_accuracy: 0.7500
Epoch 9/15
57/57 [=====] - 93s 2s/step - loss: 0.1305 - accuracy: 0.9632 - val_loss: 1.6614 - val_accuracy: 0.7917
Epoch 10/15
57/57 [=====] - 93s 2s/step - loss: 0.1301 - accuracy: 0.9643 - val_loss: 2.0759 - val_accuracy: 0.7500
Epoch 11/15
57/57 [=====] - 93s 2s/step - loss: 0.1294 - accuracy: 0.9599 - val_loss: 2.1312 - val_accuracy: 0.7083
Epoch 12/15
57/57 [=====] - 93s 2s/step - loss: 0.1328 - accuracy: 0.9599 - val_loss: 1.7550 - val_accuracy: 0.6250
Epoch 13/15
57/57 [=====] - 93s 2s/step - loss: 0.1200 - accuracy: 0.9671 - val_loss: 2.1633 - val_accuracy: 0.6250
Epoch 14/15
57/57 [=====] - 93s 2s/step - loss: 0.1206 - accuracy: 0.9682 - val_loss: 1.7448 - val_accuracy: 0.7500
Epoch 15/15
57/57 [=====] - 93s 2s/step - loss: 0.1282 - accuracy: 0.9632 - val_loss: 0.4463 - val_accuracy: 0.8333

```

Figure 5. Adjustment of synaptic weights, 83% accuracy on the test set.

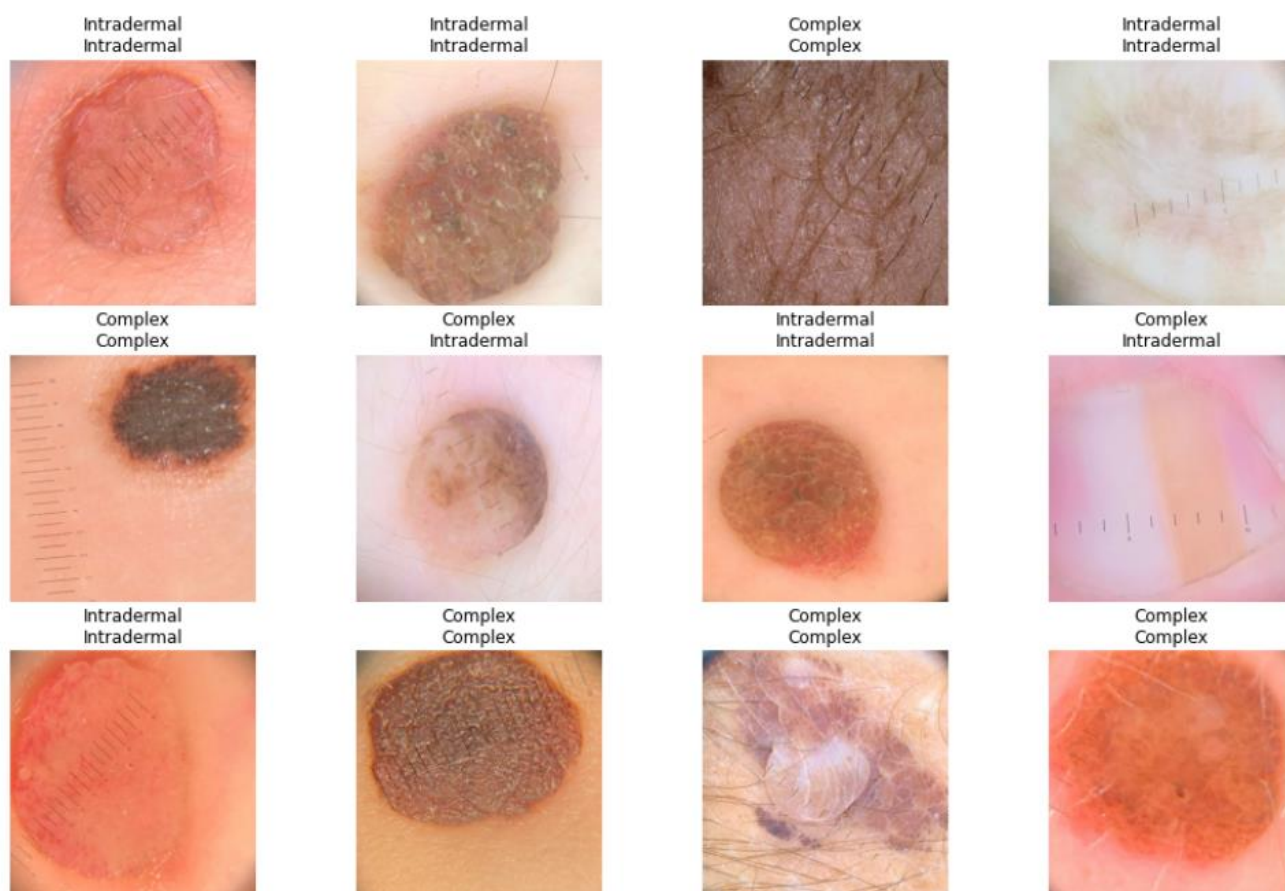


Figure 6. Neural network model after learning.

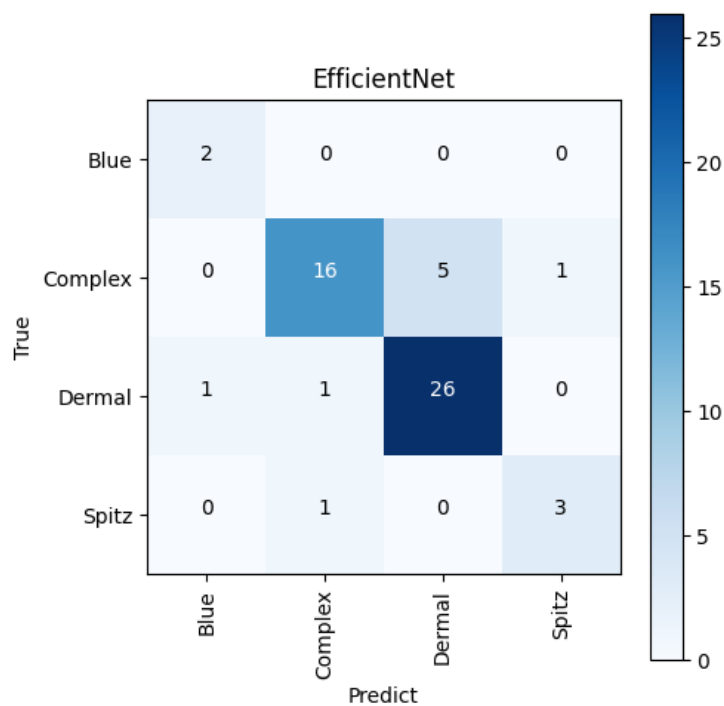


Figure 7. Confusion matrix.

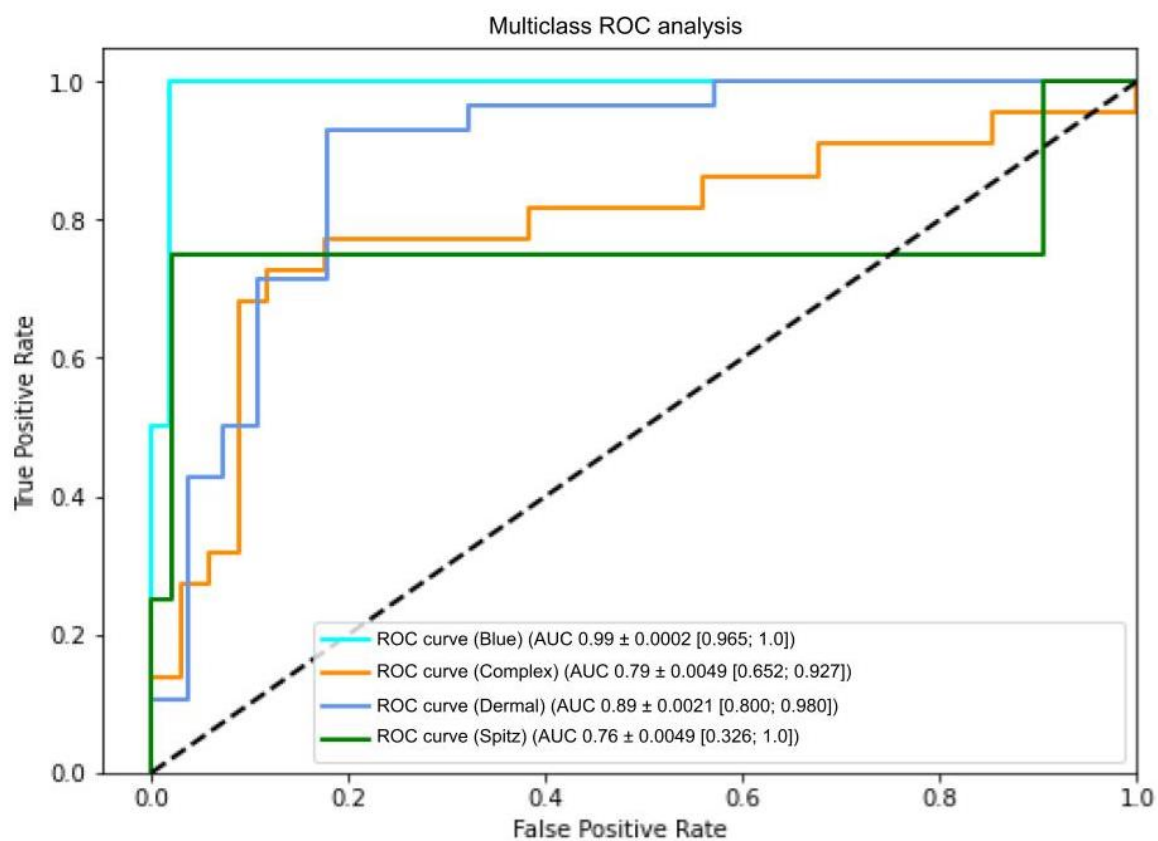


Figure 8. ROC curve, AUC. *AUC - area under the curve, (blue nevus - 0.99; complex nevus - 0.79; dermal nevus - 0.89; Spitz-nevus - 0.76).

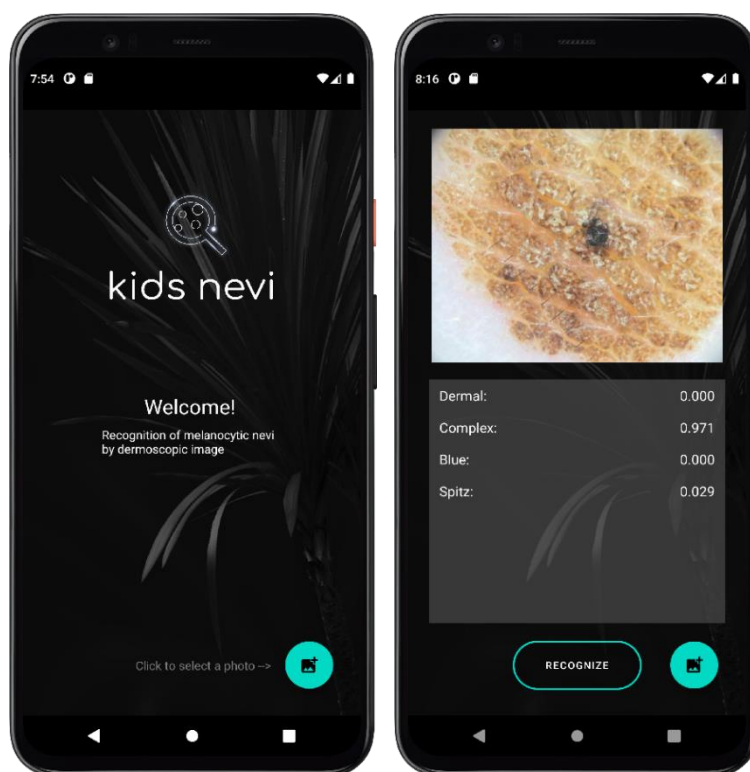


Figure 9. Mobile app “Kids nevi”.

Table 1. Dermatoscopic criteria.

Dermatoscopic criteria	abs	%
Typical pigment network	5	7,7
Atypical pigment network	3	4,6
Diameter < 2 mm	-	-
Diameter 3-5 mm	20	30,7
Diameter 6-8 mm	23	35,4
Diameter 9-10 mm	8	12,3
Diameter > 11 mm	17	26,2
Peripheral pigment network	2	3
Peripheral pigment globules	7	10,8
Network break to the periphery	1	1,5
Vascular structures in the form of comma	22	33,8
Typical globules	21	32,3
Atypical vascular pattern	1	1,5
Atypical globules	6	9,2
Homogeneous pink areas	7	10,8
Blue-white veil	3	4,6
Pigment blots	16	24,6
Pigment points	5	7,7
Asymmetry	24	36,9

Dermatoscopic criteria	abs	%
Regressive areas	2	3
Irregular borders	26	40
Central hyperpigmentation	16	24,6
Irregular coloring	51	78,5
«Cobblestone street» structure	3	4,6
Multicolor (more than 3 colors)	17	26,2
Homogeneous blue-gray pigmentation	2	3
Homogeneous area of dark brown / black color	4	6,2
Multicomponent	10	15,4
«Starburst»	3	4,6
Homogeneous structureless area	33	50,8
Ulceration	5	7,7

Table 2. Mathematical metrics (accuracy, sensitivity, specificity, f1-score).

Histological type	Artificial intelligence			
	<i>precision</i>	<i>recall</i>	<i>specificity</i>	<i>F1-score</i>
Complex nevus	0,89	0,73	0,94	0,8
Dermal nevus	0,84	0,93	0,82	0,8
Blue nevus	0,67	1	0,98	0,8
Spitz-nevus	0,75	0,75	0,98	0,75

3. Results

The NN was written in the Python programming language based on the training and test sample data. The TensorFlow 2.0 framework was used in ML. The network architecture is based on the pre-trained model “EfficientNet B7” (Figure 4). This model uses the “supervised learning” paradigm. Each element of the sample had a class affiliation.

The method of error back propagation was applied. 83% accuracy was achieved after 15 epochs of learning on the test sample (Figure 5).

The training is done on an NVIDIA 2080Ti GPU for 8 hours.

AI recognition was considered correct when there was a match (e.g., complex/complex). Prediction was incorrect when there was a mismatch (e.g., complex/dermal).

A fragment of the model developed and trained by NN is shown in Figure 6. Two morphologic variants are displayed above each dermatoscopic image: the first is the result of the NN prediction and the second is the true (histologic) type of

the neoplasm. AI recognition was considered correct if there was a match (e.g., complex/complex). Prediction was considered incorrect when there was a mismatch (e.g., complex/dermal) (Figure 6).

Visual assessment of diagnostic error was performed by comparing actual and predicted values after constructing a confusion matrix (Figure 7).

Matrix columns - predicted values of the target variable based on the results of pattern recognition by trained NN; rows - actual (resulting) indicators (morphological verification). Numbers outside the main diagonal correspond to the number of diagnostic errors. Single numbers outside the main diagonal confirm high information content of this method. The maximum diagnostic error was found in the identification of complex nevus. In five cases, the lesion was identified as a dermal nevus (Figure 7).

Figure 7 interpretation:

- 1) number of patients (n = 56) - the total value in the matrix;
- 2) the main diagonal (from left to right, from top to down) corresponds to the matching of AI and histologic diagnosis;

- 3) numbers in the main diagonal correspond to matches in the diagnoses;
- 4) numbers outside the main diagonal correspond to the number of diagnostic errors (row demonstrates the correct histologic diagnoses; column demonstrates the incorrect predicted diagnoses);
- 5) color gradation (from light to dark) corresponds to increasing numbers.

Accuracy (proportion of objects truly belonging to a given class relative to all objects), sensitivity (proportion of true positive classifications), specificity (proportion of true negative classifications) and f1-score (joint estimate of accuracy and sensitivity) are estimated using the following formulas:

Precision (PPV- positive predictive value):

$$PV = \frac{TP}{TP+FP}$$

TP – true positive rate и FP – false positive rate.

Sensitivity (recall, hit rate, TPR true positive rate):

$$TPR = \frac{TP}{TP+FN}$$

FN – false negative rate.

Specificity (selectivity, TNR):

$$TNR = \frac{TN}{FP+TN}$$

TN – true negative rate.

F1-score (joint assessment of accuracy and sensitivity):

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Calculation of mathematical metrics is performed in the Scikit-learn library (Table 2).

Sensitivity was 100% for blue nevus, 73% for complex nevus, 93% for dermal nevus, and 75% for Spitz-nevus; specificity was 98%, 94%, 82%, and 98%, respectively, despite the limited sample.

Receiver Operator Characteristic (ROC) curves were constructed and Area Under Curve (AUC) was determined to analyze the association between several determinants, technique, and morphology (the resulting characteristics were nevus blue, complex nevus, dermal nevus, and Spitz's nevus). The AUC for blue nevus was 0.99; for complex nevus, 0.79; for dermal nevus, 0.89; and for Spitz-nevus, 0.76 (Figure 8).

NN has been integrated into the “Kids nevi” mobile app. The app is available on the Android platform.

How the application works:

- 1) Launch the mobile application
- 2) Tap the "select photo" button, uploading the dermatoscopic image.
- 3) Tap the "recognize" button, analyzing the image.
- 4) Identifying the type of melanocytic nevus (blue nevus, complex nevus, dermal nevus or Spitz-nevus) with an

accuracy of one thousandth (Figure 9).

4. Discussion

The availability of large-scale data combined with advances in high-performance computing and innovative deep learning architectures has led to the rapid adoption of artificial intelligence in practicing medicine, both oncology and other specialties [9, 10].

The purpose of software developed and trained with AI in oncology is diagnosis, process staging, assessment of molecular characteristics of tumors and their microenvironment, finding new therapeutic options, prediction of treatment outcomes, etc. [11-13].

Foreign studies devoted to machine learning and NN in the diagnosis of skin neoplasms by analyzing digital images are presented in limited quantities. Most articles describe the possibilities of AI for the recognition and differential diagnosis of melanoma in adult patients. There is no similar experience in pediatric patients.

5. Conclusion

- 1) Benign skin neoplasms in children and adolescents are represented by a large variety of melanocytic, vascular and other genesis. The diversity of skin lesions explains the search for additional diagnostic possibilities adapted to age period.
- 2) The sample was limited, but artificial intelligence as an auxiliary method of melanocytic skin tumors diagnostic in children and adolescents showed high potential and great prospects. The feasibility of further improvement by expanding the database and morphological variants of tumors is beyond doubt.
- 3) The mobile app allows analyzing dermatoscopic image and recognizing nevus type. Dermatoscopic analysis of a skin tumor and a mobile app are able to provide “double control”, quick and correct clinical diagnosis and determination treatment tactics.

Abbreviations

AI	Artificial Intelligence
NN	Neural Network
ABCDE	A - Asymmetry, B - Border, C - Color, D - Diameter, E - Evolution
ML	Machine Learning
PPV	Positive Predictive Value
TP	True Positive
FP	False Positive
TPR	True Positive Rate
FN	False Negative
TN	True Negative
ROC curves	Receiver Operator Characteristic

AUC Area Under Curve
 DERM Deep Ensemble for Recognition of Melanoma

Author Contributions

Rina Khabarova: Methodology, Software, Visualization, Writing – original draft

Svetlana Kulyova: Conceptualization, Data curation, Project administration, Validation, Supervision, Writing – review & editing

Evgeniy Senchurov: Resources, Visualization

Elena Mikhailova: Visualization, Writing – original draft

Kseniya Borokshina: Visualization, Writing – original draft

Alika Kulyova: Visualization, Writing – original draft

Funding

The study was performed without external funding.

Data Availability Statement

The data supporting the outcome of this research work has been reported in this manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Zalaudek I., Hofmann-Wellenhof R., Kittler H., Argenziano G., Ferrara G., Petrillo L., Kerl H., Soyer H. P. A Dual Concept of Nevogenesis: Theoretical Considerations Based on Dermoscopic Features of Melanocytic Nevi. *J Dtsch Dermatol Ges.* 2007; 5(11): 98–92. <https://doi.org/10.1111/j.1610-0387.2007.06384>
- [2] Schaff er J. Update on melanocytic nevi in children. *Clin Dermatol.* 2015; 33(3): 368–86. <https://doi.org/10.1016/j.clindermatol.2014.12.015>
- [3] Haliasos H., Zalaudek I., Malvey J., Lanschuetzer C., Hinter H., Hofmann-Wellenhof R., Braun R., Marghoob A. A. Dermoscopy of Benign and Malignant Neoplasms in the Pediatric Population. *Semin Cutan Med Surg.* 2010; 29(4): 218–31. <https://doi.org/10.1016/j.sder.2010.10.003>
- [4] Kulyova S. A., Khabarova R. I. Diagnostic informativeness of the dermoscopic pattern of skin neoplasms in children and adolescents. *Russian Journal of Pediatric Hematology and Oncology.* 2021; 8(4): 14–19. <https://doi.org/10.21682/2311-1267-2021-8-4-14-19>
- [5] Currie G., Elizabeth H., Rohren E. Machine Learning and Deep Learning in Medical Imaging: Intelligent Imaging. *J Med Imaging Radiat Sci.* 2019; 50(4): 477–487.

<https://doi.org/10.1016/j.jmir.2019.09.005>

- [6] Merabishvili V. M. Malignant melanoma. Epidemiology, analytical indicators of the effectiveness of the oncological service (population study). *Issues of Oncology.* 2017; 63(2): 221–233. UDC 616.5-006.81: 314.4.
- [7] Barchuk A. A., Arseniev A. I., Belyaev A. M. et al. Efficiency of cancer screening. *Issues of oncology.* 2017; 63(4): 557–567. UDC 616-006.04-004.822].
- [8] Ferrari A., Brecht I., Gatta G. at al. Definding and listing very rare cancers of periadric age: consensus of the Joint Action on Rare Cancers in cooperation with the European Cooperative Study Group for Pediatric rare tumors. *Current Perspective.* 2019; 110(1): 120–126. <https://doi.org/10.1016/j.ejca.2018.12.031>
- [9] Zukotynski K., Gaudet V., Uribe C. F. Machine Learning in Nuclear Medicine: Part 2-Neural Networks and Clinical Aspects. *J Nucl Med.* 2021; 62(1): 22–29. <https://doi.org/10.2967/jnumed.119.231837>
- [10] Le Cun Y, Bengio Y, Hinton G. Deep Learning. *Nature.* 2015; 521: 436–444. <https://doi.org/10.1038/nature14539>
- [11] Barchuk A. A., Podolsky M. D., Belyaev A. M. et al. Automated diagnostics in population-based lung cancer screening. *Oncology Issues,* 2017; 63(2): 215–220. UDC 616.24-006.
- [12] Shimizu H, Nakayama KI. Artificial intelligence in oncology. *Cancer Sci.* 2020; 111: 1452–1460. <https://doi.org/10.1111/cas.14377>
- [13] Bhinder B., Gilvary C., Madhukar N. S., Elemento O. Cancer Discov. 2021; 11(4): 900–915. <https://doi.org/10.1158/2159-8290.CD-21-0090>

Biography

Rina Khabarova: Cand. of Sci. (Med.), Pediatric Oncologist of the Children Oncology Department, postgraduate student of the Research Department of Innovative Therapeutic Oncology and Rehabilitation Methods of N.N. Petrov National Medical Research Center of Oncology, Ministry of Health of the Russian Federation, Assistant of the Department of Oncology, Children's Oncology and Radiotherapy of Saint Petersburg State Pediatric Medical University.

Svetlana Kulyova: Dr. of Sci. (Med.), Associate Professor, Head of the Research Department of Innovative Therapeutic Oncology and Rehabilitation Methods, Professor of the Training and Methodology Department of N.N. Petrov National Medical Research Center of Oncology, Head of the Department of Oncology, Children's Oncology and Radiotherapy of Saint Petersburg State Pediatric Medical University, Chief Freelance Children's Specialist Oncologist of Saint-Petersburg Health Committee.

Evgeniy Senchurov: Pediatric Oncologist of the Children Oncology Department, postgraduate student of the Research Department of Innovative Therapeutic Oncology and Rehabilitation Methods of N.N. Petrov National Medical Research Center of Oncology, Ministry of Health of the Russian Federation, Assistant of the Department of Oncology, Children's Oncology and Radiotherapy of Saint Petersburg State Pediatric Medical University.

Elena Mikhailova: Pediatric Oncologist of the Children Oncology Department, postgraduate student of the Research Department of Innovative Therapeutic Oncology and Rehabilitation Methods of N.N. Petrov National Medical Research Center of Oncology, Ministry of Health of the Russian Federation.

Kseniya Borokshinova: Pediatric Oncologist of the Children Oncology Department, postgraduate student of the Research Department of Innovative Therapeutic Oncology and Rehabilitation Methods of N.N. Petrov National Medical Research Center of Oncology, Ministry of Health of the Russian Federation.

Alina Kulyova: Postgraduate Student of the N.N. Petrov National Medical Research Center of Oncology, Ministry of Health of the Russian Federation.