Skin Lesion Analysis Using Ensemble of CNN with Dermoscopic Images and Metadata

Sergey Milantev[a], Vyacheslav Olyunin[a], Natalya Milanteva[a], Ilya Bykov[a] and Igor Bessmertny[a]

[a]*ITMO University, St. Petersburg, 197101, Russia*

Abstract

Skin cancer is the most common type of cancer in the existing world, accounting for one third of all cancer cases. This paper describes approach for ISIC Skin Lesion Classification Challenge.Its main purpose is to classify skin lesions basing on dermoscopic images and additional patient metadata. A diverse dataset of 25331 images containing images from eight classes, including an additional unknown class provided by ISIC was used for training. There also were used about 8 thousand images from external private datasets: seven-point criteria database, SD-198, MED-NODE, PH2 and SKINL2v2. Ultimately there were 32748 images to train neural network models. The data has different resolutions and a large class imbalance, so different input model resizes, data augmentation and weighted cross-entropy loss was used. In this paper, there was considered the use of deep learning for skin lesions with advanced data preprocessing. There was decided to classify skin lesions using convolutional neural networks. Each image was segmented using R2U-Net and and black areas were removed from it before being fed into the convolutional network for classification. The unnecessary parts of the image was removed using segmentation method which returned accurate extraction of the skin lesion region. Shades of Grey and data augmentation processed each image. We used R2U-Net for segmentation and EfficientNet-B0-B7, SENet-154, ResNeXt-101 32x4d and Inception-ResNet-v2 to classify skin diseases. We applied a fully-connected dense layers to each model to attach metadata. Median method was applied for outlier or missing values. Label encoding was performed binary features and one-hot encoding was used for other non-binary features. We experimented with this approach and the results of this experiments were spectacular. All models with an ensemble were combined. The best ensemble can achieve a sensitivity of 81.5% and specificity of 97.7%.

Keywords

Skin Lesion Analysis, Deep Learning, Convolutional Neural Network, R2U-Net, EfficientNet, ResNeXt, SENet, Inception-ResNet,

1. Introduction

Skin cancer is the most common cancer globally. Nowadays skin cancer is an acute social problem, because it is one of the most common human diseases, affecting people of all ages and sex and progressing rapidly and rapidly around the world [1]. This problem primarily arises due

Proceedings of the 12th Majorov International Conference on Software Engineering and Computer Systems, December 10–11, 2020, Online Saint Petersburg, Russia

geerkus@gmail.com (S. Milantev); vyachaslav.olynin@gmail.com (V. Olyunin); natalya.milanteva@gmail.com (N. Milanteva); il.byckoff2011@hotmail.com (I. Bykov); bessmertny@itmo.ru (I. Bessmertny)

0000-0002-1970-5217 (S. Milantev); 0000-0002-1798-2843 (V. Olyunin); 0000-0001-9673-8349 (N. Milanteva); 0000-0002-8035-9800 (I. Bykov); 0000-0001-6711-6399 (I. Bessmertny)

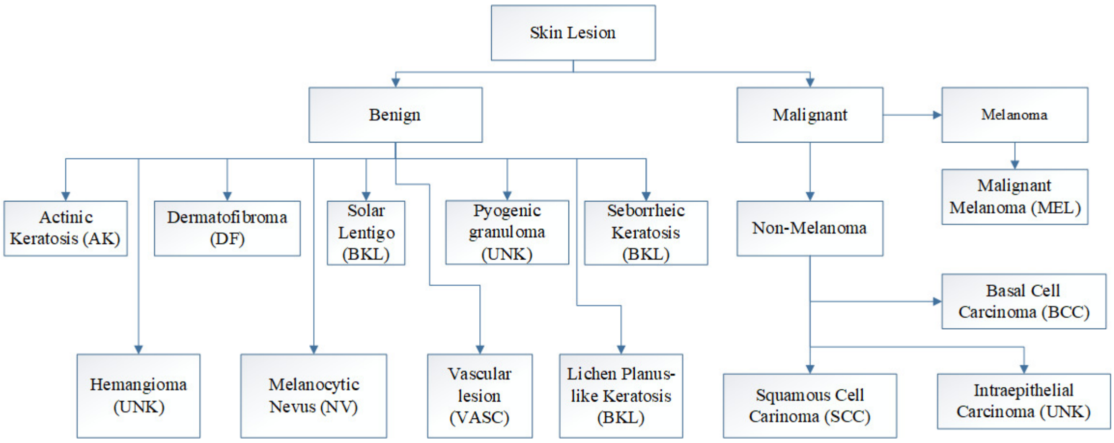

Figure 1: Skin Lesion Classification.

to expensive diagnostics, long individual patient analysis and inaccessible medical equipment which is not available in all medicine clinics.

In order to make expert knowledge more widely available, the International Skin Imaging Collaboration (ISIC) has prepared the ISIC Dataset, an international repository of dermoscopic images primarily used for technical research toward automated skin lesion analysis. ISIC hosted this dataset and requirements in ISIC Challenge. This requirements hide pitfalls that many skin lesion occur more frequently than others. It is known that a neural network must be trained on a balanced dataset for high-quality results and high accuracy of model. In this case formation of a representative dataset is a non-trivial statistical task [2].

The main goal of ISIC Challenge is to classify dermoscopic images among eight classes with unknown class (see Fig. 1).

Skin lesions are classified into the following groups:

- Melanoma (MEL) – dark colored areas with partial bleeding;
- Melanocytic Nevus (NV) – is a small convex area of brown color;
- Basal Cell Carcinoma (BCC) – is a black basal cell;
- Actinic Keratosis (AK) – is a patch with pronounced pigment, covered with scales;
- Benign Keratosis (BKL) – is a small dark pigmented spot with pronounced edges;
- Dermatofibroma (DF) – is a patch of various spectra of colors, but mostly brown;
- Vascular Lesion (VASC) – is a blood vessel growths;
- Squamous Cell Carcinoma (SCC) – is an ulcer covered with a scaly plaque;
- None of the others (UNK) – any other classes of skin lesions.

Unknown class emulates natural data that would have to be analyzed in a real medical clinic. Dermoscopy is a skin imaging method that has demonstrated improvement for diagnostics of skin lesion compared to unaided visual inspection.

ISIC challenge includes the patient metadata to further improve the diagnostic efficiency of the skin lesion problem. In addition, a large imbalance in dataset classes is still a major problem to be solved.

ISIC adds thousands of images every year to make this task more difficult. Each image is unique and may have different parameters (resolution, color spectrum, shooting quality, cropped fragments). There were used the methods for unbalanced dataset for skin lesions classification, including: data preparation, data augmentation, pre-trained on ImageNet state-of-the-art convolutional neural networks and ensemble strategy.

In this paper, there was considered the use of deep learning for skin lesions with advanced data preprocessing. Data preprocessing got an insufficient attention in state-of-the-art papers which describes this problem. In this paper, we rely heavily on the data preprocessing and data segmentation. We consistently applied data segmentation, black area removal, color change and data augmentation. Unused areas was removed using data preprocessing which has a positive impact on the accuracy of the model.

We used patient metadata with the images which are parallel included in the model using dense layers. Outliers and missing values were handled using median method.

We used fully connected dense layers. attach metadata to each model. The median method was applied for outliers or missing values. Experiments was found that using the median method for numeric features is better than skip invalid values. Label encoding was performed binary features and one-hot encoding was used for other non-binary features.

2. Materials and Methods

2.1. Datasets

The dataset consists of the main ISIC dataset and external data from SD-198, PH2, MED-NODE, Seven-Point and Light Field Image Dataset. Some of these datasets contain meta information (see Fig. 2 [3]).

It is known that some images have black areas, and metadata has missing values for some images, so data processing was applied to reduce possible errors.

In addition, additional images of other skin lesions (healthy skin, wart, miscellaneous, cysts, recurrent nevus, etc.) were included to unknown class. There was consider to increase the size of the unknown class in order to enlarge neural network's confidence in classifying main classes from unknowns.

ISIC Archive The dataset contains 25331 dermoscopic images. ISIC contains data from other datasets: HAM10000 (image size is 600x450) [4], BCN20000 (image size is 1024x1024) [5] and MSK (various image sizes) datasets. Also ISIC contains the meta information about patients' age (5 to 85 years), the anatomical site (skin lesion location) and the sex of the patient [1, 6, 7].

SD-198 The dataset is a skin lesion dataset contains 5944 clinical images of 13 skin lesions classes [8]. SD-198 is a private dataset available after signing some papers. SD-198 was captured with digital cameras and smartphones, which contains images of various size and color. This data was used for training for unknown class because dataset contains macroscopic images (images which taken without microscopic zoom).

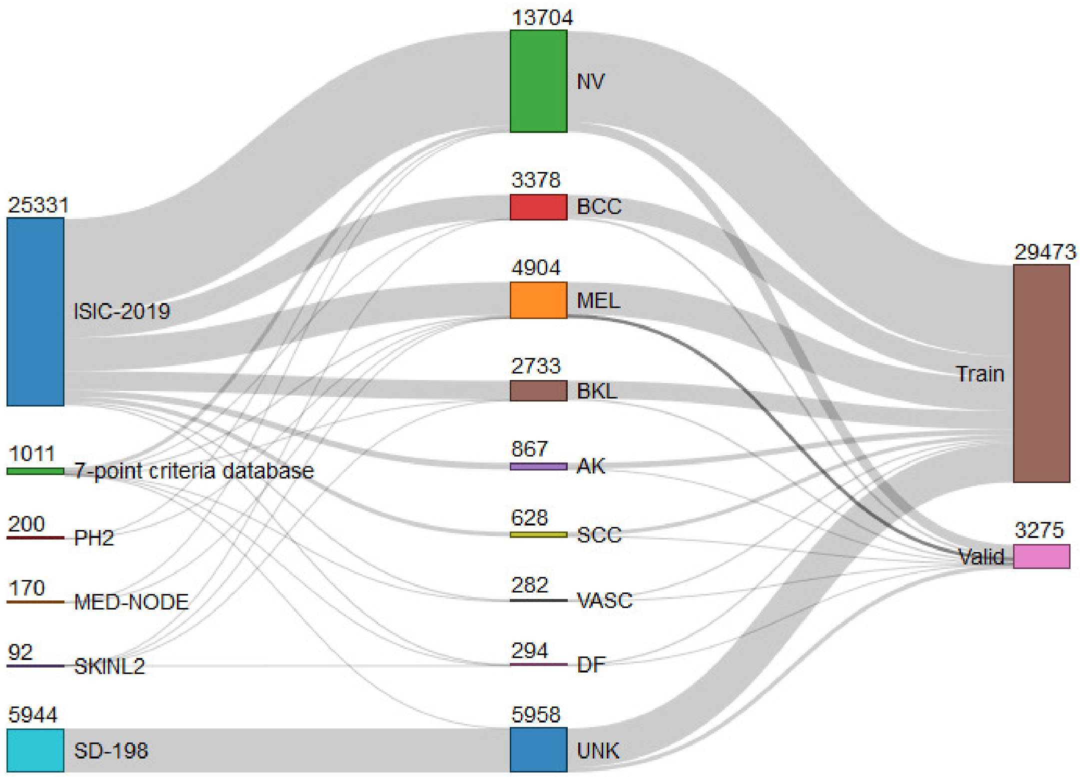

Figure 2: Skin Lesion Datasets.

PH2 The dataset is a clinical skin lesion dataset contains 200 dermoscopic images, including 80 common nevi, 80 atypical nevi, and 40 melanomas [9]. The dataset contains images of size 768x560 and structured meta information about patients' and skin lesion (pigment network, streaks, pigmentation, regression areas, dots, globules, color, blue-whitish veil, location, etc.) [10].

Seven-Point Criteria Evaluation Database The dataset is a clinical skin lesion dataset contains 1011 dermoscopic images of 7 skin lesions classes. We used two classes as unknown class which are really rare skin lesions. Seven-point is a private dataset available after signing some papers. The dataset contains images of size 768x512 and structured meta information about patients' and skin lesion (asymmetry, pigment network, dots, globules, streaks, regression areas, blue-whitish veil, color) [11].

Med-Node The dataset is a clinical skin lesion dataset contains 170 clinical images, including 100 nevi, and 70 melanomas [12].

Light Field Image Dataset of Skin Lesions The dataset is a clinical skin lesion dataset contains 92 clinical images, including melanomas, melanocytic nevus, basal cell carinoma, seborrheic keratosis, hemangioma, dermatofibroma and others. The dataset contains images of size 1920x1080 and structured meta information about patients' and skin lesion (gender, age, fototype, melanocytic) [13].

2.2. Data Preprocessing

First, black borders occuring in some images in the ISIC Dataset were cropped using ImageMagick image transformation [14]. Also the bottom strip containing information about the clinic and the image in the SD-198 Dataset.

Next, a 3% fuzz factor was applied to remove continuous black areas for each image.

Finally, Shades of Grey was used applying the expertise of Finlayson and Trezzi who investigated the performance of the illuminant estimation as a function of the Minkowski norm and there was found that the best results are obtained with p= 6 [15].

In addition, the meta information of different datasets was processed after merging. Metadata contains missing values or outlier values (for example, null values to indicate age). The median values were used for each of features to fill the non-valid data.

After this, there were chosen the label encoding for binary features (like sex and also the age of the patient) and the one-hot encoding for other features. Next, meta-features was performed using PCA method (principal component analysis) for reduce the dimensionality. This strategy allows to use even invalid data in training and validation.

2.3. Deep Learning Models

After preprocessing, R2U-Net convolutional neural network to each image was applied. The segmentation removes unnecessary parts of the image and returns accurate extraction of the skin lesion region. R2U-Net neural network allows the classification model focusing on the main features while improving the accuracy and sensitivity (see Fig. 3).

2.3.1. Model Learning

Image segmentation was added to the model for better identify patterns. The R2U-Net model for segmentation was trained on preprocessed data. The resulting R2U-Net combines with the classification convolutional neural network.

There was used a random transformation strategy which shows best results for ImageNet competition [16]. The resulting image is randomly resized and scaled from the segmented image.

In the next step data augmentation was performed. The following transformations were applied with a probability of 0.5:

- a random rotation between -45 and 45 degrees
- a random shear
- a random horizontal and vertically flip or rotated by 90 degrees
- a random zoom between 1.0 (original zoom) and 1.12

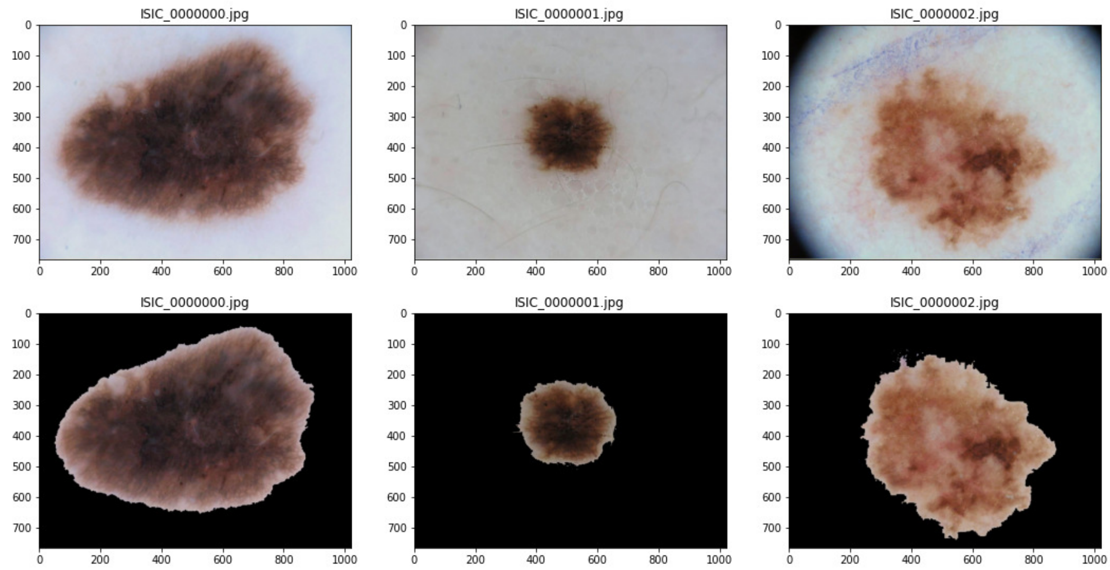

Figure 3: Skin Lesion Segmentation.

- a random brightness and contrast change no more than 10% from the original with probability of 0.15
- a random single cutout of size 16x16.

A random single cutout of size 16x16 shows excellent results for the 10 categories classification (generally, the more categories, the less cutout) [17].

Thereafter there was trained only classification CNN model on dermoscopic images. All models for 120 epochs were trained using Adam optimizer, normalized weighted cross-entropy loss function and five-fold cross-validation. The model was evaluated and saved every 5 epochs. There was use a 10% train-valid split (29473 images of train set and 3275 images of validation set).

Finally, the weights of our convolutional neural network were freezed and the dense neural network was trained to work with it. The dense neural network was trained for 60-80 epochs and scheduler with 0.1 gamma each 10 epoch. The preprocessed meta information feed to a two-layer neural network with with 256 neurons each. Each layer contains batch normalization, a ReLU activation and dropout with probability of 0.25. Features vector of neural network combines with fully-connected layer of classification convolutional neural network. The resulting features vector feed to a single-layer neural network with 1024 neurons and connect to classification layer [18]. The various pre-trained state-of-art deep learning models were used for classification dermoscopic images. The proposed model was combined with a dense neural network to analyze patients' meta data (see Fig. 4).

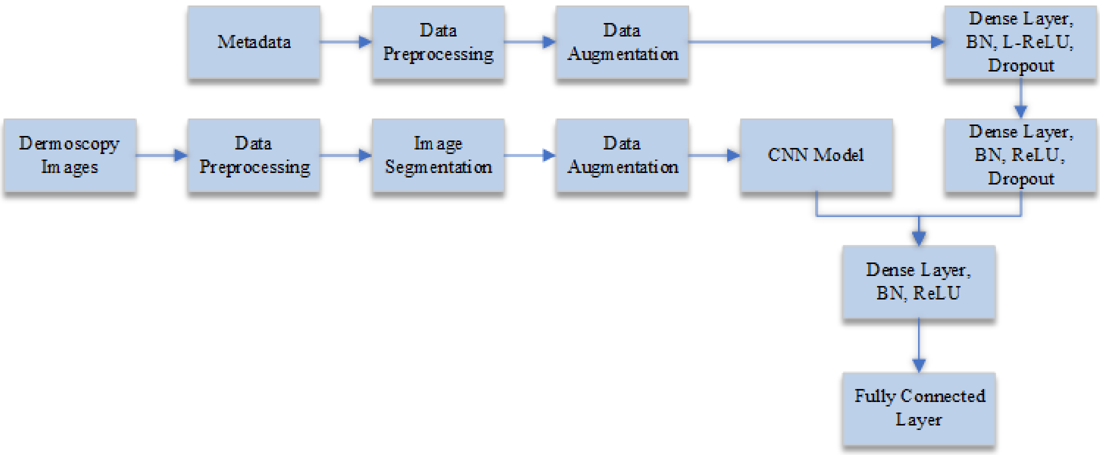

Figure 4: Skin Lesion Classification Model.

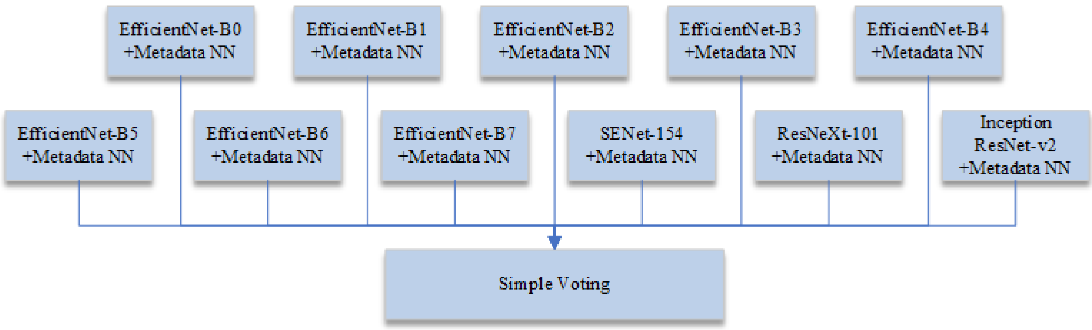

Figure 5: Skin Lesion Classification Ensemble.

2.3.2. Architecture

R2U-Net architecture with convolutional encoding and decoding units using recurrent convolutional layers based on U-Net architecture was used for image segmentation [19]. The residual units are used with RCL for R2U-Net architecture. This architecture is state-of-art in medical image segmentation.

The using strategy mainly relies on the EfficientNet model [20]. These models show good results on ImageNet and re-training. The EfficientNet family contains eight different models which differ only in depth, width and resolution scaling. There were used all types of this family: from EfficientNet-B0 (224 x 224) to EfficientNet-B7 (600 x 600) (see Fig. 5).

In the final ensemble, a SENet154 (320x320), ResNeXt-101 32x4d (320x320) and Inception-ResNet-v2 (320x320) were also included to the ensemble based on the Jie Hu and Saining Xie papers [21, 22]. Each model creates 25 predictions for each image depending on random transformation strategy and returns averaged prediction.

Table 1
The comparison of the suggested model with the state-of-the-art models in terms of sensitivity.

Model	Sensitivity	Specificity
Skin Lesion Classification Using Ensembles of Multi-Resolution EfficientNets with Meta Data [18]	$74.2\% \pm 1.1\%$	$97.3\% \pm 1.1\%$
MelaNet: A Deep Dense Attention Network for Melanoma Detection in Dermoscopy Images [23]	$78.1\% \pm 2\%$	$96.2\% \pm 2\%$
Skin Lesion Classification with Deep Learning Ensembles in ISIC 2020 [24]	80%	95%
Embedding vectors and Ensemble models [25]	77.4%	99.1%

3. Related work

Researchers in state-of-the-art papers mainly relied on data augmentation and CNNs. German researchers who won two ISIC competition at Kaggle used data augmentation and ensemble of convolutional neural networks (EfficientNet B0-B7) [18]. Dense layers were added to the ensemble for metadata processing.

Chinese researchers used modification of DenseNet. The CNN model formulated by two 169-layer convolutional neural networks. Data were applied only horizontal flipping and rotation between -30 and 30 degrees with random strategy [23].

Turkish researchers also used data augmentation and an ensemble of Xception, Inception-ResNet-V2 and NasNetLarge. Focal loss function used for Xception with different gamma parameter (γ between 1 and 4) [24]. Researchers from AirDoctor team used a similar approach [25].

Data preprocessing was absent or received little attention in each of the above works. The results of the above researchers are shown in table 1.

4. Results

The described approach has obtained fairly high sensitivity rates, which are used in many state-of-art projects for the analysis of skin lesions. The provided test data was used for this results. The results of the ensemble are shown in Table 2.

We found that R2U-Net perform well for skin lesion segmentation. Preprocessed and segmentated data was positive impact for skin lesion classification. The EfficientNets ensemble made it possible to cover various input resolutions. In general, the metadata preprocessing gave a better result.

5. Conclusion

In this paper, there was considered the use of deep learning for skin lesions with advanced data preprocessing. We used R2U-Net for segmentation and EfficientNet-B0-B7, SENet-154,

Table 2
Ensemble result for each class.

Class	Accuracy	Precision	Recall	F1-Score
Melanoma (MEL)	94.2%	0.9	0.79	0.84
Melanocytic Nevus (NV)	93.1%	0.89	0.97	0.93
Basal Cell Carcinoma (BCC)	97.2%	0.91	0.87	0.89
Actinic Keratosis (AK)	98.7%	0.81	0.81	0.81
Benign Keratosis (BKL)	94.5%	0.72	0.72	0.72
Dermatofibroma (DF)	99.5%	0.73	0.87	0.79
Vascular Lesion (VASC)	98.8%	0.95	0.75	0.84
Squamous Cell Carcinoma (SCC)	98.9%	0.87	0.73	0.79
Mean	**96.86%**	**0.85**	**0.82**	**0.83**

ResNeXt-101 32x4d and Inception-ResNet-v2 to classify skin lesion. Preprocessed metadata was incorporated into our ensemble with a dense layers. We experimented with this approach and the results of this experiments were spectacular. All models with an ensemble were combined. The best ensemble score can achieve a sensitivity of 81.5% and specificity of 97.7%

6. Future Work

It is planned to conduct experiments with various models of convolutional neural networks using the ensemble strategy, as well as add a parallel ensemble to the structure, which will receive non-segmented images as input.

Experiments are currently underway on SE-ResNet-152, SE-ResNeXt-50 (32x4d), SE-ResNeXt-101 (32x 4d) and with the removal of the black background after segmentation. The gained experience is going to be applied in ISIC Skin Lesion Challenge 2021.

References

[1] N. Codella, et al., Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC), Preprint at arXiv:1902.03368 (2019) (2019).

[2] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer (2019).

[3] Dataset Chart, 2020. URL: http://j5t.github.io/dataset-chart/.

[4] P. Tschandl, C. Rosendahl, H. Kittler, The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, Scientific Data 5 (2018). doi:10.1038/sdata.2018.161.

[5] M. Combalia, et al., BCN20000: Dermoscopic lesions in the wild, Preprint at arXiv: 1908.02288 (2019).

[6] N. Codella, 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), IEEE: 15th International Symposium on Biomedical Imaging (ISBI 2018) (2018) 168–172.
[7] H. Trong, Y. Kazuma, Resolving class imbalance in object detection with weighted cross entropy losses, Preprint at arXiv:2006.01413 (2020).
[8] X. Sun, J. Yang, M. Sun, K. Wang, A benchmark for automatic visual classification of clinical skin disease images, Springer: European Conference on Computer Vision (2016) 206–222.
[9] T. Mendonça, et al., A dermoscopic image database for research and benchmarking, 35th International Conference of the IEEE Engineering in Medicine and Biology Society (2013) 5437–5440.
[10] C. Barata, M. Ruela, et al., Two Systems for the Detection of Melanomas in Dermoscopy Images using Texture and Color Features, IEEE Systems Journal 8 (2013) 965–979.
[11] J. Kawahara, S. Daneshvar, G. Argenziano, G. Hamarneh, Seven-Point Checklist and Skin Lesion Classification Using Multitask Multimodal Neural Nets, IEEE Journal of Biomedical and Health Informatics 23 (2019) 538–546.
[12] I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman, N. Petkov, MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images, Expert Systems with Applications 42 (2015) 6578–6585. doi:10.1016/j.eswa.2015.04.034.
[13] S. M. M. De Faria, et al., Light Field Image Dataset of Skin Lesions, IEEE 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2019) 3905–3908. doi:10.1109/EMBC.2019.8856578.
[14] ImageMagick - Convert, Edit, or Compose Bitmap Images, 2020. URL: http://www.imagemagick.org.
[15] G. Finlayson, E. Trezzi, Shades of Gray and Colour Constancy, IST/SID Twelfth Color Imaging Conference (2004) 37–41.
[16] O. Russakovsky, et al., ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision (2015) 211–252. doi:10.1007/s11263-015-0816-y.
[17] T. DeVries, G. W. Taylor, Improved Regularization of Convolutional Neural Networks with Cutout, Preprint at arXiv:1708.04552 (2017).
[18] N. Gessert, et al., Skin Lesion Classification Using Ensembles of Multi-Resolution EfficientNets with Meta Data, Preprint at arXiv:1910.03910 (2019).
[19] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, V. K. Asari, Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation, NAECON 2018 - IEEE National Aerospace and Electronics Conference (2018) 228–233. doi:10.1109/NAECON.2018.8556686.
[20] M. Tan, Q. V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, Preprint at arXiv:1905.11946 (2019).
[21] J. Hu, L. Shen, S. Albanie, G. Sun, E. Wu, Squeeze-and-Excitation Networks, IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018) 7132–7141. doi:10.1109/CVPR.2018.00745.
[22] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated Residual Transformations for Deep Neural Networks, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017) 5987–5995. doi:10.1109/CVPR.2017.634.
[23] P. Zhang, MelaNet: A Deep Dense Attention Network for Melanoma Detection in Der-

moscopy Images, ISIC Leaderboards (2020).
[24] S. A. Ahmed, B. Yanikoglu, E. Aptoula, O. Goksu, Skin Lesion Classification with Deep Learning Ensembles in ISIC 2020, ISIC Leaderboards (2020).
[25] Airdoctor Team, Manuscript for ISIC Challenge 2020, ISIC Leaderboards (2020).