CVPR
LONG BEACH
CALIFORNIA
June 16-20, 2019

Melanoma Thickness Prediction Based on Convolutional Neural Network with VGG-19 Model Transfer Learning

Joanna Jaworek-Korjakowska

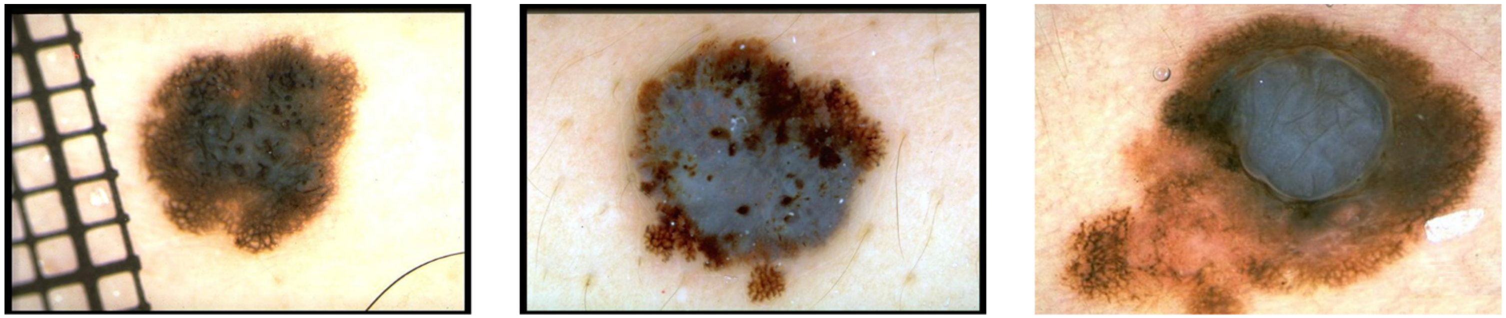
Paweł Kłeczek, Marek Gorgoń

AGH University of Science and Technhology, Krakow, Poland
Computer Vision Laboratory

Motivation

- **Thickness is one of the most important factor in melanoma prognosis**
- Is used to:
  - establish the size of the surgical margin,
  - select patients for sentinel lymph node biopsy.

Little work has concentrated on the evaluation of melanoma thickness both from the clinical as well as computer-aided diagnostic side.



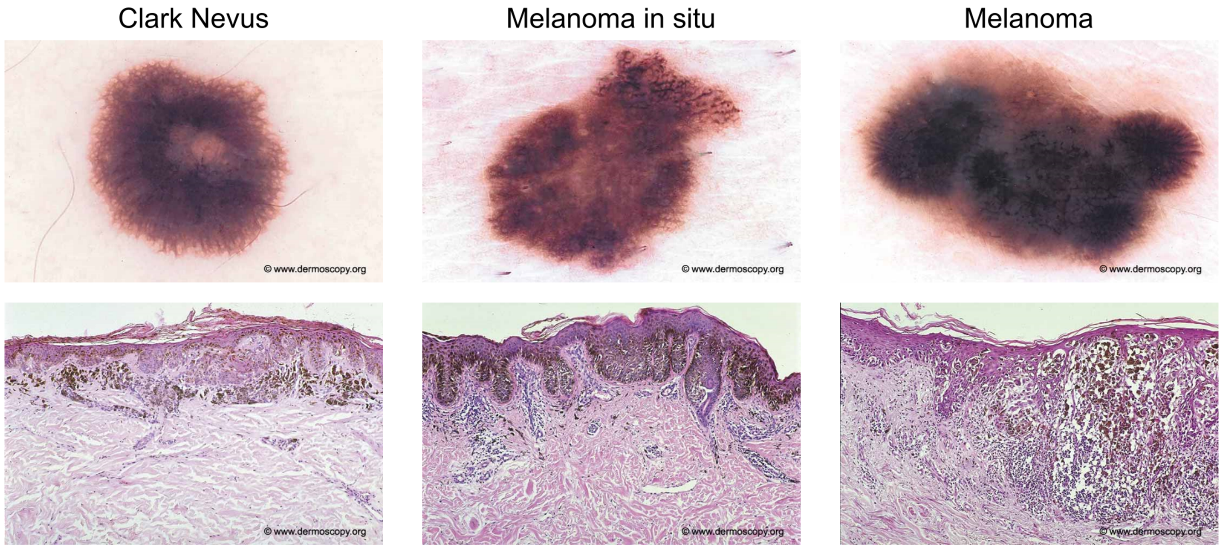
a) b) c)

Dermoscopy images of various melanoma thicknesses:
a) thin melanoma, b) intermediate melanoma, c) thick melanoma

Research goal

The main goal - classify melanocytic lesions into three classes indicating the thickness of the diagnosed skin lesion:

- **thin melanoma** including melanoma in situ with thickness less than 0.75 mm,
- **intermediate melanoma** with thickness between 0.76-1.5 mm,
- **thick melanoma** with thickness greater that 1.5 mm.



Dermoscopy and hematoxylin-eosin-stained histopathology images.
Specific underlying histopathologic correlates as well as the increasing thickness can be observed.

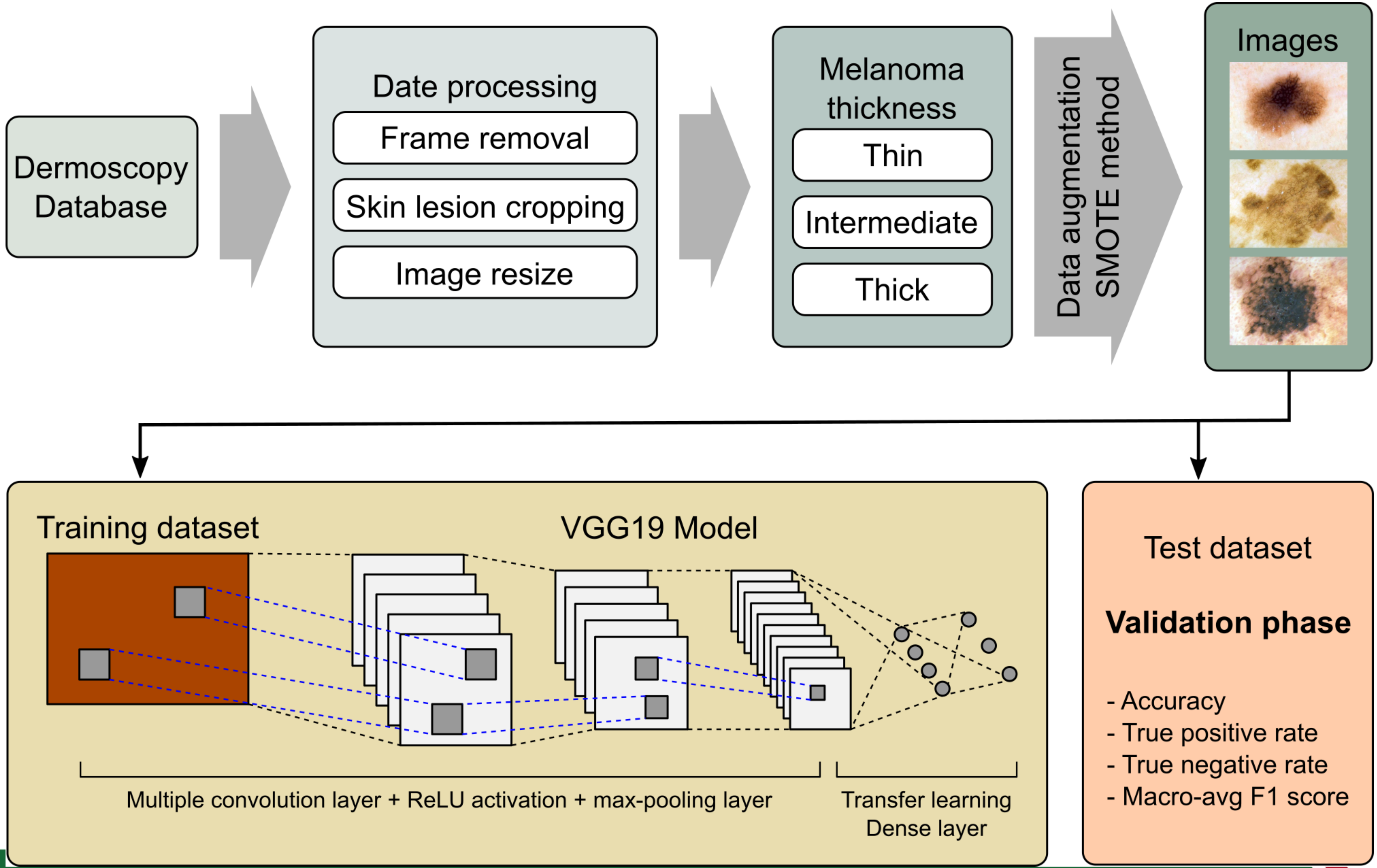
Clinical importance

The possibility of preoperative evaluation of melanoma thickness has many advantages including:

- better organization of surgical and diagnostic priorities based on distinction of lesions at high and low risk of progression,
- excision with sufficient surgical margins at the first operation,
- excision and sentinel lymph node biopsy (if needed) in a single operation, saving time and costs.

It has been reported that 13% cases in SEER Program have unknown thickness. Moreover, unknown thickness cases had a significantly increased risk of death due to melanoma than known thickness cases with an increasing trend over time.

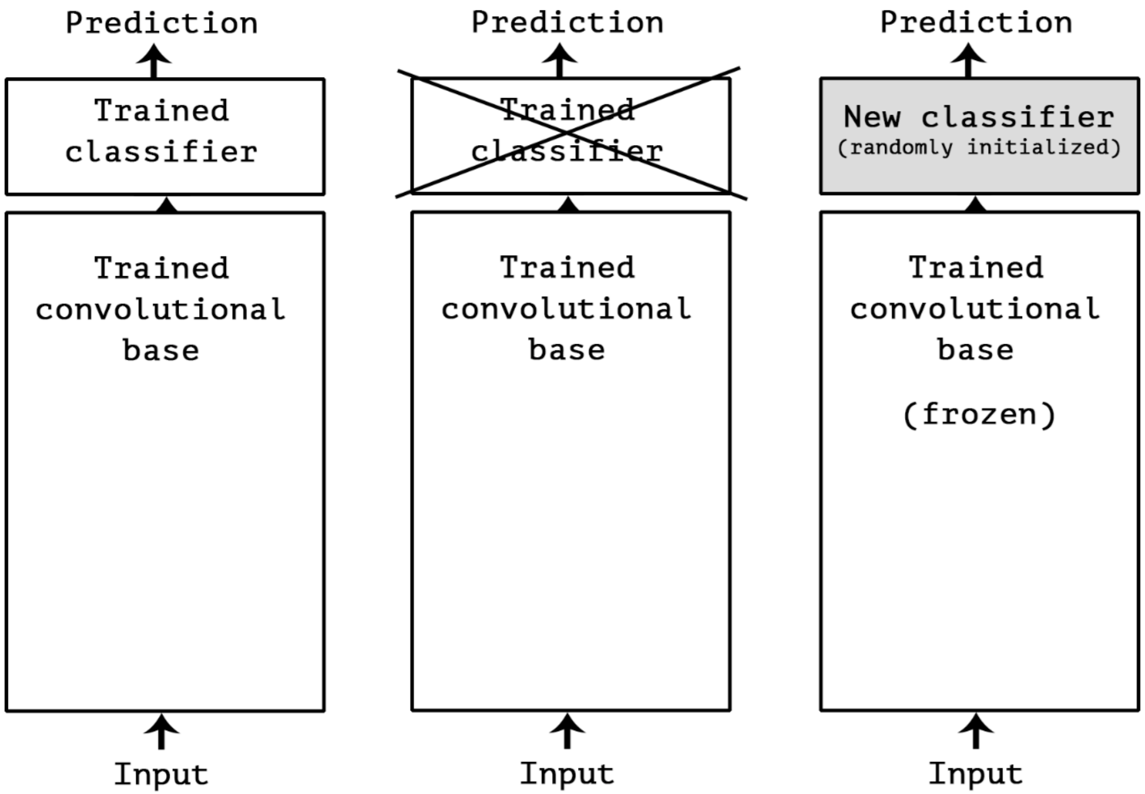
Methods



Dermoscopy Database
Date processing
Frame removal
Skin lesion cropping
Image resize
Melanoma thickness
Thin
Intermediate
Thick
Data augmentation
SMOTE method
Images
Training dataset
VGG19 Model
Multiple convolution layer + ReLU activation + max-pooling layer
Transfer learning
Dense layer
Test dataset
Validation phase
- Accuracy
- True positive rate
- True negative rate
- Macro-avg F1 score

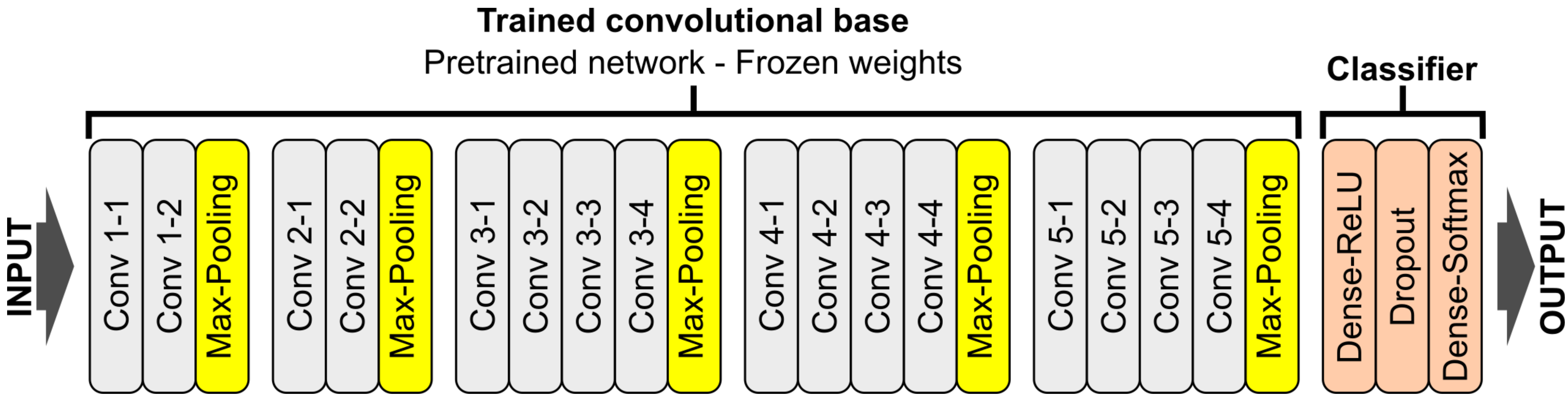
Transfer learning

- A common and highly effective approach to deep learning on small image datasets is to use a pre-trained network.
- A pre-trained network is a saved network that was previously trained on a large dataset, typically on a large-scale image-classification task.
- Its features can prove useful for many different computer-vision problems, even though these new problems may involve completely different classes than those of the original task.



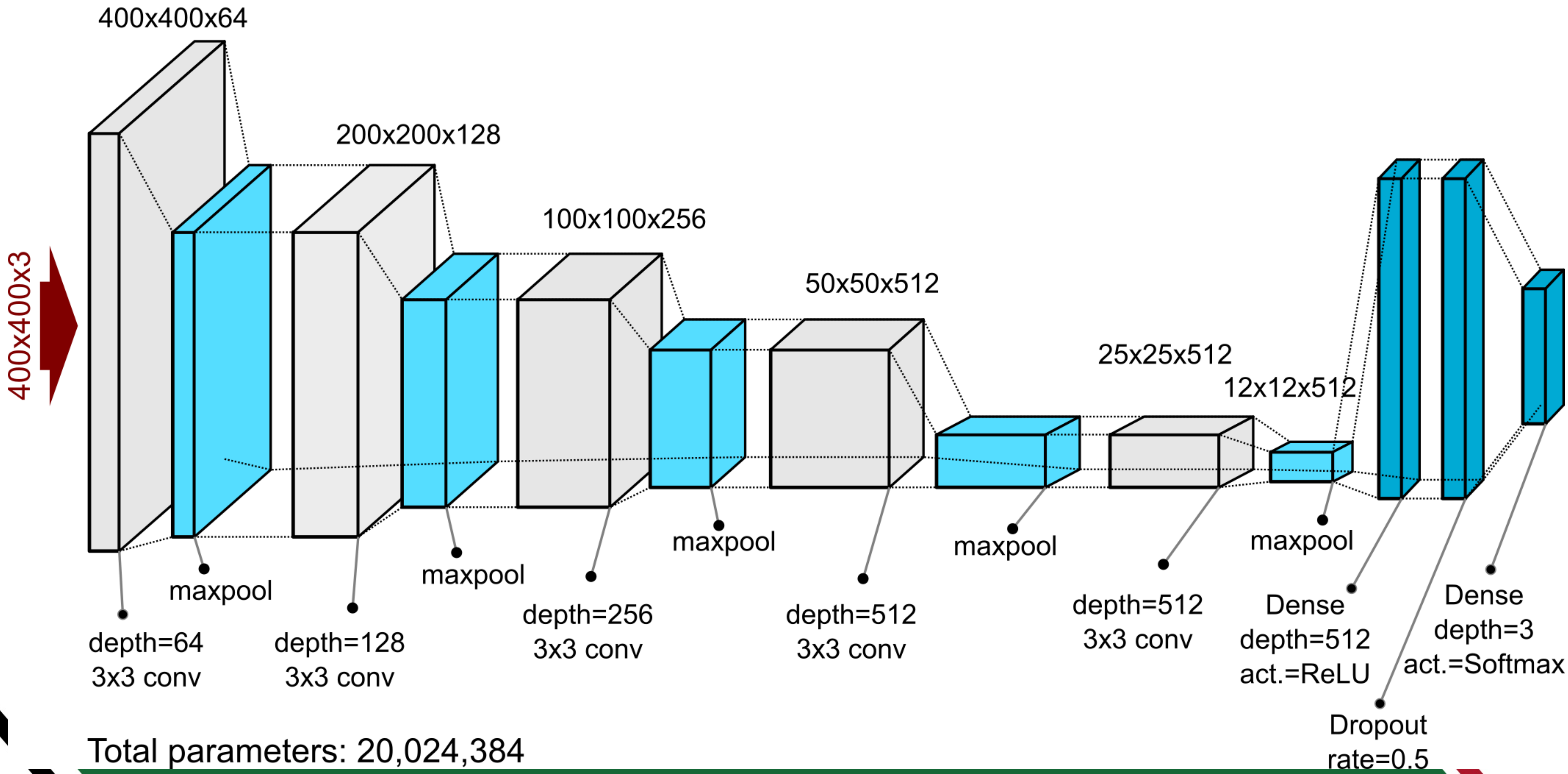
VGG-19 model

- Developed by Karen Simonyan and Andrew Zisserman (*Visual Geometry Group*) in 2014,
- VGG-19 model has roughly **143 million parameters**, where the parameters are learned from the ImageNet dataset containing **1.2 million general object images** of **1,000 different object categories** for training



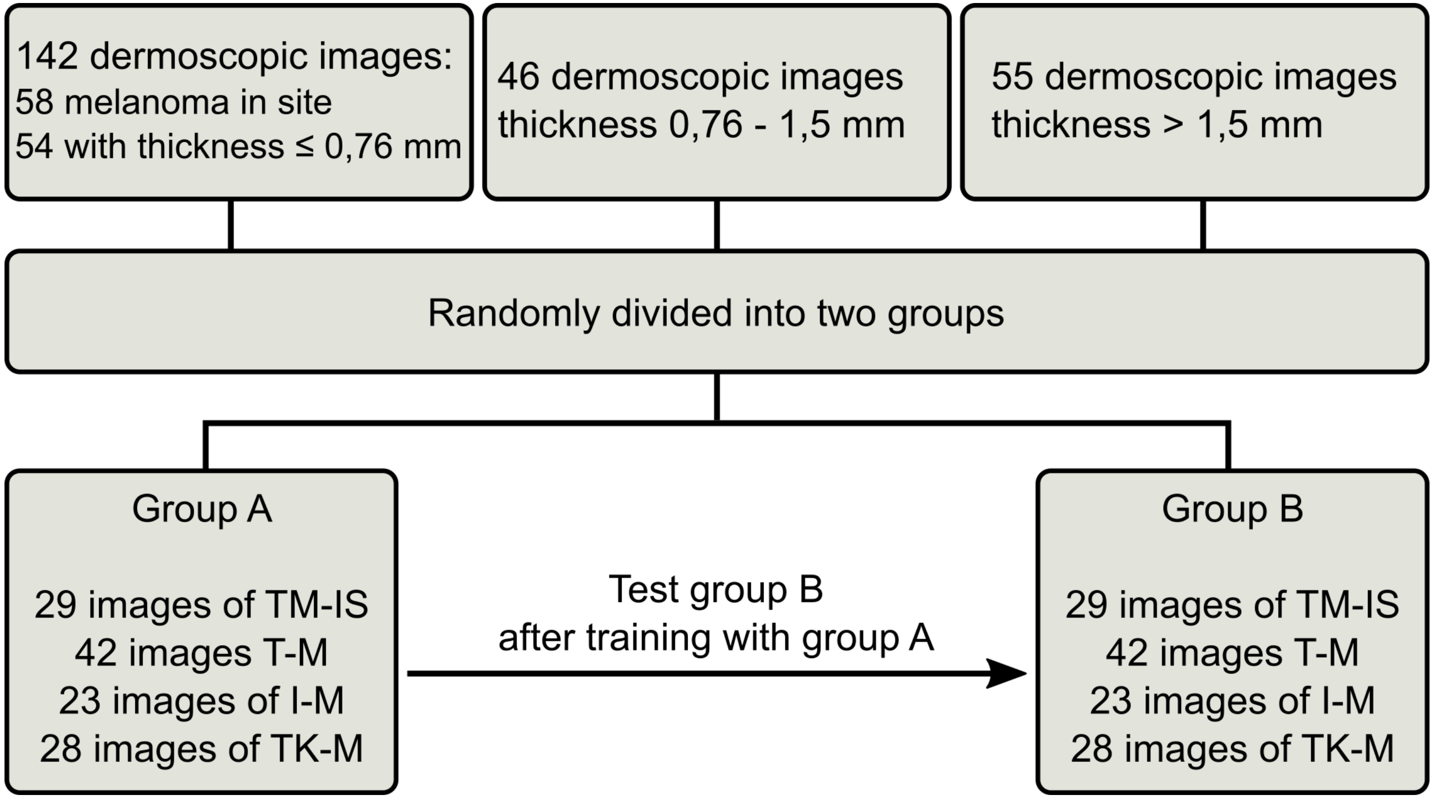
VGG-19 architecture

Illustration of the personalized network architecture of VGG-19 model.



Total parameters: 20,024,384

Image dataset



- With the binary masks we cropp the skin lesion with a tight margin (5 px) (ISIC Archive).
- Resize it to 400 × 400 px
- Handle the class imbalance by using Synthetic Minority Oversampling TEchnique (SMOTE) to generate synthetic samples

Use-cases and implementation

Main parameters:

- 50 epochs
- Batch size of 124
- As we are dealing with multi-class classification problem the

  categorical cross-entropy loss function,

  also called Softmax Loss, has been applied.
- Classification part includes densely-connected classifier

  and dropout layer (regularization)
- Adam optimization algorithm

  computationally efficient, invariant to diagonal rescale of the gradient, appropriate for problems with very noisy gradients.

Statistical analysis

To evaluate multi-class classification problem a commonly used measure is the **macro-averaged F1 (F-measure) score** which is defined as:

$$F = \frac{(\beta^2 + 1)PR}{\beta^2 P + R} \quad (8)$$

where $\beta = 1$. We specify precision P_{macro} and recall R_{macro} as follows:

$$P_{macro} = \frac{1}{|C|} \sum_{i=1}^{|C|} \frac{TP_i}{TP_i + FP_i}, \quad (9)$$

$$R_{macro} = \frac{1}{|C|} \sum_{i=1}^{|C|} \frac{TP_i}{TP_i + FN_i} \quad (10)$$

The average F1-score is 83.4% for classifying the melanoma thickness into three different categories.

Statistical analysis

Classes	TPR [%]	TNR [%]	ACC [%]
Thin	84.5	90.9	86.9
Intermediate	78.3	87.23	85.5
Thick	78.6	92.86	89.3
Average	80.5	90.3	87.2

Table 1. The performance of the melanoma thickness classification model.

Problem	Binary [ACC %]	Multi-class [ACC %]
Rubegni *et al.* [18]	86.5	—
Sáez *et al.* [20]	77.6	68.4
Our method	—	**87.2**

Table 2. Comparison with other melanoma thickness prediction methods.

Summary

AGH

The novelty of this work can be summarized as:

- we present a deep learning based solution for **the preoperative melanoma thickness prediction** into three depth classes,

- we propose **a convolutional neural network architecture with transfer learning from the VGG-19 pretrained model and an adjusted densely-connected classifier**,

- our method enables the adaptation of the **VGG-19 model** to the **melanoma thickness prediction** with a **limited data amount** based on **data augmentation** and **transfer learning method**.


Thank you!