MEDICAL UNIVERSITY OF VIENNA

Loughborough University

KTH VETENSKAP OCH KONST

TISSUEGNOSTICS PRECISION THAT INSPIRES

Investigating and Exploiting Image Resolution for Transfer Learning-based Skin Lesion Classification

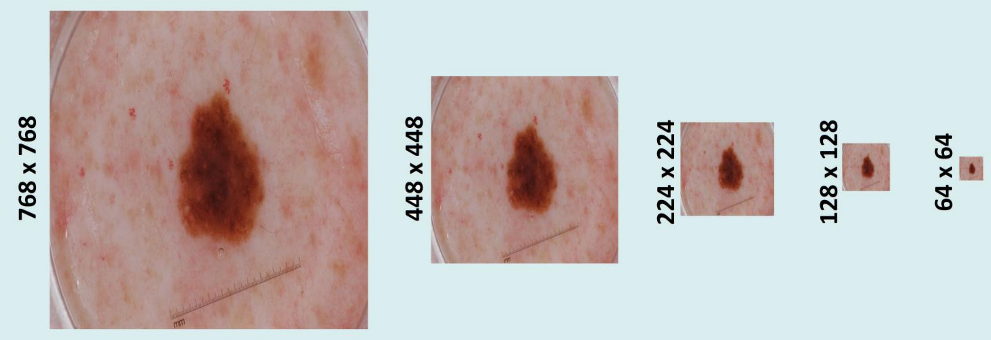
Mahbod A [1], Schaefer G [2], Wang C [3], Ecker R [4], Dorffner G [5], Ellinger I [1]

[1] Institute for Pathophysiology and Allergy Research, Medical University of Vienna, Vienna, Austria
[2] Department of Computer Science, Loughborough University, U.K
[3] Department of Biomedical Engineering and Health System, Division of Biomedical Imaging, KTH Royal Institute of Technology, Stockholm, Sweden
[4] Department of Research and Development, TissueGnostics GmbH, Vienna, Austria
[5] Section for Artificial Intelligence and Decision Support, Medical University of Vienna, Vienna, Austria

Research questions

Using pre-trained convolutional neural networks (CNN) for skin lesion classification requires down-sampling, but:

❖ What is the proper down-sampling factor?



❖ Is it useful to exploit images with different sizes for fine-tuning?

Key findings

- ❖ Using resized images with a size of larger than 128 x 128 → comparable results (89.42% to 92.37% for ResNet-18)
- ❖ Using resized images with a size of 64 x 64 → significant drop in performance (84.21% for ResNet-18)
- ❖ Increasing size → slight improvement in performance (from 89.31% to 91.44% on average for all networks)
- ❖ Fusing is the best → using multi fine-tuned CNNs and using images with different resolutions for training (92.86%)

Method

❖ Datasets (contained three skin lesion types)

Malignant Melanoma (MM) Seborrheic Keratosis (SK) Benign nevi (BN)

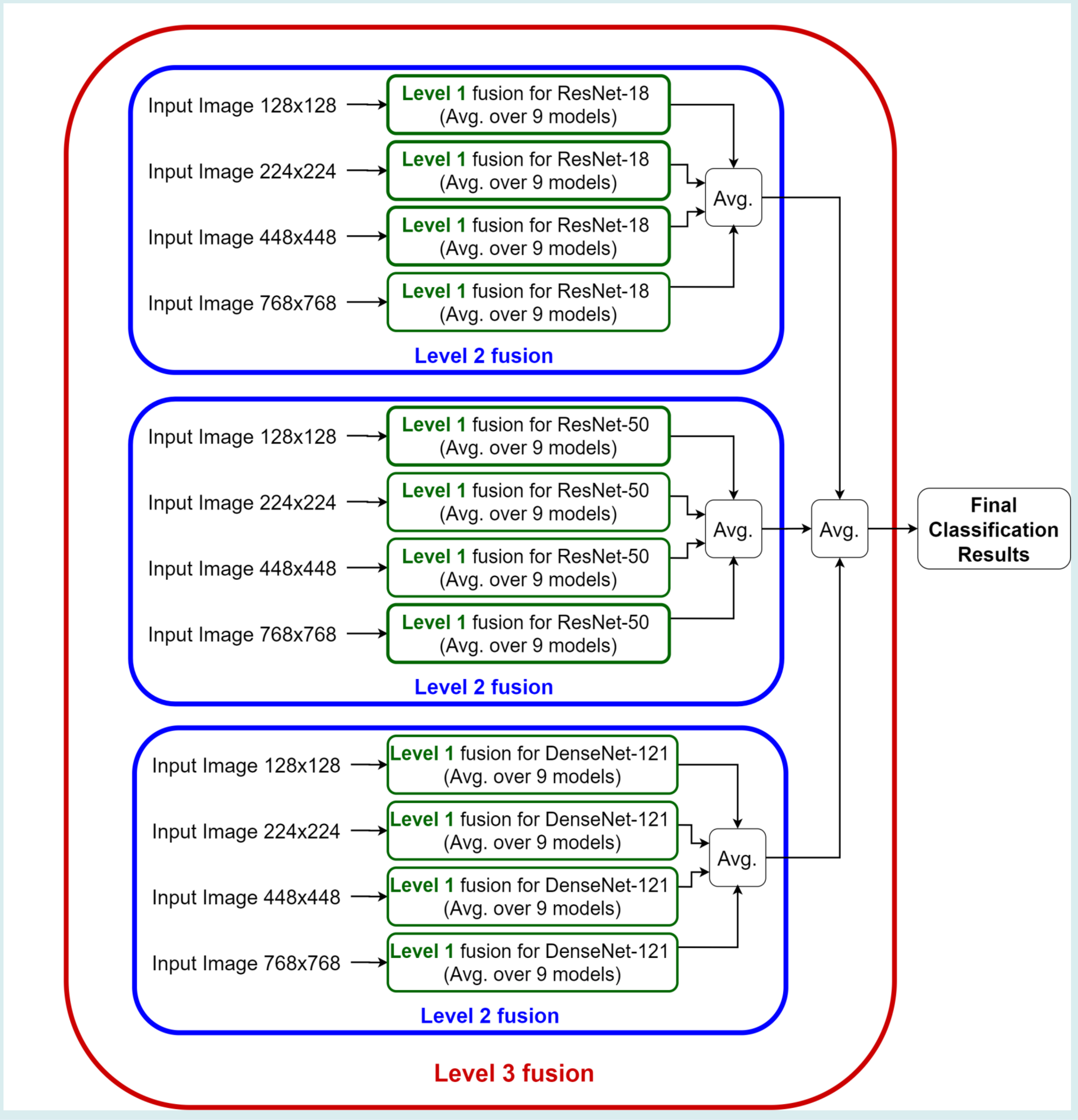
- Subset of ISIC archive [1] for training and validation data
  - ❑ **Training Data:** 2,187 images including training, validation and test images of ISIC 2016 competition as well as training and validation set of ISIC 2017 competition
  - ❑ **Test Data:** 600 test images from ISIC 2017 competition

❖ Reporting results for MM vs. all and SK vs. all classifications

❖ Pre-processing
- Gray world color constancy normalization
- ImageNet mean subtraction
- Resizing (ranging from 64 x 64 to 768 x 768)
- Data augmentation by rotation and flipping

❖ Fine-tuning 3 pre-trained CNNs, namely ResNet-18 [2], ResNet-50 [2] and DenseNet-121 [3]

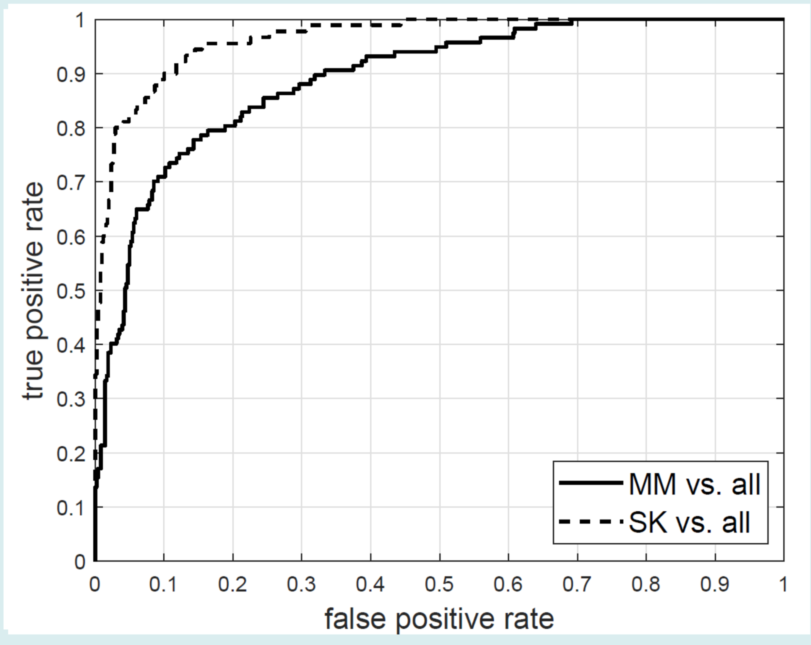
❖ Three-level fusion approach



Results

❖ Size effect on the performance for ResNet-18 (level one fusion)

Size	MM AUC (%)	SK AUC (%)	Avg. AUC (%)
64 x 64	78.86	89.55	84.21
128 x 128	85.46	93.39	89.42
224 x 224	85.37	93.81	89.59
448 x 448	89.20	95.54	92.37
768 x 768	88.89	95.85	92.37



❖ Effect of level two and level three fusion

Network	Sizes	MM AUC (%)	SK AUC (%)	Avg. AUC (%)
Fusion of ResNet-18	all	89.12	96.26	92.69
Fusion of ResNet-50	all	88.50	96.03	92.27
Fusion of DenseNet-121	all	87.69	95.77	91.73
Three-level fusion	**all**	**89.16**	**96.57**	**92.86**

❖ Comparison to the state-of-the-art methods

Network	Sizes	MM AUC (%)	SK AUC (%)	Avg. AUC (%)
Matsunaga *et al.* [4]	NA	86.8	95.3	91.1
Mahbod *et al.* [5]	224	87.3	95.5	91.4
Zhang *et al.* [6]	224	87.5	95.8	91.7
Yan *et al.* [7]	256	88.3	NA	NA
Three-level fusion	**all**	**89.16**	**96.57**	**92.86**

References:
[1] https://isic-archive.com/
[2] He et al: Deep Residual Learning for Image Recognition, CVPR, 2016
[3] Simonyan et al: Densely Connected Convolutional Networks, CVPR, 2017
[4] Matsunaga et al: Image classification of melanoma, nevus and seborrheic keratosis by deep neural network ensemble, arXiv, 2017
[5] Mahbod et al: Fusing Fine-Tuned Deep Features for Skin Lesion Classification, CMIG, 2019
[6] Zhang et al: Attention residual learning for skin lesion classification, IEEE TMI, 2019
[7] Yan et al: Melanoma recognition via visual attention, IPMI, 2019

Acknowledgement

Funded by the Horizon 2020 Framework Programme of the European Union No. 675228

amirreza.mahbod@meduniwien.ac.at