Breast Cancer Histological Image Classification Using Fine-Tuned Deep Network Fusion

Mahbod A[1,2], Ellinger I [1], Ecker R [2] , Smedby Ö [3], Wang C [3]

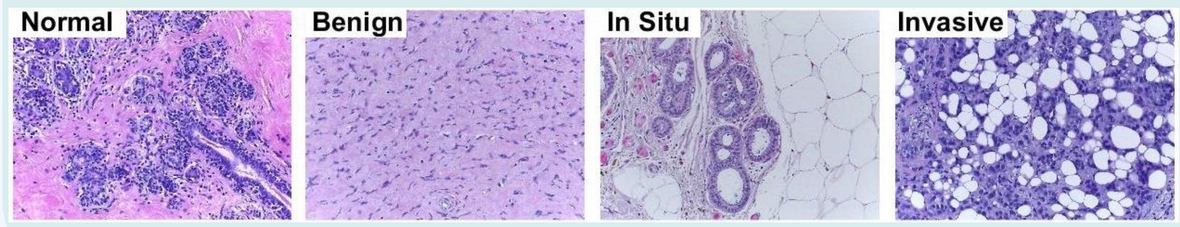
[1] Institute of Pathophysiology and Allergy Research, Medical University of Vienna, Vienna, Austria
[2] Department of Research and Development, TissueGnostics GmbH, Vienna, Austria
[3] Department of Biomedical Engineering and Health System, Division of Biomedical Imaging, KTH Royal Institute of Technology, Stockholm, Sweden

Motivation

- ❖ Breast cancer is the most commonly diagnosed cancer type & 2nd leading cause of cancer death among women
- ❖ Gold standard for diagnosis is examination of the hematoxylin and eosin (H&E) stained biopsy sections
- ❖ Using semi or fully automatic image analysis methods to help pathologists and partially substitute manual analysis is beneficial
- ❖ Deep learning-based approaches are the promising solutions

Aim

Breast cancer H&E image classification to 4 classes using fusion of fine-tuned convolutional neural networks (CNN)



Images are adapted from ICIAR 2018 grand challenge on breast cancer histology classification

Method

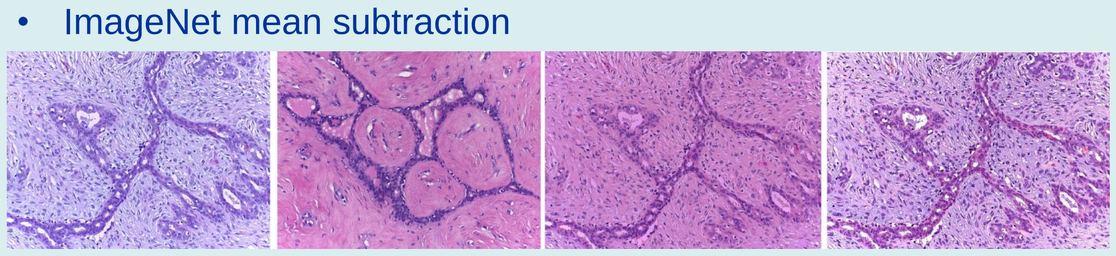
❖ **Datasets**

- BioImaging 2015 challenge dataset (BI)
  - ❑ 249 training images and 36 test images (20 initial + 20 extended cases)
- ICIAR 2018 grand challenge dataset (ICIAR)
  - ❑ 400 training images and 100 test images
- Challenge links:

ICIAR BI

❖ **Preprocessing**

- Normalization with RGB histogram matching and Macenko et al. (2009) [1] normalization method
- Background pixels remained unchanged (background detection with Otsu algorithm)
- ImageNet mean subtraction

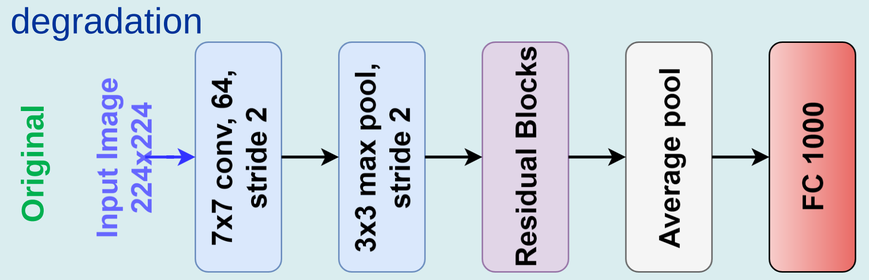


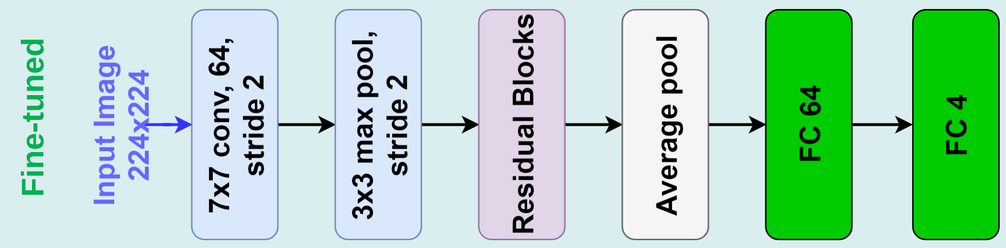
Input Image Ref Image RGB matching Macenko et al.

- Resizing (224 x 224)

❖ **CNN fine-tuning**

- ResNet-50 and ResNet-101 (2016) [2]: well-established pre-trained networks.
- Very deep networks containing residual blocks (16 for ResNet-50 and 33 for ResNet-101) with shortcut connections to prevent degradation





❖ **Implementation details:**

- Randomize initialization of the new fully connected (FC) layers
- Learning rate (LR) for the new layers 10x bigger compared to the others
- Stochastic gradient descent with momentum training (initial LR set to 0.001)
- 25 epoch total training
- Data augmentation (flip, rotate)

❖ **Fusion:**

- Taking the average over the fine-tuned networks

Results

❖ Effect of normalization on the BI dataset (ResNet-50)

Normalization	Initial acc (%)	Extended acc (%)	Overall acc (%)
RGB histogram matching	95.00	87.50	91.67
Macenko et al. (2009) [1]	90.00	81.25	86.11
No normalization	90.00	75.00	83.33

❖ Effect of normalization on the BI dataset (ResNet-101)

Normalization	Initial acc (%)	Extended acc (%)	Overall acc (%)
RGB histogram matching	95.00	75.00	86.11
Macenko et al. (2009) [1]	85.00	81.25	83.33
No normalization	80.00	75.00	77.78

❖ Comparative results

Network	Overall acc on BI dataset (%)	Overall acc on ICIAR dataset (%)
ResNet-50	91.67	87.75
ResNet-101	86.11	88.00
Fusion	97.22	88.50
Araujo et al. method (2017) [3]	77.80	N/A

Conclusion

A fully automatic method is proposed for breast cancer histological image classification based on fusion of fine-tuned deep networks applied on RGB histogram matched images which delivers excellent performance for two datasets used in this work.

References

[1] Macenko et al: A Method for Normalizing Histology Slides for Quantitative Analysis. ISBI, 2009
[2] He et al: Deep Residual Learning for Image Recognition. ICCV, 2016
[3] Araujo et al: Classification of Breast Cancer Histology Images Using Convolutional Neural Networks. PloS one, 2017

Acknowledgement

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