Multitask pretraining outperforms ImageNet in learning general representations in computational pathology

Till Nicke[1], Jan Raphael Schäfer[1], Annkristin Lange[1], Felix Thielke[1], Nadine S. Schaadt[3], Friedrich Feuerhake[3,4], Dorit Merhof[1,2], Henning Höfener[1], Johannes Lotz[1]

1) Fraunhofer MEVIS, Institute for Digital Medicine, Germany
2) Institute of Image Analysis and Computer Vision, University of Regensburg, Germany
3) Institute for Pathology, Hannover Medical School, Germany
4) Institute for Neuropathology, University Clinic Freiburg, Germany

Introduction

Effective deep learning requires vast amounts of accurately annotated training data, whereas usually, only small amounts of patch-level annotations are available in computational pathology. We investigated if histology-specific representations allow sample-efficient learning on unseen tasks by pretraining a joint model on publicly available datasets using multi-task learning.

Material and methods

We created a large-scale multi-task database comprising 10 public histology datasets covering cancer classification and segmentation, mitosis detection, and nuclei identification on a variety of different H&E stained tissue types. On this database, we trained a generalist model using multi-task learning.

Our model was evaluated by predicting inclusion of invasive breast cancer on two datasets and estimating cancer cellularity on one dataset. Results were compared to those obtained using ImageNet pretraining.

For the evaluation process, the pretrained model was frozen and the latent representations of the evaluation tasks were used as inputs to task-specific random forest predictors that learned task outputs. These predictors were trained multiple times on differently sized subsets of the unseen evaluation data. We evaluated the transfer performance using the F1-score and the mean squared error.

Results

Compared with the baseline, the mean F1-scores for the breast cancer classification improved from 63.50% to 75.22% for task one and from 40.41% to 52.53% for task two. The mean squared error in cellularity estimation improved by 11.8% from 0.0566 to 0.0506.

Conclusion

Pretraining on a diverse multi-label database generates general representations suitable for effectively detecting malignant tissue components in breast cancer.

Partly funded by the German ministry of education and research (BMBF), project: SynDICAD (01IS21067C)

Keywords: Multi-task learning, representation learning, sample-efficiency,

Notes:

Author's Index preview

F

Feuerhake Friedrich 1

H

Höfener Henning 1

L

Lange Annkristin 1
Lotz Johannes 1

M

Merhof Dorit 1

N

Nicke Till 1

S

Schaadt Nadine S. 1
Schäfer Jan Raphael 1

T

Thielke Felix 1