Texture-Based Oral Cancer Detection: A Performance Analysis of Deep Learning Approaches

UPPSALA UNIVERSITET

Jo Gay, Hugo Harlin, Elisabeth Wetzer, Joakim Lindblad, Natasa Sladoje
Uppsala University, Sweden

Introduction

- Aim: Early stage oral cancer detection by screening of brush samples acquired in dental clinics
- Automated approaches based on image classification are required
- Comparative analysis of deep learning methods for texture-based classification:

  Can Local Binary Patterns improve performance of CNNs?

- Four recently published CNNs with LBP flavor are compared w.r.t. their performance on oral cancer cell classification:

  Juefei-Xu et al. [3]
  Li et al. [4]
  Marcos et al. [5]
  Levi et al. [6], Wetzer et al. [7]

- Dataset: 10 slides from 6 patients, 3 with cancer and 3 healthy, further data collection is ongoing
- 10274 cell images, size 80x80
- Training/test set split are made on a patient level to avoid data leakage
- Labels on patient level, not for individual cells

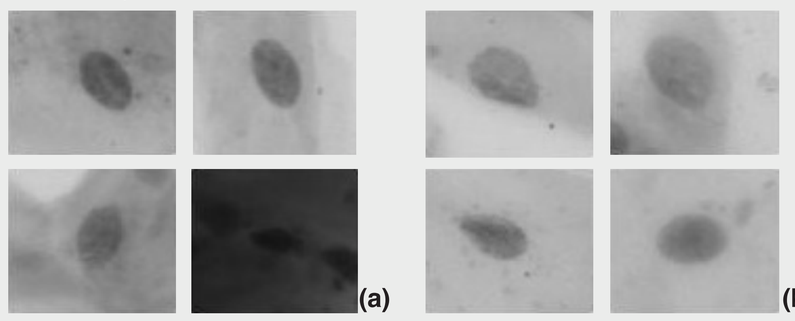


Fig. 1: Images in dataset, (a) healthy, (b) cancerous

Model I - Juefei-Xu [3]

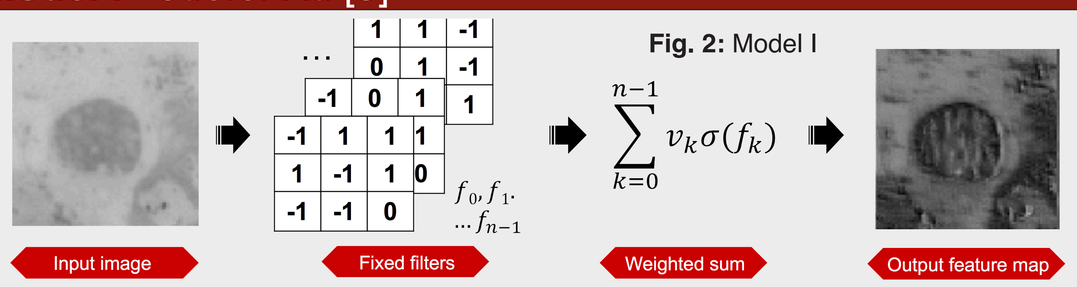


Fig. 2: Model I

- 512 fixed filters f_k with randomly positioned values ±1 and 10% zeros
- Rectified linear activation $\sigma(f_k)$
- Trainable weights v_k used in linear sum to create feature map
- Feature map as input for next layer

Model II - Li [4]

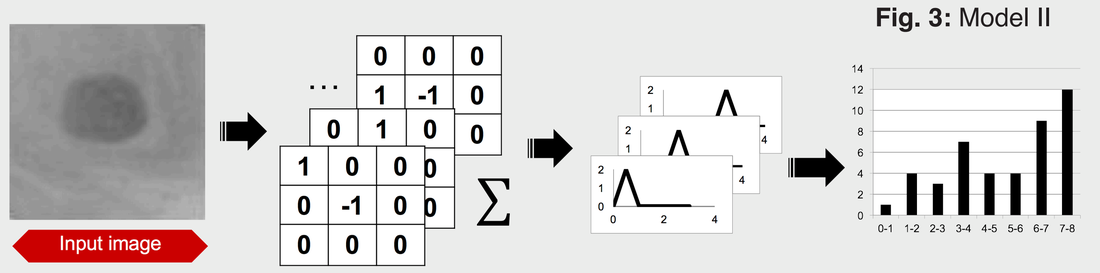


Fig. 3: Model II

- 8 fixed directional difference filters
- Sum of sigmoid activations, followed by gating functions
- Average responses are binned into histograms
- Dense layer and sigmoid activation

Model III - Marcos [5]

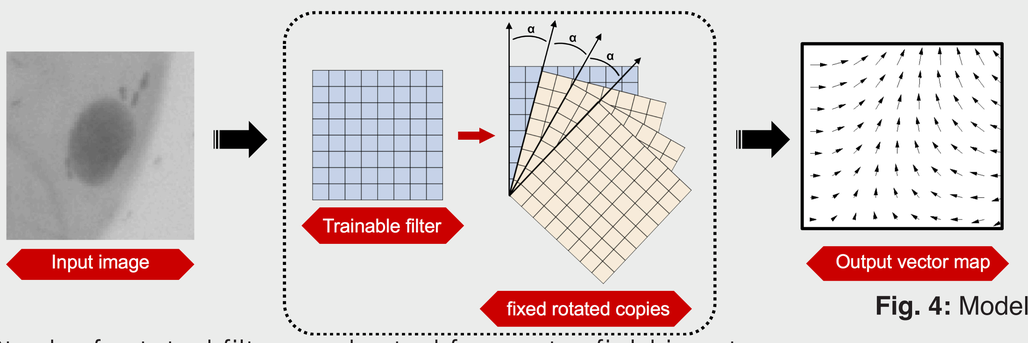


Fig. 4: Model III

- Stack of rotated filters, adapted for vector field input
- Output angle is given by the rotation of the filter with highest activation
- Activation magnitude & direction form 2D vector field output
- Batch norm & spatial pooling for vector fields

Model IV - Levi [6], Wetzer [7]

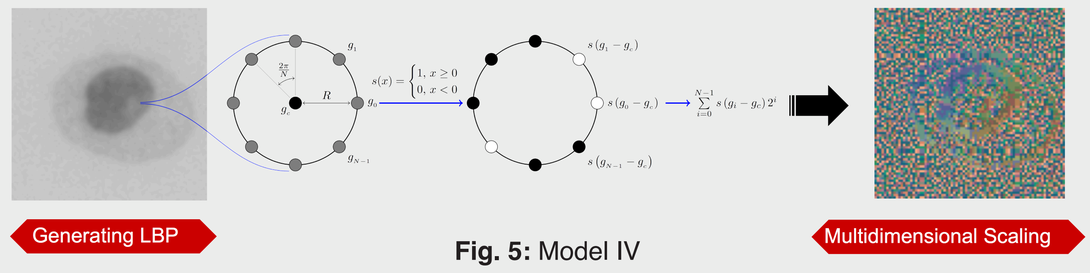
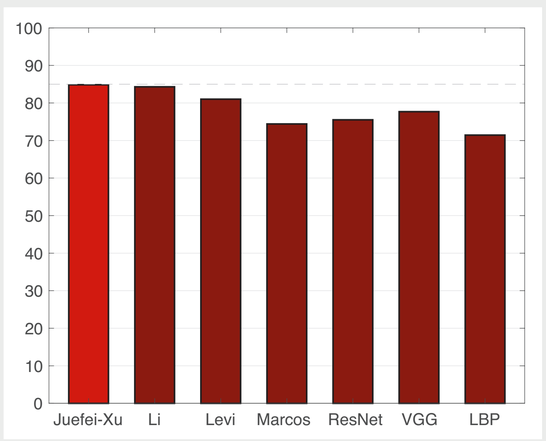


Fig. 5: Model IV

- LBP are mapped into 3D metric space by multidimensional scaling
- Two ResNet architectures are trained in parallel, one on intensity images, one on LBP maps
- Late Fusion: Softmax output is concatenated and classified by SVM

Results and Conclusions



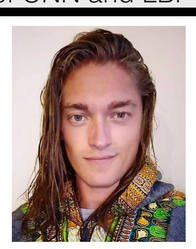
Model	Accuracy	F_1-Score
Juefei-Xu [3]	**81.03%**	**84.85%**
Li [4]	**80.70%**	**84.30%**
Levi [6]	**72.91%**	**81.01%**
Marcos [5]	**68.37%**	**74.41%**
ResNet	**78.34%**	**75.51%**
VGG	**80.66%**	**77.68%**
LBP + ANN	**75.42%**	**71.44%**

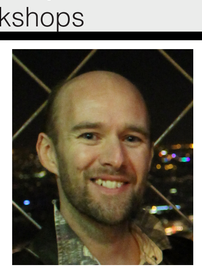
Fig. 6: Performance measured by F_1-score (imbalanced dataset)

- Rotational equivariance did not improve performance
- All three models based on LBPs exhibit higher performances w.r.t. the F1-score over conventional CNN training based on solely intensity input
- Texture imposed to a CNN in different ways improves performance

References

[1] Wieslander et al.: Deep convolutional neural networks for detecting cellular changes due to malignancy. ICCV, 2017
[2] Ojala et al.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on PAMI, 2002
[3] Juefei-Xu et al.: Local binary convolutional neural networks. CVPR, 2017
[4] Li et al.: Face spoofing detection with local binary pattern network. J. of Vis. Comm. & Img. Rep., 2018
[5] Marcos et al.: Rotation equivariant vector field networks. In ICCV, 2017
[6] Levi et al.: Emotion Recognition in the Wild via Convolutional Neural Networks and Mapped Binary Patterns. Proc. ACM ICMI, 2015
[7] Wetzer et al.: Towards Automated Multiscale Imaging Analysis in TEM: Glomerulus Detection by Fusion of CNN and LBP Maps. ECCV 2018 Workshops


JOANNA.GAY@GMAIL.COM

HUGOHARLIN@YAHOO.SE

ELISABETH.WETZER@IT.UU.SE

JOAKIM.LINDBLAD@IT.UU.SE

NATASA.SLADOJE@IT.UU.SE

Centre for Image Analysis
Uppsala University

European Women in Mathematics