IHC Classification in Breast Cancer H&E Slides with a Weakly-Supervised Approach

Sara P. Oliveira[1,2], João Ribeiro Pinto[1,2], Tiago Gonçalves[1,2],Hélder Oliveira[1,3], Jaime S. Cardoso[2,1]
[1]INESC TEC, Porto,Portugal, [2]FEUP and [3]FCUP, University of Porto, Porto, Portugal
contact: sara.i.oliveira@inesctec.pt

INTRODUCTION

The analysis of tissue sections of cancer specimens obtained by biopsy commonly starts with haematoxylin and eosin (H&E) staining, which is usually followed by immunohistochemistry (IHC), a more advanced staining technique used to highlight specific protein receptors, such as the HER2. In fact, the overexpression of HER2 is observed in 10%-20% of BCa cases and has been associated with aggressive clinical behaviour and poor prognosis. However, these cases have a better response to targeted therapies and consequent improvements in healing and overall survival. Despite the efficiency of IHC and ISH, additional costs and time spent on further testing might be avoided if HER2 overexpression could be inferred from H&E slides, as a preliminary indication of the IHC result. Thus, we propose a framework that separately processes H&E slide tiles and outputs an IHC label for the whole slide.

METHODOLOGY

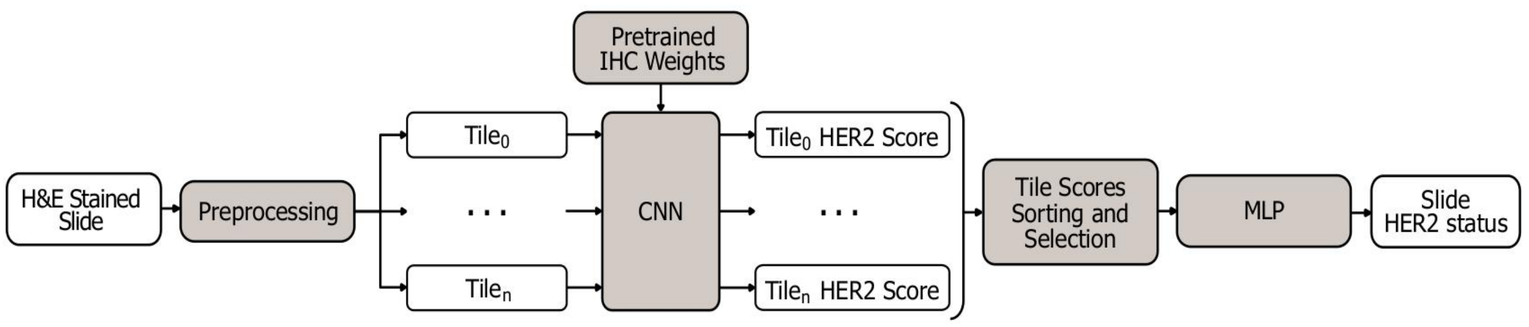

Figure 1: The proposed approach for weakly-supervised HER2 status classification on BCa H&E slides.

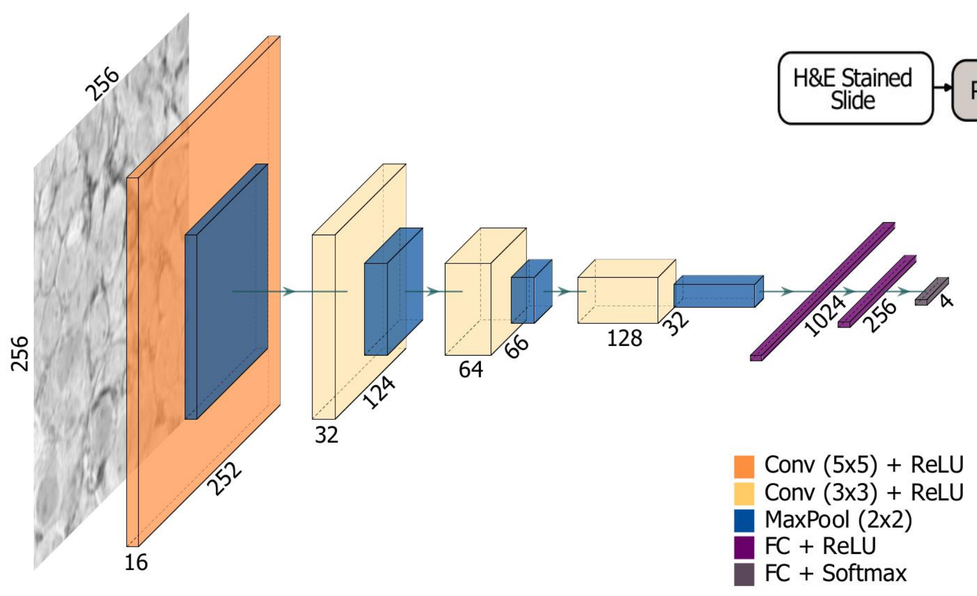

Figure 2: Architecture of the implemented convolutional neural network.

The proposed method (Figure 1) comprises a CNN (Figure 2), pre-trained for the task of HER2 scoring of IHC tiles. The pre-trained parameters are then transferred to the task of HER2 status prediction on H&E tiles, to provide the network with some knowledge of the tissue structures' appearance. Individual tile scores are then combined in a single label, returning the HER2 status for the whole slide.

DATASET

The dataset is composed of subsets of slides from two public datasets:

- HER2 Scoring Contest training set (HER2SC), with slides of 52 cases of invasive BCa, stained with both IHC and H&E;
- TCGA-TCIA-BRCA collection (BRCA), with 54 H&E slides.

All slides have the same original resolution and are weakly annotated with HER2 status (negative/positive) and score (0+, 1+, 2+, 3+), obtained from the histopathological reports.

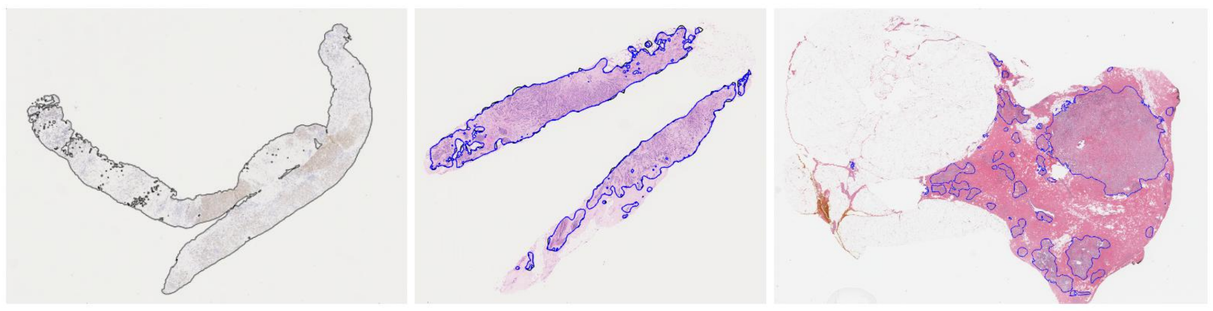

Figure 1: Image examples from used datasets: HER2SC IHC stained slides (left), HER2SC H&E stained slides (middle) and BRCA H&E stained slides (right).

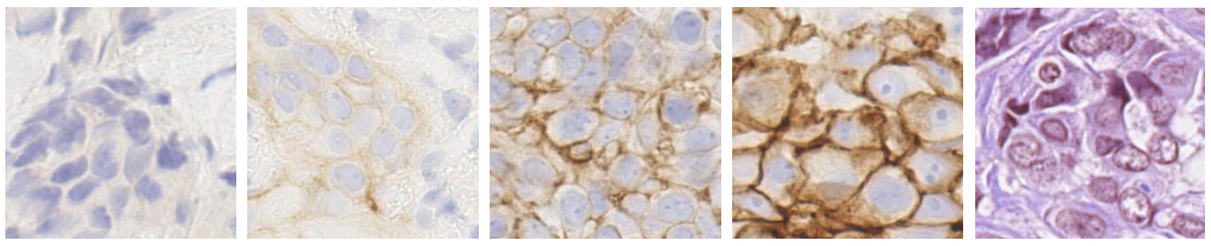

Figure 2: Tile examples extracted from IHC0+, IHC1+, IHC2+, IHC3+ and H&E slides (from left to right).

EXPERIMENTAL RESULTS

Prediction \ True Class	0	1	2	3
0	490	132	2	0
1	176	384	64	0
2	45	159	419	1
3	0	0	1	623

Table 1: CNN confusion matrix for HER2 scoring in IHC tiles.

	Accuracy	F1-score	Precision	Recall
HER2SC	83.3%	86.7%	89.6%	87.5%
BRCA	53.8%	21.5%	81.2%	31.5%

Table 2: H&E HER2 status classification results of the proposed method on the HER2SC and BRCAtest datasets.

CONCLUSION & FUTURE WORK

The evaluation results in single-database (HER2SC) and cross-database (BRCA) settings show the potential of the proposed method in standard and more challenging situations, indicating that it is possible to accurately infer BCa HER2 status solely from H&E slides.

However, further efforts should be devoted to performance improvement:

- training of the tile HER2 scoring CNN and the aggregation MLP in a single optimization process;
- aggregation of individual scores using tile location to take spatial consistency into account;
- convert the H&E tiles to the IHC space using, for example, GANs.

This work was partially funded by the Project "TAMI: TransparentArtificial Medical Intelligence" (NORTE-01-0247-FEDER-045905), co-financed by ERDF through the programs COMPETE 2020 and NORTE2020, and by FCT, under the CMU-Portugal International Partnership, and also the FCT PhD grants "SFRH/BD/139108/2018", "SFRH/BD/137720/2018" and "SFRH/BD/06434/2020"