Training and Validating an Artificial Intelligence Algorithm for HER2 Assessment in Breast Carcinoma-Stained Whole-Slide Images

B. Poulet[2], F. Thomas[1], Y. Salhi[1], C. Bossard[1,2], B. Cormier[2], L. Lambros[2], I. Chokri[2], F. Jossic[2], S. Thanguturi[2], I. de Pinieux[2], S. Salhi[1], J. Chetritt[1,2]

[1] DiaDeep, France - [2] Pathology Department, IHP Group, France

Background

HER2 amplification occurs in up to 20% of invasive breast carcinomas, impacting therapeutic response (ou treatment decision), as well as the majority of Her2 low status with the new Her2-targeted therapies. Thus, accurate assessment of HER2 status is vital (ou critical). We developed a fully-automated AI algorithm for precise HER2 protein overexpression quantification, aiding pathologists in immunohistochemistry-based scoring.

Methods

We developed and trained an end-to-end algorithm for scoring HER2 expression in immunohistochemistry-stained whole slides images (WSI) of breast carcinoma. The proposed algorithm automatically detects invasive tumor regions and allows the quantification and characterization of tumor cells on these zones. To achieve this, 3878 tiles of 512 pixels size of HER2-immunostained WSI were labeled at magnification x5 and used to train a first deep learning model able to detect invasive regions. Then, 4793 tiles annotated at the cell level, at magnification x40, were used to train a second model able to segment tumor cells as well as to analyse their cell membrane (Figure 2). Finally, this exhaustive quantification was used to assess the HER2 score in the WSI, in accordance with ASCO/CAP guidelines. The proposed pipeline provides an interpretable score along with the details of the slide analysis without requiring any annotations from pathologists. The performance of the model was investigated over a new independent dataset (Figure 1) scored by 3 senior pathologists (ground-truth). The ground truth scores were determined by these pathologists after a consensus. Besides, after a washout period of two months, the slides were rescored with the AI Assistance to compare the inter-observer agreement between these two scoring-methods (with and without AI).

The evaluation dataset was built using 68 routine cases of HER2 immunostained slides of Breast Cancer, without any further selection criteria, in order to reflect the routine practice.

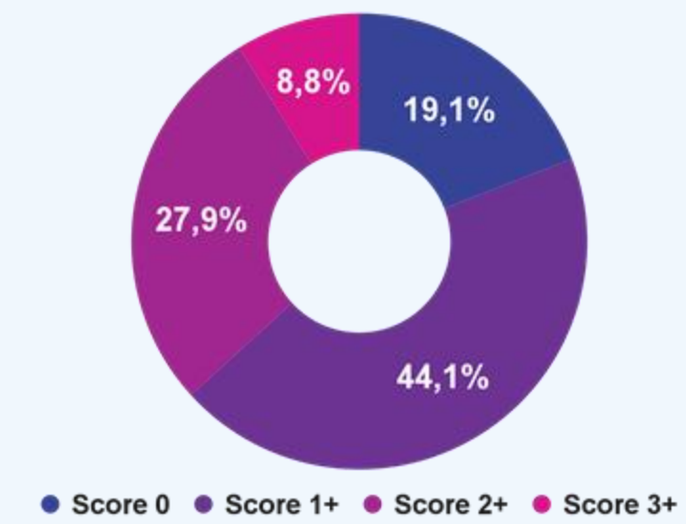

Figure 2 - Ground Truth Scores in the test dataset

Results

The model exhibited an overall balanced accuracy of 90.0% on the testing dataset (Figure 3), thus demonstrating its capability to assist pathologists in their routine practice while providing an interpretable HER2 Score. Notably, 100% of the HER2 3+ scores and 93.3% of the HER2-Low cases were identified by the model, yielding promising results to identify patients eligible to targeted therapy. Moreover, the inter-observer agreement rose from 57% to 75% with the assistance of AI (Figure 4).

Figure 1. Model Pipeline

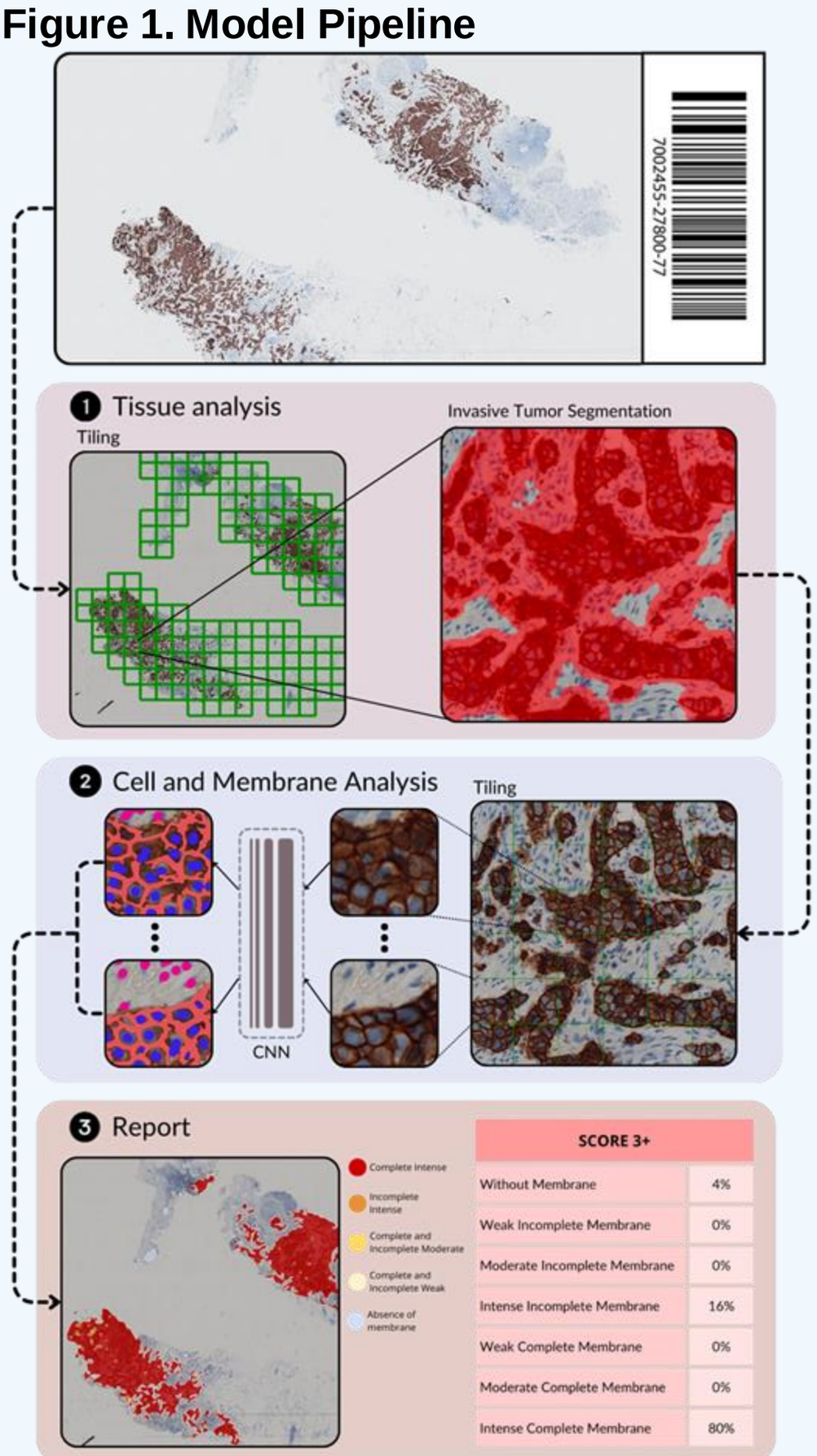

Figure 3 - Performance of the AI Algorithm vs. Ground Truth (confusion matrix)

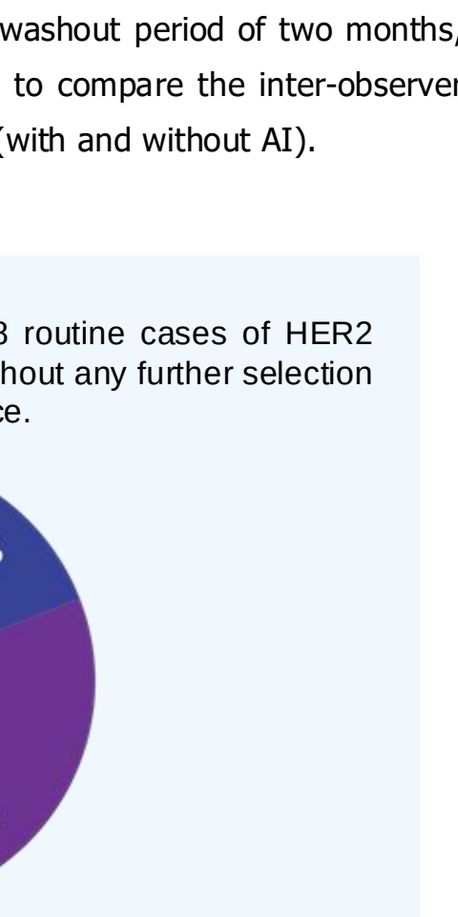

Figure 3 - Inter-pathologists agreement with and without the help of the AI

Conclusion

This study demonstrates the efficacy of the proposed fully-automated AI system in accurately assessing HER2-immunostained Whole Slide Images, in accordance with the 2018 ASCO/CAP Guidelines. By providing pathologists with easily interpretable scores, such an AI tool could be used as a regular aid in clinical decision-making, improving reproducibility and aiding in the assessment of challenging cases due to the detailed information it provides for each cases. Significantly, the introduced model exhibited reliability and consistency particularly in HER2-Low cases, eligible to new Her2-targeted therapies, setting a foundation for future advancements in precision medicine and diagnostic accuracy within HER2-related pathology.

contact@diadeep.com