Tumor Infiltrating Lymphocytes Recognition in Primary Melanoma by Deep Learning Convolutional Neural Network

P310

Filippo Ugolini[1], Francesco De Logu[2], Luigi Francesco Iannone[2], Francesca Brutti[3], Sara Simi[1], Vincenza Maio[1], Vincenzo de Giorgi[4], Anna Maria di Giacomo[5], Clelia Miracco[5], Francesco Federico[6], Ketty Peris[7], Giuseppe Palmieri[8], Antonio Cossu[9], Mario Mandalà[10], Marco Laurino[3], Daniela Massi[1]

[1]Department of Health Sciences, Section of Pathological Anatomy, University of Florence; [2]Department of Health Sciences, Section of Clinical Pharmacology and Oncology, University of Florence; [3]Institute of Clinical Physiology, National Research Council of Pisa; [4]Department of Health Sciences, Section of Dermatology, University of Florence; [5]Medical Oncology and Immunotherapy, Center for Immuno-Oncology, University of Siena; [6]Institute of Pathology, Sacred Heart Catholic University Rome; [7]Institute of Dermatology, Sacred Heart Catholic University Rome; [8]Unit of Cancer Genetics National Research Council of Sassari; [9]Department of Medical, Surgical and Experimental Sciences, University of Sassari; [10]Oncology Unit, Department of Medicine and Surgery, University of Perugia

Background

The presence of tumor-infiltrating lymphocytes (TIL) has been associated with a favorable prognosis of primary melanoma (PM). The recent development of the artificial intelligence (AI) based approach in digital pathology has been proposed for the standardized assessment of TIL on hematoxylin and eosin (H&E)-stained images (whole slide images, WSI).

To date, standardized protocols for TILs immunophenotyping and quantification are still lacking (1-3). Despite the limited availability of specific large studies, the immunophenotype seems to play a role in the efficacy of the local immune response against melanoma. Previous studies, focused mostly on T-lymphocyte populations, showed that the prevalence of CD8+ T cells is related to a better clinical outcome in contrast with T-regulatory (Treg) infiltration (4-6).

The development of a convolution neural network (CNN) for image-based automated assessment of TILs on hematoxylin and eosin (H&E)-stained sections based on an Artificial Intelligence (AI) approach could represent a useful tool. The implementation of an easy-to-use and standardized digital analysis solution able to offer fast and accurate information to pathologists is one of the great challenges of the recent years (7,8).

Methods

The study included a retrospective collection of formalin-fixed paraffin-embedded (FFPE) Stage II-III primary invasive PM (N=307), comprising a training set (N = 237 WSI) from University of Florence (N = 115, DATASET #1), University of Sassari (N = 43, DATASET #2), University of Siena (N = 15, DATASET #3), and Papa Giovanni XXIII Cancer Center Hospital Bergamo (N = 64, DATASET #4) from years 2000 to 2015 and a disjoint testing set (N = 70, DATASET #5) from Catholic University of the Sacred Heart, Fondazione Policlinico Universitario A. Gemelli IRCCS Rome from years 2016 to 2020 (Figure 1). The clinical and pathological parameters extracted from the database included sex, age (continuous variable), date of diagnosis of the primary tumor, ulceration status (absent/present), Breslow thickness (in mm), histotype, Clark's level, mitotic rate, TILs, sentinel lymph node biopsy (SLNB) status, and follow-up.

Figure 1. Examples of patches (x400 magnification) extracted from the four datasets used for the training and validation of the network, and from the disjoint dataset used for testing the model in TIL recognition.

Here, we have applied a new convolution neural network (CNN) analysis of PM WSI to automatically assess the infiltration of TIL and extract a TIL score. A CNN was trained and validated in a retrospective cohort of 307 PMs including a training set (237 WSI, 57,758 patches) and an independent testing set (70 WSI, 29,533 patches).

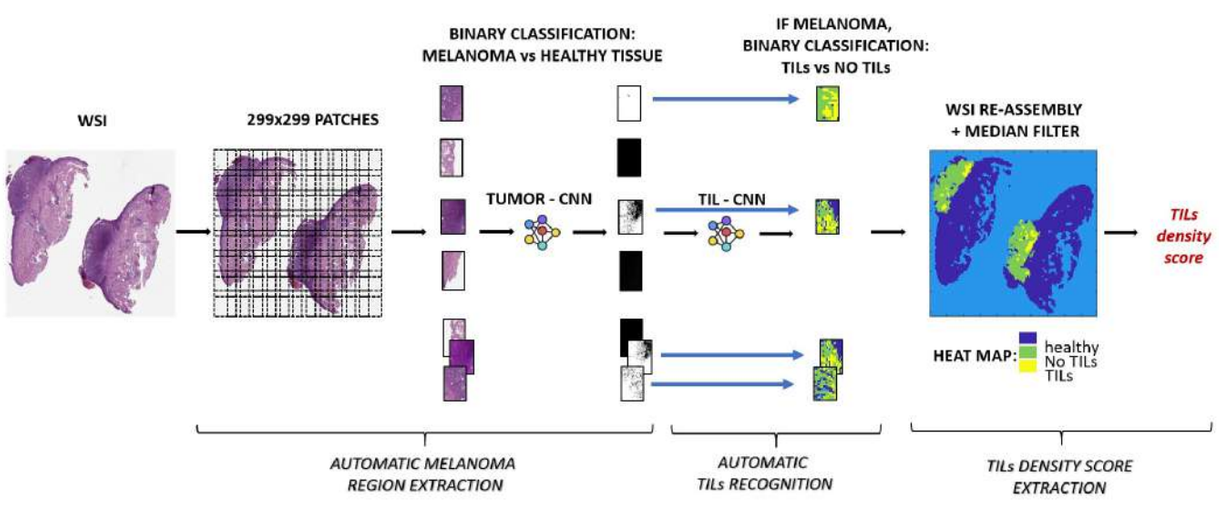

Figure 2. Pipeline of the proposed model's application. WSI example at x400 magnification.

Results

The global accuracy of the trained model for automatic recognition of patches with TILs (TIL-CNN) returned a value of 99.5% considering the validation set and a value of 100.0% considering the test set (Figure 3). For the validation set, specificity and sensitivity achieved 99.7% and 99.3% respectively. Cohen's kappa returned 0.99 and the F1 score was 0.99. For the test set, specificity, and sensitivity both achieved 100.0%. Cohen's kappa returned 1.00 and the F1 score was 1.00.

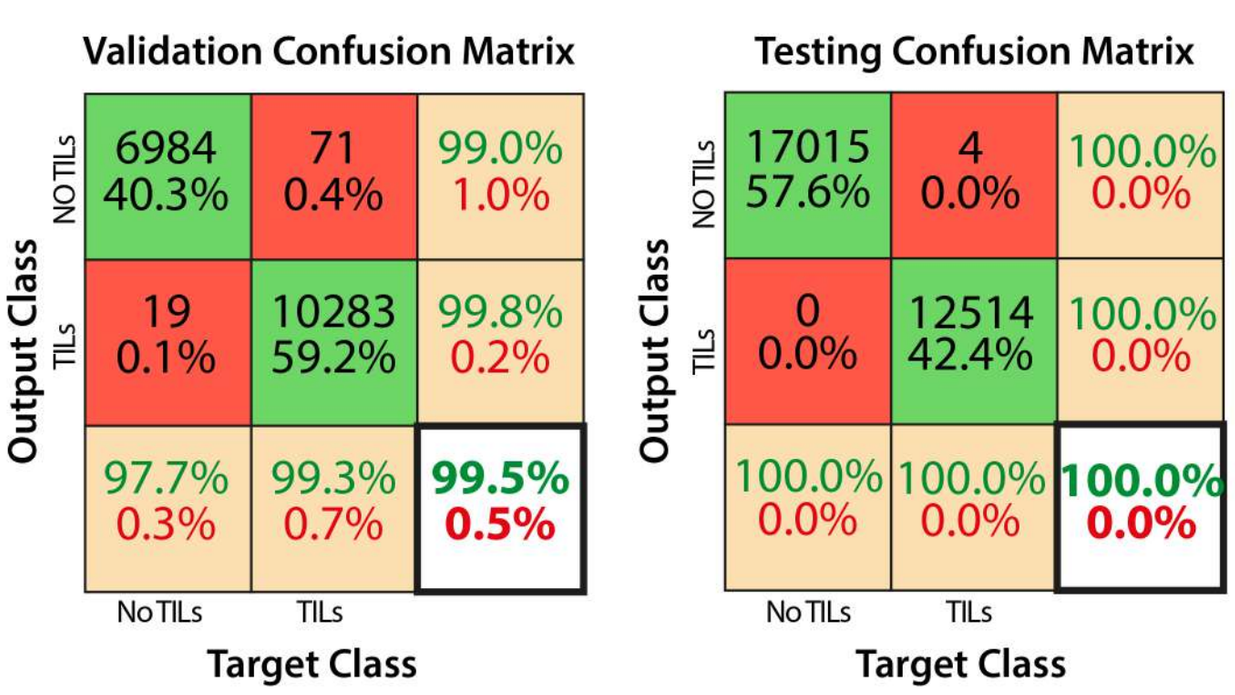

Figure 3. Validation and testing confusion matrices of the TIL-CNN for recognition of patches with (TILs) and without TILs (NO TILs). The diagonal cells correspond to observations that are correctly classified, and the off-diagonal cells to incorrectly classified observations. The column on the far right of each confusion matrix shows the percentages of all the item predicted to belong to each class that are correctly (green text) and incorrectly (red text) classified. The cell in the bottom right of each confusion matrix shows the overall accuracy (green text) and error rate (red text).

Disease-free survival analysis

The disease-free survival analysis was performed on a subset of 170 patients with complete clinical history and prognostic information. In the D1 survival model, the presence of "High TIL" (AI-TIL score >5%) in a patient presented a longer significative DFS (HR 0.6, SD 0.2, p-value=0.013) compared with "Low TIL" (AI-TIL score <5%) patients. Other covariates associated with death or relapse hazard were age (p-value<0.001) and Breslow thickness (p-value<0.01). Crude Kaplan–Meier curves with respect to dichotomous AI-TIL (low vs high) showed a quite similar DFS for AI-TIL high compared with AI-TIL low (Figure 4C). By adding the SLNB to the DFS model, although the AI-TIL index lost significance with respect to D1, the index remained statistically significant in the D11 model (p-value<0.05).

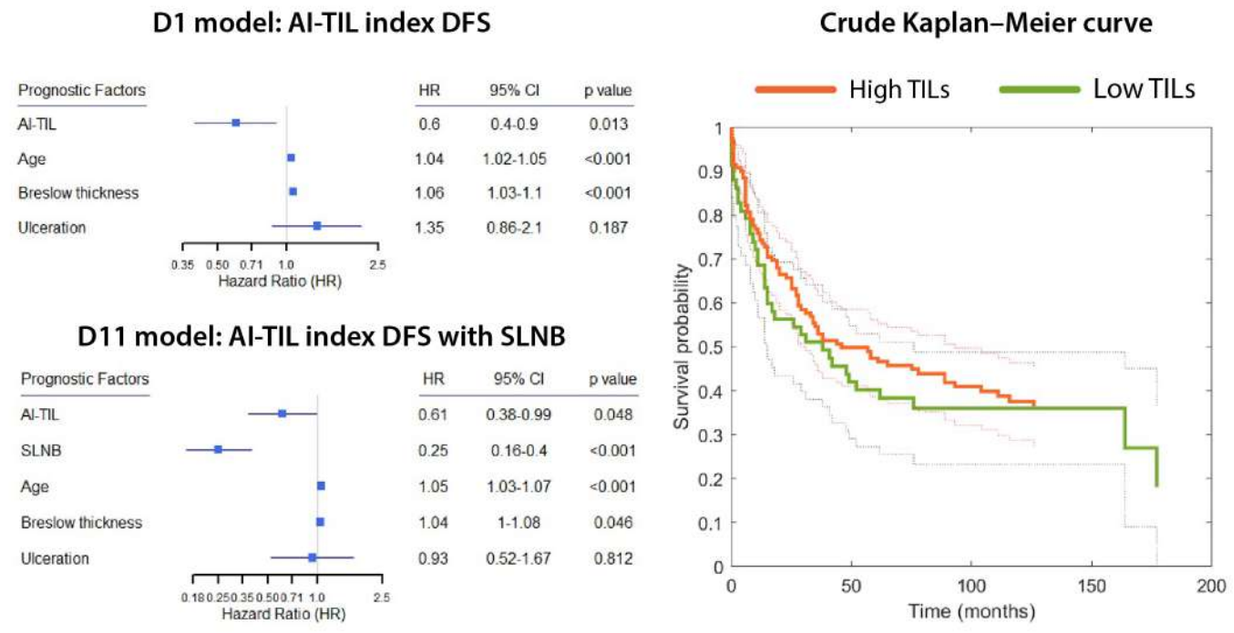

Figure 4. Forest plot of Cox Proportional Hazard regression Ratios and correspondent p-values referred to D1 model. AI-TIL is the dichotomous AI-TIL density index. The other regressors are the clinicopathological prognostic factors (top left). Forest plot of Cox Proportional Hazard regression Ratios and correspondent p-values referred to D11 model. Sentinel Lymph Node Biopsy (SLNB) has been added with respect to D1 model (bottom left). Crude Kaplan–Meier curve with respect to dichotomous AI-TIL ("Low TILs" and "High TILs") of both survival models (right).

Overall survival analysis

The overall survival analysis was performed on a subset of 170 patients with a complete clinical history and prognostic information. In the O1 survival model, a "High TIL" factor was associated with longer OS (reduced death hazard of 41%; HR 0.59, SD 0.2, p-value=0.017) with respect to a "Low TIL" factor. Other covariates associated with death hazard were: age (p-value<0.001), Breslow thickness (p-value<0.01), and ulceration (p-value<0.05). Crude Kaplan–Meier curve with respect to dichotomous AI-TIL (low vs high) showed a longer OS for AI-TIL high compared with AI-TIL low (median OS 75 and 68 months, respectively).

The addition of SLNB returned a loss of significance of the AI-TIL index for the O11 model, although its trend was consistent with O1. In fact, in the O11 survival model, the death hazard was not significantly related to the AI-TIL index (HR 0.63 SD 0.3, p=0.096). Age (p-value<0.001), Breslow thickness (p-value<0.05), and SLNB (p-value<0.001) were associated with death hazard, while ulceration lost significance with respect to the O1 model.

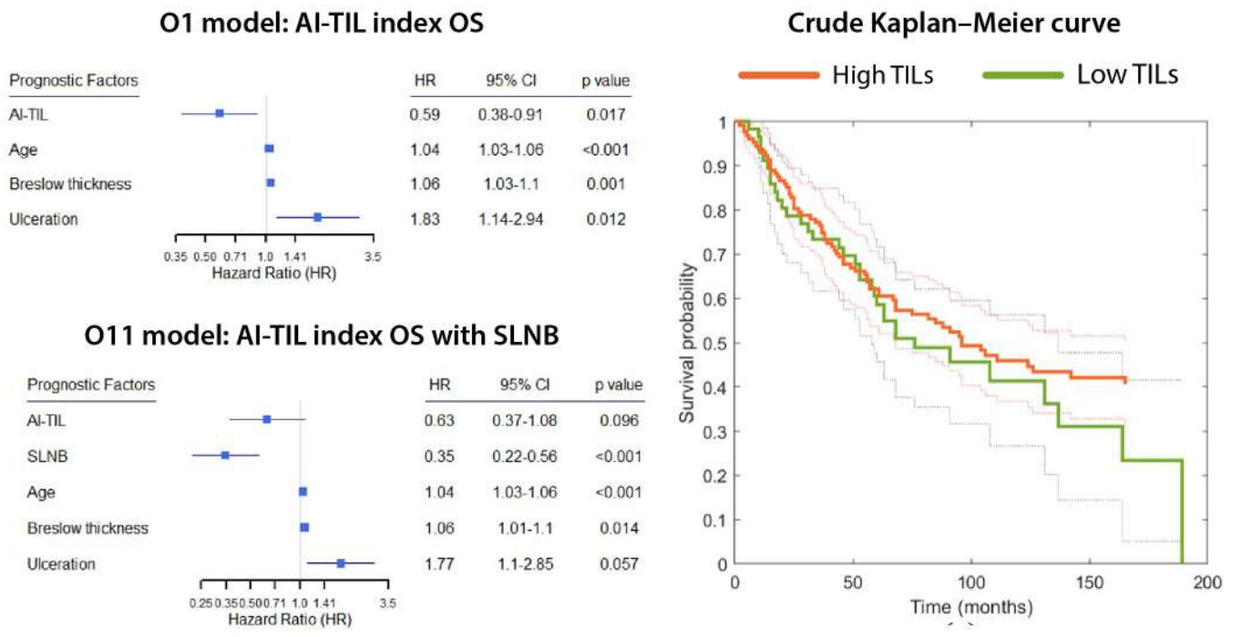

Figure 5. Forest plot of Cox Proportional Hazard regression Ratios and correspondent p-values referred to O1 model. AI-TIL is the dichotomous AI-TIL density index. The other regressors are the clinicopathological prognostic factors (top left). Forest plot of Cox Proportional Hazard regression Ratios and correspondent p-values referred to O11 model. Sentinel Lymph Node Biopsy (SLNB) has been added with respect to O1 model (bottom left). Crude Kaplan–Meier curve with respect to dichotomous AI-TIL ("Low TILs" and "High TILs") of both survival models (right).

Conclusions

The proposed pipeline for AI-TIL score extraction can increase uniformity across cohorts and clinical sites, reduce the need of monitoring by highly experienced pathologists, and prevent operator variability. Further prospective investigation is needed to determine whether AI-TIL scores should be routinely included in PM pathology reports and future AJCC staging revisions, and whether they should be applied to metastatic samples to shape the design of novel immunotherapeutic protocols.

References

1. Meneveau MO, Sahli ZT, Lynch KT, Mauldin IS, Slingluff CL: Immunotyping and Quantification of Melanoma Tumor-Infiltrating Lymphocytes. Methods Mol Biol [Internet], Methods Mol Biol, 202.
2. Maibach F, Sadozai H, Seyed Jafari SM, Hunger RE, Schenk M: Tumor-Infiltrating Lymphocytes and Their Prognostic Value in Cutaneous Melanoma. Front Immunol, Front Immunol, 2020.
3. Nishimura T, Nakui M, Sato M, Iwakabe K, Kitamura H, Sekimoto M, Ohta A, Koda T, Nishimura S: The critical role of Th1-dominant immunity in tumor immunology. Cancer Chemother Pharmacol, 2000.
4. Viguier M, Lemaître F, Verola O, Cho M-S, Gorochov G, Dubertret L, Bachelez H, Kourilsky P, Ferradini L: Foxp3 expressing CD4+ CD25(high) regulatory T cells are overrepresented in human metastatic melanoma lymph nodes and inhibit the function of infiltrating T cells. J Immunol, 2004.
5. De Logu F, de Araujo DSM, Ugolini F, Iannone LF, Vannucchi M, Portelli F, Landini L, Titiz M, De Giorgi V, Geppetti P, Massi D, Nassini R: The TRPA1 Channel Amplifies the Oxidative Stress Signal in Melanoma. Cells, 2021.
6. Sobottka B, Nowak M, Frei AL, Haberecker M, Merki S, Aebersold R,et al.: Establishing standardized immune phenotyping of metastatic melanoma by digital pathology. Lab Invest, 2021.
7. Dimitriou N, Arandjelović O, Caie PD: Deep Learning for Whole Slide Image Analysis: An Overview. Front Med (Lausanne), 2019
8. Srinidhi CL, Ciga O, Martel AL: Deep neural network models for computational histopathology: A survey. Med Image Anal, 2021.