# Sentinel lymph node status prediction using self-attention networks and contrastive learning from routine histology images of primary tumours

**Carlos Hernandez Perez**[1] CARLOS.HERNANDEZ.P@UPC.EDU
**Marc Combalia**[2,3] MCOMBALIA@CLINIC.CAT
**Josep Malvehy**[2,3] JMALVEHY@CLINIC.CAT
**Veronica Vilaplana**[1] VERONICA.VILAPLANA@UPC.EDU
[1] *Image Processing Group, Universitat Politècnica de Catalunya, Barcelona, Spain*
[2] *Athena Tech, S.L., Barcelona, Spain*
[3] *Dermatology Department, Melanoma Unit, Hospital Clínic de Barcelona, IDIBAPS, University of Barcelona, Barcelona, Spain*

**Editors:** Under Review for MIDL 2022

## Abstract

Deep learning-based computational pathology approaches are becoming increasingly prominent in histopathology image analysis. However, these images typically come with drawbacks that hamper automatic analysis, which include: labeled sample scarcity or the extremely large size of the images (ranging from $10^7$ to $10^{12}$ pixels). Nonetheless, they have proven to be a powerful tool for diagnosis and risk prevention. One such prevention is reducing the number of patients who undergo surgeries that do not benefit them. This study develops a pipeline for predicting sentinel lymph node (SLN) metastasis non-invasively from digitised Whole Slide Images (WSI) of primary melanoma tumours. Furthermore, we combine the use of a weakly supervised architecture with self-supervised contrastive pre-training. We experimentally demonstrate that 1) the use of self-attention improves sentinel lymph node status prediction and 2) self-supervised contrastive learning improves the quality of the learned representations compared to a standard ImageNet pre-training, which boosts the model's performance.
**Keywords:** Sentinel lymph node status, Semi-supervised contrastive learning, Multiple-Instance learning, Self-attention.

## 1. Introduction

Many patients with melanoma undergo sentinel lymph node biopsy, an invasive surgical procedure that checks for the presence of metastasis. However, 80% of patients do not benefit from this intervention because they have unaffected SLNs (Gershenwald et al., 2017). Thus, this intervention comes at high human and monetary costs. Nevertheless, SLN status is a vital prognostic factor for the survival of skin-cancer patients.

Artificial intelligence and, especially, Convolutional Neural Networks (CNN) have proved to be state-of-the-art methods for medical imagining analysis. We aim to introduce a new architecture and training scheme based on self-attention networks and self-supervised contrastive learning (SimCLR) (Chen et al., 2020) to improve the performance of current methods for SLN status prediction (Brinker et al., 2021).

## 2. Methodology

In this study, a total of 195 digitised (WSI) were used. Each image was extracted from a primary melanoma tumour with known sentinel node (SN) status (107 SN negative and 88 SN positive) from the Hospital Clinic of Barcelona. Ten different stratified splits were created with 80%, 10%, and 10% of the WSIs for training, validation, and testing, respectively. The tissue was segmented from the background through OTSU's thresholding method. Afterward, patches of 256x256 were cropped as seen in Figure 1. Then, 2048 features were extracted per patch using a ResNet-50 architecture pre-trained on the ImageNet dataset. Additionally, a second ResNet-50 was trained with the SimCLR framework, which employs data augmentation to group similar samples closer in the encoded feature space (Chen et al., 2020). This framework was used to extract a new set of features. Both sets remain unaltered during the training of the classifier.

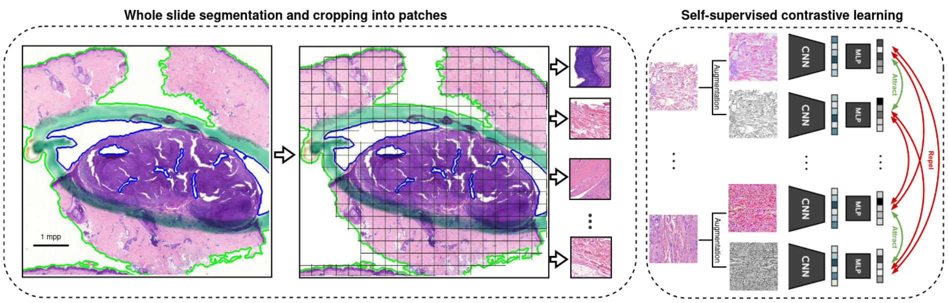


Figure 1: Segmentation and cropping of the WSI (left) and Self-supervised contrastive learning on the extracted patches' features (right).

Since only labels at the WSI level were available, we used a weakly supervised architecture based on Multiple Instance Learning (MIL). Specifically, we used an attention-based learning algorithm to identify sub-regions of high diagnostic value to accurately classify whole slides and instance-level clustering over the identified representative regions to constrain and refine the feature space (Lu et al., 2021). We also ran our experiments using a standard MIL architecture for comparison. Figure 2 shows the attention-based model as well as the resulting attention map for one WSI.

## 3. Results

Images were analyzed at 0.986 microns-per-pixel, a learning rate of $10^{-4}$ and L2 regularization ($5 \times 10^{-4}$) with stochastic gradient descent as optimizer; all parameters selected using hyperparameter optimization. For testing, we chose the weights of the model according to the best F1-score found in validation. Table 1 presents a comparison between the different models proposed in this work.

Table 1: Quantitative comparison between the proposed models and (Brinker et al., 2021). *Metrics extracted from their publication using a different dataset.

| **Pre-training** | F1-score | AUC | Balanced accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|
| Standard MIL ImageNet | 0.503 ±0.102 | 0.533 ±0.138 | 0.502 ±0.083 | 0.69 ±0.27 | 0.32 ±0.29 |
| Standard MIL SimCLR | 0.478 ±0.115 | 0.533 ±0.121 | 0.543 ±0.109 | 0.48 ±0.17 | 0.61 ±0.18 |
| Attention MIL ImageNet | 0.520 ±0.147 | 0.562 ±0.126 | 0.561 ±0.126 | 0.60 ±0.24 | 0.52 ±0.19 |
| Attention MIL SimCLR | **0.563** ±0.166 | **0.621** ±0.22 | **0.635** ±0.129 | **0.73** ±0.26 | 0.54 ±0.25 |
| Brinker et al.* | Not reported | 0.618 ±0.02 | 0.560 ±0.020 | 0.48 ±0.14 | **0.65** ±0.11 |

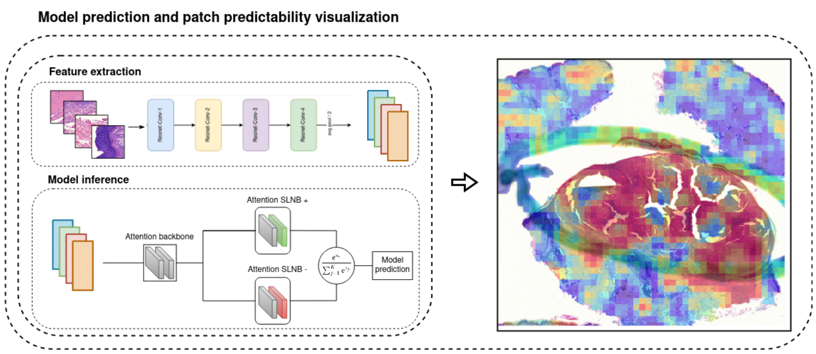


Figure 2: Schema of the model's prediction steps and patch predictability visualization.

## 4. Conclusions

Our study showed that CNN-based image classification of primary tumours to detect lymph node positivity is possible. Moreover, the use of SimCLR and self attention allows the model to improve all the considered metrics. However, further studies with more WSI are needed to validate and enhance the results. In addition, the use of other patient information complementary to the WSI may also increased the classifier's performance.

## 5. Acknowledgments

Work supported by the Spanish Research Agency (AEI) under project PID2020-116907RB-I00 of the call MCIN/ AEI /10.13039/501100011033 and the project 718/C/2019 funded by Fundació la Marato de TV3.

## References

Titus J Brinker, Lennard Kiehl, Max Schmitt, et al. Deep learning approach to predict sentinel lymph node status directly from routine histology of primary melanoma tumours. *European Journal of Cancer*, 154:227–234, 2021.

Ting Chen, Simon Kornblith, et al. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.

Jeffrey E Gershenwald et al. Melanoma staging: evidence-based changes in the american joint committee on cancer eighth edition cancer staging manual. *CA: a cancer journal for clinicians*, 67(6):472–492, 2017.

Ming Y Lu, Drew FK Williamson, et al. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature biomedical engineering*, 5(6):555–570, 2021.