A Multimodal Deep Learning Visual Interpretability Framework for Oropharnygeal Cancer Imagery

EPIDEMIUM Challenge ORL/IA : Predicting the evolution of Papillomavirus-induced ENT cancer from pathology section of tumor

Introduction

Researchers presuppose that presence of HPV viruses, interaction of immune cells etc. increase the prognostics of the disease. The challenge expects from us to present a system backed by machine learning, which is capable of shedding light on this complex set of pathological relationships by fusing:

- Various multiparametric biopsy slides
- And clinical data of patients.

Considering the intricacy and multimodality of the problem, we are convinced that **deep learning shall be at the core of the solution with visual interpretability maps as the resulting human-side interface.** Our rigorous study indicates that no existing architecture or methodology can be employed directly for this task. Certain vital reconfigurations need to be applied to state-of-the-art deep learning solutions.

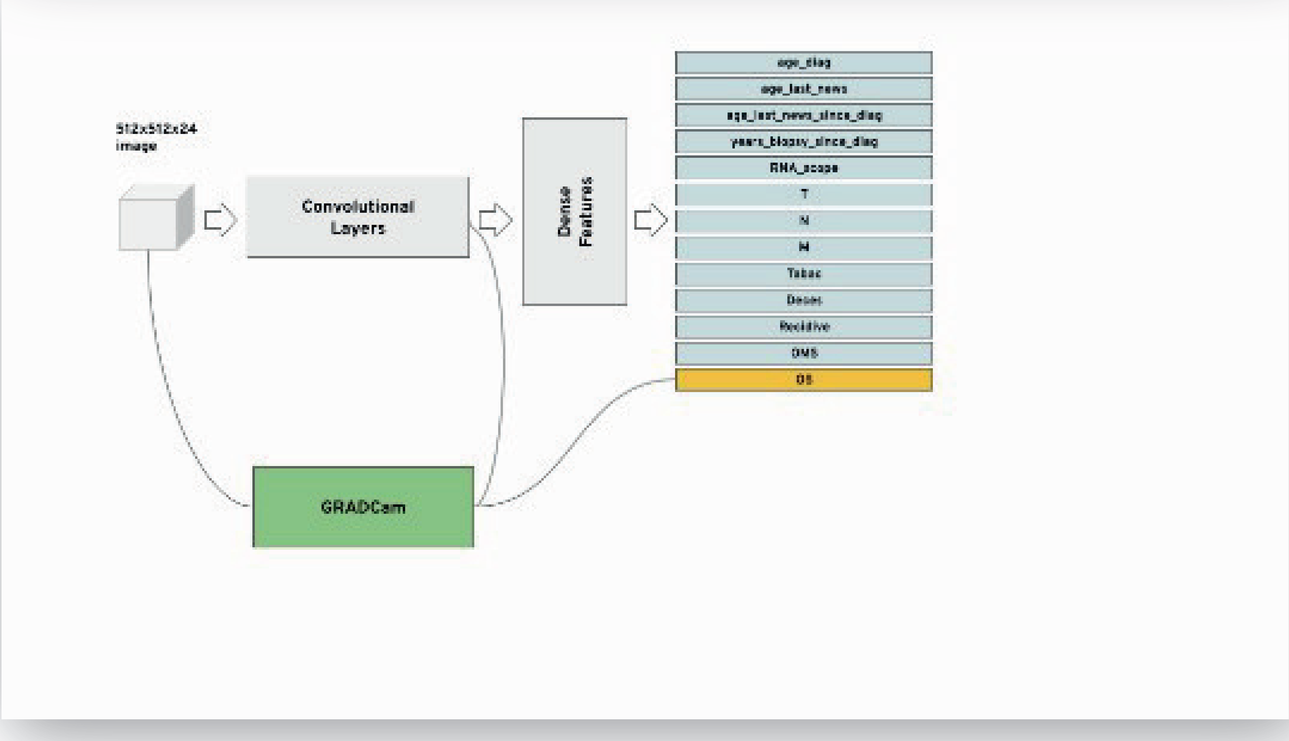
Multimodality of Clinical Statistics and Imagery : In order to develop a visual interpretability model which can direct researchers to better understand which aspects of microenvironment affects the evolution of the disease, we need to:

- Make our methodology conform to the multimodal statistical nature of the case.
- Consider these clinical parameters orthogonally, but as a compound.

Therefore, we have decided to design our methodology considering this fact.
For instance, we can not simply just take OS (Overall Survival) or whether the patient is dead or not as a single parameter and build our model.Just like the multimodal inherent nature of clinical data, the same applies to the multiplex pathology imagery of the patients. We can not interpret the effects of different markers orthogonally, **hence the proposed methodology shall incorporate the effects of the multiparametric slides coherently.**

Methodology

All the channels (multiplex stains) of an image are concatenated together to produce a single 24-channel image [512x512 pixels] which regress on 13 tabular data.



Our approach bases itself on a modified version of the GradCAM algorithm.

First of all, the main goal of the project is **to develop a visual interpretability method**, most conveniently a pixel-wise heatmap on images, to better understand the dynamics of the microenvironment around tumors on the evolution of oropharynx cancer.
In order to induce visual interpretability on the effects of microenvironment on the survivability, prognostics of the disease, we suggest **to follow gradients through only on the OS variable.**

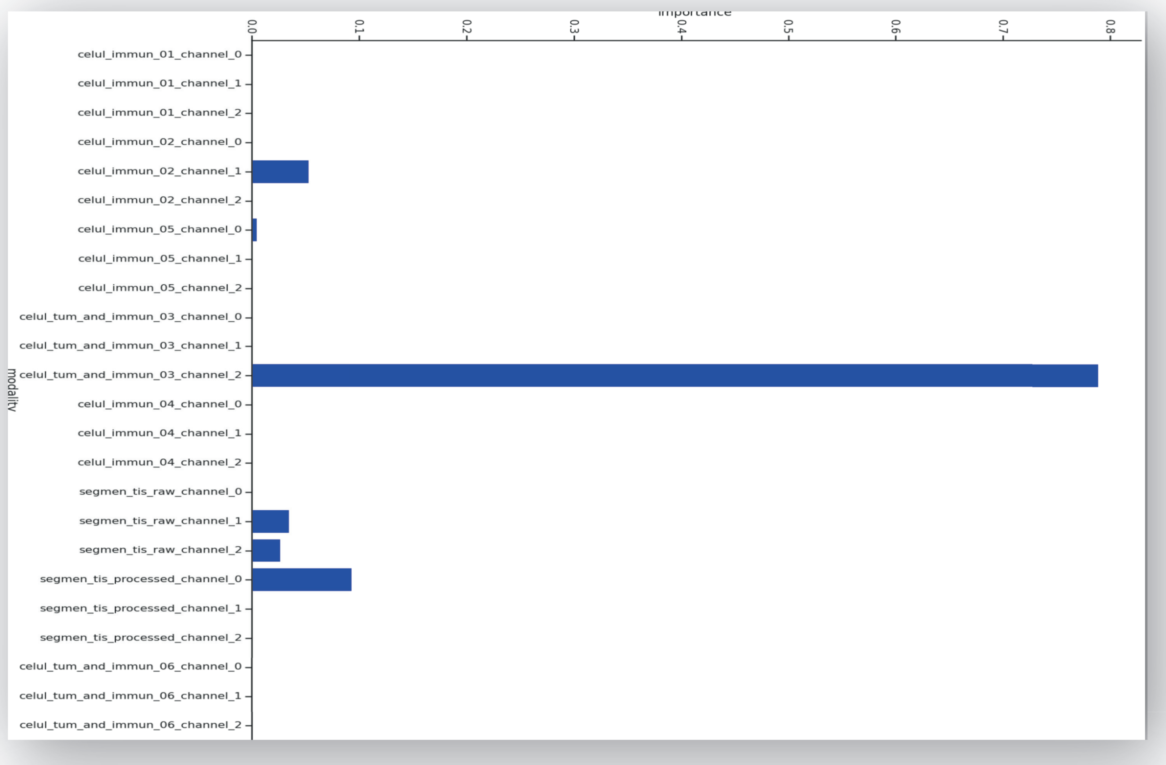
In addition to this, our model takes all the 24 channels into account, where this allows researchers to investigate the correlation between visual markers. The GradCAM algorithm was modified and fine-tuned to conform to these constraints. At the end, the higher attention values on **the pixels on the images indicate better prognostics (higher OS)** for that particular patient whilst taking into account the importance of ensemble of the tabular data.

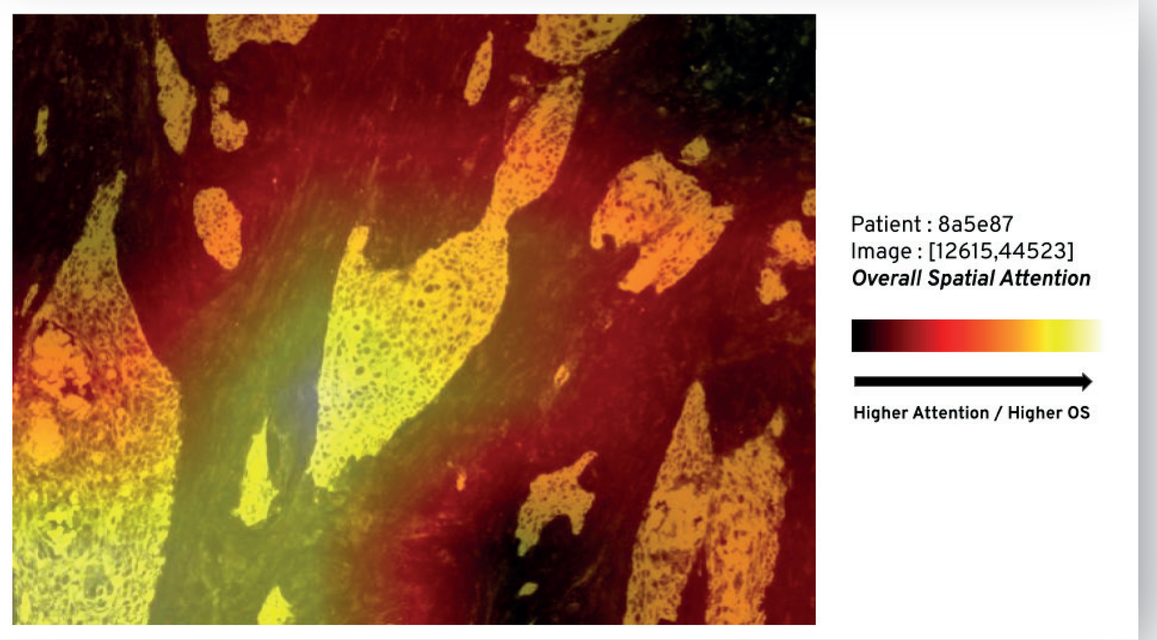
Results

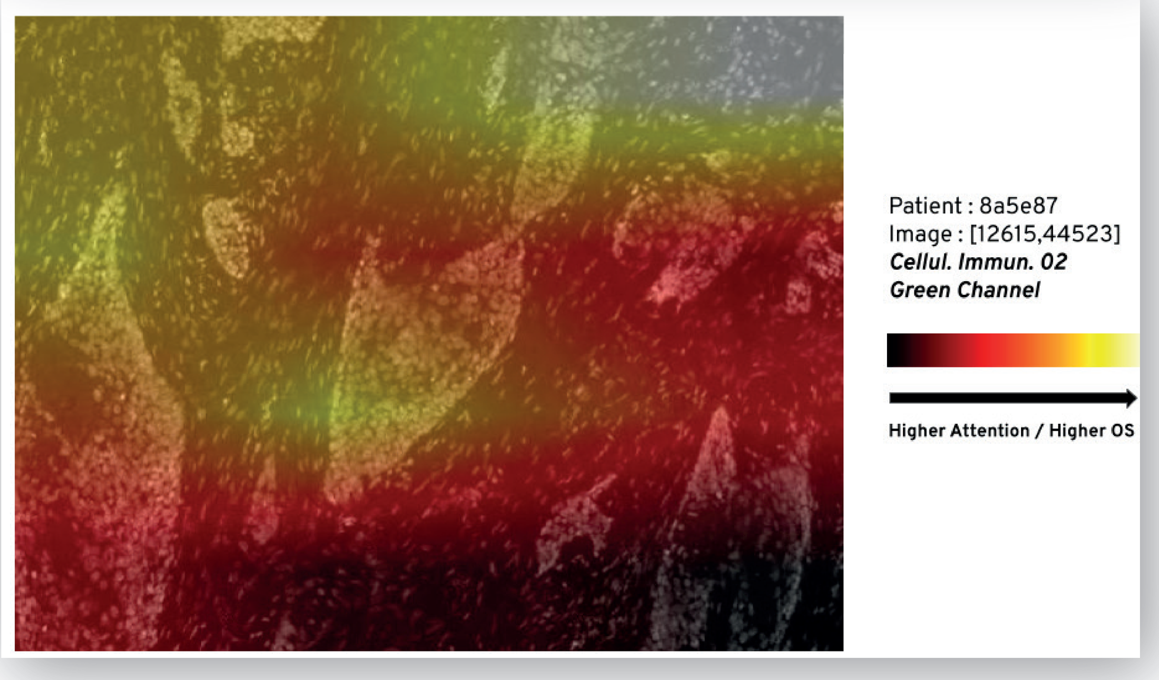
One of the advantages of our framework is the ability **to show the overall attention given on each of the 24 channels of each type of marker** due to the fact that it follows an omni-channel approach, where total positive guided gradient over separate channels can be summed and normalized. At the end, one can understand the ratio of the importance of each channel. This shows how each channel influences a patient's OS (0 : Red, 1 : Green, 2 : Blue), also considering all of his/her clinical data as our deep learning architecture is trained to regress on all the 13 numerical features.

As a further step, after having a general idea on the importance of the contributions of the channels for the overall survivability of the patient, the researcher can inspect the spatial distribution of attention among significant markers, one by one. For each image of each patient, at first we can inspect the overall spatial distribution of attention, averaged across all markers, where brighter pixels correspond to the microenvironmental regions corresponding to higher overall survivability for that particular patient.

Note that, our deep learning model regresses on all tabular data simultaneously **which makes sure that complex correlations behind the evolution of the disease are included in the process.**







Prospects, Conclusion and Future Work

We believe our deep learning architecture constitutes an original and efficient baseline thanks to:

- **Its multi-modal complete nature**
- **The carefully chosen parameters** which conform to the necessities and particularities of the problem.

It was observed that behaviour of the system is consistent between training and test instances, **showing the potential of high degree replicability.** It can be much more improved with continuous collaboration with oncological researchers in future.

Author:
Datategy team

Company:
Datategy
18 bis rue de Villiers
92300 Levallois Perret

Contact Information:
contact@datategy.net
+33 (0)176 542 897
https://datategy.net