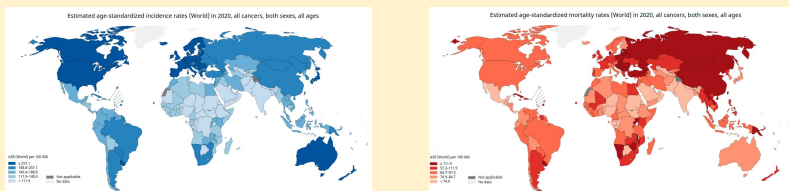


## ORLian:

A novel method to help predict survival from HPV-induced oral cancer, by training mutli-modal interpretable convolutional network based on siamese similarity pretraining model with anonymised patient pathology data and other histological data

## Introduction

While people still are worrying about waves of Covid-19 cases, tsunamis of cancer cases globally are worrying.



Total incidence 2020 est Total mortality  
Incidence and mortality maps from <https://gco.iarc.fr/>

Ear Nose and Throat (ENT) cancers caused by Human Papilloma virus (HPV) spread as a sexually transmitted disease, and we need better tools to prevent mortality in patients & database access for open science.

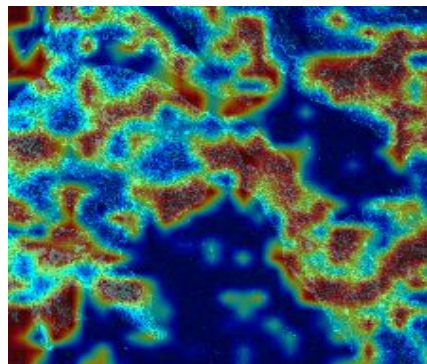
Artificial intelligence may help.

## Results

**Current Score on Macro: 5.766**

Model outperforms many publicly available models such as ResNet-18

Can capture cellular structure in images better than original baselines



GRAD-CAM visualisation

### Feature and Layer Importance:

Through manual evaluation using a medical jury, we find that in terms of layer importance:

**3 & 6 & 2 > 1 > 5 & 4**

Feature importance from the tabular can be ranked as

1. Whether the person has died
2. How old they were
3. Age of diagnosis
4. Severity of condition
5. Gender:

## Methods

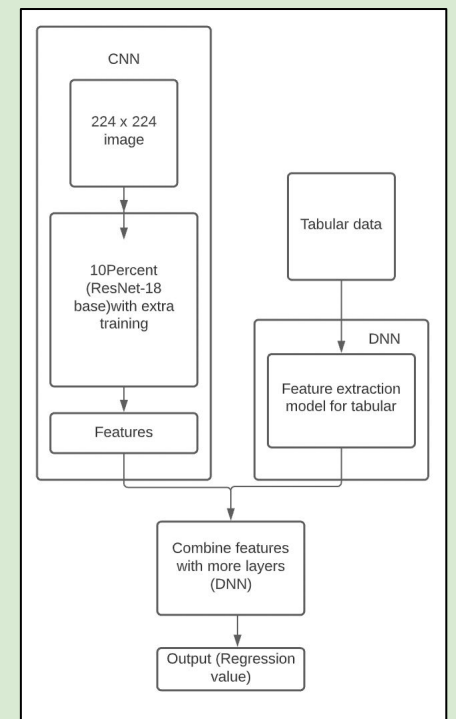
Histological sections from biopsies were used to train the El-nAI-ny model, based on a pre-trained Res-Net\*, then visualized using GRAD-CAM \*\*

Training was done using data stored on Dataiku and generated images were compared back against image stack analyses

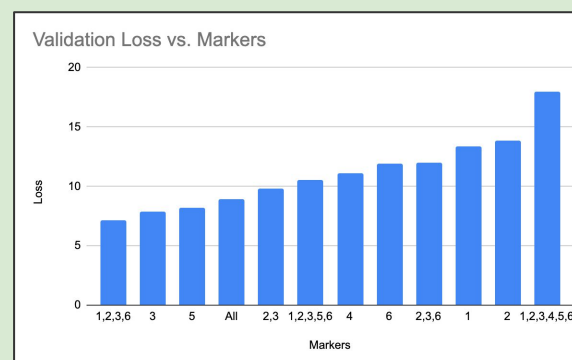
## Evaluation Methods

Model evaluation was done using two methods:

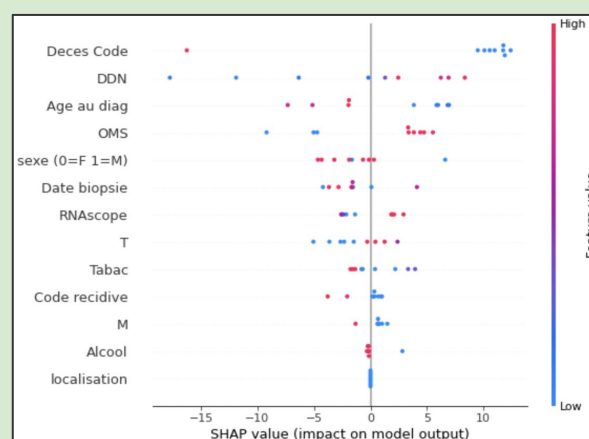
- 1) Using XGBoost and SHAP to determine feature importance from the tabular
- 2) Use Leave One Out to remove one marker feature at a time and checking performance
  - a) Medical expertise evaluate the model and rank their layers too



Model architecture



Leave One Out: Average performance on validation sets across folds for markers



Feature importance from the tabular

## Future Prospects

- Model optimization and explainability
- More clinical data
- Model deployment
- Clinical trials also in light of possibilities for vaccination now

## Acknowledgements

JOGL Epidemium 3 ORL/IA challenge  
@Nicolas Hegerle of JEDHA

**ORLIAN CORE TEAM**  
Vikramaditya Chhapwale  
Omar Osman  
Rachel Aronoff

\* Ciga, O., Xu, T. and Martel, A.L., 2022. Self supervised contrastive learning for digital histopathology. Machine Learning with Applications, 7, p.100198.

\*\* Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D., 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision (pp. 618-626).