1. Provide your team background and organization description (if applicable).

We are two PhD students from the CVN Lab (Centre for Visual Computing), an interdisciplinary research center founded in 2011. The lab is affiliated with INRIA Paris-Saclay through the OPIS joint research project team. The team focuses on developing mathematical models and computational methods for the automatic structuring, processing, and interpretation of large-scale visual data. Our research spans various areas, including optimization, machine learning, inverse problems, and biomedical image analysis.

2. Explain why you participated in the DigiLut challenge.

As PhD students specializing in the histopathology of liver, lung, and breast tissues, we were drawn to the DigiLut challenge for several reasons. While pathology across different organs employs similar analytical techniques, this challenge offers us the opportunity to deepen our expertise in detecting tissue anomalies in lung histopathology. It also allows us to apply the theoretical knowledge we've gained during our PhD studies to real-world issues, such as identifying graft rejection zones. Our strong background with this type of data made the rare and complex task of predicting graft rejection particularly appealing. Additionally, the availability of a large, annotated dataset further encouraged us to tackle this intriguing problem.

3. Describe how you built your winning model and elaborate on the technical and modeling choices you made.

a. Data preprocessing

Our first step was to extract images from the whole slide images (WSIs). We used our own custom-developed library, *prismtoolbox,* to accurately detect the lung tissue from the background within the WSIs and then generate the patches. We chose a patch size of 512x512 pixels.

b. Data processing

To distinguish positive/negative patches we used presence_of_lesions column.

- If presence_of_lesions==0: The patches from the WSI are considered as negatives so we extract 512x512 patches from the detected tissue.
- If presence_of_lesions==1: We will only extract the patches within the bounding box. If the bounding box area is less than 512x512 we resize it to match the needed size. Otherwise, we split the bounding box into windows of the needed size.

After generating the positive/negative patches we opted for 5-Folds Cross-Validation. We split our data into 5 stratified folds based on the presence_of_lesions column to ensure balanced class distribution between +/- patches in our train and test sets.

c. Modeling

To train our 5-folds models:

- We started with some data augmentations techniques (Random horizontal/vertical flips and Gaussian blurring).
- We used a pretrained model, developed by *Ozan Ciga et al* (github code and paper), as a feature extractor and we fine-tuned it by changing the last fully connected layers.
- We used Weighted Cross Entropy combined with label smoothing as loss function and Adam with weight decay as optimizer with learning rate equal to 0.001.
- For each fold, we trained our fine-tuned model for 20 epochs using two GPUs NVIDIA A100-SXM4-80GB.

d. Ensembling

After the training process, we ended up with 5 models each trained and tested on different parts of our global dataset. We chose to use an ensembling of the 5 models to create the predictions with arithmetic mean.

e. Final bounding boxes

Finally, we implemented a simple algorithm to group neighboring positive patches into clusters, and we extracted the bounding boxes coordinates of the largest clusters to construct our predictions. The positive patches were grouped recursively until the number of predicted positive areas matched the number of possible ROIs or all the positive patches were merged.

4. What GPU/CPU/RAM resources you used to build your model

Data preprocessing and processing were done on CPU. Models training was performed on GPU NVIDIA A100-SXM4-80GB