

[width=]14.jpg

Figure 1: Exemple d'une image de trace de sang.

[width=]14\_saliency.png

Figure 2: Exemple d'une image avec la carte de chaleur Grad-CAM superposée.

Afin de répondre à l'aspect "boîte noire" du modèle utilisé, nous avons implémenté de l'interprétabilité dans notre modèle à l'aide de Grad-CAM [?]. Cette méthode permet de fournir une explication visuelle vis-à-vis des décisions de classification issue de notre modèle, permettant ainsi de rajouter une certaine légitimité relative face à nos analyses de traces de sang faisant partie d'un processus judiciaire.

L'algorithme Grad-CAM (Gradient-weighted Class Activation Mapping) est une technique utilisée pour rendre les réseaux de neurones convolutifs (CNN) plus interprétables dans le domaine de la classification. Il fonctionne en générant des cartes de chaleur (heatmaps) qui mettent en évidence les zones importantes d'une image qui contribuent le plus à la prédiction d'une classe spécifique. Pour ce faire, Grad-CAM calcule les dérivées des scores de sortie de la classe cible par rapport aux caractéristiques de la dernière couche convolutive du CNN. Ces dérivées sont ensuite globalement moyennées pour obtenir les poids d'importance de chaque carte de caractéristiques. Enfin, les cartes de caractéristiques sont pondérées par les poids d'importance et combinées pour obtenir la carte de chaleur Grad-CAM, qui peut être superposée à l'image d'origine pour une visualisation plus intuitive. Cette approche permet de mieux comprendre les décisions prises par le CNN et d'améliorer la confiance dans les prédictions du modèle.

[exmple d'une image avec la carte de chaleur Grad-CAM superposée (toute simple pour montrer à quoi ca ressemble)]

Par exemple, nous avons remarqué que parfois certaines des réglottes permettant de donner l'échelle de la photo étaient détectées comme des traces de sang ou induisaient un biais dans la classification. En utilisant Grad-CAM, nous avons pu identifier que le modèle se focalisait sur ces réglottes pour prendre sa décision, ce qui nous a permis de mieux comprendre les erreurs de classification et d'améliorer la qualité des prédictions. [figure montrant une tache avec une réglotte.] Un autre avantage de Grad-CAM est qu'il permet de vérifier si le modèle se focalise sur la bonne tâche de sang dans le cas où plusieurs tâches sont présentes sur une même image.

Dans l'exemple suivant, nous voyons que c'est la tâche de gauche qui est détectée par le modèle, ce qui est confirmé par la carte de chaleur Grad-CAM qui met en évidence les zones importantes de l'image pour cette tâche. [figure montrant 2 images côte à côte avec les cartes de chaleur Grad-CAM correspondantes].