

Rapport Projet 3A

Analyse de trace de sang

Cléa Han, Yanis Labeyrie et Adrien Zabban

Ecole Centrale Méditerranée, 13013 Marseille, France

1 Introduction

Notre projet 3A s'intéresse à l'analyse de traces de sang dans le cadre du travail de l'expert criminalistique Philippe Esperança. En effet, l'objectif est de réaliser une intelligence artificielle pour assister et faciliter le travail d'analyse de scènes de crimes présentant du sang.

Ce projet s'appuie sur un ensemble de données traitées et non traitées de diverses typologies de traces de sang fournies par notre expert en criminologie, ce qui nous a permis de pencher pour une solution s'appuyant sur un apprentissage supervisé.

2 Données

Dans un premier temps, nous avons pu manipuler des données de laboratoire, c'est-à-dire des images de traces de sang reproduites en laboratoire sur des fonds réguliers, hors des scènes de crimes. Ces données de laboratoire représentent un total de 19 classes de traces de sang. Cependant, il y avait la présence d'une classe trop minoritaire parmi ces 19 classes, qui est la classe de trace d'insectes. Cette classe a donc été retirée afin de garder une certaine distribution relativement équilibrée. Nous avons donc pu travailler avec 18 classes avec les données de laboratoire.

Ces données de laboratoire ont été réparties à 80% dans nos données d'entraînement, à 10% dans nos données de validation et à 10% dans nos données de test.

Après l'élaboration de nos éventuels modèles pour la problématique traitée, nous avons pu manipuler des données dites réelles. Ce sont des données prises directement sur les scènes de crimes. Les images sont alors relativement moins consistantes et plus hétérogènes que les données de laboratoire. En effet, ces images sont donc issues de prises réalisées par la police scientifique, qui ne prend pas en compte les conditions consistantes de prise de photo respectées dans les données de laboratoire concoctée par Philippe Esperança.

Ces données réelles ont été réparties à 60% dans nos données d'entraînement, à 10% dans nos données de validation et à 30% dans nos données de test. En effet, une plus grande proportion d'images a été attribuée au test des données réelles afin d'avoir un test relativement plus fiable. Cet ensemble de données de test est composé de 73 images.

3 Modèles

Pour aborder l'analyse de trace de sang, nous avons décidé d'utiliser le modèle Resnet pré-entraîné sur Imagenet, selon diverses approches.

Nous avons dans un premier temps retiré la dernière couche dense du modèle, qui était destiné à classer sur 1000 catégories, puis effectuer du linear probing. Nous avons donc remplacé la dernière couche dense par 2 couches denses pour avoir une dernière couche dense à 18 neurones correspondant aux 18 classes identifiées à classer.

Puis, nous avons également abordé le modèle en le réentraînant sur nos données de laboratoire, c'est-à-dire effectuer du fine-tuning sur nos données.

Ensuite, nous avons aussi essayé d'effectuer du linear probing puis un fine-tuning de manière successive sur le modèle Resnet.

En effet, nous avons testé plusieurs approches vis-à-vis du modèle Resnet pré-entraîné sur Imagenet afin de maximiser nos possibles performances.

Après avoir obtenu ces différents modèles, nous les avons également fine-tuné sur l'ensemble des données réelles à notre disposition.

Afin de maximiser nos potentielles performances, nous avons effectué un Grid-search afin de chercher les meilleurs hyper-paramètres possibles.

4 Interprétabilité

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.

5 Résultats

6 Conclusion

Notre projet s'inscrit dans une volonté d'optimiser le travail d'analyse de traces de sang en criminologie, notamment pour notre expert international Philippe Esperança.

En effet, notre encadrant a pu recevoir notre projet sous la forme d'une interface qui fonctionne localement sur son appareil de travail. En effet, ce modèle d'analyse de trace de sang ne doit pas avoir accès à internet afin d'éviter tout risque d'attaque ou de fuite de données, le faire tourner en local permet de susciter la confiance de notre expert criminalistique. Il a pu apprécier et être satisfait de notre rendu.

Notre projet fait également l'objet d'une documentation afin qu'il puisse être repris aisément à l'avenir.

7 Annexe

Annexe -> Montrer une image de chaque classe