

Stage Master recherche 2024

Estimation de Pose Humaine (HPE) pour les images d'art utilisant des Réseaux Profonds

Le projet de recherche proposé se situe dans le cadre de l'estimation de pose humaine ou **HPE** (Human Pose Estimation) dans des images spécifiques par apprentissage automatique, initiée dans l'équipe IMAP. La HPE est un sujet très actuel, qui vise à fournir une représentation de la pose de personnes présentes sur des images sous forme de squelettes, en se basant sur l'identification de parties spécifiques du corps humain. De nombreuses applications sont possibles, telles que l'interaction homme-machine, l'analyse de mouvements, la réalité augmentée, ... La plus grande partie des travaux s'est cependant focalisée sur des images contemporaines, issues de photographies et/ou de vidéos, et très peu de travaux se sont intéressés à cette problématique sur des images d'art (peintures, gravures, dessins, etc.), qui mettent en défaut les méthodes actuelles du fait de la présence des caractéristiques très différentes (style, effets souhaités par l'artiste, etc.). Récemment de bons résultats ont été obtenus dans notre équipe IMAP, par l'intermédiaire de réseaux profonds, CNNs, [1], [2] pour la détection de régions dans des images d'art, images qui représentent un challenge important (par rapport aux images classiques, photos, ...). Dans la thèse de S. Bengamra [1] nous nous sommes intéressés à cette problématique grâce à l'interprétabilité, en particulier pour caractériser les attributs de bruits issus des « feature maps », et les comparer aux attributs obtenus par les méthodes classiques (saillance, filtrage, masque de bruit) [2][3]. De bons résultats ont été obtenus pour la détection de visages et de mains, et cette méthode doit être étendue à d'autres parties du corps pour proposer une méthode d'HPE de qualité. Nous avons également montré que dans le cas du déploiement visuel, la prédiction des réseaux profonds perdait en précision lorsqu'ils étaient évalués sur des tableaux de différents styles [4].

Le travail proposé consistera dans un premier temps à effectuer un état de l'art des méthodes de HPE pour les images classiques, car les recherches sur ce sujet évoluent très vite (le dernier « survey » a été effectué en 2018, [5]).

Puis une extension de la méthodologie vers des images d'art (de style réaliste) sera effectuée, pour évoluer progressivement vers des images d'art de styles plus complexes. Ce travail permettra la compréhension fine des « feature maps » pour ces images particulières. Une voie possible de compréhension de ces matrices d'attributs (« feature maps ») est le passage par une exploration (ou visualisation) de la décomposition SVD de la matrice aléatoire. La distribution associée à cette décomposition permet ensuite une interprétation des attributs aléatoires (distribution de Marchenko-Pastur, MP). Cette méthode a été exploitée en traitement de signal [6] et en analyse d'images médicales avec succès [7]. Les résultats obtenus permettent la compréhension des attributs aléatoires mais permettent aussi l'optimisation du réseau profond (par action sur la fonction d'erreur par exemple).

L'objectif du travail est d'étudier et développer les outils informatiques, pour visualiser les « feature maps » de chaque type de circuit (« backbones » utilisés) et interpréter ainsi les attributs de bruit issus de ces architectures pour aller vers une HPE fiable.

Encadrants : R. Cozot (Pr / LISIC / ULCO), A. Bigand (MCF HdR / LISIC / ULCO)

Durée : 5 mois

Financement : ANR Fablight

Eléments de bibliographie :

- [1] Contribution à la mise en place de réseaux profonds pour l'étude de tableaux par le biais de l'explicabilité. Application au style Ténébrisme (clair-obscur), Thèse Siwar Bengamra (co-tutelle), soutenue le 30/12/2023.
- [2] **A comprehensive survey on object detection in Visual Art: taxonomy and challenge**, S. Ben gamra, O. Mzoughi, A. Bigand, E. Zagrouba, *Multimedia Tools and Applications*, July 2023, DOI:[10.1007/s11042-023-15968-9](https://doi.org/10.1007/s11042-023-15968-9).
- [3] **Towards explainability in using deep learning for face detection in paintings**, S. Ben gamra, O. Mzoughi, A. Bigand, E. Zagrouba, *ICPRAM23*, 2023
- [4] **Can we accurately predict where we look at paintings ?** O. Le Meur, T. Le Pen and R. Cozot, *PLOS ONE*, Volume 15 (10),pp. 1-20, 2020. Public Library of Science, doi = 10.1371/journal.pone.0239980
- [5] **Deep learning-based human pose estimation : a survey**, C.E.ZHENG *et al.*, *J.ACM* 37, 4, 2018.
- [6] **On the Spectrum of Random Features Maps of High Dimensional Data**, Z. Liao, R. Couillet, *Proc. of the 35th Int.Conf. on ML*, PMLR 80,2018
- [7] **Explainable Deep Learning models in medical Image Analysis**, A. Singh, S. Sengupta,V. Lakshminarayanan, *Journal of Imaging*, 2020, 6.