Proposal 6

Title:
Decoding Visual Intelligence: Enhancing Transparency in Computer Vision through Explainable AI

CONTEXT

Explainable AI (XAI) is an emerging field that aims to provide transparency and understanding of AI decision-making processes. In this thesis, you will explore the application of XAI in Computer Vision, specifically in the context of image segmentation, object detection, and recognition. Deep learning models used in Computer Vision applications are complex and difficult to interpret for lay users. This thesis project will address the increasing demand for transparent and accountable computer vision-based AI systems that can be interpreted and scrutinized by non-experts. This project will require the development of AI models and their corresponding XAI methods that can explain their predictions and decisions in visual recognition tasks such as object detection, image segmentation, and scene understanding. To evaluate the effectiveness of the proposed models and XAI methods, you will need to conduct user studies with lay users from diverse backgrounds. The user studies will involve presenting visual recognition tasks and the corresponding explanations generated by the XAI models to the participants. The user studies can measure the participants' understanding, satisfaction, and trust in the models, as well as their ability to identify potential biases and errors.

AIM

The goal of this project is to design, develop, and evaluate a prototype of an explainable computer vision system.

References:

Explainable AI (XAI) in Image Segmentation in Medicine, Industry, and Other Applications. (2024). arXiv. Available at: https://arxiv.org/html/2405.01636v1

Kar, A. (2024). *Transformers and Explainable AI for Computer Vision*. Manning. Available at: https://www.manning.com/liveprojectseries/transformers-and-explainable-ai-for-computer-vision-ser

The Integration of Explainable AI Methods for the Classification of Medical Image Data. (2024). IEEE Xplore. Available at: https://ieeexplore.ieee.org/document/10645095/?utm_source=chatgpt.com

An eXplainable Image Segmentation and Classification Framework. (2023). MDPI. Available at: https://www.mdpi.com/2079-9292/12/17/3551?utm_source=chatgpt.com

Nguyen, T. T. H., et al. (2024). *LangXAI: Integrating Large Vision Models for Generating Textual Explanations to Enhance Explainability in Visual Perception Tasks*. arXiv. Available at: https://arxiv.org/abs/2402.12525?utm_source=chatgpt.com