UNIVERSIDADE
CATÓLICA
PORTUGUESA
CENTRO REGIONAL DE BRAGA

3º year Bachelor's Degree in Data Science – Deep Learning

Lucas Limpens

Diogo Rodrigues

Face Detection and Localization in Images using Deep Neural Networks

BRAGA

Índice

Introduction

Motivation

- **Face detection** is a critical area in computer vision, with applications in security, surveillance and human-computer interaction. This project aims to develop an innovative face detection and localisation system that works efficiently in a variety of image conditions, using advanced deep neural network techniques. The growing demand for more accurate and adaptable systems highlights the need for innovation in this area, making this work relevant and necessary..

Proposal

- This project proposes to create a completely new model for face detection and localisation. Although established models such as **YOLO** *(You Only Look Once)* and Faster **R-CNN** will serve as the basis of my knowledge, the main focus will be on developing an original approach that goes beyond the limitations of these architectures. By studying how these models work, it will be possible to understand their strengths and weaknesses, which will help to build a system that better adapts to the specifics of the data being used. The performance of the model will be evaluated using metrics such as accuracy and Intersection over Union (IoU) to validate its effectiveness.

Related Works

1. YOLO (You Only Look Once) for Real-Time Detection

- YOLO is a popular architecture for real-time detection due to its speed and accuracy. Detection takes place in a single pass through the image, making it efficient for low-latency applications. Although we intend to use the basic ideas of YOLO to build our model, our focus will be on developing an approach that improves accuracy in different environments.

2. Faster R-CNN for Object Detection

- Faster R-CNN is one of the most accurate methods for object detection, using a region proposal approach and CNNs. By studying its structure, we intend to learn how to improve our model in terms of accuracy, but adapt the technique for a specific focus on face detection.

3. SSD (Single Shot Multibox Detector)

- The SSD enables fast recognition, using a multi-box approach to deal with objects of different sizes. Analysis of how this works will be crucial to understanding how to deal with the variety of faces in different lighting conditions and angles.

4. Data Augmentation for Deep Learning

- Data augmentation methods are essential in computer vision, as they increase the size of the data set and improve the robustness of the model. By exploring specific techniques (such as flipping, rotation and brightness adjustment), our project will seek to understand how these practices can affect the accuracy of face recognition.

References

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. European Conference on Computer Vision.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data.