Arquiteturas de *Deep Learning* para a deteção de edifícios em Imagens Aéreas de Muito Alta Resolução

Desenvolvimento e Avaliação de arquiteturas de *Deep Learning* para Ambientes Urbanos

N.º do tema de sessão técnica: Nº. 4 - Inteligência artificial, Large Language Model, Deep Learning, etc.

Resumo

A deteção de edifícios em imagens aéreas de muito alta resolução espacial é fundamental para a criação de mapas urbanos e planeamento urbano. Normalmente, este processo é trabalhoso e propenso a erros quando realizado manualmente. Tendo por base o mencionado, o INESC TEC e a Equipa Multidisciplinar Inteligência Geoespacial (GeoInt) da Direção-Geral do Território (DGT) desenvolveram um estudo com o objetivo de explorar diferentes arquiteturas de *Deep Learning* para a deteção de edifícios nas coberturas aéreas da DGT, a fim de automatizar os processos mencionados de forma eficiente e reduzir possíveis erros. Diferentes abordagens de classificação de imagens baseadas em objetos e em pixels foram implementadas e exploradas, através da utilização da técnica de *Transfer Learning*. Dois conjuntos de dados originais foram desenvolvidos em parceria com a DGT. Para permitir um processo de aprendizagem supervisionada, um conjunto de dados foi anotado utilizando *pseudo-labelling*, enquanto o outro foi anotado utilizando dados extraídos do levantamento LiDAR de Portugal Continental de 2024. As imagens disponibilizadas para este estudo são ortofotos de diferentes municípios portugueses, datadas de 2018 e 2021, compostas por quatro bandas espectrais (RGB e infravermelho) com resolução espacial de 25 cm. O desempenho do modelo foi estudado com base em diversas métricas estatísticas, avaliando o processo de treino e teste. Um estudo comparativo foi realizado entre os desempenhos dos modelos, tendo sido possível concluir que a abordagem mais eficaz e eficiente foi o modelo de classificação baseado em pixels, SegFormer. O modelo apresentou uma Intersection over Union (IoU) média de 88,2% em todos os municípios analisados e resultados qualitativos extremamente precisos. Os resultados qualitativos consistiram na análise visual dos resultados produzidos pelos modelos, relatada pelos analistas da DGT. A metodologia está a ser adaptada para Computação de Alto Desempenho (HPC) na Fábrica de Inteligência Artificial do BSC (BSC AI Factory), que é uma iniciativa cofinanciada pela EuroHPC Joint Undertaking e dinamizada em Portugal pela Fundação para a Ciência e a Tecnologia (FCT) com o apoio do Centro Nacional de Computação Avançada (CNCA). O objetivo final é usar a metodologia para detetar os edifícios de Portugal continental observados nas imagens aéreas da DGT mais recentes, os ortofotos de 2025 a disponibilizar brevemente.

Palavras chave

Deteção de Edifícios, Mapeamento Urbano, *Deep Learning*, Imagens Aéreas de Muito Alta Resolução Espacial, *Semantic Segmentation*

Autores

Diana Cardoso
up202412203@edu.fe.up.pt
Faculdade de Engenharia da Universidade do Porto (FEUP), Portugal

João Gama
joão.gama@inesctec.pt
INESC TEC, 4200-465 Porto, Portugal

Carlos Ferreira
cgf@isep.ipp.pt
Instituto Superior de Engenharia do Porto - Instituto Politécnico do Porto (ISEP), Portugal

Rui Camacho
rcamacho@edu.fe.up.pt
Faculdade de Engenharia da Universidade do Porto (FEUP), Portugal

Hugo Costa
hcosta@dgterritorio.pt
Direção-Geral do Território (DGT), Portugal

Pedro Benevides
pbenevides@dgterritorio.pt
Direção-Geral do Território (DGT), Portugal

Ricardo Pinho
rpinho@dgterritorio.pt
Direção-Geral do Território (DGT), Portugal

Mário Caetano
mario.caetano@dgterritorio.pt
Direção-Geral do Território (DGT), Portugal