

DETECÇÃO DE PEDESTRES USANDO DEEP LEARNING EM VANTS

Fernando Vieira (f02_@hotmail.com.br)

Willian Paraguassu Amorim (willianamorim@ufgd.edu.br)

Detecção de pedestres é um problema muito abordado na atualidade, possuindo diversas aplicações com potencial para melhorar a qualidade de vida da sociedade. Algumas dessas aplicações se encontram nas áreas de sistemas de vigilância, direção autônoma e robótica. Atualmente existe um grande número de pesquisas envolvendo este assunto, trazendo muitas ramificações ao estado da arte no que diz respeito a detecção de pedestres. Mas apesar de muitos destes trabalhos terem atingido resultados interessantes, a dificuldade de aprimorar tais técnicas residia principalmente na utilização de bons descritores. Recentemente, técnicas de aprendizado de máquina baseadas em aprendizagem profunda (deep learning) vêm ganhando atenção da comunidade científica devido aos bons resultados que tais redes apresentaram em problemas clássicos de visão computacional. Este trabalho tem como objetivo realizar a detecção de pedestres usando deep learning, a partir de imagens capturadas de Veículos Aéreos Não Tripulados (VANTS). Basicamente essa tarefa é dividida em duas etapas: (i) segmentação da imagem para a seleção de regiões onde possivelmente existem pedestres, e (ii) classificação dessas regiões em imagens de pedestres ou não pedestres, utilizando para tal uma Rede Neural Convolutiva (RNC), tipo de rede profunda amplamente utilizada para reconhecimento de objetos em imagens e vídeos. As RNCs são redes do tipo feedforward e seu formato segue a seguinte estrutura de camadas: camadas convolucionais, camadas de subamostragem e a camada totalmente conectada que podem ser dispostas de diversas formas a depender do problema abordado. Os experimentos foram realizados na base de dados Town Centre, sobre 5 diferentes arquiteturas: Inception-V3, ResNet-50, VGG-16, VGG-19 e Xception, usando fine tuning e transfer learning. Os resultados mostram que as taxas de acurácia se mantiveram acima de 92% com um tempo de treinamento estável considerando a eficácia da detecção. As melhores taxas de acurácia foram extraídas das estratégias que envolvem a aplicação de fine tuning entre 75% e 100% alternadamente, com maior destaque para a estratégia de fine tuning com 75% de aplicação dos pesos sob a arquitetura ResNet-50.