

Desafio Prático de Machine Learning

Algoritmos de Visão Computacional desempenham um papel fundamental na navegação e execução de tarefas utilizando veículos aéreos não tripulados, também conhecidos como drones. Essa tecnologia permite que os drones processem e interpretarem informações visuais captadas por câmeras embarcadas, fornecendo uma percepção detalhada do ambiente ao seu redor. Com o auxílio desses algoritmos, os drones podem detectar obstáculos, reconhecer objetos, identificar áreas de interesse e até mesmo realizar tarefas complexas, como a entrega de pacotes ou inspeções de infraestruturas.

A detecção precisa de torres de energia e linhas de força durante o voo de drones é essencial para evitar acidentes ou colisões. O treinamento de modelos de Aprendizado Profundo permite que os drones detectem e reconheçam esses elementos, permitindo o planejamento de rotas seguras e evitando colisões potencialmente perigosas. Assim, a capacidade de detecção de objetos é fundamental para a segurança e a confiabilidade das operações aéreas não tripuladas próximas a infraestruturas elétricas.

Problemas de detecção de objetos envolvem a identificação e localização de objetos específicos em imagens ou vídeos. Para treinar modelos, os dados devem ser estruturados com anotações precisas, indicando as coordenadas dos objetos de interesse. As anotações podem ser fornecidas por meio de bounding boxes (caixas delimitadoras retangulares), e.g. em detecção de pedestres, as anotações podem ser bounding boxes que cercam as pessoas. A estruturação adequada dos dados de treinamento é fundamental para ensinar os modelos a reconhecer e localizar objetos com precisão em diferentes contextos visuais.

Nesta tarefa, espera-se a concepção de um modelo de Aprendizagem Profunda capaz de analisar uma imagem e detectar torres de energia e linhas de energia. Para isso, será utilizado o banco de dados [TTPLA: An Aerial-Image Dataset for Detection and Segmentation of Transmission Towers and Power Lines](#), que contém 1,1 mil imagens aéreas rotuladas. O banco de dados pode ser acessado a partir do Github utilizando o link acima. Siga as instruções do repositório para obter os rótulos relativos à detecção de objetos.



Figura: Segmentação por instância de torres de energia e linhas de força.

Atividades:

1. Escreva um algoritmo em Python que realize o treinamento de um modelo para a detecção de torres de energia e linhas de energia em imagens.
2. Siga as instruções do repositório para extração de rótulos no formato [COCO dataset](#). Utilize imagens de tamanho 640 x 360.
3. Respeite a divisão da base de dados quanto aos conjuntos de treino, validação e teste.
4. Atente-se a um código limpo, organizado, documentado e com ideias claras da solução proposta.
5. Caso mais de um modelo seja produzido, comente sobre os resultados obtidos, explicitando qual deles funcionou melhor e o porquê.
6. Na falta de uma GPU para o treinamento, recomenda-se o uso do Kaggle ou do Google Colab.

Instruções de entrega:

- A entrega será feita por meio do site GitHub (você poderá utilizar a sua conta pessoal ou criar uma nova no site).
- Prepare uma classe para encapsular o modelo. Essa classe deve carregar o modelo e disponibilizar uma função para realização de inferências.
- Você deverá “subir” as soluções para lá até a data especificada no e-mail em que foi enviado esta prova.
- Não é necessário “subir” o modelo treinado para o GitHub. Arquivos relacionados ao modelo podem ser disponibilizados pelo Google Drive.
- Você deverá enviar um e-mail (em resposta ao e-mail em que foi enviado a prova para você) informando que já subiu a solução para o Github. Neste e-mail você deve informar o endereço de acesso do seu repositório.