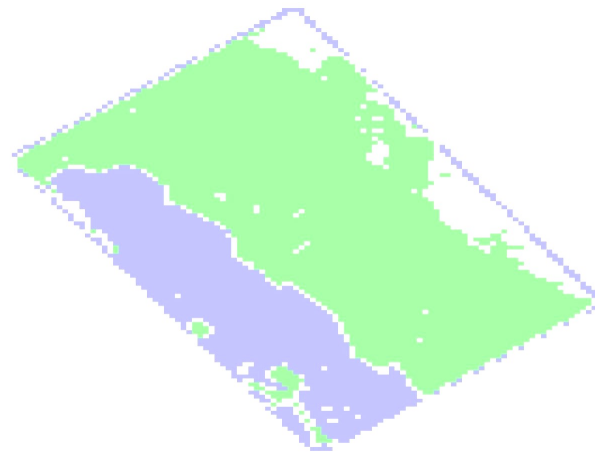
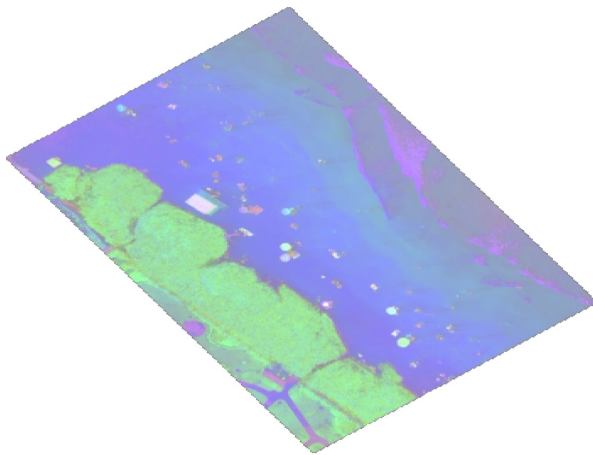


INE410121-41000025DO/ME (20242) - Visão Computacional

Análise de Áreas Costeiras com Ortoimagens de Drones: Segmentação e Classificação



Aluno: Jaceguay Zukoski

Introdução e contextualização

Este estudo aplica visão computacional a ortoimagens de drones para análise costeira. São comparados métodos de segmentação por cores e aprendizado profundo, destacando suas contribuições e limitações na gestão ambiental.

A ocupação humana altera vegetação e areia em áreas costeiras. Ortoimagens georreferenciadas permitem monitoramento preciso, orientando ações de manejo e conservação.



Figura 1 - Exemplo da progressão da área de restinga ao longo do tempo, anos 2004 até 2022. (fonte: Google Earth)

Segmentação por Cores

A segmentação por cores via ajustes de matiz, saturação e valor (HSV) evidenciou ser simples e rápida, permitindo extrair áreas vegetadas com relativa precisão.

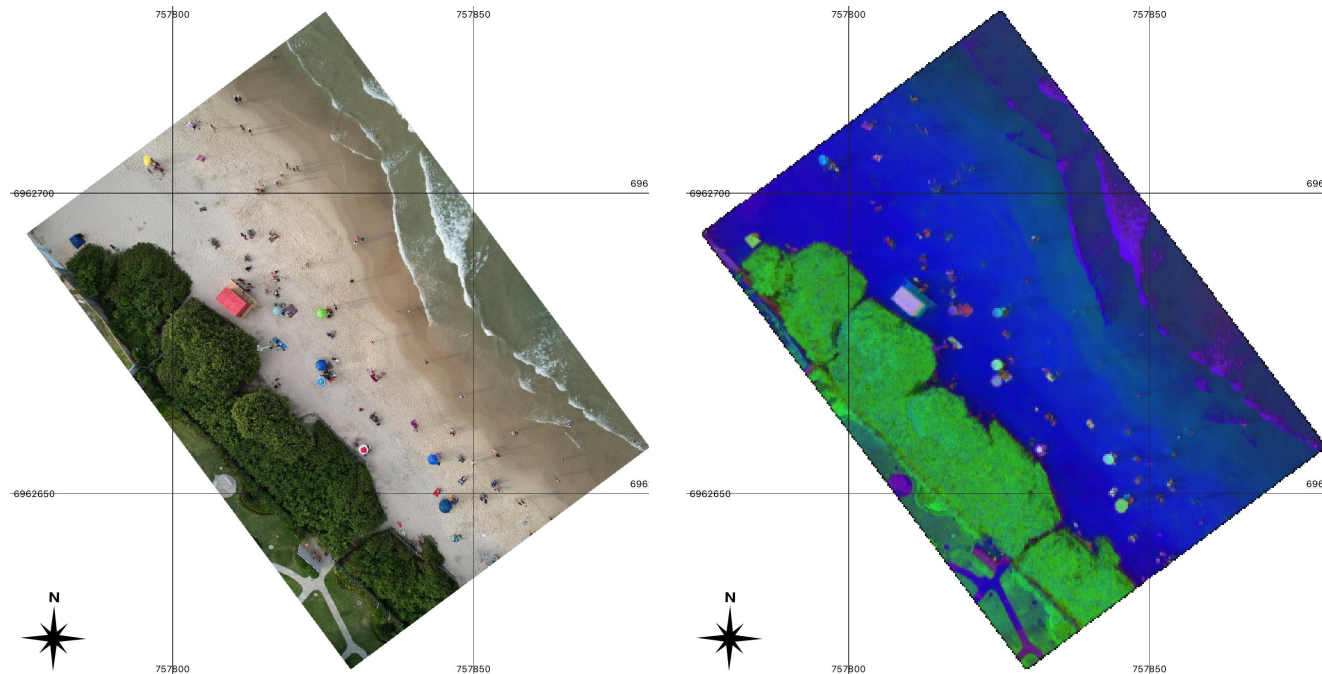


Figura 2 e 3 - Transformação da imagem do modelo de cores RGB para Hue (Matiz), Saturation (Saturação) e Value (Valor), tornando mais simples identificar e isolar tonalidades específicas. (Fonte: autor)



Figura 4 - Ferramenta utilizada para explorar intervalos no espaço de de cores adequado a cada cobertura do solo.

Segmentação por Cores

Enquanto a máscara criada para segmentar a restinga apresentou resultados satisfatórios, foi possível calcular a área de 2241,28 m² ao multiplicar o número de pixels pela resolução espacial do levantamento (0,016 m/pixel). Já na faixa de areia, a presença de objetos heterogêneos dificultou a obtenção de máscaras puras, evidenciando a sensibilidade do método a variações cromáticas e interferências visuais inesperadas.

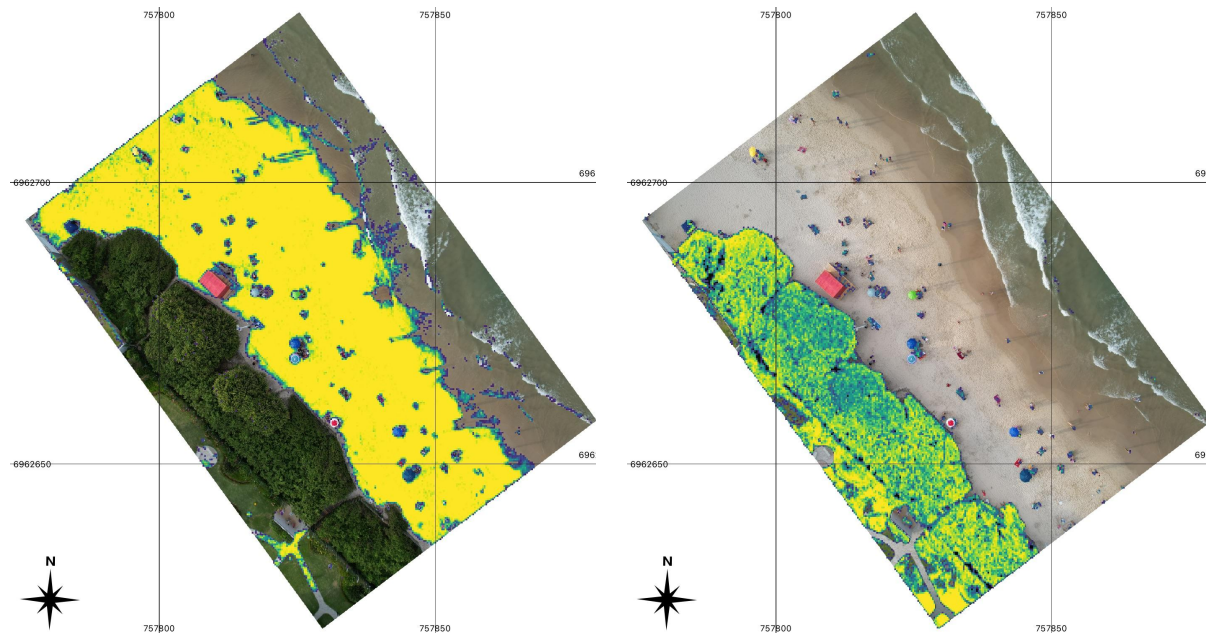


Figura 5 e 6 - Máscaras criadas para cada classe, areia e restinga.

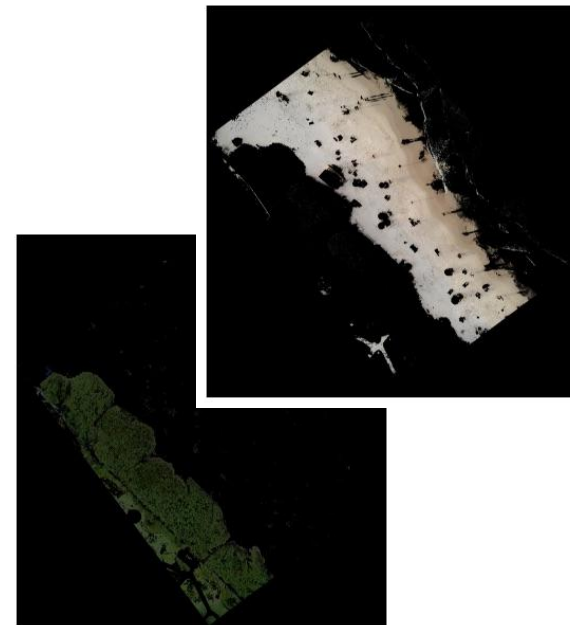


Figura 7 - Imagens obtidas a partir da interseção das máscaras com a imagem original.

Ajuste fino de um modelo pré-treinado (ResNet18)

A criação do conjunto de treinamento se aproveitou da escala espacial das ortofotografias, com uma grade de 1x1 metro usada para selecionar manualmente quadrantes de restinga e areia. Um script cortou as imagens com base nesses quadrantes, organizando os recortes em pastas por classe.



Figura 8 - Grade selecionada sobre as classes de interesse. (Fonte: autor)



Figura 9 - Amostras salvas em geotiffs em pastas distintas, representando áreas de 1x1 metro, resolução de 256x256 pixels. (Fonte: autor)

Ajuste fino de um modelo pré-treinado (ResNet18)

Uma ResNet18 pré-treinada no ImageNet foi ajustada para as classes do problema e treinada por 10 épocas com o otimizador Adam (0,001) e CrossEntropyLoss. Técnicas de *data augmentation* foram aplicadas para diversificar as amostras. O modelo alcançou rápida queda no *loss* e alta acurácia (~0,98), atribuída à boa representação prévia no modelo base e à baixa complexidade do problema.

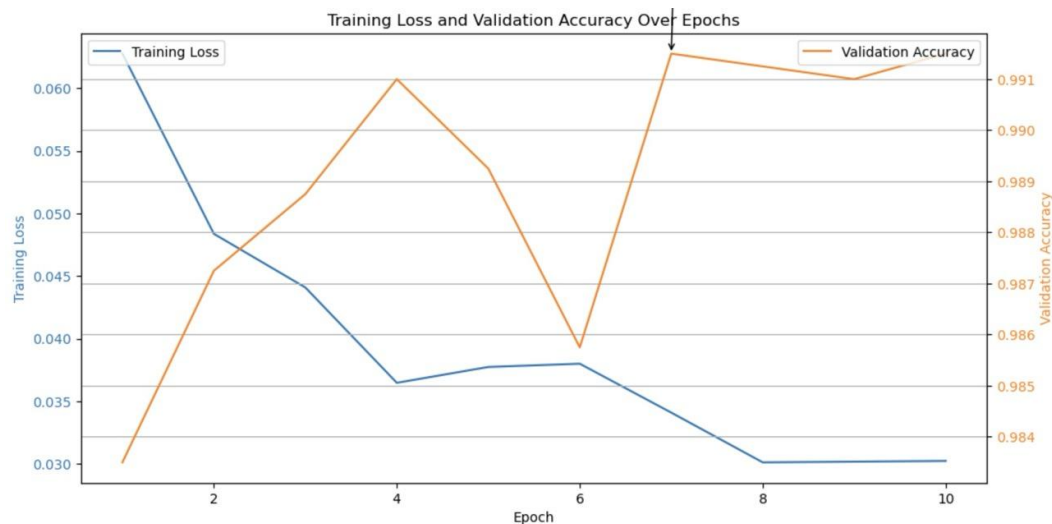


Figura 10 - Evolução do Loss e da Acurácia. (Fonte: autor)

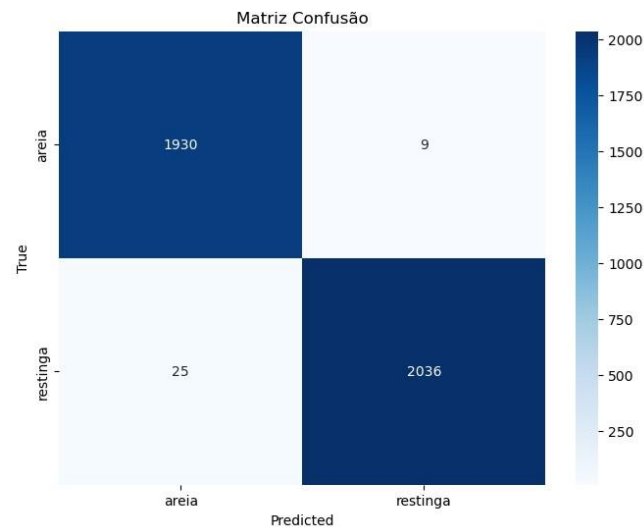
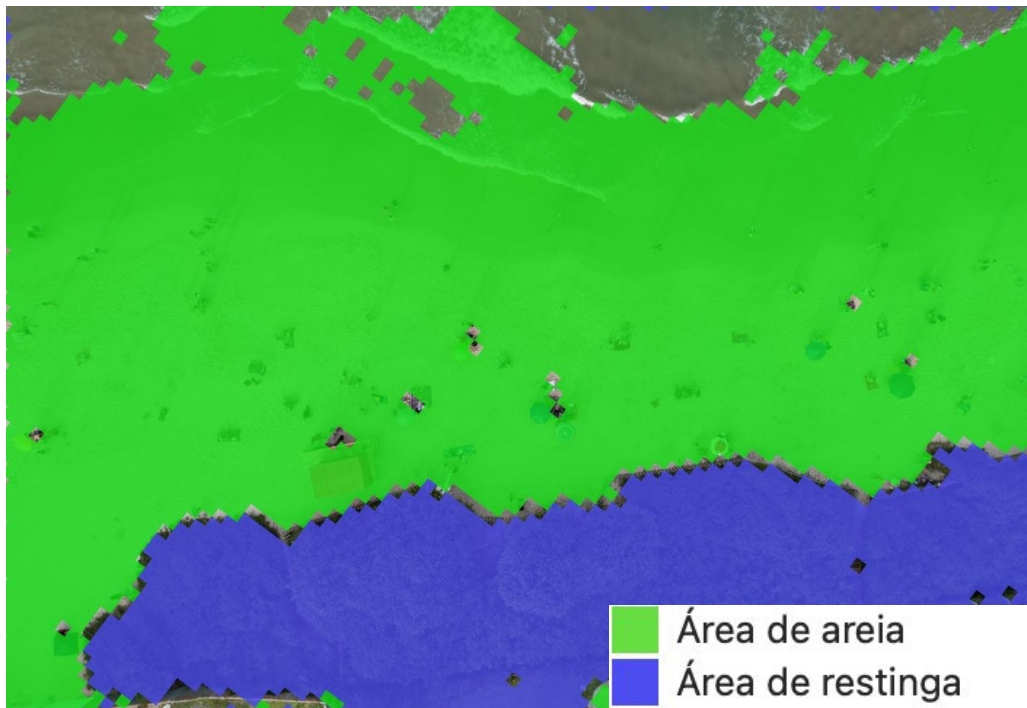


Figura 11 - Matriz de confusão. (Fonte: autor)

Ajuste fino de um modelo pré-treinado (ResNet18)

A aplicação do modelo treinado seguiu os mesmos critérios adotados durante o treinamento. A imagem avaliada foi dividida em trechos de 1x1 metro, e cada trecho foi classificado em relação às duas classes definidas. O resultado final é uma saída com informações organizadas nesse nível de granularidade.



Foram identificados trechos na área de areia que não foram corretamente classificados, este efeito foi atribuído à limitação na variedade das amostras de treinamento. A ampla diversidade de interferências presentes na faixa de areia não foi adequadamente representada nas amostras utilizadas durante o treinamento.

Figura 12 - Predição das áreas sobre cada classe em uma nova imagem, . (Fonte: autor)

Modelo YOLO para Detecção de Pessoas

Para contornar a limitação em lidar com imagens de alta resolução (6278×6631 pixels), comuns em ortoimagens, o código dividiu a ortoimagem em blocos (tiles) menores utilizando a biblioteca rasterio, de modo que cada tile pudesse ser processado separadamente. Em seguida, o YOLO foi aplicado sobre cada tile, fornecendo coordenadas dos objetos detectados.



O modelo enfrentou dificuldades diante do ângulo de visão vertical e deformações introduzidas pela fotogrametria, resultando em detecções incompletas. Ajustar um conjunto de treinamento para refletir essas condições ajudaria o YOLO a se adaptar melhor ao contexto real.

 Pessoas detectadas

Figura 13 - Detecção das pessoas na imagem, o resultado foi tratado de modo a gerar um arquivo georreferenciado (.geojson). (Fonte: autor)

Conclusão

A segmentação por cores, embora simples, é limitada, enquanto o aprendizado profundo oferece maior robustez e capacidade de abstração. A detecção de pessoas com YOLO é promissora, mas requer ajustes para lidar com especificidades das imagens.

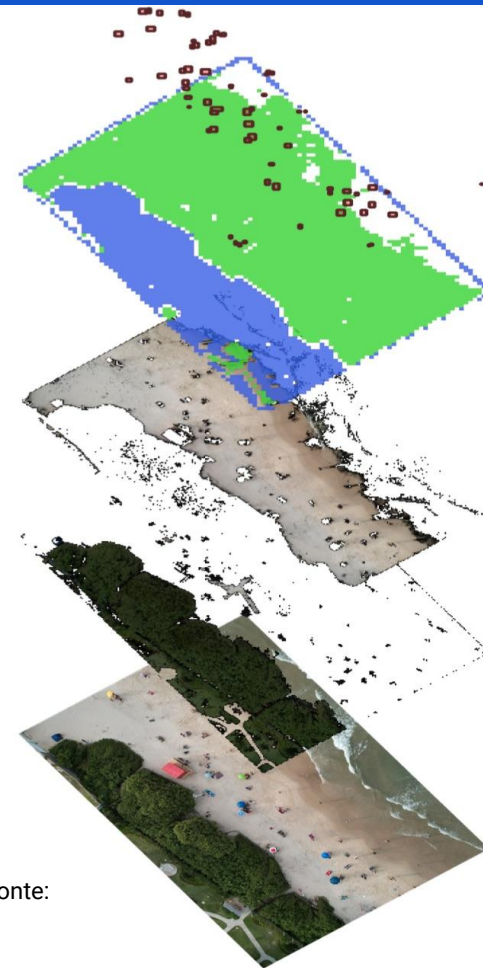
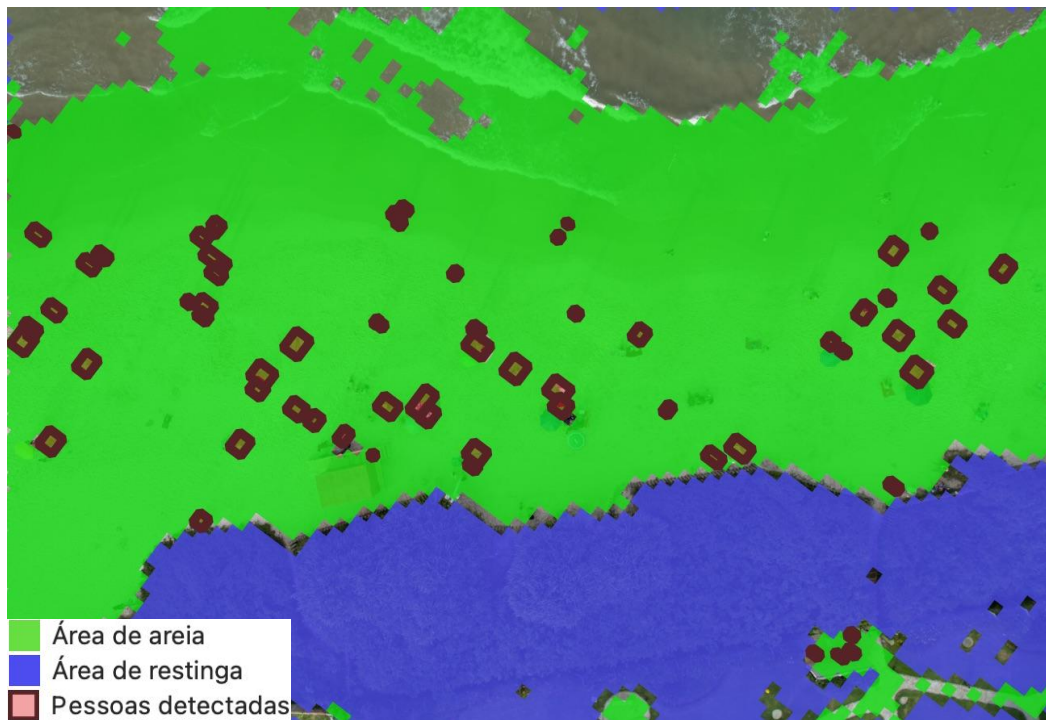


Figura 14 e 15 -
Sobreposição dos
produtos obtidos. (Fonte:
autor)

Obrigado