

ARTIGO ORIGINAL

Estimativa da Altura da Vegetação à Beira da Estrada para Tomada de Decisão de Manutenção

Estimating Roadside Vegetation Height for Maintenance Decision-Making

Mateus Reis Santos ,¹ Bruna Rocha Aguiar ,^{1,2} Enzo Rigazzo Oliveira ,^{1,2}
Thiago Meirelles Ventura ,² Raoni Florentino da Silva Teixeira ,²

¹Nova Rota do Oeste, ²Universidade Federal de Mato Grosso

mateussantos@rotadooeste.com.br; brunaaguiar@rotadooeste.com.br; enzooliveira@rotadooeste.com.br; *thiago@ic.ufmt.br;
raoni.teixeira@ufmt.br

Recebido: 02/10/2024. Revisado: 25/11/2025. Aceito: 30/11/2025.

Resumo

A manutenção da vegetação à beira de rodovias serve para garantir a segurança dos motoristas e pedestres. Entretanto, a detecção da necessidade dessa manutenção costuma ser demorada, custosa e sujeita a falha de anotações. Assim, este trabalho propõe o desenvolvimento de uma solução capaz de estimar a altura da vegetação ao longo das margens de estradas de forma automatizada. Um modelo de aprendizado de máquina foi treinado e testado com um conjunto de dados anotados manualmente e emprega uma arquitetura de rede neural convolucional, adaptada para a tarefa de classificação das alturas da vegetação. Os resultados indicam que o modelo consegue detectar as diferentes classes de altura de vegetação, com baixos índices de erro, evidenciando seu potencial para automatizar o processo de tomada de decisão para roçagem da vegetação. Este estudo contribui para o aprimoramento das técnicas de monitoramento viário, proporcionando maior eficiência operacional e redução de custos para a manutenção de rodovias.

Palavras-Chave: Aprendizado de máquina; Manutenção viária; Monitoramento de rodovias; Redes neurais convolucionais

Abstract

The upkeep of roadside vegetation is essential for ensuring the safety of both motorists and pedestrians. However, identifying the necessity for such maintenance is frequently a time-consuming and costly process, with a high potential for errors in annotation. This work therefore proposes the development of a solution capable of estimating the height of vegetation along roadsides in an automated manner. A machine learning model was developed and evaluated using a dataset of manually annotated data. The model employs a convolutional neural network architecture, adapted for the task of classifying vegetation heights. The results demonstrate that the model is able to detect the different vegetation height classes with low error rates, indicating its potential for automating the decision-making process for mowing vegetation. This study contributes to the advancement of road monitoring techniques, providing greater operational efficiency and reducing costs for road maintenance.

Keywords: Convolutional neural networks; Machine learning; Road maintenance; Road monitoring

1 Introdução

A manutenção da vegetação à beira de rodovias é uma tarefa essencial para garantir a segurança dos usuários das estradas, preservando a visibilidade e prevenindo acidentes causados por vegetação excessiva ou não controlada. As concessionárias de rodovias são frequentemente responsáveis por monitorar e tomar decisões sobre a roçagem da vegetação, tarefa que envolve custos operacionais e logísticos significativos. Atualmente, esse processo é normalmente realizado de forma manual, o que pode levar a inconsistências nas observações e atrasos na execução das atividades de manutenção.

Neste contexto, a estimativa automatizada da altura da vegetação ao longo das rodovias emerge como uma solução promissora para otimizar a tomada de decisões de manutenção. Inclusive, detecções automáticas de componentes relacionados à rodovia já foram realizadas, como o trabalho de [Bezerra et al. \(2023\)](#) que propôs um novo algoritmo para a detecção de anomalias, como buracos e rachaduras, utilizando processamento morfológico de imagens e conceitos estatísticos de intervalos de confiança. Ou como o trabalho de [Novello et al. \(2021\)](#) no qual aborda o controle de veículos autônomos combinando redes neurais convolucionais e recorrentes para mapear imagens e dados de velocidade em comandos de direção. Ou mesmo o estudo de [Buchinger and Silva \(2014\)](#) sobre ilhas de calor utilizando imagens dos satélites para analisar a temperatura de superfície do solo e o índice de vegetação normalizada diferenciada.

No entanto, em se tratando de estimativa de vegetação, desenvolver modelos de Inteligência Artificial (IA) que possam realizar essa tarefa com precisão é um desafio significativo, dada a variabilidade das condições ambientais, a diversidade da vegetação e as particularidades das imagens capturadas em campo. Considerando esses fatores, o objetivo deste trabalho é desenvolver uma solução para auxiliar na detecção em quais trechos da rodovia deve ser acionada a equipe de manutenção. Para tanto, é necessário a criação de um modelo de aprendizado de máquina capaz de estimar a altura da vegetação à beira da estrada, utilizando imagens capturadas em alta resolução ao longo dos trechos de uma rodovia, e usar tais informações para classificar um trecho como um todo. Com isso, busca-se automatizar o processo de tomada de decisão para roçagem, proporcionando maior eficiência e economia.

Este artigo está estruturado da seguinte forma: na [Seção 2](#), são apresentados os trabalhos correlatos, destacando as principais abordagens existentes para a detecção e classificação da vegetação. A [Seção 3](#) descreve os materiais e métodos utilizados no desenvolvimento do modelo, incluindo a coleta de dados, o processo de anotação das imagens e a arquitetura da rede neural empregada. Na [Seção 4](#), são apresentados os resultados obtidos, com uma análise detalhada das métricas de desempenho do modelo. Finalmente, a [Seção 5](#) traz as conclusões do trabalho e as direções futuras para a pesquisa.

2 Trabalhos correlatos

Como mencionado, o monitoramento da vegetação ao longo das rodovias é crucial para garantir a segurança de

motoristas, pedestres e animais. Por isso, diversos estudos foram conduzidos com o objetivo de inovar as técnicas de monitoramento. A [Tabela 1](#) mostra um quadro resumo dos principais trabalhos levantados e suas respectivas características.

Esses trabalhos mostram avanços significativos no uso devisão computacional para o monitoramento ambiental, mas ainda com limitações de generalização, volume de dados e automatização completa do processo. A solução aqui proposta diferencia-se por combinar um grande volume de imagens reais de rodovia, anotações manuais consistentes e um modelo otimizado para priorização de manutenção.

No trabalho de [Harbas and Subasic \(2014\)](#) foi proposto o desenvolvimento de um método para detectar vegetação na margem de rodovias. A abordagem envolve a classificação de características através de um vetor que combina componentes de cor de diferentes espectros e coeficientes da função de Wavelet. A base de dados utilizada nesta pesquisa foi feita por gravação própria nas margens de rodovias, no qual 270 imagens foram gravadas com uma filmadora HD Canon XF100 de um veículo em movimento. A taxa de precisão alcançada foi de 93,32%. No entanto, o fluxo óptico requer recursos computacionais significativos e a seleção da área de interesse limita as variações na aparência da vegetação.

O artigo de [Chowdhury et al. \(2015\)](#) propõe o uso de recurso de textura baseado em sistema especialista para identificação de vegetação. A técnica de classificação combina a extração de características de textura com classificadores como SVM e Redes Neurais. A base de dados utilizada é composta por 110 imagens de resolução 900 x 500 pixels coloridas. O método *Co-occurrence of Binary Pattern* (CBP) foi utilizado para extrair características relevantes, alcançando uma precisão de 92,72% com validação cruzada. A técnica proposta foi aplicada e avaliada em dois tipos de imagens de vegetação, gramíneas densas e esparsas. No entanto, o método é aplicável apenas em imagens de vídeo e imagens recortadas.

Em [Zhang et al. \(2016\)](#) é proposto um novo modelo de superpixel contextual espacial para classificação de vegetação em imagens naturais de beira de estrada. O modelo atua transformando a tarefa de classificação de um pixel em um domínio de superpixel para adoção mais efetiva de informações contextuais espaciais locais e globais entre superpixels em uma imagem. Atingiu-se 77,44% de precisão em um conjunto de dados com 50 imagens e houve dificuldades na classificação do solo que é frequentemente confundido com grama marrom.

Já em [Zhang et al. \(2015\)](#), propuseram a classificação de vegetação usando uma fusão de recursos de cor e textura. Os recursos de intensidade de cor são extraídos no espaço de cores aparente, enquanto a textura compreende três momentos de cor. Obteve-se uma taxa de precisão de 79% para o mesmo conjunto de dados da abordagem anterior. No entanto, as informações de cor apresentam dificuldade para distinguir a vegetação natural, particularmente aquelas com textura muito similar, continuando a gerar erros de classificação.

O método utilizado no artigo de [Harbaš et al. \(2018\)](#) usa recursos de imagem do espectro visível, permitindo o uso de câmeras coloridas comuns a bordo. O conjunto

Tabela 1: Comparação entre trabalhos relacionados e a solução proposta.

Trabalho	Abordagem	Métodos	Dataset	Resultados
Harbas and Subasic (2014)	Detecção de vegetação em margens de rodovia	Classificação por vetor de características (cor + wavelet)	270 imagens (1920×1080) capturadas com filmadora HD	Alta precisão (93,32%), porém alto custo computacional e limitação na variação da aparência da vegetação.
Chowdhury et al. (2015)	Identificação de vegetação por textura	Extração CBP + classificadores SVM e RNA	110 imagens (900×500), duas classes de vegetação (densa e esparsa)	Precisão de 92,72%, restrito a imagens de vídeo e recortes.
Zhang et al. (2015)	Classificação de vegetação por fusão de cor e textura	Momentos de cor + intensidade aparente	50 imagens naturais de beira de estrada	79% de precisão; dificuldade em diferenciar vegetação com texturas similares.
Zhang et al. (2016)	Modelo contextual espacial (superpixel)	Classificação baseada em contexto local/-global	50 imagens de margens de estrada	77,44% de precisão; confusão frequente entre solo e vegetação marrom.
Baral et al. (2022)	Segmentação de vegetação em imagens aéreas	Modelo U-Net (Deep Learning)	15 imagens aéreas de rodovias do Texas (alta resolução)	Acurácia de 92%; necessidade de mais dados e pós-processamento.
Nosso trabalho	Estimativa da altura da vegetação e tomada de decisão	CNN (ResNet-50) com Transfer Learning + pós-processamento	229.095 imagens esféricas (3072×2048) de rodovias BR-070, BR-163 e BR-364 (MT)	Acurácia de 85%; solução completa para classificação da altura da vegetação e priorização de manutenção.

de recursos utilizado consiste em recursos de cor e de textura. Com base nesse conjunto a detecção é implementada usando um algoritmo de SVM. O modelo foi treinado com 300 imagens RGB brutas, uma câmera comum foi usada para aquisição de dados, sem nenhum sensor especializado adicional. O desempenho obtido foi de 96,25% de acurácia, no entanto, ainda assim objetos problemáticos em imagens, que são semelhantes à vegetação em cor foram classificados incorretamente como vegetação na maioria das vezes.

O estudo de [Baral et al. \(2022\)](#), testa um sistema que utiliza o modelo U-Net para treinar com um conjunto de dados de imagens aéreas de alta resolução disponíveis publicamente, que foram treinados com 15 imagens de rodovias do Texas, nos Estados Unidos. Os resultados apresentam perspectivas promissoras com acurácia de 92%, porém o modelo ainda necessita de um treinamento mais robusto e da aplicação de técnicas de pós-processamento para aprimorar os resultados.

Por fim, no trabalho de [Harbas and Subasic \(2014\)](#), várias abordagens foram utilizadas para a classificação da vegetação, cada uma com diferentes níveis de complexidade e técnicas. A base de dados contém 270 imagens extraídas das sequências de vídeo gravadas em rodovias, cada uma com resolução de 1920x1080 pixels. Foram aplicados métodos com características selecionadas manualmente, abordando características de cor e textura, e um método com aprendizado de características, utilizando o algoritmo de aprendizado de máquina Support Vector Machine (SVM). Apesar dessa última abordagem obter melhores resultados, alcançando uma taxa de 96,25% de precisão, o método de detecção de vegetação não está otimizado e não possui velocidade de execução.

Os trabalhos discutidos apresentam uma variedade de abordagens para a detecção de vegetação em margens de rodovias. Embora métodos baseados em cor e textura, com classificadores SVM e Redes Neurais Artificiais, tenham alcançado resultados promissores com taxas de precisão entre 75,5% e 96,25%, algumas limitações persistem. Falhas recorrentes incluem dificuldades na diferenciação entre vegetação e objetos similares em cor e a necessidade de maior otimização para tempo real. Além disso, a falta de bases de dados padronizadas limita a comparação e o aprimoramento dos métodos existentes.

Neste contexto, a solução apresentada neste trabalho buscou adotar uma abordagem mais completa e avançada no processamento dos dados. O volume e a qualidade da base de dados utilizada para treinar o modelo são superiores em comparação com outros modelos propostos com objetivos semelhantes, uma vez que a maioria dos trabalhos apresentados fazem uso de imagens de satélite ou drones, enquanto este está preparado para obtenção de imagens na perspectiva da rodovia, obtendo mais detalhes relevantes para a classificação.

3 Materiais e Métodos

A fim de desenvolver o modelo de aprendizado de máquina para estimativa da altura da vegetação, fez-se necessário a obtenção de material para popular as bases de dados de imagens para realização do processamento do modelo, bem como a aplicação de avançadas técnicas computacionais para o desenvolvimento do modelo de aprendizados de máquina, indo desde a fase do pré ao pós-processamento dos dados, até a produção da aplicação de visualização dos

resultados. As seções a seguir descrevem todos os passos realizados para completar o objetivo do trabalho.

3.1 Obtenção e tratamento dos dados

As imagens foram obtidas utilizando a câmera de alta precisão para imagens esféricas de 360° Ladybug 5+, que foi instalada na parte superior de um veículo que percorreu trechos das rodovias BR-070, BR-163 e BR-364 no estado do Mato Grosso, o que resulta em capturas do sul ao norte do Estado, abrangendo três biomas diferentes (Amazônia, Cerrado e Pantanal). Isso faz com que diferentes tipos de vegetações sejam analisadas. As capturas foram realizadas de forma sequencial ao longo dos percursos, com a câmera configurada para registrar imagens a uma taxa de 5 quadros por segundo (FPS), garantindo uma amostragem contínua do ambiente rodoviário. No total, foram gravadas 229.095 imagens com resolução de 3072 x 2048 pixels.

Na etapa de preparação dos dados, foram selecionadas 36.765 imagens para anotação utilizando o CVAT (*Computer Vision Annotation Tool*), ferramenta de anotação de código aberto desenvolvida pela Intel. Em [Musleh et al. \(2023\)](#) é mostrado que o software oferece um conjunto abrangente de ferramentas de anotação, incluindo detecção de objetos, segmentação de imagens e anotação de vídeo.

Para isso, os dados foram categorizados em quatro classes de acordo com o nível de vegetação nas imagens: sem vegetação, baixa (até 10cm de altura), média (de 10cm a 30cm de altura) e alta (mais que 30cm de altura). A [Fig. 1](#) mostra exemplos de imagens anotadas seguindo as classes definidas.

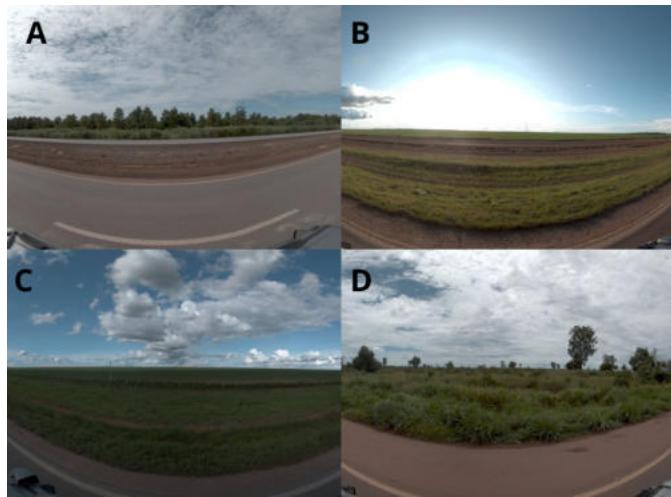


Figura 1: Imagens anotadas de acordo com o estado da vegetação: sem vegetação (A); baixa (B); média (C); alta (D).

A classificação foi realizada considerando a altura da vegetação dentro do perímetro de quatro metros a partir da linha de bordo da via analisada em cada imagem. Esta é a área principal de controle de manutenção da vegetação

em rodovias. Foram estabelecidas regras para os anotadores estimarem o tamanho da área que constitui esses quatro metros em cada imagem, considerando a distorção da imagem e a inclinação do relevo. Devido a subjetividade tanto do espaço a ser considerado e mesmo a altura estimada da vegetação, foram necessários pelo menos duas anotações coincidentes dos anotadores para classificar as imagens, contando com um terceiro anotador em casos de divergência na classificação de uma mesma imagem.

A [Tabela 2](#) apresenta as classificações obtidas através da anotação manual. A maioria das imagens foram anotadas como da classe "baixa", enquanto a menor foi da classe "média". Posteriormente houve uma seleção das imagens para que as classes fossem balanceadas, que é explicado na [Seção 3.2](#).

Tabela 2: Anotações das imagens no CVAT.

Classe	Altura	Total
Sem vegetação	-	5.047
Baixa	Até 10 cm	19.716
Média	Entre 10 e 30 cm	4.470
Alta	Acima de 30 cm	7.232

3.2 Modelo de inteligência artificial

A abordagem escolhida para desenvolver o modelo utilizado neste trabalho foi o aprendizado de máquina supervisionado, onde o algoritmo processa dados rotulados e retorna uma saída pré-definida. Para esse fim, foi utilizada a arquitetura ResNet-50, desenvolvida pela Microsoft Research e proposta em [He et al. \(2016\)](#), que se destaca por suas conexões residuais, permitindo treinar redes profundas sem o problema de degradação do gradiente, mantendo estabilidade e capacidade de generalização mesmo em cenários com grande variabilidade visual.

A escolha da ResNet-50 foi motivada por sua eficácia na extração de características e pelo amplo uso consolidado na literatura em tarefas de alto nível em visão computacional. Este modelo é utilizado em aplicações de análise de imagens médicas ([Rath et al., 2025; Lin and Wu, 2023](#)), em reconhecimento facial ([Deng et al., 2022](#)) e como backbone em arquiteturas de detecção e segmentação de objetos ([Ren et al., 2015; Carion et al., 2020](#)). Esses exemplos reforçam a robustez da arquitetura, especialmente em cenários que exigem discriminação fina de elementos em imagens, como no caso da vegetação capturada em rodovias.

Além disso, antes da definição do tipo de modelo, foram realizados experimentos preliminares com outras arquiteturas pré-treinadas, entre elas Vision Transformer (ViT) ([Dosovitskiy et al., 2021](#)) e EVA-CLIP ([Sun et al., 2023](#)). Nesses experimentos, a ResNet-50 apresentou melhores métricas de acurácia e maior estabilidade durante o treinamento, além de menor custo computacional. Assim, considerando desempenho, eficiência e maturidade da arquitetura na literatura, a ResNet-50 foi definida como a opção mais adequada para este trabalho.

A fim de especializar o modelo, foi aplicada técnica de *Transfer Learning*, no qual a camada de saída do modelo foi

configurada com quatro neurônios finais, que representam cada uma das possíveis classes para classificação da altura da vegetação presente nas imagens.

O conjunto de dados usado para treinar o modelo foi composto de 8.000 imagens para o conjunto de treinamento e 2.000 imagens para o conjunto de teste, dividido igualmente entre as quatro classes anotadas, totalizando 2.000 imagens por classe para o treinamento e 500 imagens por classe para validação do desempenho do modelo. A escolha por um total de 10.000 imagens decorreu do mínimo de imagens por classe e de experimentos preliminares que avaliaram o impacto do aumento do volume de dados sobre o desempenho do modelo. Observou-se que, a partir desse limite, o acréscimo de novas amostras não resultava em melhorias significativas nas métricas de validação, indicando uma estabilização do aprendizado. Dessa maneira, esse ponto representou um equilíbrio em desempenho computacional de treinamento e desempenho do modelo.

Três modelos distintos foram treinados. O primeiro modelo utilizou as quatro classes categorizadas nas anotações. O segundo modelo foi treinado e validado com imagens de apenas três das quatro classes, excluindo a classe "sem vegetação". Por sua vez, o terceiro modelo foi treinado com as quatro classes, porém, na fase de validação, as classes "sem vegetação" e "baixa" foram combinadas em uma única classe, uma vez que representam o mesmo grau de manutenção de vegetação. Ou seja, o terceiro modelo consiste do primeiro modelo com aplicação de pós-processamento para considerar sem vegetação e baixa como uma mesma classe.

3.3 Avaliação do modelo

Para avaliar o desempenho do modelo foi necessário calcular e analisar as métricas de avaliação atingidas durante o seu processamento. Nesta seção, serão apresentadas as métricas de Precisão, *F1 Score*, *Recall* e a acurácia obtidas pelo modelo, bem como a análise da matriz de confusão da classificação das imagens e a variação das taxas de perda (*loss*) durante o treinamento e a validação do modelo.

A taxa de precisão analisa os acertos obtidos por categoria de dados. O cálculo é feito avaliando a quantidade de verdadeiros positivos (VP) dividido pela soma de todos os valores verdadeiros e falsos positivos (FP), como é apresentado na Eq. (1):

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}} \quad (1)$$

Por sua vez, o *Recall* é a métrica que avalia a capacidade do modelo em detectar com sucesso resultados classificados como positivos. O cálculo é feito dividindo os verdadeiros positivos pela soma dos falsos negativos (NF) e verdadeiros positivos Eq. (2):

$$\text{Recall} = \frac{\text{VP}}{\text{VP} + \text{NF}} \quad (2)$$

A partir das métricas de Precisão e do *Recall* do modelo,

obtem-se o cálculo da média harmônica entre os dois valores, chamado de *F1 Score*, como é mostrado na Eq. (3):

$$\text{F1 Score} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (3)$$

Foi calculado também a acurácia. Essa métrica define o quanto próximo da realidade estão os resultados obtidos pelo modelo, representado pela Eq. (4):

$$\text{Acurácia} = \frac{\text{VP} + \text{NV}}{\text{VP} + \text{NV} + \text{FP} + \text{FN}} \quad (4)$$

4 Resultados e Discussões

Os resultados obtidos pelo treinamento do modelo para categorizar as imagens de acordo com a altura da vegetação são apresentados na Seção 4.1. Em seguida, na Seção 4.2, é aproveitado do desempenho deste modelo para a avaliação de todo um trecho. Finalmente, é possível definir o fluxo de tomada de decisão, apresentado na Seção 4.3.

4.1 Treinamento e avaliação do modelo

O treinamento dos modelos foi realizado utilizando um computador com processador Xeon w7-2495x, com memória RAM de 128 GB e duas placas de vídeo RTX A5500 de 24 GB. Os modelos foram treinados por 300 épocas. Inicialmente foram realizados dois treinamentos utilizando abordagens diferentes. A primeira abordagem treinou o modelo com as 4 classes de imagens anotadas, enquanto a segunda utilizou 3 classes, agregando a classe "sem vegetação" na classe "vegetação baixa", uma vez que demandam o mesmo nível de manutenção.

A Fig. 2 mostra a variação do índice de *loss* durante o treinamento dos modelos de classificação ao longo das 300 épocas. Enquanto o gráfico da Fig. 3 mostra a evolução da acurácia.

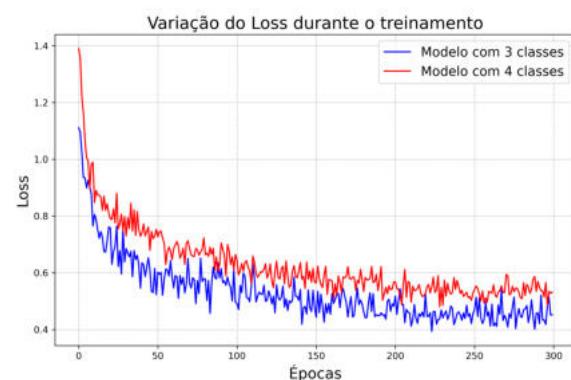


Figura 2: Medida de *loss* durante o treinamento dos modelos com 3 e 4 classes.

Durante o treinamento, observou-se que o modelo com

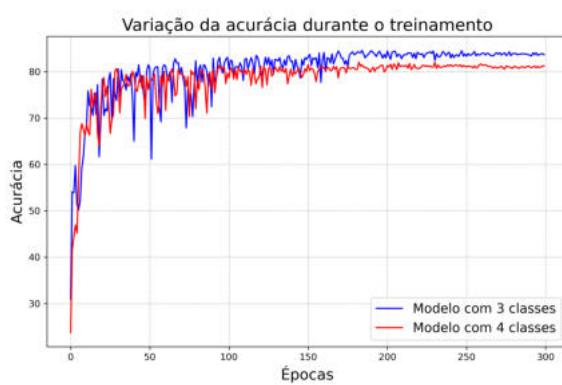


Figura 3: Medida de acurácia durante o treinamento dos modelos com 3 e 4 classes.

3 classes teve um desempenho superior ao modelo com 4 classes. O modelo de 3 classes apresentou uma redução rápida de *loss* durante as primeiras iterações e estabilizou em um valor de erro mais baixo, por volta de 0,4, indicando uma boa capacidade de generalização ao apresentar um menor erro na base de treinamento. Em contrapartida, o modelo de 4 classes iniciou com um erro maior e reduziu esse valor mais lentamente, estabilizando em torno de 0,6.

Quanto a acurácia, nota-se que ambos os modelos apresentam um rápido aumento de acurácia. Após 50 épocas o desempenho se estabiliza, com o modelo de 3 classes alcançando consistentemente uma acurácia superior, por volta de 80%, enquanto o modelo de 4 classes estabiliza por volta de 75%.

Dentro do contexto de manutenção de vegetação à beira de rodovias, as classes "baixa" e "sem vegetação" possuem o mesmo significado, uma vez que o objetivo principal é detectar trechos de vegetação que necessitam de manutenção, ou seja, onde a classificação retorna vegetação "média" e "alta". Nesse sentido, com objetivo de alcançar resultados mais precisos, foi adicionado um pós-processamento para considerar as classes "sem vegetação" e "vegetação baixa" como sendo uma só classe.

A Tabela 3 mostra o desempenho das três abordagens considerando todas as métricas de avaliação. O modelo de 3 classes com pós-processamento obteve os melhores resultados, com 85% de acurácia.

Tabela 3: Métricas de Avaliação dos Modelos

Modelo	Acurácia	F1Score	Precisão	Recall
3 classes	0,84	0,84	0,84	0,84
4 classes	0,82	0,82	0,82	0,82
Pós-processamento	0,85	0,83	0,85	0,85

Os dados de teste foram utilizados para medir o desempenho do modelo melhor avaliado durante a etapa de treinamento. Uma matriz de confusão foi gerada e está apresentada na Fig. 4. Nela é possível notar que o bom desempenho continuou na base de teste. Também é possível notar os erros principais com a classe "média", em alguns momentos se confundindo com vegetação baixa e

em outros com a vegetação alta, provavelmente por causa da subjetividade das anotações em relação a altura da vegetação.

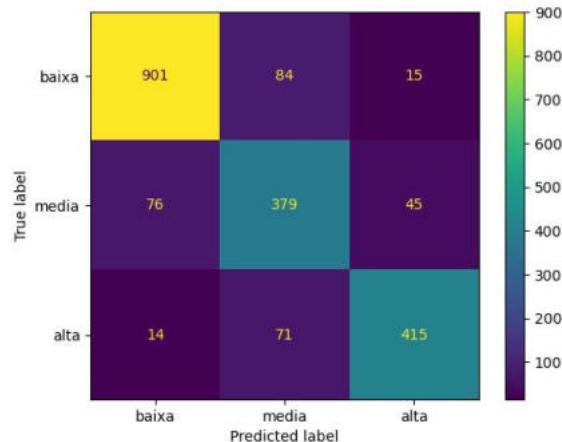


Figura 4: Matriz de confusão do modelo de quatro classes convertido em três classes no pós-processamento.

4.2 Avaliação por trecho

Para gerar um alerta sobre a necessidade de manutenção da vegetação à beira da estrada, é preciso identificar uma área contínua de vegetação alta. Para isso, é necessário o agrupamento das classificações de múltiplas imagens, de modo a formar um trecho de aproximadamente 500 metros, onde a vegetação está predominantemente classificada como "alta". Mais do que isso, é necessário um valor que represente a urgência da manutenção, para que os trechos possam ser priorizados.

Com esse propósito, foi desenvolvido um algoritmo que recebe as imagens classificadas e faz a avaliação do grau de necessidade de manutenção por trecho, retornando um valor entre 0 e 1. Quanto mais alto o valor, maior a necessidade de manutenção, indicando uma predominância maior de imagens classificadas como "alta". O Algoritmo 1 apresenta a função que calcula esse valor.

Inicialmente, é definido o peso a ser considerado na pontuação para cada classe de altura de vegetação, respectivamente: baixo, médio e alto. Além disso, são definidas variáveis para armazenar quantas imagens estão sendo consideradas. Em seguida, para cada classificação realizada no trecho, o algoritmo considera as que tiveram um grau de confiança de pelo menos 0,5. A seguir, o algoritmo calcula uma soma ponderada, multiplicando as contagens de cada classe pelo respectivo peso da classe. O resultado desta soma é normalizado pelo grau máximo de manutenção, que seria caso todas as imagens consideradas tivessem sido classificadas com a classe alta. O resultado deste cálculo representa o grau de manutenção daquele trecho.

Foi realizado um teste para validar a avaliação por trecho realizada pelo algoritmo. Para tanto, foram criados

Algorithm 1 Calcular Manutenção

```

1: procedure CALCULAR_MANUTENCAO(classificacoes)
2:   pesos  $\leftarrow$  [0,0, 0,3, 0,5]
3:   cont_classes  $\leftarrow$  [0,0,0]
4:   imagens  $\leftarrow$  0
5:   for all confianca, classe em classificacoes do
6:     if confianca  $\geq$  0,5 then
7:       cont_classes[classe]  $\leftarrow$  cont_classes[classe] + 1
8:       imagens  $\leftarrow$  imagens + 1
9:     end if
10:   end for
11:   soma  $\leftarrow$  cont_classes[0]  $\times$  pesos[0]
12:   soma  $\leftarrow$  soma + cont_classes[1]  $\times$  pesos[1]
13:   soma  $\leftarrow$  soma + cont_classes[2]  $\times$  pesos[2]
14:   manutencao  $\leftarrow$  soma / (imagens  $\times$  pesos[2])
15:   retornar manutencao
16: end procedure

```

cinco trechos com características distintas, variando a proporção de imagens das três classes (baixa, média, alta), no qual o primeiro trecho não requer manutenção e o quinto trecho requer manutenção urgente. A Fig. 5 apresenta a análise da estimativa de manutenção para esses diferentes trechos.

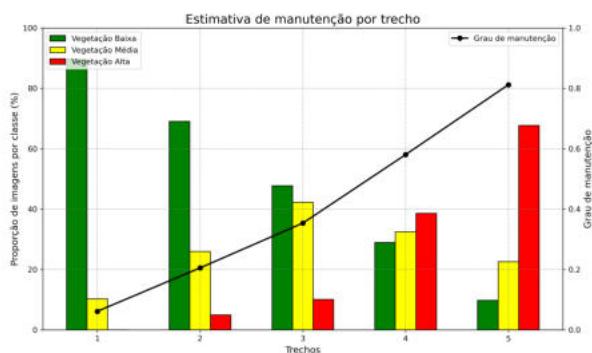


Figura 5: Comparação das classificações feitas pela rotulação e pelo modelo.

O crescimento do grau de manutenção aumentou conforme a necessidade de intervenções se intensificou, como era esperado. Nos trechos 1 e 2, a maioria das imagens foi classificada como vegetação baixa, com 184 imagens para o trecho 1 e 270 para o trecho 2. Também foram identificadas algumas áreas de vegetação média, com 21 e 105 imagens, respectivamente, enquanto a vegetação alta foi escassa, sem imagens no trecho 1 e apenas 20 no trecho 2. O grau de manutenção nesses trechos é baixo, variando entre 0,06 e 0,2, o que sugere uma necessidade mínima de intervenções.

O trecho 3 apresenta uma distribuição mais equilibrada entre os três tipos de vegetação, com uma leve predominância de vegetação baixa, contabilizando cerca de 95 imagens. Em seguida, há 84 imagens de vegetação média e 20 de vegetação alta. O grau de manutenção neste trecho está

próximo de 0,35, indicando uma necessidade moderada de intervenções.

Nos trechos 4 e 5, observa-se um aumento significativo na proporção de vegetação média e alta, enquanto a vegetação baixa diminui. Para a classe de vegetação alta, foram classificadas 152 e 132 imagens, respectivamente, enquanto a vegetação baixa teve 114 e 19 imagens. O grau de manutenção variou entre 0,58 e 0,81, indicando uma necessidade de manutenção mais intensificada.

Por meio desse teste, conclui-se que à medida que a vegetação alta se torna mais predominante em um trecho, o grau de manutenção também aumenta. Dessa forma, foi validada a estimativa gerada pelo algoritmo desenvolvido para avaliar a situação por trecho.

4.3 Fluxo para tomada de decisão

O fluxo do processo de tomada de decisão está representado na Fig. 6. O processo envolve a análise da vegetação com base em imagens coletadas pelo veículo, passando por várias etapas até a apresentação dos dados.

O processo começa com a coleta de dados por um veículo equipado com sensores e câmeras. Esses arquivos de imagens são extraídos e preparados para análise posterior. As imagens extraídas são carregadas em um banco de dados. Durante essa fase, são adicionados metadados, que são informações descritivas sobre cada imagem, como as coordenadas geográficas, lado da pista e sentido de direção do carro.

Em seguida, as imagens são organizadas e segmentadas em diferentes áreas e trechos. Utilizando o modelo criado, as imagens são classificadas de acordo com a estimativa da altura da vegetação presentes.

Com base nas classificações, é feita uma estimativa da condição da vegetação para cada área segmentada de aproximadamente 500 metros, utilizando o algoritmo desenvolvido e apresentado na Seção 4.2.

Ao final do processo, os resultados podem ser apresentados em uma interface gráfica, que apresentará todas as informações obtidas pelo processo e emitirá os alertas de necessidade de manutenção.

A implementação desse fluxo de processo para a tomada de decisões é extremamente vantajosa, especialmente em termos de automação, precisão e eficiência. A gestão proativa de rodovias é facilitada, com maior segurança e otimização no uso de recursos, ao viabilizar ações de manutenção localizadas e fundamentadas em dados concretos. A interface gráfica e os alertas automatizados aprimoram a comunicação e aceleram as respostas, resultando em um sistema mais eficiente e econômico.

5 Considerações Finais

A manutenção da vegetação nas margens das rodovias é essencial para garantir a segurança dos usuários e otimizar a gestão operacional do governo ou de concessionárias responsáveis. Normalmente, a inspeção e a roçagem dessas áreas são realizadas de forma manual, resultando em desafios logísticos e custos elevados, além de possíveis erros. Com o objetivo de contribuir para a automação desse processo, este trabalho desenvolveu um modelo de apren-

Processo de Tomada de Decisão

Vegetação

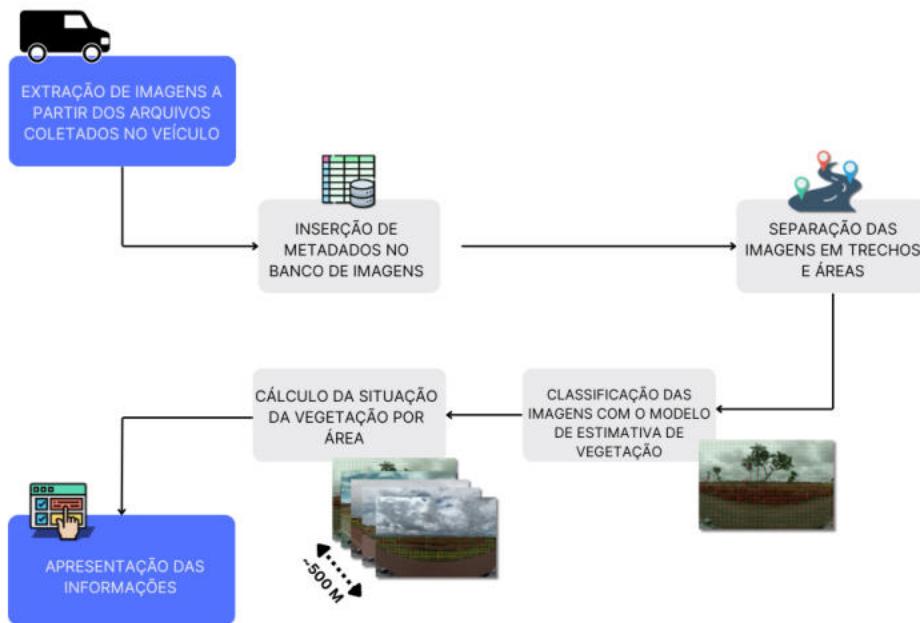


Figura 6: Diagrama do fluxo do sistema para possibilitar a tomada de decisão por parte dos usuários.

dizado de máquina capaz de estimar a altura da vegetação utilizando imagens de alta resolução capturadas ao longo de rodovias e criou um fluxo para tomada de decisão.

Para atingir esse objetivo, utilizou-se a arquitetura ResNet-50 com técnica de *Transfer Learning*, o que permitiu um desempenho robusto na categorização da vegetação em até quatro classes: "sem vegetação", "baixa", "média" e "alta". Além disso, o processo de avaliação por trecho, proposto no trabalho, possibilitou identificar áreas críticas que requerem atenção imediata. Os resultados obtidos indicaram que o modelo de quatro classes com pós-processamento (combinando "sem vegetação" e "vegetação baixa") apresentou melhor desempenho, alcançando uma acurácia de 85%. Essa abordagem demonstrou ser mais eficaz ao avaliar grandes trechos de rodovia, gerando relatórios detalhados sobre a necessidade de manutenção, permitindo maior eficiência na operação de roçagem.

Enquanto a maioria das abordagens existentes baseia-se em imagens de satélite, drones ou vídeos recortados, este trabalho utiliza imagens capturadas na perspectiva da rodovia, o que proporciona maior riqueza de detalhes e precisão na classificação da vegetação. Além do mais, o modelo desenvolvido combina técnicas de aprendizado profundo com um fluxo automatizado de decisão por trecho, permitindo não apenas a classificação pontual da vegetação, mas também a análise contínua e integrada de grandes extensões de estrada.

Como sugestão para trabalhos futuros, seria interessante explorar outras abordagens de aprendizado de máquina, como técnicas de *segmentation*, para aumentar a precisão na detecção de vegetação em diferentes cenários.

Além disso, a expansão da base de dados para incluir outras rodovias e diferentes épocas do ano pode contribuir para melhorar a robustez do modelo em situações mais variadas.

Agradecimentos

O presente trabalho foi realizado com apoio dos Recursos de Desenvolvimento Tecnológico (RDT), da Concessionária Nova Rota do Oeste, sob regulação da Agência Nacional de Transportes Terrestres (ANTT), oficializado no processo nº 50500.294497/2023-59 sob coordenação técnica de Régis De Bel e coordenação geral de Rheno Batista Tormin Filho.

Referências

- Baral, A., Nasr, M., Darghiasi, P., Abediniangerabi, B. and Shahandashti, M. (2022). Detection and classification of vegetation for roadside vegetation inspection and rehabilitation using deep learning techniques, *International Conference on Transportation and Development 2022*. <https://doi.org/10.1061/9780784484319.014>.
- Bezerra, E., Mafalda, S., Alvarez, A. and Chavez, R. (2023). Análise temporal de ilhas de calor utilizando processamento de imagens de satélite: estudo de caso rio branco, acre, *Revista Brasileira de Computação Aplicada* 15: 70–78. <https://doi.org/10.5335/rbca.v15i1.14098>.
- Buchinger, D. and Silva, A. (2014). Anomalies detection in

- asphalt pavements: a morphological image processing approach, *Revista Brasileira de Computação Aplicada* 6. <https://doi.org/10.5335/rbca.2014.3661>.
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A. and Zagoruyko, S. (2020). End-to-end object detection with transformers, *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I*, Springer-Verlag, Berlin, Heidelberg, p. 213–229. https://doi.org/10.1007/978-3-030-58452-8_13.
- Chowdhury, S., Verma, B. and Stockwell, D. (2015). A novel texture feature based multiple classifier technique for roadside vegetation classification, *Expert Systems with Applications* 42(12): 5047–5055. <https://doi.org/10.1016/j.eswa.2015.02.047>.
- Deng, J., Guo, J., Yang, J., Xue, N., Kotsia, I. and Zafeiriou, S. (2022). Arcface: Additive angular margin loss for deep face recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44(10): 5962–5979. <https://doi.org/10.1109/TPAMI.2021.3087709>.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. and Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. <https://doi.org/10.48550/arXiv.2010.11929>.
- Harbas, I. and Subasic, M. (2014). Cwt-based detection of roadside vegetation aided by motion estimation, 2014 5th European Workshop on Visual Information Processing EUVIP, IEEE;, pp. 1–6. <https://doi.org/10.1109/EUVIP.2014.7018405>.
- Harbas, I. and Subašić, M. (2014). Detection of roadside vegetation using features from the visible spectrum, 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 1204–1209. <https://doi.org/10.1109/MIPRO.2014.6859751>.
- Harbaš, I., Prentašić, P. and Subašić, M. (2018). Detection of roadside vegetation using fully convolutional networks, *Image and Vision Computing* 74: 1–9. <https://doi.org/10.1016/j.imavis.2018.03.008>.
- He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- Lin, C.-L. and Wu, K.-C. (2023). Development of revised resnet-50 for diabetic retinopathy detection, *BMC Bioinformatics* 24(1): 157. <https://doi.org/10.1186/s12859-023-05293-1>.
- Musleh, A., AlRyalat, S. A. and Qasem, A. (2023). Image annotation software for artificial intelligence applications, *High Yield Medical Reviews* 1(2). <https://doi.org/10.59707/hymrXHMX8234>.
- Novello, G., Yamamoto, H. and Cabral, E. (2021). An end-to-end approach to autonomous vehicle control using deep learning, *Revista Brasileira de Computação Aplicada* 13: 32–41. <https://doi.org/10.5335/rbca.v13i3.12135>.
- Rath, A., Mishra, B. S. P. and Bagal, D. K. (2025). Resnet50-based deep learning model for accurate brain tumor detection in mri scans, *Next Research* 2(1): 100104. <https://doi.org/10.1016/j.nexres.2024.100104>.
- Ren, S., He, K., Girshick, R. and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks, in C. Cortes, N. Lawrence, D. Lee, M. Sugiyama and R. Garnett (eds), *Advances in Neural Information Processing Systems*, Vol. 28, Curran Associates, Inc.
- Sun, Q., Fang, Y., Wu, L., Wang, X. and Cao, Y. (2023). Eva-clip: Improved training techniques for clip at scale. <https://doi.org/10.48550/arXiv.2303.15389>.
- Zhang, L., Verma, B. and Stockwell, D. (2015). Roadside vegetation classification using color intensity, in J. Dongarra, I. Foster, G. Fox, W. Gropp, K. Kennedy, L. Torczon and A. White (eds), 2015 11th International Conference on Natural Computation (ICNC), Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. <https://doi.org/10.1109/ICNC.2015.7378170>.
- Zhang, L., Verma, B. and Stockwell, D. (2016). Spatial contextual superpixel model for natural roadside vegetation classification, *Pattern Recognition* 60: 444–457. <https://doi.org/10.1016/j.patcog.2016.05.013>.