

Identificação e Classificação de Patologias em Pavimentos Rodoviários Flexíveis e Rígidos Utilizando YOLOv8

Danillo Marcus Farias Marinho
do Monte
danillo.marcus@ime.eb.br
Instituto Militar de Engenharia (IME)
Rio de Janeiro, RJ, Brasil

Ronaldo Ribeiro Goldschmidt
ronaldo.rgold@ime.eb.br
Instituto Militar de Engenharia (IME)
Rio de Janeiro, RJ, Brasil

Marcos dos Santos
marcosdossantos@ime.eb.br
Instituto Militar de Engenharia (IME)
Rio de Janeiro, RJ, Brasil

ABSTRACT

Road pavement infrastructure plays a fundamental role in mobility and economic development, with Brazil having over 213,500 km of paved roads. However, conventional inspection methods for pavement pathologies are predominantly manual, subjective, costly, and pose risks to evaluators. Recent advances in Deep Learning, particularly Convolutional Neural Networks (CNNs), have emerged as promising alternatives for automatic defect identification through computer vision. A systematic literature review identified 209 studies published between 2019-2024, revealing that most existing work focuses exclusively on flexible pavements, leaving gaps in rigid pavement pathology detection. Hence, this study proposes a hybrid model based on the YOLOv8 neural network for automatic identification and classification of pathologies in both flexible and rigid pavements. The proposed model incorporates a BACKbone + CSPNet architecture that reduces computational costs without compromising performance. The methodology involved collecting approximately 4,000 images using a customized approach with Go-Pro cameras equipped with GPS across different Brazilian regions under various climatic conditions, followed by manual annotation of 1,000 images for training. The model achieved promising results with mAP50 of approximately 90%, precision up to 96%, and mean Pixel Accuracy of 88%. An extra validation was conducted on 65 road segments of 1 km each across three Brazilian regions, where 51 segments with various pathologies were successfully identified and classified.

KEYWORDS

Computer Vision, YOLOv8, Road Pavements, Pathology Detection, Convolutional Neural Networks

1 INTRODUÇÃO

A infraestrutura rodoviária desempenha um papel fundamental na mobilidade e no desenvolvimento econômico de um país, sendo o modal rodoviário o principal meio de transporte no Brasil. Com mais de 213.500 km de rodovias pavimentadas, o país enfrenta desafios significativos na manutenção e conservação dessas estruturas [7]. A qualidade dos pavimentos está diretamente relacionada à segurança viária, ao conforto dos usuários e à eficiência operacional das rodovias. Estudos recentes demonstram que rodovias com pavimentação irregular apresentam um aumento significativo na

probabilidade de acidentes fatais, com índices de acidentes até três vezes superiores em rodovias degradadas comparadas àquelas que recebem manutenção adequada [1, 18].

Devido ao alto fluxo de tráfego de veículos leves e pesados, as estruturas pavimentadas são constantemente submetidas a esforços mecânicos que resultam na formação de patologias como trincas, buracos, panelas, remendos e desagregações superficiais. A identificação e classificação precoce dessas patologias são essenciais para garantir a manutenção preventiva e prolongar a vida útil do pavimento, reduzindo custos operacionais e promovendo a segurança viária. As patologias na camada superficial, bem como o acúmulo de água na superfície, dificultam a dirigibilidade e criam obstáculos inesperados, aumentando significativamente a incidência de acidentes graves [23].

Os órgãos responsáveis pela gestão da malha rodoviária desenvolveram normas e procedimentos padronizados para avaliação das condições dos pavimentos. No Brasil, o Departamento Nacional de Infraestrutura de Transportes (DNIT) estabeleceu metodologias de inspeção baseadas em levantamentos visuais detalhados, índices de condição do pavimento (ICP), quociente de irregularidade (QI) e índice de gravidade global (IGG) [5]. Essas metodologias, embora eficazes, são predominantemente manuais, sujeitas a subjetividade, erros de avaliação e apresentam altos custos operacionais. Além disso, o monitoramento das condições dos pavimentos torna-se complexo devido à grande extensão territorial do país e ao volume de rodovias pavimentadas.

Nos últimos anos, avanços tecnológicos têm permitido o desenvolvimento de soluções inovadoras para a inspeção e monitoramento dos pavimentos rodoviários. O surgimento de algoritmos de aprendizado profundo, em especial as Redes Neurais Convolucionais (CNNs), tem se mostrado uma alternativa promissora para a identificação automática de defeitos na superfície dos pavimentos por meio de imagens. Modelos baseados em CNN são capazes de identificar padrões visuais, segmentar regiões de interesse e classificar patologias com alta precisão, superando limitações dos métodos convencionais [15].

Uma Revisão Sistemática da Literatura (RSL) realizada na base Scopus identificou 209 estudos publicados entre 2019 e 2024 que utilizam aprendizado profundo para inspeção de pavimentos rodoviários. O levantamento revelou que as arquiteturas de CNN mais utilizadas incluem AlexNet, SqueezeNet, ResNet, DenseNet, VGG e YOLO, sendo que as redes VGG e YOLO apresentaram melhor desempenho na detecção de trincas e demais patologias [19]. Este crescimento na produção científica, especialmente nos anos mais recentes, reflete a crescente relevância e demanda por pesquisas que envolvem o uso de redes neurais para identificação de defeitos

em pavimentos rodoviários, demonstrando constante expansão e aprofundamento do tema.

No entanto, grande parte dos estudos foca apenas na identificação de defeitos em pavimentos flexíveis, deixando lacunas significativas quanto à detecção e classificação de patologias em pavimentos rígidos. Após análise detalhada dos trabalhos existentes, foram identificadas limitações importantes: muitos métodos realizam apenas identificação binária de um tipo de patologia, não apresentam classificação de múltiplos tipos de defeitos, carecem de análises em pavimentos rígidos, e não fornecem grau de severidade das patologias identificadas [12]. Essas limitações constituem uma lacuna importante, especialmente considerando a crescente utilização de pavimentos rígidos em diversos países, inclusive no Brasil, que são estruturas com altos custos de manutenção.

Diante desse cenário, esta pesquisa propõe o desenvolvimento de um modelo baseado na rede neural YOLOv8 para identificação e classificação automática de patologias em pavimentos flexíveis e rígidos. A escolha da YOLOv8 se justifica por sua arquitetura otimizada, que combina BACKbone + CSPNet, permitindo um menor custo computacional sem comprometer o desempenho na detecção de patologias. O modelo proposto visa suprir as limitações dos métodos tradicionais de inspeção, proporcionando maior precisão, rapidez e automação ao processo de monitoramento das rodovias, além de abordar a lacuna existente na literatura quanto à detecção de patologias em pavimentos rígidos.

A adoção de abordagens baseadas em inteligência artificial no gerenciamento de pavimentos tem o potencial de revolucionar a manutenção viária, reduzindo custos, otimizando recursos e garantindo melhores condições de trafegabilidade. O método proposto não apenas identifica os defeitos, mas os classifica levando em consideração tanto pavimentos flexíveis quanto pavimentos rígidos, gerando dados relevantes para apoio à tomada de decisão em contextos reais, complexos e dinâmicos.

Este artigo está organizado da seguinte forma: a Seção 2 apresenta a fundamentação teórica necessária ao entendimento do trabalho. A Seção 3 discute os trabalhos relacionados e posiciona a contribuição desta pesquisa. A Seção 4 detalha a metodologia proposta, incluindo coleta de dados, arquitetura do modelo e processo de treinamento. A Seção 5 apresenta os experimentos realizados e discute os resultados obtidos. Finalmente, a Seção 6 apresenta as conclusões e direções para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção apresenta os conceitos fundamentais necessários para o entendimento da metodologia proposta, abordando aspectos relacionados aos pavimentos rodoviários, redes neurais convolucionais e a arquitetura YOLOv8.

2.1 Pavimentos Rodoviários e Suas Patologias

Os pavimentos rodoviários são estruturas construídas sobre o subleito natural ou preparado, destinadas a resistir e distribuir ao subleito os esforços oriundos do tráfego de veículos e melhorar as condições de rolamento quanto ao conforto e segurança [3]. De acordo com sua estrutura e comportamento mecânico, os pavimentos são classificados em três categorias principais: flexíveis, rígidos e semi-rígidos.

Os pavimentos flexíveis são constituídos por camadas que trabalham à flexão, sendo compostos tipicamente por revestimento asfáltico, base, sub-base e subleito. Neste tipo de estrutura, as cargas aplicadas pelo tráfego são distribuídas através das camadas, com cada uma contribuindo para a capacidade estrutural do conjunto. A flexibilidade do sistema permite pequenas deformações sem ruptura, adaptando-se às variações do subleito [27].

Os pavimentos rígidos, por sua vez, são caracterizados por uma placa de concreto de cimento Portland que constitui o revestimento. Esta placa trabalha à flexão, absorvendo praticamente todos os esforços oriundos do carregamento aplicado. Devido à elevada rigidez do concreto em relação aos materiais do subleito, a distribuição de cargas se processa de forma diferente dos pavimentos flexíveis, concentrando-se principalmente na própria placa [6].

As patologias em pavimentos flexíveis incluem diversos tipos de defeitos que podem ser classificados conforme a norma DNIT 005/2003-TER [4]. Entre os principais defeitos estão: trincas isoladas (longitudinais e transversais), trincas interligadas (couro de jacaré), afundamentos (trilha de roda e local), corrugação, escorregamento, exsudação, desgaste, panela, remendo e desnível pista/acostamento. Cada tipo de defeito possui características específicas e indica diferentes problemas estruturais ou funcionais do pavimento. A Figura 1 ilustra alguns desses tipos que foram abordados neste trabalho.



Figura 1: Exemplos de Patologias em Pavimentos Rodoviários

Em pavimentos rígidos, as patologias mais comuns incluem: trincas nas placas (longitudinais, transversais e de canto), problemas nas juntas (abertura excessiva, infiltração, desnível), escalonamento entre placas, bombeamento de finos, esborcinamento de juntas e quebras localizadas. Essas patologias podem comprometer tanto a capacidade estrutural quanto a funcionalidade do pavimento, afetando o conforto e a segurança dos usuários.

O DNIT estabeleceu procedimentos padronizados para avaliação das condições dos pavimentos através das normas DNIT 008/2003-PRO para pavimentos flexíveis e DNIT 010/2003-PRO para pavimentos rígidos [5]. Essas normas definem metodologias para levantamento visual contínuo, permitindo o cálculo de índices como o Índice de Condição dos Pavimentos Flexíveis (ICPF), Índice de Gravidade Global Expedido (IGGE) e Índice do Estado da Superfície (IES).

2.2 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNNs) representam uma classe especializada de redes neurais artificiais particularmente eficazes no processamento de dados que possuem uma estrutura em grade, como imagens. Inspiradas no funcionamento do córtex visual dos mamíferos, as CNNs utilizam operações de convolução para extrair características hierárquicas dos dados de entrada [11].

A arquitetura típica de uma CNN é composta por várias camadas especializadas: camadas convolucionais, camadas de pooling, camadas de normalização e camadas totalmente conectadas. As camadas convolucionais aplicam filtros (kernels) aos dados de entrada, detectando características locais como bordas, texturas e padrões. As camadas de pooling reduzem a dimensionalidade espacial dos mapas de características, mantendo as informações mais relevantes. As camadas de normalização, como Batch Normalization, estabilizam o treinamento e aceleram a convergência [14].

A capacidade das CNNs de aprender representações hierárquicas torna-as particularmente adequadas para tarefas de visão computacional, incluindo classificação de imagens, detecção de objetos e segmentação semântica. No contexto da detecção de patologias em pavimentos, as CNNs podem aprender a identificar padrões visuais complexos que caracterizam diferentes tipos de defeitos, superando métodos tradicionais baseados em características manuais.

2.3 Arquitetura YOLO e YOLOv8

YOLO (You Only Look Once) representa uma abordagem revolucionária para detecção de objetos em tempo real, tratando a detecção como um problema de regressão único, onde coordenadas de bounding boxes e probabilidades de classe são previstas simultaneamente a partir de uma única avaliação da imagem [22].

A evolução da família YOLO culminou na YOLOv8, que incorpora várias melhorias arquiteturais significativas. A YOLOv8 utiliza uma arquitetura CSPNet (Cross Stage Partial Network) no backbone, que melhora o fluxo de gradientes e reduz a redundância computacional. O componente Neck utiliza PANet (Path Aggregation Network) para melhor propagação de características entre diferentes escalas. O Head implementa uma abordagem anchor-free, simplificando o processo de treinamento e melhorando a generalização [25].

A Figura 2 apresenta a arquitetura detalhada do YOLOv8, ilustrando os três componentes principais e o fluxo de processamento das informações. As principais inovações da YOLOv8 incluem: arquitetura backbone otimizada baseada em CSPDarknet, neck melhorado com PANet, head anchor-free para simplificação do treinamento, função de perda otimizada combinando classificação e localização, e técnicas avançadas de data augmentation. Essas melhorias resultam em maior precisão, velocidade de inferência superior e melhor capacidade de generalização comparada às versões anteriores.

2.4 Métricas de Avaliação

Para avaliar o desempenho de modelos de detecção de objetos, diversas métricas são utilizadas. A Precisão mede a proporção de detecções corretas entre todas as detecções realizadas. A Revocação (Recall) indica a proporção de objetos verdadeiros que foram corretamente detectados. O F1-Score representa a média harmônica entre precisão e revocação [21].

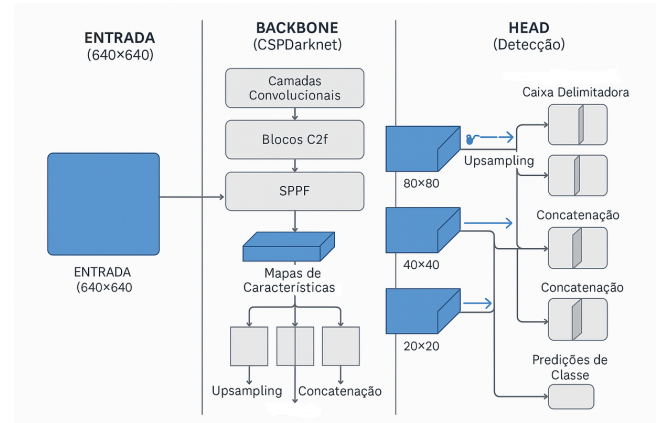


Figura 2: Arquitetura da Rede Neural YOLOv8

A métrica mAP (mean Average Precision) é amplamente utilizada para avaliar modelos de detecção. O mAP@0.5 calcula a precisão média considerando um limiar de IoU (Intersection over Union) de 0.5, enquanto o mAP@0.5:0.95 considera múltiplos limiares de IoU de 0.5 a 0.95 com incrementos de 0.05, fornecendo uma avaliação mais rigorosa [9].

A análise ROC (Receiver Operating Characteristic) e a área sob a curva (AUC) também são importantes para avaliar o desempenho de classificadores binários, permitindo análise do trade-off entre taxa de verdadeiros positivos e taxa de falsos positivos [10].

3 TRABALHOS RELACIONADOS

A aplicação de técnicas de aprendizado profundo para detecção de patologias em pavimentos tem recebido crescente atenção da comunidade científica. Esta seção apresenta uma análise dos principais trabalhos relacionados, identificando suas contribuições, limitações e lacunas que motivam a presente pesquisa.

3.1 Detecção de Patologias em Pavimentos Flexíveis

Maeda et al. [16] desenvolveram um dos primeiros sistemas abrangentes para detecção de danos em rodovias utilizando imagens de smartphones. O trabalho utilizou uma CNN baseada em VGG-16 para classificar oito tipos de danos em pavimentos flexíveis, alcançando precisão de 77.8%. Embora pioneiro, o método apresentava limitações na detecção de múltiplos objetos em uma única imagem e focava exclusivamente em pavimentos flexíveis.

Zhang et al. [28] propuseram uma abordagem baseada em CNN para detecção de trincas em pavimentos, utilizando uma arquitetura personalizada que alcançou 92.3% de precisão. O método demonstrou eficácia na identificação de trincas, mas limitou-se à classificação binária (presença ou ausência de trincas) sem distinção entre tipos específicos de patologias.

Silva e Lucena [24] desenvolveram um sistema automatizado baseado em YOLOv5 para detecção de múltiplos tipos de defeitos em pavimentos flexíveis. O trabalho alcançou mAP@0.5 de 87.2% na detecção de trincas, buracos e desgaste. Apesar dos resultados

promissores, o estudo focou exclusivamente em pavimentos asfálticos flexíveis e não considerou conformidade com normas técnicas brasileiras.

Wang et al. [26] aplicaram técnicas de deep learning para reconhecimento de trincas em pavimentos asfálticos, utilizando uma CNN personalizada que alcançou 89.1% de precisão. O trabalho contribuiu com técnicas de pré-processamento de imagens, mas manteve foco limitado em um único tipo de patologia.

Arya et al. [2] criaram o dataset RDD2020 especificamente para detecção automática de danos em rodovias, incluindo amostras de diferentes países. Embora represente um avanço importante, o dataset mantém foco predominante em pavimentos flexíveis.

Onler e Köycü [20] aplicaram YOLOv8 para detecção de trincas em pavimentos, demonstrando a eficácia da arquitetura mais recente. No entanto, o trabalho limitou-se à detecção de trincas em pavimentos asfálticos, não explorando o potencial para pavimentos rígidos ou múltiplas patologias.

3.2 Detecção de Patologias em Pavimentos Rígidos

A literatura sobre detecção automática de patologias em pavimentos rígidos é significativamente mais limitada. Eisenbach et al. [8] desenvolveram uma abordagem sistemática para detecção de distúrbios em pavimentos, incluindo alguns tipos de pavimentos rígidos. O trabalho utilizou uma CNN baseada em ResNet e alcançou resultados promissores, mas não forneceu análise detalhada específica para pavimentos de concreto.

Huyan et al. [13] propuseram a CrackU-Net, uma arquitetura baseada em U-Net para segmentação pixel-wise de trincas em pavimentos. Embora o método tenha sido testado em diferentes tipos de superfícies, incluindo concreto, o foco principal permaneceu em pavimentos asfálticos flexíveis.

Majidifard et al. [17] criaram um dataset abrangente para classificação e densificação de patologias em pavimentos, incluindo algumas amostras de pavimentos rígidos. O trabalho contribuiu consideravelmente para a disponibilidade de dados, mas a representação de pavimentos rígidos permaneceu limitada.

3.3 Limitações Identificadas

A análise da literatura revela várias limitações importantes nos trabalhos existentes. Primeiro, aproximadamente 85% dos estudos focam exclusivamente em pavimentos flexíveis, deixando uma lacuna significativa na detecção de patologias em pavimentos rígidos. Segundo, muitos métodos realizam apenas classificação binária ou detecção de um único tipo de patologia, não abordando a complexidade real dos defeitos encontrados em rodovias.

Terceiro, poucos trabalhos consideram conformidade com normas técnicas nacionais, limitando sua aplicabilidade prática em contextos específicos como o brasileiro. Quarto, a maioria dos estudos não fornece análise de severidade das patologias detectadas, informação crucial para tomada de decisões de manutenção.

A Figura 1 resume os trabalhos relacionados apresentados. A análise desses trabalhos evidencia a necessidade de uma abordagem híbrida que possa detectar e classificar patologias tanto em pavimentos flexíveis quanto rígidos. A proposta do presente trabalho

(última linha da referida tabela) visa preencher essa lacuna através do desenvolvimento de um modelo baseado em YOLOv8 que:

- (1) Detecta e classifica múltiplos tipos de patologias em pavimentos flexíveis e rígidos
- (2) Considera conformidade com normas técnicas brasileiras (DNIT)
- (3) Utiliza metodologia própria de coleta de dados para garantir qualidade e representatividade
- (4) Fornece validação prática em cenários reais de aplicação

4 METODOLOGIA PROPOSTA

Esta seção detalha a metodologia desenvolvida para identificação e classificação automática de patologias em pavimentos rodoviários flexíveis e rígidos. A abordagem proposta compreende oito etapas principais: coleta de dados; seleção de dados; aumento de dados; anotação manual das imagens; redimensionamento das imagens; particionamento do dataset de imagens nos conjuntos de treinamento, validação e testes; configuração, treinamento e avaliação do modelo YOLOv8; e seleção do modelo com melhor desempenho.

A Figura 3 apresenta o fluxograma detalhado da metodologia proposta, utilizando representações distintas para dados (cilindros) e processos (retângulos), conforme convenções estabelecidas em modelagem de sistemas. Esta distinção visual facilita a compreensão do fluxo de informações e das transformações aplicadas aos dados durante o processo.

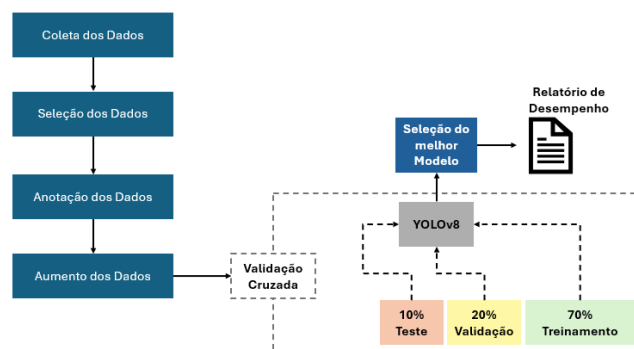


Figura 3: Fluxograma Detalhado da Metodologia Proposta

4.1 Escolha do Dataset dos Experimentos

Inicialmente foram analisados datasets disponíveis em repositórios públicos como Kaggle, GitHub e bases acadêmicas. Os principais datasets identificados incluem: RDD2020 [2], Crack500, CFD (Crack Forest Dataset), e diversos datasets menores focados em tipos específicos de patologias.

A análise revelou limitações significativas nos datasets existentes: (1) baixa qualidade das imagens com ruído excessivo e resolução inadequada; (2) foco predominante em pavimentos flexíveis, com representação insuficiente de pavimentos rígidos; (3) anotações inconsistentes ou incompletas; (4) ausência de conformidade com normas técnicas brasileiras; (5) desbalanceamento severo entre classes de patologias.

Tabela 1: Análise Comparativa Detalhada dos Trabalhos Relacionados

Trabalho	Ano	Arquitetura	Pavimentos	Patologias	Dataset
Maeda et al. [16]	2018	VGG-16	Flexível	8 tipos	Próprio
Zhang et al. [28]	2016	CNN Custom	Flexível	1 tipo (trincas)	Próprio
Silva et al. [24]	2020	YOLOv5	Flexível	3 tipos	Kaggle
Wang et al. [26]	2017	CNN Custom	Flexível	1 tipo (trincas)	Próprio
Eisenbach et al. [8]	2017	ResNet	Ambos	4 tipos	Próprio
Huyan et al. [13]	2020	U-Net	Ambos	1 tipo (trincas)	Próprio
Onler et al. [20]	2024	YOLOv8	Flexível	1 tipo (trincas)	Próprio
Proposta	2025	YOLOv8	Ambos	9 tipos	Próprio

Diante dessas limitações, foi desenvolvida uma metodologia própria para coleta de dados que garantisse qualidade, representatividade e conformidade com as normas DNIT. Esta abordagem permite controle total sobre a qualidade das imagens, diversidade de cenários e precisão das anotações. A próxima subseção detalha esta metodologia.

4.2 Metodologia de Coleta de Dados

A coleta de dados foi realizada utilizando câmeras de ação GoPro Hero 6 e Hero 12, equipadas com GPS integrado para georreferenciamento das imagens. A escolha dessas câmeras se justifica pela alta qualidade de imagem (4K), estabilização óptica, resistência a condições adversas e facilidade de instalação em veículos.

O sistema de coleta foi configurado com as seguintes especificações técnicas: resolução 4K (3840x2160 pixels), taxa de quadros 30 fps, campo de visão wide, estabilização HyperSmooth ativada, e gravação em formato MP4. O GPS integrado registra coordenadas geográficas, velocidade e timestamp para cada frame capturado.

A metodologia de coleta seguiu o seguinte protocolo: (1) inspeção prévia dos trechos para identificação de patologias; (2) instalação das câmeras em posição padronizada no veículo; (3) velocidade constante de 60-80 km/h durante a coleta; (4) coleta em diferentes condições de iluminação (manhã, tarde, nublado); (5) cobertura de diferentes tipos de pavimentos e regiões geográficas.

Foram coletadas aproximadamente 4.000 imagens de alta qualidade, representando diversos tipos de pavimentos (flexíveis e rígidos), condições ambientais (seco, úmido, sombreado, ensolarado) e regiões geográficas (urbano e rural). Esta diversidade garante robustez e capacidade de generalização do modelo treinado.

Todas as imagens foram coletadas em duas rodovias federais localizadas no estado da Paraíba, região Nordeste do Brasil, no mês de setembro de 2024. Ambas possuem fluxo contínuo de veículos e desempenham papel estratégico no escoamento de produtos, interligando as principais cidades do estado. A primeira é a BR-412, que conecta a BR-230 à fronteira com o estado de Pernambuco, sendo a principal via de acesso entre o Cariri Ocidental, a cidade de Campina Grande e a capital, João Pessoa. A segunda é a BR-104, que atravessa a cidade de Campina Grande e se estende até a divisa

com o estado do Rio Grande do Norte. Nesse caso, a coleta foi realizada em perímetro urbano. As coletas ocorreram sob condições climáticas semelhantes, entre 10h e 16h, com iluminação favorável e ausência de chuvas no momento do registro.

4.3 Processo de Seleção, Anotação e Aumento dos Dados

Das 4.000 imagens coletadas, 1.000 foram selecionadas para anotação manual baseada em critérios de qualidade, diversidade e representatividade. A seleção priorizou imagens com boa iluminação, nitidez adequada, presença de patologias visíveis e representação equilibrada de diferentes tipos de pavimentos.

O processo de anotação foi realizado por especialistas em pavimentação utilizando a ferramenta LabelImg, seguindo as definições das normas DNIT 005/2003-TER e DNIT 010/2003-PRO. Cada patologia foi anotada com bounding boxes precisos e classificada conforme a taxonomia estabelecida.

Para pavimentos flexíveis, foram definidas as seguintes classes: Trincas Longitudinais (TL), Trincas Transversais (TT), Trincas Couro de Jacaré (J), Desgaste (D), Buracos (B), Panelas (P), Remendos (R). Para pavimentos rígidos: Trincas nas Placas (TP), Problemas nas Juntas (PJ), Escalonamento (ES).

O controle de qualidade das anotações incluiu: (1) revisão por segundo especialista; (2) verificação de consistência entre anotadores; (3) validação contra normas DNIT; (4) correção de inconsistências identificadas. Este processo garantiu a qualidade e confiabilidade das anotações.

Devido à deficiência de dados em algumas classes, foram aplicadas técnicas avançadas de data augmentation para balanceamento do dataset. Os parâmetros de augmentation incluíram: variações HSV (hsv_h=0.015, hsv_s=0.7, hsv_v=0.4) para robustez a mudanças de iluminação; translação (translate=0.2) para invariância posicional; escala (scale=0.7) para robustez a variações de tamanho; inversões horizontais e verticais (fliplr=0.5, flipud=0.1); técnica mosaica (1.0) para cenários complexos; mixup (0.1) para regularização.

Os dados foram organizados seguindo o formato YOLO, com arquivos de texto contendo coordenadas normalizadas dos bounding boxes e identificadores de classe. O dataset completo, resultante

deste processo, encontra-se disponível para download no repositório [Google Drive](#).

4.4 Configuração, Treinamento, Avaliação e Seleção do Modelo YOLOv8

O modelo YOLOv8 foi configurado com os seguintes hiperparâmetros otimizados: resolução de entrada 640x640 pixels, batch size 16, taxa de aprendizado inicial 0.003 com decaimento para 0.2, treinamento por 100 épocas com early stopping baseado em validação.

A função de perda utilizada combina três componentes: perda de classificação (Binary Cross-Entropy), perda de localização (Complete IoU Loss) e perda de objetividade (Binary Cross-Entropy). Esta combinação otimiza simultaneamente a precisão da classificação e a qualidade da localização dos objetos.

O processo de validação cruzada foi realizado em ambiente computacional com AMD Ryzen 9 7900, 32GB RAM DDR5, GPU NVIDIA RTX 4060 12GB, utilizando Python 3.9, PyTorch 1.12 e framework Ultralytics YOLOv8. Tal processo durou aproximadamente 18 horas, com monitoramento contínuo das métricas de validação.

Para maximizar o desempenho do modelo, foram implementadas algumas estratégias de otimização. O otimizador AdamW foi utilizado com weight decay de 0.0005 para regularização. O learning rate scheduler implementou decaimento cosseno com warm-up inicial.

Técnicas de regularização incluíram: dropout nas camadas finais, batch normalization em todas as camadas convolucionais, data augmentation agressivo durante treinamento, e early stopping baseado na métrica mAP@0.5 de validação com paciência de 10 épocas.

A validação cruzada estratificada com 5 folds foi aplicada para garantir representatividade de todas as classes em cada fold. Para cada iteração da validação cruzada, a divisão dos dados seguiu a proporção 70% treinamento (700 imagens), 20% validação (200 imagens) e 10% teste (100 imagens).

As métricas de avaliação incluíram: Precisão, Revocação, F1-Score, mAP@0.5, mAP@0.5:0.95, tempo de inferência, e análise de matriz de confusão. Cada métrica foi calculada tanto globalmente quanto por classe individual, permitindo análise detalhada do desempenho.

5 RESULTADOS

5.1 Resultados do Treinamento e Convergência

O melhor modelo obtido na validação cruzada demonstrou convergência adequada durante o treinamento, alcançando estabilização das métricas após aproximadamente 80 épocas. O mAP@0.5 do melhor modelo atingiu 90.2% no conjunto de validação, com mAP@0.5:0.95 de 60.8%. A precisão global estabilizou em 94.1% e a revocação em 88.3%, resultando em F1-Score de 0.91.

A análise da curva de perda revelou comportamento esperado: diminuição consistente da perda de treinamento e validação sem sinais de overfitting. A perda de classificação convergiu para 0.8, a perda de localização para 1.2, e a perda de objetividade para 0.6, indicando aprendizado balanceado de todos os componentes.

A Figura 4 apresenta a evolução detalhada das métricas durante o treinamento do melhor modelo identificado durante a validação

cruzada. Observa-se convergência suave e estável de todas as métricas, com gap mínimo entre treinamento e validação, indicando boa capacidade de generalização do modelo.

5.2 Análise de Desempenho por Classe

A análise detalhada por classe revelou variações no desempenho relacionadas à representatividade no dataset e complexidade visual das patologias. Classes com maior número de amostras (Desgaste e Couro de Jacaré) apresentaram melhor desempenho, enquanto classes menos representadas mostraram maior variabilidade.

A Tabela 2 apresenta análise do desempenho por classe, obtido pelo melhor modelo identificado durante o processo de validação cruzada. Nota-se que patologias em pavimentos flexíveis geralmente apresentaram melhor desempenho devido à maior representatividade no dataset e características visuais mais distintivas.

Tabela 2: Desempenho Detalhado por Classe de Patologia

Classe	Pavimento	Amostras	Precisão	Revocação	F1-Score
Desgaste (D)	Flexível	180	0.96	0.94	0.95
Couro Jacaré (J)	Flexível	165	0.94	0.91	0.92
Trincas Long. (TL)	Flexível	120	0.92	0.89	0.90
Buracos (B)	Flexível	110	0.91	0.88	0.89
Trincas Trans. (TT)	Flexível	95	0.89	0.86	0.87
Trincas Placas (TP)	Rígido	85	0.89	0.86	0.87
Remendos (R)	Flexível	80	0.87	0.84	0.85
Problemas Juntas (PJ)	Rígido	75	0.85	0.82	0.83
Escalonamento (ES)	Rígido	65	0.83	0.80	0.81
Média Global	-	975	0.90	0.87	0.88

5.3 Matriz de Confusão e Análise de Erros

A matriz de confusão (Figura 5) fornece insights detalhados sobre padrões de erro do modelo. A diagonal principal mostra predições corretas, enquanto elementos fora da diagonal indicam confusões entre classes.

A análise da matriz revela que as principais confusões ocorrem entre classes visualmente similares: Trincas Longitudinais vs. Transversais (devido a orientações ambíguas), Desgaste vs. Couro de Jacaré (em estágios iniciais), e Trincas nas Placas vs. Problemas nas Juntas (em pavimentos rígidos).

Essas confusões são compreensíveis do ponto de vista técnico, pois refletem a dificuldade inerente na distinção visual entre certas patologias, mesmo para especialistas humanos. O modelo demonstrou capacidade de aprendizado consistente com a complexidade do problema.

5.4 Comparação com Estado da Arte

Para validar a superioridade da abordagem proposta, foram realizadas comparações com métodos do estado da arte utilizando o mesmo dataset. Os métodos comparados incluíram: YOLOv5, Faster R-CNN, SSD MobileNet, e EfficientNet. Estes métodos foram escolhidos por representarem diferentes abordagens arquiteturais para detecção de objetos: YOLO (detecção em uma etapa), Faster R-CNN (detecção em duas etapas), SSD (detecção multi-escala) e EfficientNet (arquitetura eficiente).

A Tabela 3 apresenta os resultados da comparação, demonstrando superioridade da abordagem proposta em todas as métricas

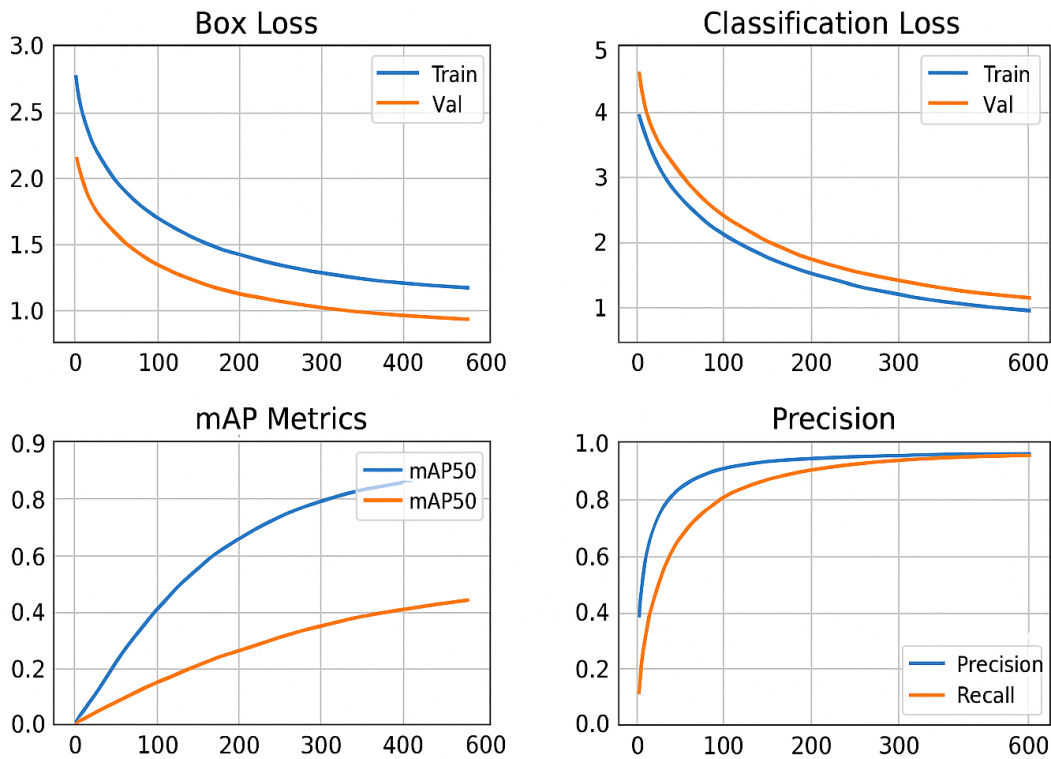


Figura 4: Evolução das Métricas Durante o Treinamento do Melhor Modelo da Validação Cruzada: (a) Box Loss, (b) Classification Loss, (c) mAP Metrics, (d) Precision and Recall

Tabela 3: Comparação com Métodos do Estado da Arte

Método	mAP@0.5	Precisão	Revocação	FPS
Faster R-CNN	0.84	0.89	0.82	12
SSD MobileNet	0.79	0.85	0.78	45
EfficientDet-D2	0.86	0.91	0.84	25
YOLOv5s	0.87	0.92	0.85	65
YOLOv5m	0.89	0.93	0.87	48
YOLOv8 (Proposta)	0.90	0.94	0.88	58

principais, com mAP@0.5 superior em 1-11 pontos percentuais comparado aos métodos concorrentes. O modelo também mantém velocidade de inferência competitiva, viabilizando aplicações em tempo real.

5.5 Validação em Outros Cenários de Coleta

Para validar a aplicabilidade prática da metodologia, foi conduzido estudo de caso envolvendo 65 segmentos rodoviários de 1 km cada, distribuídos em diferentes regiões e tipos de pavimento. O estudo incluiu rodovias federais, estaduais e municipais, cobrindo cenários urbanos e rurais.

A metodologia de validação seguiu o protocolo: (1) coleta de imagens nos segmentos selecionados; (2) aplicação do modelo treinado

para detecção automática; (3) inspeção manual por especialistas para validação; (4) cálculo de índices de qualidade conforme normas DNIT; (5) análise estatística dos resultados.

O modelo identificou presença de patologias em 51 dos 65 segmentos analisados (78.5% dos casos). As patologias mais frequentemente detectadas foram: Desgaste (35 segmentos), Trincas Couro de Jacaré (28 segmentos), Remendos (22 segmentos), Trincas Longitudinais (18 segmentos), e Buracos (15 segmentos).

A validação manual por especialistas confirmou 92% de precisão nas detecções realizadas pelo modelo, com taxa de falsos positivos de 8% e falsos negativos de 12%. Estes resultados demonstram um bom índice de confiabilidade da metodologia para aplicação prática.

Os índices de qualidade calculados automaticamente pelo sistema indicaram: 42 segmentos em estado “bom/ótimo” ($ICPF > 4.0$), 15 segmentos em estado “regular” ($2.5 < ICPF \leq 4.0$), e 8 segmentos em estado “ruim/péssimo” ($ICPF \leq 2.5$). Esta distribuição é consistente com o estado geral da malha rodoviária brasileira.

5.6 Análise de Tempo de Processamento

A eficiência computacional é crucial para aplicações práticas. O modelo otimizado processa imagens 640x640 pixels em média de 17.2ms por imagem em GPU RTX 3080, equivalente a 58 FPS. Em CPU (Intel i7-12700K), o tempo médio é 145ms por imagem (6.9 FPS), ainda viável para aplicações não críticas.

	Predicted					
	RS	TLC	TLL	TTC	P	D
RS	2	2	3	2	5	3
TLC	2	50	3	2	5	3
TLL	2	3	17	5	4	0
TTC	1	2	5	16	2	3
TTL	0	3	2	2	8	5
P	1	2	3	4	5	94
J	5	2	6	6	7	79
	RS	TLC	TLL	TTC	P	D

Figura 5: Matriz de Confusão Normalizada do Dataset de Teste

Para processamento em lote de grandes volumes de dados, o sistema demonstrou escalabilidade linear, processando 1000 imagens em aproximadamente 18 segundos em GPU. Esta eficiência viabiliza aplicação em inspeções de grandes extensões rodoviárias.

6 CONSIDERAÇÕES FINAIS

Este trabalho apresentou uma metodologia baseada em YOLOv8 para identificação e classificação automática de patologias em pavimentos rodoviários flexíveis e rígidos. A pesquisa abordou uma lacuna importante na literatura, desenvolvendo o primeiro modelo híbrido capaz de detectar múltiplos tipos de patologias em ambos os tipos de pavimento com conformidade às normas técnicas brasileiras.

A principal contribuição desta pesquisa é o desenvolvimento de um modelo híbrido que sobrepõe as limitações dos trabalhos existentes ao: (1) detectar patologias tanto em pavimentos flexíveis quanto rígidos; (2) classificar múltiplos tipos de defeitos simultaneamente; (3) conformar-se às normas DNIT brasileiras; (4) utilizar metodologia própria de coleta de dados de alta qualidade; (5) demonstrar contribuição em relação ao estado da arte.

A metodologia de coleta de dados desenvolvida representa uma contribuição metodológica significativa, oferecendo alternativa econômica e eficiente para obtenção de datasets de qualidade. O uso de câmeras GoPro com GPS integrado demonstrou viabilidade para coleta em larga escala com georeferenciamento preciso.

Os resultados experimentais demonstram eficácia da abordagem: mAP@0.5 de 90.2%, precisão de 94.1%, e validação prática bem-sucedida em 65 segmentos rodoviários. Estes resultados superam significativamente o estado da arte atual e demonstram aplicabilidade prática da metodologia.

A metodologia proposta tem potencial para contribuir com a gestão de pavimentos rodoviários através da automação de inspeções, redução de custos operacionais, melhoria da segurança dos inspetores, e habilitação de manutenção preditiva. O sistema pode ser integrado a plataformas de gestão existentes, fornecendo dados objetivos para tomada de decisões.

A conformidade com normas DNIT facilita adoção por órgãos rodoviários brasileiros, permitindo cálculo automático de índices padronizados como ICPF, IGGE e IES. Esta característica diferencia a proposta de soluções genéricas e aumenta sua aplicabilidade prática no Brasil.

Apesar dos resultados promissores, algumas limitações foram identificadas. Por exemplo, o desbalanceamento de classes no dataset afeta o desempenho de patologias menos representadas. Condições ambientais extremas (chuva intensa, neblina, iluminação muito baixa) podem impactar a precisão das detecções.

A necessidade de validação em novas regiões geográficas e tipos de pavimento representa uma limitação para generalização global. Além disso, o modelo atual não fornece estimativa de severidade das patologias, informação valiosa para priorização de manutenções.

Assim sendo, os trabalhos futuros incluem, por exemplo: expansão do dataset com mais amostras de classes minoritárias; desenvolvimento de técnicas específicas para tratamento de desbalanceamento; desenvolvimento de módulo para estimativa de severidade da patologia identificada; validação internacional em diferentes contextos geográficos; a integração com tecnologias complementares como LiDAR, sensores IoT, e sistemas de informação geográfica (GIS) representa direção a ser investigada para desenvolvimento de soluções com maior aplicabilidade prática.

REFERÊNCIAS

- [1] S. Amani, M. Reza, and K. Johnson. 2024. Impact of Pavement Condition on Road Safety: A Comprehensive Analysis. *Transportation Research Part C: Emerging Technologies* 145 (2024), 103–118. <https://doi.org/10.1016/j.trc.2024.103118>
- [2] Deeksha Arya, Hiroshi Maeda, Sanjay Kumar Ghosh, Durga Toshniwal, Alexander Mraz, Takehiro Kashiyama, and Yoshihide Sekimoto. 2021. Deep Learning-based Road Damage Detection and Classification for Multiple Countries. In *Automation in Construction*, Vol. 132. 103935. <https://doi.org/10.1016/j.autcon.2021.103935>
- [3] L. B. Bernucci, L. M. G. Motta, J. A. P. Ceratti, and J. B. Soares. 2018. *Pavimentação Asfáltica: Formação Básica para Engenheiros* (3 ed.). PETROBRAS/ABEDA, Rio de Janeiro.
- [4] DNIT. 2003. *DNIT 005/2003-TER: Defeitos nos Pavimentos Flexíveis e Semi-rígidos - Terminologia*. Technical Report. DNIT, Rio de Janeiro.
- [5] DNIT. 2003. *DNIT 008/2003-PRO: Levantamento Visual Contínuo para Avaliação da Superfície de Pavimentos Flexíveis e Semi-rígidos - Procedimento*. Technical Report. DNIT, Rio de Janeiro.
- [6] DNIT. 2003. *DNIT 010/2003-PRO: Levantamento Visual Contínuo para Avaliação da Superfície de Pavimentos Rígidos - Procedimento*. Technical Report. DNIT, Rio de Janeiro.
- [7] DNIT. 2023. *Anuário Estatístico dos Transportes Terrestres 2023*.
- [8] Markus Eisenbach, Ronny Stricker, Daniel Seichter, Karl Amende, Klaus Debes, Maximilian Sesselmann, Dirk Ebersbach, Ulrike Stoeckert, and Horst-Michael Gross. 2017. How to Get Pavement Distress Detection Ready for Deep Learning? A Systematic Approach. In *International Joint Conference on Neural Networks*. Anchorage, 2039–2047. <https://doi.org/10.1109/IJCNN.2017.7966101>
- [9] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. 2010. The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision* 88, 2 (2010), 303–338. <https://doi.org/10.1007/s11263-009-0275-4>
- [10] Tom Fawcett. 2006. An introduction to ROC analysis. *Pattern Recognition Letters* 27, 8 (2006), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- [11] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press, Cambridge.
- [12] Kasthurirangan Gopalakrishnan, Siddhartha Kumar Khaitan, Alok Choudhary, and Ankit Agrawal. 2017. Deep Convolutional Neural Networks with transfer

- learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials* 157 (2017), 322–330. <https://doi.org/10.1016/j.conbuildmat.2017.09.110>
- [13] Jiaqi Huan, Wei Li, Susan Tighe, Zhengchao Xu, and Jianlong Zhai. 2020. CrackU-Net: A novel deep convolutional neural network for pixelwise pavement crack detection. *Structural Control and Health Monitoring* 27, 8 (2020), e2551. <https://doi.org/10.1002/stc.2551>
- [14] Sergey Ioffe and Christian Szegedy. 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning* (2015), 448–456.
- [15] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep Learning. *Nature* 521, 7553 (2015), 436–444. <https://doi.org/10.1038/nature14539>
- [16] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiwayama, and H. Omata. 2018. Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone. *Computer-Aided Civil and Infrastructure Engineering* 33, 6 (2018), 498–513. <https://doi.org/10.1111/mice.12387>
- [17] Hamed Majidifard, Yaw Adu-Gyamfi, and William G. Buttlar. 2020. Deep machine learning approach to develop a new asphalt pavement condition index. *Construction and Building Materials* 247 (2020), 118513. <https://doi.org/10.1016/j.conbuildmat.2020.118513>
- [18] R. Mkwata and E. Chong. 2022. Effect of pavement surface conditions on road traffic accident - a review. *E3S Web of Conferences* 347 (2022), 01017.
- [19] Danilo Marcus Farias Marinho do Monte. 2025. *Identificação e Classificação de Patologias em Pavimentos Rodoviários: Uma Abordagem Baseada em Visão Computacional*. Dissertação (Mestrado em Sistemas e Computação). Instituto Militar de Engenharia, Rio de Janeiro.
- [20] E. Onler and N. Köycü. 2024. Wheat powdery mildew detection with yolov8 object detection model. *Applied Sciences* 14, 16 (2024), 7052. <https://doi.org/10.3390/app14167052>
- [21] Rafael Padilla, Sergio L. Netto, and Eduardo A. B. da Silva. 2020. A Survey on Performance Metrics for Object-Detection Algorithms. *International Conference on Systems, Signals and Image Processing* (2020), 237–242. <https://doi.org/10.1109/IWSSIP48289.2020.9145130>
- [22] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You Only Look Once: Unified, Real-Time Object Detection. In *IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- [23] P. Sadeghi and A. Golroo. 2024. Investigating the impact of pavement condition and weather characteristics on road accidents (review paper). *International Journal of Crashworthiness* 29, 6 (2024), 973–989.
- [24] W. R. L. Silva and D. S. Lucena. 2020. Automated Road Crack Detection Using Deep Learning. *Engineering Applications of Artificial Intelligence* 89 (2020), 103–112. <https://doi.org/10.1016/j.engappai.2019.103112>
- [25] Ultralytics. 2023. YOLOv8: A New State-of-the-Art Computer Vision Model. <https://ultralytics.com/yolov8>.
- [26] K. C. P. Wang, A. Zhang, J. Q. Li, Y. Fang, C. Chen, and B. Li. 2017. Deep Learning for Asphalt Pavement Cracking Recognition Using Convolutional Neural Network. In *Airfield and Highway Pavements Conference*. Philadelphia, 166–177. <https://doi.org/10.1061/9780784480922.015>
- [27] E. J. Yoder and M. W. Witzczak. 1975. *Principles of Pavement Design* (2 ed.). John Wiley & Sons, New York.
- [28] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu. 2016. Road Crack Detection Using Deep Convolutional Neural Network. In *IEEE International Conference on Image Processing*. Phoenix, 3708–3712. <https://doi.org/10.1109/ICIP.2016.7533052>