

# Estudo Empírico de Técnicas de Aprendizagem de Máquina na Classificação e Identificação de Imagens de Depressões em Vias Rodoviárias.

Jhonatan Oliveira de Souza<sup>1</sup>, Ingrid Mendes<sup>2</sup>, Otávio Noura Teixeira<sup>1</sup>

<sup>1</sup>Faculdade de Engenharia de Computação (FECOMP), Campus Universitário de Tucuruí (CAMTUC), Universidade Federal do Pará (UFPA)

<sup>2</sup>PPGEE - Programa de Pós Graduação em Engenharia Elétrica (ITEC / UFPA)  
ITEC-Instituto de Tecnologia da UFPA

{jhonatanoliveira315, onoura}@gmail.com, ingrid.mendes@tucuruui.ufpa.br

**Abstract.** This study presents an experimental comparative evaluation of machine learning techniques applied to the classification of road images with and without potholes. The algorithms K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Convolutional Neural Networks (CNN) were analyzed under the same standardized *Dataset* and experimental protocol. Each model was executed 50 times using different random seeds to ensure statistical robustness. Performance was assessed using Accuracy, Precision, Recall, and F1-Score metrics, along with computational cost analysis. The results indicate superior performance and greater stability of traditional classifiers (KNN and SVM) compared to CNN in the evaluated scenario, highlighting the impact of *Dataset* size and model complexity on learning effectiveness.

**Resumo.** Este trabalho apresenta uma avaliação experimental comparativa de técnicas de aprendizagem de máquina aplicadas à classificação de imagens de vias rodoviárias com e sem buracos. Foram analisados os algoritmos *K-Nearest Neighbors* (KNN), Máquinas de Vetores de Suporte (MVS) e Redes Neurais Convolucionais (RNC), utilizando os mesmos conjuntos de dados e protocolo experimental padronizado. Cada modelo foi executado 50 vezes, variando a semente aleatória para garantir maior robustez estatística. O desempenho foi avaliado por meio das métricas de Acurácia, Precisão, *Recall* e *F1-Score*, além da análise do custo computacional. Os resultados indicaram desempenho superior e maior estabilidade dos classificadores tradicionais (KNN e MVS) em comparação à RNC no cenário avaliado, destacando a influência do tamanho do *Dataset* e da complexidade do modelo no processo de aprendizagem.

## 1. Introdução

A integridade das infraestruturas rodoviárias exerce papel fundamental no suporte à mobilidade urbana e interurbana, bem como no desenvolvimento econômico, especialmente em países fortemente dependentes do transporte rodoviário, como o Brasil. De acordo com a Confederação Nacional do Transporte, uma parcela significativa da malha viária brasileira apresenta algum nível de deficiência estrutural, impactando diretamente os custos logísticos, o consumo de combustível e a segurança dos usuários (CNT, 2023). Nesse contexto, a identificação precoce de defeitos no pavimento torna-se essencial para a manutenção da qualidade das vias e a redução de acidentes.

Entre os principais defeitos observados em pavimentos asfálticos destacam-se as depressões, popularmente conhecidas como buracos (*potholes*). Essas irregularidades

resultam de processos físicos e mecânicos associados à infiltração de água, fadiga dos materiais, variações térmicas e tráfego intenso, levando à perda progressiva da capacidade estrutural da via (HUANG, 2004). A presença de buracos compromete o conforto e a segurança da circulação, estando associada ao aumento do risco de acidentes, danos aos veículos e elevação dos custos de manutenção corretiva.

Tradicionalmente, a identificação de buracos em pavimentos rodoviários é realizada por meio de inspeções visuais manuais ou técnicas semi-automatizadas, que exigem elevado esforço humano, apresentam alto custo operacional e estão sujeitas à subjetividade do avaliador, além de não acompanharem de forma eficiente a extensão da malha viária (HUANG, 2004; CNT, 2023). Em função dessas limitações, abordagens baseadas em visão computacional e aprendizagem de máquina têm sido amplamente investigadas como alternativas para a automação do processo de avaliação das condições do pavimento, possibilitando maior precisão, rapidez e padronização na detecção de defeitos estruturais (ESPÍNDOLA *et al.*, 2021a; ASAD *et al.*, 2022).

Nos últimos anos, diversos estudos têm explorado a aplicação de algoritmos de aprendizagem de máquina para a detecção e classificação de defeitos em pavimentos a partir de imagens, incluindo métodos clássicos, como *K-Nearest Neighbors* (KNN) e Máquinas de Vetores de Suporte (MVS), bem como abordagens mais recentes baseadas em Redes Neurais Convolucionais (RNC) (ESPÍNDOLA *et al.*, 2021a; ASAD *et al.*, 2022). Esses métodos apresentam diferenças quanto à complexidade computacional, à necessidade de dados e à capacidade de extração de padrões, o que torna relevante sua comparação em cenários experimentais controlados.

Dessa forma, este Trabalho de Conclusão de Curso tem como objetivo aplicar e comparar técnicas de aprendizagem de máquina na identificação e classificação de depressões em imagens de vias rodoviárias. Foram avaliados os algoritmos KNN, MVS e RNC, analisando-se seu desempenho preditivo, estabilidade e custo computacional, com o intuito de contribuir para a compreensão das vantagens e limitações de cada abordagem no contexto da detecção automática de buracos em pavimentos.

## **2. Fundamentação Teórica**

### **2.1 Inteligência Artificial**

A Inteligência Artificial (IA) compreende um conjunto de técnicas e métodos da ciência da computação voltados ao desenvolvimento de sistemas capazes de executar tarefas que, tradicionalmente, demandariam inteligência humana, como reconhecimento de padrões, tomada de decisão e aprendizado a partir de dados. Ao longo das últimas décadas, a IA tem sido amplamente incorporada em aplicações de engenharia e infraestrutura, principalmente em cenários que exigem automação, eficiência e redução de custos operacionais.

No contexto da avaliação de pavimentos rodoviários, técnicas de IA permitem automatizar processos de inspeção e monitoramento, superando limitações associadas à subjetividade e à baixa escalabilidade das abordagens manuais, especialmente quando aplicadas a grandes extensões viárias (RUSSELL; NORVIG, 2021).

### **2.2 Redes Neurais Artificiais e Convolucionais**

Dentro desse campo, as Redes Neurais Artificiais (RNAs) destacam-se como modelos computacionais inspirados no funcionamento do sistema nervoso biológico. Essas redes são formadas por unidades interconectadas, denominadas neurônios artificiais, que ajustam seus pesos de forma iterativa durante o processo de treinamento, permitindo o aprendizado de relações complexas e não lineares entre dados de entrada e saída. Devido a essa capacidade de generalização, as RNAs têm sido amplamente utilizadas em tarefas de classificação e reconhecimento de padrões, incluindo aplicações na engenharia civil e na análise de imagens, onde relações não lineares são frequentemente observadas (HAYKIN, 2009; BISHOP, 2006).

As Redes Neurais Convolucionais (RNCs) constituem uma especialização das redes neurais artificiais desenvolvida para o processamento de dados com estrutura espacial, como imagens digitais. Diferentemente dos modelos tradicionais baseados em vetores, as RNCs preservam a organização bidimensional das imagens por meio de camadas convolucionais, que aplicam filtros locais capazes de extrair automaticamente características hierárquicas, como bordas, texturas e padrões visuais mais complexos. Essa característica torna as RNCs particularmente adequadas para tarefas de visão computacional, incluindo a detecção de defeitos em pavimentos rodoviários. No entanto, a literatura ressalta que o desempenho dessas redes está fortemente associado à disponibilidade de grandes volumes de dados e ao uso de arquiteturas suficientemente profundas, uma vez que o elevado número de parâmetros exige diversidade amostral para o aprendizado eficaz de representações discriminativas (GOODFELLOW; BENGIO; COURVILLE, 2016; LECUN; BENGIO; HINTON, 2015).

Em contraste com abordagens baseadas em *deep learning*, algoritmos clássicos de aprendizado de máquina continuam relevantes em diversos cenários práticos. O algoritmo *K-Nearest Neighbors* (KNN), por exemplo, é um método supervisionado baseado em instâncias, no qual a classificação de uma amostra é determinada a partir das classes predominantes entre seus  $k$  vizinhos mais próximos no espaço de características. O KNN não realiza uma fase de treinamento explícita, mantendo os dados de treinamento armazenados e realizando os cálculos apenas no momento da classificação. Apesar de sua simplicidade conceitual, o algoritmo pode apresentar bom desempenho em problemas nos quais as classes estão bem separadas no espaço vetorial, especialmente quando precedido por etapas adequadas de pré-processamento dos dados (BISHOP, 2006; MITCHELL, 1997).

De forma complementar, as Máquinas de Vetores de Suporte (MVS) representam uma abordagem supervisionada voltada à construção de modelos de classificação robustos, baseados na determinação de um hiperplano ótimo que maximize a margem entre classes distintas. Uma das principais vantagens das MVSs reside em sua capacidade de lidar com dados de alta dimensionalidade, especialmente quando associadas a funções de kernel que permitem a separação não linear dos dados. Essas características fazem com que as MVSs sejam amplamente utilizadas em problemas de classificação de imagens e detecção de defeitos estruturais, apresentando um equilíbrio favorável entre desempenho preditivo, estabilidade e custo computacional (CORTES; VAPNIK, 1995; BISHOP, 2006).

Dessa forma, a literatura evidencia que tanto métodos clássicos quanto abordagens baseadas em redes neurais profundas possuem vantagens e limitações específicas, sendo sua adequação fortemente dependente das características do problema, do volume de dados disponível e das restrições computacionais. Esse panorama reforça a relevância de estudos comparativos, como o desenvolvido neste trabalho, que avaliam diferentes técnicas sob condições experimentais controladas, contribuindo para uma compreensão mais abrangente de seu desempenho na detecção automática de buracos em vias rodoviárias.

### 2.3 Métricas de Avaliação

Em problemas de classificação supervisionada, diferentes métricas são utilizadas para avaliar o desempenho dos modelos preditivos. Entre as métricas mais utilizadas destacam-se a Acurácia, Precisão, *Recall* e *F1-Score*, amplamente empregadas na literatura para avaliar classificadores aplicados a imagens e reconhecimento de padrões (BISHOP, 2006; GOODFELLOW *et al.*, 2016).

A Acurácia representa a proporção total de previsões corretas realizadas pelo modelo em relação ao número total de amostras avaliadas. Trata-se de uma métrica global de desempenho, útil quando as classes do conjunto de dados estão balanceadas.

A Precisão indica a proporção de instâncias corretamente classificadas como positivas em relação ao total de instâncias que o modelo classificou como positivas. Essa métrica avalia a confiabilidade das detecções realizadas pelo modelo.

O *Recall*, também conhecido como sensibilidade, mede a capacidade do modelo em identificar corretamente todas as instâncias positivas presentes no conjunto de dados, sendo especialmente importante em aplicações nas quais a não detecção de um evento pode gerar impactos significativos.

Por fim, o *F1-Score* corresponde à média harmônica entre Precisão e *Recall*, sendo utilizado como uma métrica de equilíbrio entre essas duas medidas, especialmente em cenários nos quais existe desbalanceamento entre classes.

## 3. Trabalhos correlatos

A detecção automática de defeitos em pavimentos rodoviários tem sido amplamente investigada na literatura, especialmente em razão dos impactos econômicos e de segurança associados à degradação das vias. Diferentes abordagens vêm sendo propostas, variando desde métodos tradicionais de aprendizado de máquina até técnicas mais recentes baseadas em redes neurais profundas, empregando distintos tipos de dados, como imagens, sinais de vibração e informações sensoriais.

Estudos iniciais e revisões sistemáticas indicam que algoritmos clássicos de aprendizado de máquina, como Máquinas de Vetores de Suporte (MVS) e *K-Nearest Neighbors* (KNN), têm sido utilizados com sucesso em tarefas de monitoramento e avaliação de pavimentos, especialmente quando aplicados a conjuntos de dados estruturados ou imagens previamente processadas (CANO-ORTIZ; PASCUAL-MUÑOZ; CASTRO-FRESNO, 2022). Esses trabalhos demonstram que, apesar de não explorar explicitamente relações espaciais complexas, tais métodos

podem alcançar resultados satisfatórios quando os padrões discriminativos entre as classes são suficientemente evidentes.

Com o avanço da visão computacional, abordagens baseadas em imagens passaram a ganhar destaque. Espíndola *et al.* (2021a) aplicaram Redes Neurais Convolucionais para a detecção de buracos e remendos em pavimentos asfálticos, obtendo bons resultados em termos de acurácia e *F1-score*, especialmente quando utilizados conjuntos de dados mais robustos. Espíndola *et al.* (2021b) também investigaram a classificação automática do tipo de pavimento por meio de RNC, evidenciando o potencial do *deep learning* na automação de sistemas de avaliação viária. De forma semelhante, Asad *et al.* (2022) propuseram um sistema de detecção de buracos baseado em *deep learning* com foco em aplicações em tempo real e computação embarcada, reforçando o potencial das RNCs para esse tipo de problema, embora destacando o elevado custo computacional associado ao treinamento desses modelos.

Outros trabalhos exploram arquiteturas ainda mais complexas, combinando diferentes tipos de dados. Abordagens que utilizam imagens RGB associadas a imagens térmicas demonstram ganhos de desempenho na identificação de defeitos, porém exigem sensores especializados e maior complexidade computacional, o que limita sua aplicabilidade em cenários de baixo custo (CHEN; CHANDRA; SEO, 2022). De maneira alternativa, alguns estudos investigam a detecção de buracos a partir de dados de vibração veicular, utilizando redes neurais convolucionais para analisar sinais temporais, o que evidencia a diversidade de estratégias existentes para o mesmo problema (OZOGLU; GÖKGÖZ, 2023).

Revisões recentes da literatura destacam que não existe uma abordagem universalmente superior para a detecção de buracos em pavimentos, sendo o desempenho fortemente influenciado pelo tipo de dado, pelo tamanho do conjunto de treinamento e pela complexidade do modelo adotado (SAFYARI; MAHDIPANAH; SHIRI, 2024). Em particular, estudos comparativos indicam que, em cenários com conjuntos de dados limitados, métodos tradicionais como MVS podem apresentar desempenho semelhante ou até superior ao de redes neurais profundas, além de maior estabilidade e menor custo computacional (BASHA *et al.*, 2024).

Nesse contexto, observa-se que, embora as Redes Neurais Convolucionais apresentem grande potencial para a detecção automática de defeitos em pavimentos, seu desempenho está fortemente condicionado à disponibilidade de grandes volumes de dados e a arquiteturas adequadamente profundas, conforme discutido na literatura recente. Por outro lado, algoritmos clássicos permanecem relevantes, especialmente em aplicações onde a simplicidade, a interpretabilidade e a eficiência computacional são fatores determinantes.

Diferentemente de muitos trabalhos que avaliam apenas uma única abordagem, o presente estudo propõe uma análise comparativa entre métodos clássicos (KNN e MVS) e uma abordagem baseada em *deep learning* (RNC), utilizando os mesmos conjuntos de dados e protocolo experimental padronizado. Essa comparação direta permite avaliar não apenas o desempenho preditivo, mas também a estabilidade dos modelos e o custo computacional associado, contribuindo para uma compreensão mais abrangente das

vantagens e limitações de cada técnica no contexto da classificação automática de buracos em vias rodoviárias.

#### 4. Metodologia

A metodologia adotada foi estruturada de forma a garantir uma comparação justa entre os algoritmos selecionados, utilizando a mesma base de dados, métricas de avaliação padronizadas e múltiplas execuções experimentais.

As etapas do processo experimental iniciam-se com o preparo do *Dataset*, seguido pelo pré-processamento das imagens, divisão dos dados, treinamento dos modelos e avaliação por métricas de desempenho (Figura 1).



Figura 1 – Fluxo metodológico do experimento

#### 5. Conjunto de Dados

Para a realização dos experimentos foram utilizadas três bases públicas de imagens de vias rodoviárias contendo exemplos de pavimentos com e sem depressões. As bases foram obtidas na plataforma *Kaggle*, amplamente utilizada para compartilhamento de conjuntos de dados voltados a aplicações em ciência de dados e aprendizagem de máquina.

As bases utilizadas foram o *Pothole Detection Dataset* (ATULYAKUMAR, 2018), o *Pothole and Plain Road Images* (VIRENBR11, 2020) e o *MIIA Pothole Image Classification Challenge* (HAMMADI, 2023).

No total, foram utilizadas imagens provenientes dessas três bases, totalizando diferentes cenários de pavimentos rodoviários. Cada base contém imagens classificadas em duas categorias principais: presença de depressões na via (*potholes*) e pavimento em condições normais.

Os conjuntos de dados utilizados apresentaram tamanhos distintos, permitindo avaliar o impacto do volume amostral no desempenho dos modelos. O *Dataset 1* é composto por 681 imagens, o *Dataset 3* por 740 imagens e o *Dataset 4* por 5.679 imagens. Essa variação possibilita analisar o comportamento dos algoritmos tanto em cenários com volume reduzido quanto ampliado de dados, aspecto particularmente relevante para modelos baseados em *deep learning*.

O *Dataset 4* contém originalmente 5.679 imagens. No entanto, apenas 4026 imagens possuem rótulos disponíveis no arquivo *train\_ids\_labels.csv*, sendo estas utilizadas no treinamento e avaliação dos modelos. As demais imagens pertencem ao conjunto de teste da competição original e não possuem rótulos públicos, não sendo utilizadas neste estudo.

Inicialmente, quatro bases de dados foram consideradas para a realização dos experimentos. No entanto, uma das bases analisadas, denominada *Dataset 2*, apresentou

uma quantidade reduzida de imagens, totalizando apenas 60 amostras, valor significativamente inferior ao das demais bases avaliadas.

Devido ao baixo volume de dados disponíveis, essa base foi descartada da etapa experimental, pois conjuntos de dados muito pequenos podem comprometer a capacidade de generalização dos modelos de aprendizagem de máquina e introduzir viés nos resultados obtidos. Dessa forma, os experimentos foram conduzidos utilizando apenas as três bases com maior quantidade de imagens, garantindo maior robustez na avaliação dos algoritmos.

## 6. Algoritmos Avaliados

Três técnicas foram selecionadas para a classificação: o KNN, que atua por similaridade de vizinhança; o MVS, focado na definição de um hiperplano ótimo de separação; e a RNC, que utiliza camadas convolucionais para extrair automaticamente as características das imagens. Enquanto os classificadores tradicionais (KNN e MVS) foram estruturados com a biblioteca *scikit-learn*, a implementação da rede profunda (RNC) foi realizada sobre o *framework TensorFlow/Keras*.

## 7. Pré-processamento das Imagens

Antes do processamento, as imagens passaram por uma etapa de padronização, definidas de acordo com as exigências de cada algoritmo avaliado, considerando a forma de representação dos dados e o custo computacional. Algoritmos clássicos, como KNN e MVS, requerem que as imagens sejam transformadas em vetores numéricos, não explorando diretamente relações espaciais entre pixels, enquanto as Redes Neurais Convolucionais (RNC) exigem entradas matriciais com preservação da estrutura espacial da imagem, o que justifica a adoção de diferentes estratégias de pré-processamento (BISHOP, 2006; GOODFELLOW *et al.*, 2016).

Nos algoritmos clássicos (KNN e MVS), foi aplicado o processo de *flattening*, convertendo as imagens em vetores unidimensionais para permitir a aplicação de classificadores baseados em espaço vetorial (Figura 2).

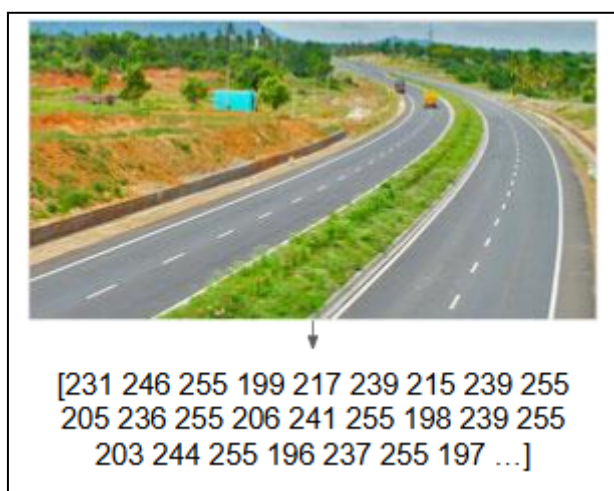


Figura 2 - Representação vetorial utilizada por KNN e MVS

Para o algoritmo *K-Nearest Neighbors* (KNN), as imagens foram redimensionadas para  $250 \times 250$  pixels e convertidas em vetores unidimensionais por

meio do processo de *flattening*. Foi adotado  $n\_neighbors = 5$ , utilizando a métrica de distância Euclidiana padrão da biblioteca scikit-learn.

No caso das Máquinas de Vetores de Suporte (MVS), as imagens foram convertidas para escala de cinza, redimensionadas para  $100 \times 100$  pixels e posteriormente transformadas em vetores por meio do processo de *flattening*. Foi utilizado o classificador *SVC* da biblioteca scikit-learn com kernel *RBF* (padrão) e parâmetro de regularização  $C = 1.0$  (Figura 2).

A Rede Neural Convolutiva foi estruturada com imagens redimensionadas para  $64 \times 64$  pixels com três canais de cor (RGB). A arquitetura adotada consistiu em duas camadas convolucionais com 32 filtros e kernel  $3 \times 3$ , ativação *ReLU*, cada uma seguida por camadas de *MaxPooling* ( $2 \times 2$ ). Após as camadas convolucionais, aplicou-se a operação de *Flatten*, seguida de uma camada densa com 128 neurônios (*ReLU*) e uma camada de saída com um neurônio e ativação sigmoide para classificação binária (Figura 3).



Figura 3 - Representação espacial utilizada pela RNC

O modelo foi treinado durante 25 épocas, com batch size igual a 32, utilizando o otimizador *Adam* e a função de perda *Binary Crossentropy*. Foi adotado *validation\_split* de 20% para avaliação durante o treinamento.

O conjunto foi dividido em 80% para treinamento e 20% para teste. No caso da RNC, uma fração do conjunto de treinamento foi separada automaticamente para validação ( $validation\_split = 0,2$ ) durante o processo de ajuste.

## 8. Protocolo Experimental

Para garantir a validade estatística e mitigar o impacto da aleatoriedade na separação dos dados, cada algoritmo foi submetido a 50 execuções independentes. Em cada iteração, utilizou-se uma semente aleatória (random seed) distinta para o particionamento entre treino e teste, permitindo observar a estabilidade dos modelos frente a diferentes cenários. O desempenho foi mensurado por meio de um conjunto de métricas consolidadas: Acurácia, Precisão, *Recall* e *F1-Score*. Tais indicadores, calculados individualmente a cada rodada, serviram de base para uma análise comparativa fundamentada no comportamento médio e na robustez de cada técnica.

A escolha dessas métricas está em acordo com trabalhos recentes da literatura voltados à detecção de buracos em pavimentos, nos quais esses indicadores são amplamente utilizados para avaliar o desempenho de modelos de classificação aplicados a imagens viárias (ESPÍNDOLA *et al.*, 2021a; ASAD *et al.*, 2022)

Os experimentos foram conduzidos em ambiente local para assegurar o controle das variáveis de execução e a reprodutibilidade dos testes. A configuração de hardware utilizada dispõe de um processador AMD Ryzen 7 5700G, 16 GB de memória RAM e uma placa de vídeo NVIDIA GeForce GTX 1060 (3 GB). No que diz respeito ao software, os algoritmos foram implementados em *Python* (versão 3.10.11) sobre o sistema operacional Windows 11.

Durante os testes, o sistema foi mantido operando apenas com os processos essenciais para evitar interferências externas no desempenho. Essas especificações ajudam a contextualizar os resultados e os tempos obtidos nas 50 rodadas de execução.

Os códigos utilizados para implementação dos algoritmos RNC, KNN e MVS, bem como os scripts responsáveis pelo processamento das bases de dados e execução dos experimentos, estão disponíveis publicamente em um repositório online, acessível em: <https://github.com/Jhonatandbz/pothole-classification-cnn-knn-svm>.

## 9. Resultados

Após a realização das 50 execuções independentes para cada algoritmo, foram obtidos os resultados quantitativos referentes às métricas de Acurácia, Precisão, *Recall* e *F1-Score*, além do tempo de execução. Os resultados apresentados a seguir referem-se aos valores médios observados ao longo das execuções para os algoritmos RNC, KNN e MVS, considerando os diferentes conjuntos de dados avaliados.

A Tabela 1 apresenta o desempenho médio dos algoritmos em cada *Dataset*. Observa-se que os valores variam de acordo com o conjunto de dados utilizado, com destaque para o *Dataset 4*, que apresentou os maiores índices de desempenho para todos os algoritmos.

<b>Algoritmo</b>	<b><i>Dataset</i></b>	<b>Tempo (s)</b>	<b>Acuracia</b>	<b><i>F1-Score</i></b>	<b>Precisao</b>	<b><i>Recall</i></b>
RNC	<i>Dataset 4</i>	272,551	0,5178	0,405	0,412	0,399
RNC	<i>Dataset 3</i>	165,306	0,4992	0,493	0,492	0,495
RNC	<i>Dataset 1</i>	150,531	0,5038	0,494	0,485	0,506
MVS	<i>Dataset 4</i>	28,857	0,8889	0,851	0,950	0,772
KNN	<i>Dataset 4</i>	18,939	0,9273	0,909	0,941	0,880
KNN	<i>Dataset 3</i>	9,449	0,7406	0,757	0,708	0,815
KNN	<i>Dataset 1</i>	8,551	0,7239	0,728	0,698	0,764
MVS	<i>Dataset 3</i>	6,959	0,7372	0,750	0,711	0,796
MVS	<i>Dataset 1</i>	6,367	0,7277	0,731	0,703	0,763

Tabela 1 – Desempenho médio dos algoritmos RNC, KNN e MVS nos diferentes conjuntos de dados

Observa-se a evolução das métricas ao longo das 50 execuções. Nota-se que os algoritmos KNN e MVS mantiveram comportamento mais estável ao longo das rodadas, enquanto a RNC apresentou maior variação, especialmente nas métricas de Precisão e *Recall* (Figuras 4–7).

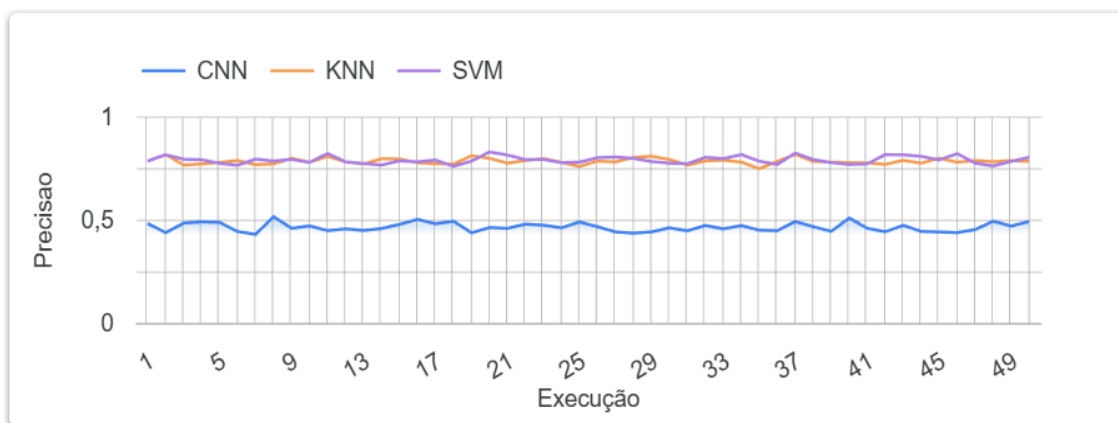


Figura 4 – Evolução da Precisão ao longo das 50 execuções para os algoritmos RNC, KNN e MVS

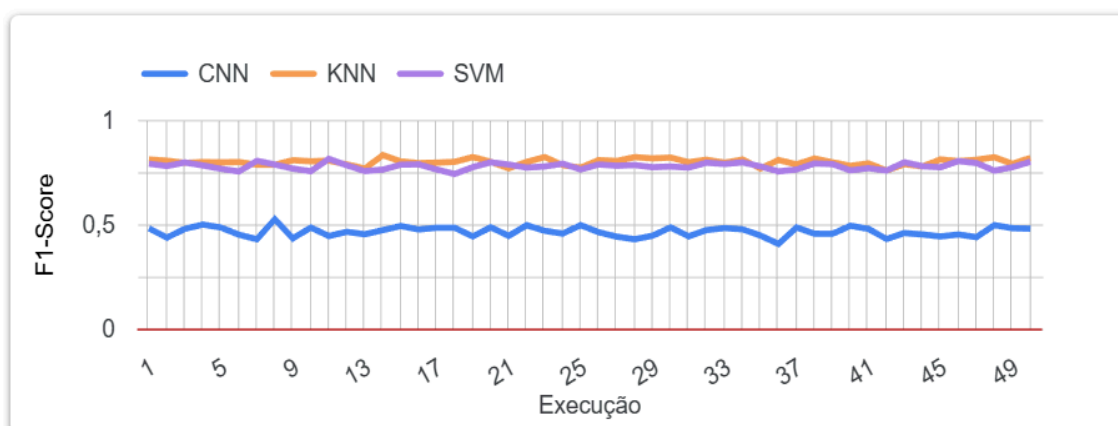


Figura 5 – Evolução do *F1-Score* ao longo das 50 execuções para os algoritmos RNC, KNN e MVS

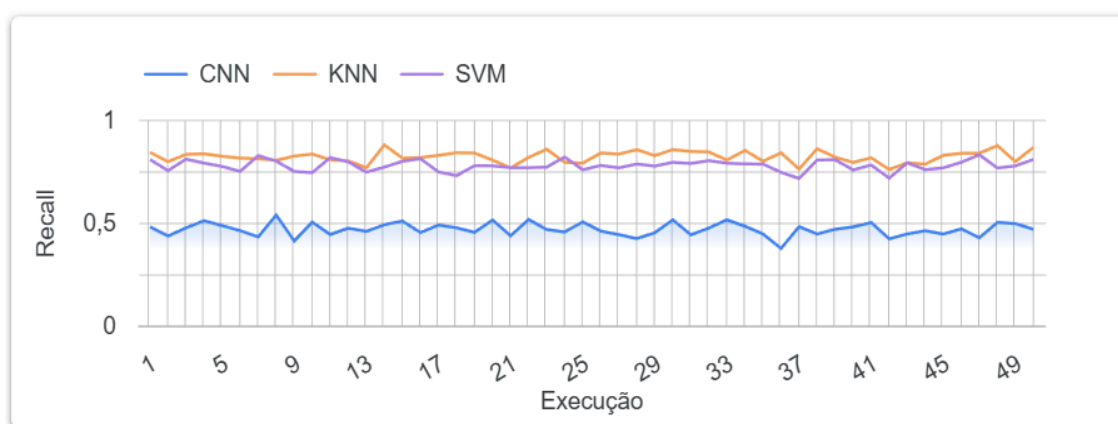


Figura 6 – Evolução do *Recall* ao longo das 50 execuções para os algoritmos RNC, KNN e MVS

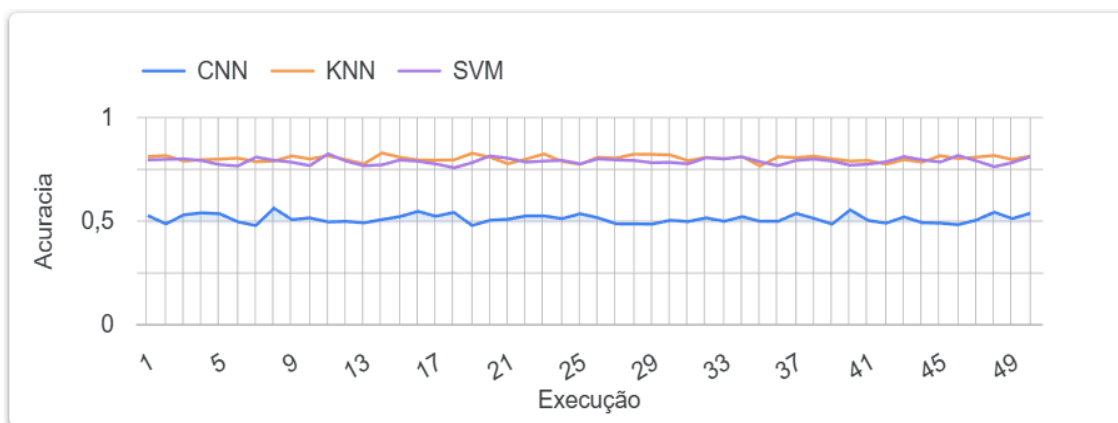


Figura 7 – Evolução da Acurácia ao longo das 50 execuções para os algoritmos RNC, KNN e MVS

Em relação ao tempo de execução, os resultados indicam diferenças expressivas entre os algoritmos. Enquanto KNN e MVS concluíram suas execuções em poucos segundos, a RNC demandou tempos significativamente maiores (Figura 8).

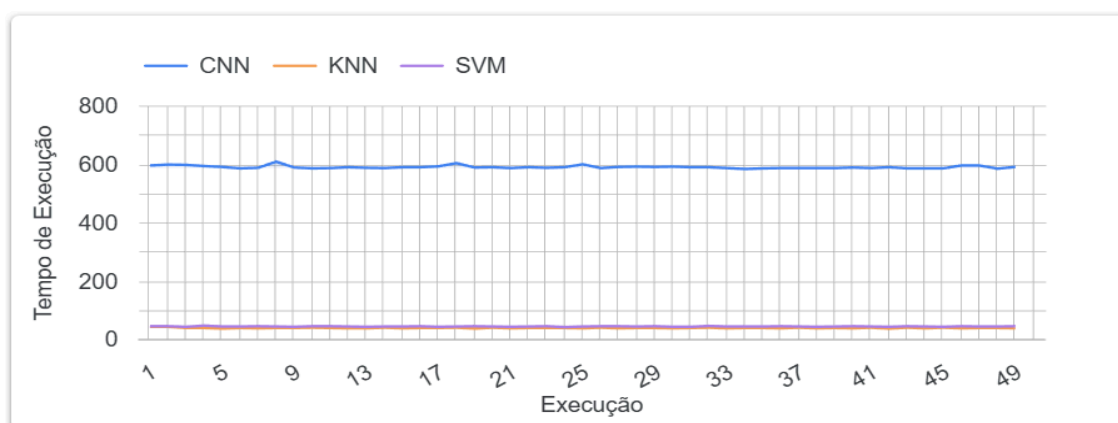


Figura 8 – Tempo de execução dos algoritmos RNC, KNN e MVS ao longo das 50 execuções

## 10. Discussões

Os resultados obtidos evidenciam diferenças claras de desempenho e estabilidade entre os algoritmos avaliados, permitindo uma análise crítica de suas aplicações no contexto da classificação de imagens de buracos em vias rodoviárias. De modo geral, os classificadores tradicionais KNN e MVS apresentaram desempenho superior ao da RNC em todas as métricas analisadas, além de maior estabilidade ao longo das 50 execuções independentes.

Na literatura, o desempenho de modelos aplicados à detecção de defeitos em pavimentos é considerado satisfatório quando métricas como Precisão, *Recall* e, principalmente, *F1-Score* apresentam valores elevados e comportamento estável entre diferentes execuções experimentais. Estudos como Espíndola *et al.* (2021a) e Asad *et al.* (2022) adotam essas métricas como critérios centrais de avaliação, destacando que valores consistentes de *F1-Score* e *Recall* são fundamentais em aplicações de segurança viária, nas quais falhas de detecção possuem impacto direto. Nesse contexto, os

resultados obtidos para os algoritmos KNN e MVS neste trabalho situam-se dentro da faixa de desempenho considerada adequada pela literatura, apresentando boa capacidade de detecção e estabilidade.

A superioridade do KNN e do MVS pode ser explicada, em parte, pela natureza do problema e pelas características do conjunto de dados utilizado. Ambos os algoritmos operam diretamente sobre vetores de características e, no presente trabalho, utilizaram imagens previamente redimensionadas e transformadas por meio do processo de *flattening*. Essa abordagem, embora elimine informações espaciais, mostrou-se suficiente para capturar padrões discriminativos entre imagens com e sem buracos, resultando em valores elevados e consistentes de Precisão, *Recall* e *F1-Score*.

Por outro lado, a Rede Neural Convolutiva apresentou desempenho inferior e maior variabilidade entre as execuções. Esse comportamento pode estar associado a fatores como o tamanho limitado do conjunto de dados, a simplicidade da arquitetura adotada e o número reduzido de épocas de treinamento, que podem não ter sido suficientes para que a rede aprendesse representações mais complexas e generalizáveis. A literatura aponta que Redes Neurais Convolutivas tendem a apresentar melhores resultados quando treinadas com grandes volumes de dados e arquiteturas mais profundas, uma vez que o elevado número de parâmetros exige maior diversidade amostral para o aprendizado eficaz de representações discriminativas. Em cenários com conjuntos de dados limitados, tais modelos podem apresentar desempenho inferior, conforme discutido por Goodfellow *et al.* (2016) e LeCun *et al.* (2015), o que ajuda a explicar os resultados observados neste estudo.

A análise das métricas de Precisão e *Recall* revela aspectos relevantes para aplicações práticas. Enquanto a Precisão indica a confiabilidade das detecções realizadas, o *Recall* reflete a capacidade do modelo em identificar corretamente os buracos presentes nas imagens, sendo esta última métrica particularmente importante em cenários de segurança viária. Nesse sentido, os resultados mostram que KNN e MVS mantiveram valores de *Recall* elevados, reduzindo a probabilidade de falhas por omissão, o que reforça sua adequação para aplicações em que a não detecção de defeitos pode acarretar riscos significativos.

Em relação ao custo computacional, observou-se que os algoritmos KNN e MVS apresentaram tempos de execução significativamente inferiores aos da RNC em todos os conjuntos de dados avaliados. Esse comportamento já era esperado, visto que o treinamento de redes neurais profundas exige um elevado custo computacional, inerente ao caráter iterativo de seu processo de otimização ao longo das épocas. Assim, considerando simultaneamente desempenho preditivo e eficiência computacional, os classificadores tradicionais demonstraram uma relação custo-benefício mais favorável no contexto experimental analisado.

Ressalta-se que os resultados obtidos estão condicionados às características dos conjuntos de dados utilizados, à arquitetura simplificada da RNC e à ausência de técnicas de aumento de dados (*data augmentation*). Assim, os achados não visam desqualificar o potencial das redes neurais profundas, mas contextualizar seu desempenho no cenário experimental analisado.

Por fim, a variação de desempenho observada entre os diferentes conjuntos de dados reforça a influência da qualidade, padronização e distribuição das imagens no aprendizado dos modelos. O melhor desempenho obtido em determinados *Datasets* sugere que características específicas dos dados podem facilitar a discriminação entre classes, aspecto que deve ser considerado em aplicações reais. Dessa forma, os resultados obtidos não apenas comparam os algoritmos, mas também evidenciam a importância do conjunto de dados na eficácia de sistemas automáticos de detecção de buracos.

## 11. Considerações Finais

Este trabalho teve como objetivo aplicar e comparar diferentes técnicas de aprendizagem de máquina na identificação e classificação de buracos em imagens de vias rodoviárias, avaliando os algoritmos *K-Nearest Neighbors* (KNN), Máquinas de Vetores de Suporte (MVS) e Redes Neurais Convolucionais (RNC) sob condições experimentais controladas. Para isso, foi adotado um protocolo padronizado de pré-processamento, particionamento dos dados e avaliação, permitindo uma análise justa e reproduzível do desempenho de cada abordagem.

Os resultados obtidos evidenciaram que, no cenário analisado, os classificadores tradicionais KNN e MVS apresentaram desempenho superior ao da RNC em todas as métricas avaliadas, além de maior estabilidade ao longo das 50 execuções independentes. Esses métodos também se destacaram pelo baixo custo computacional, apresentando tempos de execução significativamente inferiores aos observados no RNC.

A rede neural convolucional demonstrou desempenho mais modesto e maior variabilidade entre as execuções, especialmente em métricas críticas como *Recall* e *F1-Score*, o que indica limitações em sua capacidade de generalização nas condições adotadas.

Dessa forma, os resultados reforçam que, para problemas de classificação de buracos em pavimentos com conjuntos de dados limitados e restrições computacionais, métodos clássicos de aprendizagem de máquina podem oferecer uma relação custo-benefício mais favorável, sem comprometer a confiabilidade das detecções.

Apesar dos resultados consistentes, algumas limitações devem ser consideradas na interpretação dos achados. Primeiramente, a RNC empregada neste estudo utilizou uma arquitetura relativamente simples e um número restrito de épocas de treinamento, o que pode ter limitado sua capacidade de aprendizado de representações mais complexas. Além disso, não foram aplicadas técnicas de aumento de dados (*data augmentation*), que poderiam contribuir para a ampliação da diversidade amostral e para a melhoria do desempenho de modelos baseados em deep learning.

Outro aspecto relevante diz respeito às características dos conjuntos de dados utilizados, que apresentaram variações quanto à quantidade de imagens, padronização e condições de captura. Essas diferenças podem influenciar diretamente o desempenho dos algoritmos e restringem a generalização dos resultados para outros cenários não avaliados. Ademais, o uso do processo de *flattening* nos métodos KNN e MVS, embora eficaz neste contexto, implica a perda de informações espaciais que poderiam ser exploradas por abordagens mais avançadas.

Como trabalhos futuros, sugere-se a ampliação deste estudo por meio da utilização de conjuntos de dados maiores e mais diversificados. A investigação de arquiteturas de RNC mais profundas e modernas, bem como a aplicação de técnicas de data augmentation e transfer learning, também representa uma linha promissora para avaliar o real potencial das redes neurais convolucionais neste domínio.

Além disso, futuras pesquisas podem explorar a integração de informações provenientes de múltiplas fontes de dados, como sensores inerciais, imagens térmicas ou dados de vídeo, visando aumentar a robustez dos sistemas de detecção. Por fim, a avaliação dos algoritmos em ambientes de execução embarcados ou em tempo real pode contribuir para aproximar as soluções propostas de aplicações práticas no contexto da gestão e manutenção de infraestruturas rodoviárias.

### **Referências Bibliográficas**

- ASAD, M. H. et al. Pothole Detection Using Deep Learning: A Real-Time and AI-on-the-Edge Perspective. *Advances in Civil Engineering*, v. 2022, Article ID 7864623, 2022.
- BISHOP, C. M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- CONFEDERAÇÃO NACIONAL DO TRANSPORTE (CNT). *Pesquisa CNT de Rodovias 2023*. Brasília: CNT; SEST SENAT; ITL, 2023.
- CORTES, C.; VAPNIK, V. Support-Vector Networks. *Machine Learning*, v. 20, n. 3, p. 273–297, 1995.
- ESPÍNDOLA, A. C. et al. Pothole and Patch Detection on Asphalt Pavement Using Deep Convolutional Neural Network. In: *Proceedings of the XLII Ibero-Latin-American Congress on Computational Methods in Engineering (CILAMCE-PANACM)*. Rio de Janeiro, 2021a.
- ESPÍNDOLA, A. C. et al. Pavement Surface Type Classification Based on Deep Learning to the Automatic Pavement Evaluation System. In: *Proceedings of the XLII Ibero-Latin-American Congress on Computational Methods in Engineering (CILAMCE-PANACM)*. Rio de Janeiro, 2021b.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge: MIT Press, 2016.
- HAYKIN, S. *Neural Networks and Learning Machines*. 3. ed. Upper Saddle River: Pearson Education, 2009.
- HUANG, Y. H. *Pavement Analysis and Design*. 2. ed. Upper Saddle River: Pearson Prentice Hall, 2004.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep Learning. *Nature*, v. 521, n. 7553, p. 436–444, 2015.
- MITCHELL, T. M. *Machine Learning*. New York: McGraw-Hill, 1997.
- RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 4. ed. Hoboken: Pearson, 2021.

- CANO-ORTIZ, S.; PASCUAL-MUÑOZ, P.; CASTRO-FRESNO, D. Machine learning algorithms for monitoring pavement performance. *Automation in Construction*, v. 139, p. 104309, 2022.
- CHEN, C.; CHANDRA, S.; SEO, H. Automatic Pavement Defect Detection and Classification Using RGB-Thermal Images Based on Hierarchical Residual Attention Network. *Sensors*, v. 22, n. 15, p. 5781, 2022.
- SAFYARI, Y.; MAHDIPANAH, M.; SHIRI, H. A review of vision-based pothole detection methods using computer vision and machine learning. *Sensors*, v. 24, n. 17, p. 5652, 2024.
- OZOGLU, F.; GÖKGÖZ, T. Detection of road potholes by applying CNN based on road vibration data. *Sensors*, v. 23, 2023.
- BASHA, S. et al. Precision crack analysis in concrete structures using CNN and SVM. *Revista Matéria*, 2024.
- ATULYAKUMAR. Pothole Detection Dataset. Kaggle, 2018. Disponível em: <https://www.kaggle.com/datasets/atulyakumar98/pothole-detection-dataset/data>. Acesso em: 20 nov. 2025.
- VIRENBR11. Pothole and Plain Road Images. Kaggle, 2020. Disponível em: <https://www.kaggle.com/datasets/virenbr11/pothole-and-plain-rode-images>. Acesso em: 20 nov. 2025.
- HAMMADI, S. MIIA Pothole Image Classification Challenge. Kaggle, 2023. Disponível em: <https://www.kaggle.com/datasets/salimhammadi07/miia-pothole-image-classification-challenge>. Acesso em: 20 nov. 2025.