



UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUITA FILHO”

---

FCT - Faculdade de Ciências e Tecnologia  
DMC - Departamento de Matemática e Computação  
Bacharelado em Ciência da Computação

Trabalho de Conclusão de Curso  
(Modalidade Trabalho Acadêmico)

Revisão bibliográfica

**Aprendizado profundo na detecção de anomalias em dados ferroviários**

Orientador: Prof. Dr. Cassio Machiaveli Oishi

Autor: Miguel de Campos Rodrigues Moret

Presidente Prudente, 20 de novembro de 2025

## **Aprendizado profundo na detecção de anomalias em dados ferroviários**

Revisão bibliográfica apresentado ao curso de Ciência da Computação do Departamento de Matemática e Computação da Universidade Estadual Júlio de Mesquita Filho - Faculdade de Ciências e Tecnologia como requisito para a aprovação na disciplina de Projeto Científico I.

Presidente Prudente – São Paulo  
2025

# **Aprendizado profundo na detecção de anomalias em dados ferroviários**

Trabalho aprovado, Presidente Prudente, 2025:

---

**Prof. Dr. Cassio Machiaveli Oishi**

Orientador

---

**Prof. Dr. Nome Avaliador**

Convidado 1

Presidente Prudente – São Paulo

2025

## **RESUMO**

A detecção automática de defeitos ferroviários tem se tornado um tema de grande relevância devido à necessidade de aumentar a segurança, reduzir custos de manutenção e melhorar a eficiência operacional das ferrovias. Métodos tradicionais de inspeção, realizados manualmente, são caros, lentos e sujeitos a erros humanos, o que motiva o uso de técnicas de visão computacional e aprendizado profundo. Este trabalho apresenta uma análise comparativa de diferentes modelos de deep learning aplicados à detecção de anomalias em trilhos ferroviários por meio de imagens. São investigadas arquiteturas clássicas de redes convolucionais, métodos híbridos e transformadores visuais, além de modelos especializados projetados para o domínio ferroviário. O estudo inclui etapas de pré-processamento e transferência de aprendizado, seguindo as melhores práticas encontradas na literatura. Os resultados esperados envolvem melhor desempenho em relação aos métodos tradicionais, maior robustez em condições variadas e identificação das arquiteturas mais adequadas para a tarefa. O objetivo geral é contribuir para o desenvolvimento de tecnologias mais precisas e acessíveis para manutenção preditiva da infraestrutura ferroviária.

## **PALAVRAS-CHAVES**

Detecção de anomalias; Visão computacional; Aprendizado profundo; Transferência de aprendizado; Transformadores visuais; Infraestrutura ferroviária.

## **ABSTRACT**

The automatic detection of railway defects has become increasingly relevant due to the growing demand for improving safety, reducing maintenance costs, and enhancing the operational efficiency of railway systems. Traditional inspection methods, performed manually, are often slow, costly, and prone to human error, which motivates the adoption of computer vision and deep learning techniques. This work presents a comparative analysis of several deep learning models applied to the detection of rail anomalies using image data. The study investigates classical convolutional neural networks, hybrid approaches, visual transformers, and architectures specifically designed for railway inspection. It also includes preprocessing strategies and transfer learning techniques based on best practices from the literature. The expected outcomes include superior performance compared to traditional methods, improved robustness under varying imaging conditions, and the identification of the most suitable architectures for automatic defect detection. The overall goal is to support the development of accurate and accessible technologies for predictive maintenance of railway infrastructure.

## **KEYWORDS**

Rail defect detection; Computer vision; Deep learning; Transfer learning; Vision Transformers; Railway infrastructure.

## LISTA DE FIGURAS

1	Exemplificação das múltiplas camadas de um modelo de aprendizado profundo	10
2	Exemplificação de uma rede neural convolucional . . . . .	11
3	Exemplificação de transferencia de aprendizado . . . . .	12
4	Exemplificação do funcionamento de um transformador visual . . . . .	13
5	Exemplificação do funcionamento de uma arquitetura SVM+CNN . . . . .	14
6	Conjunto de imagens do segundo dataset em grade 1x2 . . . . .	15
7	Conjunto de imagens do primeiro dataset em grade 2x2 . . . . .	16
8	Imagen mostrando o exemplo de uma matriz de confusão . . . . .	18

## SUMÁRIO

<b>1</b>	<b>Introdução</b>	<b>8</b>
1.1	Justificativa . . . . .	8
1.2	Objetivos . . . . .	9
<b>2</b>	<b>Fundamentação teórica</b>	<b>9</b>
2.1	Aprendizado profundo . . . . .	9
2.2	Rede neural convolucional . . . . .	10
2.3	Transferência de aprendizado . . . . .	11
2.4	Transformadores visuais . . . . .	12
2.5	SVM . . . . .	14
<b>3</b>	<b>Metodologia</b>	<b>14</b>
3.1	Datasets . . . . .	14
3.2	Modelos . . . . .	16
3.3	Resultados Esperados . . . . .	17
3.4	Avaliação dos Resultados . . . . .	18
<b>4</b>	<b>Referências</b>	<b>19</b>

## 1 Introdução

O setor ferroviário possui papel estratégico no transporte de cargas e passageiros em escala continental. Estudos recentes destacam o potencial do modal ferroviário para aumentar a eficiência logística e reduzir custos operacionais, além de contribuir para a mobilidade sustentável. Entretanto, a infraestrutura ferroviária enfrenta problemas recorrentes, como desgaste de trilhos, falhas mecânicas, degradação de componentes estruturais e atrasos operacionais, os quais comprometem a segurança e a confiabilidade do sistema (Tiong; Ma; Palmqvist, 2023).

Tradicionalmente, inspeções ferroviárias são realizadas por equipes especializadas, de forma manual e periódica. Embora essenciais, tais procedimentos tendem a ser lento, subjetivos e onerosos, podendo representar mais de 50% dos custos totais de manutenção em alguns países (Huang *et al.*, 2018). Além do mais, a dependência exclusiva da inspeção humana dificulta a detecção precoce de falhas, especialmente em redes extensas ou de difícil acesso.

Com o avanço da automação e das tecnologias associadas à Indústria 4.0, métodos de inspeção não destrutiva (NDE) têm ganhado destaque. Dispositivos embarcados, sensores, sistemas ópticos, varredura a laser e imageamento de alta resolução permitem monitoramento contínuo e preciso da infraestrutura sem interromper a operação ferroviária (Tiong; Ma; Palmqvist, 2023). Esses avanços tornaram viável o desenvolvimento de sistemas inteligentes capazes de identificar anomalias com maior agilidade e confiabilidade.

No entanto, métodos clássicos de processamento de imagens apresentam limitações diante da elevada variabilidade visual do ambiente ferroviário e das taxas significantemente altas de falsos positivos. (Sarhani; Voß, 2024). Esse cenário evidencia a necessidade de técnicas mais robustas e adaptáveis, capazes de lidar com a complexidade e diversidade das imagens coletadas em campo.

### 1.1 Justificativa

Dante desses avanços, este trabalho concentra-se na investigação de métodos de aprendizado profundo aplicados à detecção automática de anomalias em trilhos ferroviários a partir de imagens. O objetivo é avaliar diferentes modelos, analisar seu desempenho e discutir suas potencialidades como ferramentas de apoio à manutenção preventiva da infraestrutura ferroviária. Por meio da revisão, sistematização e análise crítica da literatura, busca-se identificar tendências, limitações e oportunidades de pesquisa, contribuindo para o desenvolvimento de soluções mais eficientes e acessíveis para o setor.

## 1.2 Objetivos

Sendo assim, este trabalho tem como objetivo investigar o uso de técnicas de aprendizado profundo para a detecção de anomalias em dados ferroviários, buscando desenvolver e avaliar modelos capazes de identificar automaticamente falhas estruturais em trilhos por meio de imagens. Para isso, serão conduzidas etapas de pré-processamento e análise exploratória dos dados, implementação de diferentes arquiteturas de aprendizado profundo, avaliação de desempenho por meio de métricas apropriadas e comparação dos resultados com estudos relacionados da literatura.

## 2 Fundamentação teórica

Durante a revisão da literatura, foram identificados modelos e arquiteturas relevantes. No entanto, observa-se que muitos deles derivam de estruturas fundamentais semelhantes, incorporando otimizações, variações ou módulos adicionais para aprimorar o desempenho. Entre os conceitos mais relevantes destacam-se:

### 2.1 Aprendizado profundo

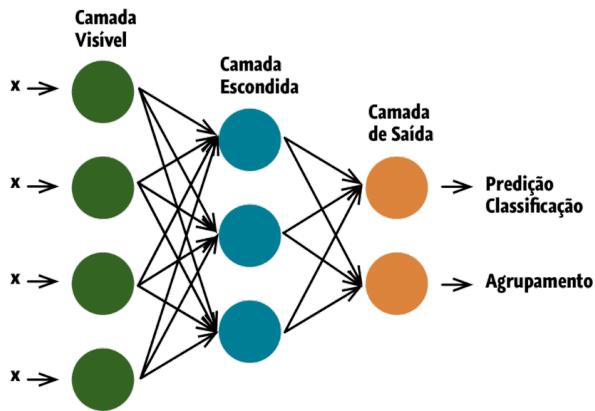
O aprendizado profundo é um subcampo do aprendizado de máquina que se baseia no uso de redes neurais artificiais com múltiplas camadas, capazes de modelar relações complexas em dados. Ao contrário de algoritmos tradicionais de aprendizado de máquina, que frequentemente dependem de extração manual de características, o aprendizado profundo consegue aprender automaticamente representações hierárquicas dos dados.

No contexto de imagens, por exemplo, redes profundas podem aprender a reconhecer padrões visuais em diferentes níveis de abstração: camadas iniciais detectam bordas e contornos, camadas intermediárias capturam formas e texturas, e camadas mais profundas identificam objetos completos ou padrões de defeitos. Fazendo com que o modelo também identifique defeitos sutis.

As redes profundas são compostas por várias unidades chamadas neurônios artificiais, organizadas em camadas. Cada neurônio realiza operações matemáticas sobre os dados de entrada, aplicando pesos e funções de ativação não lineares, o que possibilita à rede modelar relações complexas entre variáveis. O treinamento de uma rede profunda envolve a otimização desses pesos, geralmente por meio do algoritmo de retropropagação, que ajusta os parâmetros para

minimizar o erro entre a saída prevista e a saída real.

Figura 1: Exemplificação das múltiplas camadas de um modelo de aprendizado profundo



Fonte: Scientific Figure on ResearchGate (s.d.)

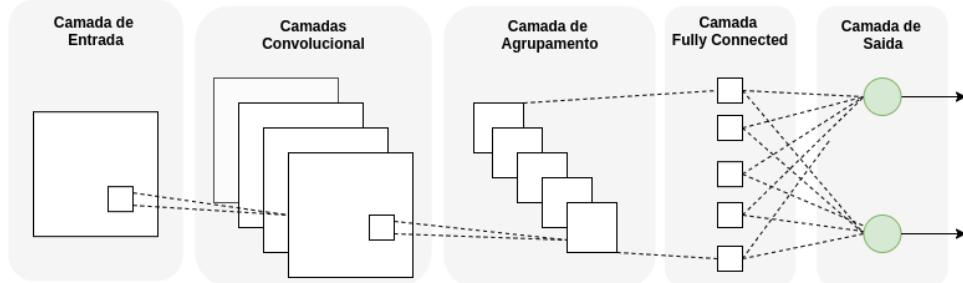
## 2.2 Rede neural convolucional

As Redes Neurais Convolucionais (CNNs) são uma arquitetura específica de aprendizado profundo projetada para lidar com dados estruturados espacialmente, onde o valor de um elemento está relacionado com seus vizinhos, por exemplo, imagens.

Uma CNN típica é composta por:

- **Uma camada de Entrada:** Recebem os dados em formato de matriz.
- **N Camadas Convolucionais:** Aplicam filtros que detectam padrões locais (Kernels), como bordas, curvas e texturas. Cada filtro responde a um tipo específico de característica, produzindo mapas de características que destacam essas informações. Nesses mapas, são adicionados bias e aplicadas funções de ativação (ex: ReLU), com o intuito de introduzir não linearidade ao sistema e permitir o aprendizado de padrões mais complexos.
- **N Camadas de Agrupamento:** Reduzem a dimensionalidade da matriz resultante (mapas de ativação), preservando informações importantes e tornando o modelo menos sensível a pequenas variações na posição dos objetos.
- **Uma Camada Totalmente Conectadas:** Recebem as características extraídas em forma de um vetor, processa elas e são produzem a saída final.

Figura 2: Exemplificação de uma rede neural convolucional



Fonte: Barbosa *et al.* (2021)

No caso da detecção de anomalias em trilhos ferroviários, as CNNs permitem que o modelo aprenda automaticamente padrões de falhas, como deformações, fissuras, ou desgaste, sem a necessidade de extração manual de características.

### 2.3 Transferência de aprendizado

Um dos principais desafios para o uso aprendizado profundo é a necessidade de grandes volumes de dados rotulados para treinar redes neurais com desempenho satisfatório. Esse problema é ainda mais evidente no meio acadêmico, onde bancos de dados extensos nem sempre estão disponíveis, tornando inviável o treinamento de modelos complexos a partir do zero. Para contornar essa limitação, emprega-se a técnicas de transferência de aprendizado.

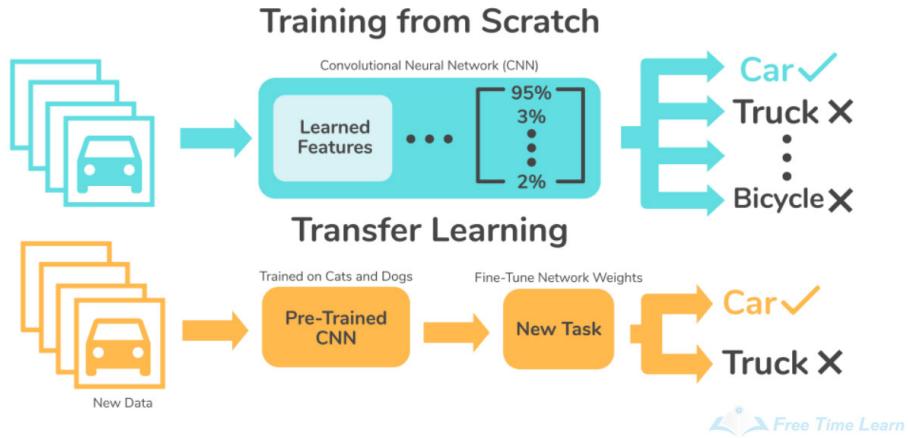
A transferência de aprendizado consiste em reutilizar o conhecimento adquirido por um modelo previamente treinado em um grande dataset genérico, como ImageNet ou COCO, aplicando-o a uma nova tarefa que dispõe de uma quantidade limitada de dados. Esses modelos pré-treinados já aprenderam a identificar padrões visuais fundamentais, como bordas, formas e texturas, que tendem a ser úteis em diversos domínios.

Na prática, o processo de transferência de aprendizado pode ser dividido em duas etapas principais:

1. **Congelamento do modelo base:** o modelo pré-treinado, frequentemente chamado de modelo professor, mantém seus pesos nas camadas iniciais e intermediárias, preservando a capacidade de extrair padrões gerais já aprendidos.
2. **Adaptação à nova tarefa:** as camadas finais do modelo são modificadas ou substituídas para se ajustarem ao novo problema. Essas camadas são então treinadas com o dataset

específico, permitindo que o modelo aprenda características relevantes para a tarefa alvo sem necessidade de treinamento integral. Em modelos híbridos, essa etapa pode incluir a combinação de arquiteturas, como utilizar a saída da camada totalmente conectada de uma CNN como entrada para um classificador externo, por exemplo uma SVM.

Figura 3: Exemplificação de transferencia de aprendizado



Fonte: FreeTime Learning (s.d.)

Na Figura 3 é demonstrado a diferença do processo de treinamento de uma CNN treinada do zero e treinamento de uma CNN pré-treinada, sendo necessário apenas modificar camada de saída da rede e realizar o aperfeiçoamento dos parâmetros.

Neste trabalho, será utilizada uma redes neural convolucional pré-treinada para aproveitar o conhecimento adquirido por grandes modelos treinados em bases extensas e minimizar os impactos causados pelo uso de um dataset reduzido.

A arquitetura escolhida foi a MobileNetV3, sendo esta uma melhoria do mobileNetV2 utilizada por Zheng *et al.* (2021) em seu trabalho de identificação de defeitos em pinos de fixação e amplamente recomendada para tratar de datasets pequenos, como é o caso.

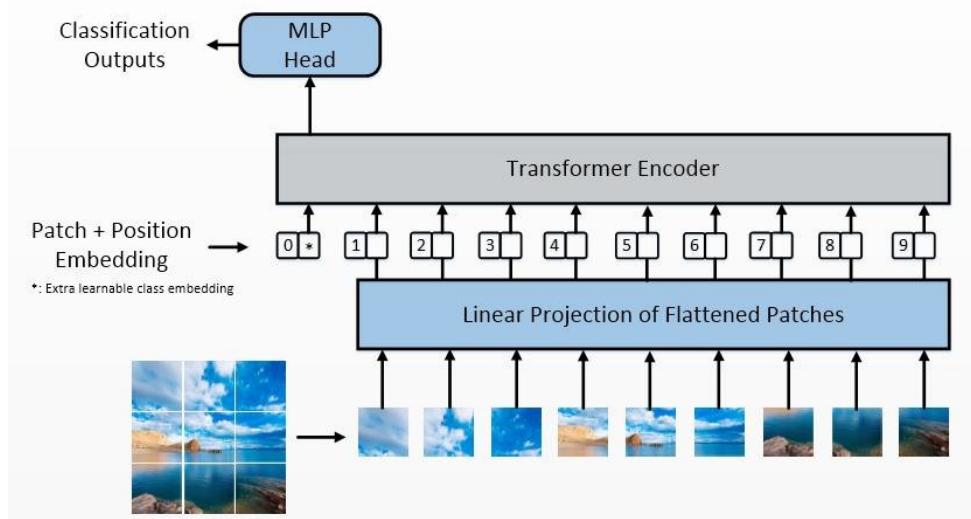
## 2.4 Transformadores visuais

Os Transformadores Visuais (ViTs) representam uma abordagem mais recente no processamento de imagens, inspirada nos Transformers utilizados originalmente em tarefas de linguagem natural. Enquanto as CNNs focam em padrões locais, os transformadores capturam relações globais na imagem por meio de mecanismos de atenção, permitindo que o modelo considere o contexto completo ao tomar decisões.

O funcionamento básico de um Vision Transformer envolve:

- Divisão da imagem em patches menores.
- Transformação de cada patch em vetores de embedding, que codificam informações visuais.
- Passagem dos embeddings por camadas de autoatenção, que ponderam a importância relativa de cada patch em relação aos outros, integrando contexto global.
- Classificação ou detecção baseada nos embeddings processados, permitindo identificar padrões complexos e sutis, como pequenas fissuras ou deformações.

Figura 4: Exemplificação do funcionamento de um transformador visual



Fonte: Liang, Wang e Ling (2021)

Transformadores visuais têm se mostrado altamente eficazes em tarefas onde a variabilidade de cenários é grande ou os datasets são limitados, pois conseguem aproveitar melhor a informação contextual e evitar confusões causadas por objetos próximos ou fundos complexos.

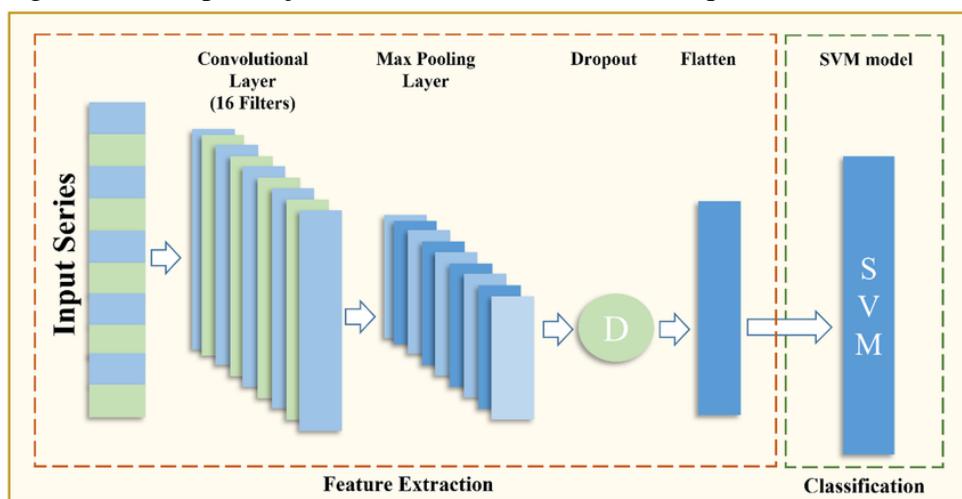
Como observado no trabalho de Shahin *et al.* (2024), modelos híbridos, combinando CNNs e ViTs, também têm sido utilizados para aproveitar a extração local de CNNs e a consciência global dos transformadores, alcançando desempenho superior em detecção de defeitos ferroviários. Isso é feito, treinando CNNs em grandes datasets de imagens e utilizando-as para treinar os ViTs.

## 2.5 SVM

Embora máquinas de vetores de suporte (SVM) não sejam técnicas de aprendizado profundo, será utilizado SVM de modelo híbrido. Para isso, a CNN pré-treinada será utilizada como extratora de características para a SVM. As últimas camadas convolucionais irão produzir vetores representativos do conteúdo visual da imagem, que serão utilizados como input pela SVM, evitando a necessidade de extração de características manuais.

Ela será utilizada como baseline para os outros métodos, visto que, métodos de aprendizado profundo devem produzir resultados melhores que métodos tradicionais como SVM.

Figura 5: Exemplificação do funcionamento de uma arquitetura SVM+CNN



Fonte: Wang e Zhu (2023)

## 3 Metodologia

### 3.1 Datasets

Um dos datasets utilizados neste trabalho será *Railway Track fault Detection Resized* ( $224 \times 224$ ), disponibilizado na plataforma Kaggle por Gerry *et al.* (2022). Trata-se de uma versão re-dimensionada do conjunto original *Railway Track Fault Detection* disponibilizado por Hossain *et al.* (2021), também na plataforma Kaggle.

As imagens originais possuem alta resolução, o que torna o pré-processamento custoso, por esse motivo o dataset foi recriado com todas as imagens reduzidas para  $244 \times 244 \times 3$ , facilitando seu uso, também foi incluído um arquivo *rails.csv*, que simplifica o carregamento e a organização das amostras (Gerry *et al.*, 2022).

Também será utilizado o dataset *Railway Track Fault Detection | Dataset2(Fastener)*, disponibilizado na plataforma Kaggle por Adnan (2021). Porém, ao contrário do conjunto anterior, esse conjunto contém apenas imagens de pinos de fixação e não foi redimensionado.

Ambos os datasets são compostos pelo mesmo número de imagens com e sem anomalias. O conjunto de Gerry *et al.* (2022) contém 384 imagens de trilhos divididas entre 300 imagens de treinamento, 22 imagens de teste e 62 imagens de validação. Esse conjunto contém três tipos de anomalias: ausência de pinos de fixação, deformações e rachaduras nos trilho e danos na fundação do trilho (parte de madeira transversal ao trilho). Enquanto que o conjunto de Adnan (2021) contém 1399 imagens de pinos de fixação divididas entre 980 imagens de treinamento, 140 imagens de teste 280 imagens de validação

Como mencionado anteriormente, o segundo dataset não foi previamente redimensionado. Portanto suas imagens serão ajustadas para o formato 224x244x3, a fim de padronizar as dimensões e minimizar a heterogeneidade entre os diferentes conjuntos de dados.

Figura 6: Conjunto de imagens do segundo dataset em grade 1x2

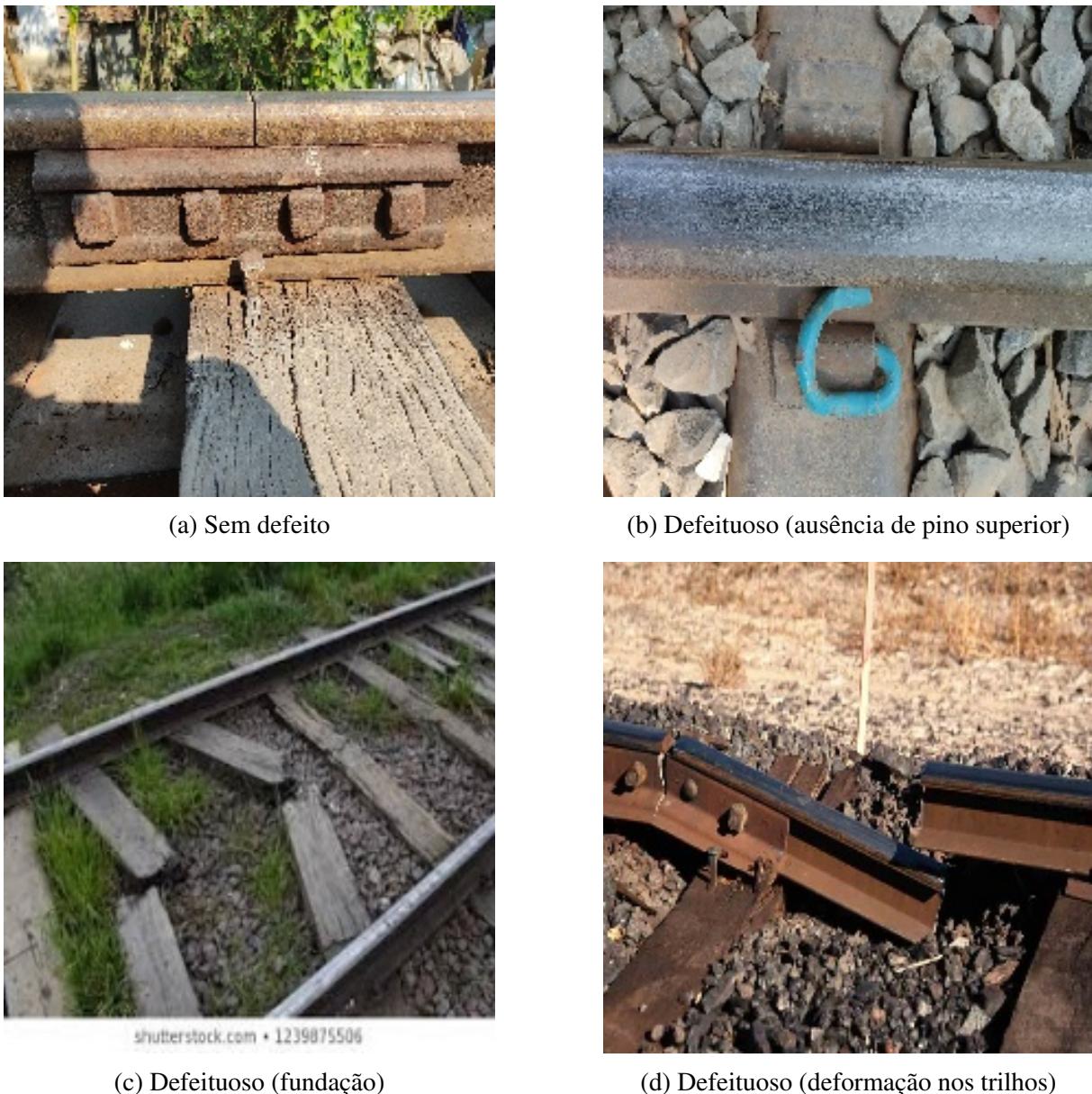


(a) Defeituoso



(b) Sem defeito

Figura 7: Conjunto de imagens do primeiro dataset em grade 2x2



Fonte: Gerry *et al.* (2022)

### 3.2 Modelos

Os seguintes métodos serão analisados:

- **ResNet** (emprega conexões residuais para facilitar o treinamento de redes profundas);
- **InceptionV3** (utiliza blocos com múltiplos filtros paralelos para capturar padrões em diferentes escalas);

- **MobileNetV3** (adota convoluções separáveis e blocos otimizados para dispositivos de baixo custo computacional);
- **CNN + SVM** (a CNN extrai características e a SVM classifica, funcionando como baseline simples e interpretável);
- **ViT – Vision Transformer** (divide a imagem em patches e usa autoatenção global para capturar dependências longas);
- **DeiT** (versão otimizada do ViT com treinamento mais eficiente e menor necessidade de dados);
- **Swin Transformer** (utiliza janelas de autoatenção deslizantes e hierárquicas para reduzir custo computacional);
- **ConViT** (combina convoluções com autoatenção, introduzindo “soft convolution” para transição suave entre ambos);
- **MobileViT** (integra blocos MobileNet com módulos tipo Transformer, oferecendo baixo custo e atenção global);
- **Rail-Former** (usa encoder–decoder com autoatenção e módulo Criss-Cross para preservar detalhes finos de defeitos ferroviários);
- **TrackNet** (aplica autoatenção multi-head e transferência de aprendizado para melhorar detecção em trilhos com dados limitados).

### 3.3 Resultados Esperados

Espera-se que os modelos desenvolvidos sejam capazes de identificar, com precisão alta, os diferentes tipos de defeitos presentes nas imagens do conjunto de dados. De forma geral, os resultados esperados incluem:

- Os modelos baseados em aprendizado profundo obtenham desempenho superior aos métodos clássicos.
- As arquiteturas que utilizam transferência de aprendizado apresentem resultados mais estáveis e eficientes mesmo com um conjunto de dados limitado.

- Os modelos devem demonstrar capacidade de generalização, consistente mesmo diante de variações de iluminação, ângulo ou ruído.
- A análise comparativa identificar quais arquiteturas são mais adequadas para detecção automática de defeitos ferroviários e justificar suas vantagens técnicas.

### 3.4 Avaliação dos Resultados

Os resultados dos testes serão obtidos através de uma matriz de confusão, que organiza as predições do modelo em quatro categorias fundamentais: verdadeiros positivos (TP) falsos positivos (FP), verdadeiros negativos (TN) e falsos negativos (FN). Essa estrutura permite avaliar não apenas o desempenho geral do classificador, mas também entender em quais situações ele tende a errar.

Figura 8: Imagem mostrando o exemplo de uma matriz de confusão

		Positivo (1)	Negativo (0)
Valores Previstos	Positivo (1)	TP (verdadeiro Positivo)	FP (Falso Positivo)
	Negativo (0)	FN (Falso Negativo)	TN (Verdadeiro Negativo)

Fonte: Lopes (2023)

A partir da matriz de confusão, quatro métricas amplamente utilizadas na avaliação de modelos de classificação serão calculadas:

- **Acurácia (Accuracy)**

Mede a proporção de classificações corretas em relação ao total de previsões realizadas, sendo utilizada para analisar o desempenho dos modelos de forma geral.

$$A = \frac{TP + TN}{TP + FP + FN + TN}$$

- **Precisão (Precision)**

Mede a proporção de positivos verdadeiros entre todas as previsões positivas, sendo utilizada para avaliar o grau de confiabilidade das detecções positivas.

$$P = \frac{TP}{TP + FP}$$

- **Revocação (Recall)**

Mede a capacidade do modelo de recuperar os positivos reais, avaliando o quanto o modelo deixa de detectar casos positivos.

$$R = \frac{TP}{TP + FN}$$

- **F1-score**

Média harmônica entre precisão e revocação. Resume ambas as métricas em um único valor equilibrado. Amplamente utilizada na literatura de machine learning.

$$F1 = 2 \cdot \frac{P \cdot R}{P + R}$$

Após a obtenção dos resultados, será realizada uma comparação com os estudos presentes na literatura, analisando como os diferentes modelos se comportaram no contexto deste trabalho em relação ao desempenho observado em suas aplicações originais. Essa análise permitirá identificar em que medida fatores como domínio dos dados, condições de captura das imagens e complexidade das anomalias influenciam o desempenho das arquiteturas, além de evidenciar quais modelos apresentam maior capacidade de generalização para o cenário de detecção de falhas em trilhos ferroviários.

## 4 Referências

ADNAN, Ashik. **Railway Track Fault Detection Dataset (Fastener Version)**. [S. l.: s. n.], 2021. Kaggle. Disponível em: <https://www.kaggle.com/datasets/ashikadnan/railway-track-fault-detection-dataset2fastener>. Acesso em: 20 nov. 2025.

BARBOSA, Guilherme *et al.* Segurança em Redes 5G: Oportunidades e Desafios em Detecção de Anomalias e Predição de Tráfego Baseadas em Aprendizado de Máquina. In: [s. l.: s. n.], out. 2021. p. 145–189. ISBN 9786587003658. DOI: 10.5753/sbc.7165.8.4.

FREETIME LEARNING. **Explain Transfer Learning in the Context of Deep Learning.** [S. l.: s. n.]. <https://www.freetimelearning.com/software-interview-questions-and-answers.php?Explain-transfer-learning-in-the-context-of-deep-learning.&id=4184>. Accessed: 2025-11-20.

GERRY *et al.* **Railway Track Fault Detection (Resized 224x224).** [S. l.: s. n.], 2022. Kaggle Dataset. Accessed: 20 Nov. 2025. Disponível em: <https://www.kaggle.com/datasets/gpiosenka/railway-track-fault-detection-resized-224-x-224>.

HOSSAIN, Shahriar *et al.* **Railway Track Fault Detection.** [S. l.]: Kaggle, 2021. Kaggle Dataset. Accessed: 20 Nov. 2025. DOI: 10.34740/KAGGLE/DSV/1884733. Disponível em: <https://www.kaggle.com/dsv/1884733>.

HUANG, Lang *et al.* Big-data-driven safety decision-making: A conceptual framework and its influencing factors. **Safety Science**, v. 109, p. 46–56, 2018. ISSN 0925-7535. DOI: <https://doi.org/10.1016/j.ssci.2018.05.012>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0925753518300973>.

LIANG, Jingxin; WANG, Dong; LING, Xufeng. Image Classification for Soybean and Weeds Based on ViT. **Journal of Physics: Conference Series**, v. 2002, p. 012068, ago. 2021. DOI: 10.1088/1742-6596/2002/1/012068.

LOPES, André. **Medidas de performance de modelos de classificação.** [S. l.: s. n.], 2023. <https://brains.dev/2023/medidas-de-performance-modelos-de-classificacao/>. Acessado em: 20 Nov 2025.

SARHANI, Malek; VOSS, Stefan. Prediction of rail transit delays with machine learning: How to exploit open data sources. **Multimodal Transportation**, v. 3, n. 2, p. 100120, 2024. ISSN 2772-5863. DOI: <https://doi.org/10.1016/j.multra.2024.100120>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2772586324000017>.

SCIENTIFIC FIGURE ON RESEARCHGATE. **Inteligência Artificial: Uma era de abundância ou o fim da espécie humana? – Visão geral das redes neurais no aprendizado profundo.** [S. l.: s. n.]. [https://www.researchgate.net/figure/Figura-4-Visao-geral-das-redes-neurais-no-aprendizado-profundo\\_fig4\\_319212925](https://www.researchgate.net/figure/Figura-4-Visao-geral-das-redes-neurais-no-aprendizado-profundo_fig4_319212925). Acessado em: 20 Nov 2025.

SHAHIN, Mohammad *et al.* Improving the Concrete Crack Detection Process via a Hybrid Visual Transformer Algorithm. **Sensors**, v. 24, n. 10, 2024. ISSN 1424-8220. DOI: 10.3390/s24103247. Disponível em: <https://www.mdpi.com/1424-8220/24/10/3247>.

TIONG, Kah Yong; MA, Zhenliang; PALMQVIST, Carl-William. A review of data-driven approaches to predict train delays. **Transportation Research Part C: Emerging Technologies**, v. 148, p. 104027, 2023. ISSN 0968-090X. DOI: <https://doi.org/10.1016/j.trc.2023.104027>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0968090X23000165>.

WANG, Jujie; ZHU, Shuzhou. A Novel Stock Index Direction Prediction Based on Dual Classifier Coupling and Investor Sentiment Analysis. **Cognitive Computation**, v. 15, p. 1–19, abr. 2023. DOI: 10.1007/s12559-023-10137-4.

ZHENG, Danyang *et al.* A Defect Detection Method for Rail Surface and Fasteners Based on Deep Convolutional Neural Network. **Computational Intelligence and Neuroscience**, v. 2021, n. 1, p. 2565500, 2021. DOI: <https://doi.org/10.1155/2021/2565500>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1155/2021/2565500>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1155/2021/2565500>.