

# Aprendizado profundo na detecção de anomalias em dados ferroviários

Miguel de Campos Rodrigues Moret

Orientador: Prof. Dr. Cassio Machiaveli Oishi

Universidade Estadual Paulista "Júlio de Mesquita Filho"  
Faculdade de Ciência e Tecnologia

11 de novembro de 2025

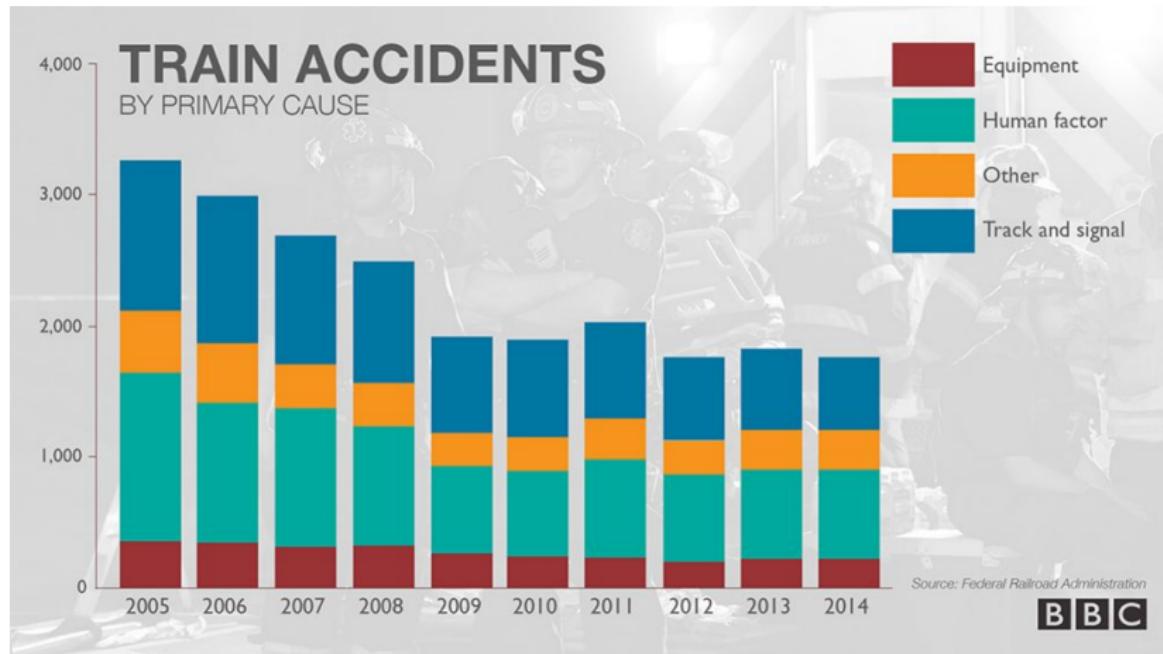
- 1 Formulação do problema
- 2 Objetivos do projeto
- 3 Justificativa do projeto
- 4 Metodologia e plano de trabalho
- 5 Equipamento e material
- 6 Cronograma de execução

# Formulação do problema

27.9% dos acidentes ferroviárias são causados por defeitos nos trilhos  
(Short Line Safety Institute, 2020).

# Formulação do problema

Acidentes ferroviários por causa (EUA, 2005-2014)



Fonte: Federal Railroad Administration (BBC), 2015

# Formulação do problema

As malhas ferroviárias são inspecionadas manualmente, o que as torna demoradas e suscetíveis a erros humanos.

Manutenção manual



Fonte: Massa, 2020

# Formulação do problema

Avanços tecnológicos recentes tem impulsionado a busca por soluções automatizadas para inspeção e manutenção ferroviária.

Manutenção autônoma



Fonte: Redação CIMM, 2025

Investigar o uso de técnicas de aprendizado profundo para a detecção de anomalias em dados ferroviários:

- Realizar o pré-processamento e a análise exploratória dos datasets ferroviários escolhidos.
- Implementar e testar diferentes modelos de aprendizado profundo aplicados à detecção de anomalias.
- Avaliar o desempenho dos modelos por meio de métricas adequadas (Acurácia, Precisão, Revocação, F1-score).
- Comparar os resultados com estudos semelhantes da literatura, destacando vantagens e limitações.

Por que utilizar aprendizado profundo?

- Modelos clássicos como SVM e Random Forest exigem extração manual de características.
- O aprendizado profundo aprende representações diretamente dos dados, capturando padrões sutis e não lineares presentes nos trilhos ferroviários.

Vantagens para o contexto ferroviário:

- Permite integrar diferentes tipos de dados (imagens, vibrações, sons e outras medições estruturais).
- Apresenta resultados superiores em tarefas de detecção de anomalias industriais, aumentando a precisão e reduzindo inspeções manuais.

# Metodologia e plano de trabalho

Será feito um aprofundamento sobre:

- Técnicas de visão computacional aplicadas à detecção de defeitos em trilhos ferroviários.
- Modelos de aprendizado profundo, com foco em:
  - Redes neurais;
  - Vit;
  - DeiT.
- Estratégias de aprendizado por transférencia em conjunto de dados reduzidos.

# Metodologia e plano de trabalho

- Coleta e organização dos datasets ferroviários.
- Pré-processamento das imagens (normalização, balanceamento, aumento de dados).
- Implementação dos modelos de aprendizado profundo (CNN, ViT, DeiT).
- Treinamento e validação dos modelos.
- Avaliação de desempenho com métricas adequadas (F1, AUC, etc.).
- Comparação com resultados da literatura.

# Metodologia e plano de trabalho

Exemplo de trilho sem pino



Fonte: <https://www.kaggle.com/datasets/gpiosenka/railway-track-fault-detection-resized-224-x-224/data?select=test>

# Metodologia e plano de trabalho

Exemplo de trilho quebrado



Fonte: <https://www.kaggle.com/datasets/gpiosenka/railway-track-fault-detection-resized-224-x-224/data?select=test>

# Metodologia e plano de trabalho

Exemplo de trilho sem suporte



Fonte: <https://www.kaggle.com/datasets/gpiosenka/railway-track-fault-detection-resized-224-x-224/data?select=test>

- Artigos, livros, monografias para a aquisição da fundamentação teórica para a elaboração do projeto.
- Utilização da plataforma *Kaggle* para obtenção dos *datasets* utilizados.
- Linguagem de programação utilizada será *Python* e as bibliotecas Numpy, Pytorch, entre outras.
- Processamento dos treinos utilizará os computadores do lab 6.
- Elaboração do relatório, revisão bibliográfica e artigo será utilizado *LATEX*.

# Cronograma de execução

As atividades a serem executadas estão listadas a seguir:

- ① Anteprojeto;
- ② Revisão bibliográfica;
- ③ Análise e preparação dos *datasets* a serem utilizados;
- ④ Análise e implementação dos métodos de aprendizado profundo a serem utilizados;
- ⑤ Análise acerca dos resultados obtidos;
- ⑥ Elaboração do trabalho;

# Cronograma de execução

O cronograma será dividido em bimestres:

Atividade	1º	2º	3º	4º	5º	6º	7º	8º	9º
1	•	•	•						
2	•	•	•						
3		•	•	•					
4			•	•	•	•			
5						•	•	•	
6							•	•	•

# Referências

- SARHANI, M.; VOSS, S. Prediction of rail transit delays with machine learning: How to exploit open data sources. *Multimodal Transportation*, v. 3, n. 2, p. 100120, 1 jun. 2024.
- TIONG, K. Y.; MA, Z.; PALMQVIST, C.-W. A review of data-driven approaches to predict train delays. *Transportation Research Part C: Emerging Technologies*, v. 148, p. 104027, 1 mar. 2023.
- ZHANG, J.; ZHANG, J. Artificial Intelligence Applied on Traffic Planning and Management for Rail Transport: A Review and Perspective. *Discrete Dynamics in Nature and Society*, v. 2023, p. e1832501, 27 jul. 2023.
- SHI, L. et al. Data-Driven Bayesian Network Analysis of Railway Accident Risk. *IEEE Access*, v. 12, p. 38631–38645, 1 jan. 2024.
- MOHAMMADI, S. et al. Rail Defect Classification with Deep Learning Method. *Green Energy and Intelligent Transportation*, p. 100332–100332, 1 jun. 2025.