

UNIVERSIDADE DE SÃO PAULO  
ESCOLA POLITÉCNICA  
DEPARTAMENTO DE ENGENHARIA MECATRÔNICA E DE SISTEMAS MECÂNICOS  
ENGENHARIA MECÂNICA DE PROJETO E FABRICAÇÃO

Modesto Valci Moreira Lopes

**Utilização de redes neurais para avaliação do contato  
roda-trilho baseada na análise de perfis**

São Paulo  
2023

MODESTO VALCI MOREIRA LOPES

**Utilização de redes neurais para avaliação do contato  
roda-trilho baseada na análise de perfis**

— Versão Corrigida —

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo como parte dos requisitos para obtenção do título de Mestre em Ciências.

Área de Concentração: Engenharia Mecânica de Projeto e Fabricação

Orientador: Prof. Dra. Larissa Driemeier

São Paulo  
2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 29 de Novembro de 2023

Assinatura do autor: Modesto Valci Moreira Lopes

Assinatura do orientador: Priscila Drummond

#### Catálogo-na-publicação

Lopes, Modesto Valci Moreira

Utilização de redes neurais para avaliação do contato roda-trilho baseada na análise de perfis / M. V. M. Lopes -- versão corr. -- São Paulo, 2023.  
80 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos.

1.Inteligência Artificial 2.Mecânica do Contato 3.Roda-Trilho 4.Desgaste 5.Redes Convolucionais I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos II.t.

Dedico este trabalho a:

Meus pais, João Modesto e Ediran, por sempre apoarem meus sonhos.

Tio Paulo (*in memoriam*), que hoje me acompanha lá de cima.

# Agradecimentos

Primeiramente quero agradecer a Deus por ter me dado saúde, força e por provar sua existência para mim nos momentos em que minha fé era posta a prova.

Agradeço ao meu pai, João Elias Moreira, por sempre ter me aconselhado e me apoiado nessa minha nova trajetória.

Agradeço a minha mãe, Maria Eidiran Lopes, que representa tudo de maravilhoso nesse mundo e que sempre apoiou inteiramente meus sonhos, especialmente este, de me tornar mestre.

Agradeço ao meu irmão, João Elias Moreira Filho, que sempre me apoiou nas minhas decisões e que sempre esteve do meu lado me ajudando nos mais diversos problemas da vida.

Agradeço também a todos os meus familiares que sempre me apoiaram e torceram pelo meu sucesso, e em especial, meu tio Paulo (*in memoriam*) e tia Lara por terem me recebido em São Paulo e por me darem sempre o suporte que precisei.

Agradeço à minha orientadora Dra. Larissa Driemeier por ter me acompanhado durante todo o mestrado e por sempre ter tido a paciência, empatia e atenção que um excelente orientador deve ter. Seu bom senso de humor sempre me fez me sentir confortável e seus conselhos foram essenciais para o meu desenvolvimento profissional e pessoal.

Agradeço também aos professores Dr. Roberto Martins Souza, Dra. Izabel Fernanda Machado, Dr. Francisco José Profito e ao Dr. Newton Kyioshi Fukumasu e Eng. Vinícius Alves de Lima, pelas contribuições a este trabalho e por toda a ajuda fornecida.

Agradeço também aos meus amigos/colegas de trabalho André, Arnaldo, Marcos Dantas, João Carlos, Davi, Leandro, Chico, Alejandro, Santiago, Fábio, Vanessa, Marcionila, Danny, Antonio, Samira, Samara, Jaimer, Daimer e Javier, pelos bons momentos vividos no laboratório.

De forma especial, quero agradecer a Arnaldo por sempre ter estado disponível para me ajudar com o Abaqus; a André Freire por ter me ensinado diversas coisas sobre Abaqus e ao Felipe, por ter me ajudado a revisar/escrever essa dissertação.

Agradeço também aos meus amigos Ana, Felipe, Valério, Melissa (Mel), Bárbara, Thárcylla, Leo, Kamila, Gabi, Isaac, Tobias, Edison, Dedéu, Dênis, Cori, João, Laura, Vitão, Matheus, Olivia, Nathaly, Natália, Viny, Michele, Virginia, Wasim, Bruno, Sven, Kevin, Steven, Nele, Mirco, Deren, Jannis, Melzinha, Young Jin, Solena, Adele, Adelaide, Glácio, João Antônio, Isaac, Grazi, César, Pedro(s), Artur, Leandro, Juan Diego, Darlynn, Duda, Suelen, Alissa, Naju, Malu, Jake, Emily, Amelie, Anne, João Bosco, Hildegardis, Jonathan, Enzo, Priscila, Cristina, Vivian, Bia, Guilherme, Alpino, Laiza, Amanda e Ricardo por todos os bons momentos juntos e por terem estado ao meu lado nos piores e melhores momentos, fazendo São Paulo ser para mim não só uma passagem, mas também um lar.

Temos que continuar aprendendo.  
Temos que estar abertos.  
E temos que estar prontos,  
para espalhar nosso conhecimento  
a fim de chegar a uma compreensão  
mais elevada da realidade.

— *Thich Nhat Hanh*

# Sumário

**Sumário** • *v*

**Resumo** • *vii*

**Abstract** • *viii*

**Lista de Figuras** • *ix*

**Lista de Tabelas** • *xi*

**1 Introdução** • *1*

**2 Objetivos** • *4*

**3 Fundamentação Teórica e Revisão da Literatura** • *5*

3.1 Histórico e Panorama das Ferrovias • *5*

3.2 Composição do Sistema Ferroviário • *7*

3.3 Mecânica do Contato Roda-Trilho • *11*

3.3.1 Modelagem do Contato Geral • *13*

3.3.2 Análises do Contato Roda-Trilho • *15*

3.3.3 Métrica para Avaliação do Tipo de Contato • *20*

3.4 Aspectos Teóricos da Inteligência Artificial • *21*

3.4.1 Redes Neurais Convolucionais • *27*

3.5 Revisão da Literatura acerca da utilização de IA na área ferroviária • *29*

**4 Metodologia** • *35*

4.1 Banco de Dados • *35*

4.1.1 Modelo Virtual • *36*

4.1.2 Automatização • *42*

4.2 Modelo de Inteligência Artificial • *46*

4.2.1 Tratamento das Imagens • *47*

4.2.2 Arquitetura e seus Hiperparâmetros • *48*

4.2.3 Hiperparâmetros de Treinamento • *50*

4.2.4 Otimização dos Hiperparâmetros • *51*

4.3 Treinamento da Rede Neural • *53*

**5 Resultados e Discussão** • *56*

**6 Conclusão • 63**

**Referências • 65**



## Resumo

A análise do contato entre roda e trilho tem sido amplamente incentivada nos últimos anos, devido à sua influência significativa na segurança e nos custos de transporte. Tradicionalmente, um dos tópicos investigados é a formação de defeitos nas superfícies de rolamento roda-trilho, corroborando na necessidade de inspeções e manutenção. Com o objetivo de minimizar tempo e custos e melhorar essas análises, técnicas de aprendizado de máquina têm sido cada vez mais utilizadas para esse fim. A avaliação da conformidade do contato entre roda e trilho auxilia na determinação do tipo de dano predominante no sistema, podendo ele ser do tipo desgaste ou fadiga de contato por rolamento (RCF, do inglês *“rolling contact fatigue”*). RCF será predominante em contatos Não-Conformes, enquanto o desgaste prevalecerá em contatos Conformes e Quase Conforme. Para a determinação do tipo de contato, é possível utilizar uma métrica chamada separação máxima (ou parâmetro *“s”*), que compreende a distância entre a superfície do trilho e da roda ao longo do centroide do trilho após o contato. O objetivo do presente trabalho é propor um modelo de Inteligência Artificial, por meio de Redes Neurais Convolucionais (CNNs), para classificar a conformidade do contato roda-trilho. Os dados de entrada são imagens dos perfis do trilho e da roda, obtidos diretamente de medições de campo, e a saída é uma classificação de acordo com o nível do contato. A base de dados foi gerada a partir de uma análise validada de elementos finitos quasi-estática em 2D (MEF), realizada no *software* Abaqus. A análise também exigiu o desenvolvimento de um *script* em *Python* para automatizar o processo de simulação de várias combinações de diferentes perfis de roda e trilho. Finalmente, após treinamento e validação, a CNN apresentou resultados satisfatórios, com acurácia de 93%, permitindo a avaliação direta do contato entre os perfis, sem análises complexas e custosas por meio de *software* de elementos finitos.

**Palavras-chave:** Inteligência Artificial. Mecânica do Contato. Roda-Trilho. Desgaste. Redes Convolucionais

## Abstract

The analysis of wheel-rail contact has been widely encouraged in recent years, due to its significant influence on safety and transportation costs. Traditionally, one of those investigated is the formation of defects in the wheel-rail bearing surfaces, corroborating the need for inspections and maintenance. In order to minimize time and costs and improve analyses, machine learning techniques have been increasingly used for this purpose. The evaluation of the conformity of the contact between wheel and rail helps to take advantage of the predominant type of damage in the system, allowing it to be wear or rolling contact fatigue (RCF). RCF will predominate on Non-Conformal contacts, while wear will prevail on Conformal and Closely Conformal contacts. To determine the type of contact, it is possible to use a metric called maximum separation (or "s" parameter), which comprises the distance between the surface of the rail and the wheel along the centroid of the rail after contact. The objective of this work is to propose an Artificial Intelligence model, through Convolutional Neural Networks (CNNs), to classify the conformity of the wheel-rail contact. The inputs data are rail and wheel profiles images, obtained directly from field controls, and the output is a classification according to the level of contact. The database was generated from a validated 2D quasi-static finite element (FE) analysis performed in the Abaqus software. The analysis also tracks the development of a Python script to automate the process of simulating various combinations of different wheel and rail profiles. Finally, after training and validation, CNN presented advanced results, with an accuracy of 93%, allowing the direct evaluation of the contact between the profiles, without complex and costly analyzes using finite element software.

**Keywords:** Artificial Intelligence. Contact Mechanics. Wheel-Rail. Wear. Convolutional Networks

# Lista de Figuras

1.1	Panorama mundial das malhas ferroviárias, onde as linhas férreas estão representadas pelas linhas pretas. . . . .	1
3.1	“Locomotion”, a primeira locomotiva a transportar passageiros através da linha férrea. . . . .	5
3.2	Densidade das malhas ferroviárias no mundo. . . . .	6
3.3	Infraestrutura e superestrutura de um sistema ferroviário. . . . .	7
3.4	Representação da bitola e do ângulo <i>cant</i> em uma linha férrea. . . . .	8
3.5	Perfil de trilho tipo 136RE padrão AREMA. . . . .	9
3.6	Truque de três peças. . . . .	10
3.7	Vista isométrica (a), lateral (b) e em corte (c) de uma roda AREMA. . . . .	11
3.8	Contato do tipo Conforme (a) e do tipo Não-Conforme (b). . . . .	12
3.9	Medições dos perfis de um trilho e de uma roda através do miniprof. . . . .	13
3.10	Perfis de roda (a) e de trilho (b) . . . . .	13
3.11	Contato geral de Hertz entre dois corpos: a) contato entre os sólidos, b) área de contato elíptica. . . . .	14
3.12	Carregamento normal e tangencial em um contato entre corpos. . . . .	15
3.13	Configuração de um contato centralizado entre roda e trilho e sua respectiva área de contato. . . . .	16
3.14	Diferentes regiões de contato roda-trilho. . . . .	16
3.15	Contato em dois pontos a) e em um único ponto b) entre roda e trilho. . . . .	17
3.16	Exemplo de contato na Região C entre roda e trilho. . . . .	17
3.17	Remoção de material resultante de uma situação de desgaste por rolamento/deslizamento. . . . .	18
3.18	Dano por fadiga na superfície de uma roda. . . . .	19
3.19	Abordagem sistêmica da interface de gerenciamento e pesquisa para o tribossistema roda-trilho. . . . .	19
3.20	Representação das métricas no contato roda-trilho. . . . .	20
3.21	Linha do tempo resumida da inteligência artificial. . . . .	22
3.22	Algoritmos para aprendizagem supervisionada e não-supervisionada. . . . .	22
3.23	Ilustração de uma rede neural. . . . .	23
3.24	Perceptron e seus componentes. . . . .	24
3.25	Gráficos da função ReLU e Leaky ReLU respectivamente. . . . .	26
3.26	Extração de características de imagem no processo de convolução. . . . .	28
3.27	Redução da dimensão do mapa de características via processo de <i>pooling</i> . . . . .	28
3.28	Rede neural convolucional LeNet-4. . . . .	29
3.29	Levantamento de artigos que trazem aplicações de IA em ferrovias. . . . .	30

3.30	Modelo de IA para a identificação da falha por aprendizado semi-supervisionado utilizando autotreinamento. . . . .	31
3.31	Esquema do teste disco-disco (twin disc rig). . . . .	31
3.32	Resumo dos principais aspectos do estado da arte. . . . .	34
4.1	Metodologia geral para o desenvolvimento do modelo de IA. . . . .	36
4.2	Metodologia macro para a construção do banco de dados. . . . .	36
4.3	Perfis de rodas e trilhos fornecidos pela Vale S.A. . . . .	37
4.4	Exemplo da divisão dos perfis de trilho em famílias de acordo com geometria e altura do boleto. . . . .	37
4.5	Modelo 2D do rodeiro. . . . .	38
4.6	Modelo 3D do rodeiro: a) Vista isométrica e b) Vista frontal. . . . .	38
4.7	Simplificação da geometria do trilho para a seção marcada em vermelho. . . . .	39
4.8	Simplificação da geometria da roda para a seção marcada em vermelho. . . . .	39
4.9	Distribuição da pressão de contato em simulação do modelo 3D. . . . .	41
4.10	Resultados obtidos através da simulação do contato entre a roda e o trilho a), onde: b) é o gradiente de tensão de von Mises, c) a deformação plástica equivalente e d) a pressão de contato na região destacada pelo quadro vermelho em a). . . . .	42
4.11	Sequência de comandos responsável por gerar o modelo na automatização. . . . .	43
4.12	Sequência de comandos no <i>script</i> , responsável por obter os <i>outputs</i> das simulações. . . . .	45
4.13	Sequência de comandos no <i>script</i> , responsável por extrair os valores de <i>s</i> das simulações. . . . .	45
4.14	Exemplo de dados de imagem dos perfis da roda e do trilho. . . . .	46
4.15	Exemplo de dados de imagens originais a) e com suavização Canny b). . . . .	48
4.16	Arquitetura da Rede Neural Convolutiva. . . . .	48
4.17	Inativação de perceptrons através da técnica Dropout. . . . .	51
4.18	Etapa de pré-processamento das imagens para inserção no modelo de IA, onde: a) refere-se a um exemplo de imagens que foram utilizadas no treinamento do primeiro modelo e b) para o treinamento do segundo modelo. . . . .	54
4.19	Treinamento dos modelos de IA. . . . .	55
5.1	Curva da função perda do treinamento da primeira rede neural. . . . .	57
5.2	Matriz de Confusão para o conjunto esquerdo. . . . .	58
5.3	Matriz de Confusão para o conjunto direito. . . . .	58
5.4	Curva da função perda do treinamento da segunda rede neural. . . . .	59
5.5	Matriz de Confusão para o conjunto esquerdo. . . . .	59
5.6	Matriz de Confusão para o conjunto direito. . . . .	60
5.7	Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado direito. . . . .	61
5.8	Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado direito. . . . .	61
5.9	Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado esquerdo. . . . .	61
5.10	Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado esquerdo. . . . .	62

# Lista de Tabelas

3.1	Definição de Conforme, Não-Conforme e Quase Conforme. . . . .	20
4.1	Propriedades mecânicas da roda AAR Class C e do trilho AREMA TR68. . . .	37
4.2	Tabela com hiperparâmetros das camadas de convolução . . . . .	50
4.3	Tabela com hiperparâmetros das camadas densas e das camadas de saída . .	50
4.4	Hiperparâmetros de treinamento. . . . .	50
4.5	Hiperparâmetros otimizados pelo método RSF. . . . .	52
4.6	Distribuição do tipo de contato no banco de dados para 1700 casos. . . . .	53
4.7	As dez melhores configurações de hiperparâmetros alcançadas via RSF. . . .	53
4.8	Distribuição do tipo de contato no banco de dados para 10.750 casos. . . . .	53
5.1	Amostra do banco de dados elaborado através da automatização da simulação.	56

## Introdução

Sendo consideradas como um dos principais meios de transporte público da atualidade, as linhas férreas, desde o seu surgimento, são sinônimo de desenvolvimento cultural, econômico e político (ZHUANG et al., 2018). O transporte ferroviário é o precursor na movimentação de cargas no mundo, exercendo um papel fundamental na logística mundial. Além disso, a ferrovia tem sido um dos principais meios de transporte em vários países (Figura 1.1), principalmente, nos que se enquadram nas maiores economias do mundo (ANTF, 2023).

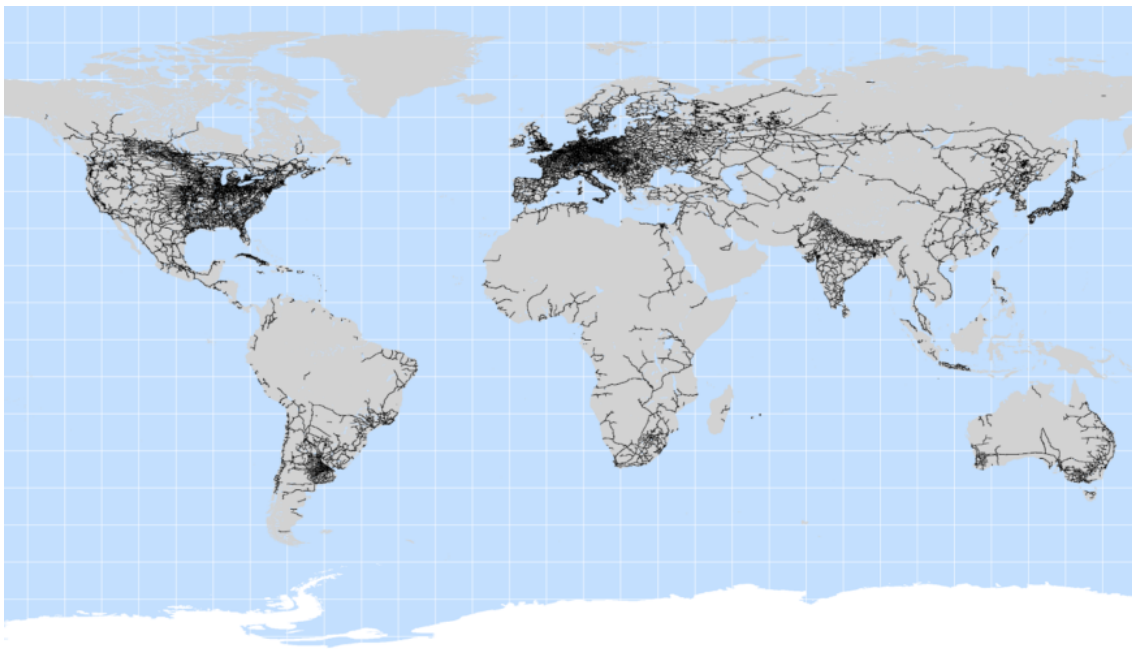


Figura 1.1: Panorama mundial das malhas ferroviárias, onde as linhas férreas estão representadas pelas linhas pretas. Extraído de MASSA Indústria (2023)

Nos últimos anos, a indústria ferroviária vem experimentando uma crescente demanda em suas operações de serviços devido o avanço do comércio global, o que corroborou

para o aumento de alguns parâmetros no sistema ferroviário, tais como: velocidade, peso, frequência de carga, entre outros (LASISI; ATTOH-OKINE, 2018).

Segundo Jamshidi et al. (2018), o aumento do tráfego de trens e das suas respectivas cargas causaram uma mudança na vida útil da infraestrutura férrea, sendo assim, o monitoramento e a manutenção das estruturas estão entre as principais preocupações dos gestores e afins, para manter e melhorar o desempenho das operações ferroviárias. Ainda de acordo com os autores, para manter o funcionamento operacional das linhas férreas em um nível eficaz, é necessário que se tenha um sistema capaz de acompanhar a integridade e as respectivas evoluções dos possíveis defeitos que podem surgir nos trilhos.

De acordo com Cárdenas-Gallo et al. (2017), que traz dados da *US Federal Railroad Administration Office of Safety Analysis*, os defeitos que ocorrem nos trilhos são uma das principais causas de acidentes ferroviários nos Estados Unidos, onde cerca de 577 acidentes (quase 33% do total) ocorreram no ano de 2012 devido a esses defeitos, resultando em um custo de aproximadamente US\$ 102,9 milhões de dólares. Lasisi e Attoh-Okine (2018) ressaltam que os defeitos em trilhos (principalmente os que desconfiguram as geometrias dos mesmos) constituem uma porção significativa no número de acidentes ferroviários que acontecem ao redor do mundo, onde, só nos Estados Unidos, foram responsáveis por 40% dos acidentes que ocorreram na década passada.

Dentre os principais tipos de defeitos que afetam o sistema roda-trilho presentes em literatura, pode-se citar: *squats*, corrosão, *head checks*, soldas danificadas, corrugações, defeitos na geometria, entre outros. Em suma, a maioria dos defeitos presentes nas linhas férreas são ocasionados pelo desgaste e também pela fadiga de contato entre os dois elementos (roda e trilho).

Magel e Kalousek (2002) traz em uma discussão acerca da mecânica do contato entre as rodas e os trilhos, como o tipo de contato influencia e direciona a formação do tipo de defeito no sistema. Segundo o autor, contatos do tipo conforme e quase conforme viabilizam o desgaste do sistema, ao passo que o contato do tipo não-conforme promove a fadiga de contato, tendo em vista a presença de maiores tensões. O mesmo autor propõe ainda uma métrica que é capaz de rotular o tipo de contato em conjunto roda-trilho. Essa métrica será abordada posteriormente tendo em vista sua relevância no alcance do objetivo deste trabalho.

De acordo com Shebani e Iwnicki (2018) o desgaste das rodas e trilhos é um problema significativo nos sistemas ferroviários. A previsão precisa desse desgaste pode melhorar a economia, o conforto de condução, a prevenção de descarrilhamento e o planejamento de intervenções de manutenção. A má previsão pode resultar em falha e consequente atraso e aumento de custos se não for controlada de forma eficaz. No entanto, a previsão

do desgaste de rodas e trilhos ainda é um grande desafio para engenheiros e operadores ferroviários.

Tendo isso em vista, houve a necessidade de se desenvolverem metodologias para o monitoramento, avaliação de defeitos e predição de falhas para o desenvolvimento de planejamentos de manutenções cada vez mais assertivos (ANDRADE; TEIXEIRA, 2012). Dentre as diversas abordagens aplicadas para tal existentes em literatura, há o uso de Inteligência Artificial (IA).

Segundo Haenlein e Kaplan (2019) a inteligência artificial é uma técnica da ciência que estuda e desenvolve teorias, métodos, técnicas e sistemas de aplicações para a simulação e extensão da inteligência humana. Atualmente a IA se tornou um curso interdisciplinar que abrange vários campos, seja ele da lógica, linguística, ciência da computação, ciência cognitiva, entre outros.

As redes neurais artificiais (RNA) são atualmente utilizadas para resolver uma vasta gama de problemas complexos de engenharia (SHEBANI; IWNICKI, 2018). Uma RNA tem a capacidade de aprender correlações que seja difícil de descrever com modelos físicos ou outras abordagens matemáticas (KHUDHAIR; TALIB, 2010). Embora a previsão perfeita raramente seja possível, as redes neurais podem ser utilizadas para fazer boas previsões em diferentes tipos de casos.



---

## Objetivos

O objetivo principal desse trabalho é a avaliação do tipo de contato roda-trilho a partir da análise de seus perfis geométricos utilizando Inteligência Artificial.

Para isso, essa pesquisa busca inicialmente entender os fenômenos físicos que existem no contato roda-trilho, bem como compreender a influência dos perfis geométricos das rodas e dos trilhos nas respostas tribológicas do contato de acordo com a literatura. Sendo assim, o trabalho contempla como objetivo específico o estudo e seleção de um critério que possibilite classificar o contato roda-trilho através da avaliação dos seus perfis.

Outro objetivo específico presente é o levantamento dos principais trabalhos que já foram desenvolvidos no que concerne à aplicação de IA na área ferroviária, de modo a mapear as melhores estratégias e metodologias que podem ser utilizadas para este fim.

Por fim, outros objetivos específicos contemplados neste trabalho são:

- Elaboração de banco de dados que relacione os perfis geométricos das rodas e dos trilhos em contato com a respectiva classificação do tipo de contato entre as geometrias;
- Desenvolver modelo de IA que seja capaz de classificar o tipo de contato existente entre roda e trilho através da análise das imagens dos perfis geométricos.

---

# Fundamentação Teórica e Revisão da Literatura

## 3.1 Histórico e Panorama das Ferrovias

A Revolução Industrial, que se deu na metade do século XVIII na Europa, desencadeou meios de produção em massa devido a evolução das máquinas e equipamentos dos processos de fabricação. Com o aumento no volume de produção, houve a necessidade de transportá-las com rapidez para os mercados consumidores, forçando os empresários ingleses a adotarem a ideia de George Stephenson (1781-1848) que apresentou sua primeira locomotiva em 1814. Entretanto foi somente em 1825 que o mundo foi apresentado à primeira locomotiva a vapor capaz de transportar pessoas e cargas, a “Locomotion” (Figura 3.1), cuja máquina tracionou um conjunto ferroviário em um percurso de 15 quilômetros a uma velocidade de 20 km/h pela estrada de ferro Stockton & Darlington. Esse feito gerou a primeira linha férrea do mundo e fazendo seu inventor fundar a primeira fábrica de locomotivas, se tornando o precursor do sistema ferroviário no mundo (DNIT, 2023).

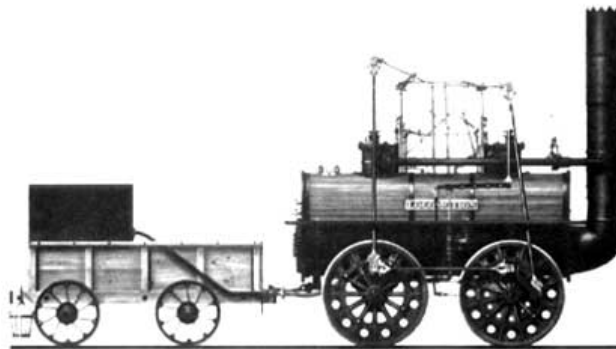


Figura 3.1: “Locomotion”, a primeira locomotiva a transportar passageiros através da linha férrea. Extraído de ANTF (2023)

No Brasil, as linhas férreas se desenvolveram logo depois e sempre estiveram ligadas intimamente às políticas governamentais, onde, em 1852, o pioneiro empreendedor brasileiro Irineu Evangelista de Souza (1813-1889), também conhecido como Barão de Mauá, recebeu a concessão do Governo Imperial para a construção de uma linha férrea no Rio de Janeiro. O projeto permitiu a integração do transporte aquaviário e ferroviário, iniciando a primeira operação intermodal do Brasil (ANTF, 2023).

Embora historicamente a Inglaterra receba o título de “berço do sistema ferroviário”, muitos países europeus merecem destaque quanto a suas extensas malhas ferroviárias. Isso é explicado pelo fato dessas nações serem pequenas em extensão territorial comparadas a países como China, Rússia, Estados Unidos, Canadá, Índia e Brasil. Atualmente, a maior densidade ferroviária do mundo fica nos Estados Unidos (Figura 3.2), com 23 km de ferrovia a cada 1000 km<sup>2</sup>, seguido da Índia com cerca de 20,8 km/1000 km<sup>2</sup>. O Brasil aparece em décimo lugar no ranking mundial possuindo 3,4 km/1000 km<sup>2</sup>, totalizando cerca de 29,18 mil km de malha ferroviária (ANTF, 2023).

**DENSIDADE DAS MALHAS FERROVIÁRIAS**

	Área (milhões km <sup>2</sup> )	Ferrovias (mil km)	Ferrovias/Áreas (km/ 1.000 km <sup>2</sup> )
 EUA	9,83	293,56	29,8
 Índia	3,29	68,53	20,8
 África do Sul	1,22	20,99	17,2
 Argentina	2,78	36,92	13,3
 China	9,60	124,00	13,2
 México	1,96	15,39	7,8
 Canadá	9,98	77,93	7,8
 Rússia	17,1	87,16	5,1
 Austrália	7,74	36,97	4,8
 Brasil	8,52	29,18	3,4

Figura 3.2: Densidade das malhas ferroviárias no mundo. Extraído de ANTF (2023)

Nesse sentido, o sistema ferroviário é extremamente importante ao mercado interno que representa a maior parte do Produto Interno Bruto (PIB) nacional. As exportações de insumos ou serviços também tem sua contribuição, representando cerca de 14% do PIB brasileiro. E mesmo que o principal meio de transporte para exportação seja portuário, para que os produtos cheguem aos portos, é necessário a utilização do modal rodoviário e ferroviário. Em dados, cerca de 61% do total transportado é por via rodoviária e 22% via ferroviária. Apesar da diferença, os desafios para o crescimento de setores produtivos da economia, como o agronegócio e indústrias, exigem que haja um equilíbrio entre as duas vias de transporte, resultando numa logística mais eficaz e rentável. Isso pode ser explicado pela capacidade das ferrovias serem capazes de carregar quantidades de cargas maiores, com frentes ausentes de pedágios e completar viagens mais rapidamente (MASSA INDÚSTRIA, 2023).

No Brasil, o sistema ferroviário tem sido amplamente empregado para o transporte de minérios e grãos, enquanto que o transporte de passageiros está principalmente concentrado no transporte público urbano de algumas capitais. A Estrada de Ferro Vitória-Minas (EFVM) e a Estrada de Ferro de Carajás (EFC), que integram o sistema multimodal de logística da VALE, são uma das maiores e mais importantes linhas férreas do país, com 905 quilômetros de extensão (2.141,68 quilômetros de linha total) e 892 quilômetros com trecho majoritariamente tangente, respectivamente, sendo responsáveis pelo transporte de diferentes tipos de produtos e pelo traslado de milhares de pessoas. Vale destacar também a Ferrovia Norte-Sul (EF-151), ferrovia que será capaz (após a conclusão do seu projeto) de transportar produtos e pessoas por todas as 5 regiões do Brasil, possuindo 4.155 quilômetros de extensão.

Tendo em vista a relevância que o setor ferroviário tem no desenvolvimento do país, colaborando no funcionamento do mercado interno de insumos e na movimentação de pessoas, diversas pesquisas e estudos têm sido promovidos e realizados (principalmente nas áreas de projetos, monitoramento e manutenção) com o intuito de desenvolver, melhorar e garantir a segurança do sistema.

### 3.2 Composição do Sistema Ferroviário

O conjunto de elementos que formam as linhas férreas (Figura 3.3) podem ser subdivididos em dois sistemas: infra-estrutura e superestrutura. A infra-estrutura corresponde às estruturas de assentamento (subleito) que servem para dar suporte aos demais elementos da via, enquanto que a superestrutura é composta pelo sublastro, lastro, dormentes, trilhos e os acessórios de fixação do trilho ao dormente.

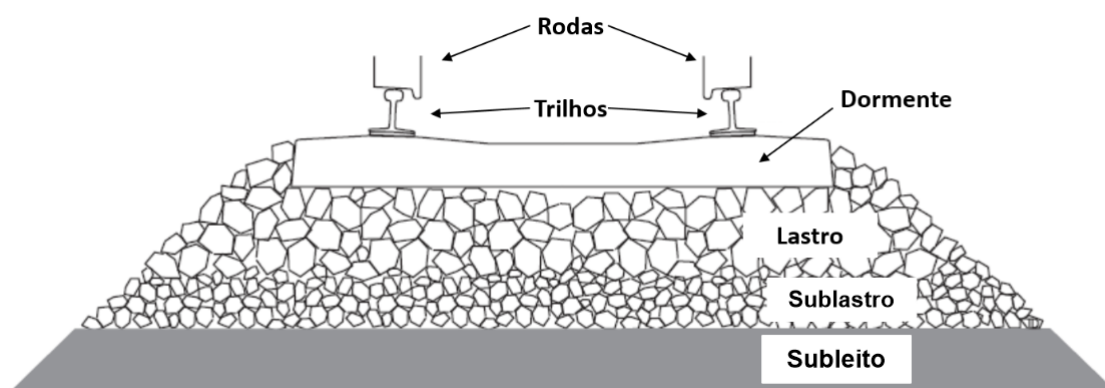


Figura 3.3: Infraestrutura e superestrutura de um sistema ferroviário. Adaptado de Iwnicki, Björklund e Enblom (2009)

O lastro é uma zona formada por um material pétreo que recebe a carga do dormente e a transmite uniformemente para o sublastro, além de: estabilizar a via, garantir o amortecimento das ações dinâmicas das cargas dos trens, proteger o sistema da umidade devido a drenagem da água da chuva e permitir o nivelamento das vias. O sublastro consiste em uma camada formada por materiais naturais (cascalhos, lateritas, solos arenosos, etc.) que possui a função de absorver as cargas transmitidas pelo lastro e transferi-los para a infraestrutura.

Os dormentes, produzidos em aço, concreto ou madeira, são vigas transversais que apoiam os trilhos, mantendo-os fixos e transmitindo a sua carga para o lastro de maneira uniforme. A fixação garantida pelos dormentes também auxilia na configuração do espaçamento entre os trilhos, chamada de bitola (*gauge*), parâmetro este que sofreu diversas mudanças ao longo dos anos, variando de acordo com a origem dos fornecedores de material rodante e da sua aplicação. No Brasil, é comum encontrar bitolas estreitas de 1000 mm (EFVM por exemplo), a bitola padrão de 1435 mm e a bitola irlandesa de 1600 mm (EFC por exemplo). Além disso, é empregado um pequeno ângulo conhecido como "rail cant" na conexão entre o trilho e o dormente, com a finalidade de alinhar o trilho com a inclinação das rodas (ângulo  $\beta$  da Figura 3.4).

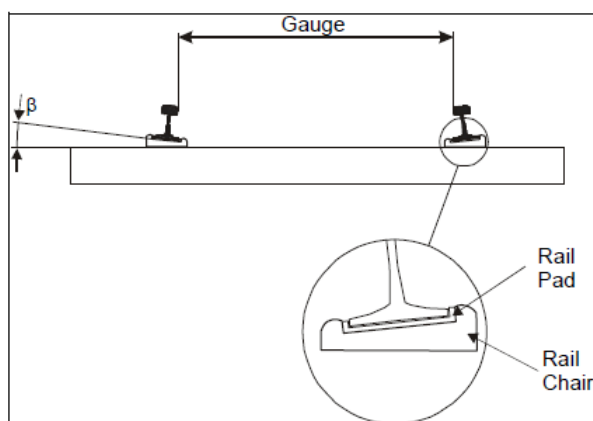


Figura 3.4: Representação da bitola e do ângulo *cant* em uma linha férrea. Extraído de Zakharov et al. (2001)

Os trilhos são elementos longitudinais (geralmente feito de aço) que possuem como principal função permitir o rolamento e servir de suporte para as cargas transportadas pelas locomotivas. De acordo com Rives, Pita e Puente (1977), o trilho é o elemento ativo da superestrutura ferroviária e, sobre ele, são aplicadas diversas ações procedentes do material rodante. A geometria atual dos trilhos está representada na Figura 3.5, cujo formato garante maior resistência à flexão, uma vez que a maior parte da massa dos trilhos está concentrada na região onde haverá maior solicitação normal, otimizando o desempenho do componente. O trilho pode ser dividido em três seções denominadas:

boleto, alma e patim. O boleto consiste na parte de cima do trilho, cuja função é servir de pista de rolamento para as rodas dos trens. O patim corresponde à base do trilho, conferindo apoio e fixação do trilho no dormente. Já a alma, consiste na porção de material que conecta o patim ao boleto, fornecendo altura e exercendo papel fundamental no momento de inércia da geometria. Geralmente, os trilhos são fabricados com ligas de aços, cuja composição consiste em combinar o ferro (Fe) com vanádio (V), cromo (Cr), Molibdênio (Mo), Titânio (Ti) e Nióbio (Nb). Esses elementos de ligas são utilizados uma vez que possuem elevada tensão de escoamento e ruptura, bem como composição química uniforme.

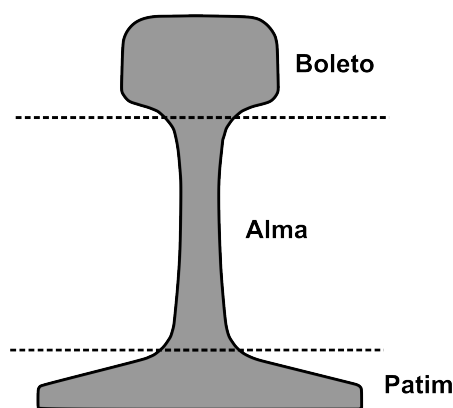


Figura 3.5: Perfil de trilho tipo 136RE padrão AREMA. (Autoria Própria)

Todo o sistema descrito acima tem como única função, permitir a movimentação de composições ferroviárias de transportes de cargas (trens) que são constituídos por locomotivas (veículos motorizados) e vagões conectados, cujo número pode chegar até 400.

Cada vagão de carga é comumente construído com dois conjuntos de truques, que desempenham um papel essencial como sistemas de suspensão e direção. Esses truques são conectados ao vagão de forma articulada, permitindo maior mobilidade em curvas e transições para vagões mais longos em comparação com uma configuração de eixos fixos diretamente ao vagão. Dentre os diferentes truques que existem, o de três peças é um dos mais utilizados no transporte de carga (Figura 3.6). Ele é composto por uma travessa central e duas vigas laterais, o que explica o seu nome. O vagão se apoia articuladamente na travessa central, que, por sua vez, é suportada pelas vigas laterais utilizando um conjunto de molas e cunhas de fricção para amortecimento. As vigas laterais se conectam por meio de mancais aos rodeiros, permitindo a livre rotação dos mesmos. Cada rodeiro ferroviário é composto por duas rodas de aço rigidamente conectadas por um eixo de aço, fazendo com que as rodas direita e esquerda girem juntas (ZAKHAROV et al., 2001).

Para garantir a estabilidade do trem sobre os trilhos, as rodas ferroviárias possuem duas características distintas: pistas de rolamento cônicas e flanges no lado interno

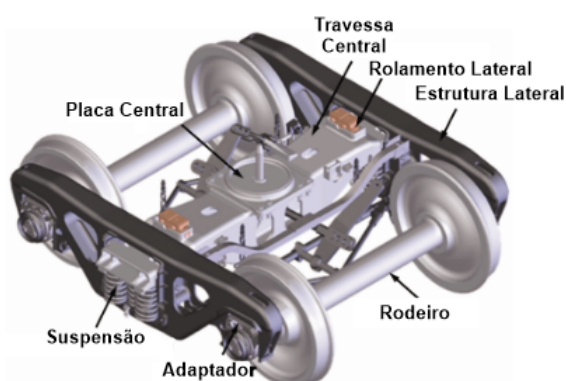


Figura 3.6: Truque de três peças. Extraído de Wu et al. (2018)

(Figura 3.7). A pista cônica desempenha um papel crucial no alinhamento do rodeiro com a via em trechos retos. Se houver algum deslocamento lateral do rodeiro, a conicidade das rodas permite que cada roda tenha um raio de rolamento diferente. Mesmo com a mesma velocidade angular, isso resulta em velocidades de translação distintas para cada roda, o que provoca um movimento de retorno do rodeiro à sua posição central na via. Essa diferença nos raios de rolamento proporcionada pela conicidade das rodas também é benéfica ao descrever curvas, permitindo que as rodas externas tenham uma velocidade de translação maior do que as rodas internas. Assim, evita-se um desgaste excessivo dos componentes, uma vez que o escorregamento entre rodas e trilhos é reduzido.

Já a flange da roda atua como um limitador dos movimentos laterais do rodeiro quando a conicidade das rodas não é suficiente para garantir o alinhamento adequado. Isso ocorre, por exemplo, em curvas que exigem uma diferença de raios de rolamento maior do que a proporcionada pela conicidade ou um deslocamento lateral do rodeiro maior do que a folga entre a flange e o trilho. Nessas situações, a flange entra em contato com a superfície lateral do trilho para permitir a realização da curva. No entanto, esse tipo de contato resulta em deslizamento significativo entre as partes, o que é considerado prejudicial nas ferrovias devido ao desgaste acelerado tanto da roda quanto do trilho.

A interação entre a roda e o trilho desempenha um papel fundamental na resolução dos desafios relacionados ao transporte de cargas pesadas. Essa interface requer a minimização do atrito, a fim de possibilitar a movimentação suave de cargas pesadas com o mínimo de resistência. Ao mesmo tempo, é necessário garantir a existência de um nível adequado de atrito, a fim de permitir a tração, frenagem e direcionamento eficientes do trem.

O contato entre a roda e o trilho resulta em mínima deformação dos corpos sob carga, levando a um contato de rolamento com perdas de energia por atrito reduzidas ao longo da área de contato e mínima dissipação de energia nos materiais envolvidos.

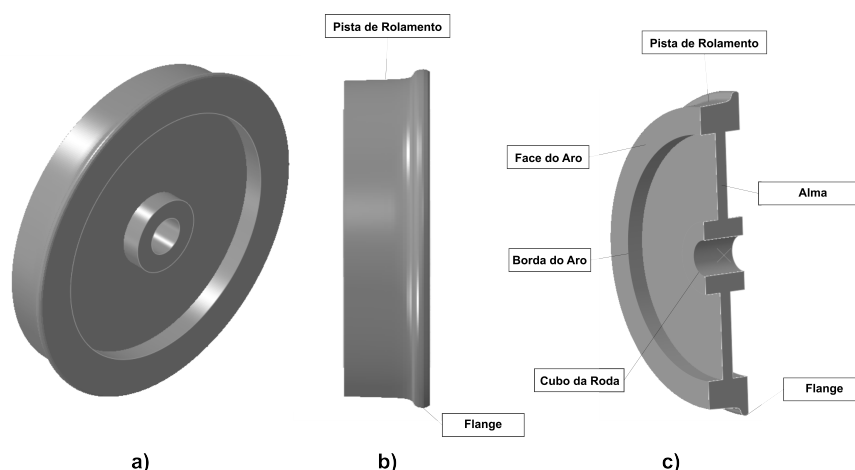


Figura 3.7: Vista isométrica (a), lateral (b) e em corte (c) de uma roda AREMA. (Autoria Própria)

Essa propriedade permite que as ferrovias apresentem uma resistência ao rolamento extremamente baixa, viabilizando o transporte eficiente de cargas de grande magnitude.

Os principais parâmetros que regem as forças de contato, as tensões e deformações, bem como as pressões de contatos e a posição do ponto de contato no conjunto roda-trilho, são os perfis das rodas e dos trilhos. Logo, qualquer alteração que ocorra nessas geometrias corrobora diretamente na modificação do comportamento do sistema ferroviário. Sendo assim, diversos estudos acerca da mecânica do contato são realizados com o intuito de aprimorar o sistema ferroviário bem como encontrar soluções que permitam aumentar a vida útil do sistema.

### 3.3 Mecânica do Contato Roda-Trilho

Segundo Iwnicki, Björklund e Enblom (2009), quando dois corpos entram em contato, as forças aplicadas são transferidas através da área de contato, a qual geralmente é significativamente menor em relação ao tamanho dos corpos. Os principais objetivos em uma análise de mecânica do contato são o cálculo das tensões e deformações na interface de contato bem como no interior dos corpos.

De acordo com Czichos e Winer (1978), os principais pontos que devem ser considerados em uma análise são: o número de corpos envolvidos no contato, as geometrias dos corpos (dimensionalidade, rugosidade), as propriedades dos materiais, o modo de deformação (puramente elástica, elasto-plástica ou puramente plástica), o tipo de movimento relativo (estático, rolamento, deslizamento, etc), a velocidade do movimento relativo e as forças de contato (normais e tangenciais).



De acordo com Mattei e Di Puccio (2022) a geometria do contato pode ser categorizada em dois tipos: conforme e não-conforme. No do tipo conforme (Figura 3.8.b) as geometrias em contato apresentam superfícies com curvaturas muito próximas, corroborando na formação de maiores áreas de contato (tamanho comparável às dimensões dos corpos). Como a área é maior, a pressão resultante do carregamento distribuído é menor. Já do tipo não-conforme (Figura 3.8.a) consiste no contato entre corpos cujas superfícies não possuem curvaturas que se encaixam, gerando pequenas áreas de contato e, conseqüentemente, maiores pressões de contato. Contudo, a pressão interfacial e as tensões no interior e na vizinhança da zona de contato são muito maiores do que as tensões distribuídas no restante do corpo, logo, problemas do tipo não-conforme podem ser estudados isoladamente.

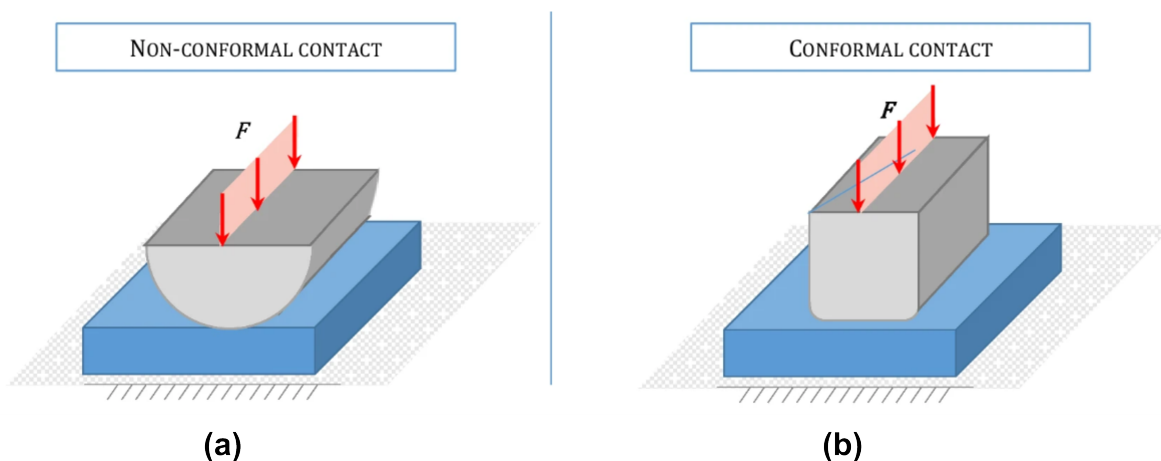


Figura 3.8: Contato do tipo Conforme (a) e do tipo Não-Conforme (b). Extraído de Mattei e Di Puccio (2022)

De acordo com Iwnicki, Björklund e Enblom (2009), o ponto de partida para uma análise do contato roda-trilho é a identificação do tamanho e da forma da área de contato. No entanto, para isso, deve-se ter informações importantes acerca dos perfis dos trilhos e das rodas, de modo que se possa observar também, a forma com que as geometrias se desviarão significativamente do projeto inicial após utilização (desgaste ou outros danos). Para tal, é possível realizar medições precisas dos perfis usando dispositivos mecânicos, como o MiniProf (Figura 3.9), ou a laser. Perfis típicos de rodas e trilhos são apresentados na Figura 3.10.

A teoria clássica do contato foi elaborada por Hertz (1882) que provou que a área de contato em um contato não-conforme entre dois corpos de revolução é elíptica, além de propor formas de calcular os semi-eixos da elipse e a distribuição de pressão no contato. Embora a teoria de Hertz desconsidere a presença de atrito e se limite a casos onde os corpos são perfeitamente elásticos, o seu modelo ainda é amplamente utilizado em



Figura 3.9: Medições dos perfis de um trilho e de uma roda através do miniprof. Extraído de Iwnicki, Björklund e Enblom (2009)

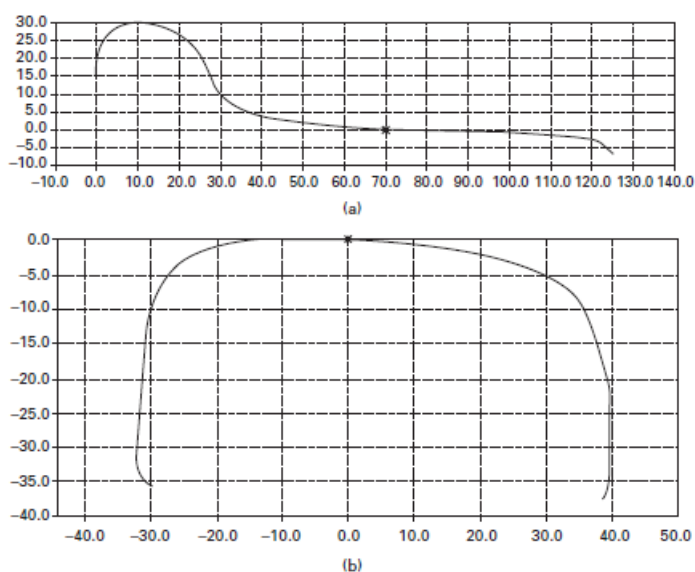


Figura 3.10: Perfis de roda (a) e de trilho (b). Extraído de Iwnicki, Björklund e Enblom (2009)

estudos iniciais em diversos casos de contato, além de servir de base para a modelagem do contato geral (IWNICKI; BJÖRKLUND; ENBLUM, 2009).

### 3.3.1 Modelagem do Contato Geral

O principal objetivo da modelagem do contato é a obtenção das tensões e deformações originadas do contato entre dois corpos sujeitos a cargas normais e/ou tangenciais ao plano.

No que concerne ao contato sujeito a cargas normais, o modelo de contato proposto por Hertz (1882) é utilizado para tal, onde o mesmo não considera a presença de cargas tangenciais. A Figura 3.11 apresenta um contato Hertziano usual entre dois corpos, onde:

- As direções  $x$  e  $y$  correspondem as direções dos raios principais de cada corpo;

- $R_{1x}$ ,  $R_{1y}$ ,  $R_{2x}$  e  $R_{2y}$  representam os raios de curvatura das geometrias;
- O  $\theta$  consiste no ângulo entre os eixos  $x_1$  e  $x_2$ ;

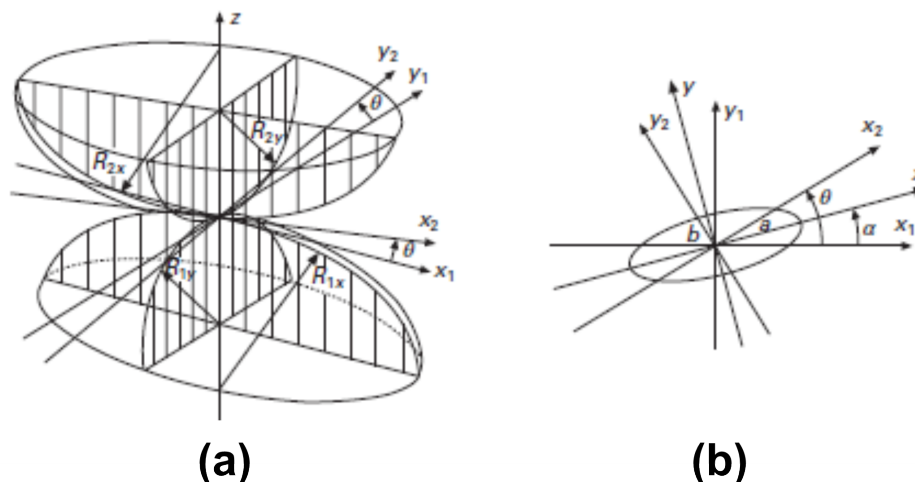


Figura 3.11: Contato geral de Hertz entre dois corpos: a) contato entre os sólidos, b) área de contato elíptica. Extraído de Iwnicki, Björklund e Enblom (2009)

É importante salientar que a forma da área de contato Hertziano será sempre uma elipse com o eixo maior estando entre os eixos  $x_1$  e  $x_2$ . Embora não haja equações gerais para o modelo de Hertz, as pressões e as deformações podem ser calculadas seguindo equações específicas, a partir dos dados geométricos dos corpos, das propriedades dos materiais dos corpos (módulo de elasticidade e coeficiente de Poisson) e da carga normal aplicada (IWNICKI; BJÖRKLUND; ENBLÖM, 2009).

De acordo com Iwnicki, Björklund e Enblom (2009), quando um corpo é sujeito a um carregamento normal e a uma força tangencial crescente, ocorrerá escorregamento entre os corpos em contato. Esse escorregamento se inicia nas extremidades do contato, onde a pressão é baixa, caracterizando o micro deslizamento. A região de aderência estará localizada no centro do contato, assumindo uma forma retangular para um contato linear e uma forma circular para um contato circular. A região de deslizamento, representada na Figura 3.12, circunda a região de aderência e possui formato anelar para contatos circulares.

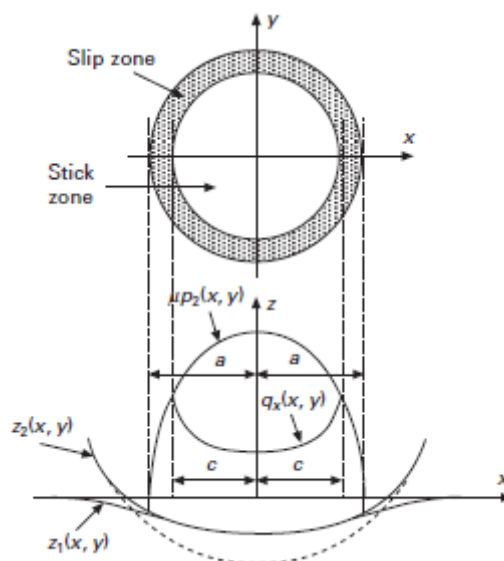


Figura 3.12: Carregamento normal e tangencial em um contato entre corpos. Extraído de Iwnicki, Björklund e Enblom (2009)

### 3.3.2 Análises do Contato Roda-Trilho

De acordo com Iwnicki, Björklund e Enblom (2009), o contato entre a roda e o trilho é do tipo rolamento-deslizamento, envolvendo tensões normais e tangenciais. Normalmente, as distribuições de tensões normais e de cisalhamento estão interligadas e não são independentes. No entanto, é viável analisar o contato roda-trilho decompondo as soluções para ambas distribuições de tensões.

A incidência do contato roda-trilho em via reta se dá maior no centro da banda de rodagem (Figura 3.13). O contato na região é mais definido quando há perfis de rodas cônicos puros em contato com perfis de trilho que possuem centro de curvatura acentuado no boleto. Já em curvas, o contato no perfil das rodas é invariavelmente simétrico enquanto que no trilho é assimétrico e depende do sentido da curva. Geralmente, em curvas, a roda que fará menor raio de curvatura (de dentro) se desloca de modo que a flange entre em contato com a superfície lateral do boleto do trilho, auxiliando na manutenção do carrilhamento do trem. Entretanto, esse tipo de contato corrobora em maior desgaste dos perfis, além de possuir maiores tensões de contato e fluência (ZAKHAROV et al., 2001).

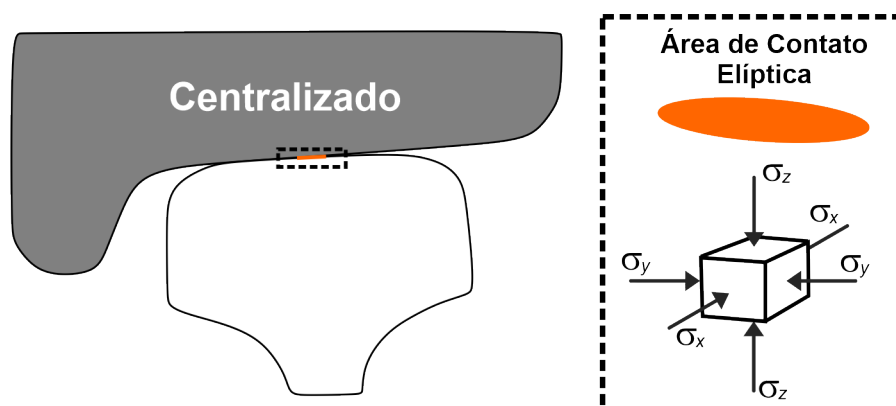


Figura 3.13: Configuração de um contato centralizado entre roda e trilho e sua respectiva área de contato. Adaptado de Zakharov et al. (2001)

Zakharov et al. (2001) divide a região de contato entre roda e trilho da seguinte maneira (Figura 3.14):

- Região A: Contato entre pista de rolamento e topo do boleto;
- Região B: Contato entre flange da roda com o canto do boleto;
- Região C: Contato entre as laterais da roda e do trilho;

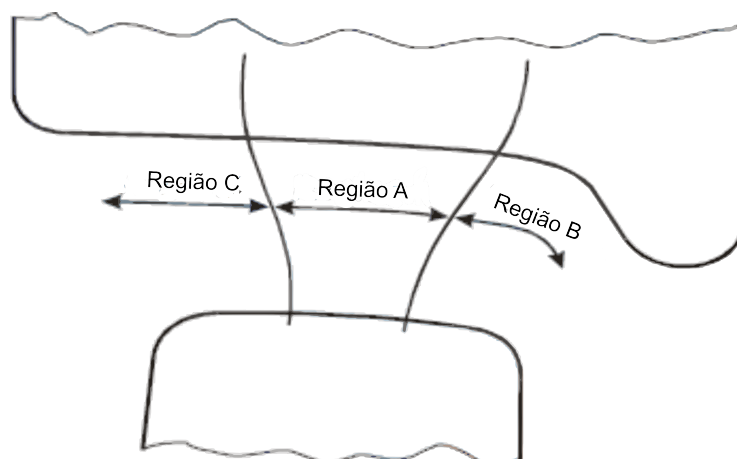


Figura 3.14: Diferentes regiões de contato roda-trilho. Adaptado de Zakharov et al. (2001)

Em geral, o contato entre roda e trilho é estabelecido na Região A, principalmente em trechos de via reta ou com raios de curvaturas pequenos. Nesse contato há menores pressões e forças laterais, ao passo que as velocidades dos veículos são maiores (Figura 3.13).

A área de contato na Região B é menor do que nas outras regiões, consistindo, geralmente, no contato da flange da roda com a superfície lateral do boleto do trilho. O contato para este caso pode ser em apenas um ponto (Figura 3.15b), viabilizando a

presença de altas tensões com escorregamento longitudinal, bem como pode ocorrer em dois pontos (Figura 3.15a), onde o desgaste e a fluência do material serão predominantes.

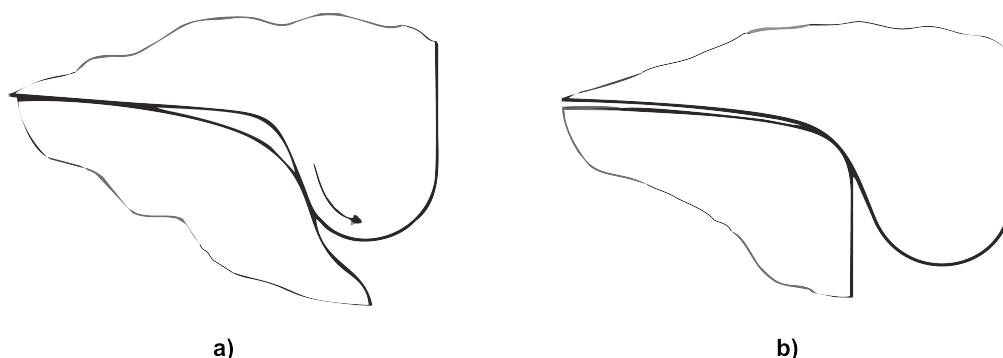


Figura 3.15: Contato em dois pontos a) e em um único ponto b) entre roda e trilho. Adaptado de Zakharov et al. (2001)

Já a Região C consiste no contato menos frequente de ocorrer, apresentando altas pressões de contato, escorregamento longitudinal e elevado desgaste seguido de fluência do material devido o esterçamento incorreto dos eixos (Figura 3.16).

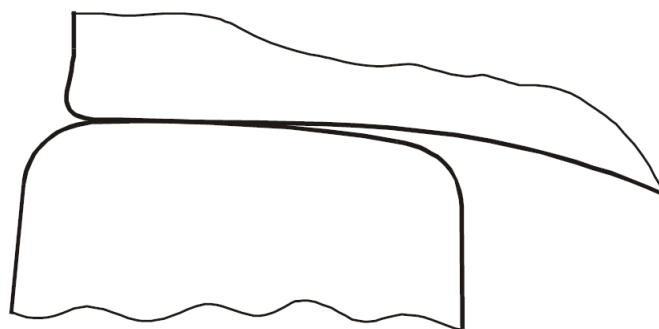


Figura 3.16: Exemplo de contato na Região C entre roda e trilho. Adaptado de Zakharov et al. (2001)

No contato roda-trilho, diversos defeitos podem surgir com o passar do tempo, tais como: *squats*, corrosão, *head checks* e corrugação. Os *squats* consistem em danos resultantes do fluxo plástico associado a altas cargas dinâmicas sobre pequenos defeitos triviais (juntas, soldas). Os *head checks* representam os problemas predominantes em trilhos e geralmente ocorrem na lateral do boleto. Esse defeito é resultado do acúmulo de deformação plástica, cuja região de nucleação de trincas é favorecida pelas cargas e atrito elevados, levando à formação de fissuras superficiais. A corrosão é outra preocupação significativa, especialmente em regiões úmidas (túneis por exemplo), levando à degradação do material ferroso, comprometendo a integridade da linha. Já as corrugações consistem no desgaste oscilatório dos trilhos causado pelas interações roda-trilho a longo prazo, causando problemas como desconforto para passageiros e ruídos.

Esses e outros defeitos surgem e são concebidos através de dois tipos de danos que o sistema pode apresentar: desgaste e fadiga de contato de rolamento (RCF - *rolling contact fatigue*).

Segundo Lewis e Olofsson (2009), o desgaste consiste na perda ou desprendimento de material de uma superfície que está em contato (Figura 3.17). Os materiais podem experimentar desgaste através de vários mecanismos ou modos, e o tribossistema engloba a composição do material, condições do ambiente e a existência de qualquer contaminante, como partículas provenientes do desgaste (*wear debris*). No contexto do contato entre roda e trilho, também podem estar presentes modificadores de atrito e substâncias que o intensificam, como areia, folhas, entre outros.

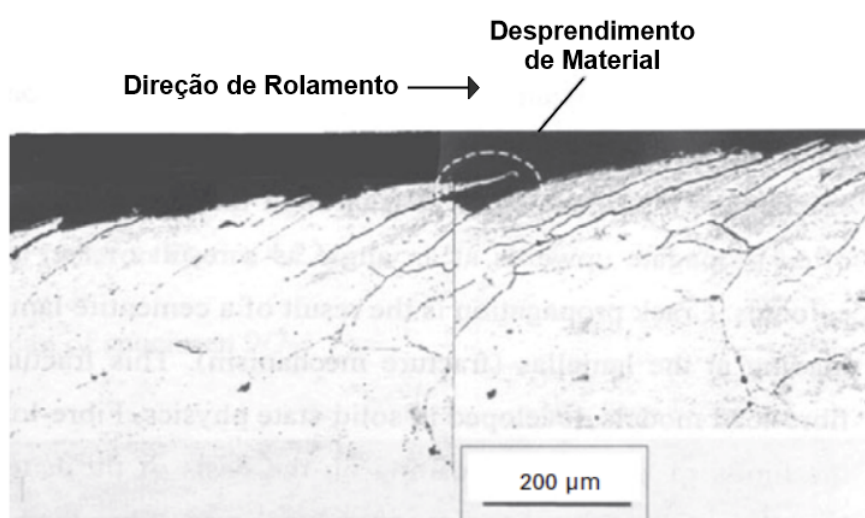


Figura 3.17: Remoção de material resultante de uma situação de desgaste por rolamento/deslizamento. Adaptado de Lewis e Olofsson (2009)

A fadiga de contato consiste na formação e propagação de trincas na roda e no trilho com o passar do tempo devido aos ciclos de utilização. As falhas podem ser superficiais (na superfície) ou subsuperficiais (abaixo da superfície). Segundo Lewis e Olofsson (2009), a fadiga na superfície da roda e nos trilhos resulta do acúmulo de deformação plástica (*ratcheting*) que posteriormente sofre danos de natureza frágil, com formação de trincas, devido o encruamento. Essas trincas se propagam para o interior do material até atingir cerca de 5 mm de profundidade, ocorrendo a fratura quando elas se ramificam ao longo da pista de rolamento da roda. A Figura 3.18 ilustra a aparência típica de uma falha por fadiga na superfície da roda. Trincas iniciadas em regiões subsuperficiais, formadas a vários milímetros abaixo da superfície, também são responsáveis por fraturas catastróficas e pela fragmentação da roda, após a sua propagação.

Os problemas críticos no contato roda-trilho estão interligados de certa forma (Figura 3.19). Segundo Lewis e Olofsson (2009), o desgaste e a fadiga de contato (RCF) estão



Figura 3.18: Dano por fadiga na superfície de uma roda. Extraído de Lewis e Olofsson (2009)

conectados em situações de danos no contato roda-trilho. Quando o crescimento das trincas é impedido, o desgaste tende a aumentar, ao passo que, se o desgaste for retardado, a fadiga de contato será favorecida. A utilização de lubrificantes e modificadores de atrito é comumente empregada para reduzir o desgaste, porém, eles também podem influenciar no crescimento de trincas do tipo RCF, tendo em vista que a evolução desse tipo de dano é ditada pela pressão de contato e pelo atrito. Há ainda a possibilidade de controlar a magnitude do atrito através dos modificadores de atrito.

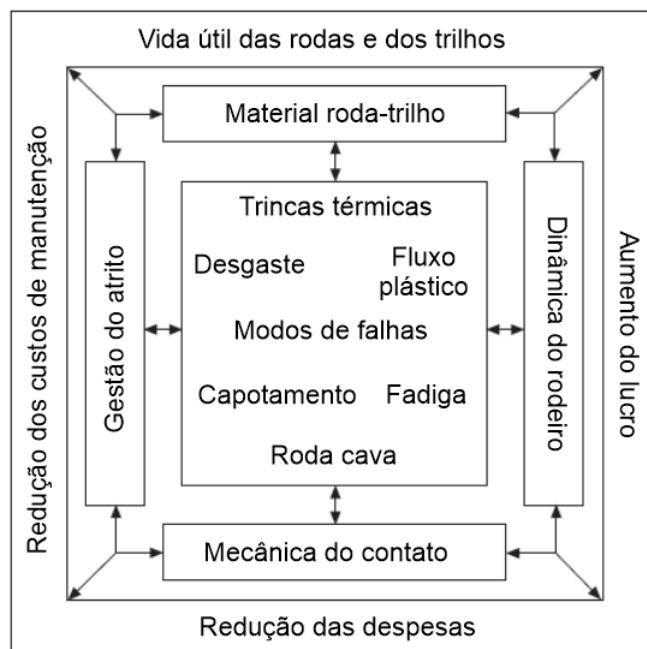


Figura 3.19: Abordagem sistêmica da interface de gerenciamento e pesquisa para o tribossistema roda-trilho. Adaptado de Lewis e Olofsson (2009)



### 3.3.3 Métrica para Avaliação do Tipo de Contato

Como mencionado na Introdução, Magel e Kalousek (2002) propõe duas métricas que são capazes de classificar o tipo de contato entre roda e trilho: separação máxima em um único ponto ( $s$ ) e separação máxima em dois pontos ( $d$ ). A primeira corresponde à distância entre as superfícies da roda e do trilho ao longo da linha do centróide da geometria do trilho (Figura 3.20). Já a segunda, consiste na distância entre a superfície lateral do boleto do trilho e a flange da roda. O autor aborda ainda um outro tipo de contato chamado de quase conforme, que compreende um contato de transição entre o conforme e o não-conforme.

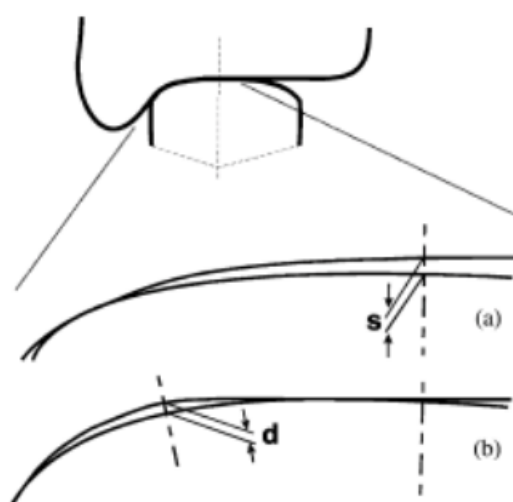


Figura 3.20: Representação das métricas no contato roda-trilho. Extraído de Magel e Kalousek (2002)

Segundo o autor, um contato será considerado quase conforme, quando os valores de  $s$  ou  $d$  forem menores que 0,1. Quando as distâncias estiverem na faixa entre 0,1 e 0,4, o contato será conforme, enquanto que será não-conforme quando os valores forem maiores que 0,4. A Tabela 3.1 expõe o que foi descrito.

Tabela 3.1: Definição de Conforme, Não-Conforme e Quase Conforme.

Não-Conforme	$(s \text{ ou } d) > 0,4 \text{ mm}$
Conforme	$0,1 \text{ mm} < (s \text{ ou } d) < 0,4 \text{ mm}$
Quase Conforme	$(s \text{ ou } d) < 0,1 \text{ mm}$

Ainda de acordo com o autor, nos contatos do tipo conforme e quase conforme o desgaste será o dano predominante no sistema, ao passo que o contato do tipo não-conforme favorecerá a fadiga de contato.

## 3.4 Aspectos Teóricos da Inteligência Artificial

Segundo Franco (2017), o primeiro trabalho científico que surgiu acerca de inteligência artificial foi em 1943, através da elaboração de um modelo de neurônios artificiais, proposto por Warren McCulloth (1898-1969) e Walter Pitts (1923-1969), que serviu como precursor da abordagem conexionista. Outro autor muito importante que contribuiu também para os primeiros estudos na área de IA foi o Alan Turing (1912-1954), matemático britânico que ficou conhecido por sua colaboração na decodificação da criptografia de mensagens do exército alemão na Segunda Guerra Mundial, através dos seus conhecimentos acerca de inteligência de máquinas computacionais. Em 1950, Turing propõem em um artigo um teste intitulado de “Teste de Turing” que ponderava sobre a possibilidade de uma máquina inteligente ser capaz de se comportar da mesma forma que um ser humano.

Entretanto o mundo só é apresentado pela primeira vez acerca do conceito de Inteligência Artificial (IA) em 1956, na qual Jonh McCarthy a definiu como “A ciência e engenharia de fazer máquinas inteligentes, especialmente programas de computadores inteligentes”.

De acordo com Franco (2017), os primeiros anos da IA foram bastante promissores, onde, dado as tecnologias primitivas da época, os computadores conseguiam resolver problemas e efetuar atividades que consideravam inteligentes. Além disso, os pesquisadores na área previam o desenvolvimento de uma máquina que seria capaz de superar a inteligência humana em um curto período de tempo. Entretanto essas ideias logo cessaram, uma vez que os computadores não conseguiam lidar com problemas mais complexos, na qual os limitados modelos de representação e o baixo poder computacional eram os principais obstáculos.

Todavia, por volta de 1980, dentro do campo da IA, surge o conceito de *Machine Learning* (ML) ou Aprendizagem de Máquina (Figura 3.21), cuja ciência é focada no estudo de como os computadores podem obter novos conhecimentos ou habilidades através da simulação da forma como os seres humanos aprendem, bem como também reorganizar as arquiteturas de conhecimento já existentes para a melhoria da performance. ML também pode ser descrita como o estudo dos algoritmos de aprendizagem capaz de resolver problemas cuja solução seja demasiada complexa para se solucionar através de métodos convencionais (métodos numéricos, análise de dados etc). Os algoritmos de machine learning podem atuar através de classificação (para valores discretos), regressão (para valores contínuos) e clusterização (categorização de dados) para a predição dos resultados. Atualmente existem três formas de aprendizagem de máquinas, sendo elas: aprendizagem supervisionada, aprendizagem não-supervisionada e aprendizagem por reforço.

A aprendizagem supervisionada consiste no treinamento de um modelo através da utilização de dados experimentais rotulados, isto é, para cada dado inserido no modelo, o

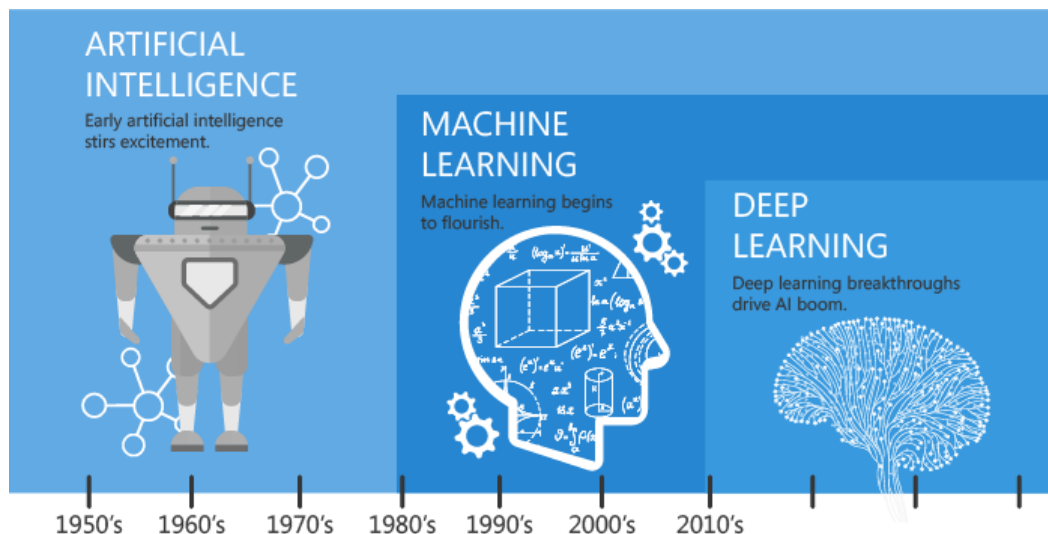


Figura 3.21: Linha do tempo resumida da inteligência artificial. Extraído de Lee e Yoon (2021)

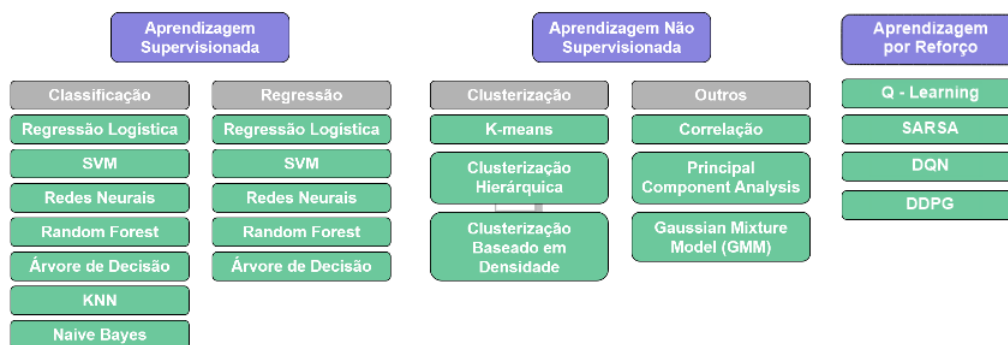


Figura 3.22: Algoritmos para aprendizagem supervisionada e não-supervisionada. (Autoria Própria)

algoritmo será informado acerca do resultado esperado para aquela determinada combinação de características. É muito comum neste tipo de aprendizagem a regressão e classificação como modelos de predição. Já a aprendizagem não-supervisionada consiste no treinamento de um modelo através de um banco de dados não-rotulados, isto é, não há um direcionamento dos resultados através da combinação das características dos dados. Dessa forma o algoritmo deverá, através do cálculo da similaridade entre os dados, aprender e dar um resultado adequado. A clusterização é muito utilizada nesse caso. Na Figura 3.22 são apresentados os principais algoritmos para cada tipo de aprendizagem.

A aprendizagem por reforço é uma área da aprendizagem de máquina que se preocupa em quais ações um modelo deve tomar para maximizar a sua performance de aprendizagem através de sinais equivalentes a recompensa e punição. Os sinais de reforço originados pelo ambiente são utilizados para avaliar se a ação executada pela máquina

trará um benefício ou malefício, notificando, portanto, ao sistema de aprendizagem qual é a forma correta de atuar.

Anos mais tarde, um novo campo da pesquisa da área de IA, chamada Redes Neurais, é encabeçada no desenvolvimento de modelos de aprendizagem. Uma rede neural (Figura 3.23) é um sistema de computador composto por uma simples e vasta conexão de elementos que processam informações através de uma resposta dinâmica à dados externos. As redes neurais são formadas por camadas de neurônios artificiais (perceptrons) conectados entre si que simulam o funcionamento do cérebro humano, possuindo características e funções como: processamento, aprendizado, associação, classificação e memória.

A primeira camada de uma rede neural é chamada de camada de entrada. Nela, como o próprio nome sugere, estão os dados de entrada ou as *features* que serão processadas e associadas a uma resposta. Cada neurônio corresponde a um dado de entrada. A última camada por sua vez é chamada de camada de saída, e é nela, que a resposta final é obtida, seja ela um único valor ou um vetor de resultados. O número de neurônios na camada de saída corresponde ao número de respostas (sejam eles valores contínuos ou discretos). Já as camadas que ficam entre as de entrada e saída são chamadas de camadas ocultas. Nelas, ocorrem os cálculos que tentam estabelecer as correlações entre os dados de entrada e de saída.

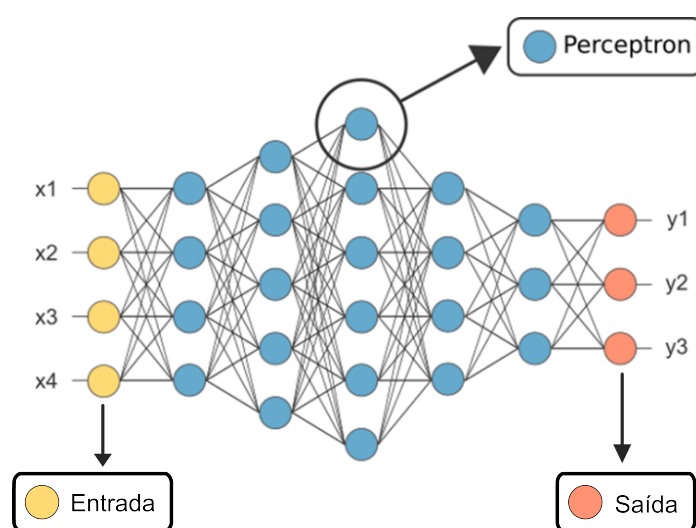


Figura 3.23: Ilustração de uma rede neural. Agrawal et al. (2021)

O neurônio matemático (Figura 3.24) desempenha um papel fundamental no processamento de informações nas redes neurais artificiais, permitindo a modelagem e a aprendizagem de padrões complexos. Sua estrutura e funcionamento proporcionam uma base sólida para o desenvolvimento de sistemas inteligentes e algoritmos de aprendizado de máquina.

O neurônio matemático é um modelo simplificado do neurônio biológico, sendo fundamental para o funcionamento das redes neurais artificiais. Ele recebe um ou mais sinais de entrada, representados matematicamente pelas sinapses, e produz um único sinal de saída, que pode ser uma saída final ou uma outra entrada para neurônios posteriores. Esse processo envolve a multiplicação dos sinais de entrada ( $x_i$ ) pelos pesos correspondentes ( $w_i$ ), gerando as entradas ponderadas. Em seguida, ocorre a função de combinação, na qual as entradas ponderadas são somadas para obter um resultado único.

Esse resultado é então submetido a uma função de ativação, que desempenha um papel crucial ao decidir se o neurônio deve ser ativado ou não. Ela evita o acréscimo progressivo dos valores de saída ao longo das camadas da rede, já que possui valores máximos e mínimos definidos em intervalos determinados. O uso de funções de ativação não-lineares confere à rede neural uma capacidade poderosa de processamento, uma vez que permite o modelo generalizar ou se adaptar a uma variedade de dados com um espaço maior de hipóteses.

Além disso, o perceptron pode incluir um viés de entrada, que é somado ao somatório da função de ativação. Isso permite que um neurônio tenha uma saída não nula mesmo quando todas as suas entradas são nulas, aumentando portanto a flexibilidade da função de ativação e, conseqüentemente, a capacidade de aproximação da rede. Sendo assim, um neurônio apresenta os seguintes componentes de acordo com a Figura 3.24.

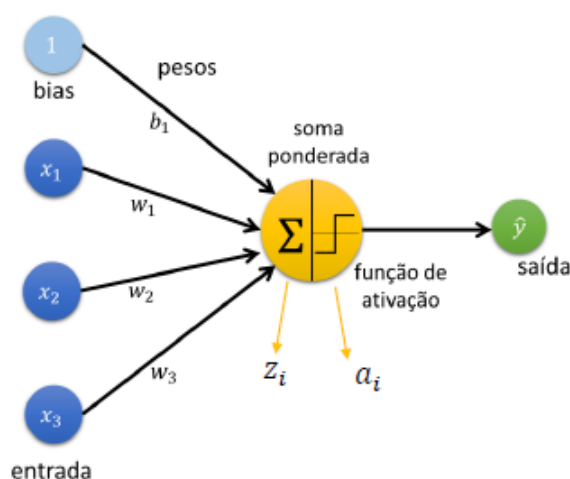


Figura 3.24: Perceptron e seus componentes. (Autoria Própria)

- Sinais de Entrada ( $x_1, x_2, \dots, x_n$ ): São os dados de entrada, comumente normalizados, que alimentam o modelo;

- Pesos ( $w_1, w_2, \dots, w_n$ ): São os coeficientes de ponderação que são aplicados aos respectivos dados de entrada. Os pesos são recalculados durante o treinamento da rede;
- Combinação linear ( $\sum$ ): Soma de todos os sinais de entrada que foram ponderados;
- Bias ( $b_i$ ): Valor adicionado à combinação linear;
- $z_i$ : Resultado da soma da combinação linear com o bias (Equação 3.1);
- $a_i$ : Resultado de  $z_i$  aplicado à função de ativação  $f(z_i)$ ;
- Saída da rede ( $\hat{y}$ ): É o dado de saída de um neurônio, podendo ser um sinal para um próximo neurônio ou a resposta de uma rede neural.

$$z^{(i)} = \sum_j (x_j^{(i)} \cdot w_j) + b^{(i)} = w^T \cdot x^{(i)} + b^{(i)} \quad (3.1)$$

Como mencionado anteriormente as funções de ativação podem agregar às redes neurais a não-linearidade, proporcionando-as a capacidade de resolver problemas complexos. Dentre as principais funções de ativação, é possível citar: Função Linear, Sigmóide, Tanh, Softmax, ReLU e *Leaky ReLU*.

A função de ativação Linear (embora não seja não-linear optou-se por trazê-la aqui) nada mais é do que a aplicação de uma transformação linear às entradas (Equação 3.2) para gerar uma saída. Ela é ideal para ser aplicada em problemas simples e sem muita complexidade.

$$\hat{y} = az \quad (3.2)$$

onde  $a$  é o coeficiente angular da função. A função Sigmóide (Equação 3.3) é suave, continuamente diferenciável e não-linear, sendo que suas respostas varia entre 0 e 1 (com a curva em formato de S). A função é amplamente utilizada em redes neurais de classificação, uma vez que ela aproxima os valores de  $\hat{y}$  para os extremos. Uma limitação dessa função, por exemplo, é que ela não retorna valores negativos nas saídas (fenômeno desejável em algumas situações).

$$\hat{y} = \frac{1}{1 + e^{-x}} \quad (3.3)$$

A função tangente hiperbólica (Equação 3.4) é semelhante à função sigmóide. Entretanto ela soluciona a limitação das saídas negativas, sendo simétrico à origem, variando no intervalo de -1 a 1. Entre outras características, ela é não-linear, contínua e diferenciável em todos os pontos.

$$\hat{y} = \frac{2}{1 + e^{-2x}} - 1 \quad (3.4)$$

A função Softmax (Equação 3.5), diferentemente da sigmóide e da tanh, pode ser utilizada para problemas de multiclassificação. Isso é possível devido sua capacidade de calcular probabilidades de uma entrada  $z_i$  pertencer a uma classe dentre várias.

$$\hat{y} = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (3.5)$$

onde  $z$  é um vetor de escalares e  $z_i$  e  $z_j$  são o  $i$ -ésimo e o  $j$ -ésimo elementos do vetor para  $i \in \{1, 2, \dots, n\}$  e  $j \in \{1, 2, \dots, n\}$ , respectivamente.

A função ReLU (Unidade Linear Retificada) nada mais é do que uma função não-linear que permite a inativação de neurônios, caso a saída seja um valor negativo, e a ativação do mesmo, caso a saída seja positiva. Essa habilidade da função pode tornar a rede esparsa, facilitando a sua computação e melhorando a eficiência nos cálculos e análises.

$$\hat{y} = \max(0, z) \quad (3.6)$$

A função de ativação *Leaky ReLU* (Equação 3.7) é uma derivação da função ReLU (Figura 3.25 e é amplamente utilizada em diversos modelos de IA. Sua principal diferença é a aplicação de um coeficiente  $\alpha$  nos casos em que os valores  $z$  forem menores que o, impedindo assim a inativação não saudável dos perceptrons (*dying* ReLU). Dessa forma, a função garante que a rede continue esparsa (sem perda de eficiência nos cálculos e na computação) e elimina o problema com gradientes nulos.

$$\hat{y} = \begin{cases} \alpha \cdot z, & z < 0 \\ z, & z \geq 0 \end{cases} \quad (3.7)$$

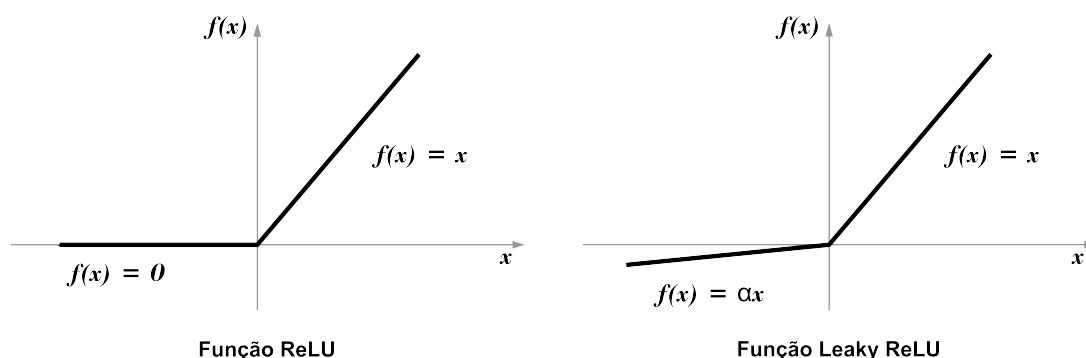


Figura 3.25: Gráficos da função ReLU e Leaky ReLU respectivamente. Adaptado de Li et al. (2022)

### 3.4.1 Redes Neurais Convolucionais

Atualmente existem diversos tipos de redes neurais que podem ser utilizadas para os mais diversos fins, dentre elas:

- Redes Neurais Densas (FNN, do inglês *Fully Connected Networks*): Aplicação em problemas de classificação, regressão, reconhecimento de padrões, etc;
- Redes Neurais Convolucionais (CNN, do inglês *Convolutional Neural Network*): Possui aplicações na área da visão computacional (reconhecimento de imagens, detecção de objetos, segmentação de imagem), processamento de imagem, etc;
- Redes Neurais Recorrentes: Aplicada no processamento de linguagem natural (tradução automática, geração de texto), previsão de séries temporais e análise de sequências;
- Redes Adversárias Generativas (GAN, do inglês *Generative Adversarial Networks*): Aplicações em geração de imagens, super-resolução de imagens, transferência de estilo e geração de texto.

Dentre os modelos de IA apresentados acima, convém destacar as redes neurais convolucionais (CNN). Sendo um dos modelos mais estudados nos últimos anos no que concerne ao reconhecimento e processamento de imagens, as CNNs têm flexibilidade considerável para permitir que os usuários façam modificações de acordo com suas necessidades, através da configuração e combinação dos seus componentes, sendo eles: filtros e camadas de *pooling* (YEH et al., 2023).

Os filtros (Figura 3.26), ou *kernel*, atuam como detectores de características e são aplicados nas regiões da imagem de entrada. Cada filtro é responsável por aprender características específicas, como bordas, texturas ou padrões complexos. Durante o treinamento, a rede ajusta os valores dos filtros para melhor identificar as características relevantes nas imagens. Essa convolução repetida dos filtros pela imagem de entrada permite que a CNN capture hierarquicamente informações de baixo nível (como bordas) até informações de alto nível (como características complexas).

Já as camadas de *pooling* são responsáveis por reduzir a dimensionalidade dos mapas de características obtidos através dos filtros, mantendo e agrupando as informações mais importantes em regiões menores (Figura 3.27). Existem dois tipos de camada de *pooling*: o *Max Pooling*, o mais comumente utilizado, que retém o valor máximo em cada região destacando as características mais proeminentes, bem como o *Average Pooling*, que calcula a média dos valores na região e ajudam a reduzir o número de parâmetros na rede, tornando-a mais eficiente e menos suscetível ao *overfitting*.



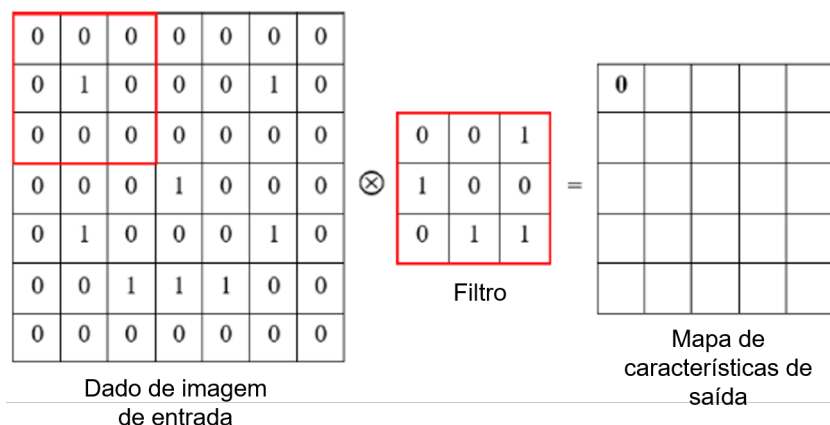


Figura 3.26: Extração de características de imagem no processo de convolução. Adaptado de Yeh et al. (2023)

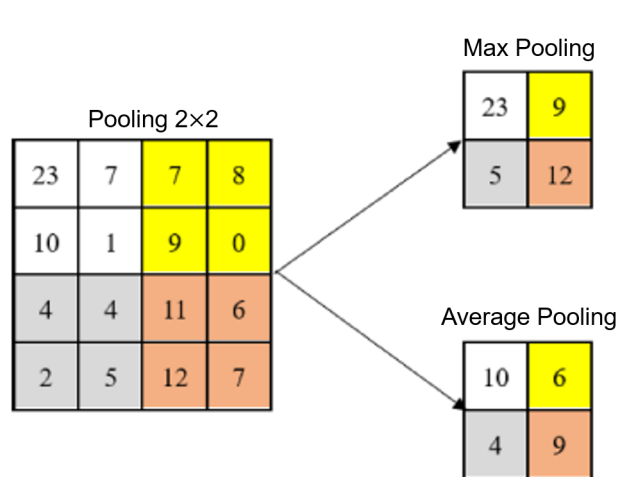


Figura 3.27: Redução da dimensão do mapa de características via processo de *pooling*. Adaptado de Yeh et al. (2023)

Nas camadas de convolução, há ainda a possibilidade de incluir camadas de ativação (funções de ativação) que introduzem não-linearidade na rede, permitindo a modelagem de relações mais complexas entre as características. A escolha do tipo de função de ativação, bem como o número de filtros e suas dimensões, o tipo de *pooling* e as dimensões das suas camadas, podem influenciar diretamente na performance da CNN e no custo computacional para o seu treinamento.

Vale salientar ainda que as redes neurais convolucionais, geralmente, são combinadas com outros tipos de redes neurais de modo a alcançar um determinado fim. Um exemplo bastante comum é a combinação de redes neurais convolucionais com redes neurais densas. Enquanto as CNNs são altamente eficazes na extração e identificação de características em imagens, as FNNs são especializadas em aprender representações globais dessas características para tomar decisões finais. A transição das CNNs para as FNNs ocorre

geralmente após as camadas de convolução e *pooling*, onde os recursos extraídos são "achatados" em um vetor unidimensional. Esses vetores são, então, alimentados em uma ou mais camadas de FNNs para realizar tarefas específicas, como classificação e/ou regressão.

A Figura 3.28 traz uma rede neural convolucional chamada de LeNet-4. Sendo um dos primeiros modelos convolucionais desenvolvidos cuja função é reconhecer a escrita de números a partir de imagens com 1 canal (preto, branco e tons de cinza), sua arquitetura conta com 2 camadas de convolução (6 filtros de dimensão 5 x 5 na primeira e 16 filtros de dimensão 5 x 5 na segunda), 2 camadas de *pooling* (onde cada camada possui dimensão de 2 x 2, reduzindo o tamanho dos mapas de características pela sua metade) e é conectada com duas camadas de redes neurais densas, sendo uma com 120 neurônios e a última (camada de *output*) com 10 neurônios, responsável por classificar o número presente em cada dado de imagem.

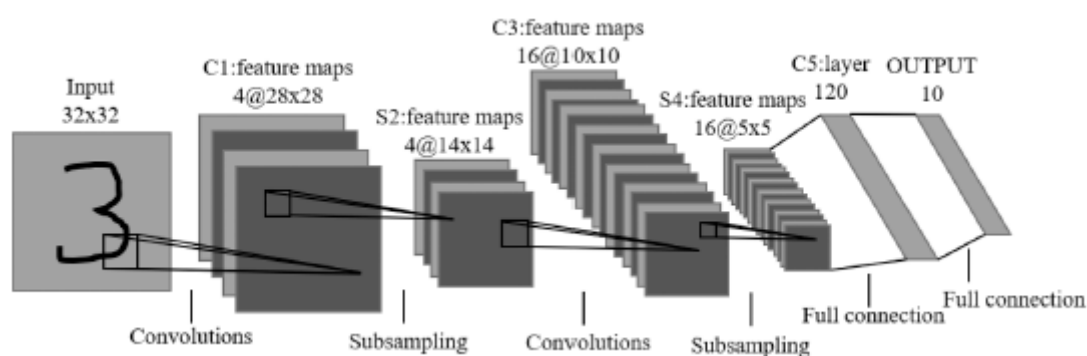


Figura 3.28: Rede neural convolucional LeNet-4. Extraído de Yeh et al. (2023)

### 3.5 Revisão da Literatura acerca da utilização de IA na área ferroviária

A Figura 3.29 apresenta um levantamento de artigos científicos, publicados nos últimos 5 anos, que trazem diversas aplicações da inteligência artificial no setor ferroviário. Esses trabalhos serviram de base para auxiliar na seleção do tipo de estratégia que será utilizado para o treinamento da rede neural.

Hajizadeh, Núñez e Tax (2016) utilizaram aprendizagem semi-supervisionada de classificação para definir se há ou não *squat* em um determinado trecho de via, através da análise e processamento de imagens. Os autores possuíam, inicialmente, 21.979 trechos de vias rotulados e 718.520 não rotulados. Apenas 3% de todos os dados apresentavam *squat*. Uma das metodologias utilizadas pelos autores é brevemente apresentada na Figura 3.30. Dados pré-processados referem-se a duas características (*features*) extraídas das imagens de segmentos de trilho: resultado da convolução do segmento com um

Título	Autores	Ano	Metodologia
MSRConvNet: Classification of railway track defects using multi-scale residual convolutional neural network.	Acikgoz e Korkmaz	2023	Utiliza um modelo que emprega conexões de salto com blocos de aprendizado residual e convoluções em múltiplas escalas para extrair características relevantes dos trilhos de modo a identificar e classificar defeitos. Aplica também data augmentation para auxiliar no balanceamento de dados.
Integration of Computer Vision and Convolutional Neural Networks in the System for Detection of Rail Track and Signals on the Railway.	Petrovic et al.	2022	Combina técnicas tradicionais de visão computacional, incluindo a suavização Canny, com o modelo de inteligência artificial "You Only Look Once" (YOLO), baseado em redes neurais convolucionais (CNNs) para a detecção de objetos de infraestrutura ferroviária.
A Bayesian machine learning approach for online detection of railway wheel defects using track-side monitoring.	Y. Ni et al.	2021	Desenvolvimento de um método probabilístico de Bayes. Sparse Bayesian Learning (SBL).
Analysis of Railway Track Irregularities with Convolutional Autoencoders and Clustering Algorithms.	Nietling et al.	2020	Detecção de anomalias através de métodos tradicionais de processamento de sinais com autoencoders e algoritmos de clusterização
Artificial neural networks applied to the measurement of lateral wheel-rail contact force: A comparison with a harmonic cancellation method.	P. Urda et al.	2020	Testa e compara vários algoritmos para uma rede neural para se obter a força lateral de contato do sistema roda-trilho. A acurácia da rede é comparada com a acurácia de um método chamado de Harmonic Cancellation. Os dados para treinar e validar a rede são retirados através de uma simulação FEM. Após isso, a rede é testada com dados experimentais.
Implementation of machine learning techniques for prognostics for railway wheel flange wear.	C.J. Fourie et al.	2020	Regressão logística, redes neurais artificiais e random forest são utilizados para prognóstico de desgaste na flange da roda.
Prediction of wheel and rail wear under different contact conditions using artificial neural networks	A. Shebani et al.	2018	Utilização de Nonlinear Autoregressive models with exogenous input neural network (NARXNN) para prever o desgaste de rodas e trilhos através de dados experimentais.

Figura 3.29: Levantamento de artigos que trazem aplicações de IA em ferrovias. (Autoria Própria)

conjunto de parametrizações da família de funções Gabor (método comum de extração de características no processamento de imagens) e DCT (do inglês *Discrete Cosine Transform*). Um modelo de aprendizado de máquinas de classificação SVM (do inglês, *Support Vector Machine*) é treinado utilizando-se os dados rotulados. Depois de treinado, o modelo é utilizado para gerar rótulos nos dados não rotulados. Os novos dados rotulados como positivos (para diminuir o problema de desbalanceamento) são utilizados para retreinar o classificador, assumindo que os novos dados ajudarão a melhorar seu desempenho.

Niebling, Baasch e Kruspe (2020) combinam métodos tradicionais de processamento de sinais com *autoencoders* e algoritmos de clusterização. Os autores primeiramente utilizaram uma metodologia de detecção de anomalias nos resultados obtidos com sensor ABA (*axle box acceleration*) nos trilhos da linha férrea. Em seguida, dividem as anomalias em 9 *clusters* através de técnicas de IA, porém não discutem com profundidade as diferenças entre os clusters encontrados.



Figura 3.30: Modelo de IA para a identificação da falha por aprendizado semi-supervisionado utilizando autotreinamento. (Autoria Própria)

Shebani e Iwnicki (2018) elaboraram uma rede neural NARXNN (do inglês, *nonlinear autoregressive with external input*) com o objetivo de prever o desgaste da roda e do trilho (mudança da geometria do contato roda-trilho). Esse tipo de rede é comumente utilizado para sistemas dinâmicos não lineares, cuja estratégia permite a previsão de uma série temporal a partir de valores anteriores da mesma série. Os dados para o treinamento da rede foram obtidos através de um teste disco-disco (*twin disc rig* – Figura 3.31), onde foi possível obter valores de desgaste e rugosidade para o sistema, seja à seco, úmido, lubrificado ou lixado. Como dados de entrada para o treinamento da rede, os autores utilizaram os seguintes parâmetros: perfis do par roda-trilho, carregamento, velocidade, ângulo de guinada e primeira e segunda derivadas dos perfis da roda e do trilho. Os dados de saída foram os valores de desgaste da roda e do trilho. A rede se mostrou eficiente na predição dos desgastes. A métrica definida foi o erro absoluto percentual médio.

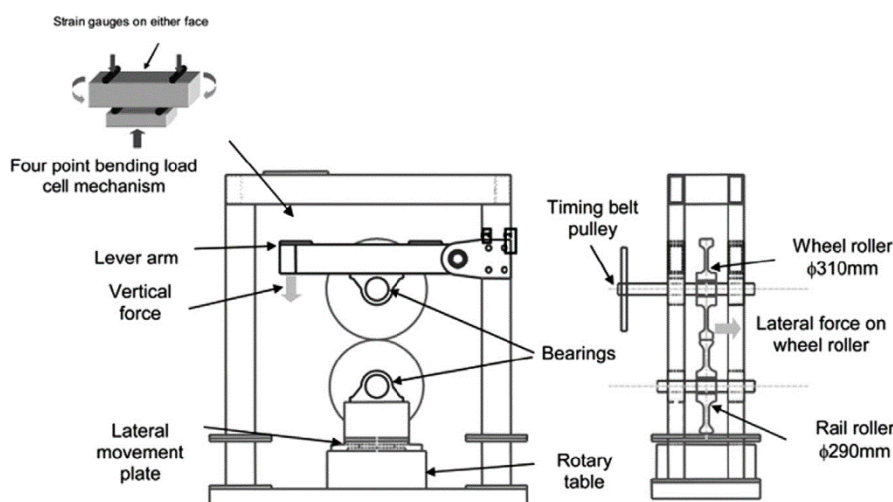


Figura 3.31: Esquema do teste disco-disco (*twin disc rig*). Extraído de Shebani e Iwnicki (2018)

Urda et al. (2020) desenvolveram uma rede neural que determina, de forma eficiente, a força lateral no contato roda-trilho através de dados da deformação radial da roda. Neste

trabalho, a rede neural é treinada e validada através de dados obtidos de simulação por MEF. Além disso, os autores compararam ainda diversos algoritmos, tais como *Resilient Backpropagation*, Levenberg-Marquardt, entre outros. O que apresentou melhor eficiência foi o Levenberg-Marquardt. Ademais, a rede foi testada com dados obtidos de forma experimental e apresentou desempenho satisfatório.

O trabalho desenvolvido por Jamshidi et al. (2018) propõe uma metodologia para auxiliar na tomada de decisões no que se refere à manutenção. Uma vez que haja a coleta de dados (sinais e imagens de vídeos) através do carro controle, a abordagem proposta possui cinco passos para o alcance do objetivo. O primeiro passo consiste na utilização de redes convolucionais profundas (DCNNs, do inglês *Deep Convolutional Neural Networks*) para a análise do grau de severidade e a posição dos defeitos presentes na linha férrea. O segundo passo, através do conhecimento prévio acerca dos trilhos (perfis, irregularidades, velocidade), há a elaboração de fatores (pesos) que vão influenciar diretamente na integridade da malha férrea. No terceiro passo, há a realização da análise de interdependência entre os defeitos e os fatores de influência (ambos obtidos nos dois passos iniciais). Já no quarto passo, um modelo de inferência *Fuzzy* é utilizado para elaborar regras acerca da integridade dos trilhos baseado nos fatores de influência. Por fim, no passo 5, depois da estimação das condições físicas e estruturais dos trilhos, há as respectivas avaliações e tomada de decisões acerca da manutenção. Vale ressaltar que, para a execução deste trabalho, os autores consideraram como dados defeituosos apenas os defeitos do tipo *squat*; entretanto, os autores afirmam que outros defeitos, como *head checks* e *corrugations*, por exemplo, também podem ser considerados com esta metodologia.

Já Cárdenas-Gallo et al. (2017) trazem como foco os defeitos de geometria que podem estar presentes em uma linha férrea e propõem a construção de um classificador para a previsão da degradação de trilhos, através da análise de defeitos do tipo *hump* e depressões. Como dados de entrada, foram utilizados sinais de amplitude do defeito, tempo que o defeito foi avaliado, severidade e carga em toneladas do vagão. Como dado de saída, a estratégia elaborada tem a capacidade de informar quanto tempo um determinado defeito levará para atingir um determinado nível de severidade.

Mohammadi et al. (2019) desenvolveram uma metodologia robusta, baseada em IA, de previsão de parâmetros de geometria da linha que mais contribuem para a ocorrência de defeitos ferroviários. O conjunto de dados de 100 milhas de geometria, coletadas entre 2012 e 2017, inclui 61 tipos de medições de geometria dos trilhos, defeitos nos trilhos e informações sobre o peso total das mercadorias e carros que passam em cada segmento em MGT (do inglês, *Million Gross Ton*).

Petrović et al. (2022) propoem um algoritmo para a detecção de objetos de infraestrutura ferroviária, como trilhos e sinais, com o objetivo de identificar os sinais relevantes ao longo da rota do trem. A metodologia combina técnicas tradicionais de visão computacional, incluindo a suavização Canny, com o modelo de inteligência artificial conhecido como "You Only Look Once"(YOLO), baseado em redes neurais convolucionais (CNNs). Essa abordagem permite que o sistema de IA reconheça e classifique de forma precisa os sinais específicos relacionados à via em que o trem está se movendo. Os resultados dos testes realizados demonstraram alta confiabilidade, com taxa de detecção de sinais relevantes de até 99,7%. Esse avanço tecnológico é especialmente relevante para cenários complexos, como estações e pátios, onde múltiplas vias estão presentes.

Acikgoz e Korkmaz (2023) desenvolveram um sistema automatizado de classificação de defeitos em linhas ferroviárias, visando garantir a segurança do transporte ferroviário por meio de monitoramento e inspeção periódicos. O modelo proposto, denominado MSRConvNet, emprega conexões de salto com blocos de aprendizado residual e convoluções em múltiplas escalas para extrair características relevantes dos trilhos. Essa abordagem permite a extração de mapas de características em diferentes escalas, melhorando a capacidade do modelo em identificar e classificar os diversos tipos de defeitos. Além disso, é aplicado um método de aumento de dados para equilibrar a distribuição das classes e mitigar o impacto de conjuntos de dados desbalanceados. Os resultados obtidos demonstram que o modelo proposto alcançou uma taxa de acurácia, precisão, sensibilidade, especificidade, pontuação F1 e coeficiente de correlação de Matthews de 99,83%, evidenciando seu desempenho superior em relação a modelos de aprendizado profundo de referência.

Os trabalhos de Cárdenas-Gallo et al. (2017) e Mohammadi et al. (2019) são trabalhos considerados relevantes para este projeto e que serão avaliados em detalhe, pois focam na influência da geometria dos perfis do par roda trilho na severidade do contato.

Finalmente, conforme exposto nesta revisão da literatura (Figura 3.29), a avaliação de um sistema roda-trilho usando uma abordagem em IA é um tema cada vez mais frequente. A Figura 3.32 resume os principais aspectos da literatura que serviram de direcionamento para a atualização da proposta deste projeto.

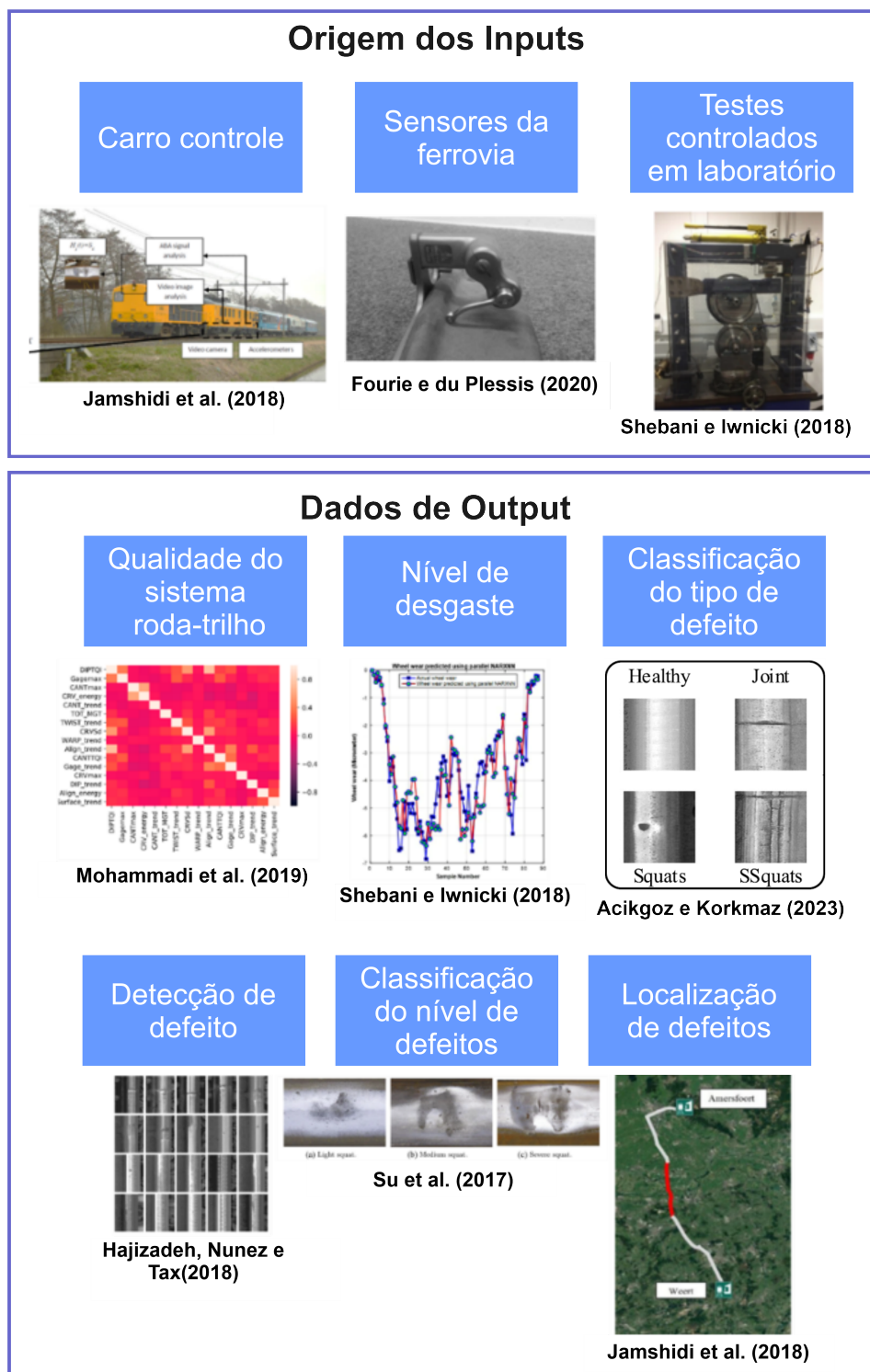


Figura 3.32: Resumo dos principais aspectos do estado da arte. (Autoria Própria)

---

## Metodologia

O desenvolvimento deste trabalho consistiu na elaboração de um banco de dados formado por resultados de simulações computacionais que fornecessem informações acerca do tipo de contato existente entre rodas e trilhos com o intuito de treinar um modelo de IA.

Tendo em vista que a métrica  $s$  proposta por Magel e Kalousek (2002) fornece uma classificação acerca do tipo de contato existente entre roda e trilho, optou-se então por se trabalhar com esse classificador para alcançar o objetivo deste trabalho. No entanto, para que a extração deste parâmetro fosse possível, foi necessário desenvolver um modelo virtual em MEF que permitisse o cálculo das distâncias entre as superfícies após o sistema atingir o estado de equilíbrio. A Figura 4.1 apresenta a metodologia geral deste trabalho.

Inicialmente, foi necessário o levantamento dos perfis de rodas e trilhos da Vale S.A. Esses dados foram inseridos no Modelo Virtual para realizar simulações numéricas e para compor o *dataset* que é utilizado para treinar o modelo de IA. De modo a aumentar o volume de simulações e resultados para alimentar o banco de dados, foi necessário automatizar o Modelo Virtual. Com o *dataset* bem consolidado, uma vez que houve a escolha do tipo de IA, o mesmo foi treinado e por fim validado. Mais detalhes acerca de cada etapa serão apresentados nas seções posteriores.

### 4.1 Banco de Dados

A metodologia macro de construção do banco de dados está ilustrada na Figura 4.2. Sua elaboração consistiu inicialmente em levantar dados de perfis de roda e trilho (Figura 4.3) e posterior pré-processamento dos trilhos, separando-os em grupos de acordo com semelhanças geométricas e alturas do boleto (Figura 4.4). Em seguida, as geometrias foram importadas para o *software* Abaqus CAE de modo a obter o parâmetro  $s$ . Feito isso, o banco de dados foi então elaborado através da construção de uma planilha que possui as combinações de rodas e trilhos e os respectivos tipos de contato entre eles. Os dados



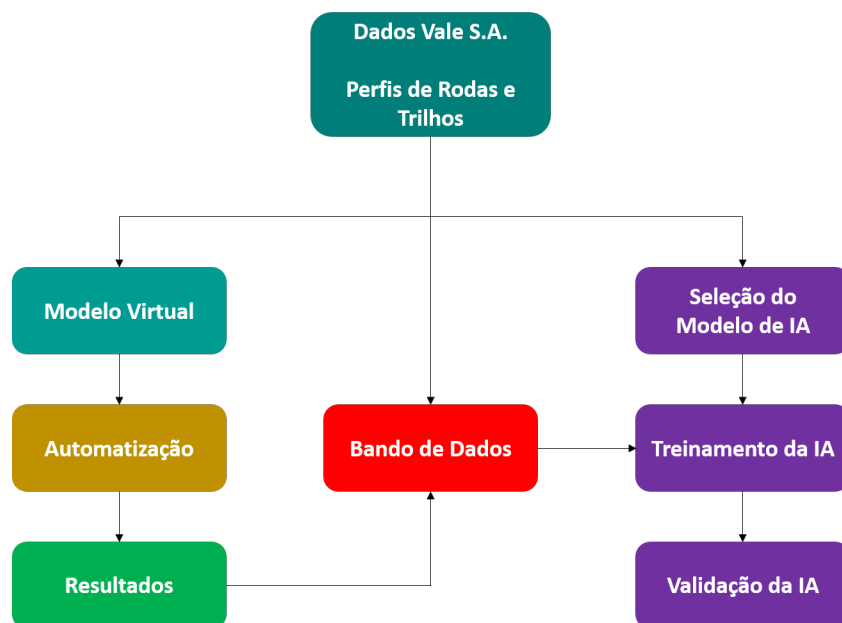


Figura 4.1: Metodologia geral para o desenvolvimento do modelo de IA. (Autoria Própria)

de perfis foram fornecidos pela Vale S.A. Visando fornecer uma compreensão mais ampla e detalhada da elaboração do banco de dados, nas próximas seções será apresentado, respectivamente, o modelo virtual utilizado para a extração do *output* de interesse e a rotina em *Python* desenvolvida para automatizar o processo de simulação e extração dos resultados do modelo.

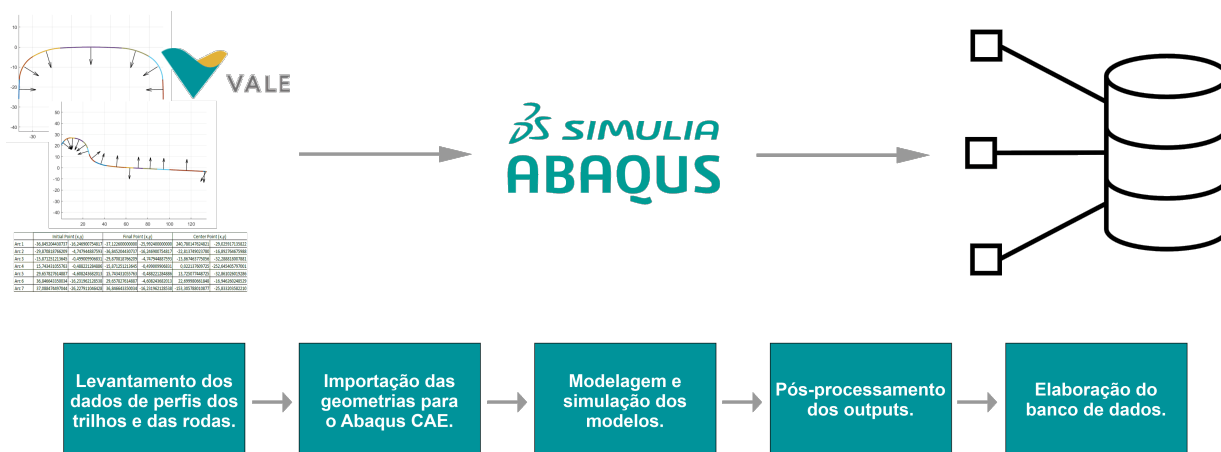


Figura 4.2: Metodologia macro para a construção do banco de dados. (Autoria Própria)

#### 4.1.1 Modelo Virtual

O modelo virtual utilizado para avaliar o contato entre perfis de roda e trilho (Figura 4.5) foi desenvolvido no *software* Abaqus e consiste na representação 2D de um rodeiro (metade de um truque) de um vagão em condições de tráfego tangente, cujo perfil de roda

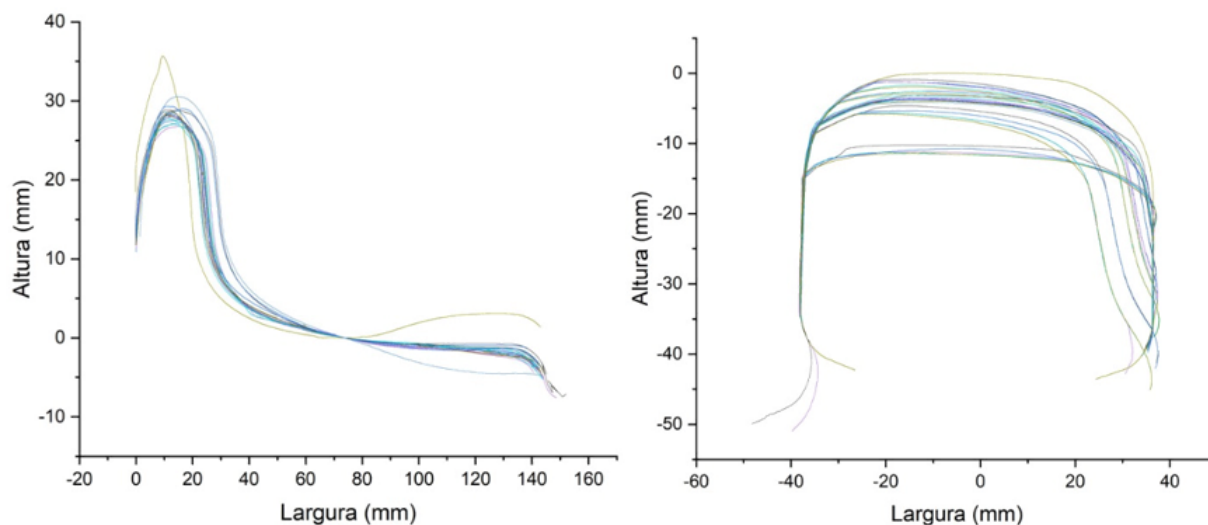


Figura 4.3: Perfis de rodas e trilhos fornecidos pela Vale S.A. (Autoria Própria)

utilizado é AAR Classe C e perfil de trilho é AREMA TR68, ambos novos (sem desgaste) e perfeitamente lisos. As propriedades mecânicas das geometrias dos materiais são apresentadas na Tabela 4.1. Vale ressaltar que o material foi considerado elastoplástico sem propriedades de encruamento (perfeitamente plástico). O detalhamento da metodologia utilizada para validar o respectivo modelo virtual via simulação 3D (Figura 4.6) pode ser vista em Machado et al. (2023).

Tabela 4.1: Propriedades mecânicas da roda AAR Class C e do trilho AREMA TR68.

	Tensão de Escoamento	Módulo de Young	Coefficiente de Poisson
AAR Class C	900 MPa	206 GPa	0,3
AREMA TR68	890 MPa	207 GPa	0,3

O modelo 3D utilizado para a validação do 2D consiste na representação de um rodeiro de um vagão em condições de tráfego tangente, dinâmica e com velocidade linear do trem constante de  $70 \text{ km/h}$  e carga de  $269 \text{ kN}$  (aproximadamente 27 toneladas) para o

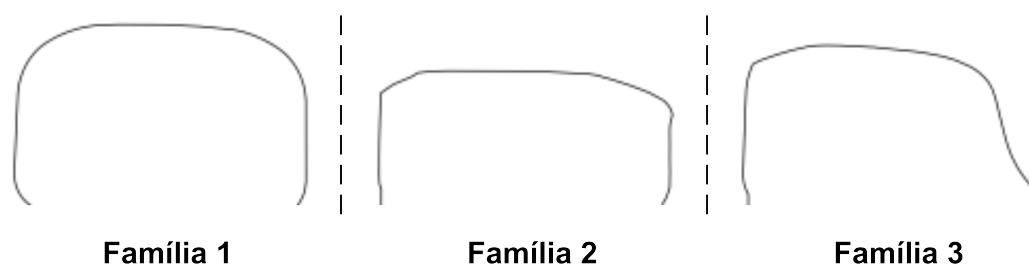


Figura 4.4: Exemplo da divisão dos perfis de trilho em famílias de acordo com geometria e altura do boleto. (Autoria Própria)



Figura 4.5: Modelo 2D do rodeiro. Extraído de Machado et al. (2023)

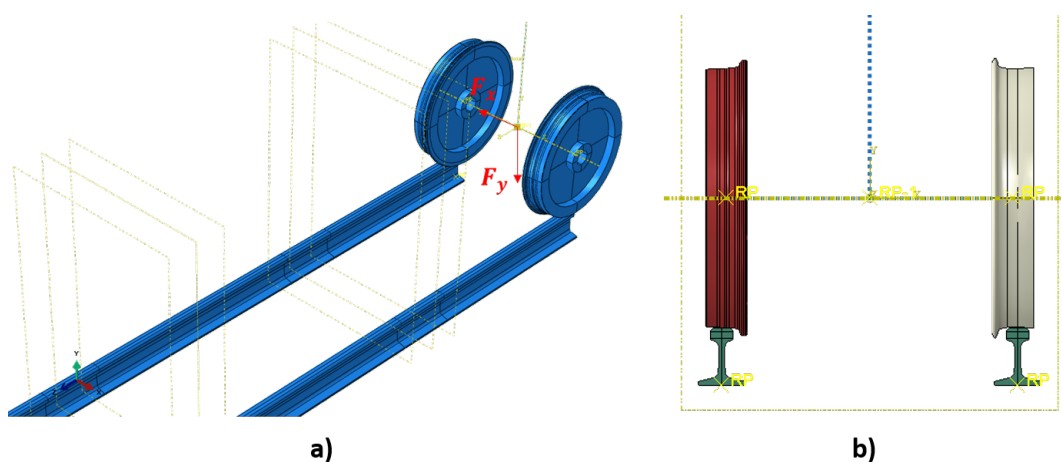


Figura 4.6: Modelo 3D do rodeiro: a) Vista isométrica e b) Vista frontal. Extraído de Machado et al. (2023)

rodeiro. Os trilhos possuem restrição de engaste com inclinação de  $0,025 \text{ rad}$  (devido o *cant*) e as rodas estão conectadas entre si no seus centros por um elemento de viga (rígido), que restringe o grau de liberdade delas, garantindo que possuam o mesmo movimento de rotação e translação. Além disso, atrito de  $0,3$  também foi considerado, haja vista as condições de escorregamento e rolamento das rodas. Os perfis utilizados, suas características e respectivas propriedades mecânicas são iguais as do utilizado no 2D.

Embora o modelo 3D se mostre completo e possa fornecer diversas respostas acerca da mecânica do contato roda-trilho, o alto custo computacional de simulação (12 horas) justifica a utilização de modelos mais simplificados (Figura 4.5), haja vista a necessidade de gerar dados para treinar uma IA. Sendo assim, tendo em vista a utilização do modelo 2D, maiores detalhes serão abordados a seguir.

Por se tratar de uma análise em Elementos Finitos, optou-se por simplificar a geometria das rodas e dos trilhos de forma a reduzir a área da malha, corroborando assim em menores

tempos de processamento, permitindo inclusive aumentar a discretização do problema sem custos consideráveis. Assim, nenhuma parte abaixo do boleto do trilho foi considerada e apenas uma pequena parte da seção da roda foi representada (Figura 4.7 e 4.8).

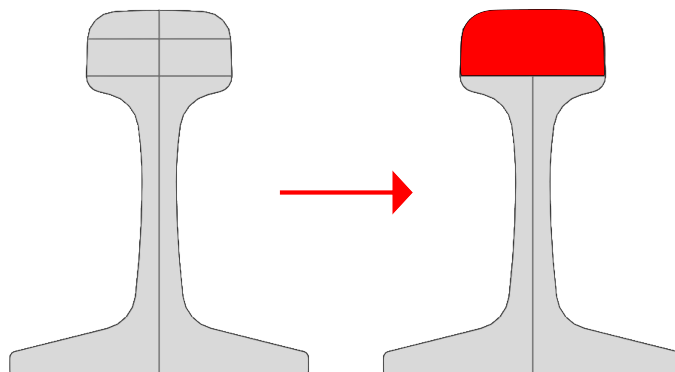


Figura 4.7: Simplificação da geometria do trilho para a seção marcada em vermelho. (Autoria Própria)

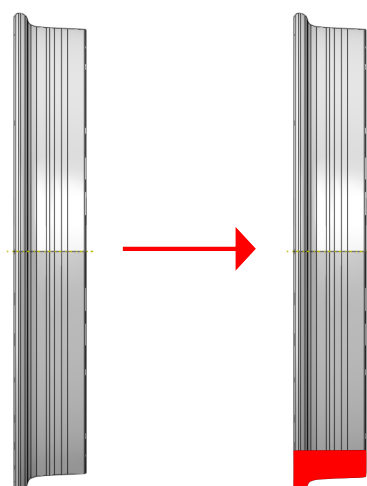


Figura 4.8: Simplificação da geometria da roda para a seção marcada em vermelho. (Autoria Própria)

A simulação quase estática assimétrica do modelo permitiu as seções das rodas transladarem no plano OXY e girarem no eixo OZ. O contato com os trilhos ocorreu após a aplicação da carga, que representa o peso do vagão, no centro do rodeiro. No modelo, os trilhos são fixados através da restrição do tipo engaste ao longo de sua base e com inclinação de  $0,025$  radianos de forma a representar o *cant*, enquanto as rodas são acopladas ao centro do rodeiro através de um elemento de viga (rígida), condicionando os graus de liberdade das duas rodas, obrigando-as ao mesmo movimento de translação. Essa configuração foi adotada para permitir que o centro de massa se mova e encontre o ponto onde a resultante das forças (externas e internas do sistema) seja nula, isto é, alcance o ponto de equilíbrio para as cargas aplicadas.

A carga aplicada na simulação foi de 3420 N. Este valor foi validado através de um estudo da distribuição da pressão de contato obtida com a simulação 3D. De acordo com Lima (2023), uma certa direção do caminho do contato roda-trilho pode ser selecionada e o modelo 2D quase-estático pode ser empregado para calcular o carregamento naquela direção. A partir do modelo tridimensional, a pressão máxima de contato ( $p_0^{3D}$ ) e a metade da largura do contato roda-trilho ( $a_{3D}$ ) para o ponto de contato, é possível chegar na equação do carregamento 2D equivalente ao do modelo tridimensional:

$$F_{2D_{roda}} = 1,569 \cdot a_{3D} \cdot p_0^{3D} \quad (4.1)$$

Através da Figura 4.9 é possível averiguar que a pressão de contato ( $p_0^{3D}$ ) obtida com a simulação 3D (considerando todas as condições descritas anteriormente) foi de 1,09 *GPa* e o comprimento ( $2 \cdot a_{3D}$ ) da área de contato (região de pressão máxima) foi de 2 *mm*. Logo, foi possível calcular a força equivalente para o rodeiro do modelo bidimensional ( $F_{2D_{rodeiro}}$ ) aplicando os valores de pressão e largura da área de contato na Equação 5.1:

$$F_{2D_{roda}} = 1,569 \cdot 1 \cdot 1,09 \cdot 10^3 = 1710N \quad (4.2)$$

$$F_{2D_{rodeiro}} = 2 \cdot F_{2D_{roda}} = 2 \cdot 1710 = 3420N \quad (4.3)$$

Em relação à configuração de contato entre as superfícies das geometrias, foi adotado um contato seco e sem lubrificação, onde também foi desconsiderado o atrito, uma vez que o mesmo não contribui significativamente no sistema devido à condição de rolamento puro no sentido longitudinal.

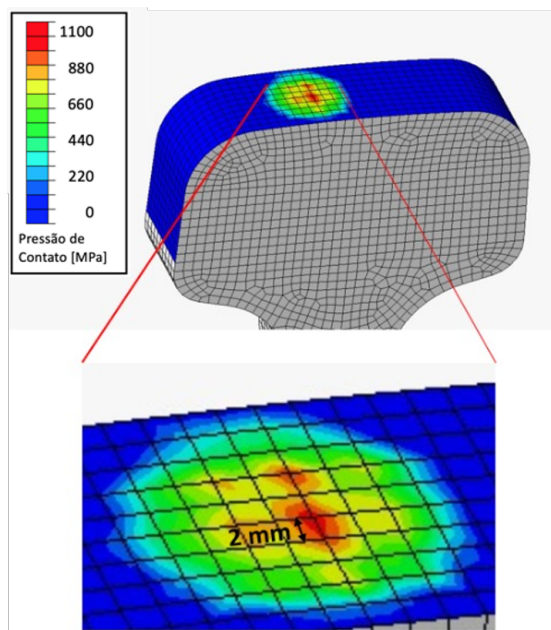


Figura 4.9: Distribuição da pressão de contato em simulação do modelo 3D. Extraído de Lima (2023)

Para o desenho das geometrias das rodas e dos trilhos foram utilizadas curvas *splines*, o que proporcionou uma discretização contínua por meio de uma malha com elementos quadrangulares e triangulares em estado plano de deformação. A discretização da superfície foi feita com arestas de 1 mm de comprimento, apresentando melhor relação custo/benefício entre a representação dos detalhes do perfil geométrico e o tempo necessário para a simulação, além de ter convergido para uma solução estável. Foram utilizados aproximadamente 6.400 elementos em cada roda e 2.800 elementos em cada trilho, totalizando 18.400 elementos. Os elementos utilizados foram do tipo CPE4 (*quad-dominated*, combinação de elementos quadriláteros e triangulares), indicando a utilização de 4 pontos de integração em estado plano de deformação. Vale ressaltar que os critérios de parada para a simulação foram dois: tempo de simulação e/ou alcance do limite de escoamento do material.

O desenvolvimento deste modelo apresenta vantagens como: representatividade significativa do modelo real; obtenção do comportamento mecânico através de tensões e deformações; permitir a análise da influência de variações de perfis geométricos na resposta mecânica e tribológica do contato roda-trilho; obtenção do ponto de contato em equilíbrio e menor tempo de simulação quando comparado aos modelos 3D.

Deste modelo é possível extrair diversos resultados (Figura 4.10) que fornecem informações sobre a mecânica do contato roda-trilho, tais como: tensão de von Mises e outros invariantes, deformação plástica equivalente (se houver), pressão de contato ao longo da superfície, área de contato, ponto de contato e os valores da distância máxima em ponto

único  $s$  (calculada através dos deslocamentos das geometrias após atingir o estado de equilíbrio) que foi discutido anteriormente.

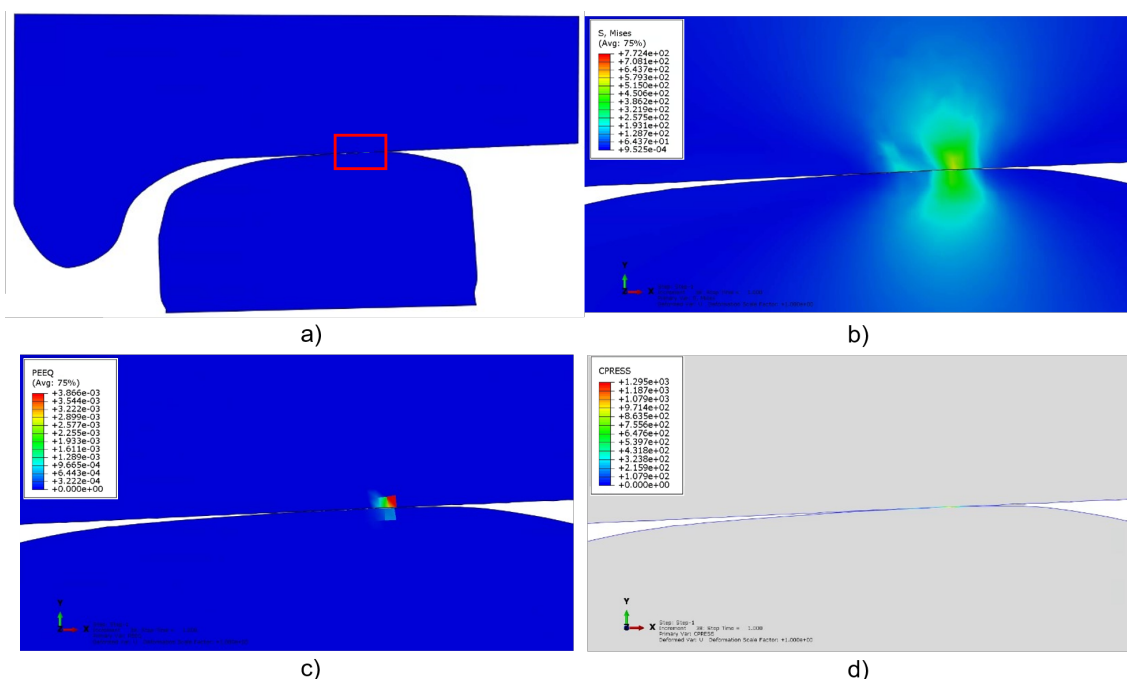


Figura 4.10: Resultados obtidos através da simulação do contato entre a roda e o trilho a), onde: b) é o gradiente de tensão de von Mises, c) a deformação plástica equivalente e d) a pressão de contato na região destacada pelo quadro vermelho em a). (Autoria Própria)

Embora a execução do modelo 2D tenha um baixo custo computacional (tempo de simulação de 30 segundos), o mesmo possui um alto consumo de tempo no que concerne à montagem do modelo, devido ao processo manual envolvido nesta etapa. Contudo, isso representa um gargalo para o desenvolvimento do banco de dados, cujo tamanho precisa ser grande o suficiente para treinar de maneira apropriada o modelo de IA proposto por este trabalho.

Dessa forma, foi desenvolvida uma rotina que garantiu a automatização da construção e simulação dos modelos no software de modo a dispensar todo o processo de introdução manual de dados. Detalhes sobre a automatização serão abordados na próxima seção.

#### 4.1.2 Automatização

Para o desenvolvimento da rotina, foram analisadas diversas estratégias para que a automatização pudesse ser alcançada. Dentre as estratégias, optou-se por desenvolver um *script* com linguagem de programação em *Python*.

O Abaqus possui um compilador capaz de executar códigos na linguagem de programação *Python*, onde as ferramentas do programa podem ser acessadas e executadas

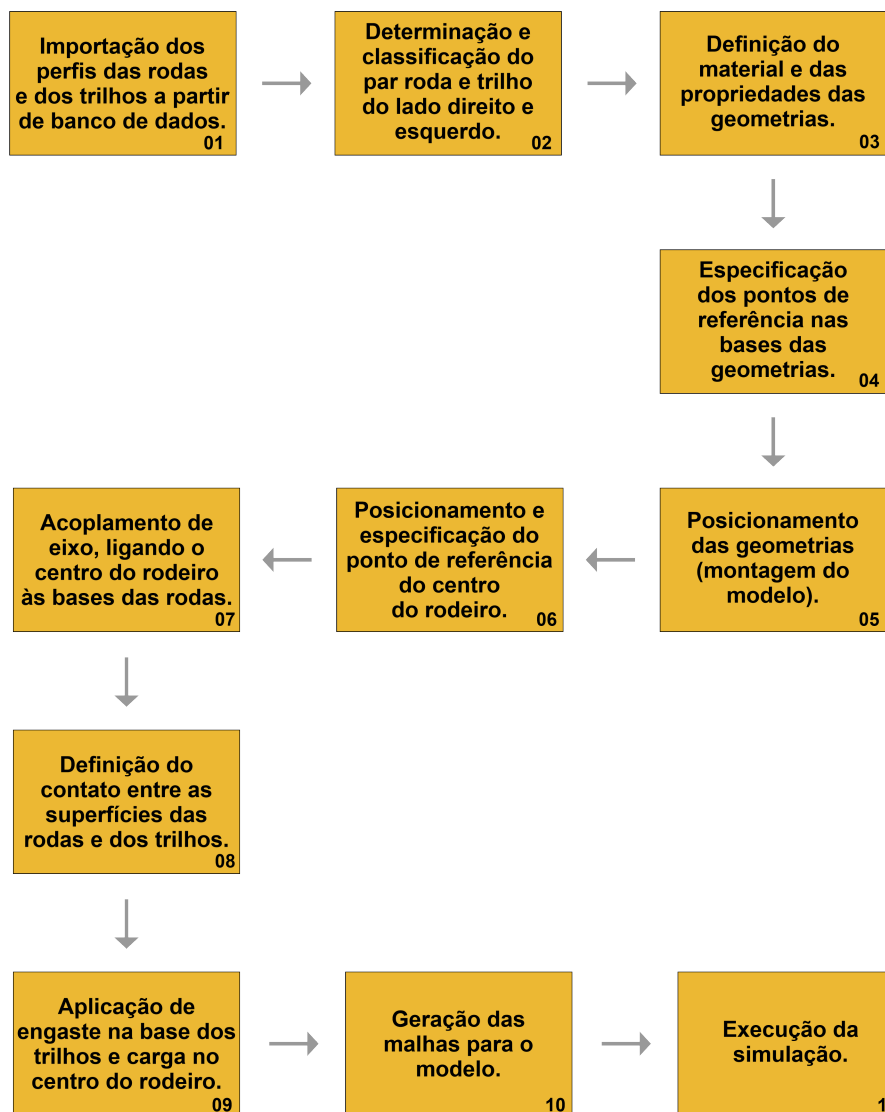


Figura 4.11: Sequência de comandos responsável por gerar o modelo na automatização. (Autoria Própria)

de forma precisa através das linhas de comando presentes no *script*. Assim, por meio dele, foi possível importar as geometrias para o Abaqus CAE, seguidas de sua posterior modelagem, montagem do modelo e simulação.

Todos os passos que são apresentados na Figura 4.11 são inseridos dentro de um loop, onde cada loop representa a geração e processamento de um modelo (dois pares roda-trilho).

Inicialmente, os arquivos das geometrias das rodas e dos trilhos, que contém os pontos do Miniprof (em formato de arquivo de texto), são aleatoriamente selecionados e enviados para o Abaqus CAE. Uma vez importados, os pontos são então plotados dentro do *software* e, através deles, são geradas as geometrias (Bloco 1 da Figura 4.11).



Após isso, é feita a combinação aleatória dos conjuntos roda-trilho e a posterior definição de qual conjunto será o lado direito e o lado esquerdo (Bloco 2 da Figura 4.11). Em seguida, os materiais e as propriedades mecânicas das geometrias são definidas, tais como: Módulo de Elasticidade ( $E$ ), Coeficiente de Poisson ( $\nu$ ) e encruamento (Bloco 3 da Figura 4.11).

Tendo finalizado essas etapas iniciais, o *script* entra então no processo de montagem do modelo. Inicialmente, são definidos pontos de referências nos centros das bases das geometrias, bem como nas suas extremidades (Bloco 4 da Figura 4.11). Esses pontos de referências nada mais são do que artifícios utilizados para facilitarem diversos processos, tais como: translação e rotação de geometrias, aplicação de condições de contorno, cargas, etc.

Posteriormente, há a translação e posicionamento das geometrias de acordo com a norma de bitola (Bloco 5 da Figura 4.11). Neste caso, as distâncias são configuradas a partir dos pontos de referências das extremidades, de modo a ajustar uma distância de 928 mm entre as rodas e, os trilhos, à uma distância de 1000 mm. Ajusta-se também uma distância vertical (que pode variar a depender do nível de desgaste dos perfis) entre a roda e o trilho de cada conjunto (direto e esquerdo), garantindo que não haja penetração (intersecção) entre as geometrias após a montagem, o que geraria problemas durante a simulação.

Realizado o posicionamento das geometrias, é então inserido no centro do modelo, alinhado com as bases das rodas, um ponto e seu respectivo ponto de referência, de modo a representar o centro do rodeiro (Bloco 6 da Figura 4.11). Em seguida, são aplicadas *wires* (linhas), conectando o centro do rodeiro aos pontos de referências situados no centro das bases das rodas (Bloco 7 da Figura 4.11), representando o eixo que conecta as rodas.

Uma vez que os passos de montagem foram realizados, são especificados quais superfícies entrarão em contato durante a simulação (Bloco 8 da Figura 4.11). Em seguida, são inseridas as condições de contorno (Bloco 9 da Figura 4.11), onde é aplicada à base de cada trilho uma restrição do tipo engaste (restrição dos 6 graus de liberdade), bem como a carga no ponto de referência do centro do rodeiro, caracterizando a força peso do vagão, suportada por um rodeiro.

Tendo concluído todas essas etapas, o *script* então inicia o processo de construção das malhas para cada geometria (Bloco 10 da Figura 4.11). Neste passo, configurações como: tipo de malha, tamanho de malha e fator de desvio, são especificadas dentro do código.

Por fim, o *script* executa a simulação do modelo, isto é, a análise em elementos finitos. Durante a análise, são gerados arquivos de saída com resultados de interesse. Esses arquivos são nomeados automaticamente de acordo com a ordem de simulação (Bloco 11

da Figura 4.11). Toda a sequência descrita referente à Figura 4.11 dura aproximadamente 40 segundos.

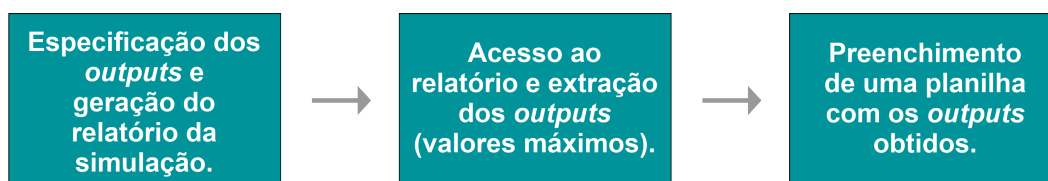


Figura 4.12: Sequência de comandos no *script*, responsável por obter os *outputs* das simulações. (Autoria Própria)

A extração de resultados da simulação tais como: von Mises, pressão de contato, deformação plástica, pontos de contato e as distâncias  $s$ , também é alcançada utilizando a rotina em *python* (Figura 4.12). Para isso, foi necessário exportar relatórios que contivessem os respectivos *outputs* e, em seguida, explorá-los via *script*. Vale ressaltar que esse processo também é automatizado, sendo capaz de gerar os relatórios de todas as simulações realizadas e extrair os seus respectivos resultados.

Vale salientar que o parâmetro  $s$  não é um *output* padrão do Abaqus, isto é, para obter esse valor, foi necessário elaborar uma rotina que o calculava a partir dos deslocamentos das geometrias dados pelo *software* em associação às coordenadas pertencentes às superfícies do trilho e da roda colineares ao centroide do trilho (Figura 4.13)

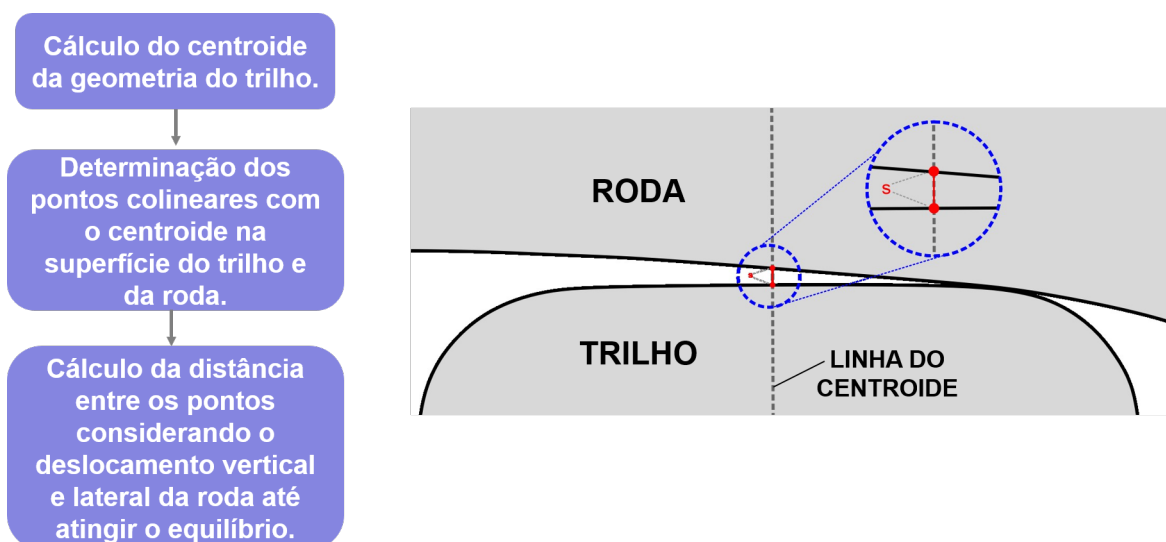


Figura 4.13: Sequência de comandos no *script*, responsável por extrair os valores de  $s$  das simulações. (Autoria Própria)

Todos os processos descritos acima foram capazes de construir uma planilha, cuja estrutura, permite visualizar as combinações de pares roda-trilho e as suas respectivas respostas oriundas das simulações: tensão de von Mises, deformação plástica, pressão de contato máxima, pontos de contato e a distância  $s$ . Embora o *script* seja capaz de fornecer



Figura 4.14: Exemplo de dados de imagem dos perfis da roda e do trilho.

todos esses resultados, o único *output* usado para construir o banco de dados foi o *s*, uma vez que ele é o classificador que informa o tipo de contato existente entre roda e trilho.

Tendo em vista que o número de perfis de roda e de trilhos adquiridos foram 24 e 41 respectivamente, através da automatização, foi possível gerar mais de 10.000 combinações de perfis diferentes, isto é, foi possível desenvolver um *dataset* com mais de 10.000 dados.

Todos os *scripts* aqui mencionados para a elaboração do banco de dados podem ser acessados via *GitHub* em Lopes (2023).

## 4.2 Modelo de Inteligência Artificial

No que concerne à seleção de uma abordagem de *machine learning* que seja capaz de avaliar o contato roda-trilho utilizando a análise das combinações de diversas geometrias de rodas e trilhos, inicialmente foi necessário realizar a definição da natureza dos dados, onde:

- Dados de entrada: Imagens de pares de perfis de roda e trilho, plotados através dos pontos dos perfis extraídos em medições de campo (Figura 4.14). As imagens possuem resolução de 148 x 96 pixels e apresentam apenas 1 canal de cores (tons de cinza).
- Dados de saída: Classificação sobre a conformidade do contato em 3 categorias: Contato Conforme, Contato Não Conforme, Contato Quase Conforme.

Uma vez que os dados de entrada são imagens, optou-se por utilizar as Redes Neurais Convolucionais (CNN), tendo em vista sua capacidade de processá-las. Além disso, a rede é do tipo multiclassificação, uma vez que os dados de saída são classificações acerca do contato.

Vale salientar que a rede neural desenvolvida tem como função avaliar o contato roda-trilho referente a um rodeiro (tal como no modelo virtual apresentado), ou seja, a IA recebe 4 dados de imagens, sendo duas rodas e dois trilhos e retorna como resposta 2 dados de saída (classificação do contato para o lado esquerdo e direito). Essa análise do

conjunto do rodeiro é necessária, dada a influência que o par roda-trilho de um lado pode influenciar no contato do par do outro lado e vice-versa.

Todos os códigos de desenvolvimento das etapas que serão descritas adiante (tratamento das imagens, elaboração da rede neural e a otimização dela) foram programados em linguagem de programação *python* na plataforma *Google Colaboratory*, com a utilização do *Tensorflow* como *framework* e o *Keras* como API, auxiliando portanto na elaboração do modelo de IA de maneira mais fácil e intuitiva.

### 4.2.1 Tratamento das Imagens

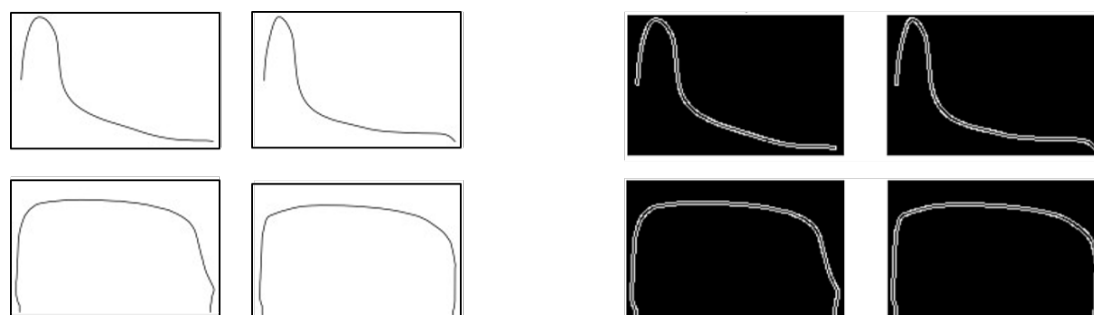
Através da análise dos *inputs* apresentados na Figura 4.14 é possível notar que os dados tratam-se de imagens de contorno. Tendo em vista que a rede neural convolucional deverá classificar o tipo de contato através da análise dos perfis, se faz necessária a aplicação de técnicas que auxiliem o modelo a detectar os contornos através do destaque dos perfis, auxiliando também no descarte de informações não relevantes (espaços em branco).

Uma técnica clássica de detecção de bordas que é bastante utilizada para tal é o método de Canny. Segundo Ding e Goshtasby (2001), uma implementação típica desta técnica segue as seguintes etapas:

- Aplicação de filtro Gaussiano para suavização através da aplicação de pesos a cada pixel da imagem, reduzindo o ruído.
- Determinação da magnitude do gradiente e a direção do gradiente em cada pixel através da aplicação do operador de Sobel.
- Avaliação da magnitude dos gradientes dos pixels. Se o gradiente de um pixel for maior do que a de seus vizinhos na direção do gradiente, o pixel será marcado como borda. Caso contrário, será rotulado como plano de fundo.
- Remoção das bordas fracas por meio de limiarização por histerese. Nessa etapa, são definidos dois limites: um mínimo e um máximo. Os valores de gradientes que estiverem abaixo dos limite mínimo são considerados não-bordas e conseqüentemente são eliminados. Valores acima do limite máximo são rotulados de bordas fortes. Já os gradientes que estão entre os limiares são considerados bordas fracas, podendo elas serem descartadas caso não estejam conectadas diretamente com bordas fortes.

Por fim, o resultado final após a aplicação da suavização Canny é uma imagem binária com pixels que são classificados como bordas fortes ou não-bordas.

Esta técnica foi aplicada nos dados de imagem do banco de dados (Figura 4.15) com o intuito de destacar os contornos dos perfis e assim facilitar a aprendizagem do modelo.



(a) Dados de imagem original.

(b) Dados de imagem com suavização Canny.

Figura 4.15: Exemplo de dados de imagens originais a) e com suavização Canny b).

Os valores determinados para a limiarização por histerese foram de 100 e 200 para os limites mínimo e máximo respectivamente.

#### 4.2.2 Arquitetura e seus Hiperparâmetros

A arquitetura do modelo possui duas etapas: a primeira consiste nas etapas convolucionais, que é responsável por processar os dados da imagem, e a segunda compreende a sequência de camadas densas e a camada de saída, responsáveis por calcular as correlações existentes entre os dados de imagem e as classificações do contato. A arquitetura completa da CNN pode ser vista na Figura 4.16.

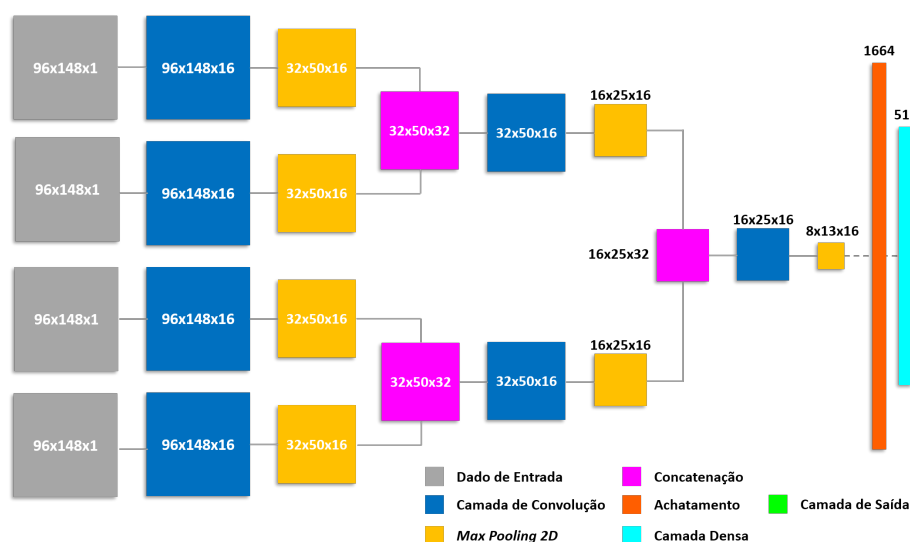


Figura 4.16: Arquitetura da Rede Neural Convolucional. (Autoria Própria)

No estágio convolucional, há três etapas de processamento de imagem, onde, na primeira etapa, cada camada convolucional extrai características de cada imagem do perfil da roda e do trilho (uma camada por imagem). Depois disso, os parâmetros calculados pela roda e trilho de cada lado são concatenados e então processados por uma segunda

etapa da camada convolucional. Essa etapa se faz necessária, uma vez que a partir dela, a IA obterá características resultantes da combinação do contato entre roda e trilho. Em seguida, os resultados de cada lado são concatenados e depois processados por uma camada convolucional (terceiro passo), obtendo-se características que correspondem à combinação das geometrias no modelo como um todo. Esta última etapa é necessária face à interdependência existente, em termos de contato, entre os perfis do conjunto direito e esquerdo. Por fim, a matriz de resultados é achatada a um vetor, podendo ser processada pelas camadas densas no segundo estágio.

Dezesseis filtros foram usados em cada camada de convolução, com tamanho de  $3 \times 3$  na primeira etapa e  $1 \times 1$  na segunda e terceira. O número de filtros e o seu tamanho foram determinados com base na rede neural convolucional LeNet-5, cuja arquitetura possui poucos filtros e o tamanho do *kernel* decresce na terceira camada de convolução (YEH et al., 2023).

A LeNet-5 foi utilizada como referência para o desenvolvimento da CNN deste trabalho, devido à semelhança das propriedades dos seus dados de entrada.

Uma técnica de *pooling* também foi utilizada para reduzir a quantidade de informações obtidas pelos filtros. O *2D Max Pooling* foi selecionado haja vista sua capacidade de destacar características mais relevantes em problemas de detecção de contorno em imagens.

A função de ativação utilizada foi a *Leaky ReLU* devido à sua ampla aplicação e por oferecer vantagens como: baixo custo computacional por não haver cálculos complexos, convergência rápida e evitar o problema do "*dying ReLU*" (inativação não saudável de perceptrons).

A segunda etapa consiste em uma sequência de sete camadas densas que são capazes de calcular a correlação entre as imagens e o tipo de classificação de contato. O número de perceptrons segue a sequência descendente de  $2^n$  através das camadas. Os hiperparâmetros da rede neural podem ser vistos nas Tabelas 4.2 e 4.3. A função de ativação utilizada nas camadas ocultas também é a *Leaky ReLU*. Em seguida, as camadas de saídas, que consistem em duas camadas diferentes com três perceptrons em cada uma, classificam o contato com base nos parâmetros calculados pelas camadas ocultas. Nesta etapa, cada camada é responsável por classificar um lado (esquerdo e direito) e cada perceptron calcula a probabilidade de um contato pertencer a uma classe. Isso é possível devido à utilização da função *Softmax* nas camadas de saída.

Alguns dos hiperparâmetros da rede neural foram otimizados através de uma técnica de otimização clássica chamada *Random Search Fit*, cuja metodologia de aplicação será abordada posteriormente.

Hiperparâmetros das Camadas de Convolução					
Primeira Camada		Segunda Camada		Terceira Camada	
Nº de Filtros	16	Nº de Filtros	16	Nº de Filtros	16
Dimensão	3x3	Dimensão	1x1	Dimensão	1x1
<i>Pooling</i>	Max 2D	<i>Pooling</i>	Max 2D	<i>Pooling</i>	Max 2D
Dimensão	3x3	Dimensão	2x2	Dimensão	2x2
Função	<i>LeakyReLU</i>	Função	<i>LeakyReLU</i>	Função	<i>LeakyReLU</i>

Tabela 4.2: Tabela com hiperparâmetros das camadas de convolução

Hiperparâmetros das Camadas Densas e de Saída					
Primeira Camada Densa		Segunda Camada Densa		Terceira Camada Densa	
Perceptrons	512	Perceptrons	256	Perceptrons	128
Função	<i>LeakyReLU</i>	Função	<i>LeakyReLU</i>	Função	<i>LeakyReLU</i>
Quarta Camada Densa		Quinta Camada Densa		Sexta Camada Densa	
Perceptrons	64	Perceptrons	32	Perceptrons	16
Função	<i>LeakyReLU</i>	Função	<i>LeakyReLU</i>	Função	<i>LeakyReLU</i>
Sétima Camada Densa		Camada de Saída - 1		Camada de Saída - 2	
Perceptrons	8	Perceptrons	3	Perceptrons	3
Função	<i>LeakyReLU</i>	Função	<i>Softmax</i>	Função	<i>Softmax</i>

Tabela 4.3: Tabela com hiperparâmetros das camadas densas e das camadas de saída

### 4.2.3 Hiperparâmetros de Treinamento

Os hiperparâmetros de treinamento selecionados podem ser vistos na Tabela 4.4. A função perda escolhida para o acompanhamento e avaliação do aprendizado da rede foi a *Categorical Crossentropy*, uma vez que os rótulos foram inseridos na IA na forma de *one-hot encoding* (técnica que converte valores escalares em vetores binários) já que as variáveis categóricas não possuem relação ordinal. Assim, as classes terão os seguintes valores: Quase Conforme - [1,0,0], Conforme - [0,1,0] e Não Conforme - [0,0,1]. A *Accuracy* foi a métrica escolhida, haja vista a possibilidade de averiguar a capacidade de previsão assertiva do modelo.

Tabela 4.4: Hiperparâmetros de treinamento.

Epocas	200
Função Custo	<i>Categorical Crossentropy</i>
Métrica	<i>Accuracy</i>
Otimizador	Adam
Treino-Validação	80% - 20%
Regularização	<i>Dropout,L2</i>

O otimizador utilizado foi o Adam (*Adaptive Moment Estimation*) dado que: é mais eficiente computacionalmente do que outros métodos; ajusta automaticamente a taxa de

aprendizado durante o treinamento; adequado para problemas com grande quantidade de dados e/ou parâmetros; combina estimativas de gradientes de primeira ordem (como o AdaGrad) e momentos de segunda ordem (como o RMSProp) para a atualização da taxa de aprendizado (KINGMA; BA, 2017).

Para evitar problemas com *overfitting*, foram utilizadas a técnica de regularização L2 (*Ridge Regression*) e a de *Dropout*. A primeira consiste na aplicação de uma penalização equivalente ao valor ao quadrado da magnitude dos pesos na função perda, impedindo que a IA calcule pesos que a façam se ajustar demais ao banco de dados de treinamento. Quando a regularização é aplicada, o cálculo da função perda é então calculada da seguinte forma:

$$Loss = \sum_{i=1}^n [y_i - (w * x_i + b)]^2 + \alpha \sum_{j=1}^m w_j^2 \quad (4.4)$$

Onde  $n$  é o número total de dados de treinamento,  $m$  é o número de features (ou variáveis),  $y_i$  é o valor de saída esperado para o caso  $i$ ,  $x_i$  é o  $i$  – *esimo* vetor de atributos de entrada,  $w$  é o vetor de pesos a serem aprendidos durante o treinamento,  $b$  é o bias e  $\alpha$  é o hiperparâmetro de regularização.

A segunda, por outro lado, consiste na inativação de alguns perceptrons durante o treinamento para que alguns pesos não sejam melhorados, impedindo assim que a rede neural se ajuste aos ruídos dos dados de treinamento (Figura 4.17).

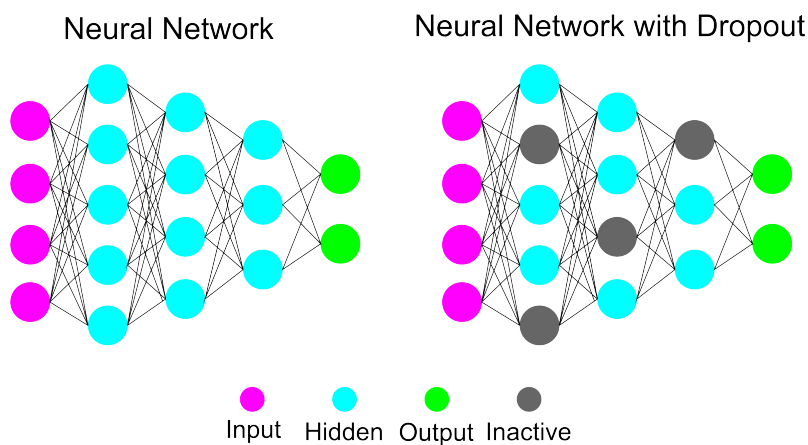


Figura 4.17: Inativação de perceptrons através da técnica *Dropout*. (Autoria Própria)

Foi definida também a divisão do banco de dados em 80% para treinamento e 20% para validação.

#### 4.2.4 Otimização dos Hiperparâmetros

Visando alcançar o melhor modelo possível para a arquitetura desenvolvida, bem como selecionar de maneira apropriada alguns hiperparâmetros de treinamento, foi aplicada uma



técnica clássica de otimização chamada *Random Search Fit* para encontrar a combinação ótima entre alguns hiperparâmetros da arquitetura e de treinamento.

Segundo Liashchynskiy e Liashchynskiy (2019) a técnica *Random Search Fit* (RSF) é um método de busca aleatória usado para encontrar soluções ótimas ou próximas do ótimo em problemas de otimização. É uma abordagem que não requer informações específicas sobre a função objetivo ou restrições do problema, tornando-a flexível e aplicável a uma ampla gama de problemas.

A principal finalidade do RSF é explorar o espaço de busca de soluções de forma aleatória, avaliando a qualidade de cada solução encontrada. A estratégia consiste em gerar soluções aleatórias dentro do espaço de busca e avaliar seu desempenho. Com base nessa avaliação, o algoritmo seleciona as soluções que apresentam melhor desempenho e continua explorando o espaço em torno delas até atingir um determinado critério de parada.

No que concerne à arquitetura da IA e aos hiperparâmetros de treinamento, o RSF foi aplicado visando a otimização através da testagem de valores pré-estabelecidos (Tabela 4.5) para os filtros das camadas de convolução; para os perceptrons da quinta, sexta e sétima camadas densas; para o coeficiente de penalização da técnica de regularização L2; para a porcentagem de inativação do *Dropout* e para o *Batch Size*.

Tabela 4.5: Hiperparâmetros otimizados pelo método RSF.

Nº de Filtros da Primeira Camada de Convolução	8,16,32
Nº de Filtros da Segunda Camada de Convolução	8,16,32
Nº de Filtros da Terceira Camada de Convolução	8,16,32
Nº de Perceptrons da Quinta Camada Densa	16,32,64
Nº de Perceptrons da Sexta Camada Densa	8,16,24
Nº de Perceptrons da Sétima Camada Densa	4,8,12
Batch Size	16,32,64,128
Regularização L2	0,0001/0,001/0,01/0,1
Dropout	0,1/0,15/0,2/0,25/0,3/0,35/0,4/0,45

Na aplicação do método, foram utilizados 1.700 dados aleatórios devidamente balanceados (Tabela 4.6) de modo a reduzir o custo computacional e o tempo de treinamento. Como critério de parada, foi utilizada a técnica *Early Stop* para 10 épocas, bem como o valor da *Accuracy* para 0,85 para os dados de validação. Os demais hiperparâmetros de treinamento aplicados correspondem aos apresentados na Tabela 4.4.

Tabela 4.6: Distribuição do tipo de contato no banco de dados para 1700 casos.

Tipo de Contato	Conjunto Esquerdo	Conjunto Direito
Quase Conforme	33%	35%
Conforme	35%	33%
Não Conforme	32%	32%

Durante o processo de otimização, foram treinados cerca de 1.296 modelos com diferentes configurações. A Tabela 4.7 traz as 10 melhores combinações de hiperparâmetros encontradas pela RSF. Ao analisá-la, é possível concluir que a melhor configuração para a rede neural consiste em: 16 filtros para as três camadas de convolução; 32, 16 e 8 perceptrons para a quinta, sexta e sétima camadas densas respectivamente; 0,1 como porcentagem de inativação *Dropout*; 0,01 como regularizador L2 e 128 como *Batch Size*.

Configuração de Hiperparâmetros Otimizados					
Acurácia Esq.	Acurácia Dir.	Filtros	Dropout	Perceptrons (5-6-7)	L2
80,01 %	81,79 %	8	0,15	48-24-12	0,001
75,53 %	72,25 %	8	0,2	48-24-12	0,01
74,43 %	73,25 %	16	0,25	32-16-8	0,01
79,34 %	77,34 %	8	0,15	48-24-12	0,01
77,46 %	77,46 %	8	0,2	48-24-12	0,001
72,05 %	72,23 %	8	0,25	16-8-4	0,01
82,27 %	83,57 %	16	0,1	32-16-8	0,01
78,21 %	78,15 %	16	0,2	32-16-8	0,01
77,01 %	76,79 %	16	0,2	48-24-12	0,01
80,14 %	80,21 %	32	0,1	32-16-8	0,01

Tabela 4.7: As dez melhores configurações de hiperparâmetros alcançadas via RSF.

### 4.3 Treinamento da Rede Neural

Uma vez tendo determinada a melhor arquitetura e os demais hiperparâmetros, o seu treinamento então fora realizado com 10.750 dados, cuja distribuição está apresentada na Tabela 4.8. Como a divisão de treino e validação é de 80% para treino e 20% para validação, 8.600 dados serão utilizados para treinar o modelo e 2.150 serão usados para validação.

Tabela 4.8: Distribuição do tipo de contato no banco de dados para 10.750 casos.

Tipo de Contato	Conjunto Esquerdo	Conjunto Direito
Quase Conforme	39%	40%
Conforme	38%	38%
Não Conforme	23%	22%

As Figuras 4.18 e 4.19 ilustram os processos de treinamento da rede neural. É importante notar que foram treinados dois modelos de IA que se distinguem na etapa de inserção das imagens. O primeiro, Figura 4.18.a), foi treinado utilizando as imagens padrões obtidas pela plotagem dos pontos dos perfis. Já no segundo (Figura 4.18.b)), optou-se por rotacionar as imagens das rodas em  $180^\circ$  com o intuito de verificar se o direcionamento dos perfis de acordo com a realidade proporcionaria um aprendizado melhor para a rede. A Figura 4.19 descreve desde a etapa de inserção das imagens pré-processadas no modelo até a classificação final do contato para cada lado do conjunto (esquerdo e direito).

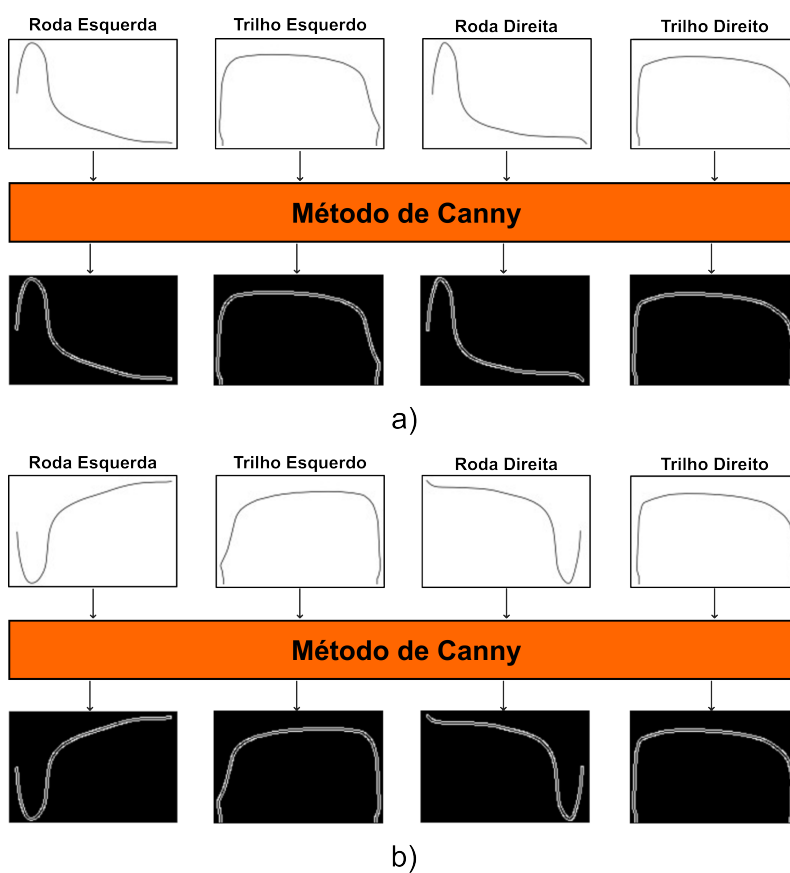


Figura 4.18: Etapa de pré-processamento das imagens para inserção no modelo de IA, onde: a) refere-se a um exemplo de imagens que foram utilizadas no treinamento do primeiro modelo e b) para o treinamento do segundo modelo. (Autoria Própria)

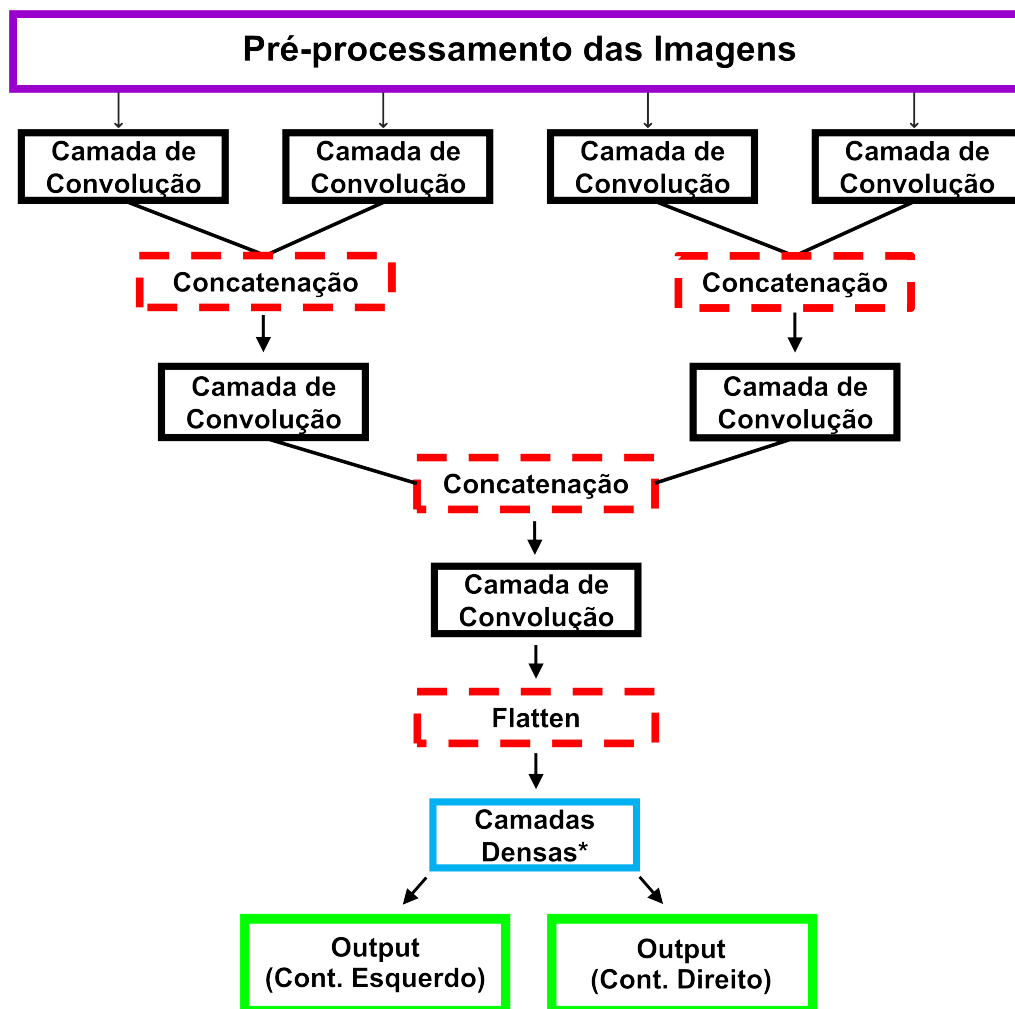


Figura 4.19: Treinamento dos modelos de IA. (Autoria Própria)

## Resultados e Discussão

A Tabela 5.1 apresenta uma amostra do banco de dados obtido através do processo de automatização do modelo virtual descrito neste trabalho. Os rótulos das rodas e dos trilhos foram inseridos com números de modo a garantir o sigilo da caracterização do projeto dos perfis pertencentes à Vale S.A.

Banco de Dados					
Roda Esq.	Trilho Esq.	Roda Dir.	Trilho Dir.	s - Esquerdo	s - Direito
Roda 18	Trilho 41	Roda 15	Trilho 26	0.187	0.004
Roda 11	Trilho 5	Roda 4	Trilho 22	1.705	0.335
Roda 20	Trilho 11	Roda 12	Trilho 19	0.157	0.0
Roda 13	Trilho 39	Roda 9	Trilho 2	0.0	0.015
Roda 14	Trilho 3	Roda 20	Trilho 28	0.0	0.051
Roda 15	Trilho 37	Roda 8	Trilho 41	0.0	0.242
Roda 20	Trilho 19	Roda 12	Trilho 3	0.138	0.099
Roda 2	Trilho 23	Roda 24	Trilho 2	0.0	0.097
Roda 24	Trilho 28	Roda 2	Trilho 2	0.125	0.009
Roda 20	Trilho 6	Roda 6	Trilho 38	0.013	0.069

Tabela 5.1: Amostra do banco de dados elaborado através da automatização da simulação.

De modo a acompanhar a performance da rede durante o treinamento do primeiro modelo, é apresentada, inicialmente, a curva da função perda na Figura 5.1 para os dados de treinamento e para os dados de validação.

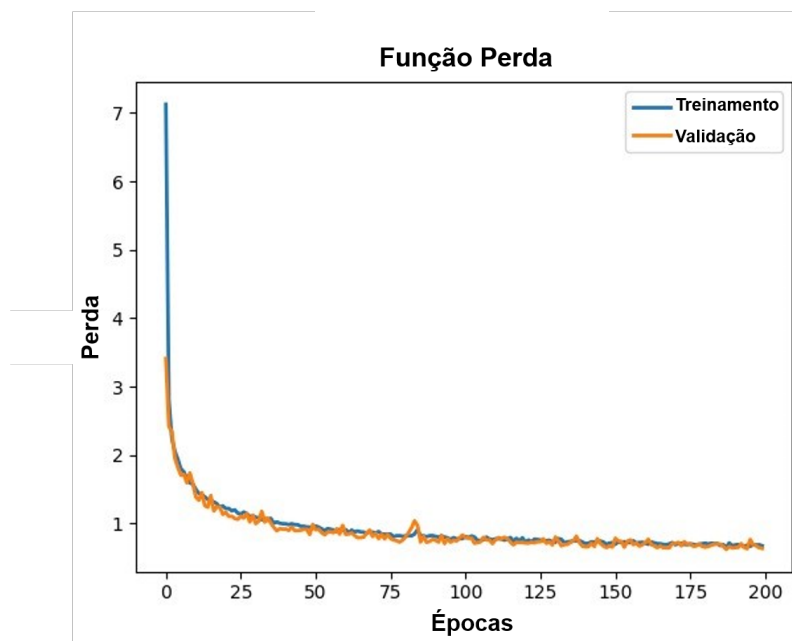


Figura 5.1: Curva da função perda do treinamento da primeira rede neural. (Autoria Própria)

Ao avaliar a curva azul no gráfico, que representa o erro obtido ao prever os resultados dos dados de treinamento, é possível afirmar que a previsão da IA melhorou à medida que o treinamento avançou. Em outras palavras, a rede neural foi capaz de calcular pesos que permitiram reduzir a função perda e, conseqüentemente, diminuir o erro de previsão. Ao avaliar a curva laranja, que representa o erro obtido ao prever os dados de validação, é possível afirmar que a IA também mostrou melhoria em suas previsões à medida que o treinamento progrediu. Esse tipo de resultado é importante, pois mostra que a IA apresentou baixo erro de previsão para novos casos.

Uma vez que as curvas da função de custo para treinamento e validação diminuem juntas e a distância entre elas é praticamente zero, pode-se dizer que não há *overfitting*, ou seja, a rede aprendeu correlações que permitem identificar o tipo de contato para casos além daqueles apresentados a ela.

As Figuras 5.2 e 5.3 mostram as matrizes de confusão para os conjuntos esquerdo e direito, respectivamente, em relação aos acertos e erros de previsão da rede sobre o tipo de contato para os dados de validação (2150 casos).

Contato	Quase Conforme	Conforme	Não Conforme
Quase Conforme	96%	4%	0%
Conforme	8%	91%	1%
Não Conforme	0%	5%	95%

Figura 5.2: Matriz de Confusão para o conjunto esquerdo. (Autoria Própria)

Contato	Quase Conforme	Conforme	Não Conforme
Quase Conforme	95%	5%	0
Conforme	10%	88%	2%
Não Conforme	0	7%	93%

Figura 5.3: Matriz de Confusão para o conjunto direito. (Autoria Própria)

De acordo com a Figura 5.2, para o conjunto esquerdo, dos 853 contatos do tipo Quase Conforme, 96% foram previstos corretamente como Quase Conforme e outros 4% equivocadamente como Conforme. Em relação aos 810 contatos do tipo Conforme, 91% foram previstos corretamente, sendo que 8% foi previsto como Quase Conforme e 1% como Não-Conforme. Dos 487 casos Não-Conforme, 95% foi classificado como Não-Conforme e 5% como Conforme.

De acordo com a Figura 5.3, para o conjunto da direita, há 870 contatos do tipo Quase Conforme, dos quais 95% foram previstos corretamente e 5% foram erroneamente classificados como Conforme. Em relação aos contatos do tipo Conforme, há 808 contatos dentre os quais 88% foram previstos corretamente, 10% foram previstos como Quase Conforme e 2% como Não Conforme. Dos 472 casos Não Conformes no banco de dados, 93% foram classificados como Não Conforme e 7% como Conforme.

A rede neural apresentou uma precisão de 93,91% para o conjunto da esquerda e 91,91% para o conjunto da direita, resultando em uma acurácia geral de 93% para o primeiro modelo.

Em se tratando do segundo modelo de IA, a Figura 5.4 apresenta a curva da função custo referente ao seu treinamento.

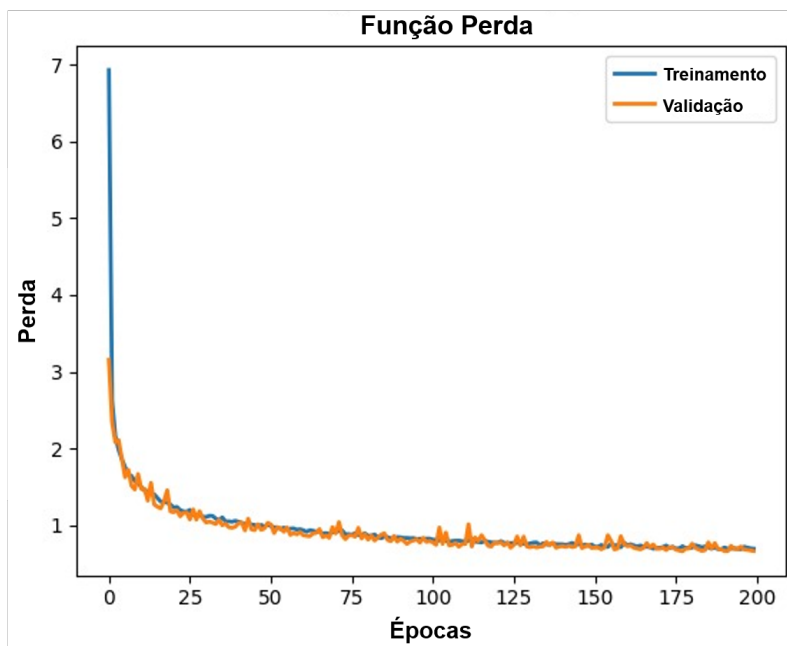


Figura 5.4: Curva da função perda do treinamento da segunda rede neural. (Autoria Própria)

Assim como no primeiro modelo, a previsão da IA melhorou à medida que o treinamento avançou, evidenciando também um aprendizado sem *overfitting* uma vez que ambas as curvas decresceram durante o processo.

As Figuras 5.5 e 5.6 mostram as matrizes de confusão para cada lado em relação aos acertos e erros de previsão da rede sobre o tipo de contato nos dados de validação (2150 casos).

Contato	Quase Conforme	Conforme	Não Conforme
Quase Conforme	811	42	0
Conforme	64	741	5
Não Conforme	0	35	452

Figura 5.5: Matriz de Confusão para o conjunto esquerdo. (Autoria Própria)



Contato	Quase Conforme	Conforme	Não Conforme
Quase Conforme	815	55	0
Conforme	62	719	27
Não Conforme	0	32	440

Figura 5.6: Matriz de Confusão para o conjunto direito. (Autoria Própria)

De acordo com a Figura 5.5, para o conjunto esquerdo, há 853 contatos do tipo Quase Conforme, dos quais 811 foram previstos corretamente como Quase Conforme e 42 foram equivocadamente previstos como Conforme. Em relação aos contatos do tipo Conforme, há 810 contatos, porém 741 foram previstos corretamente, sendo 64 previstos como Quase Conforme e mais 5 como Não-Conforme. Dos 487 casos de Não-Conforme no banco de dados, 452 foram classificados como Não-Conforme e 35 como Conforme.

De acordo com a Figura 5.6, para o conjunto da direita, há 870 contatos do tipo Quase Conforme, dos quais 815 foram previstos corretamente e 55 foram erroneamente classificados como Conforme. Em relação aos contatos do tipo Conforme, há 808 contatos dentre os quais 719 foram previstos corretamente, 62 foram previstos como Quase Conforme e 27 como Não Conforme. Dos 472 casos Não Conformes no banco de dados, 440 foram classificados como Não Conforme e 32 como Conforme.

Portanto, a rede neural apresentou uma precisão de 93,21% para o conjunto da esquerda e 91,81% para o conjunto da direita, resultando em uma acurácia geral de 92,5% para o segundo modelo.

Uma vez que os dois modelos treinados apresentam praticamente a mesma acurácia, é possível concluir então que a rotação das imagens das rodas não fornece vantagens no que concerne à classificação do contato através da análise dos perfis. Contudo, tendo em vista que o primeiro modelo expressou acurácia ligeiramente maior que o segundo modelo, as demais análises que serão realizadas nesta seção serão relacionadas à primeira rede neural.

Dentre os casos que a IA não foi capaz de prever, convém aqui apresentar alguns casos de modo que se possa entender as razões pela qual o modelo não conseguiu classificar de maneira correta o contato. A Figura 5.7 traz um caso cujo contato do conjunto esquerdo é do tipo Não Conforme e do conjunto direito é do tipo Conforme (vide o valor do  $s$  de cada caso). Embora a rede tenha acertado o contato do lado esquerdo, o mesmo não ocorreu para o lado direito, classificando-o como Não Conforme. O mesmo acontece para o caso da Figura 5.8 (cujo contato para o lado esquerdo e direito é Não Conforme e Conforme

respectivamente) onde o modelo previu corretamente para o lado esquerdo mas errou para o lado direito, classificando-o como Quase Conforme.

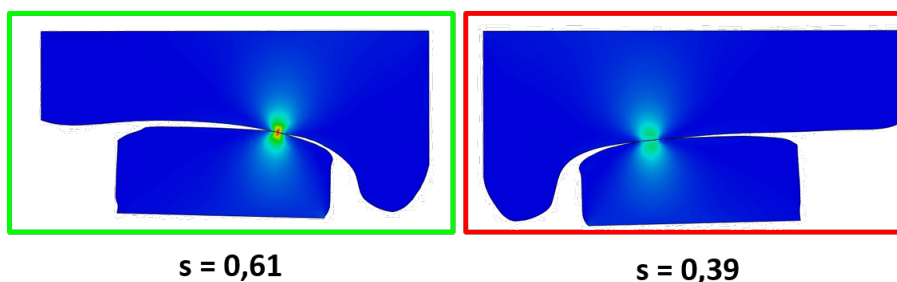


Figura 5.7: Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado direito. (Autoria Própria)

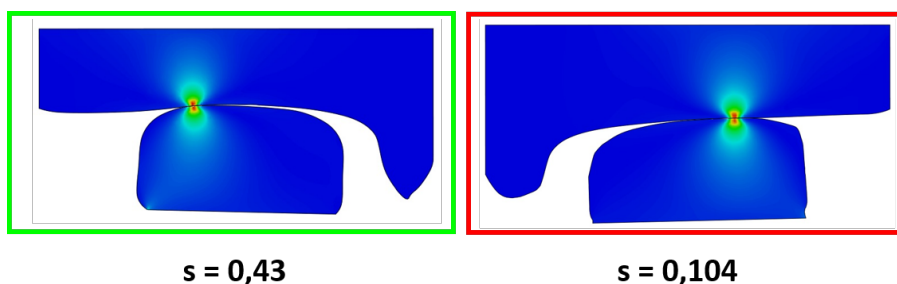


Figura 5.8: Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado direito. (Autoria Própria)

A Figura 5.9 traz um caso cujo contato é Não Conforme para ambos os conjuntos. Embora a rede neural tenha acertado na previsão do contato para o lado direito, a mesma falhou para o lado esquerdo, classificando o contato como do tipo Conforme. Já a Figura 5.10 traz um caso cujo contato é do tipo Conforme para o lado esquerdo e Quase Conforme para o lado direito. Na análise deste caso, o modelo classificou o lado direito corretamente mas errou para o lado esquerdo, prevendo-o como do tipo Quase Conforme.

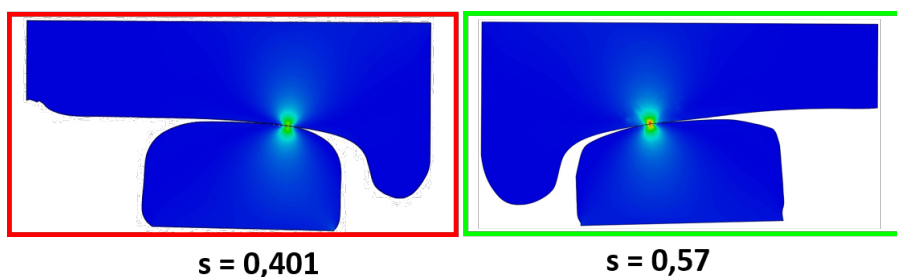


Figura 5.9: Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado esquerdo. (Autoria Própria)

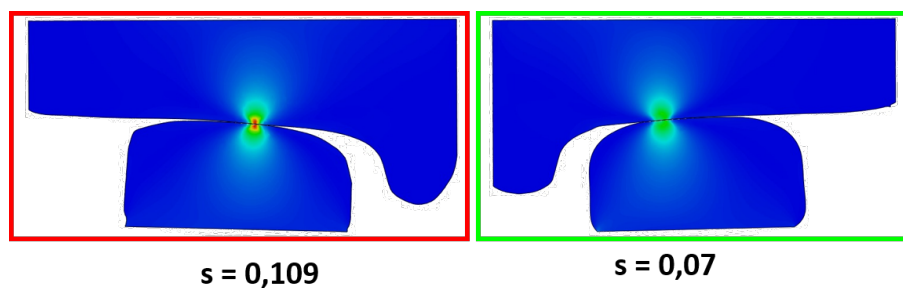


Figura 5.10: Exemplo de combinação onde o modelo não foi capaz de prever o contato no lado esquerdo. (Autoria Própria)

Tendo em vista o que foi exposto acima bem como os valores dos  $s$  de cada caso, é possível concluir que a rede neural apresentou dificuldade (em sua maioria) em classificar os casos que apresentaram valores de  $s$  muito próximos da fronteira de transição entre Quase Conforme e Conforme ( $s = 0,1$ ) e Conforme e Não Conforme ( $s = 0,4$ ). Esse tipo de conclusão pode auxiliar na tomada de decisões acerca do retreinamento do modelo, aplicando planos de ações que auxiliem a contornar este problema de previsão.

---

## Conclusão

Este trabalho propôs uma metodologia, utilizando redes neurais convolucionais, que permite a avaliação do contato roda-trilho fornecendo informações sobre o tipo de contato: Conforme, Não Conforme e Quase Conforme, cuja precisão atingiu 93%. Um dos diferenciais do treinamento desse modelo é o uso de um banco de dados que foi desenvolvido por meio de resultados obtidos por FEM, garantindo assim menores custos financeiros e de tempo.

No entanto, um dos fatores que contribuíram para a obtenção do banco de dados foi o desenvolvimento de *scripts* em *Python* que foram capazes de acessar o *software* Abaqus e executá-lo, permitindo que o processo de automação ocorresse, garantindo assim a geração de milhares de dados em um curto período de tempo.

Outro ponto que também chama a atenção é o fato de a rede neural convolucional ter sido capaz de encontrar uma correlação entre os contornos dos perfis das rodas e trilhos com o respectivo tipo de contato (o que conseqüentemente fornece informações sobre o tipo de dano que será prevalente no sistema). Esse fato levanta discussões e até permite estudos adicionais sobre quais outros fenômenos (tensões de von Mises, pressão de contato, etc.) as redes neurais convolucionais podem prever ao avaliar os perfis das geometrias, ou mesmo ajudar no estudo e compreensão de quais fenômenos são influenciados ou podem ser explicados por meio da avaliação dos perfis.

Particularmente, a Rede Neural Convolucional que foi desenvolvida neste trabalho permite sua aplicação direta para auxiliar na análise e prevenção de possíveis problemas que possam existir em um sistema ferroviário (trilhos e trens), além de poder ser usada como uma ferramenta da qualidade em diferentes setores.

Embora tenha se alcançado uma alta acurácia de previsão do tipo de contato, o modelo apresenta dificuldade em avaliar os casos cujo valor de  $s$  está próximo às zonas de transição entre os tipos de contato ( $s = 0,1$  e  $s = 0,4$ ). Isso também está relacionado à

limitação que a própria métrica apresenta em definir o tipo de contato para os casos com  $s = 0,1$  e  $s = 0,4$ .

Como sugestões de continuidade, diferentes aspectos da pesquisa podem ser aprimorados, destacando-se: modelo de elementos finitos, metodologia de IA e utilização de outros fenômenos (von Mises, pressão de contato, etc). Em se tratando do modelo virtual, diversas outras situações podem ser implementadas, tais como: curvas, aspectos dinâmicos, mudança da distância de bitola, novos modelos de vagão e conjuntos de roda e trilho. A inserção dessas novas condições podem fornecer informações acerca da mecânica do contato para diferentes situações, bem como possibilitar a obtenção de fenômenos que não são possíveis alcançar através da simulação quase-estática (a aplicação de dinamicidade ao modelo pode fornecer respostas de desgaste devido à rotação com escorregamento, por exemplo).

Quanto à metodologia de IA, novas abordagens podem ser aplicadas para aumentar a acurácia do modelo proposto, tais como: usar algoritmos de lógica fuzzy para auxiliar nos problemas de fronteira abordados e usar outras técnicas de processamento de imagens que podem auxiliar a rede a analisar os *inputs*. Ainda se tratando de inteligência artificial, uma outra possível metodologia seria o treinamento de uma rede neural de regressão, de modo que a IA pudesse ser capaz de calcular o valor do  $s$  através da análise dos perfis, onde a classificação do contato seria proveniente de pós-processamento. Essa abordagem poderia superar a limitação encontrada pela rede neural proposta por este trabalho em classificar os casos com valores de  $s$  muito próximos da fronteira de transição do tipo de contato.

E em relação a utilização dos demais *outputs* obtidos via simulação de elementos finitos, diversos estudos podem ser feitos de modo a investigar quais outros fenômenos um modelo semelhante a este pode desenvolver correlações com os perfis das rodas e trilhos, de forma que, no final, seja possível, por exemplo, desenvolver um grau de severidade que possa fornecer informações mais completas sobre o contato roda-trilho.

---

## Referências

- ACIKGOZ, Hakan; KORKMAZ, Deniz. MSRConvNet: Classification of railway track defects using multi-scale residual convolutional neural network. **Engineering Applications of Artificial Intelligence**, v. 121, p. 105965, 2023. ISSN 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2023.105965>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0952197623001495>>.
- AGRAWAL, Ishika et al. Emotion Recognition from Facial Expression using CNN. In: 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC). [S.l.: s.n.], 2021. P. 01-06. DOI: 10.1109/R10-HTC53172.2021.9641578.
- ANDRADE, António Ramos; TEIXEIRA, Paulo Fonseca. A Bayesian model to assess rail track geometry degradation through its life-cycle. **Research in Transportation Economics**, v. 36, n. 1, p. 1-8, 2012. Selected papers from the 12th WCTR Topic Area Transport Economics and Finance. ISSN 0739-8859. DOI: <https://doi.org/10.1016/j.retrec.2012.03.011>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0739885912000297>>.
- ANTF. **Associação Nacional de Transportes Ferroviários**. [S.l.: s.n.], 2023. Disponível em: <<https://www.antf.org.br/>>. Acesso em: 12 mai. 2023.
- CÁRDENAS-GALLO, Iván et al. An ensemble classifier to predict track geometry degradation. **Reliability Engineering & System Safety**, v. 161, p. 53-60, 2017. ISSN 0951-8320. DOI: <https://doi.org/10.1016/j.res.2016.12.012>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0951832016310109>>.
- CZICHOS, Horst; WINER, Ward Otis. Tribology: A Systems Approach to the Science and Technology of Friction, Lubrication and Wear (Tribology Series, 1). **Journal of Lubrication Technology**, v. 100, p. 513-514, 1978.

DING, Lijun; GOSHTASBY, Ardeshir. On the Canny edge detector. **Pattern Recognition**, v. 34, n. 3, p. 721–725, 2001. ISSN 0031-3203. DOI: [https://doi.org/10.1016/S0031-3203\(00\)00023-6](https://doi.org/10.1016/S0031-3203(00)00023-6). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0031320300000236>>.

DNIT. [S.l.: s.n.], 2023. Acesso em: 2023-05-12. Disponível em: <<https://www.gov.br/dnit/pt-br/assuntos/ferrovias/historico>>. Acesso em: 12 mai. 2023.

FRANCO, Cristiano Roberto. **Inteligência Artificial**. [S.l.: s.n.], 2017. P. 180. ISBN 978-85-515-0066-8.

HAENLEIN, Michael; KAPLAN, Andreas. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. **California Management Review**, v. 61, n. 4, p. 5–14, 2019. DOI: 10.1177/0008125619864925.

HAJIZADEH, Siamak; NÚÑEZ, Alfredo; TAX, David M.J. Semi-supervised Rail Defect Detection from Imbalanced Image Data. **IFAC-PapersOnLine**, v. 49, n. 3, p. 78–83, 2016. 14th IFAC Symposium on Control in Transportation SystemsCTS 2016. ISSN 2405-8963. DOI: <https://doi.org/10.1016/j.ifacol.2016.07.014>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2405896316302099>>.

IWNICKI, S.; BJÖRKLUND, S.; ENBLÖM, R. 3 - Wheel–rail contact mechanics. In: LEWIS, R.; OLOFSSON, U. (Ed.). **Wheel–Rail Interface Handbook**. [S.l.]: Woodhead Publishing, 2009. P. 58–92. ISBN 978-1-84569-412-8. DOI: <https://doi.org/10.1533/9781845696788.1.58>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9781845694128500033>>.

JAMSHIDI, Ali et al. A decision support approach for condition-based maintenance of rails based on big data analysis. **Transportation Research Part C: Emerging Technologies**, v. 95, p. 185–206, 2018. ISSN 0968-090X. DOI: <https://doi.org/10.1016/j.trc.2018.07.007>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0968090X18309859>>.

KHUDHAIR, Ali; TALIB, N. NEURAL NETWORK ANALYSIS FOR SLIDING WEAR OF 13%CR STEEL COATINGS BY ELECTRIC ARC SPRAYING. In.

KINGMA, Diederik P.; BA, Jimmy. **Adam: A Method for Stochastic Optimization**. [S.l.: s.n.], 2017. arXiv: 1412.6980 [cs.LG].

LASISI, Ahmed; ATTOH-OKINE, Nii. Principal components analysis and track quality index: A machine learning approach. **Transportation Research Part C: Emerging Technologies**, v. 91, p. 230–248, 2018. ISSN 0968-090X. DOI:

<https://doi.org/10.1016/j.trc.2018.04.001>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0968090X18304303>>.

LEE, DonHee; YOON, Seong. Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges. **International Journal of Environmental Research and Public Health**, v. 18, p. 271, jan. 2021. DOI: 10.3390/ijerph18010271.

LEWIS, R.; OLOFSSON, U. 2 - Basic tribology of the wheel-rail contact. In \_\_\_\_\_. **Wheel-Rail Interface Handbook**. Edição: R. Lewis e U. Olofsson. [S.l.]: Woodhead Publishing, 2009. P. 34-57. ISBN 978-1-84569-412-8. DOI: <https://doi.org/10.1533/9781845696788.1.34>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9781845694128500021>>.

LI, Z. et al. cardiGAN: A Generative Adversarial Network Model for Design and Discovery of Multi Principal Element Alloys, fev. 2022.

LIASHCHYNSKYI, Petro; LIASHCHYNSKYI, Pavlo. **Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS**. [S.l.: s.n.], 2019. arXiv: 1912.06059 [cs.LG].

LIMA, Vinicius Alves. Efeito das variáveis dos campos de tensões no desgaste e em danos superficiais no contato roda-trilho usando o MEF. **Dissertação de Mestrado**, 2023.

LOPES, Modesto Valci Moreira. **AbaqusAutomatization**. 2023. Disponível em: <<https://github.com/ModsVal/AbaqusAutomatization>>.

MACHADO, I.F. et al. Integração de Modelo de Elementos Finitos (MEF) e Aprendizado de Máquina (ML) para a Análise do Grau de Severidade no Sistema Roda-Trilho. **Relatório Técnico apresentado à Vale**, 2023.

MAGEL, Eric E; KALOUSEK, Joseph. The application of contact mechanics to rail profile design and rail grinding. **Wear**, v. 253, n. 1, p. 308-316, 2002. CM2000 S.I. ISSN 0043-1648. DOI: [https://doi.org/10.1016/S0043-1648\(02\)00123-0](https://doi.org/10.1016/S0043-1648(02)00123-0). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0043164802001230>>.

MASSA INDÚSTRIA. **MASSA Indústria**. [S.l.: s.n.], 2023. Disponível em: <<https://massa.ind.br/transporte-ferroviario-no-mundo/>>. Acesso em: 12 mai. 2023.

MATTEI, Lorenza; DI PUCCIO, Francesca. Frictionless vs. Frictional Contact in Numerical Wear Predictions of Conformal and Non-conformal Sliding Couplings. **Tribology Letters**, v. 70, set. 2022. DOI: 10.1007/s11249-022-01657-5.



MOHAMMADI, Reza et al. Exploring the impact of foot-by-foot track geometry on the occurrence of rail defects. **Transportation Research Part C: Emerging Technologies**, v. 102, p. 153–172, 2019. ISSN 0968-090X. DOI:

<https://doi.org/10.1016/j.trc.2019.03.004>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0968090X18314001>>.

NIEBLING, Julia; BAASCH, Benjamin; KRUSPE, Anna. Analysis of Railway Track Irregularities with Convolutional Autoencoders and Clustering Algorithms. In: [s.l.: s.n.], ago. 2020. P. 78–89. ISBN 978-3-030-58461-0. DOI: 10.1007/978-3-030-58462-7\_7.

PETROVIĆ, Aleksandar Dragan et al. Integration of Computer Vision and Convolutional Neural Networks in the System for Detection of Rail Track and Signals on the Railway. **Applied Sciences**, v. 12, n. 12, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/12/6045>>.

RIVES; PITA, A. L.; PUENTE, M. J. M. **Tratado de ferrocarriles I via**. Madrid: Rueda, 1977.

SHEBANI, A.; IWNICKI, S. Prediction of wheel and rail wear under different contact conditions using artificial neural networks. **Wear**, v. 406-407, p. 173–184, 2018. ISSN 0043-1648. DOI: <https://doi.org/10.1016/j.wear.2018.01.007>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0043164817304350>>.

URDA, Pedro et al. Artificial neural networks applied to the measurement of lateral wheel-rail contact force: A comparison with a harmonic cancellation method. **Mechanism and Machine Theory**, v. 153, p. 103968, 2020. ISSN 0094-114X. DOI: <https://doi.org/10.1016/j.mechmachtheory.2020.103968>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0094114X20301890>>.

WU, Qing et al. Methodology to optimize wedge suspensions of three-piece bogies of railway vehicles. **Journal of Vibration and Control**, v. 24, n. 3, p. 565–581, 2018. DOI: 10.1177/1077546316645698.

YEH, Wei-Chang et al. Simplified swarm optimization for hyperparameters of convolutional neural networks. **Computers & Industrial Engineering**, v. 177, p. 109076, 2023. ISSN 0360-8352.

ZAKHAROV, Sergey et al. **Guidelines to best practices for heavy haul railway operations: wheel and rail interface issues**. [S.l.: s.n.], mai. 2001. ISBN 5 89277 037 0.

ZHUANG, Li et al. Automated vision inspection of rail surface cracks: A double-layer data-driven framework. **Transportation Research Part C: Emerging Technologies**, v. 92, p. 258–277, 2018. ISSN 0968-090X. DOI: <https://doi.org/10.1016/j.trc.2018.05.007>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0968090X18306260>>.