



GABRIEL DE SOUZA LIMA

**DETECÇÃO DE ANOMALIAS EM PAVIMENTOS
UTILIZANDO SENSORES INERCIAIS**

SÃO PAULO
2024

GABRIEL DE SOUZA LIMA

**DETECÇÃO DE ANOMALIAS EM PAVIMENTOS
UTILIZANDO SENSORES INERCIAIS**

Versão Corrigida

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

Orientador:
Prof. Dr. José Roberto Castilho Piqueira

Área de Concentração:
Engenharia de Sistemas

SÃO PAULO

2024

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, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Lima, Gabriel

Detecção de anomalias em pavimentos utilizando sensores inerciais / G.
Lima -- versão corr. -- São Paulo, 2024.

143 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Telecomunicações e Controle.

1.Inteligência Artificial 2.Sistemas de Controle I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Telecomunicações e Controle II.t.

Este trabalho é dedicado à Maria Aparecida dos Reis Lima e Wilson de Souza Lima, que desde criança me ensinaram o valor do conhecimento.

AGRADECIMENTOS

Durante toda minha trajetória, tive a oportunidade de ter pessoas que de alguma maneira contribuíram para minha formação acadêmica e profissional. Nesse momento, lembrando de todo o percurso para a realização desse trabalho só consigo expressar o meu sentimento de gratidão para todos que até aqui estiveram comigo.

Agradeço primeiramente a Deus por ser meu baluarte, minha luz e o meu caminho, me presenteando a cada dia com a oportunidade de estar vivo.

Agradeço a minha família, em especial meus pais Wilson Lima e Maria Reis por me guiarem para o caminho do trabalho, do estudo e do amor ao próximo, aos meus irmãos Andrea, Anderson, Fabio e Fabiana por sempre estarem comigo e me servirem como exemplo durante toda a minha vida, e a minha avó Aparecida Thereza (*In Memoriam*) que ficou muito feliz quando eu iniciei o mestrado e certamente estaria orgulhosa de mim agora.

Ao meu orientador José Piqueira, o meu muito obrigado por todo o conhecimento transmitido e por ser a minha inspiração na vida acadêmica.

Agradeço também aos professores Angelo Zanini e Flavio Maranhão pela confiança e oportunidade em mim depositada, ao professor Nuncio Perrella que enxergou o meu potencial quando eu mesmo não o via e aos professores que também marcaram a minha carreira: Sandro Martini, Elaine Martini, Robson Colantonio, Gediael Felipe, Eduardo Botter e Humberto Feliz.

Aos meus amigos Enio Blay, Mauro Chiozzotto, Jonathan Zappile, Marcio Paixão, Giovanni Schiffini, Tiago Marum, Renan Andrade, Bruno Marques, Renan Bassetto, Sandro Sakata, Adrian Lozano e Gustavo Mansur, que foram fundamentais no dia a dia, compartilhando momentos e dividindo aprendizado.

Agradeço a minha noiva Karoline Fragoso por todo apoio e compreensão para que eu pudesse viver esse sonho e a Mel, a *pet* responsável por discutir comigo durante as inúmeras madrugadas de trabalho todos os conceitos aqui escrito.

*“Não vos amoldeis às estruturas deste mundo, mas transformai-vos pela renovação da mente, a fim de distinguir qual é a vontade de Deus: o que é bom, o que Lhe é agradável, o que é perfeito.
(Bíblia Sagrada, Romanos 12, 2))*

LIMA, G. S.. **Detecção de Anomalias em Pavimentos Utilizando Sensores Inerciais**. 2024. 142f. Dissertação (Mestrado em Ciências) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2024.

RESUMO

Sistemas para detecção automática de anomalias no pavimento são de suma importância para gestores públicos, seja para a realização de reparos de buracos, para monitoramento de ativos ou para mapear a deterioração das vias. Na última década tem crescido a utilização de automóveis providos de sensores inerciais embarcados para a realização dessa tarefa, no qual os dados de campo são captados e transmitidos para análises que comumente empregam algoritmos de aprendizado supervisionado para a identificação dos elementos no asfalto, entretanto um problema intrínseco com essa abordagem está no fato de que a coleta dos dados anômalos é em sua essência difícil e desbalanceada, o que potencialmente aumenta a geração de falsos positivos no momento da predição. Outro problema comum nessa área é a não utilização de um método para o estabelecimento da frequência de amostragem para esses sensores, fazendo com que muitos trabalhos sobre amostragem o sinal de captura, inviabilizando a criação de sistemas descentralizados devido ao alto número de amostras por veículo. Portanto, o trabalho aqui descrito tem como objetivo criar um sistema para a detecção de anomalias, onde um método para a definição da frequência de amostragem é estabelecido e um modelo de aprendizado não supervisionado testado para o módulo de detecção. Resultados apontam que é possível trabalhar com uma frequência de amostragem menor que a frequência de Nyquist, subamostrando assim o sinal no momento da coleta e, a partir de um módulo de aprendizado não supervisionado pode-se detectar amostras anômalas com uma assertividade de 94%.

Palavras-chave: Detecção de Anomalias. Autoencoder. Sensores Inerciais.

LIMA, G. S.. **Pavement Anomaly Detection Using Inertial Sensors**. 2024. 142p. Master's Thesis (Master in Science in Computer Science) – University of São Paulo, São Paulo, 2024.

ABSTRACT

Systems for automatically detecting pavement anomalies are extremely important for public managers, whether for repairing potholes, monitoring assets or mapping road deterioration. In the last decade, the use of automobiles equipped with on-board inertial sensors has grown to carry out this task, in which field data is captured and transmitted for analyzes that commonly employ supervised learning algorithms to identify elements on the asphalt, however a problem Intrinsic to this approach is the fact that collecting anomalous data is essentially difficult and unbalanced, which potentially increases the generation of false positives at the time of prediction. Another common problem in this area is the failure to use a method to establish the sampling frequency for these sensors, causing many studies to sample the capture signal, making the creation of decentralized systems unfeasible due to the high number of samples per vehicle. Therefore, the work described here aims to create a system for anomaly detection, where a method for defining the sampling frequency is established and an unsupervised learning model tested for the detection module. Results indicate that it is possible to work with a sampling frequency lower than the Nyquist frequency, thus sub-sampling the signal at the time of collection and, using an unsupervised learning module, it is possible to detect anomalous samples with an accuracy of 94 %.

Keywords: Anomaly Detection. Autoencoder. Inertial Sensors.

LISTA DE ILUSTRAÇÕES

Figura 1 – Métodos para detecção de anomalias em pavimentos	18
Figura 2 – Técnicas para detecção de anomalias em pavimentos utilizando sensores inerciais	19
Figura 3 – Dashboard - Sistema de gestão de Pavimentos Geovista	21
Figura 4 – Geobox	22
Figura 5 – Framework para detecção de anomalias	23
Figura 6 – Revisão bibliométrica - Método de pesquisa	26
Figura 7 – Rede de palavras relacionadas com as palavras chaves processadas pelo software VOSviewer	27
Figura 8 – Agrupamento das palavras relacionadas processadas pelo software VOSviewer	28
Figura 9 – Principais publicações processadas pelo software VOSviewer	30
Figura 10 – Frequência de publicação de artigos de detecção de anomalias em vias através de sensores de vibração utilizando machine learning	31
Figura 11 – Relação das áreas de interesse - Detecção de anomalias em pavimentos	33
Figura 12 – Taxonomia - Detecção de anomalias	34
Figura 13 – Métricas de distância para um conjunto de dados contendo dois pontos.	35
Figura 14 – Grupo de pontos sem classificação prévia, categorizados em dois grupos pelo algoritmo, onde a anomalia é definida como sendo o ponto que tem a maior distância até os centroides	37
Figura 15 – Demonstração da inferência de um novo ponto p através do algoritmo KNN.	38
Figura 16 – Pontos separados em dois grupos pelo hiperplano ótimo após a utilização do algoritmo SVM.	38
Figura 17 – Representação do funcionamento do algoritmo PCA, onde os dados normais de treino podem ser observados em verde, vermelho e azul e as anomalias que são introduzidas na inferência estão na cor preta. . .	43
Figura 18 – Funcionamento de uma ANN <i>autoencoder</i> , onde o conceito fundamental dessa técnica é ilustrado por meio de um diagrama de blocos representando as camadas de compressão, gargalo e reconstrução.	44
Figura 19 – Modelo 1DOF utilizado para representar um veículo passando sobre um obstáculo com uma velocidade longitudinal v	49
Figura 20 – Modelos 2DOF utilizado para representar um veículo passando sobre um obstáculo com uma velocidade longitudinal v	50
Figura 21 – Geobox instalado em um veículo	53
Figura 22 – Método Proposto para a Detecção de Anomalias.	54

Figura 23 – Diagrama de corpo livre - Modelo quarto de carro	55
Figura 24 – Obstáculo padronizado	58
Figura 25 – Aplicativo para Coleta de Dados de Treino.	59
Figura 26 – Vias boas - Sistema Geovista	60
Figura 27 – Aplicativo para Coleta de Dados de Teste.	60
Figura 28 – Carro passando sobre uma anomalia - Análise do pitch, roll, yaw	62
Figura 29 – Fluxo para Pré Processamento.	62
Figura 30 – Criação das janelas de tempo	63
Figura 31 – RNA - Autoencoder LSTM	64
Figura 32 – Diagrama de Bode da região não amortecida	68
Figura 33 – Análise transitória no domínio do tempo - 10 km/h	70
Figura 34 – Análise transitória no domínio da frequência- 10 km/h	71
Figura 35 – Análise transitória no domínio do tempo - 20 km/h	72
Figura 36 – Análise transitória no domínio da frequência- 20 km/h	73
Figura 37 – Análise transitória no domínio do tempo - 30 km/h	74
Figura 38 – Análise transitória no domínio da frequência- 30 km/h	75
Figura 39 – Análise transitória no domínio do tempo - 40 km/h	76
Figura 40 – Análise transitória no domínio da frequência- 40 km/h	77
Figura 41 – Análise transitória no domínio do tempo - 50 km/h	78
Figura 42 – Análise transitória no domínio da frequência- 50 km/h	79
Figura 43 – Resposta Transitória amostrada a 20 Hz - 10 km/h.	81
Figura 44 – Resposta Transitória amostrada a 20 Hz - 20 km/h.	82
Figura 45 – Resposta Transitória amostrada a 20 Hz - 30 km/h.	83
Figura 46 – Resposta Transitória amostrada a 20 Hz - 40 km/h.	84
Figura 47 – Resposta Transitória amostrada a 20 Hz - 50 km/h.	85
Figura 48 – Regiões para frequência de amostragem	86
Figura 49 – Vias boas para treino de máquina	87
Figura 50 – Análise do arquivo 001 - Rua da Mooca	88
Figura 51 – Análise do arquivo 002 - Rua dos Trilhos	89
Figura 52 – Análise do arquivo 003 - Rua da Mooca	90
Figura 53 – Análise do arquivo 004 - Rua dos Trilhos	91
Figura 54 – Análise do arquivo 005 - Rua da Mooca	92
Figura 55 – Análise do arquivo 006 - Avenida Paes de Barros	93
Figura 56 – Histórico do treino de máquina	95
Figura 57 – Histograma para determinação do limiar do erro	95
Figura 58 – Matriz de confusão	97

LISTA DE TABELAS

Tabela 1 – Metadados da primeira pesquisa para revisão bibliométrica	26
Tabela 2 – Metadados da segunda pesquisa para revisão bibliométrica	29
Tabela 3 – Metadados da terceira pesquisa para revisão bibliométrica	30
Tabela 4 – Principais algoritmos - Detecção de anomalias (Métodos baseados em distância)	36
Tabela 5 – Principais algoritmos - Redes neurais artificiais	40
Tabela 6 – Algoritmos baseados em reconstrução	45
Tabela 7 – Publicações que utilizam métodos baseados em vibrações para detecção de anomalias no pavimento.	48
Tabela 8 – Principais Especificações Honda City 2013	52
Tabela 9 – Parâmetros Simulação Quarto de Carro	66
Tabela 10 – Relação velocidade, amostragem e tempo de estabilização	80
Tabela 11 – Arquivos de treino	87
Tabela 12 – Quantidade de amostras - Dataset de teste	94
Tabela 13 – Métricas do sistema de inteligência artificial - Resultado	97

LISTA DE ABREVIATURAS E SIGLAS

ALARM	<i>Annual Local Authority Road Maintenance</i>
ANN	Redes Neurais Artificiais - Do inglês <i>Artificial Neural Network</i>
COF	Conectividade Baseada em Fator de Outliers - Do inglês <i>Connectivity-Based Outlier Factor</i>
CNN	Redes Neurais Convolucionais - Do inglês <i>Convolutional Neural Network</i>
CV	Cavalo vapor - Unidade de medida de potência
DOF	Grau de Liberdade - Do inglês <i>Degree of Freedom</i>
FFT	Transformada Rápida de Fourier - Do inglês <i>Fast Fourier Transform</i>
FN	Falso Negativo - Do inglês <i>False Negative</i>
GIS	Sistema de Informação Geográfica - Do inglês <i>Geographic Information System</i>
GRU	Unidade Recorrente com Portas- Do inglês <i>Gated Recurrent Unit</i>
IOT	Internet das Coisas - Do inglês <i>Internet of Things</i>
IRI	Índice de Rugosidade Internacional
Kg	Quilo grama - Unidade de medida de massa
KNN	K Vizinhos mais Próximos - Do inglês <i>K Nearest Neighbor</i>
LIDAR	Detecção de Luz e alcance - Do inglês <i>Light Detection and Ranging</i>
LOF	Fator Local de Outlier - Do inglês <i>Local Outlier Factor</i>
LSTM	Memória de Longo e Curto Prazo - Do inglês <i>Long-Short Term Memory</i>
MLP	Perceptron de Múltiplas Camadas - Do inglês <i>Multi Layer Perceptron</i>
MSE	Erro Médio Quadrático - Do inglês <i>Mean Square Error</i>
OCSVM	Máquina de Vetores de Suporte de Uma Classe- Do inglês <i>One Class Support Vector Machine</i>
PCA	Análise das Componentes Principais - Do inglês <i>Principal Component Analysis</i>

RGB	Vermelho, Verde, Azul - Do inglês <i>Red, Green, Blue</i>
RMSE	Raiz do Erro Médio Quadrático <i>Root Mean Square Error</i>
RNN	Redes Neurais Recorrentes - Do inglês <i>Recurrent Neural Network</i>
SL	Aprendizado Supervisionado - Do inglês <i>Supervised Learning</i>
SSL	Aprendizado Semi Supervisionado - Do inglês <i>Semi-Supervised Learning</i>
SVM	Máquina de Vetores de Suporte - Do inglês <i>Support Vector Machine</i>
UL	Aprendizado não Supervisionado - Do inglês <i>Unsupervised Learning</i>
WoS	Web of Science

LISTA DE SÍMBOLOS

\mathcal{L}	Libra esterlina - Moeda da Grã-Bretanha
M, m	Massa suspensa e não suspensa
c	Constante de amortecimento da região suspensa
z, \dot{z}, \ddot{z}	Deslocamento, velocidade e aceleração vertical da região suspensa
$z_u, \dot{z}_u, \ddot{z}_u$	Deslocamento, velocidade e aceleração vertical da região não suspensa
$z_r, \dot{z}_r, \ddot{z}_r$	Deslocamento, velocidade e aceleração vertical imposta pela via
k, k_t	Constante de rigidez da região suspensa e não suspensa
ω_{n1}, ω_{n2}	Frequência natural da região suspensa e não suspensa
ζ_s, ζ_u	Relação de amortecimento da região suspensa e não suspensa

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Justificativa da Pesquisa	20
1.2	Objetivos	21
1.3	Contexto da Pesquisa	21
1.4	Metodologia	22
1.5	Estrutura da Dissertação	24
2	REVISÃO BIBLIOMÉTRICA	25
2.1	Fase 1	26
2.2	Fase 2	29
2.3	Conclusão da Revisão Bibliométrica	32
3	REVISÃO BIBLIOGRÁFICA	33
3.1	Algoritmos de detecção de anomalias	33
3.1.1	Métodos baseados em distância	34
3.1.2	Métodos baseados em densidade	36
3.1.3	Métodos baseados em modelos	36
3.1.3.1	Algoritmos de agrupamento	36
3.1.3.2	Algoritmos de classificação	37
3.1.3.3	Algoritmos baseados em reconstrução	42
3.2	Métodos de detecção de anomalias no pavimento	46
3.2.1	Métodos baseados em visão computacional 2D	46
3.2.2	Métodos baseados em reconstrução 3D	46
3.2.3	Métodos baseados em vibrações	47
3.3	Métodos analíticos para estudo da vibração em suspensões	49
3.4	Conclusão da Revisão Bibliográfica	51
4	PROGRAMA EXPERIMENTAL	52
4.1	Instalação dos sensores	52
4.2	Método para Definição da Frequência de Amostragem	53
4.2.1	Modelo Matemático	54
4.2.2	Resposta do Sistema	57
4.3	Coleta do <i>Dataset</i> de Treino	59
4.4	Coleta do <i>Dataset</i> de Teste	60
4.5	Escolha das <i>Features</i> do sistema	61
4.5.1	Velocidade	61

4.5.2	Aceleração	61
4.5.3	Velocidade Angular	61
4.6	Módulo de Pré Processamento	62
4.6.1	Análise da integridade do <i>Dataset</i>	63
4.6.2	Criação da janela de tempo	63
4.6.3	Normalização dos dados	63
4.7	Módulo de Inteligência Artificial	64
4.7.1	Treino de máquina	64
4.7.2	Definição do limiar do erro	65
5	RESULTADOS E DISCUSSÕES	66
5.1	Caracterização do Sistema	66
5.2	Simulação	67
5.2.1	Resposta em Estado Estacionário	67
5.2.2	Resposta em Estado Transitório	69
5.2.3	Definição da Frequência de Amostragem	80
5.3	Análise no <i>Dataset</i> de Treino	86
5.4	Análise no <i>Dataset</i> de Teste	94
5.5	Módulo de Inteligência Artificial	94
5.5.1	Treino de Máquina	94
5.5.2	Definição do Limiar de Erro	95
5.5.3	Teste Controlado	96
6	CONCLUSÕES E CONSIDERAÇÕES FINAIS	98
6.1	Sugestões para Trabalhos Futuros	99
	REFERÊNCIAS	100
	APÊNDICES	109
	APÊNDICE A – REVISÃO BIBLIOMÉTRICA - QUERIES	110
	APÊNDICE B – REVISÃO BIBLIOMÉTRICA - GRUPO DE PALAVRAS RELACIONADAS	111
	APÊNDICE C – CÓDIGO - SIMULAÇÃO DA RESPOSTA ESTACIONÁRIA	123
	APÊNDICE D – CÓDIGO - SIMULAÇÃO DA RESPOSTA TRANSITÓRIA	125

APÊNDICE E – ANÁLISE - DATASET DE TESTE 126

ANEXOS 140

ANEXO A – ESPECIFICAÇÕES DO VEÍCULO UTILIZADO 141

1 INTRODUÇÃO

Nos grandes centros urbanos a qualidade do asfalto é um item de suma importância, fazendo com que gestores em diversos países priorizem ações para reparo e manutenção.

Apenas em 2022 a cidade de Perinth na Austrália registrou cerca de 1300 chamados para tapa buraco [1], no mesmo ano a ALARM (*Annual Local Authority Road Maintenance*) survey mostrou que a Grã-Bretanha teve um custo de £12 bilhões para reparar suas estradas esburacadas e as obras podem levar mais de nove anos para serem concluídas [2, 3].

Diversos fatores contribuem para que haja deterioração no pavimento, onde podemos destacar as condições climáticas, a sobrecarga de tráfego, a idade da infraestrutura, os materiais e métodos utilizados em sua construção [4].

Não apenas buracos, mas diversas patologias geram prejuízos aos cofres públicos e afetam o conforto dos usuários como é o caso de remendos realizados de forma incorreta, fissuras geradas pelo desgaste do pavimento, entre outros [5].

Sendo assim, um sistema que reporta tais defeitos de forma automatizada torna-se importante, visto que tal tecnologia reduzirá o trabalho de gerenciar buracos (em muitos centros urbanos fica a cargo dos cidadãos informar tais problemas no asfalto) e trará contribuições para o planejamento de manutenção de vias [6], principalmente quando se tem a necessidade de descobrir a origem do evento, podendo então responsabilizar empresas que realizam intervenções e não concluem de forma adequada.

Outra demanda latente está na área de engenharia de tráfego, possibilitando engenheiros obterem mapas das lombadas existentes na cidade, encontrando assim pontos carentes de tal ativo. Desse modo, para gestores públicos, tão importante quanto detectar anomalias é classificá-las, permitindo diferentes tomadas de decisões para inúmeros propósitos.

Nesse cenário, entende-se por anomalias como sendo qualquer elemento físico de natureza patológica (buraco, trincas, remendo) ou na forma de ativo (lombada, sarjetão, tampas de bueiro) que possa perturbar o conforto do motorista [7].

Verifica-se então, conforme mostra a Figura 1 que para a construção desse sistema de detecção e/ou classificação de anomalias em pavimentos há três métodos possíveis na literatura.

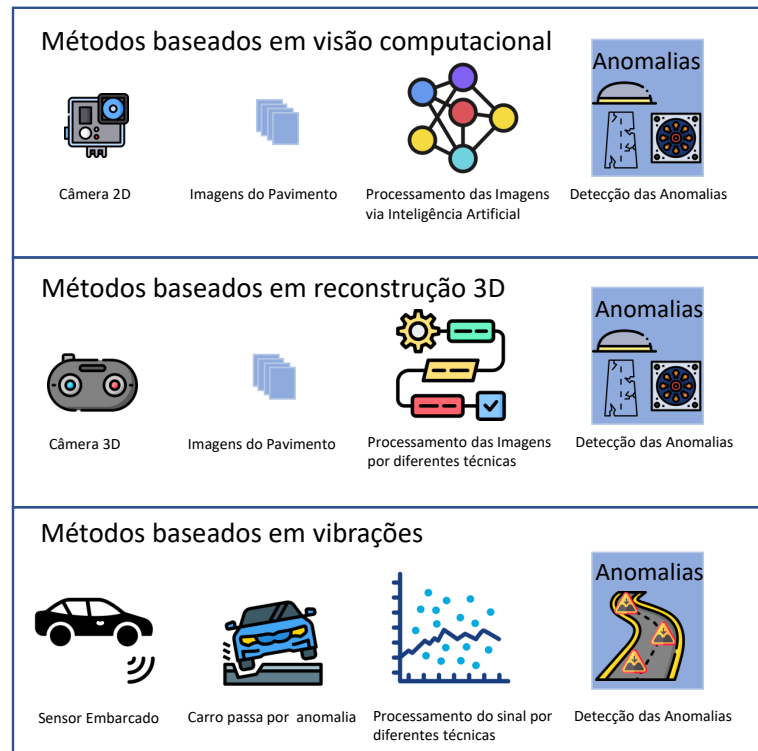


Figura 1 – Métodos para detecção de anomalias em pavimentos. (Fonte: Autor 2024)

Em síntese podemos elencar que os métodos baseados em visão computacional 2D utilizam apenas imagens oriundas de câmeras 2D para a geração de dados e algoritmos de inteligência artificial para processamento, enquanto métodos baseados em reconstrução 3D utilizam diversas técnicas de escaneamento para a criação de um modelo tri dimensional, e métodos baseados em vibrações utilizam sensores inerciais para a fase de captação [4].

Pode-se definir sensores inerciais como sendo todo dispositivo que utiliza a lei da inércia de Newton para medir grandezas como aceleração (acelerômetro) e velocidade angular (giroscópio) [8], estando estes disponíveis em uma vasta gama de smartphones, tendo pouca necessidade de banda para transmissão e possuindo o melhor custo benefício frente a outras técnicas, se tornando assim adequados para sistemas de tempo real [6, 9]. Porém um problema intrínseco da utilização desses sensores para a detecção de anomalias no pavimento é o alto número de FN (Falso Negativo - Do inglês *False Negative*) que essa solução possui, principalmente quando a classe de interesse é buraco, visto que o motorista tende a desviar de tal elemento [10].

Portanto, para mitigar esse problema e aumentar a probabilidade de detecção podem ser utilizadas soluções híbridas, unindo esta tecnologia com a área de visão computacional [9], ou ainda soluções colaborativas [11], onde diversos carros formam uma rede *crowdsensing* percorrendo a mesma via inúmeras vezes.

Outro fator chave para a utilização de sensores inerciais é a escolha da frequência de amostragem, visto que se escolhida de forma correta esta permitirá uma representação fidedigna dos dados de coleta, não gerando amostras de forma redundante ou de forma simplificada, possibilitando economia no armazenamento das informações, principalmente no caso em que a coleta dos dados acontece com inúmeros veículos. Contudo, atualmente na literatura não há uma definição concreta em torno deste valor [4], fazendo com que a escolha dessa grandeza seja realizada sem um método ou uma análise aprofundada.

No tocante a métodos baseados em vibrações a partir de sensores inerciais é possível verificar três maneiras predominantes de se detectar e/ou classificar anomalias [4] conforme mostra a Figura 2.

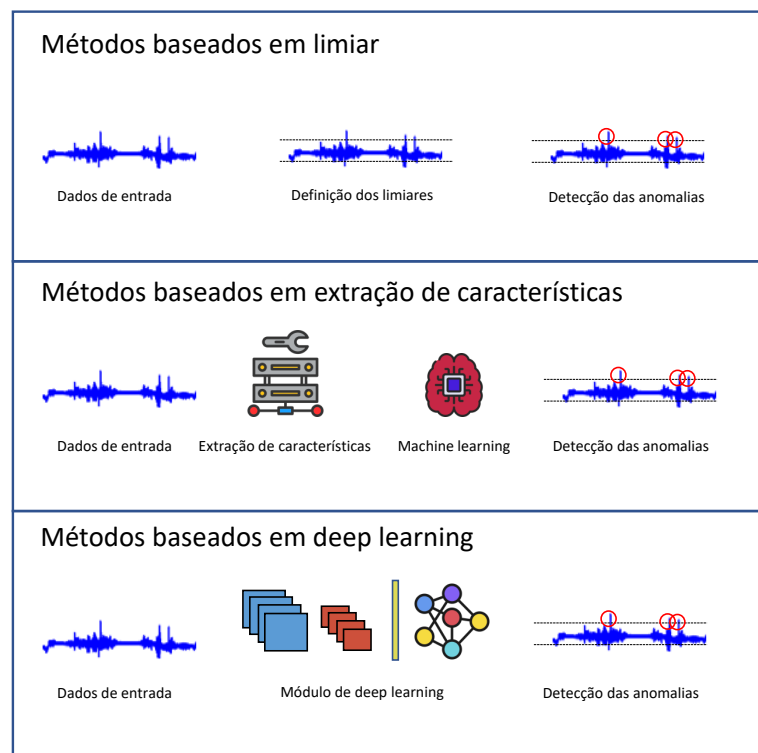


Figura 2 – Técnicas para detecção de anomalias em pavimentos utilizando sensores inerciais. (Fonte: Autor 2024)

Métodos baseados em limiar utilizam regras condicionais para realizar a detecção de anomalias, não as classificando, onde pode-se analisar apenas o eixo vertical do acelerômetro ou combinar com as demais direções dos sensores [12].

Por sua vez, métodos baseados em extração de características utilizam diversas técnicas para pré-processamento a fim de reduzir a dimensionalidade do conjunto de dados, sintetizando as informações em uma base processada para um algoritmo de aprendizado de máquina que pode tanto classificar como detectar as anomalias. Em compensação métodos baseados em *deep learning* realizam o aprendizado sem a necessidade dessa etapa, ou seja,

os dados são utilizados em sua forma bruta [4].

Técnicas que utilizam aprendizado de máquina vem ganhando cada vez mais espaço não apenas para detecções de anomalias como buracos, mas também para a estimação do IRI (Índice de Rugosidade Internacional) [12], onde tais abordagens podem ser aplicadas utilizando algoritmos de SL (Aprendizado Supervisionado - Do inglês *Supervised Learning*) ou UL (Aprendizado não Supervisionado - Do inglês *Unsupervised Learning*).

O que difere os algoritmos SL de UL é a forma de se construir a base de dados para o aprendizado de máquina, uma vez que o SL requer que os dados sejam anotados para treinamento, inserindo não apenas as características de uma ocorrência como entrada do sistema, mas também a informação de sua classe. Porém essa necessidade acaba dificultando a coleta dos dados, visto que a classe anômala é em sua essência rara [13].

Muitos estudos para classificação de anomalias em pavimentos utilizam um passageiro para rotular em tempo real os dados coletados durante o percurso [14], porém esse método está sujeito a falhas humanas podendo comprometer a coleta e conseqüentemente o treino, gerando um alto número de FP (Falso Positivo - Do inglês *False Positive*) na fase de predição.

Em contrapartida algoritmos de UL tentam aprender padrões ou estruturar os dados sem depender de bases pré rotuladas [15] detectando *outliers* ou segmentos anômalos que destoam da maioria dos dados no espaço de características [16].

Diversas áreas atualmente utilizam UL ou SSL (Aprendizado Semi Supervisionado - Do inglês *Semi-Supervised Learning*) para a detecção de anomalias como a medicina [17, 18, 19], cyber segurança [20, 21, 22] e *supply chain* [23], demonstrando o potencial desta técnica que ainda não foi aplicada para área de pavimentos.

1.1 Justificativa da Pesquisa

Com base no cenário anteriormente mencionado nota-se a necessidade de esforços para a construção de um sistema automático de detecção e classificação de anomalias no pavimento, sendo então proposto a utilização em cascata de dois algoritmos de aprendizado de máquina a fim de minimizar o problema recorrente da geração de FP em algoritmos de SL, dado que em um sistema desse tipo um algoritmo de UL pode ser treinado a partir de uma grande e irrestrita base de dados e então fornecer apenas os trechos preditos como anomalia para um segundo algoritmo que por ter em sua entrada informações já filtradas pode tomar decisões mais assertivas conforme demonstrado em [24]. Por fim, destaca-se que o foco do presente trabalho está no módulo de detecção uma vez que já há pesquisas voltadas para a área de classificação [7, 25].

1.2 Objetivos

O principal objetivo deste trabalho é propor um sistema de detecção de anomalias em pavimentos a partir de sensores inerciais embarcados em automóveis.

Para a realização de tal feito tem-se como objetivos específicos:

- Definir um método para a obtenção da frequência de aquisição de sensores inerciais.
- Avaliar o potencial do aprendizado não supervisionado para a detecção de anomalias no pavimento.

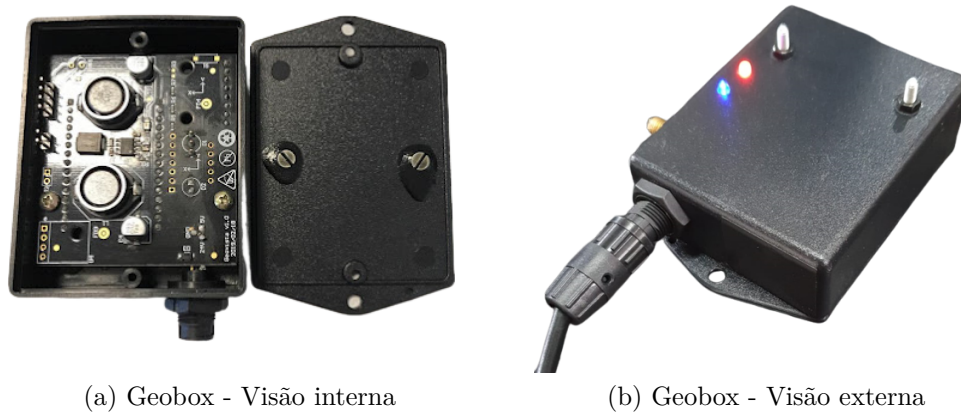
1.3 Contexto da Pesquisa

A presente pesquisa tem como base dados oriundos do projeto Geovista, onde pesquisadores da Universidade de São Paulo e Instituto Mauá de Tecnologia em parceria com a FDTE fornecem soluções focadas em zeladoria urbana para cidades inteligentes a partir da geração de uma grande massa de dados por meio de uma aplicação *crowdsourcing* que é capaz de captar informações de diversos sensores inerciais instalados em uma frota de aproximadamente 150 carros que trafegam por diversas cidades do Brasil [26], gerando *dashboards* interativos com diversas análises como o da qualidade superficial do pavimento mostrado na Figura 3.



Figura 3 – Dashboard - Sistema de gestão de Pavimentos Geovista. (Fonte: Autor 2024)

Desse modo será utilizado o *hardware* já desenvolvido pela empresa denominado Geobox, sendo este um circuito integrado dotado de um acelerômetro e um giroscópio em uma caixa de PVC blindada contra humidade, poeira e solavancos como ilustrado na Figura 4.



(a) Geobox - Visão interna

(b) Geobox - Visão externa

Figura 4 – Geobox. (Fonte: Autor 2024)

Por não conter sensores de GPS (Sistema de Posicionamento Global - Do inglês *Global Positioning System*) e a comunicação ser apenas por *bluetooth* tal dispositivo necessita realizar conexões com um *smartphone* para armazenamento, transmissão via internet e geolocalização dos dados, sendo comumente utilizado um aplicativo instalado no aparelho do motorista para a realização desta tarefa.

Por fim pode-se destacar que o autor tem atuado como engenheiro de dados nesse projeto, desenvolvendo estratégias para análise, manipulação e processamento de séries temporais para detecção e classificação de anomalias no pavimento.

1.4 Metodologia

Em conformidade com o objetivo principal a Figura 5 detalha a arquitetura utilizada para a construção do sistema proposto, onde carros de passeio com o dispositivo Geobox instalado são utilizados para a coleta e envio dos dados para um servidor em nuvem que através de um *pipeline* de inteligência artificial realiza a análise, treinamento, inferência e georreferenciamento dos dados, apresentando como produto final em um dashboard as anomalias detectadas. Esta arquitetura (mesmo que sem o módulo contendo o algoritmo de SL) possibilita compreender a fundo como se dá a dinâmica de captura dos dados oriundos de sensores inerciais e permite atingir os objetivos específicos propostos, proporcionando uma flexibilidade ao fluxo para que futuramente se desejado possa ser acoplado um módulo em série para receber as informações já processadas pelo algoritmo de UL e realizar a classificação a posteriori.

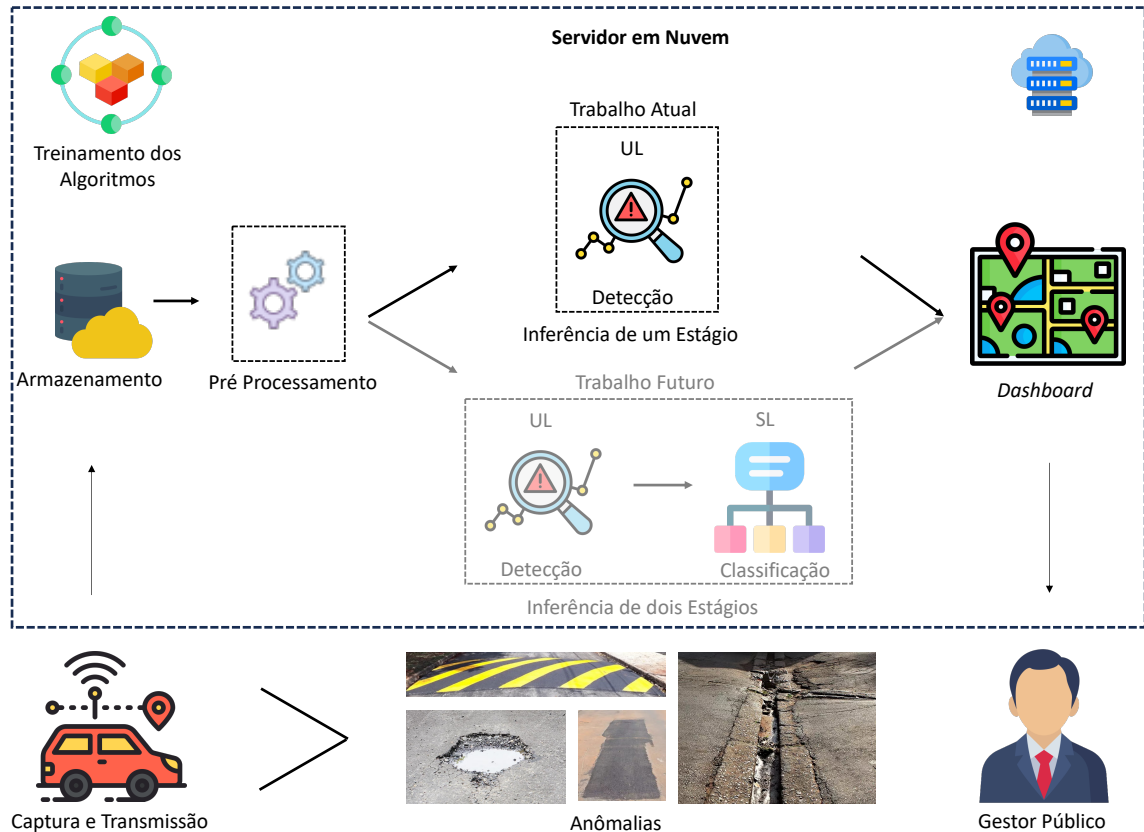


Figura 5 – Arquitetura para detecção de anomalias. (Fonte: Autor 2024)

Para atingir os objetivos detalhados anteriormente a atual pesquisa foi dividida em duas partes:

A primeira de fundamentação teórica traz a base da pesquisa, sendo realizado uma revisão bibliométrica para mapear os principais autores, publicações e áreas relacionadas com o tema de interesse, possibilitando gerar *insights* a partir de uma base de dados contendo inúmeros trabalhos, delimitando assim as condições de contorno e entendendo qual é o estado da arte. Logo após, a partir dos principais artigos encontrados pelo método anterior têm-se a revisão bibliográfica, sendo esta exclusivamente para compreensão e discussão dos principais métodos e técnicas realizadas pela academia.

A Segunda parte traz o programa experimental visando aplicar de forma prática os conceitos anteriormente obtidos, visto que a partir da instalação do sensor no carro são coletados dados de forma controlada para a criação de um *dataset* de anomalias para algoritmos SL e um *dataset* de não anomalias para algoritmos de UL, onde a coleta para esses dois tipos faz-se necessária apenas para comparação, pois o modelo utilizado para o módulo de detecção é baseado em *autoencoder* [27] e portanto pertencente aos algoritmos de UL. É importante ressaltar que todas análises realizadas e o fluxo descrito anteriormente pode ser aplicado para múltiplos carros, porém nesse trabalho apenas um veículo será utilizado para melhor observação e controle. Por fim, destaca-se que em

paralelo é realizado de forma analítica a definição da frequência de amostragem, gerando um método baseado em equações e na simulação oriunda do modelo denominado quarto de carro (Do inglês *quarter car model*) [28], possibilitando uso futuro dessa técnica por arquitetos de soluções para detecção e classificação de anomalias em pavimentos.

1.5 Estrutura da Dissertação

O capítulo 1, acima apresentado, traz a contextualização da pesquisa, a justificativa, os objetivos a serem alcançados e uma breve análise da metodologia empregada para a realização do trabalho.

O capítulo 2 visa prover uma base teórica para a pesquisa, analisando não apenas os principais autores e publicações, mas também as tendências e vieses na literatura. Desse modo para a contextualização do estado da arte de forma quantitativa é realizado uma revisão bibliométrica junto de suas análises e conclusões.

O capítulo 3 realiza uma extensa revisão bibliográfica, continuando o estudo realizado no capítulo 2, só que de forma aprofundada.

O capítulo 4 detalha os métodos utilizados no programa experimental, explicitando estratégias, algoritmos, técnicas e dados utilizados para o cumprimento dos objetivos.

O capítulo 5 apresenta os resultados e discussões a fim de se comprovar as conjecturas previamente assumidas, validar a arquitetura do sistema definido na metodologia e assegurar que o método proposto para a definição da frequência de amostragem pode ser aplicado de forma prática.

Por fim, o capítulo 6 traz as considerações finais e sugestões para trabalhos futuros.

2 REVISÃO BIBLIOMÉTRICA

Com o intuito de descobrir padrões de pesquisa, principais autores, áreas correlatas e subáreas do tema de interesse foi realizado uma revisão bibliométrica utilizando os metadados gerados nas principais plataformas de revisão sistemática da academia.

Essa abordagem pode ser utilizada em complemento da revisão bibliográfica, guiando o estudo a partir da identificação dos principais tópicos e potenciais direções para uma determinada área [29].

Diversos autores em diferentes campos de atuação recorrem a revisão bibliométrica para mapear o cenário de sua pesquisa em um determinado período e encontrar o estado da arte daquele assunto, como é o caso de [30] para a área de machine learning em econometria e [31] para a área de GIS.

Isto posto, neste trabalho para a realização de tal revisão foi utilizado a plataforma Scopus e WoS (*Web of Science*) para pesquisar palavras chaves do tema de interesse por meio de *queries* contendo operações *booleanas* a fim de se obter informações relevantes de cada publicação como o ano, autor, palavras relacionadas, etc., sintetizando assim em uma estrutura tabular um grande volume de dados que não seria possível ser analisado sem este método pré definido. Logo após, a partir dessa base gerada têm-se a exploração dos dados por meio do software VOSviewer, possibilitando realizar operações de agrupamento para diversos propósitos.

A Figura 6 mostra o método de pesquisa realizado, estando este dividido em duas partes (Fase 1 e Fase 2), permitindo segmentar o campo de estudo em área e sub área de interesse, onde em um primeiro momento uma busca com palavras chaves genéricas foi realizada para se ter uma compreensão macro do cenário atual, obtendo assim novas palavras que se relacionam fortemente com a área de pesquisa, obtendo então as tecnologias, métodos, processos e algoritmos relacionados.

Logo após, novas buscas menos abrangente foram realizadas nas bases indexadas com o intuito de se encontrar os principais autores, trabalhos realizados e tendências de publicações, gerando não apenas insumos para a fase de revisão bibliográfica, mas também *insights* para a continuação do trabalho.

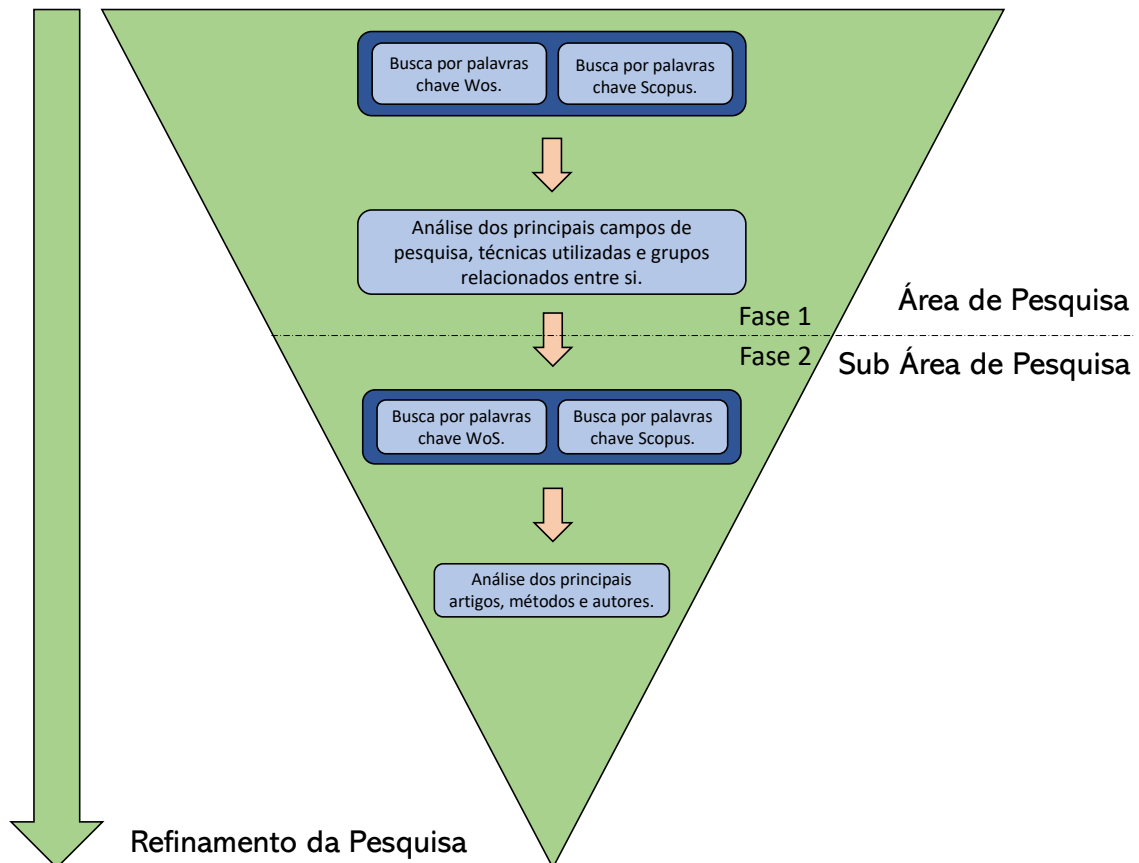


Figura 6 – Revisão bibliométrica - Método de pesquisa. (Fonte: Autor 2024)

2.1 Fase 1

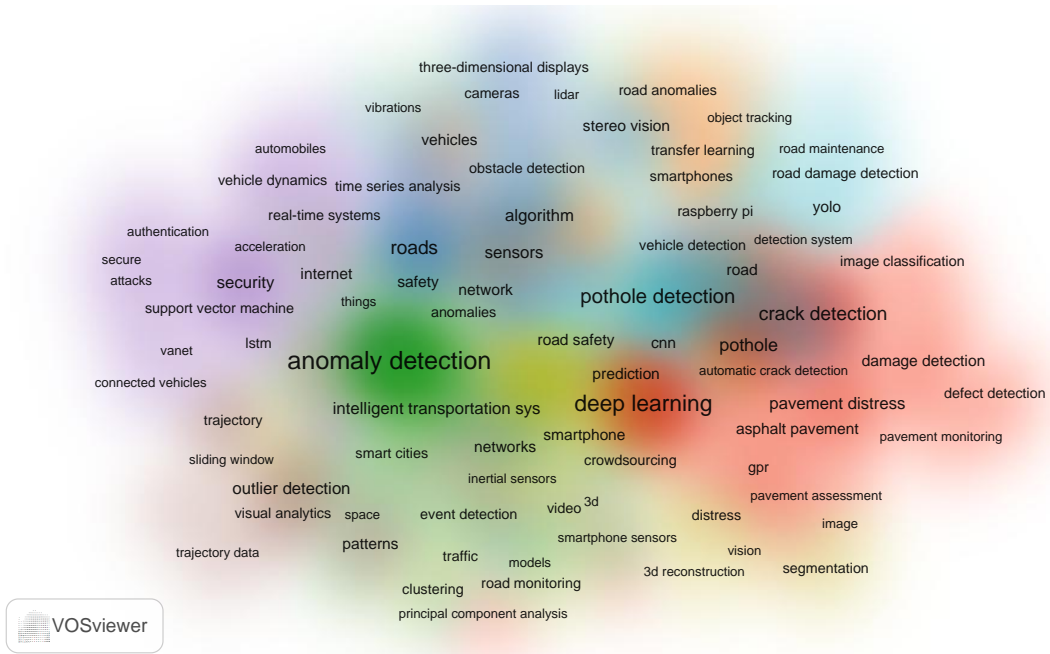
Para a fase 1 foi utilizado as palavras chave “road”, “pavement”, “street”, “anomaly”, “pothole”, “bump”, “gutter”, “detection”, “classification”, estando estas organizadas em formato de *query* (vide Apêndice A) para busca nas bases mencionadas, onde o resultado é apresentado na Tabela 1.

Tabela 1 – Metadados da primeira pesquisa para revisão bibliométrica. (Fonte: Autor 2024)

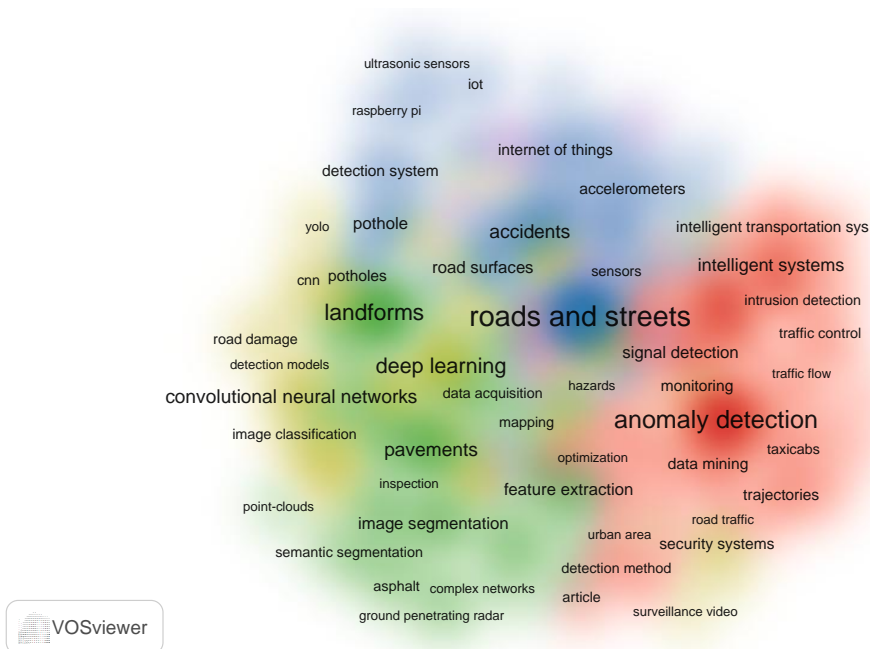
Base de dados indexada	WoS	Scopus
Período da pesquisa	1989-2023	1997-2023
Número de Artigos	1068	1369
Número de Autores	3808	4796
Idioma	Inglês	Inglês
Data da pesquisa	12 de Outubro de 2023	

Com os dados em mãos pode-se observar de imediato através do software VOSviewer as 250 principais palavras relacionadas com o tema de interesse, estando estas separadas em grupos de similaridade conforme demonstrado na Figura 7.

Analisando apenas o agrupamento de tais expressões podemos observar 9 grupos para os dados da WoS e 5 grupos formados a partir do resultado da base Scopus como mostra a Figura 8.



(a) Metadados da WoS



(b) Metadados da Scopus

Figura 8 – Agrupamento das palavras relacionadas processadas pelo software VOSviewer. (Fonte: Autor 2024)

Os grupos formados pelas palavras conglomeradas (vide Apêndice B) mostram diferentes vertentes de pesquisa, onde pode-se verificar que expressões como “*image seg-*

mentation”, “*stereo vision*”, e “*accelerometers*” estão localizadas em conjuntos diferentes, nos dando assim um indício de que são métodos distintos para o mesmo propósito (detecção de anomalias em vias).

2.2 Fase 2

Para a fase 2 uma busca mais restrita foi realizada conforme descrito na Tabela 2, visando analisar o campo de detecção de anomalias a partir de métodos baseados em vibrações, sendo utilizado as palavras chave “*road*”, “*pavement*”, “*street*”, “*anomaly*”, “*pothole*”, “*bump*”, “*gutter*”, “*vibration*” e “*accelerometer*”, estando estas organizadas em formato de *query* (vide Apêndice A) para a realização da busca nas bases anteriormente mencionadas.

Tabela 2 – Metadados da Segunda Pesquisa para revisão bibliométrica. (Fonte: Autor 2024)

Base de dados indexada	WoS	Scopus
Período da pesquisa	1996-2023	1997-2023
Número de Artigos	463	598
Número de Autores	1477	1977
Idioma	Inglês	Inglês
Data da pesquisa	13 de Outubro 2023	

A partir dos dados obtidos nas bases indexadas podemos inferir quais são os principais artigos já publicados nesse campo, onde essa análise tem grande valia, pois tais artigos e autores serão utilizados como referência ao decorrer da pesquisa.

A Figura 9 mostra um mapa temporal, relacionando a data de publicação dos principais artigos com o número de citações que eles tiveram até então, onde os trabalhos mais citados aparecem com um tamanho maior.

Como podemos ver, artigos sobre aplicações que utilizam sensores inerciais para a detecção de anomalias (Erikson et al. [32] e Mohan et al. [33]) são os que tem mais citações nas duas bases, porém é possível verificar também trabalhos sobre o estudo da dinâmica e controle de suspensões (Sun et al. [34], Nguyen et al. [35] e Du et al. [36]), sendo estes deveras importante para a fase de definição da frequência de amostragem, dado que o entendimento e modelamento matemático do fenômeno é crucial.

Um fato interessante é que dentre os artigos mais citados o mais recente data de 2012 [37], o que é estranho em um primeiro momento dado a inovação do tema, entretanto ao observar a fundo estas publicações percebe-se que as técnicas empregadas para a detecção de anomalias não são técnicas recentes, mas sim métodos clássicos para processamento de dados.

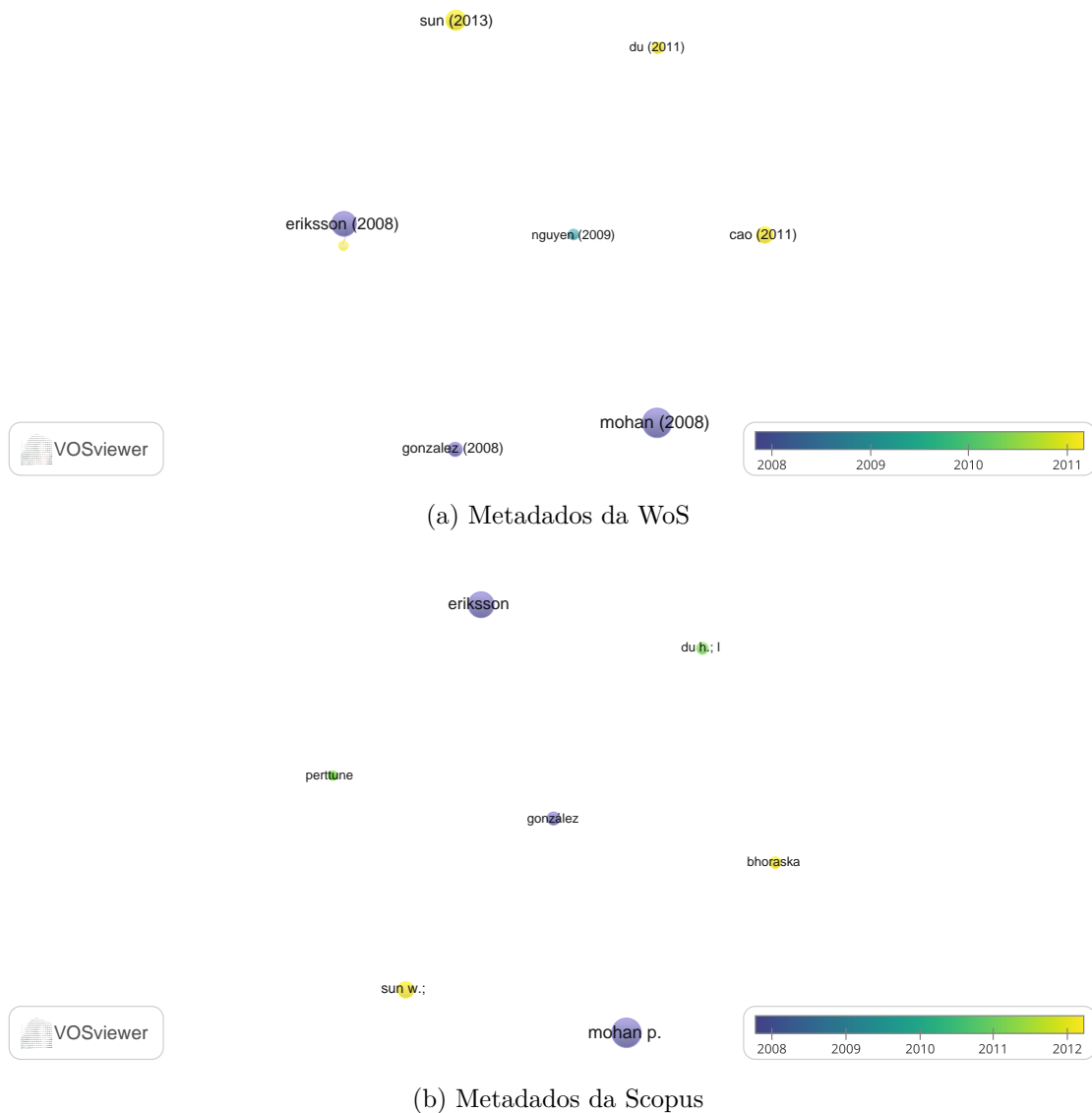


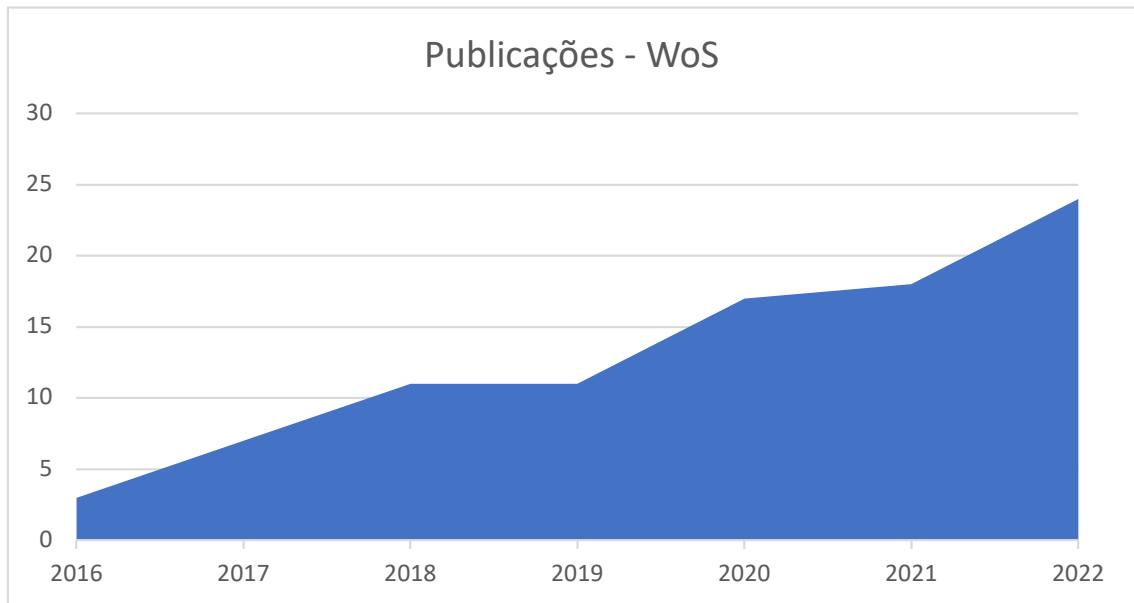
Figura 9 – Principais publicações processadas pelo software VOSviewer. (Fonte: Autor 2024)

Desse modo uma nova busca foi realizada com o propósito de se obter técnicas atuais, adicionando então novos termos como “*machine learning*” e “*deep learning*”, estando estes organizados em formato de *query* (vide Apêndice A), onde o resultado da pesquisa pode ser encontrado na Tabela 3.

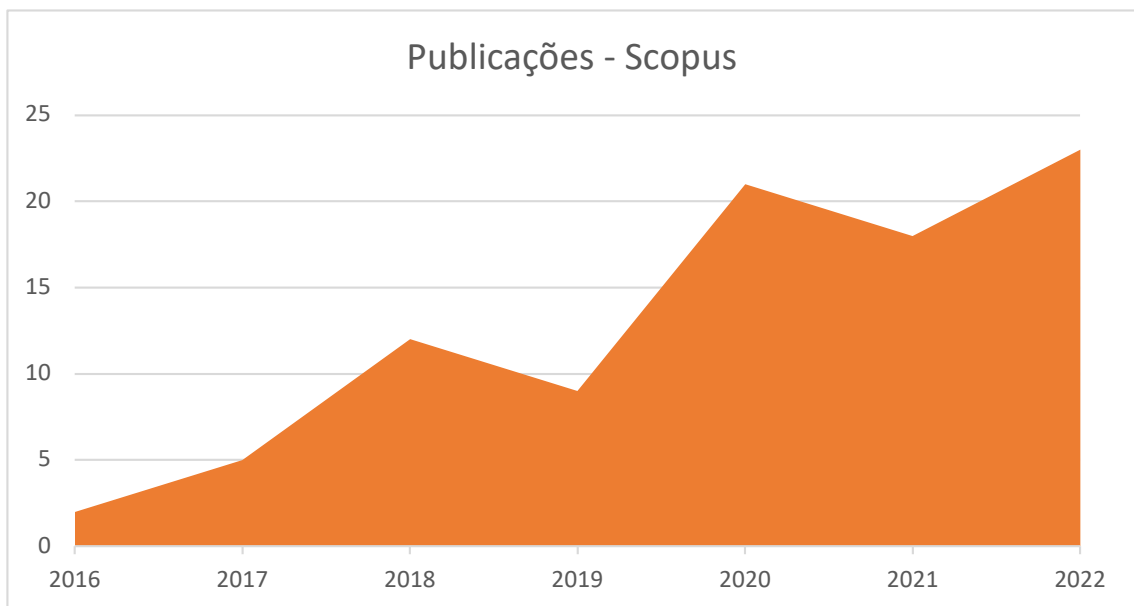
Tabela 3 – Metadados da terceira pesquisa para revisão bibliométrica. (Fonte: Autor 2024)

Base de dados indexada	WoS	Scopus
Período da pesquisa	2008-2023	2009-2023
Número de Artigos	112	111
Número de Autores	462	399
Idioma	Inglês	Inglês
Data da pesquisa	13 de Outubro de 2023	

De fato tais métodos vem sendo adotados de forma acelerada ao longo dos últimos 7 anos como mostra a Figura 10.



(a) WoS



(b) Scopus

Figura 10 – Frequência de publicação de artigos de detecção de anomalias em vias através de sensores de vibração utilizando machine learning. (Fonte: Autor 2024)

Entretanto o número baixo de publicações para esse tema evidencia que estamos na fronteira do conhecimento.

2.3 Conclusão da Revisão Bibliométrica

Em suma podemos verificar que o campo de estudo aqui apresentado é vasto e complexo, onde o mesmo se apoia em variadas tecnologias, como câmeras que capturam imagens 2D/3D, sensores que geram dados na forma de séries temporais, etc., nos dando assim inúmeros caminhos a serem seguidos. Não apenas tecnologias, mas diversas áreas do conhecimento são base para esse campo, como as ciências da computação que fornece algoritmos para a fase de análise dos dados e a engenharia mecânica que através de análises do perfil de vibração de regiões do automóvel ajuda na etapa de estudo e entendimento do fenômeno.

Indo de encontro com os objetivos anteriormente traçados vemos que nos últimos anos os algoritmos de aprendizado de máquina vem ganhando cada vez mais espaço, dado o alto desempenho que essa abordagem tem em outras áreas, as tornando assim uma solução com elevado potencial para a pesquisa.

Por fim destaca-se que os artigos que foram apontados como relevantes, assim como outros ainda não mencionados serão analisados na revisão bibliográfica para se obter o cenário completo das contribuições existentes até então.

3 REVISÃO BIBLIOGRÁFICA

Em posse dos principais artigos é realizado então a revisão bibliográfica, estando essa dividida em três áreas de atuação conforme ilustrado na Figura 11, onde o conjunto central ilustra o tema principal e as regiões de interseção representam assuntos importantes para o estudo.

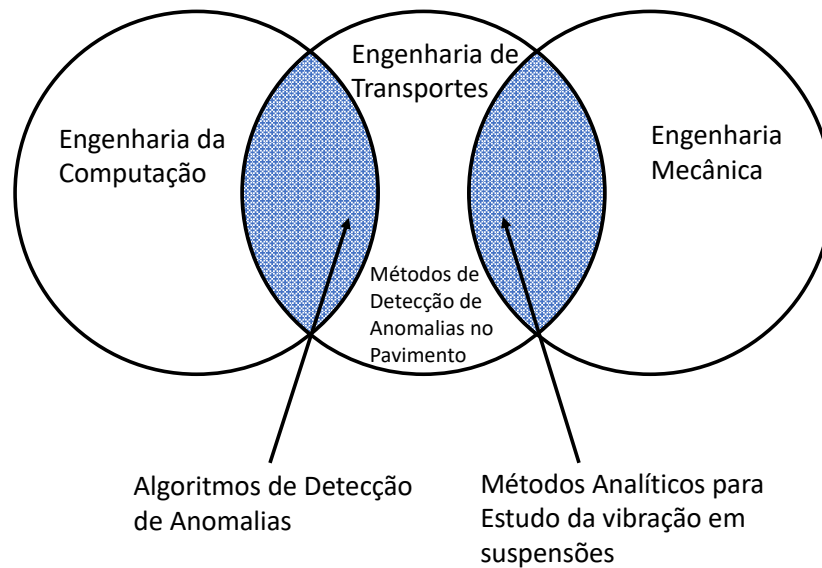


Figura 11 – Relação das áreas de interesse - Detecção de anomalias em pavimentos.
(Fonte: Autor 2024)

3.1 Algoritmos de detecção de anomalias

Detecção de anomalias (Do inglês "*Anomaly detection*"), também referenciado na literatura como "*novelty detection*" ou "*outlier detection*" é o método que possibilita identificar dados que destoam de um certo grupo, podendo estes representar padrões ou comportamentos incongruentes [38].

Liu et al [39] divide o campo de detecção de anomalias em: métodos baseados em distância, métodos baseados em densidade e métodos baseados em modelos, conforme evidenciado na taxonomia ilustrada na Figura 12 , onde diversos algoritmos de SL e UL podem ser alocados como subdivisões dessas três classes.

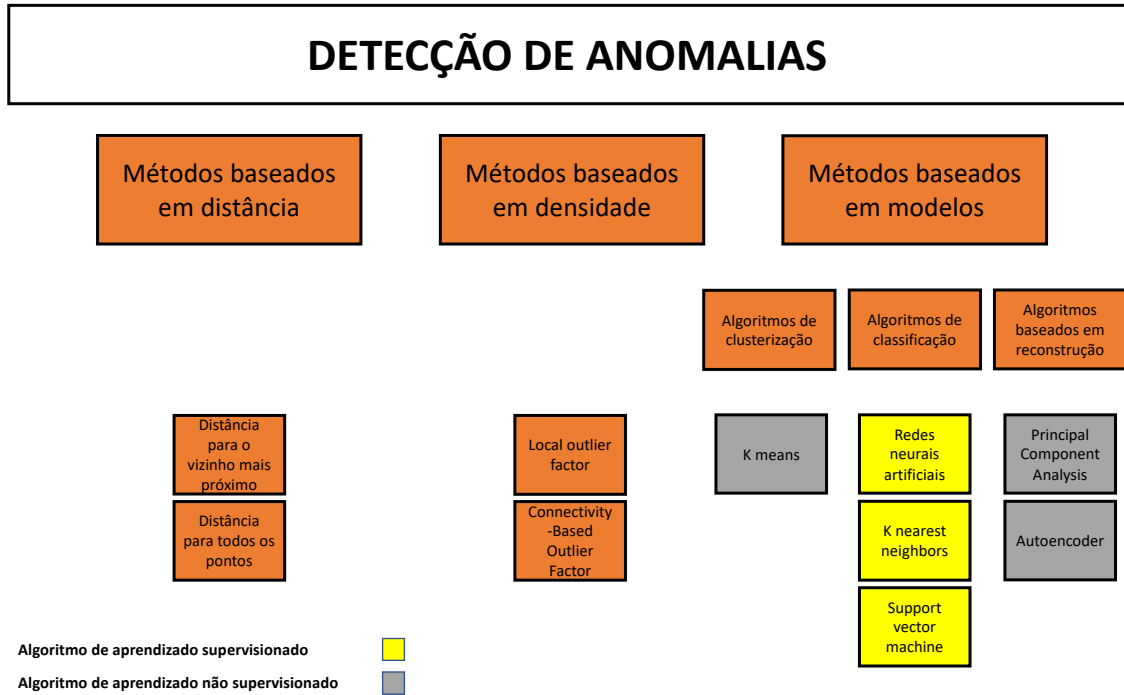


Figura 12 – Taxonomia - Detecção de anomalias. (Fonte: Autor 2024)

3.1.1 Métodos baseados em distância

Métodos baseados em distância detectam anomalias em um determinado conjunto de dados \mathcal{D} através da análise de métricas que relaciona cada ponto com os seus vizinhos, sendo a distância e a similaridade os critérios mais utilizados na literatura.

Pode-se definir a similaridade entre dois conjuntos de pontos p, q como o inverso de sua distância [40], avaliando assim quão próximo os dados estão, sendo a similaridade cosseno uma das mais difundidas no meio acadêmico, onde sua equação é expressa por:

$$s(p, q) = \frac{\sum_{i=1}^d p_i q_i}{\sqrt{\sum_{i=1}^d p_i^2 \sum_{i=1}^d q_i^2}} \quad (3.1)$$

Em relação a medida de distância, podemos elencar a distância de Minkowski, a distância de Manhattan, a distância Euclidiana e a distância de Chebyshev como sendo as métricas mais conhecidas.

A distância de Minkowski entre um conjunto de pontos p, q pode ser entendida como :

$$d(p, q) = \left(\sum_{i=1}^d |p_i - q_i|^l \right)^{\frac{1}{l}} \quad (3.2)$$

Sendo a Equação 3.2 genérica e de ordem $\frac{1}{l}$.

Para o caso em que $l = 1$ a distância de Minkowski é conhecida como distância de Manhattan:

$$d(p, q) = \left(\sum_{i=1}^d |p_i - q_i| \right) \quad (3.3)$$

Para $l = 2$ a distância de Minkowski é conhecida como distância Euclidiana :

$$d(p, q) = \sqrt{\sum_{i=1}^d (p_i - q_i)^2} \quad (3.4)$$

Por fim, no limite, quando $l \rightarrow \infty$ têm-se a distância de Chebyshev.

A Figura 13 ilustra a compreensão gráfica das diferentes métricas de distância.

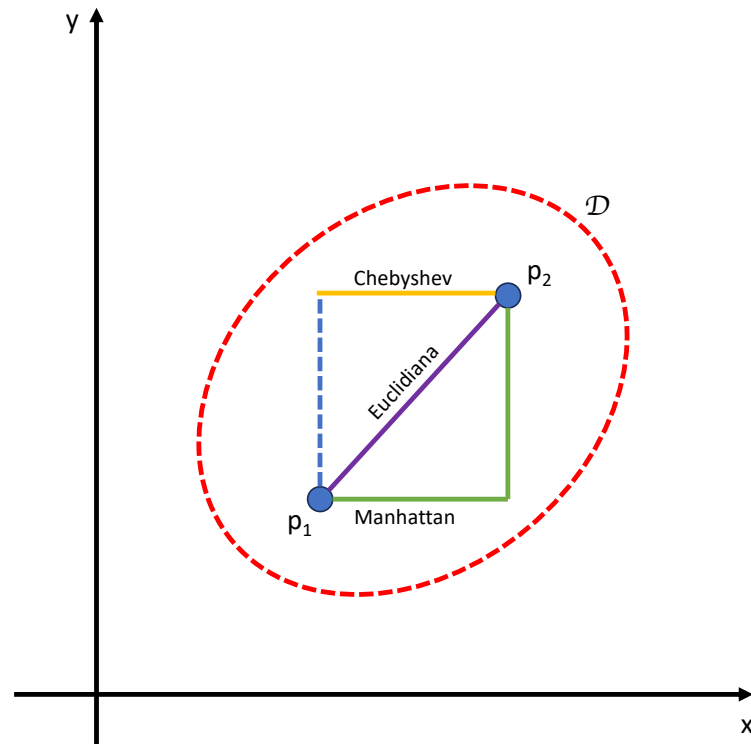


Figura 13 – Métricas de distância para um conjunto de dados contendo dois pontos (Fonte: Adaptado de [41] 2024)

Diversos métodos empregando diferentes métricas podem ser utilizados para detectar anomalias, ficando a cargo do executor analisar o conjunto de dados e escolher a abordagem apropriada.

Os principais algoritmos dessa área estão listados na Tabela 4, onde uma análise aprofundada pode ser encontrada em [40].

Tabela 4 – Principais algoritmos - Detecção de anomalias (Métodos baseados em distância). (Fonte: Autor 2024)

Algoritmo	Objetivo	Métrica
Distância para todos os pontos	Calcular a soma da distância de cada ponto para todos os pontos do conjunto de dados. O valor anômalo é o que possui o maior somatório.	$\alpha(p) = \sum_{q \in \mathcal{D}} d(p, q)$
Distância para o vizinho mais próximo	Calcular a distância para todos os vizinhos. O valor anômalo é o que possui a maior distância para o vizinho mais próximo.	$\alpha(p) = \min_{q \in \mathcal{D}, q \neq p} d(p, q)$

É importante ressaltar que mesmo não precisando de treinamento de máquina esses métodos requerem um grande poder computacional não sendo indicado para bases de dados com muitas características (alta dimensionalidade) [42].

3.1.2 Métodos baseados em densidade

Métodos baseados em densidade partem do princípio que pontos anômalos estão afastados de regiões de alta densidade [39]. Tais métodos são semelhantes aos métodos baseados em distância, porque utilizam métricas como a distância Euclidiana, herdando assim suas vantagens e desvantagens.

Um dos principais algoritmos nessa área é o LOF (*Local outlier Factor*) [43], sendo este um algoritmo que calcula quão esparsa cada ponto p em um conjunto de dados \mathcal{D} está em relação aos seus vizinhos. Entretanto este algoritmo falha quando a densidade do ponto anômalo é muito próxima da densidade de seus pontos adjacentes. Visando sanar tal problema Tang et al [44] propõe um algoritmo denominado COF (*Connectivity-Based Outlier Factor*) o tornando uma melhoria do algoritmo original.

3.1.3 Métodos baseados em modelos

Métodos baseados em modelos descrevem um determinado conjunto de dados \mathcal{D} por meio de modelos computacionais previamente treinados [40].

Este campo está dividido em : algoritmos de agrupamento, algoritmos de classificação e algoritmos de reconstrução.

3.1.3.1 Algoritmos de agrupamento

Um dos principais algoritmos de agrupamento é denominado *K means*, sendo este originário de técnicas do campo de processamento de sinal [45], onde a ideia central desse procedimento é agrupar j pontos p pertencentes a um conjunto de dados \mathcal{D} em k grupos

com centroides $C = \{c_1, \dots, c_k\}$ a partir de iterações, atualizando então o valor do conjunto C a cada repetição. Ao término desse processo é calculado a distância $d(p, c_l)$ e o ponto anômalo é o que possui o maior valor.

A Figura 14 ilustra o funcionamento do método.

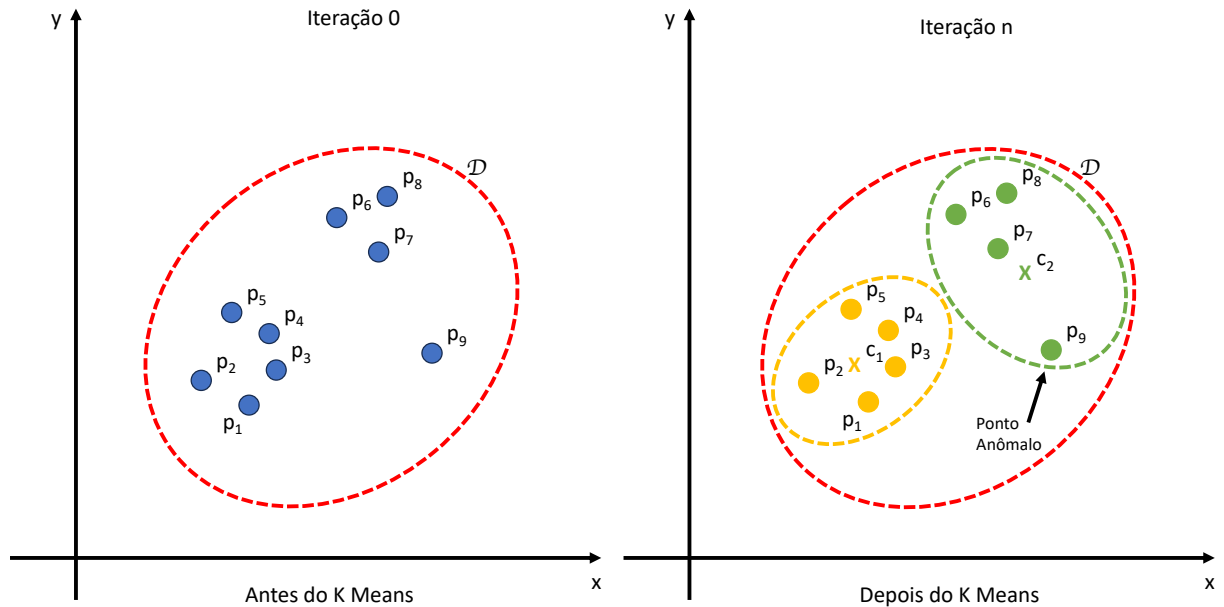


Figura 14 – Grupo de pontos sem classificação prévia, categorizados em dois grupos pelo algoritmo, onde a anomalia é definida como sendo o ponto que tem a maior distância até os centroides. (Fonte: Autor 2024)

3.1.3.2 Algoritmos de classificação

Algoritmos de classificação para detecção de anomalias utilizam técnicas tradicionais de categorização a fim de distinguir a classe anômala da classe não anômala, sendo comumente utilizado algoritmos de SL, dado sua vasta gama de possibilidades.

Neste contexto destacam-se os algoritmos, KNN (*K Nearest Neighbor*), SVM (*Support Vector Machine*) e ANN (*Artificial Neural Network*).

O algoritmo denominado KNN [46] utiliza métricas de distância para determinar a classe de um novo ponto p a partir de um conjunto de dados \mathcal{D} com p_j pontos previamente rotulados, onde é calculado a distância $d(p, p_j)$ e a classe desse novo ponto é definida pela classe de maior ocorrência dos k vizinhos mais próximos.

A Figura 15 ilustra o funcionamento do algoritmo.

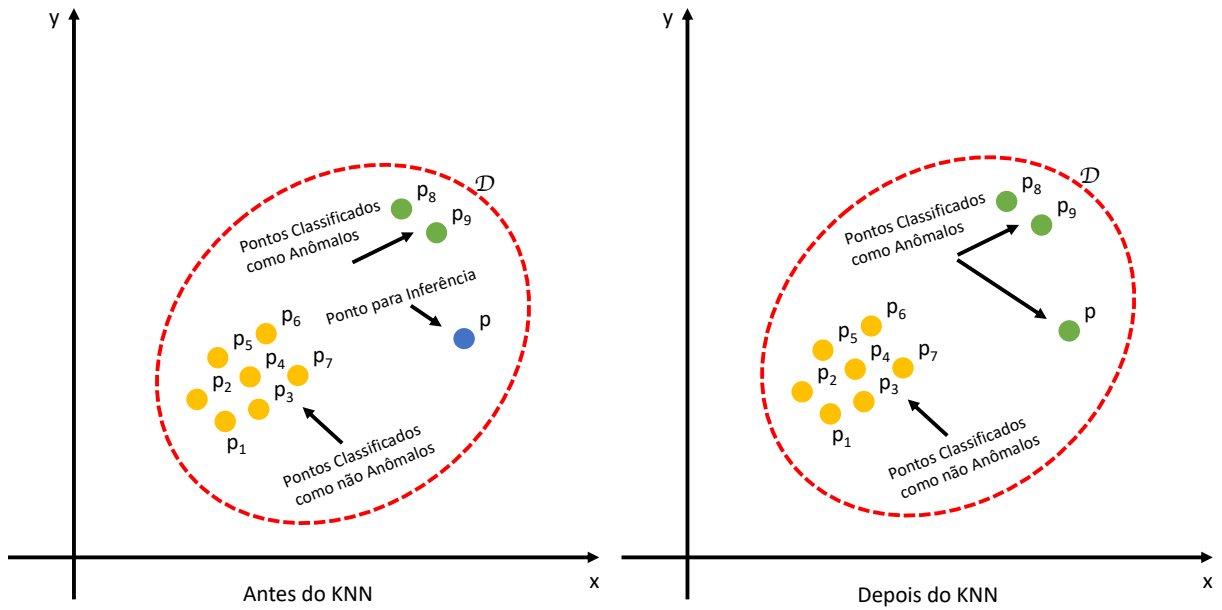


Figura 15 – Demonstração da inferência de um novo ponto p através do algoritmo KNN. (Fonte: Autor 2024)

Já o algoritmo SVM [47] pode ser utilizado tanto para regressão como para classificação, tendo como objetivo central encontrar um hiperplano que melhor separa um conjunto de dados pré rotulados \mathcal{D} no espaço de características a partir de uma função denominada *kernel*, criando assim uma borda de decisão como a ilustrada na Figura 16.

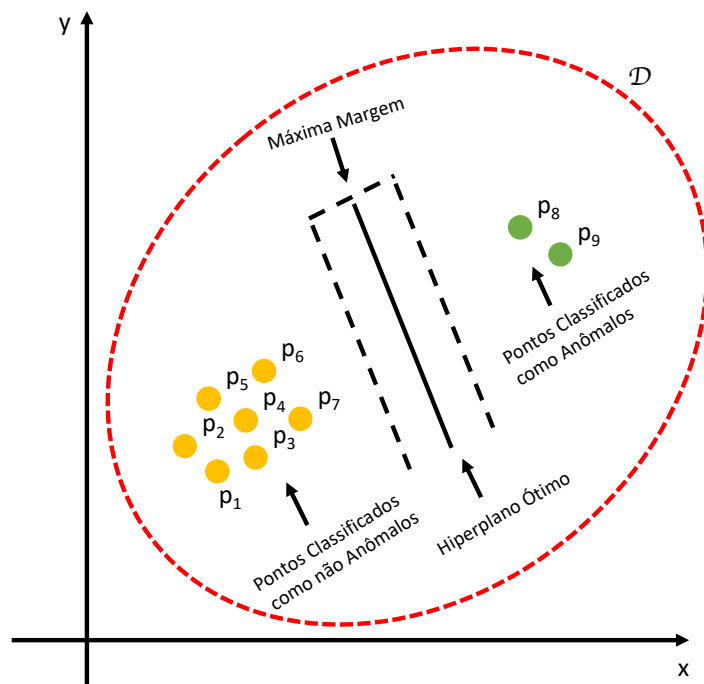


Figura 16 – Pontos separados em dois grupos pelo hiperplano ótimo após a utilização do algoritmo SVM. (Fonte: Autor 2024)

Uma variante importante desse algoritmo é o OCSVM (*One Class Support Vector Machine*), sendo este de UL amplamente utilizado para detecções de anomalias [48], onde seu grande diferencial está no fato de que este foi projetado para trabalhar com conjuntos de dados altamente desbalanceado, o que é essencial para essa aplicação, dado que o evento anômalo é raro [40], o que dificulta o processo de coleta de *datasets* para treinamento de máquina e gera um alto índice de FP na fase de inferência para os modelos tradicionais de classificação.

Diferente dos outros exemplos, o OCSVM tem em seu treino apenas amostras consideradas normais, excluindo assim a anomalia dessa etapa, fazendo então com que o modelo aprenda o padrão do conjunto não anômalo, delimitando os dados com uma borda de decisão otimizada definida por um *kernel*, onde qualquer ponto que caia fora desse hiperplano é identificado como anomalia.

Por fim pode-se destacar que as ANN's são técnicas computacionais de inteligência artificial inspiradas em redes neurais biológicas, tendo assim a capacidade de gerar um modelo matemático que compreende padrões e associações complexas a partir de um conjunto de dados de treino.

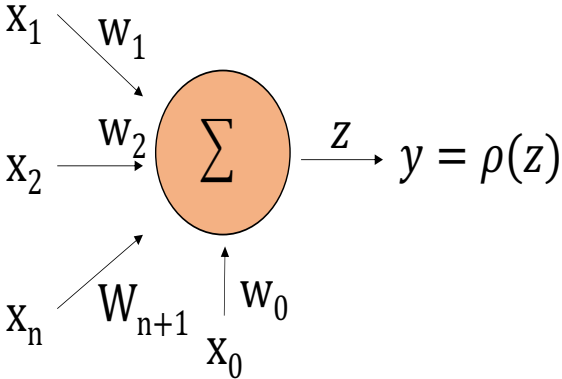
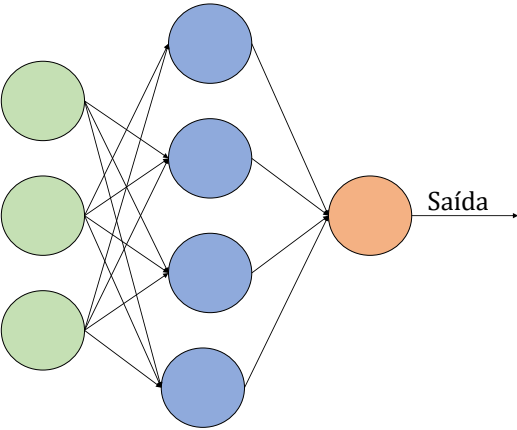
Esse tipo de algoritmo possui inúmeras variantes que podem ser combinadas com diversos métodos [49], não se limitando apenas a classificação mas podendo também resolver problemas de regressão, agrupamento, reconstrução, otimização, etc.

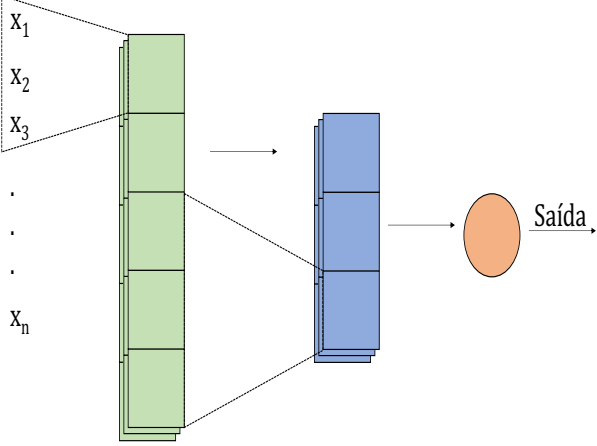
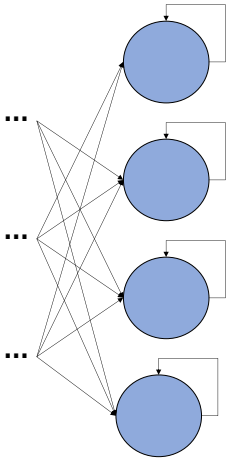
De forma genérica uma ANN para classificação de anomalias recebe um conjunto de amostras pré rotuladas $X = \{x_1, \dots, x_n\}$ e tem como principal objetivo encontrar um conjunto ótimo de pesos $W = \{w_1, \dots, w_{n+1}\}$ a partir de um processo denominado treino, gerando assim previsões binárias para novos dados de entrada. Independente da topologia trabalhada a saída desses modelos contam com um neurônio binário $\hat{y} = \{0, 1\}$ definido pela função *sigmoid* $\rho(z)$.

Quando projetadas para trabalhar com detecção de anomalias as ANN's enfrentam os mesmos problemas mencionados anteriormente, principalmente em relação ao desequilíbrio das classes, necessitando uma maior cautela na fase de concepção.

A Tabela 5 apresenta as ANN's mais utilizadas para detecção de anomalias a partir de classificação, sendo elas do tipo *perceptron*, MLP (*Multi Layer Perceptron*), CNN (*Convolution Neural Network*) e RNN (*Recurrent Neural Network*), todas pertencentes ao campo de SL.

Tabela 5 – Principais algoritmos - Redes neurais artificiais. (Fonte: Autor 2024)

Algoritmo	Descrição	Topologia
<p data-bbox="286 563 439 595"><i>Perceptron</i></p>	<p data-bbox="512 443 1256 719">O perceptron [50] é uma ANN que possui apenas uma camada e um neurônio, sendo este um modelo elementar que serve como base para arquiteturas mais complexas. A partir de um vetor X de n valores, é possível determinar um conjunto W de pesos de modo que a saída y seja aproximada do resultado esperado.</p>	
<p data-bbox="327 1066 398 1098">MLP</p>	<p data-bbox="512 826 1256 1337">A ANN do tipo MLP possui uma topologia com neurônios totalmente conectados entre si, podendo conter diversas camadas. A informação passa por esse sistema em apenas um sentido, por isso essa rede também é denominada de <i>feedforward</i> pois não há uma realimentação. Hornik et al [51] demonstrou que com apenas uma camada esse tipo de rede consegue aproximar qualquer função contínua, provando que MLP's são aproximadores universais, ou seja, essa topologia é capaz de representar funções complexas por meio dados.</p>	

Algoritmo	Descrição	Topologia
CNN	<p>ANN convolucionais [52] são redes especializadas para processamento de séries temporais, imagens e audio.</p> <p>Essa arquitetura possui tal terminologia devido as operações realizadas em suas camadas, onde <i>kernels</i> (filtros cujo parâmetros são definidos em treinamento) são aplicados nos dados de entrada em uma operação denominada de convolução, extraindo as informações importantes e propagando-as para as camadas subsequentes. Quando aplicada para classificação a saída dessa rede pode conter um ou mais neurônios do tipo <i>perceptron</i>.</p>	 <p>The diagram illustrates the topology of a Convolutional Neural Network (CNN). It shows an input layer on the left with nodes labeled $x_1, x_2, x_3, \dots, x_n$. These nodes are connected to a hidden layer in the middle, represented by a stack of blue rectangular blocks. An arrow points from the input layer to the hidden layer. From the hidden layer, another arrow points to an output layer on the right, which consists of a single orange oval node labeled 'Saída'.</p>
RNN	<p>ANN recorrentes [53] ao contrario das MLP's possuem uma arquitetura que permite a realimentação (<i>feedback</i>) dos neurônios, conectando assim, a saída de um estado com sua entrada. O diferencial dessas redes é a capacidade de processar dados de iterações passadas, possuindo então uma poderosa habilidade denominada de memória, o que torna essa topologia ideal para conjunto de dados que tenham dependência temporal. Ademais, existem diversas variações dessa arquitetura, onde as mais conhecidas são a LSTM (<i>Long Short Term Memory</i>) e GRU (<i>Gated Recurrent Unit</i>).</p>	 <p>The diagram illustrates the topology of a Recurrent Neural Network (RNN). It shows a sequence of four blue circular nodes arranged vertically. Each node is connected to a common set of input nodes on the left, indicated by three vertical ellipses (...). Each node also has a feedback loop from its output back to its input, represented by a curved arrow on the right side of each node. The nodes are connected sequentially from top to bottom.</p>

3.1.3.3 Algoritmos baseados em reconstrução

Algoritmos de detecção de anomalias baseados em reconstrução assumem que é possível replicar os valores de entrada do modelo em sua saída a partir de funções de base [54], de modo que o sistema aprenda o que é não anomalia dado um conjunto de treino \mathcal{D} com dimensão m (m características) contendo apenas amostras normais. Ao fim do processo espera-se que novos pontos anômalos p_a sejam detectados pelo algoritmo devido ao alto erro de reconstrução que haverá frente a incapacidade do sistema treinado lidar com dados desse tipo.

Uma das técnicas mais utilizados nessa área chama-se PCA (Análise das Componentes Principais - Do inglês *Principal Component Analysis*) [55], na qual o algoritmo de UL tem como finalidade reduzir a dimensionalidade do conjunto de treino e então reconstruir o subespaço original através das componentes principais S_{normal} , obtendo assim um erro de reconstrução associado a cada amostra, onde idealmente os pontos anômalos terão mais dificuldade de projeção, o que elevará seu valor de erro.

Desse modo, para quantificar o erro entre o valor real e valor reconstruído pode-se empregar diversas métricas, como o MSE (Erro Médio Quadrático - Do inglês *Mean Square Error*) ou RMSE (Raiz do Erro Médio Quadrático - Do inglês *Root Mean Square Error*), cujas equações são definidas como:

$$MSE = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \hat{x}_{ij})^2 \quad (3.5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \hat{x}_{ij})^2} \quad (3.6)$$

sendo:

x = valor de entrada

\hat{x} = valor de saída

m = numero de características (dimensões) do *dataset* de treino

n = numero de amostras

A Figura 17 ilustra o funcionamento deste algoritmo, onde pode-se observar no momento da inferência a alta dispersão dos pontos anômalos no subespaço reconstruído, o que não acontece no subespaço original, elevando assim o erro entre entrada e saída para esses pontos.

Nota-se também que o subespaço denominado S_{normal} tem como finalidade representar os dados de treino (não anômalos) de forma condensada nas componentes principais.

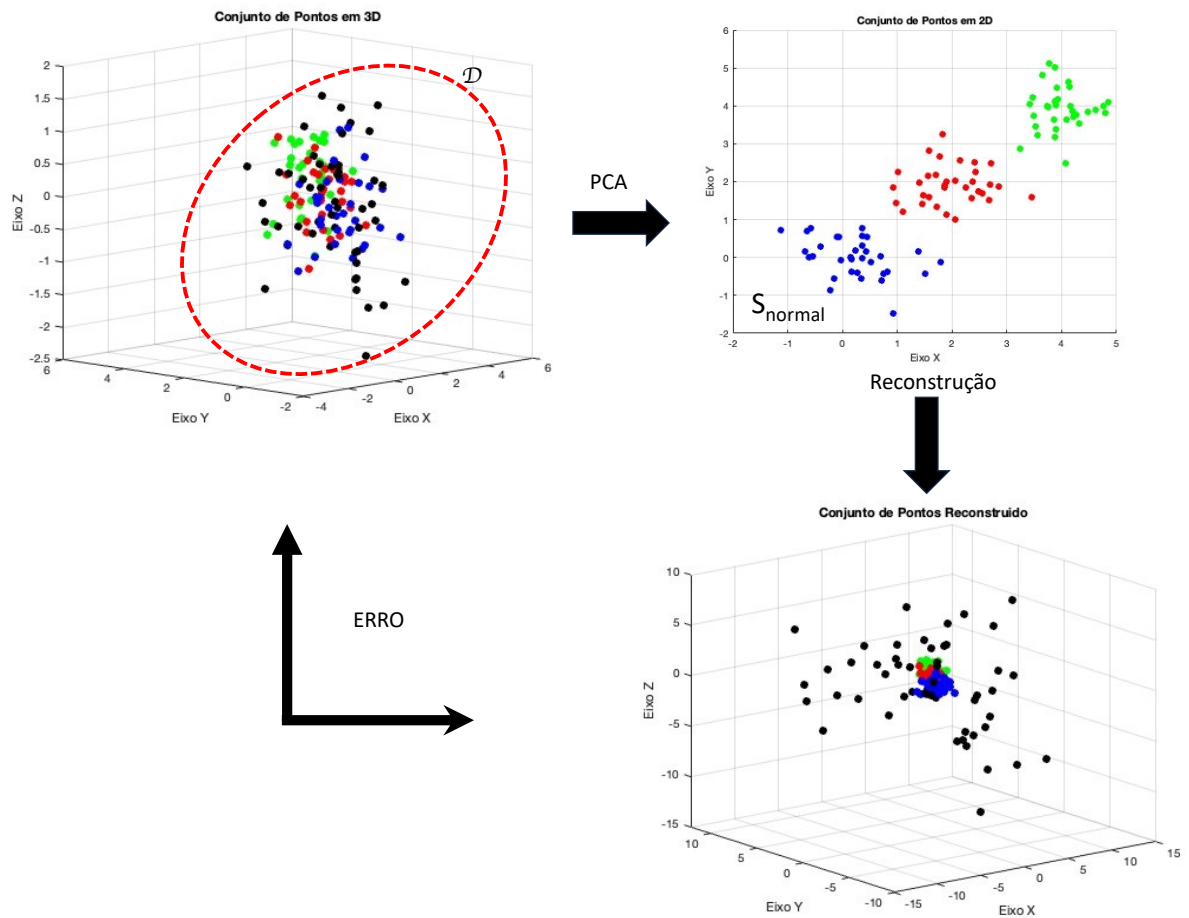


Figura 17 – Representação do funcionamento do algoritmo PCA, onde os dados normais de treino podem ser observados em verde, vermelho e azul e as anomalias que são introduzidas na inferência estão na cor preta. (Fonte: Autor 2024)

Outra técnica amplamente utilizada são as ANN's do tipo *autoencoder*, que utilizam o conceito de *encoder* e *decoder* para reduzir e aumentar a dimensionalidade dos dados de treino em sua arquitetura [56], ou seja, diferente do algoritmo PCA, esse método utiliza neurônios artificiais para fazer esse processo, o que em termos práticos se traduz em diminuir a quantidade de nós em cada camada, concentrando as principais características do conjunto em uma região denominada gargalo e a partir desse ponto aumentar simetricamente as unidades neurais, reconstruindo novamente os dados, de modo que possa se recuperar a entrada do sistema em sua saída.

Em redes desse tipo o objetivo do treino é minimizar o erro de reconstrução a partir de um conjunto de dados \mathcal{D} com n amostras não anômalas $\{x_1, x_2, \dots, x_n\}$, gerando uma rede ótima para os casos de treino.

A Figura 18 demonstra o fluxo de trabalho de ANN's do tipo *autoencoder*, onde as anomalias são detectadas a partir da ideia de compressão e reconstrução, tornando o gargalo uma região que representa os dados normais na forma comprimida.

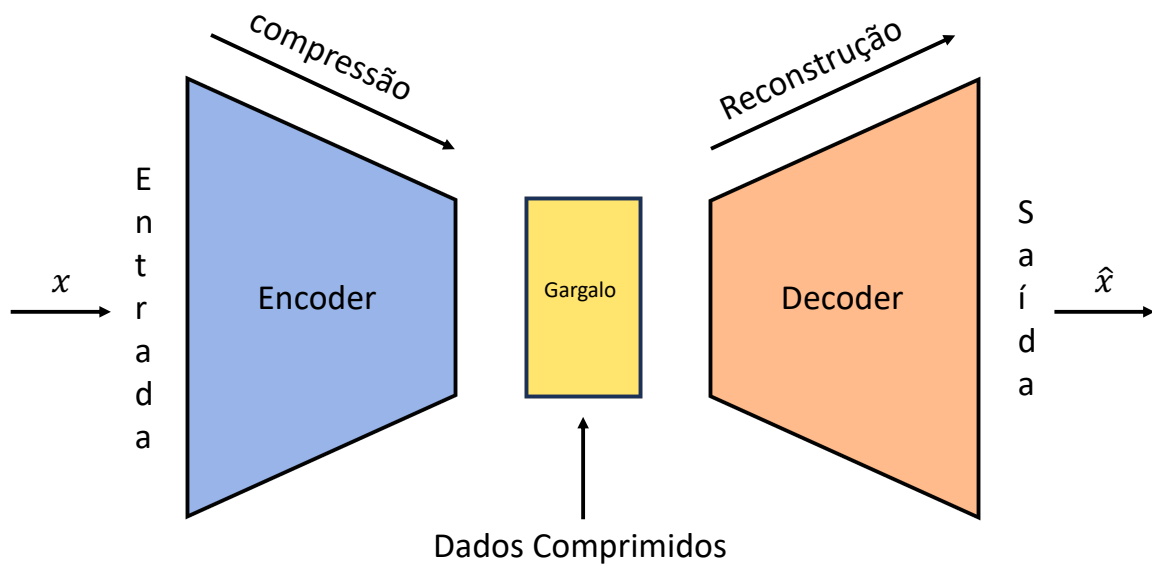


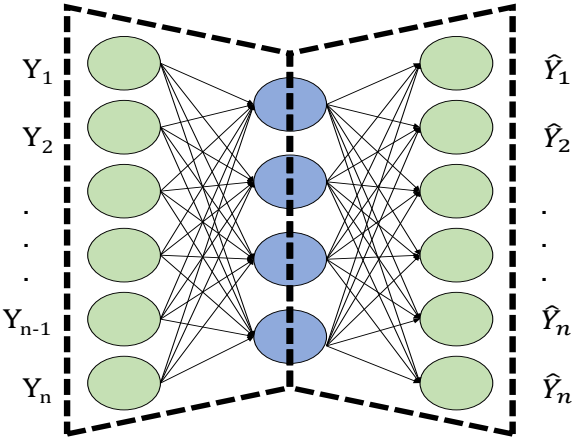
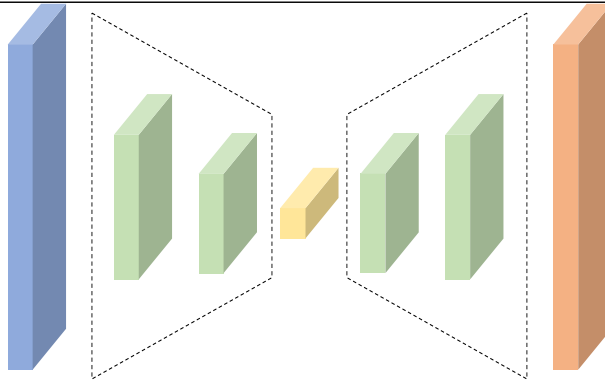
Figura 18 – Funcionamento de uma ANN *autoencoder*, onde o conceito fundamental dessa técnica é ilustrado por meio de um diagrama de blocos representando as camadas de *encoder*, gargalo e *decoder*. (Fonte: Autor 2024)

Idealmente, a rede treinada conseguirá reconstruir com um baixo erro os dados normais, o que não fará com os pontos anômalos (assim como acontece no algoritmo PCA), possibilitando detectar essas anomalias através de um limiar de erro pré definido, segregando assim as classes através de uma borda de decisão.

Pode-se elencar inúmeras variantes de ANN's que utilizam *autoencoder* para os mais variados propósitos, desde redução da dimensionalidade de um conjunto de dados para melhor compreensão e entendimento deles [57] a detecção de anomalias, onde a principal diferença de cada arquitetura está no tipo de neurônio utilizado e seu modo de ativação.

A Tabela 6 traz uma breve descrição das duas principais topologias de ANN's utilizadas na literatura para detecção de anomalias (*Vanilla autoencoder* e *LSTM autoencoder*), sendo estas pertencentes a classe de algoritmos de UL, uma vez que os dados de treino não contém uma classificação prévia.

Tabela 6 – Algoritmos baseados em reconstrução. (Fonte: Autor 2024)

Algoritmo	Descrição	Topologia
<p><i>Vanilla Auto Encoder</i></p>	<p>A ANN do tipo <i>vanilla autoencoder</i> é uma arquitetura elementar do campo de algoritmos de reconstrução, dado que sua estrutura conta com neurônios do tipo <i>perceptron</i> para realizar as operações em todas as camadas da rede, fazendo essa topologia parecer com a arquitetura MLP, com excessão que neste caso as camadas de <i>encoder</i>, gargalo e <i>decoder</i> devem seguir o arranjo anteriormente mencionado.</p>	
<p><i>LSTM Auto Encoder</i></p>	<p><i>Autoencoders LSTM</i> são ANN's projetadas para detecção de anomalias, onde os módulos de <i>encoder</i> e <i>decoder</i> são compostos por neurônios do tipo LSTM [23], o que possibilita a sua utilização em dados de natureza sequencial [58] como as séries temporais, uma vez que seus neurônios dotados de memórias de longo prazo possibilitam a aprendizagem de padrões temporais complexos.</p>	

3.2 Métodos de detecção de anomalias no pavimento

Os métodos de detecção de anomalias no pavimento utilizam os algoritmos anteriormente citados para a verificação da condição de logradouros, unindo assim áreas como engenharia da computação e engenharia de transportes, auxiliando na identificação de diversos elementos de interesse em vias urbanas a partir de diferentes tipos de sensores.

A revisão bibliográfica para este campo está dividida em três métodos bases conforme descreve [6], onde cada sub seção trará a análise das principais contribuições já publicadas até o presente momento.

3.2.1 Métodos baseados em visão computacional 2D

Métodos baseados em visão computacional 2D utilizam câmeras tradicionais para a detecção de anomalias no pavimento, onde tais câmeras podem ou não estar embarcadas em veículos, tendo como principal ganho a geração da evidência visual da anomalia no momento de sua detecção.

Conforme descreve Nachuan et al [59] a análise das imagens 2D captadas em campo pode ser realizada através de processamento clássico ou de aprendizado de máquina, possibilitando assim a utilização de diferentes algoritmos para a detecção dessas anomalias.

No processamento clássico técnicas para extrair informações da matriz de cor RGB (*Red, Green, Blue*) são aplicadas, normalmente na forma de um processamento sequencial [60] combinando diversas análises como filtro de cores, agrupamento espectral [61], limiar de contraste [62], extração de textura [63] etc.

Já os métodos que utilizam aprendizado de máquina são capazes não apenas de detectar que há anomalias, mas também localizá-las na própria imagem, circundando o objeto pixel a pixel [64] ou anotando-o por meio de um polígono [65], onde tradicionalmente algoritmos de SL são utilizados para esta tarefa [66].

Por fim, pode-se destacar que um problema intrínseco dos métodos baseados em visão computacional 2D é a incapacidade de se representar completamente a estrutura espacial do objeto de interesse, visto que uma dimensão é perdida nesse processo [67].

3.2.2 Métodos baseados em reconstrução 3D

Métodos baseados em reconstrução 3D (também conhecido por métodos baseados em visão estéreo) tem sido amplamente pesquisado nos últimos anos [59], abrindo novas possibilidades para o campo de detecção, uma vez que este viabiliza estimar o volume das anomalias no pavimento.

Para sua realização há três abordagens possíveis, estando elas divididas em: Escaneamento a laser, Sensores passivos e *Microsoft Kinect* [67].

Os métodos de escaneamento a laser utilizam equipamentos que emitem múltiplos feixes em diferentes comprimentos de onda e capturam as reflexões para construir o modelo 3D do objeto [68, 69], onde normalmente os dispositivos são embarcados em veículos dedicados [70], impossibilitando a construção de uma rede de coleta, dado o alto preço e tamanho dos equipamentos. Atualmente a utilização de sensores LIDAR (*Light Detection and Ranging*) tem sido uma inovação para área, uma vez que este tipo de sensor ocupa menos espaço e possui uma acurácia elevada [71], superando assim os problemas anteriormente elencados, porém as pesquisas nesse campo estão em fases embrionárias e necessitam serem validadas.

Por outro lado métodos oriundos de sensores passivos utilizam múltiplas câmeras tradicionais para a captura das três dimensões, o que faz desse método ser potencialmente utilizado em aplicações de campo, dado a sua facilidade de instalação e baixo custo comparado aos métodos anteriores. Seu funcionamento se dá a partir da geração de uma matriz de disparidade entre os equipamentos [68, 72] que permite reconstruir a nuvem de pontos do objeto detectado [73].

Os métodos que utilizam a câmera desenvolvida pela *Microsoft* através do sensor *Kinect* são semelhantes ao anterior, possibilitando não apenas detectar os objetos em 3D, mas também medi-los [74, 75], estimando suas dimensões por meio de inúmeros algoritmos, gerando uma gama de aplicações para análise do pavimento.

Em suma pode-se elencar que métodos que utilizam a reconstrução 3D tem grande potencial a longo prazo, dado as soluções que eles trazem para os problemas em questão, porém, atualmente eles são caros, possuem um tempo de processamento maior que os outros métodos devido a grande massa de dados gerada [76] e em alguns casos são menos efetivos em ambientes externos [68].

3.2.3 Métodos baseados em vibrações

Métodos baseados em vibrações utilizam sensores inerciais para a detecção de anomalias, onde estes podem ser acoplados em diferentes partes do automóvel para a realização da tarefa de forma dedicada (pela confecção de um *hardware* exclusivo) ou através da utilização de *smartphones* que possuam a tecnologia.

Conforme descreve Martinez et al [4] a detecção de anomalias nesse campo pode ser classificada em métodos baseados em limiar, métodos de extração de características ou métodos baseados em deep learning.

Métodos baseados em limiar são métodos diretos pois os mesmos não dependem de um treinamento de máquina ou processamento prévio dos dados para a detecção das anomalias, entretanto um grande desafio nesse campo é definir o valor numérico de tais fronteiras, visto que este valor é influenciado diretamente por diversos parâmetros como

a velocidade do automóvel e a posição do sensor [12].

Métodos que utilizam a extração de características de forma manual tendem a reduzir a dimensão do conjunto de dados (redução de *features*) persistindo assim apenas as informações relevantes para a aprendizagem de máquina, eliminando então as redundâncias e as características não importantes, entretanto tal método demanda bastante experiência do executor devido a sua complexidade [4].

Em contrapartida métodos baseados em deep learning dispensam a utilização de extração de característica de forma manual, possibilitando assim o treinamento de máquina com os dados na sua forma bruta [77].

Na última década, diversos artigos têm sido publicados utilizando os três métodos apresentados para essa área, conforme demonstrado no Capítulo 2, porém o que chama a atenção ao analisar os trabalhos é a disparidade em relação a frequência de amostragem dos sensores inerciais, uma vez que distintos valores são utilizados sem uma análise aprofundada, gerando em muitos casos sobre amostragem dos dados e inviabilizando uma futura utilização do sistema com múltiplos carros na forma *crowdsensing*. A Tabela 7 apresenta os principais trabalhos com suas respectivas taxas de amostragem, evidenciando assim o problema em questão.

Tabela 7 – Publicações que utilizam métodos baseados em vibrações para detecção de anomalias no pavimento. (Fonte: Autor 2024)

Artigo	Método Utilizado	Frequência de amostragem [Hz]
Fox et al. [78]	Extração de características	1
Badurowicz et al. [79]	Extração de características	10
Bansal et al. [80]	Extração de características	10
Gurdit et al. [81]	Limiar	10
Seraj et al. [82]	Extração de características	47
Varona et al. [83]	<i>deep learning</i>	50
Pandey et al. [84]	<i>deep learning</i>	100
Mednis et al. [85]	Limiar	100
Guangtao et al. [86]	Limiar	100
Wang et al. [87]	Limiar	124
Eriksson et al. [32]	Extração de características	380
Anaissi et al. [88]	Extração de características	500

3.3 Métodos analíticos para estudo da vibração em suspensões

Métodos para o estudo da dinâmica das suspensões de carros têm sido amplamente utilizados ao longo dos anos, principalmente para a análise e melhoria do conforto dos passageiros [89, 90, 91], onde técnicas oriundas da engenharia mecânica são aplicadas para a construção de um modelo matemático a fim de se realizar simulações e análises a posteriori. Dentre essas técnicas destacam-se os modelos de parâmetros concentrados (Do inglês *lumped models*) que assumem uma representação simplificada do sistema, concentrando diversos parâmetros em um único elemento através de métodos empíricos para a definição de seus valores [92], descrevendo assim o comportamento do conjunto por meio de equações de menor complexidade.

No contexto automotivo essa simplificação se dá majoritariamente pelo DOF (Grau de Liberdade - Do inglês *Degree of Freedom*) escolhido para modelar dinamicamente a suspensão, pois a medida que são adicionadas coordenadas independentes para a definição de seu estado ou posição, mais difícil se torna a análise, entretanto é importante ressaltar que este não é o único elemento que define o grau de simplicidade do modelo, havendo assim outros detalhes que devem ser analisados no estudo [93].

O modelo de 1DOF é o mais simples que pode ser utilizado para descrever a dinâmica de suspensões [94], sendo este representado por um sistema de massa M , mola k e amortecedor c se movimentando em apenas um eixo que descreve o movimento vertical do sistema. Sua arquitetura é demonstrada na Figura 19, onde pode-se observar seu alto nível de abstração, condensando diversas variáveis em apenas três parâmetros.

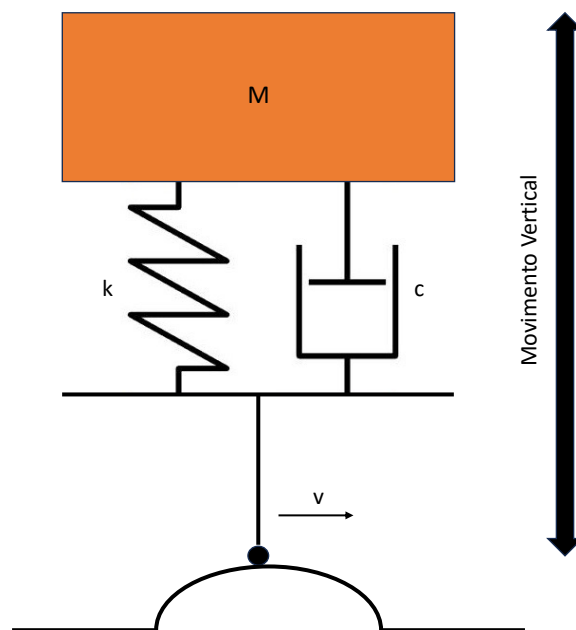


Figura 19 – Modelo 1DOF utilizado para representar um veículo passando sobre um obstáculo com uma velocidade longitudinal v . (Fonte: Autor 2024)

Já o modelo com 2DOF possui duas representações válidas [95], são elas: modelo quarto de carro e modelo metade de carro (Do inglês *half car model*), conforme ilustrado na Figura 20.

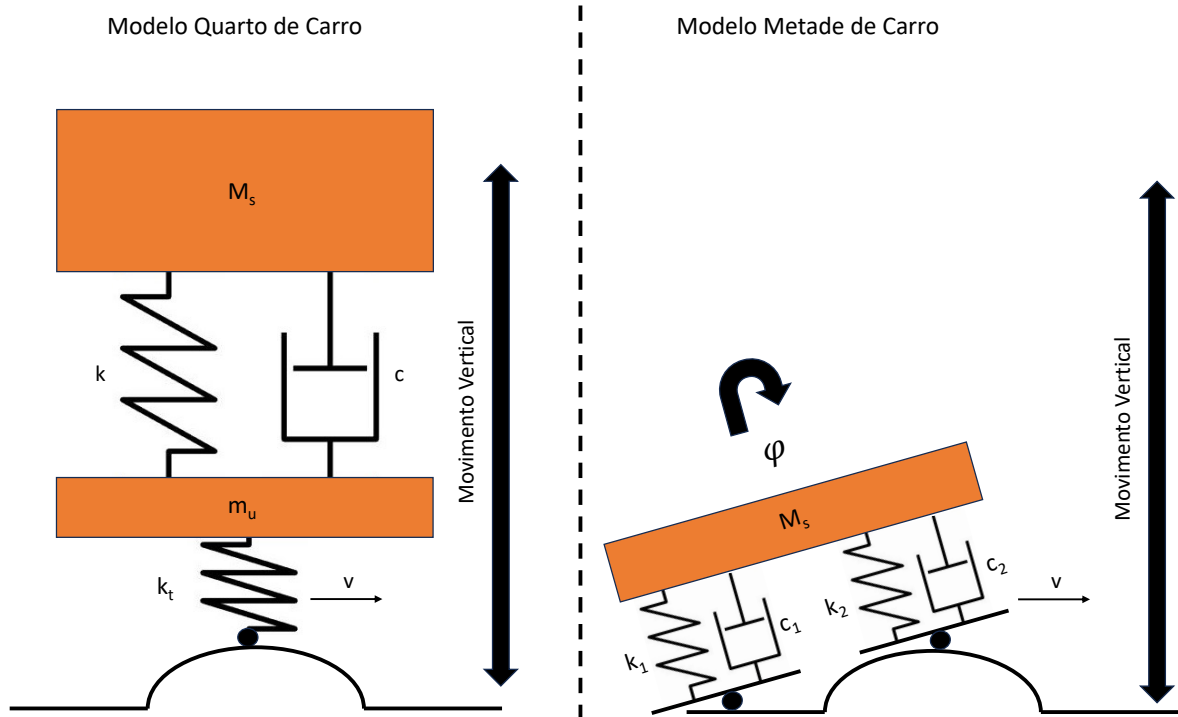


Figura 20 – Modelos 2DOF utilizado para representar um veículo passando sobre um obstáculo com uma velocidade longitudinal v . A esquerda o modelo quarto de carro e a direita o modelo metade de carro. (Fonte: Autor 2024)

O modelo quarto de carro assume que a massa suspensa M_s e a massa não suspensa m_u do veículo podem ser modelados como dois corpos rígidos, estando interconectados pela suspensão (mola k e amortecedor c), descrevendo a reação vertical do conjunto (*bounce movement*) de apenas um semi eixo. A rigidez do pneu é representado por uma mola k_t adicionada abaixo da massa suspensa.

Por outro lado o modelo metade de carro pode ser compreendido como um corte transversal no veículo que considera apenas a massa suspensa M_s , o que possibilita a análise da inclinação do automóvel (φ - *pitch motion*) além do movimento vertical do bloco. O sistema de amortecimento é representado pelo conjunto mola (k_1, k_2), amortecedor (c_1, c_2) para a roda dianteira e traseira respectivamente.

Por fim, pode-se elencar que diversos outros modelos de complexidade superior com nDOF pode ser utilizado para descrever inúmeros outros fatores que tem influência no movimento do veículo, como a massa dos passageiros, o motor, etc. [96], porém é importante ressaltar que o custo computacional para realizar tais simulações e resolver as equações diferenciais oriundas do modelo cresce consideravelmente [93].

3.4 Conclusão da Revisão Bibliográfica

Para concluir podemos afirmar que na literatura há diversos meios para que se possa atingir os objetivos anteriormente elencados, onde a utilização de algoritmos da área das ciências da computação faz-se necessário, um vez que métodos que utilizam *deep learning* são adequados para a resolução do problema em questão.

Dentre esses métodos, podemos destacar as ANN's LSTM *autoencoder*, que tem uma topologia especializada em detecção de anomalias em séries temporais [97], pois os mesmos mantêm a relação no domínio do tempo dos dados de entrada [55], o que faz desse algoritmo de UL adequado para a utilização no módulo de detecção anteriormente mencionado.

Para a construção do método que define a frequência de amostragem dos sensores inerciais, pode-se utilizar ferramentas conceituadas na literatura para o estudo da região de instalação do dispositivo, o que nos possibilita entender o perfil de vibração a partir de simulações.

No tocante as simulações nota-se que o modelo de 2DOF quarto de carro se mostra promissor, dado sua baixa complexidade e custo computacional, tendo ainda um bom desempenho quando testado com anomalias pontuais, como lombadas e buracos [98], o tornando ideal para o trabalho aqui descrito.

4 PROGRAMA EXPERIMENTAL

Para que se possa atingir os objetivos anteriormente propostos é realizado um programa experimental que está dividido nas seguintes etapas: (i) instalação dos sensores, (ii) método para definição da frequência de amostragem, (iii) coleta do *dataset* de treino, (iv) coleta do *dataset* de teste, (v) módulo de pré processamento e (vi) módulo de inteligência artificial.

4.1 Instalação dos sensores

A instalação dos sensores é uma fase de suma importância, uma vez que há diversos pontos que demandam tomadas de decisões que influenciarão diretamente na qualidade dos dados de coleta e conseqüentemente no trabalho realizado, como a região para acoplamento, o modelo do veículo, o método de comunicação dos equipamentos, o fundo de escala dos sensores, etc.

Desse modo, para que se possa eliminar variáveis condicionadas a utilização de diferentes automóveis optou-se pelo estudo em apenas um carro sedã da marca Honda, cujo suas principais especificações podem ser verificadas na tabela Tabela 8 e as demais informações no manual do proprietário Honda City [99] que se encontra no Anexo A.

Tabela 8 – Principais Especificações Honda City 2013. (Fonte: Retirado do manual do proprietário Honda City 2013 [99])

Veículo	
Modelo	City
Montadora	Honda
Versão	EX FLEX
Ano	2013
Peso Bruto Total	1540 Kg
Potência	115 CV

Logo após, foi escolhido a bandeja inferior da suspensão do eixo dianteiro no lado do motorista como sendo o lugar para o acoplamento do dispositivo Geobox conforme ilustra a Figura 21, deixando o sistema posicionado na região denominada não amortecida, recebendo integralmente os esforços impostos pela via e tendo apenas o pneu como filtro natural.



Figura 21 – Geobox instalado em um veículo. (Fonte: Adaptado de [26])

Devido ao *hardware* estar localizado abaixo do sistema de suspensão a coleta dos dados acontece sem os efeitos de amortecimento, permitindo assim a análise bruta das medições, o que pode gerar *insights* e detecções de anomalias que não seriam possíveis com a localização do dispositivo na cabine. Outro ganho gerado pelo posicionamento fixo dos sensores é o fato do mesmo sempre manter os eixos de medições constante, o que não acontece quando se utiliza os sensores do próprio *smartphone*, visto que sua orientação comumente varia devido ao seu manuseio, necessitando assim de um algoritmo extra para correção de sua posição [87].

No tocante ao fundo de escala dos equipamentos foi definido de forma arbitrária 4G para os sensores de aceleração e $34,9 \frac{rad}{s}$ para os sensores de velocidade angular, proporcionando assim maior resolução para regiões de menor amplitude, melhorando então a precisão para a aquisição dos dados.

Por fim, pode-se elencar que apesar da coleta do sistema ser realizada por um dispositivo externo e longe da ação humana, sua comunicação se dá através de *bluetooth* com o celular do condutor, sendo este responsável apenas por geolocalizar os dados e transmitir para a nuvem através de um aplicativo para sistemas *android*.

4.2 Método para Definição da Frequência de Amostragem

Para a definição da frequência de amostragem é proposto um método que contempla uma etapa analítica e uma fase de simulação conforme evidenciado na Figura 22, sendo sugerido sua utilização por arquitetos de soluções no momento da concepção de sistemas, analisando assim o comportamento dos sensores quando estes estão em operação, o que possibilita compreender o fenômeno em sua totalidade para tomadas de decisões

mais assertivas em relação ao valor de aquisição dos dados, permitindo a economia de banda para armazenamento e transmissão sem perder a informação do sinal captado.

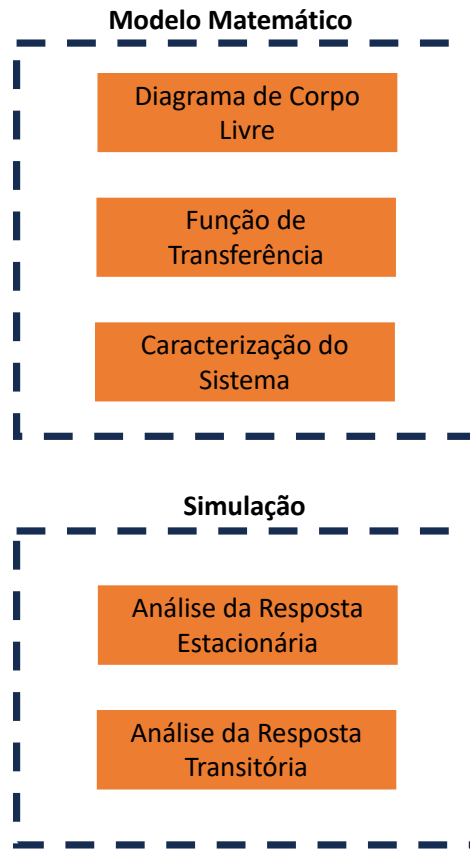


Figura 22 – Método Proposto para a Detecção de Anomalias. (Fonte: Autor 2024)

A etapa analítica consiste em um modelo matemático de 2DOF denominado quarto de carro, que utiliza como base equações da mecânica newtoniana para descrever a dinâmica da região de instalação do sensor, por meio de uma função de transferência que relaciona a aceleração imposta pela via com a aceleração da região de interesse. A caracterização desse modelo se dá pela análise de alguns parâmetros como a frequência natural da região analisada, o coeficiente de amortecimento, etc.

Por fim, a fase de simulação computacional parte do equacionamento anteriormente realizado para descrever a resposta estacionária e a resposta transitória desse sistema.

4.2.1 Modelo Matemático

A partir do modelo quarto de carro podemos obter a dinâmica da região de interesse descrevendo as forças que atuam nesse sistema. Para isso um diagrama de corpo livre é mostrado na Figura 23.

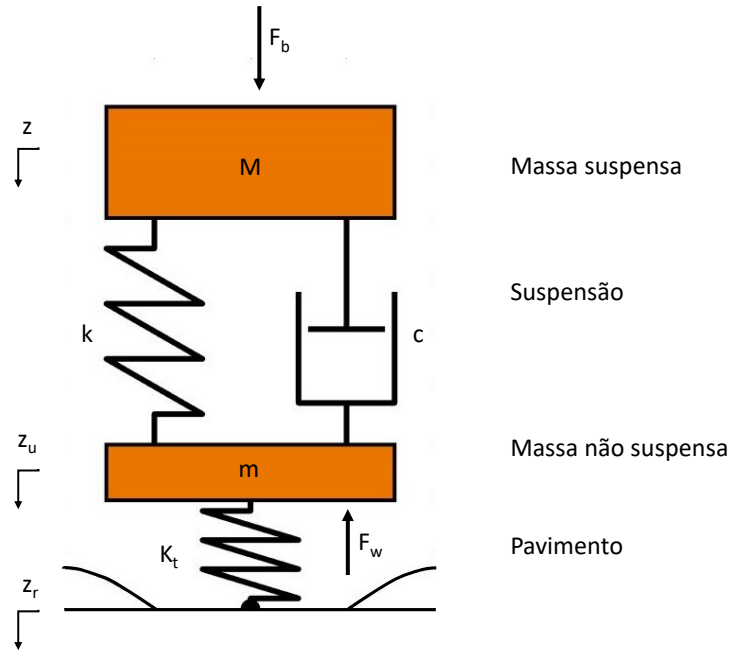


Figura 23 – Diagrama de corpo livre - Modelo quarto de carro. (Fonte: Autor 2024)

Onde :

M, m = Massa da região suspensa e não suspensa

k, k_t = Constante de rigidez da região suspensa e não suspensa

c = Constante de amortecimento

F_b, F_w = Força externa aplicada na região suspensa e não suspensa

$z_r, \dot{z}_r, \ddot{z}_r$ = Deslocamento, velocidade e aceleração vertical imposta pela via

z, z_u = Deslocamento vertical da região suspensa e não suspensa

Analisando então as forças incidentes através da segunda lei de Newton $\forall t \in \mathbb{N}$:
 $z_r(t) \neq 0$ podemos descrever as equações de movimento no domínio do tempo como:

$$M\ddot{z}(t) = c\dot{z}_u(t) + kz_u(t) - c\dot{z}(t) - kz(t) + F_b \quad (4.1)$$

$$m\ddot{z}_u(t) = c\dot{z}(t) + kz(t) - c\dot{z}_u(t) - z_u(t)(k + k_t) + k_t z_r(t) + F_w \quad (4.2)$$

Sendo, Equação 4.1 pertencente a região suspensa e Equação 4.2 a região não suspensa.

Admitindo $F_b = 0$, $F_w = 0$ e reescrevendo as equações de forma que apenas a aceleração seja a grandeza de interesse, temos :

$$M\ddot{z}(t) = c \int (\ddot{z}_u(t) - \ddot{z}(t))dt + k \int \int (\ddot{z}_u(t) - \ddot{z}(t))dt^2 \quad (4.3)$$

$$m\ddot{z}_u(t) = c \int (\ddot{z}(t) - \ddot{z}_u(t))dt + k \int \int (\ddot{z}(t) - \ddot{z}_u(t))dt^2 + k_t \int \int (\ddot{z}_r(t) - \ddot{z}_u(t))dt^2 \quad (4.4)$$

Logo, para se obter a função de transferência do sistema é admitido que não há condições iniciais de aceleração e então utilizado a transformada unilateral de Laplace na Equação 4.3 e Equação 4.4 para se trabalhar no domínio da frequência por questão de simplicidade.

$$\frac{\ddot{z}_u(s)}{\ddot{z}_r(s)} = k_t \frac{(Ms^2 + cs + k)}{(Ms^2 + cs + k)(ms^2 + cs + k + k_t) - (cs + k)^2} \quad (4.5)$$

Onde o módulo de $\frac{\ddot{z}_u(s)}{\ddot{z}_r(s)}$ é denominado ganho do sistema, ou transmissibilidade da aceleração entre região não suspensa e pavimento, conforme afirma [100].

Isso posto e com a Equação 4.5 definida pode-se então realizar a caracterização do sistema a partir de alguns parâmetros importantes, como a frequência natural da região de interesse e seu coeficiente de amortecimento, a fim de se entender a dinâmica da resposta. É importante ainda ressaltar que ao analisar a função de transferência obtida percebe-se que esta é uma função de quarta ordem, dado o grau de seu denominador, o que gera quatro polos de interesse que eventualmente também podem ser úteis para descrever o comportamento do sistema e gerar a resposta no domínio do tempo.

Em condições que não há excitação externa o sistema vibra de forma assíncrona, onde a região não suspensa possui uma frequência natural simbolizado como ω_n , medida em $\frac{rad}{s}$ conforme demonstrado na Equação 4.6 e descrito por [101].

$$\omega_n = \sqrt{\frac{(k + k_t)}{m}} \quad (4.6)$$

Onde, para fins práticos é definido a frequência natural normalizada, medida em Hz a partir da Equação 4.6:

$$\hat{\omega}_n = \frac{\omega_n}{2\pi} \quad (4.7)$$

Caso haja tal perturbação é possível descrever como se dá o amortecimento do sistema, onde uma constante adimensional é atribuída para a região não suspensa (ζ_u), sendo ela descrita como:

$$\zeta_u = \frac{c}{2\sqrt{(k + k_t)m}} \quad (4.8)$$

Por fim, pode-se reescrever a Equação 4.5 em termos dos parâmetros aqui definidos:

$$\frac{\ddot{z}_u(s)}{\ddot{z}_r(s)} = \frac{k_t}{m} \frac{(Ms^2 + cs + k)}{[(Ms^2 + cs + k)(s^2 + 2\zeta_u\omega_n s + \omega_n^2) - (2\zeta_u\omega_n s + \frac{k}{m})^2]} \quad (4.9)$$

4.2.2 Resposta do Sistema

Com os parâmetros para análise e a função de transferência já definidos, pode-se obter a resposta do sistema, sendo ela composta por uma parte estacionária e outra transitória conforme descrito na equação Equação 4.10 e evidenciado em [102].

$$R(t) = R_{tr}(t) + R_{ss}(t) \quad (4.10)$$

Para a resposta estacionária ($R_{ss}(t)$) admite-se que não há mais transientes residuais e o sistema está em regime (quando $t \rightarrow \infty$), ou seja, nesta parcela pode-se observar o ganho para situações na qual os parâmetros de entrada não variam no tempo.

Já para a resposta transitória ($R_{tr}(t)$) se têm a reação do sistema a partir de uma entrada que afete seu equilíbrio, fazendo assim com que o mesmo responda de forma oscilatória até atingir o seu estado final.

De forma análoga, quando um automóvel passa por um obstáculo no pavimento a região não amortecida responde a essa entrada de forma transitória, onde essa oscilação momentânea do sistema pode indicar para um algoritmo de aprendizado de máquina que há uma anomalia, trazendo assim uma grande importância para esse transiente, dado que entender sua dinâmica pode ser útil na fase de definição da frequência de amostragem.

Desse modo, adotando um obstáculo genérico assim como o ilustrado na Figura 24 onde sua largura é denominada l_b , sua altura h_b e seu perfil cossenoidal por conveniência matemática, podemos escrever o deslocamento vertical da via como:

$$z_r(t) = \frac{hb}{2} [1 - \cos(\frac{2\pi}{t_p}t)] \quad 0 \leq t \leq t_p \quad (4.11)$$

Sendo t_p o tempo em que o veículo demora para atravessar o obstáculo a partir de uma velocidade longitudinal v definida por :

$$v = \frac{l_b}{t_p} \quad (4.12)$$

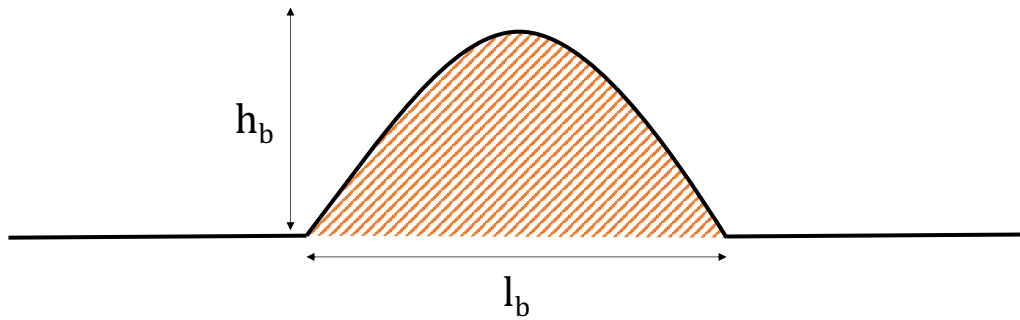


Figura 24 – Obstáculo padronizado. (Fonte: Autor 2024)

Realizando a primeira derivada em função do tempo na Equação 4.11 podemos obter a velocidade vertical que o pavimento impõe sobre o sistema (carro):

$$\dot{z}_r(t) = \frac{\pi h_b}{t_p} \sin\left(\frac{2\pi}{t_p}t\right) \quad 0 \leq t \leq t_p \quad (4.13)$$

Derivando então a Equação 4.13, sendo esta agora a segunda derivada da Equação 4.11 e por consequência a aceleração vertical imposta pela via, temos:

$$\ddot{z}_r(t) = \frac{2h_b\pi^2}{t_p^2} \cos\left(\frac{2\pi}{t_p}t\right) \quad 0 \leq t \leq t_p \quad (4.14)$$

Definindo $\omega = \frac{2\pi}{t_p}$ como sendo a velocidade angular do sinal de entrada (medida em $\frac{rad}{s}$), podemos reescrever a Equação 4.14 no domínio do tempo como:

$$\ddot{z}_r = \frac{h_b\omega^2}{2} \cos(\omega t) \quad 0 \leq t \leq t_p \quad (4.15)$$

E, no domínio da frequência:

$$\ddot{z}_r(s) = \frac{h_b\omega^2 s}{2(s^2 + \omega^2)} \quad (4.16)$$

Por fim, para se obter a resposta do sistema no momento em que o carro passa por um obstáculo basta utilizar a Equação 4.5 em conjunto com a Equação 4.16, encontrando assim:

$$\ddot{z}_u(s) = \frac{k_t h_b \omega^2 (Ms^2 + cs + k)s}{2(s^2 + \omega^2)[(Ms^2 + cs + k)(s^2 + 2\zeta_u \omega_n s + \omega_n^2) - (2\zeta_u \omega_n s + \frac{k}{m})^2]} \quad (4.17)$$

Onde, é necessário a realização da transformada inversa de Laplace da Equação 4.17 para a obtenção da resposta no domínio do tempo.

4.3 Coleta do *Dataset* de Treino

Para a obtenção dos dados treino foi realizado uma coleta de campo visando o treinamento para algoritmos de UL, gerando assim um *dataset* que servirá como base para o trabalho aqui descrito e estando disponível para pesquisas futuras de aprendizado de máquina, podendo ser incrementado de forma rápida, uma vez que não há o processo de classificação das amostras.

Com o *Geobox* devidamente acoplado no veículo e a taxa de amostragem definida, um aplicativo também desenvolvido pela empresa Geovista denominado *Geopavement* foi instalado no celular do condutor, conforme demonstrado na Figura 25, onde o motorista foi responsável apenas em iniciar a gravação e termina-lá, abolindo a necessidade de se utilizar um passageiro para rotular os dados, mitigando assim o problema de falhas humanas descrita em [14].

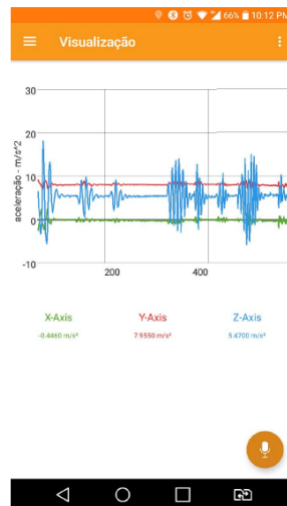


Figura 25 – Aplicativo para Coleta de Dados de Treino. (Fonte: Autor 2024)

O critério de escolha das vias para coleta de dados parte da premissa de que *datasets* para treinamento de algoritmos baseados em reconstrução não pode conter elementos considerados anômalos, desse modo, foi escolhido no sistema Geovista vias classificadas como boas ou ótimas conforme ilustra a Figura 26 , onde uma pré inspeção foi realizada em loco, assegurando que nenhuma informação anômala seria considerada como asfalto regular.



Figura 26 – Vias boas - Sistema Geovista. (Fonte: Autor 2024)

4.4 Coleta do *Dataset* de Teste

Para a conferência da assertividade do modelo optou-se por coletar novas amostras de anomalias no pavimento, de modo que agora a presença de um passageiro para realizar a classificação se faz necessário, construindo assim um *dataset* de SL que poderá também ser utilizado para treinamento de máquina em pesquisas futuras.

Desse modo, foi utilizado um segundo aplicativo de celular denominado *Geohunter* que possui as funcionalidades detalhadas na Figura 27.

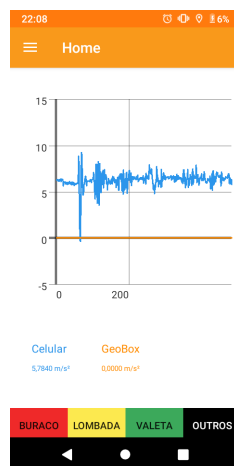


Figura 27 – Aplicativo para Coleta de Dados de Teste. (Fonte: Autor 2024)

As amostras coletadas pela aplicação são denominadas como *ground truth* e possuem alta confiabilidade, pois cada elemento coletado provém de uma indicação manual, onde é informando o tipo de anomalia encontrada, gerando uma gravação de três segundos do momento em que o carro passa sobre a localização de interesse no asfalto.

4.5 Escolha das *Features* do sistema

Em um primeiro momento, para testes e análises, inúmeras características foram coletadas em campo, gerando uma tabela com diversas colunas que descrevem a dinâmica do automóvel a luz de diferentes variáveis, entretanto, para economizar banda de transmissão e evitar o crescimento da dimensionalidade dos dados de forma exacerbada, fez-se necessário uma etapa de escolha de variáveis que tem como objetivo filtrar informações que possam atrapalhar o aprendizado de máquina.

Sendo assim, foram definidas sete *features* derivadas de três grandezas distintas que têm alta influência na assertividade do sistema de inteligência artificial, são elas: velocidade, aceleração e velocidade angular.

4.5.1 Velocidade

A velocidade longitudinal do carro é uma grandeza importante nesse contexto, pois ela influencia diretamente a resposta de aceleração no eixo z, sendo possível notar para a mesma anomalia, leituras com padrões diferentes pelo simples fato de alterar a velocidade do veículo [103].

Desse modo, utilizar o dado de velocidade do automóvel no treino de máquina pode ajudar não apenas na detecção das anomalias, mas também a identificar o real motivo da intensidade da aceleração.

4.5.2 Aceleração

A aceleração da região não suspensa é medida nas três direções (x, y, z) pelo sensor, onde o eixo z possui uma média de $10 \frac{m}{s^2}$ devido a ação da força gravitacional.

De fato, a leitura com maior importância para o sistema de detecção está no eixo z, dado que há uma reação direta na aceleração vertical no momento em que o carro passa por uma anomalia, entretanto as outras medições podem trazer informações importantes que descrevem o modo de direção do motorista.

Optou-se então por incluir integralmente a leitura dos acelerômetros no treino de máquina para que assim as ações de freadas, conversões bruscas e mudanças de direções não fossem identificadas como anomalias, mitigando a detecção de FP.

4.5.3 Velocidade Angular

Utilizar os dados de velocidade angular a priori pode parecer redundante, uma vez que a medição é realizada nas mesmas direções que a aceleração, porém, a partir desse dado pode-se obter movimentos característicos do automóvel como o *pitch*, *roll* e *yaw*, conforme ilustrado na Figura 28.

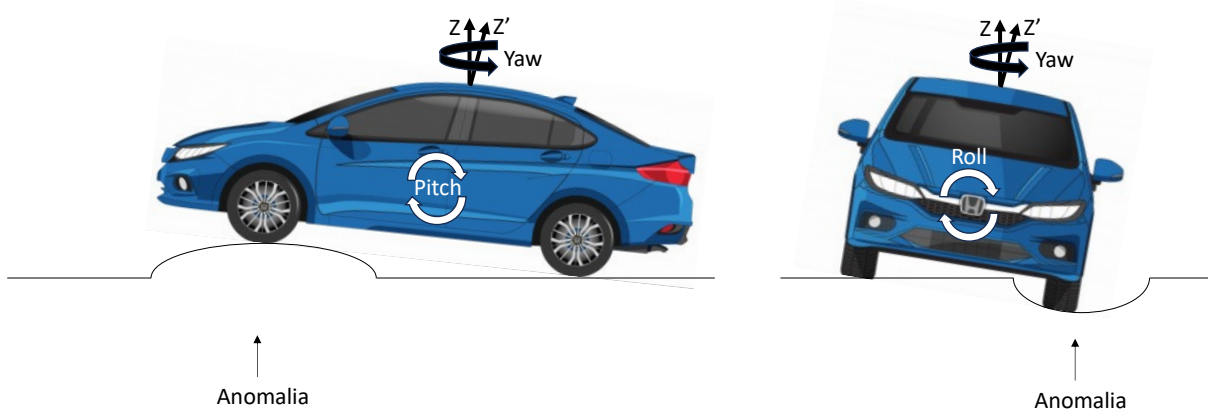


Figura 28 – Carro passando sobre uma anomalia - Análise do *pitch*, *roll*, *yaw*. (Fonte: Autor 2024)

Segundo Seraj et al [82] essas três medidas podem variar significativamente conforme o carro passa sobre as anomalias e utilizá-las pode ajudar no aprendizado de máquina, aumentando a taxa de classificação do sistema, o que nos incentiva a inserir essa característica no processo de treinamento, ainda que o método de definição para a taxa de aquisição do sistema atual não leve em conta essa variável, uma vez que tal estudo demandaria a utilização de um modelo matemático com uma maior complexidade.

4.6 Módulo de Pré Processamento

Após a coleta de campo, os dados em sua forma bruta são enviados para um servidor em um arquivo do tipo CSV (Valores Separados por Vírgula - Do inglês *Comma Separated Values*), onde um pré processamento se faz necessário a fim de eliminar inconsistências e adequar as informações para que ocorra a fase de treinamento. Essa tarefa inclui: análise da integridade do *dataset*, criação da janela de tempo e normalização dos dados conforme ilustrado na Figura 29.

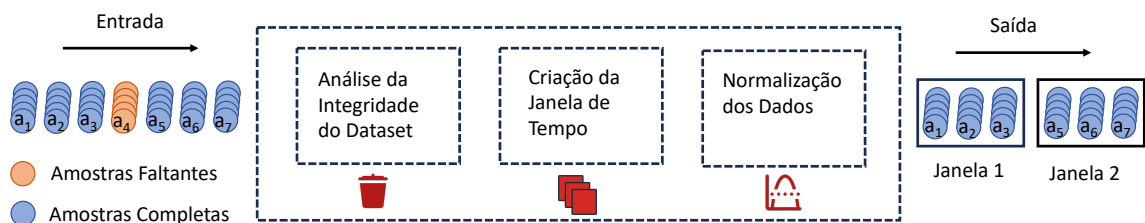


Figura 29 – Fluxo para Pré Processamento. (Fonte: Autor 2024)

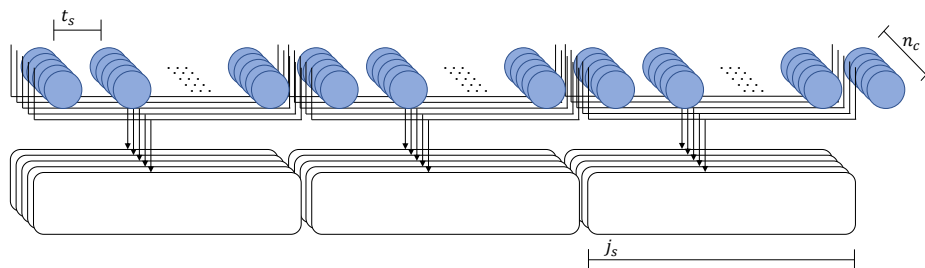
De forma geral, os pontos coletados passam por cada etapa na forma serial antes de serem utilizados pelos algoritmos de treino e inferência, sendo este pré processamento indispensável para o bom funcionamento do sistema.

4.6.1 Análise da integridade do *Dataset*

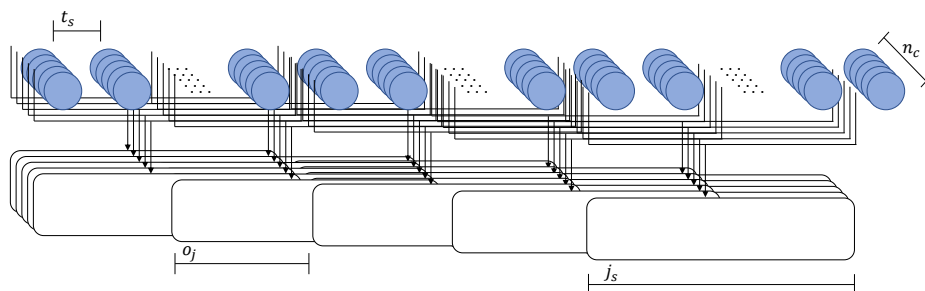
A primeira análise a ser realizada é a inspeção do *dataset*, na qual tal etapa consiste em verificar cada linha de dado coletado a fim de encontrar incongruências de medições, como valores nulos, fora de ordem, dentre outros, o que em alguns casos acaba gerando o descarte dos dados, devido a alguma falha na coleta ou transmissão.

4.6.2 Criação da janela de tempo

O conjunto de dados que inicialmente contém medidas pontuais para n_c canais espaçadas temporalmente a um valor t_s (tendo por consequência uma frequência f_s) é convertido para um conjunto de janelas de tempo de tamanho j_s , para que assim tal conjunto esteja em conformidade com a entrada da RNA, sendo que essas janelas podem ou não ter superposição o_j conforme mostra a Figura 30.



(a) Janelas de tempo - Sem sobreposição



(b) Janelas de tempo - Com sobreposição

Figura 30 – Criação das janelas de tempo. (Fonte: Autor 2024)

Desse modo, a fase de inferência e treinamento ocorre com os dados dispostos em grupos de pontos, dando assim um contexto temporal para cada amostra, uma vez que há uma relação de dependência em cada elemento medido.

4.6.3 Normalização dos dados

Após a criação da janela de tempo os dados são normalizados a fim de se obter um novo valor para cada ponto x , onde $x \in I$, sendo I o intervalo $[0, 1]$.

A equação que realiza essa transformação é descrita a seguir:

$$\bar{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (4.18)$$

Vale ressaltar que a normalização é deveras importante para as fases subsequentes, uma vez que essa etapa ajuda no aprendizado de máquina quando utilizado RNA's.

4.7 Módulo de Inteligência Artificial

O módulo de inteligência artificial é composto por uma RNA recorrente do tipo LSTM baseada em *autoencoders*, onde cada ponto está associado a uma janela de tempo com um tamanho j_s pré-determinado, conforme ilustrado na Figura 31.

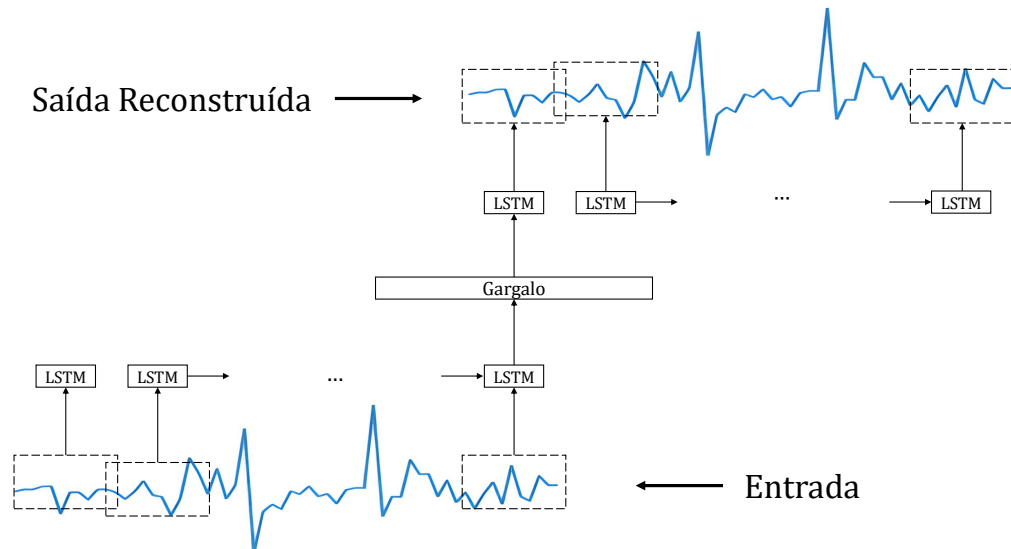


Figura 31 – RNA - Autoencoder LSTM. (Fonte: Autor 2024)

Portanto, ao fornecer as variáveis alvo o algoritmo tentará reconstruí-las, o que fará muito bem para os pontos não anômalos, entretanto espera-se que pontos contendo anomalias tenham um erro absoluto médio alto.

4.7.1 Treino de máquina

O treino de máquina consiste em iterar por inúmeras vezes os dados não anômalos que previamente foram processados, disponibilizando-os para a RNA de modo que o objetivo principal seja minimizar a função de custo, onde espera-se que ao término do treino exista um erro de reconstrução baixo, dado que o sistema nesse momento consegue reconstruir dados não anômalos de forma ótima.

4.7.2 Definição do limiar do erro

Enfim, após o término do treino é estudado o valor do limiar do erro para a detecção da anomalia, onde se utiliza os próprios dados contidos no aprendizado para realizar a inferência na rede já treinada e plotar a distribuição do seu erro, no qual o limiar ideal é um limiar maior que o maior valor encontrado nesse processo, julgando assim como anômalos os pontos com valores de erro de reconstrução maiores que esse valor delimitado.

5 RESULTADOS E DISCUSSÕES

A partir da base teórica explicitada, do método para coleta de campo e treinamento de máquina definidos, faz-se necessário uma análise nos resultados obtidos, para que se tenha uma discussão detalhada sobre a real aplicação dos objetivos propostos.

Como o modelo matemático escolhido baseia-se em parâmetros concentrados, fez-se necessário a definição dos mesmos, onde após uma consulta no manual do proprietário [99] foram definidos os valores descritos na Tabela 9.

Tabela 9 – Parâmetros da Simulação Quarto de Carro. (Fonte: Autor 2024)

Grandeza	Símbolo	Valor
Massa suspensa	M	330
Massa não suspensa	m	62
Rigidez da massa não suspensa	k_t	200000
Rigidez da massa suspensa	k	20000
Constante de amortecimento da massa suspensa	c	2000

Vale ressaltar que para cada tipo de carro há a necessidade de verificar se os valores escolhidos são aderentes, uma vez que as respostas obtidas variam significativamente com a alteração desses parâmetros. Desse modo, para fins de comparação foi analisado o trabalho [104], tendo como conclusão que as constantes escolhidas estão dentro do valor aplicado na academia para carros com suspensão do tipo *MacPherson*.

Sendo assim, para a definição da frequência de amostragem é utilizado o método apresentado no Capítulo 4, onde a partir das equações pré estabelecidas é realizado a caracterização do sistema e simulação das respostas de aceleração para a região de instalação do sensor, obtendo a dinâmica do fenômeno para algumas situações apresentadas.

Ao final da fase de simulação, espera-se que o método proposto possa ajudar na escolha de um valor concreto para a frequência de amostragem, onde este será utilizado no processo de captura dos dados, e treinamento do módulo de inteligência artificial, no qual é avaliado a capacidade de aprendizado do algoritmo de UL a partir de testes controlados.

5.1 Caracterização do Sistema

A fim de compreender o comportamento da região de instalação do sensor de forma matemática, pode-se aplicar os valores definidos na Tabela 9 na Equação 4.7 e Equação 4.8,

obtendo respectivamente a frequência natural normalizada do sistema $\hat{\omega}_n = 9,48$ Hz e seu coeficiente de amortecimento $\zeta_u = 0,27$.

Desse modo, pode-se então reescrever de forma numérica a Equação 4.9 como:

$$\frac{\ddot{z}_u(s)}{\ddot{z}_r(s)} = \frac{6.6e7s^2 + 4e8s + 4e9}{20460s^4 + 784000s^3 + 7.384e7s^2 + 4e8s + 4e9} \quad (5.1)$$

Uma vez que $0 < \zeta_u < 1$ é possível afirmar que a região de interesse quando está em sua fase transitória retorna para a estabilidade de forma sub amortecida, mesmo não sofrendo uma ação do sistema de amortecimento do veículo.

5.2 Simulação

Para a realização da simulação foi utilizado o software Matlab, gerando assim a análise da resposta em estado estacionário e estado transitório a partir do modelo de 2 DOF quarto de carro, no qual pode-se encontrar os códigos de programação realizados no Apêndice C e Apêndice D.

5.2.1 Resposta em Estado Estacionário

Admitindo apenas a parte estacionária da resposta do sistema, é possível plotar um diagrama que relaciona a entrada de sinais de diferentes frequências com o ganho obtido para tais, sendo este diagrama denominado diagrama de Bode.

Essa análise é de suma importância, pois ela permite identificar algumas frequências críticas do sistema além de descrever o seu comportamento perante diversas situações na qual já não há mais transitórios, evidenciando assim frequências de corte f_c , ressonâncias, áreas com alto ganho, etc.

Ao observar a resposta apresentada na Figura 32, podemos ver que o sistema descrito pela Equação 4.9 atua como um filtro passa baixa com uma frequência de corte muito próxima da frequência natural, atenuando assim as componentes de alta frequência.

A atenuação intrínseca que ocorre no sistema quando este está em estado estacionário prejudica consideravelmente a leitura dos dados para frequências maiores que f_c , fazendo com que essas componentes fiquem com uma relação sinal ruído baixa, o que impossibilita a compreensão dos dados em uma fase posterior. Sendo assim, pode-se afirmar que para o estado estacionário nada adianta utilizar frequências de amostragem elevadas, pois o sinal entregue para o sensor de aceleração é naturalmente filtrado devido a sua localização de instalação. Desse modo, pode-se elencar que analisando apenas a resposta estacionária, amostrar $2f_c$ ou $2\hat{\omega}_n$ (uma vez que $f_c \approx \hat{\omega}_n$) é a melhor escolha pensando em economia de banda e armazenamento de dados.

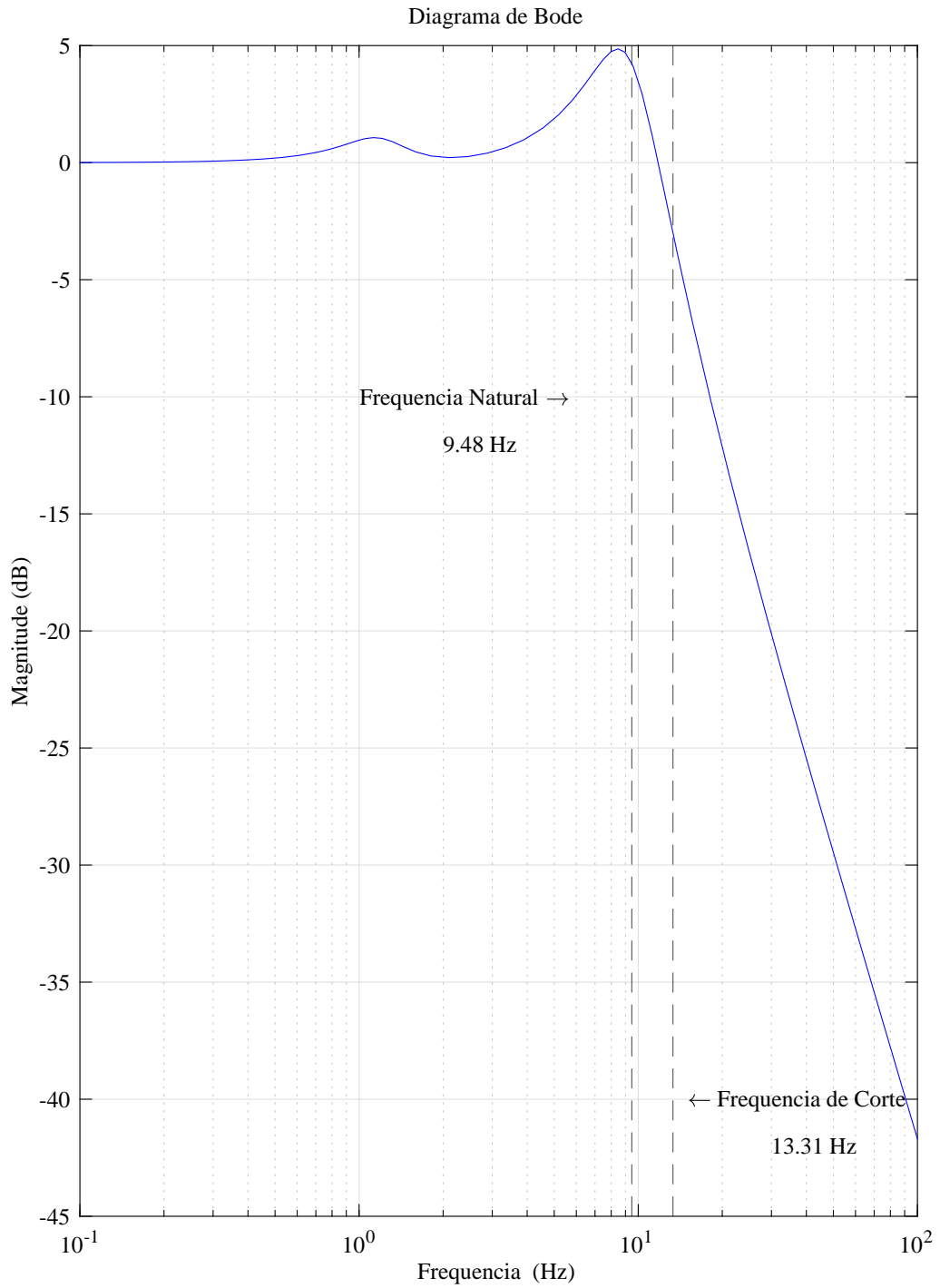


Figura 32 – Diagrama de Bode da região não amortecida. (Fonte: Autor 2024)

5.2.2 Resposta em Estado Transitório

A idéia por trás da simulação da resposta transitória é compreender como se dá o momento em que o carro passa por uma anomalia no pavimento, observando assim alguns pontos importantes do fenômeno. Desse modo, é definido um obstáculo padronizado, no qual o mesmo pode ser descrito pela Equação 4.11, onde sua altura está convencionada como sendo $h_b = 0.1m$ e seu comprimento $l_b = 0.5m$.

Admite-se ainda que o veículo trafega em diferentes velocidades, estendendo assim a análise para condições variadas.

No tocante a entrada do sistema, é importante elencar que na área de engenharia de controle o impulso unitário é uma das funções mais utilizadas para se obter a resposta transitória, porém, a escolha de simular um obstáculo padronizado tem como base a necessidade de definição da frequência de amostragem, uma vez que será necessário se aproximar das condições reais do problema para a correta tomada de decisão.

É essencial ressaltar que a análise aqui exposta está dividida em análise transitória no domínio do tempo e análise transitória no domínio da frequência, onde olhar o problema sob os ângulos de diferentes domínios pode gerar *insights* e revelar nuances nunca antes vistas.

Sendo assim, pode-se elencar que a análise no domínio do tempo investiga não apenas a maneira como o sistema retorna para a estabilidade, mas também a duração de tal acomodação, o que permite mensurar o período transitório, onde este é deveras importante para a detecção de anomalias, pois é nesse momento em que as perturbações podem ser detectadas. Logo, a Figura 33, Figura 35, Figura 37, Figura 39 e Figura 41 demonstram tal simulação no domínio do tempo para $10\frac{km}{h}$ ($2,78\frac{m}{s}$), $20\frac{km}{h}$ ($5,55\frac{m}{s}$), $30\frac{km}{h}$ ($8,33\frac{m}{s}$), $40\frac{km}{h}$ ($11,11\frac{m}{s}$) e $50\frac{km}{h}$ ($13,89\frac{m}{s}$) respectivamente.

Ja a análise no domínio da frequência nos permite obter, através de uma FFT (Transformada Rápida de *Fourier* - Do inglês *Fast Fourier Transform*) um gráfico que expande o sinal em suas componentes de frequência, explicitando assim a composição espectral da resposta transitória em função da amplitude. A Figura 34, Figura 36, Figura 38, Figura 40 e Figura 42 trazem a mesma simulação no domínio da frequência para as velocidades descritas anteriormente.

Por fim, analisando as figuras apresentadas podemos observar no domínio do tempo que a medida em que se aumenta a velocidade longitudinal do veículo o tempo de resposta transitória diminui, ocorrendo o inverso com a amplitude de resposta, onde a magnitude da aceleração pode chegar a $20 G$ por um breve período de tempo (Figura 41). No domínio da frequência nota-se que a frequência de corte f_c aumenta gradativamente com o acréscimo da velocidade até chegar em $50 Hz$ (Figura 42).

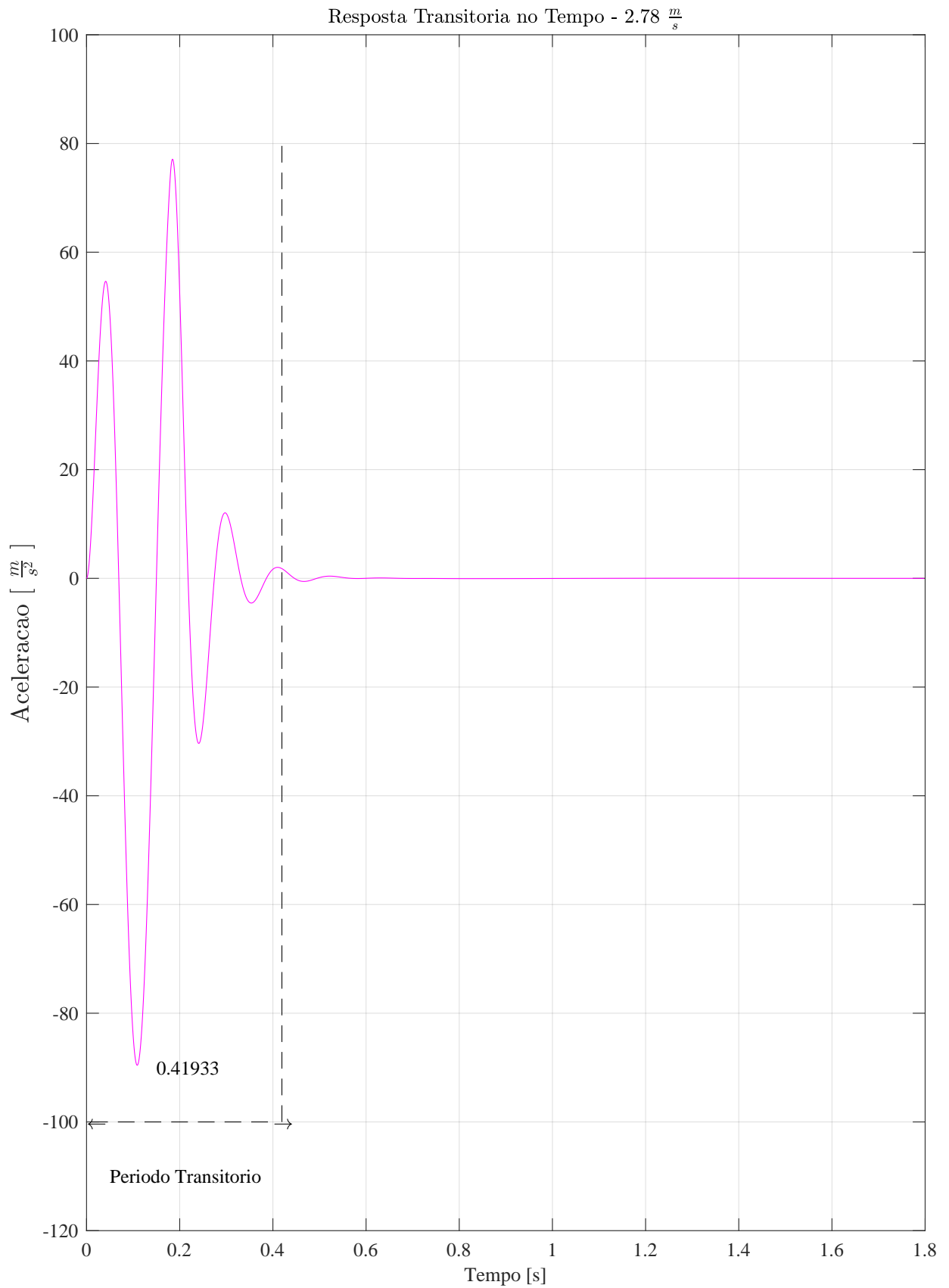


Figura 33 – Análise transitória no domínio do tempo - 10 km/h. (Fonte: Autor 2024)

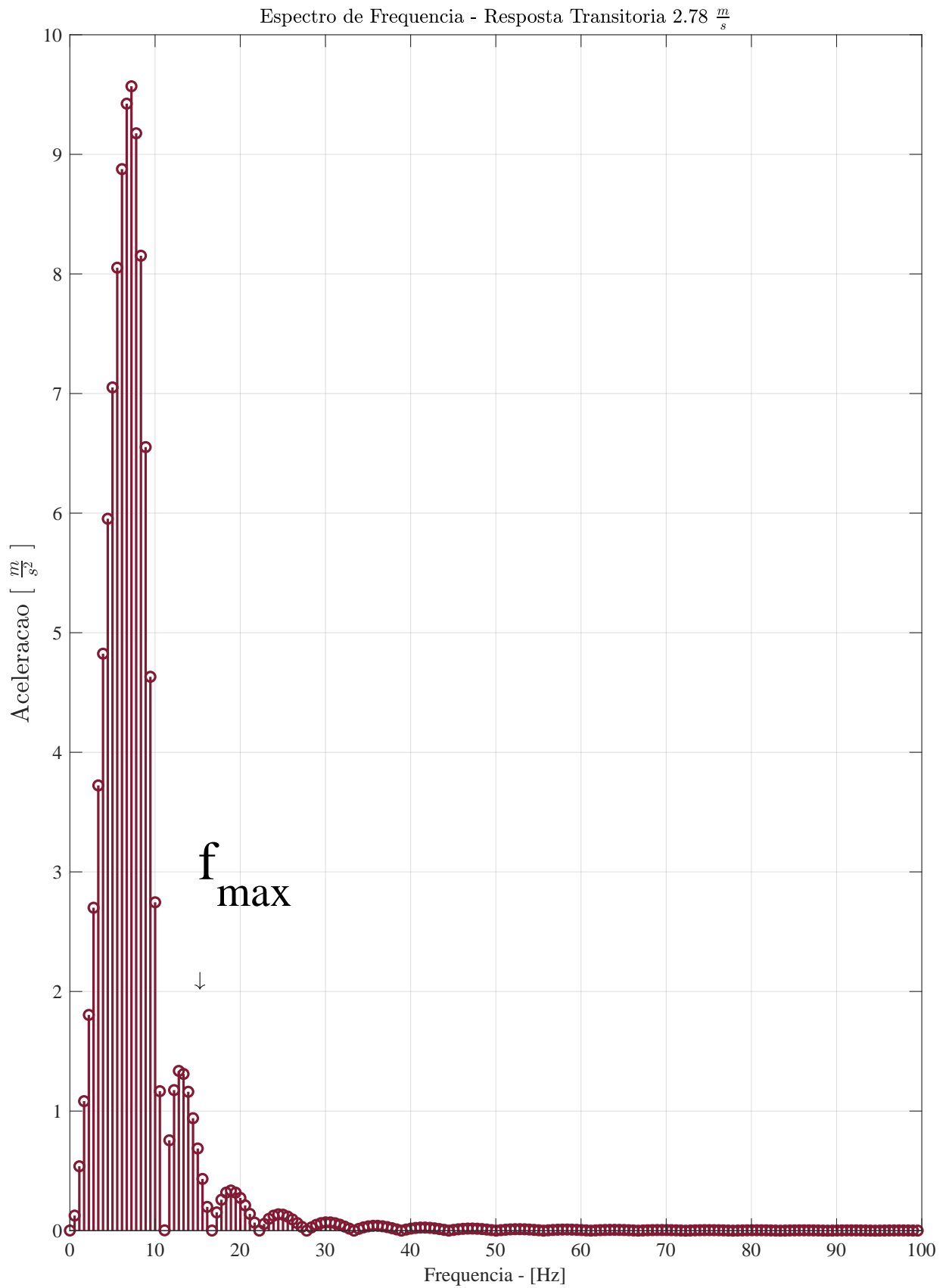


Figura 34 – Análise transitória no domínio da frequência - 10 km/h. (Fonte: Autor 2024)

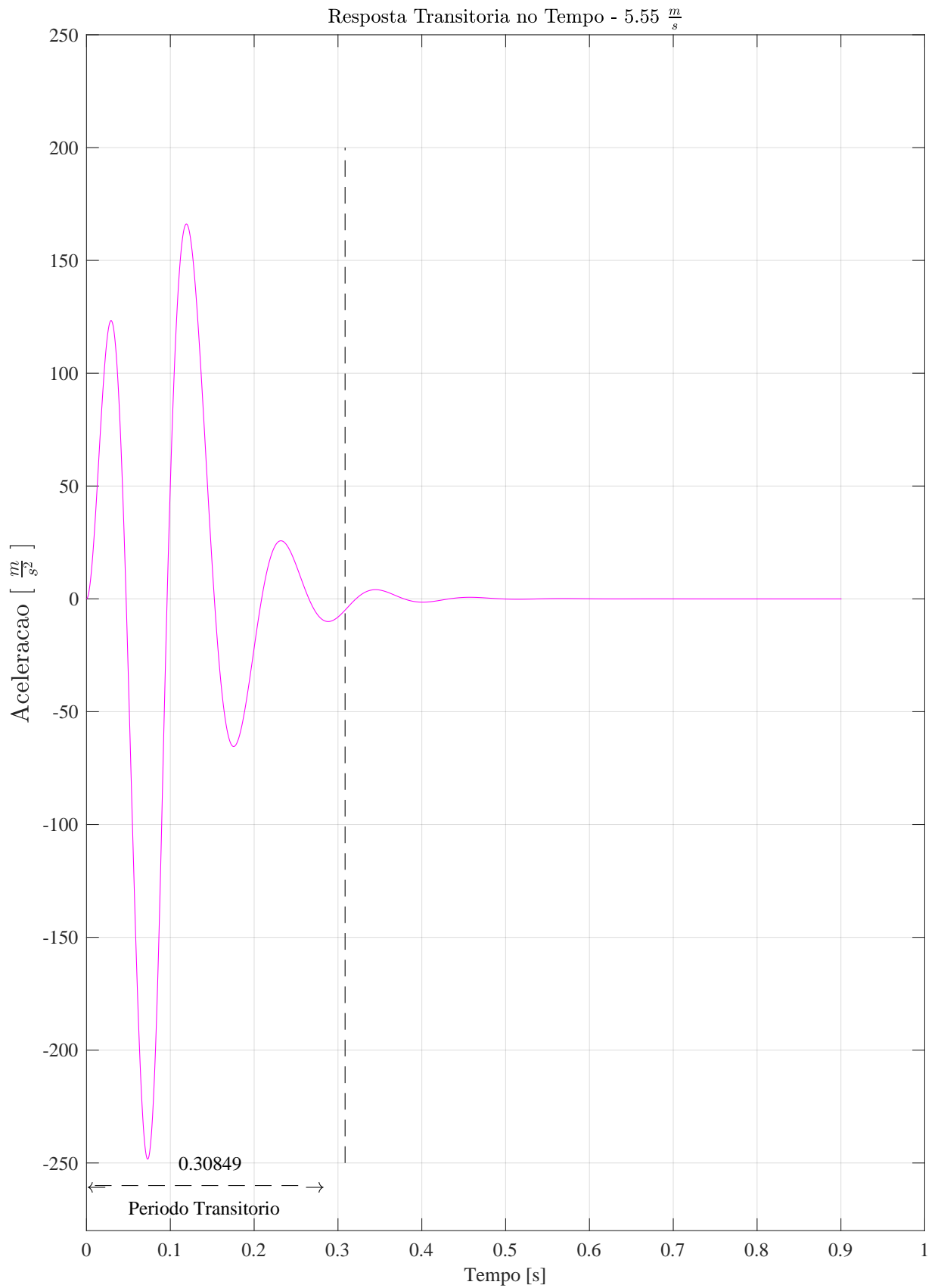


Figura 35 – Análise transitória no domínio do tempo - 20 km/h. (Fonte: Autor 2024)

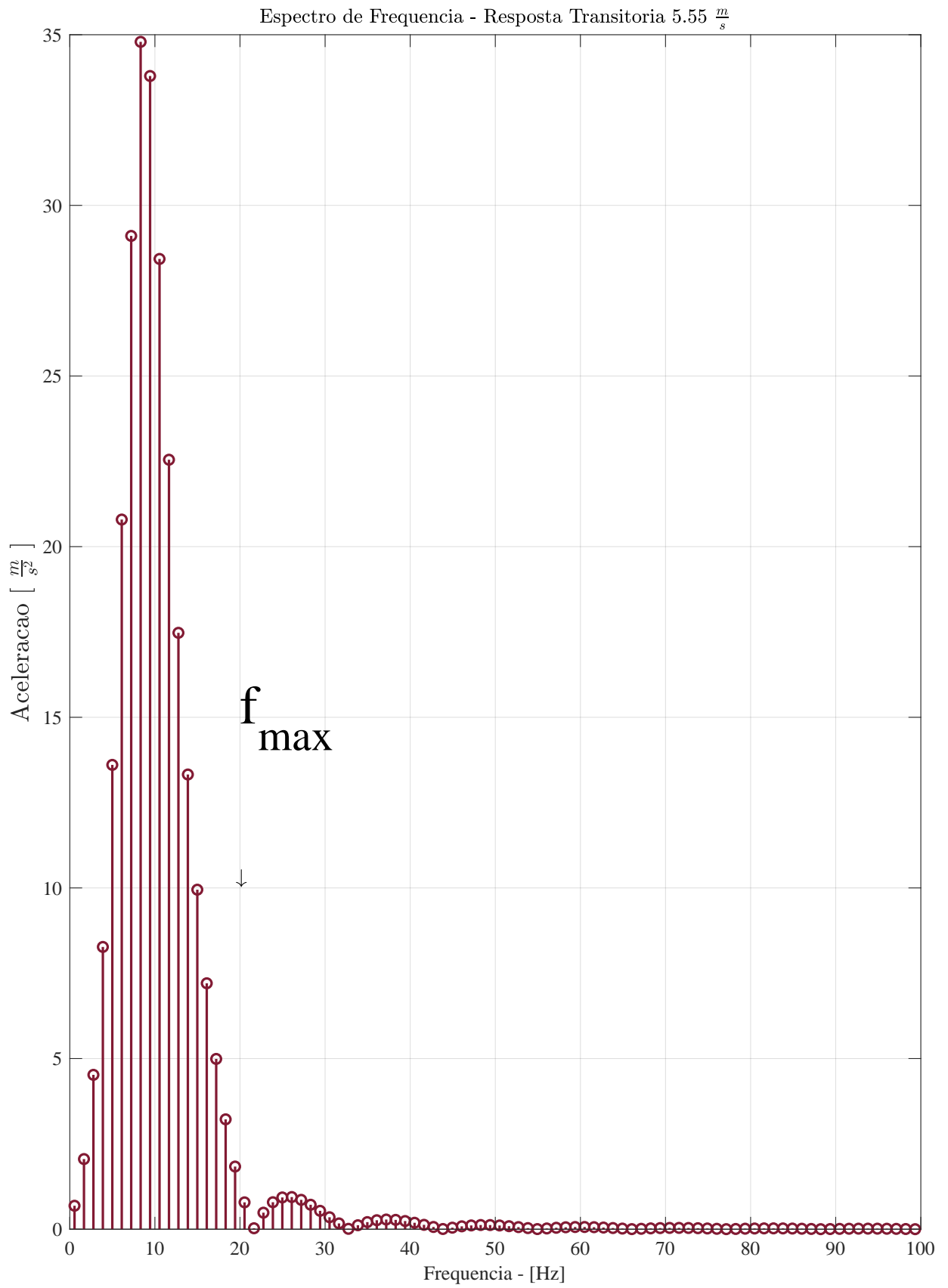


Figura 36 – Análise transitória no domínio da frequência - 20 km/h. (Fonte: Autor 2024)

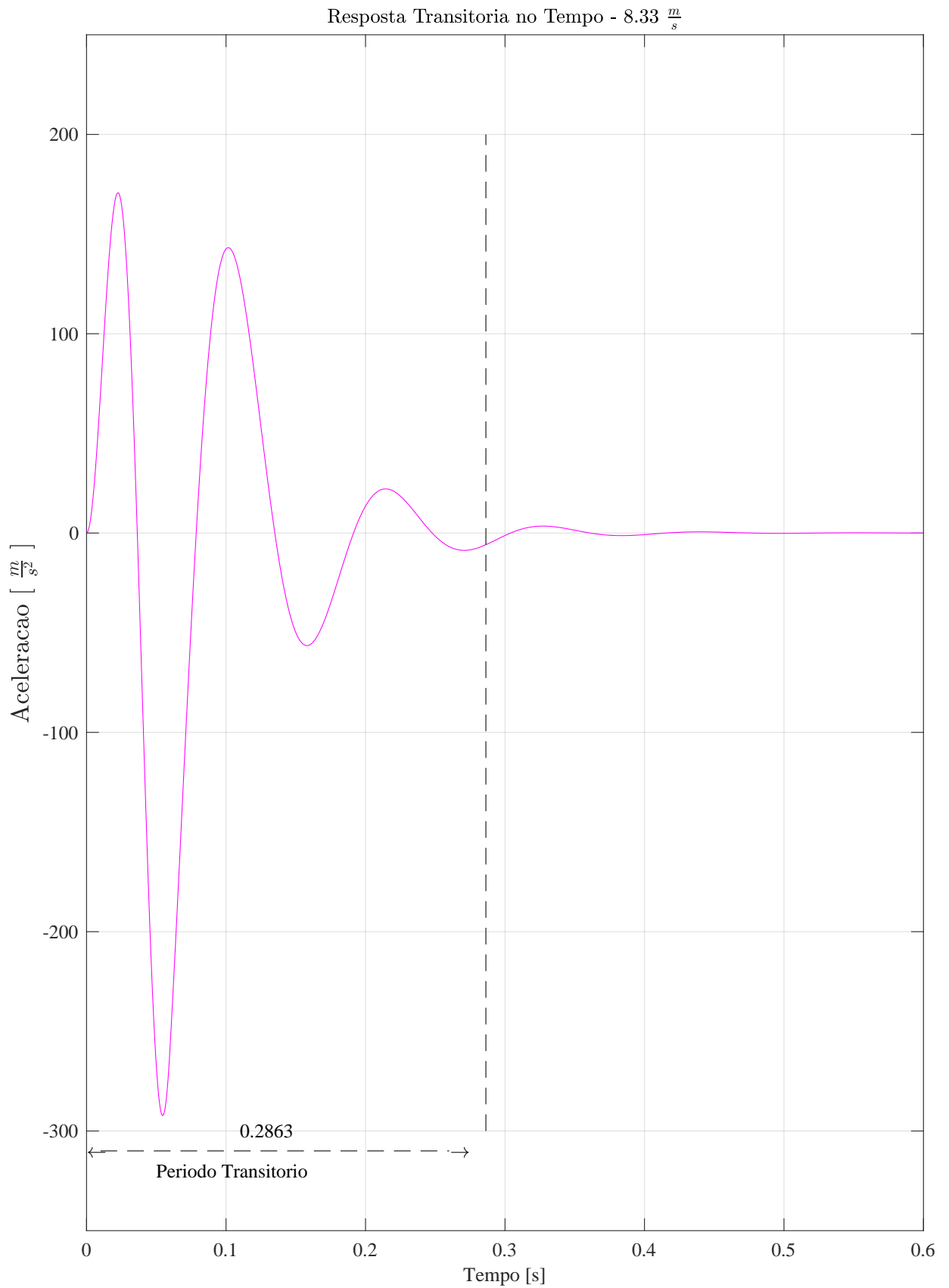


Figura 37 – Análise transitória no domínio do tempo - 30 km/h. (Fonte: Autor 2024)

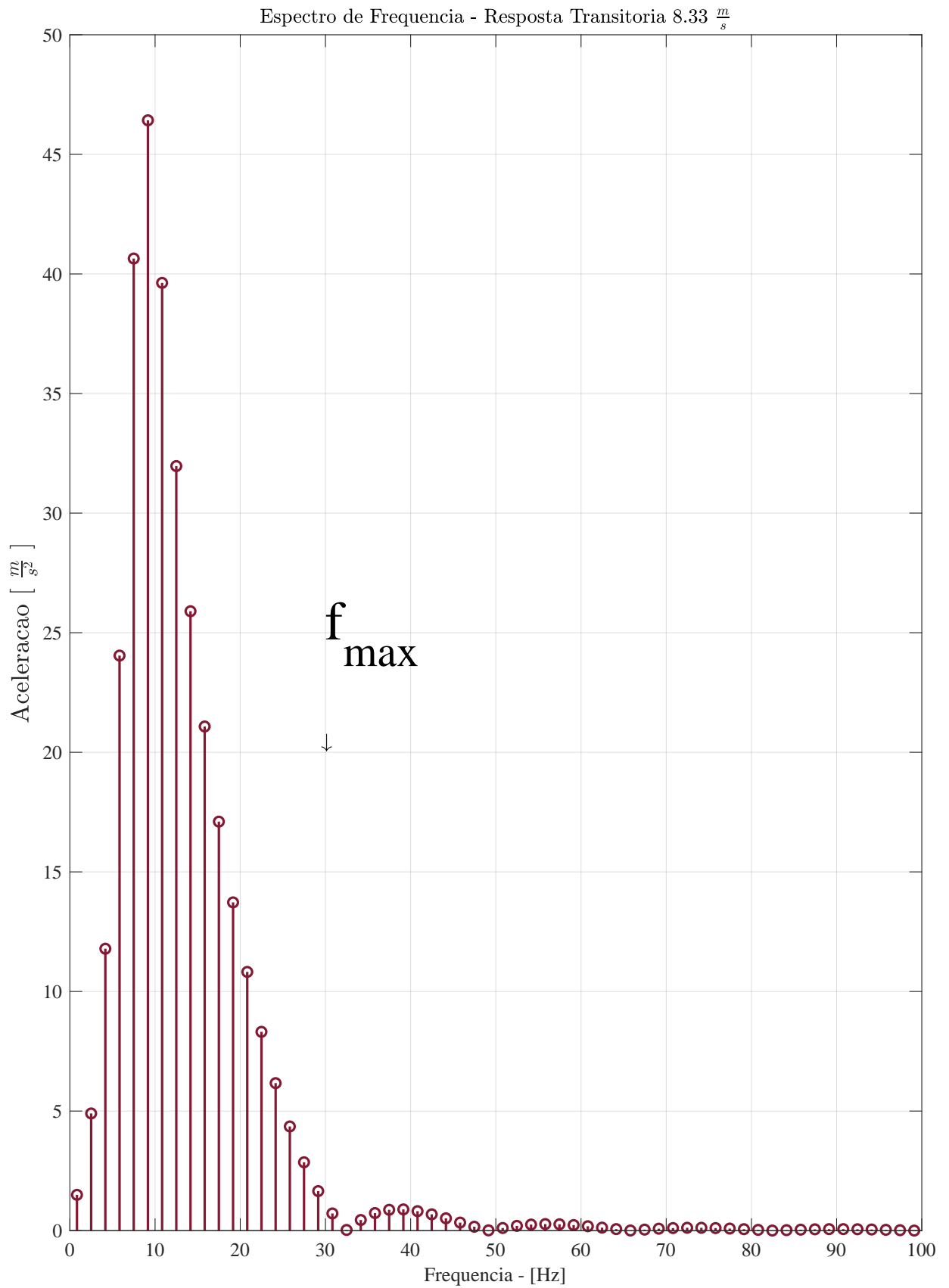


Figura 38 – Análise transitória no domínio da frequência - 30 km/h. (Fonte: Autor 2024)

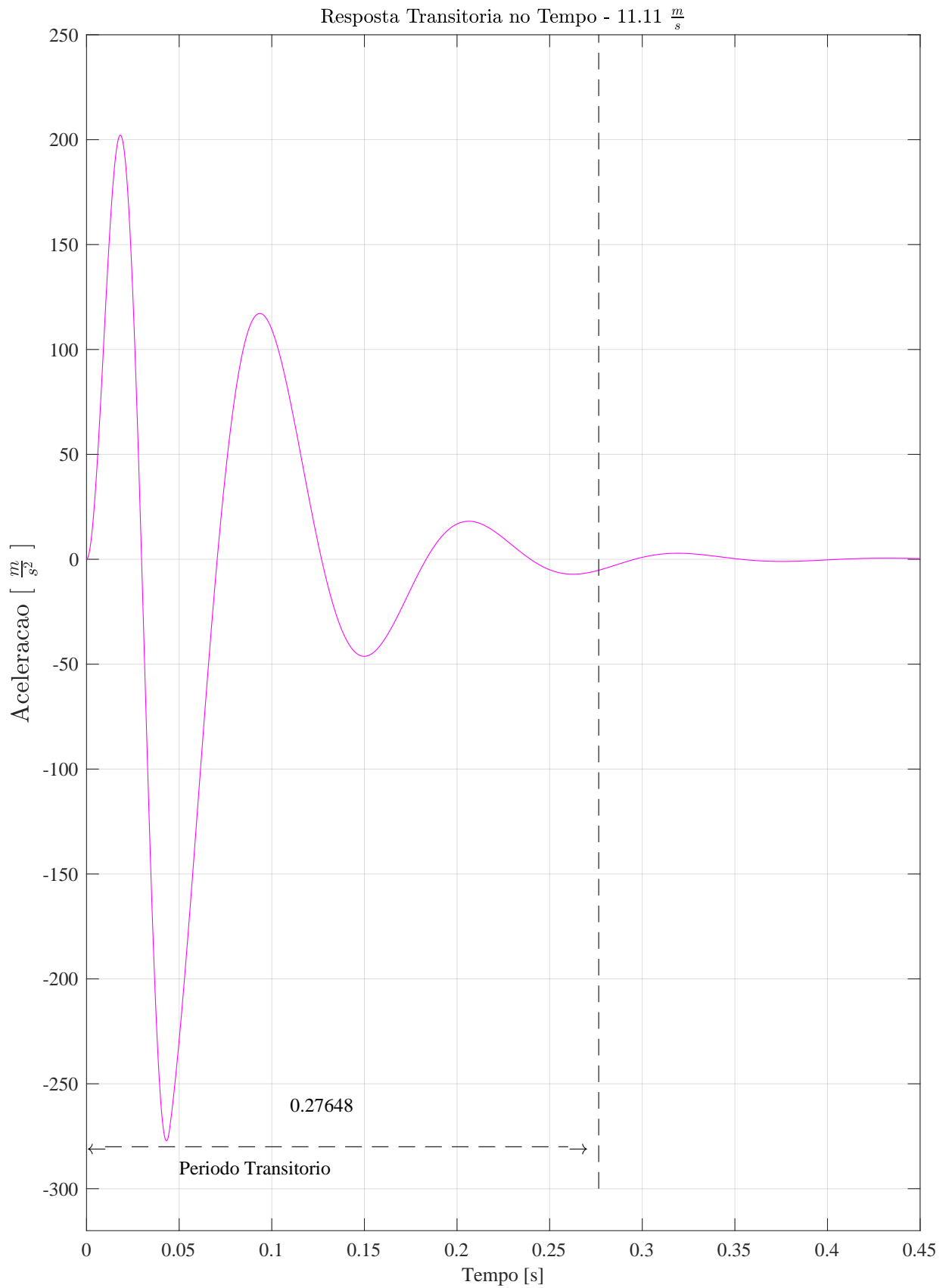


Figura 39 – Análise transitória no domínio do tempo - 40 km/h. (Fonte: Autor 2024)

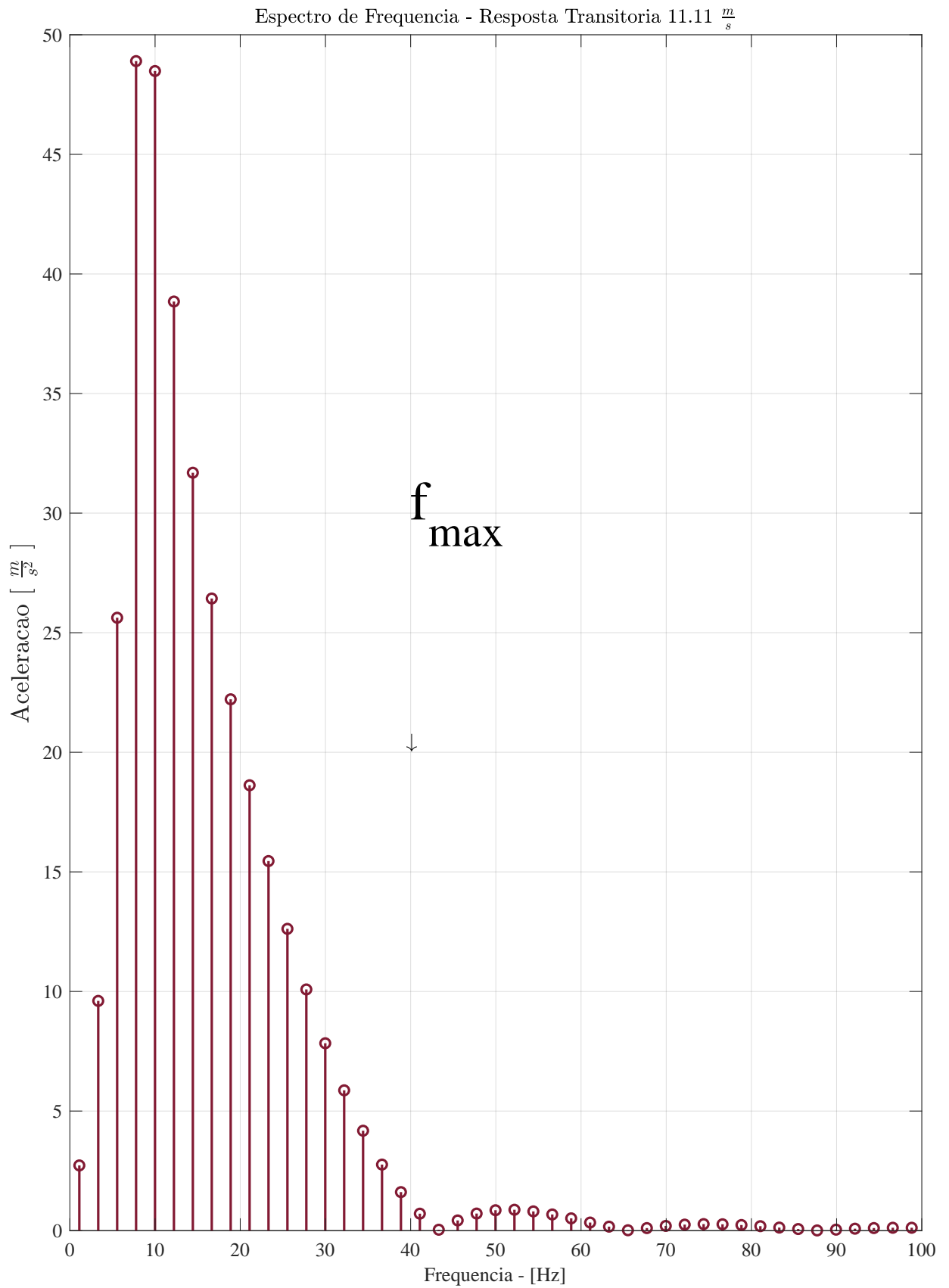


Figura 40 – Análise transitória no domínio da frequência - 40 km/h. (Fonte: Autor 2024)

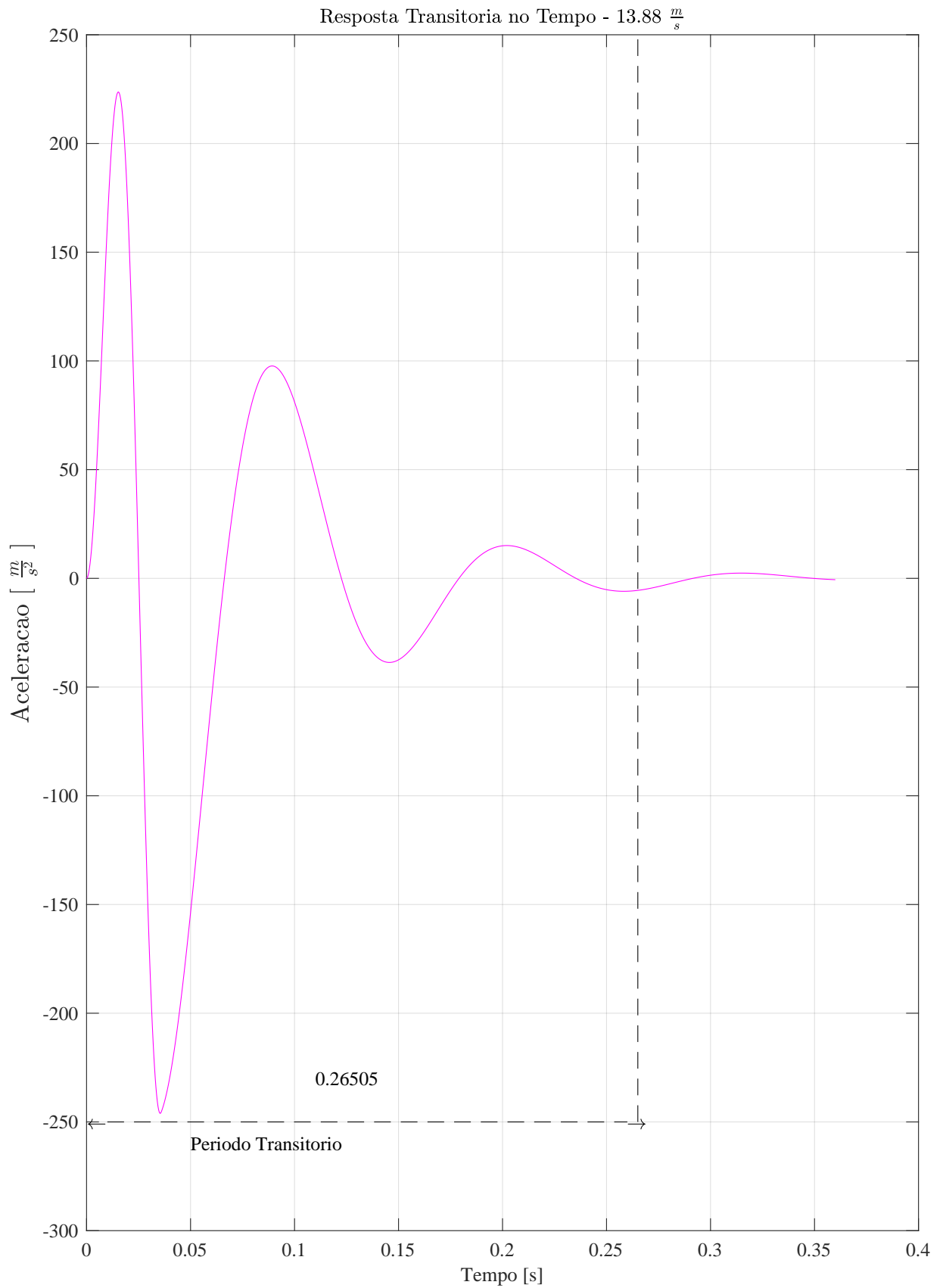


Figura 41 – Análise transitória no domínio do tempo - 50 km/h. (Fonte: Autor 2024)

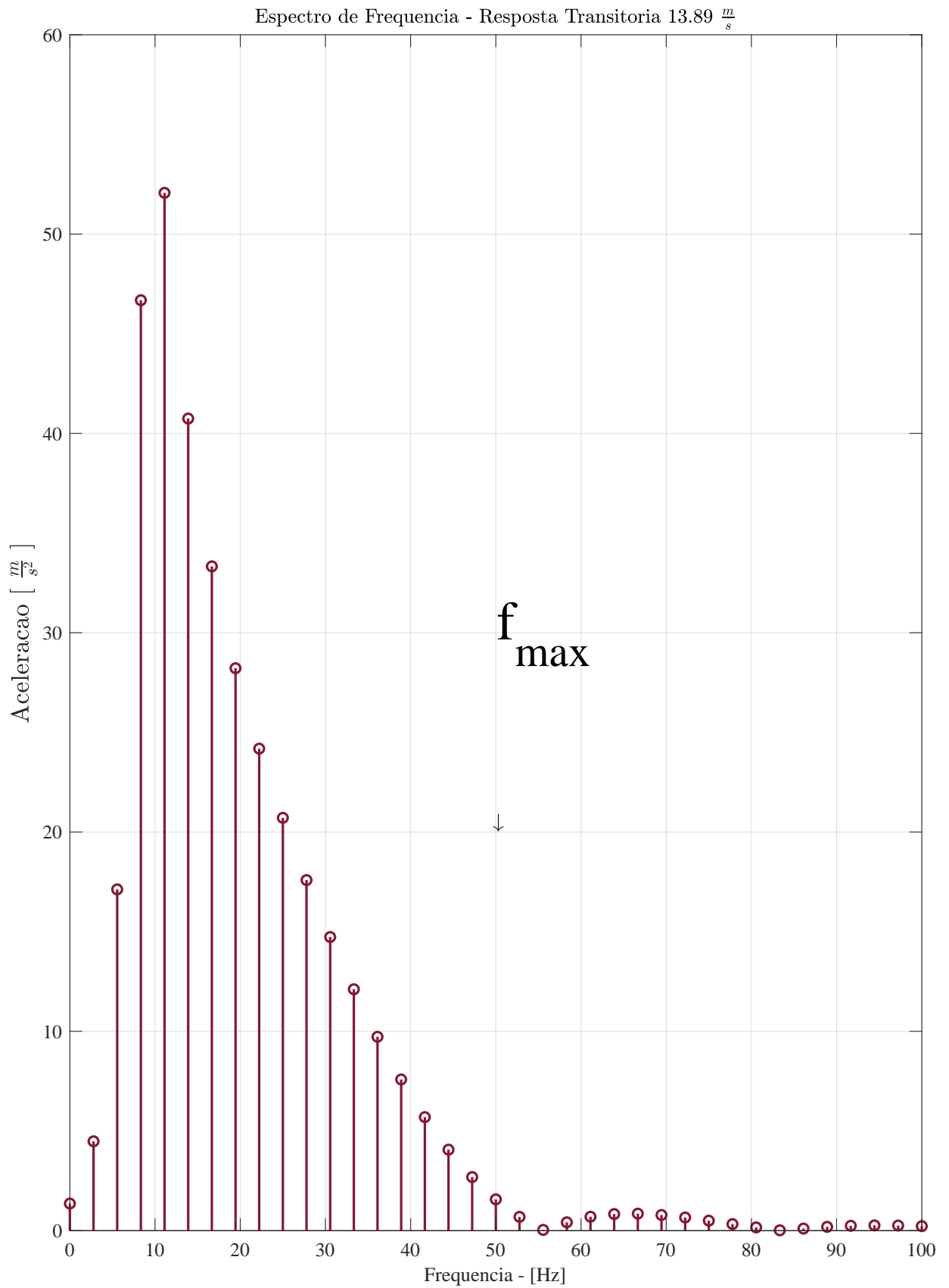


Figura 42 – Análise transitória no domínio da frequência - 50 km/h. (Fonte: Autor 2024)

5.2.3 Definição da Frequência de Amostragem

Após a realização do método proposto há insumos suficientes para a determinação do valor que será utilizado nas próximas etapas do trabalho. Admitindo então como ponto de partida a frequência de amostragem f_s sugerida para o estado estacionário, temos $2\hat{\omega}_n \approx 20Hz$, o que pode servir muito bem para a parte estacionária, porém, para que se tenha validade, é necessário verificar quais os ganhos e perdas de se utilizar $f_s = 2\hat{\omega}_n$ para o estado transitório.

Logo, a partir dos resultados temporais da resposta transitória podemos ver que o tempo de estabilização do modelo para todas as condições de entrada em diferentes velocidades é maior que o tempo de amostragem para $f_s = 20Hz$ conforme mostra a Tabela 10, nos dando a certeza de que o sensor será capaz de coletar a tempo pelo menos cinco pontos desse período.

Tabela 10 – Relação velocidade, amostragem e tempo de estabilização. (Fonte: Autor 2024)

Velocidade do Carro	Tempo de Estabilização do sistema	Tempo de Amostragem	Número de Amostras coletadas no período
$10 \frac{km}{h}$	0.42s	0.05s	8
$20 \frac{km}{h}$	0.31s		6
$30 \frac{km}{h}$	0.29s		5
$40 \frac{km}{h}$	0.28s		5
$50 \frac{km}{h}$	0.26s		5

Entretanto, fica claro na análise espectral que ao amostrar a 20 hz, inúmeras componentes de frequência do momento em que o carro passa por uma anomalia são perdidas, inviabilizando assim a captura completa do sinal, pois a aquisição total só seria possível amostrando a um valor superior a $2f_{max}$ dado o teorema de Nyquist. Porém, para a economia de recursos do armazenamento dos dados de diversos veículos conjectura-se que detectar anomalias com sinais sub amostrados pode ter o mesmo efeito que amostrar a frequência de Nyquist conforme demonstra Kawaguchi et al [105].

Desse modo a Figura 43, Figura 44, Figura 45, Figura 46 e Figura 47, demonstram uma simulação de como será o sinal captado que o sensor de aceleração receberá admitindo 20 Hz como a frequência amostragem sobre a resposta transitória no tempo, onde pode-se constatar que não há uma perda expressiva em relação ao comportamento do sinal, mas sim apenas uma redução na amplitude da aceleração o que em termos práticos nada mudará, dado que o sensor está configurado para ter seu fundo de escala em 4G.

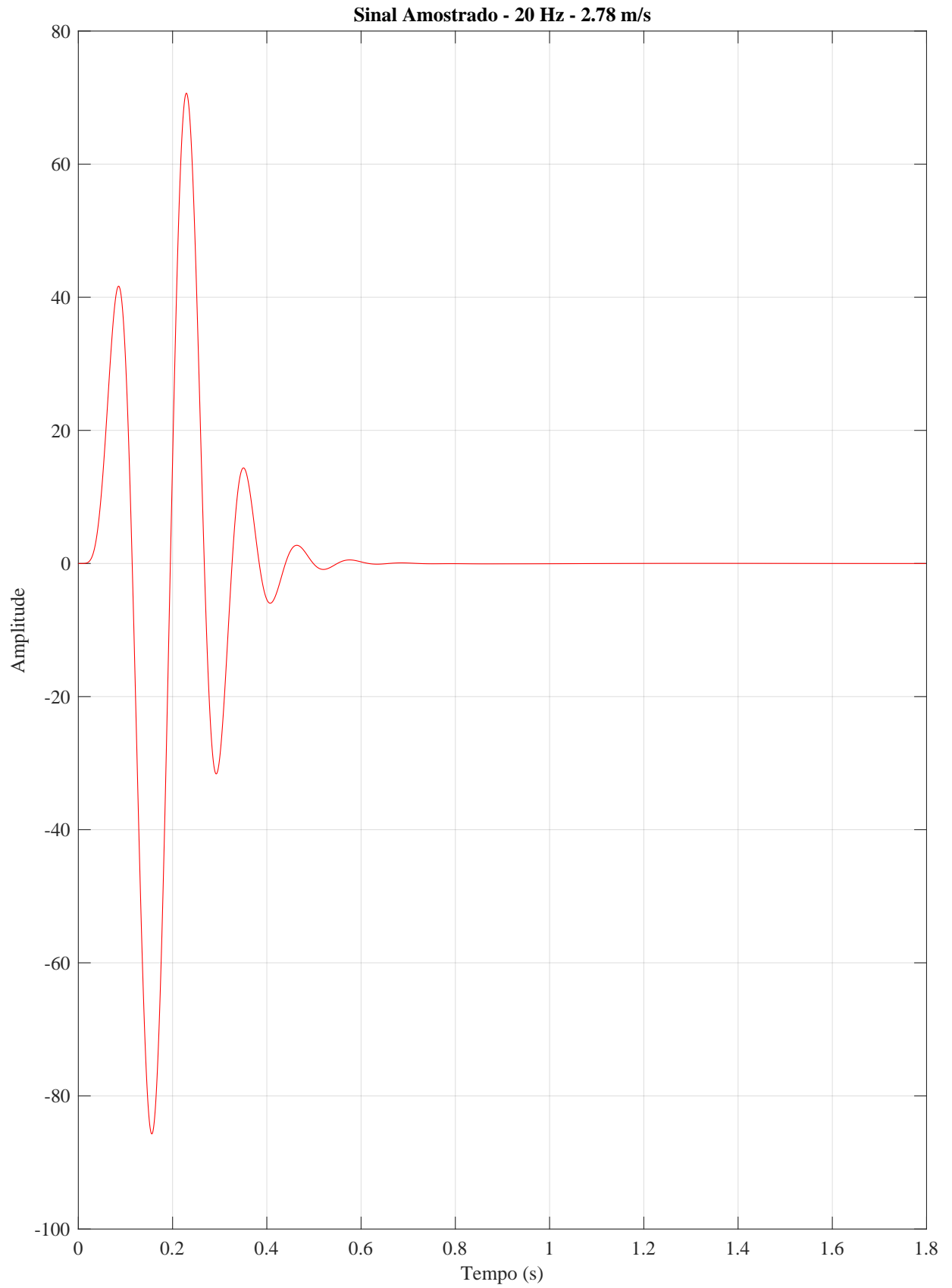


Figura 43 – Resposta Transitória amostrada a 20 Hz - 10 km/h. (Fonte: Autor 2024)

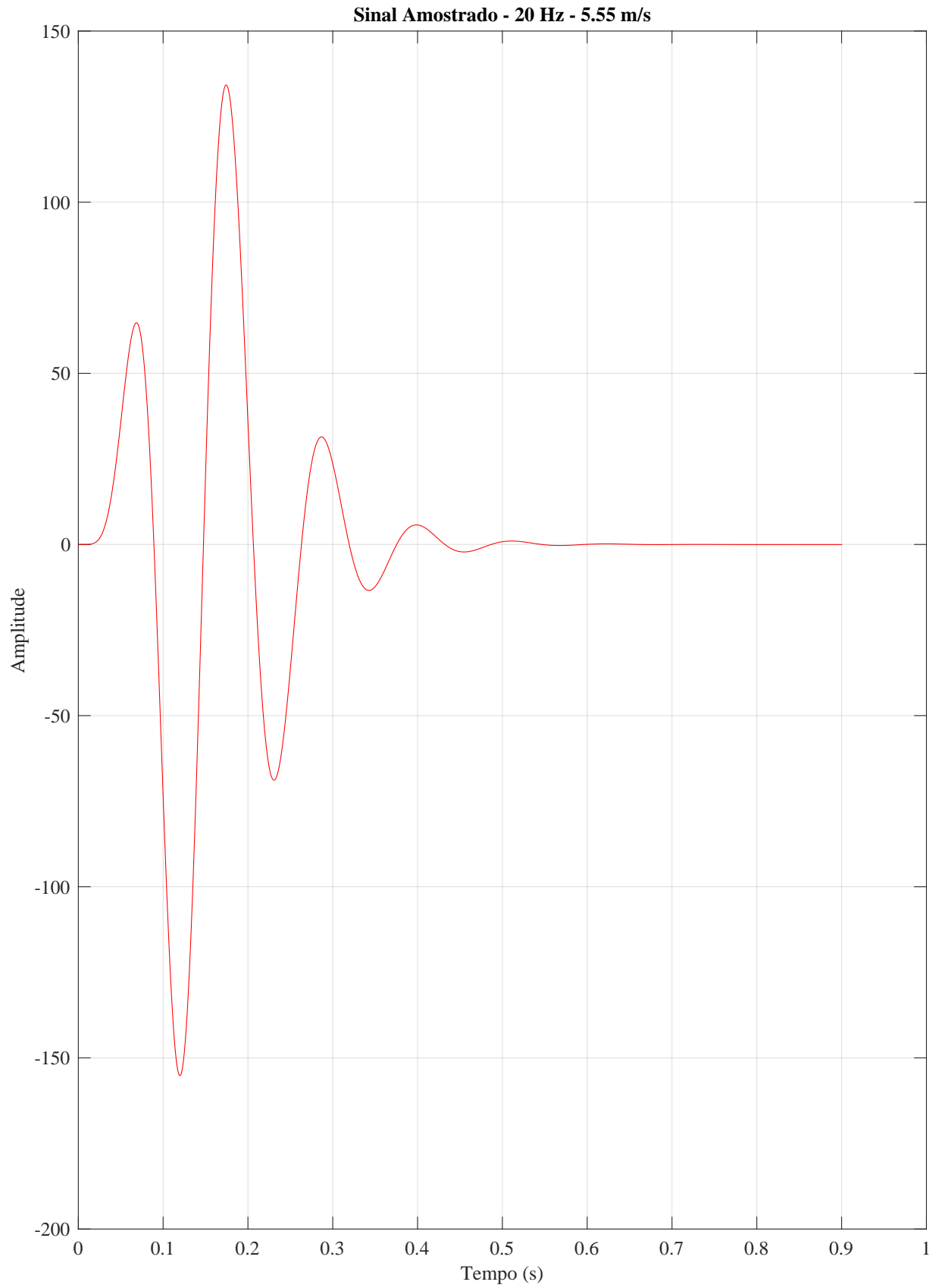


Figura 44 – Resposta Transitória amostrada a 20 Hz - 20 km/h. (Fonte: Autor 2024)

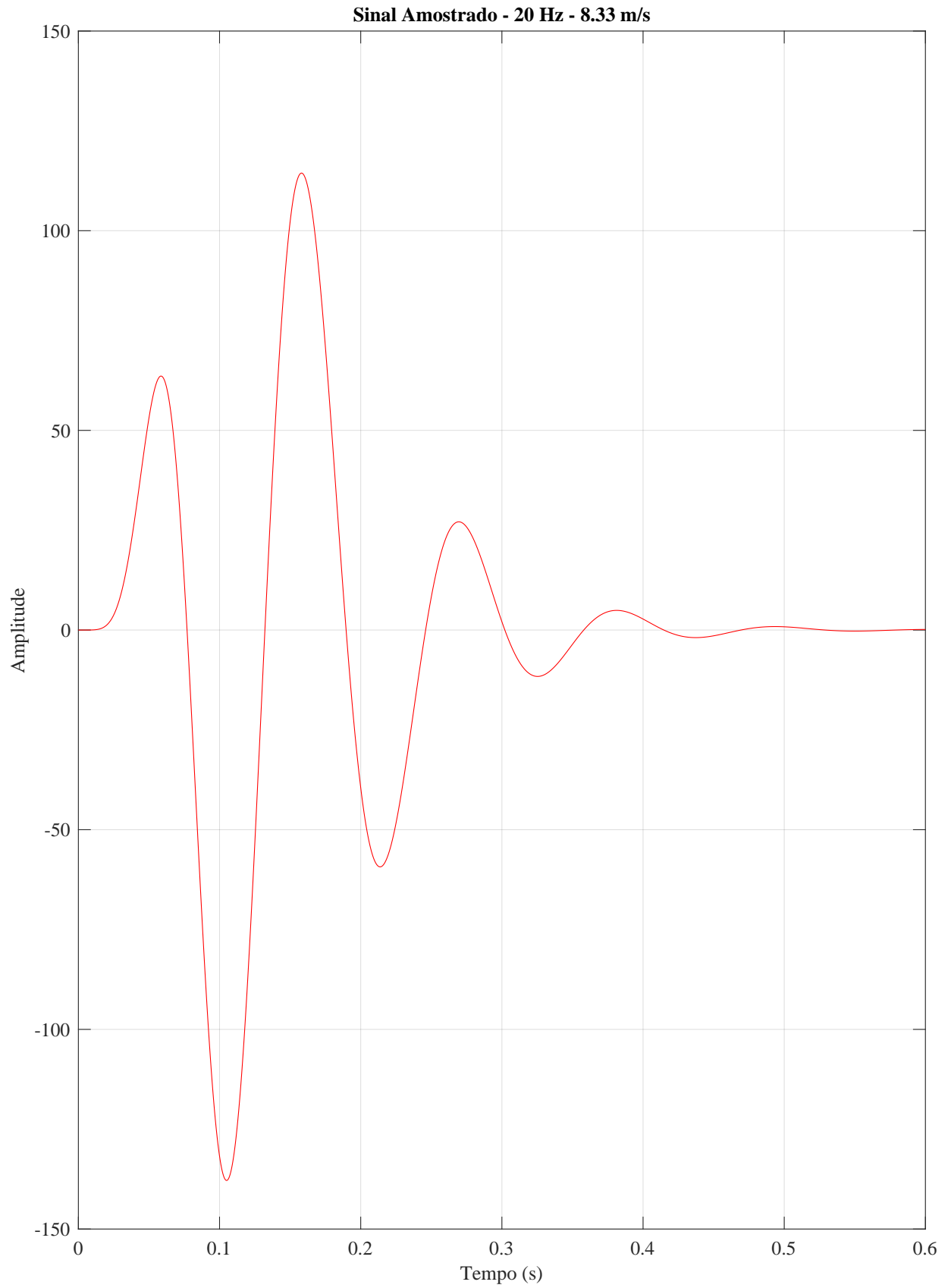


Figura 45 – Resposta Transitória amostrada a 20 Hz - 30 km/h. (Fonte: Autor 2024)

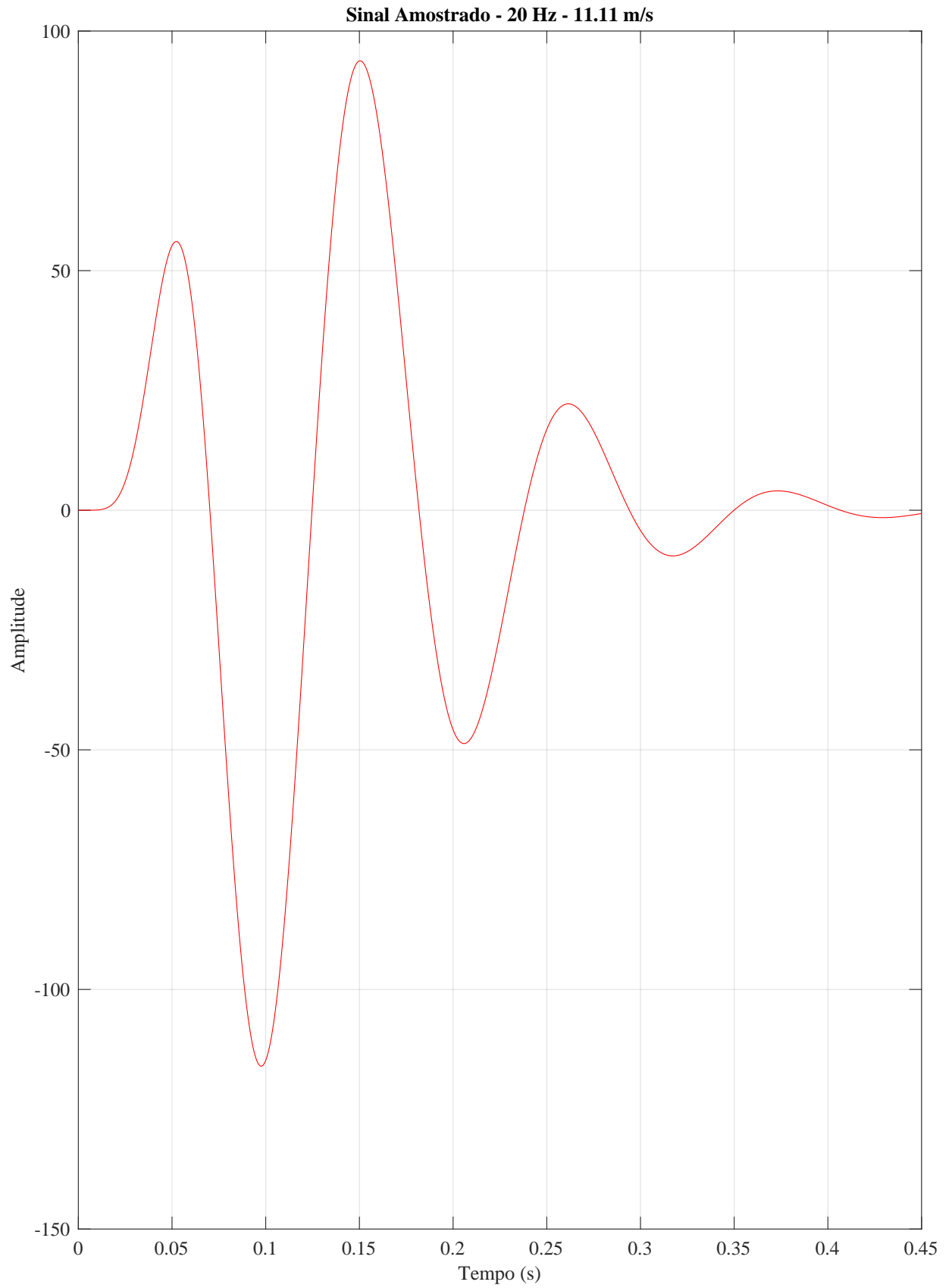


Figura 46 – Resposta Transitória amostrada a 20 Hz - 40 km/h. (Fonte: Autor 2024)

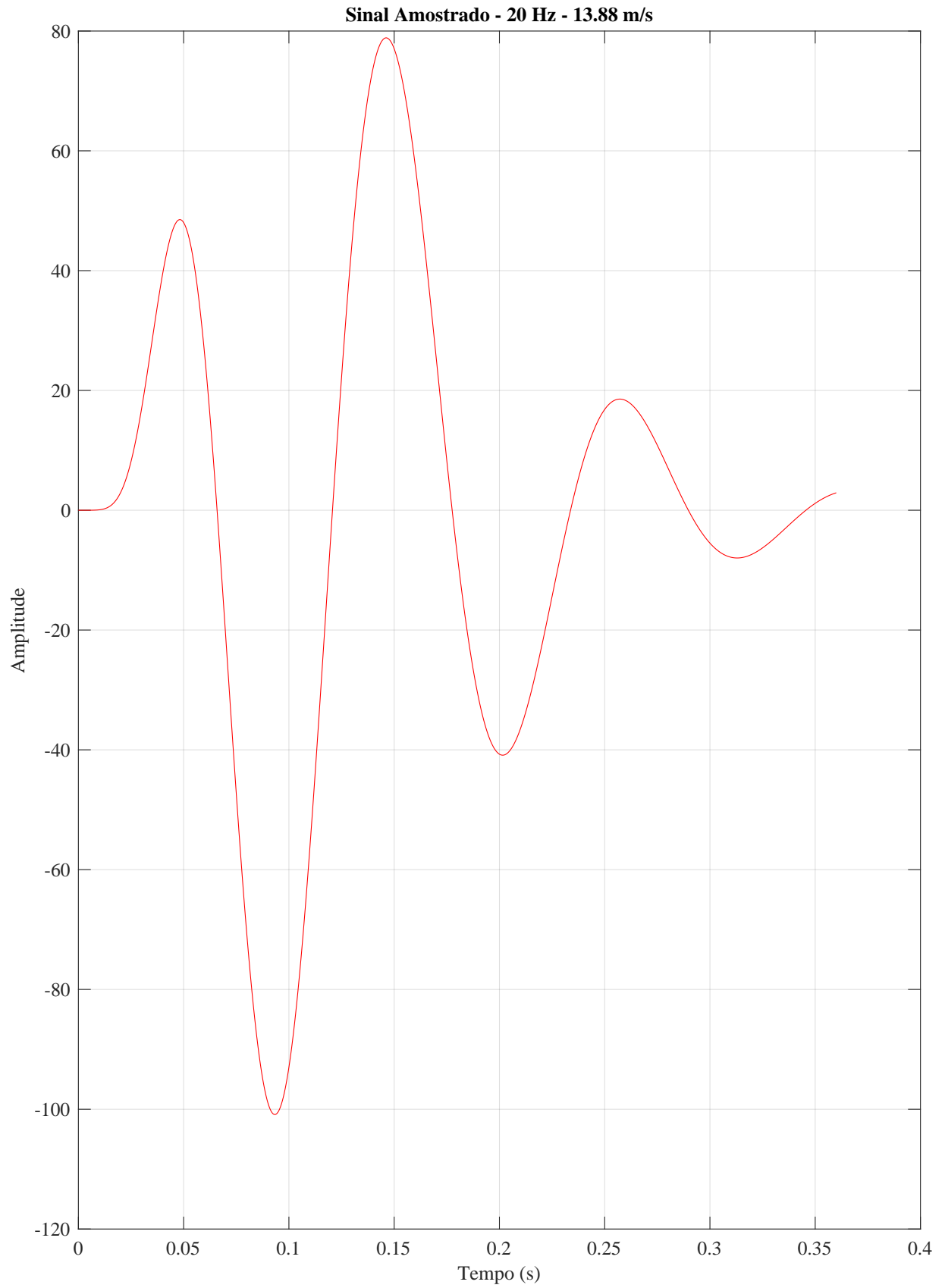


Figura 47 – Resposta Transitória amostrada a 20 Hz - 50 km/h. (Fonte: Autor 2024)

Para sumarizar a análise realizada a Figura 48 sugere três regiões de frequência para a detecção de anomalias utilizando sensores inerciais acoplados na região não suspensa do veículo em questão, onde fica a cargo do arquiteto da solução escolher a faixa de trabalho conforme sua necessidade, sendo altamente recomendado utilizar a frequência de Nyquist apenas quando a necessidade é a captura completa do sinal.

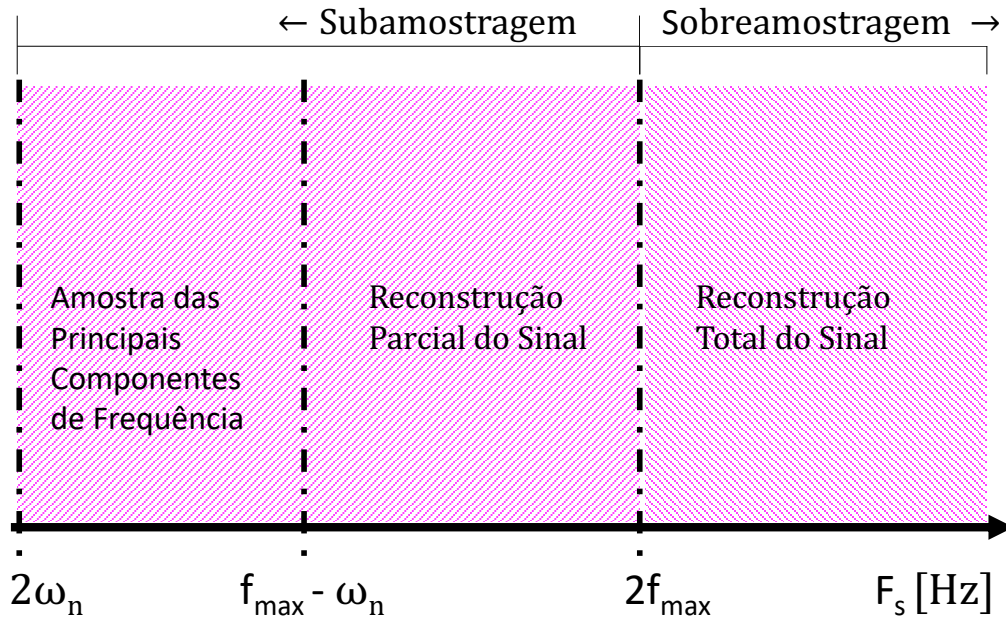


Figura 48 – Regiões para frequência de amostragem. (Fonte: Autor 2024)

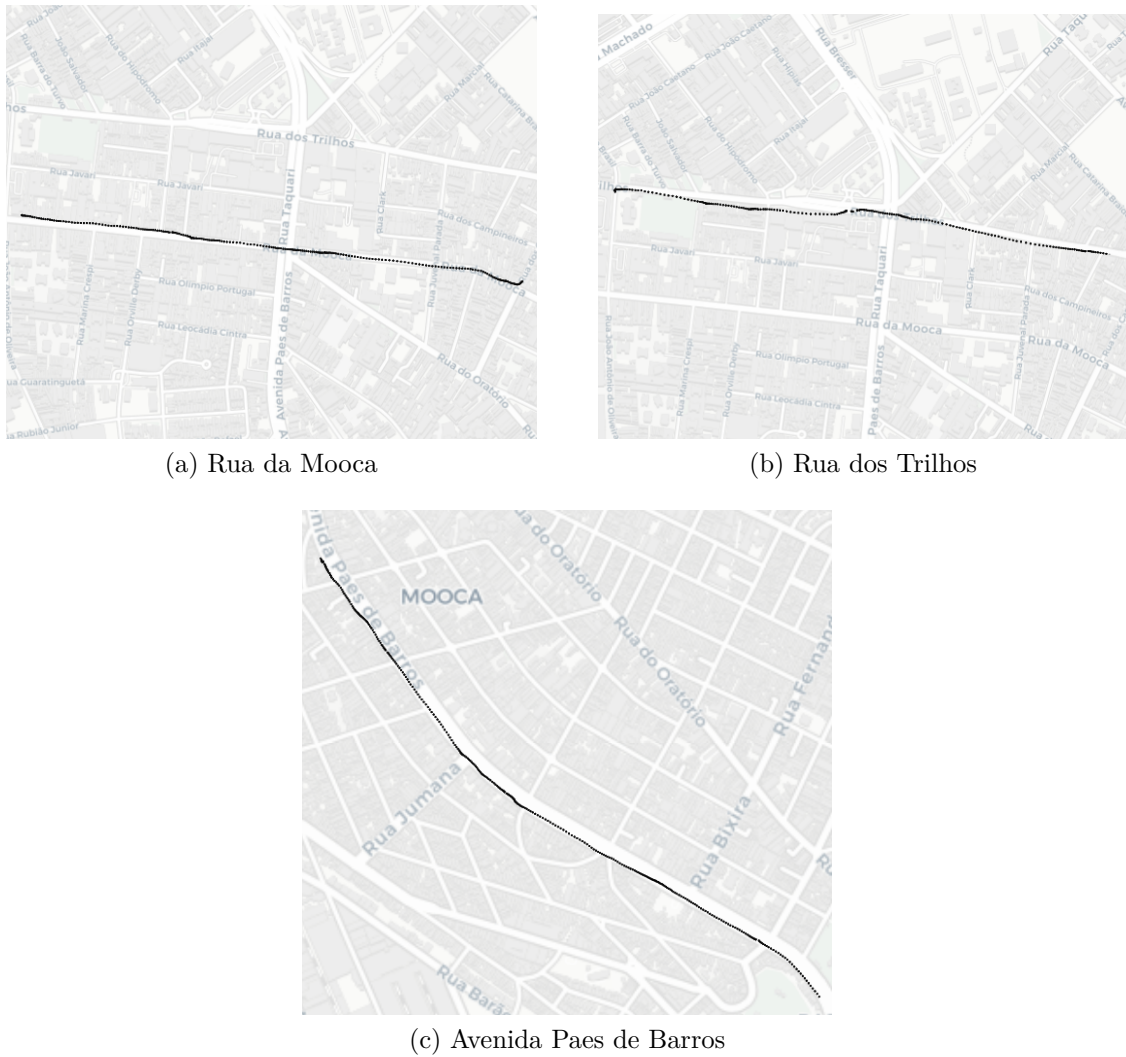
Portanto, para fins de comprovação no presente trabalho será utilizado a menor frequência proposta ($f_s = 20Hz$) para a detecção de anomalias no pavimento.

5.3 Análise no *Dataset* de Treino

Para a coleta do *dataset* de treino foi realizado um passeio em três ruas da cidade de São Paulo conforme mostra a Figura 49, no qual o carro equipado com os sensores coletando dados a $f_s = 20Hz$ trafegou com diferentes velocidades em vias que seguramente não possuem anomalias, criando assim uma base de dados para treino de algoritmos UL baseados em reconstrução.

No momento da aquisição dos dados o motorista teve a necessidade de habitar a gravação no início da viagem e desabilitar ao término, gerando assim um arquivo para cada rua trafegada.

É importante ressaltar que esse *dataset* não possui classificação de anomalias.



(a) Rua da Mooca

(b) Rua dos Trilhos

(c) Avenida Paes de Barros

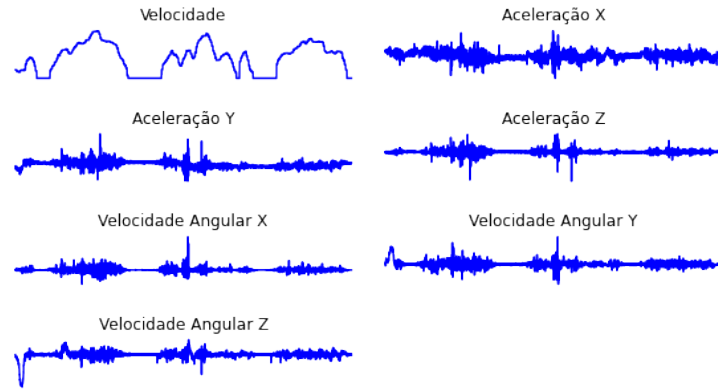
Figura 49 – Vias boas para treino de máquina. (Fonte: Autor 2024)

Os dados gerados foram armazenadas em seis arquivos distintos como mostra a Tabela 11.

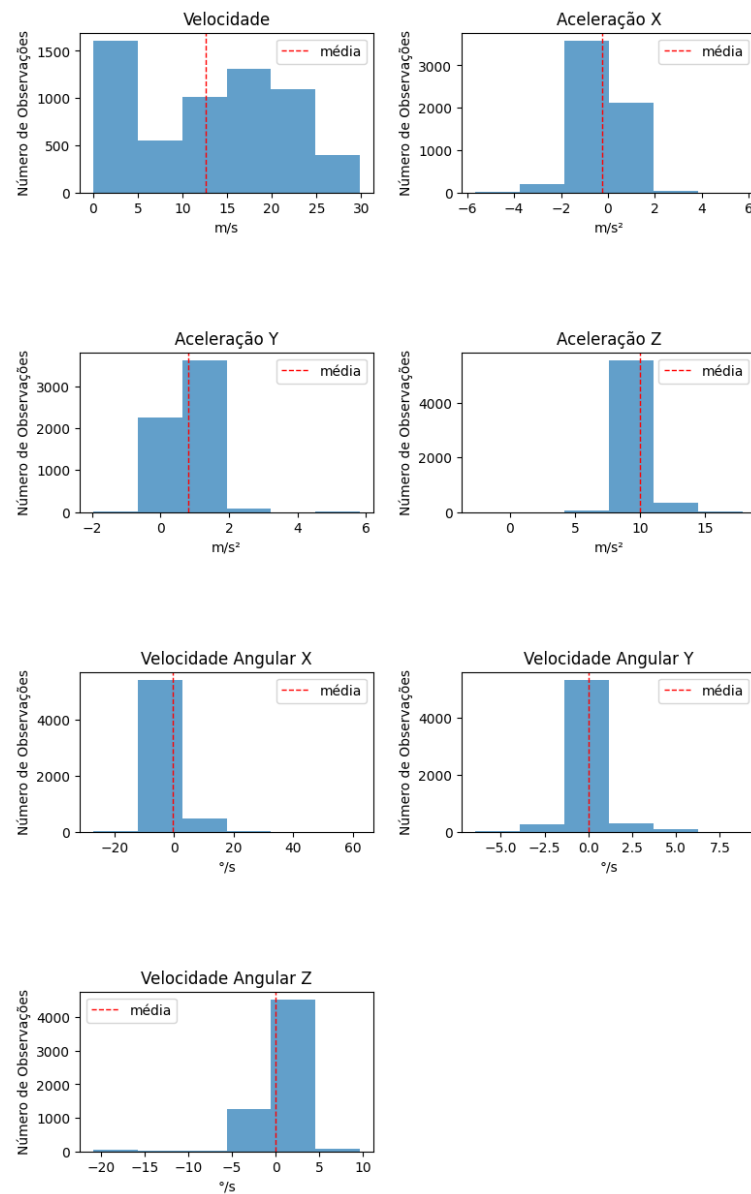
Tabela 11 – Arquivos de treino. (Fonte: Autor 2024)

Via	Nome do Arquivo	Nº Linhas
Rua da Mooca	001-SP-BOM-001	5978
Rua da Mooca	001-SP-BOM-003	4731
Rua da Mooca	001-SP-BOM-005	5805
Rua dos Trilhos	001-SP-BOM-002	7144
Rua dos Trilhos	001-SP-BOM-004	5805
Avenida Paes de Barros	001-SP-BOM-006	8683

A Figura 50, Figura 51, Figura 52, Figura 53, Figura 54 e Figura 55, mostram as análises de cada viagem, explicitando o histograma das variáveis de interesse acompanhado pelo seu valor médio.

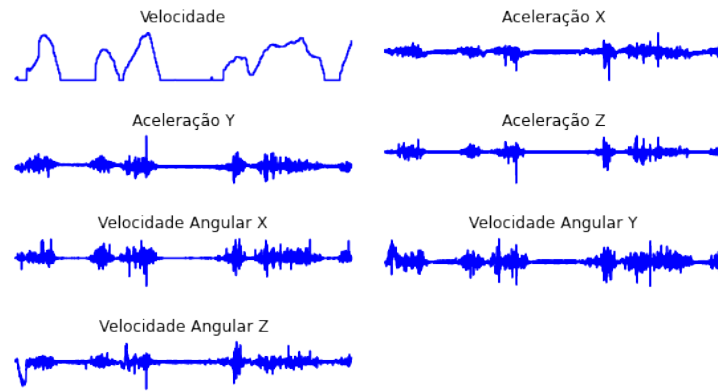


(a) Features - Rua da Mooca

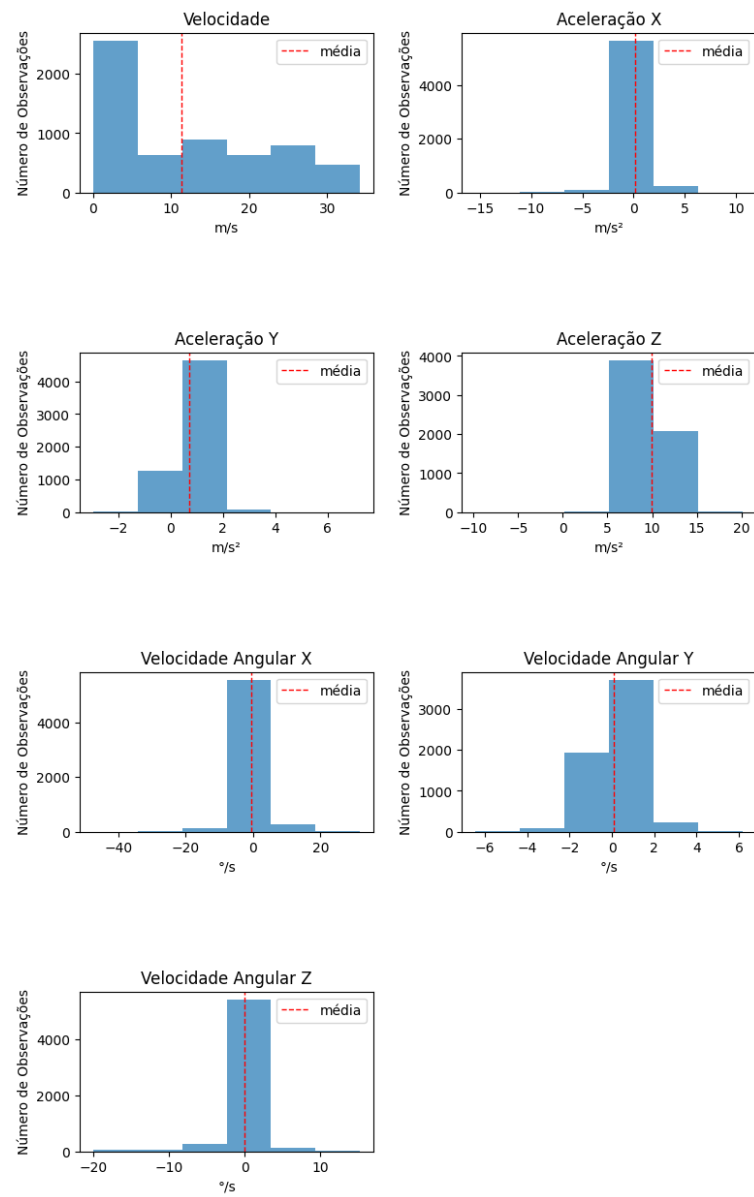


(b) Histograma das Features - Rua da Mooca

Figura 50 – Análise do arquivo 001 - Rua da Mooca. (Fonte: Autor 2024)

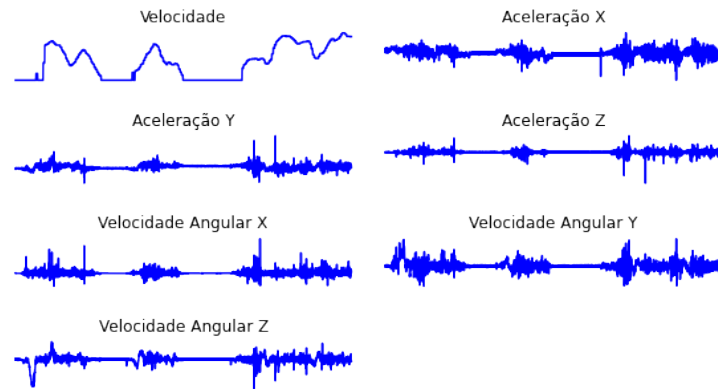


(a) Features - Rua dos Trilhos

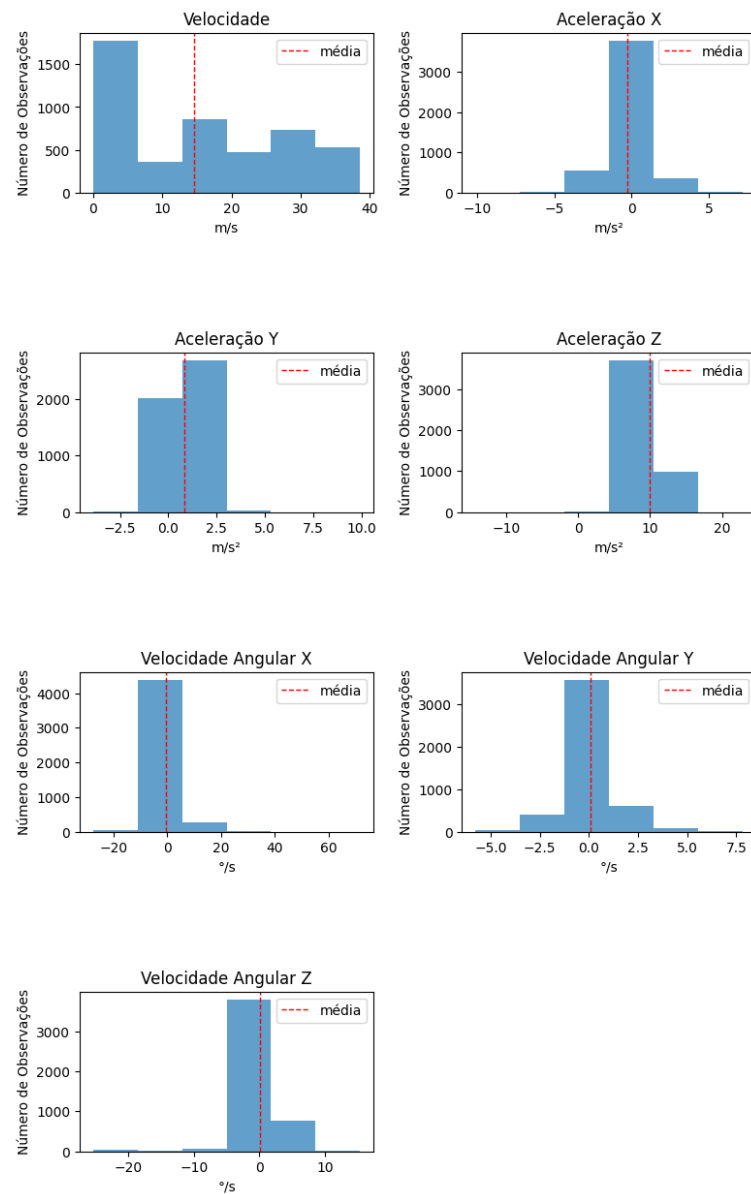


(b) Histograma das Features - Rua dos Trilhos

Figura 51 – Análise do arquivo 002 - Rua dos Trilhos. (Fonte: Autor 2024)

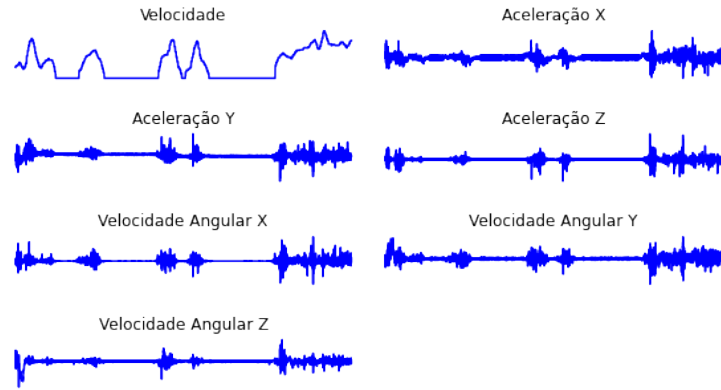


(a) Features - Rua da Mooca

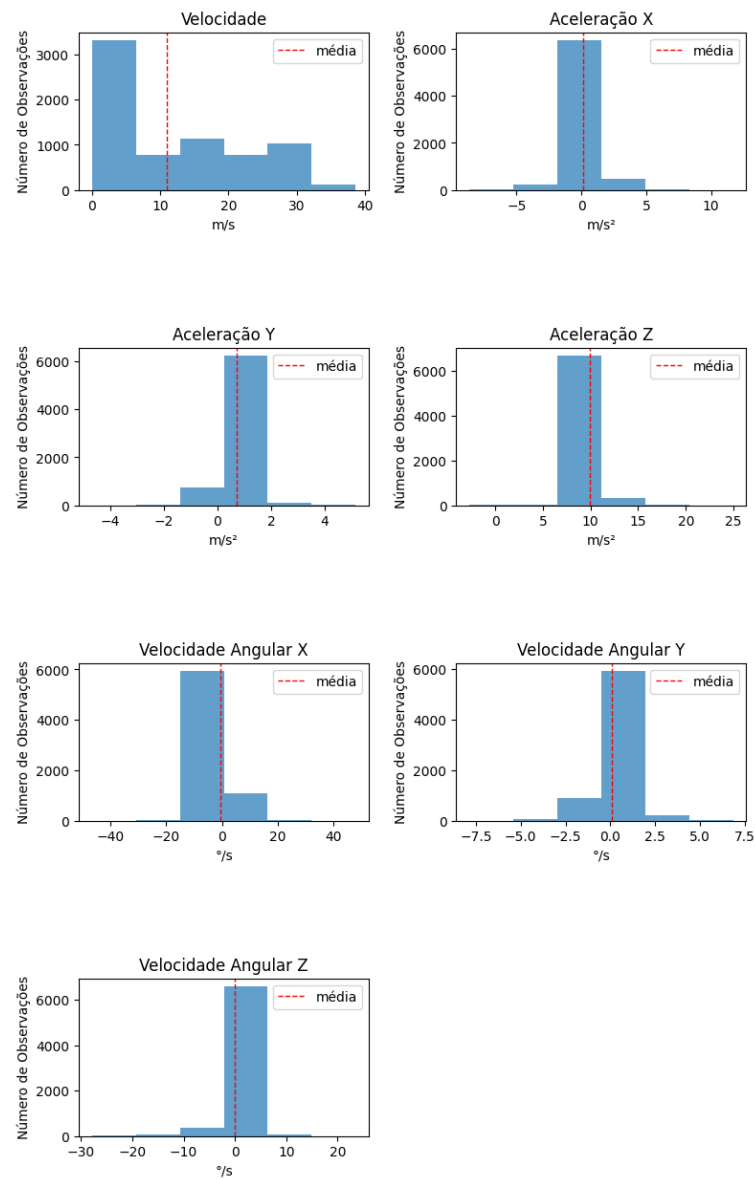


(b) Histograma das Features - Rua da Mooca

Figura 52 – Análise do arquivo 003 - Rua da Mooca. (Fonte: Autor 2024)

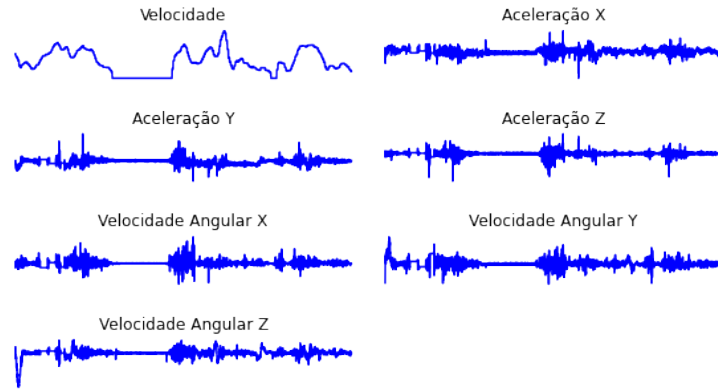


(a) Features - Rua dos Trilhos

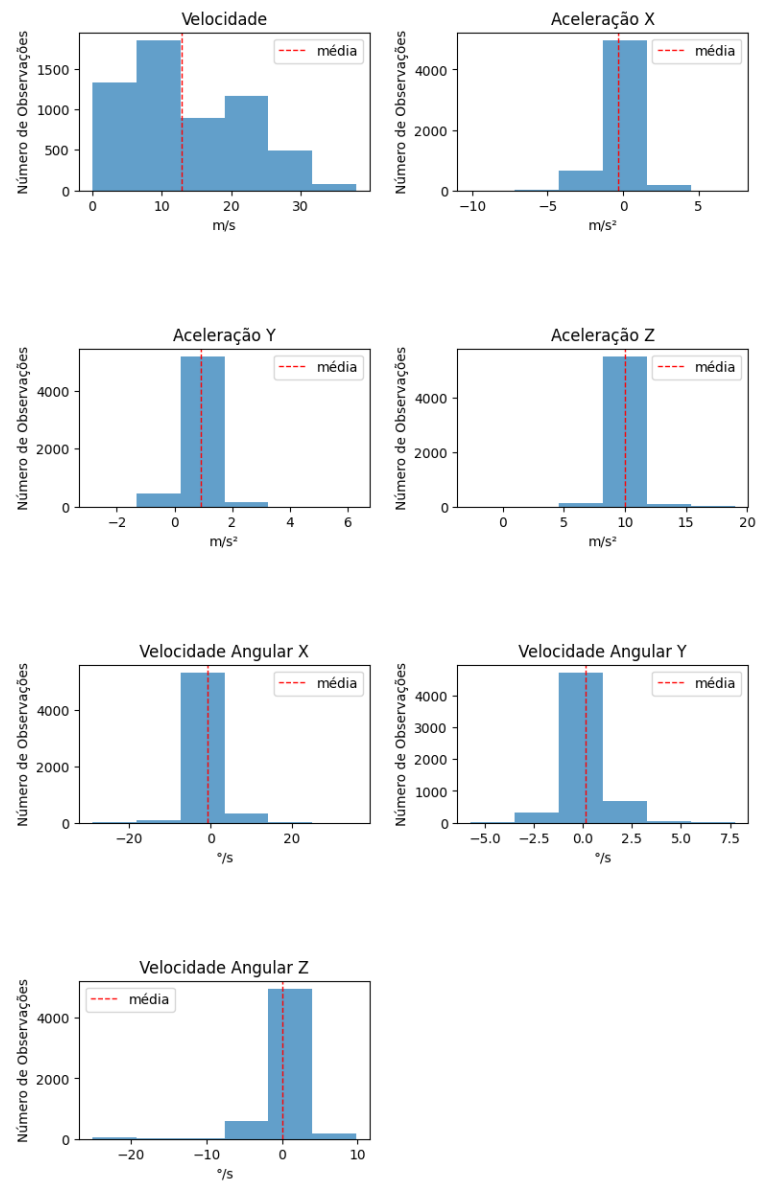


(b) Histograma das Features - Rua dos Trilhos

Figura 53 – Análise do arquivo 004 - Rua dos Trilhos. (Fonte: Autor 2024)

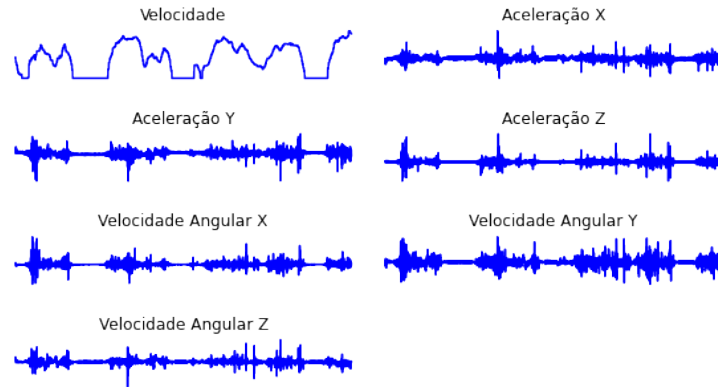


(a) Features - Rua da Mooca

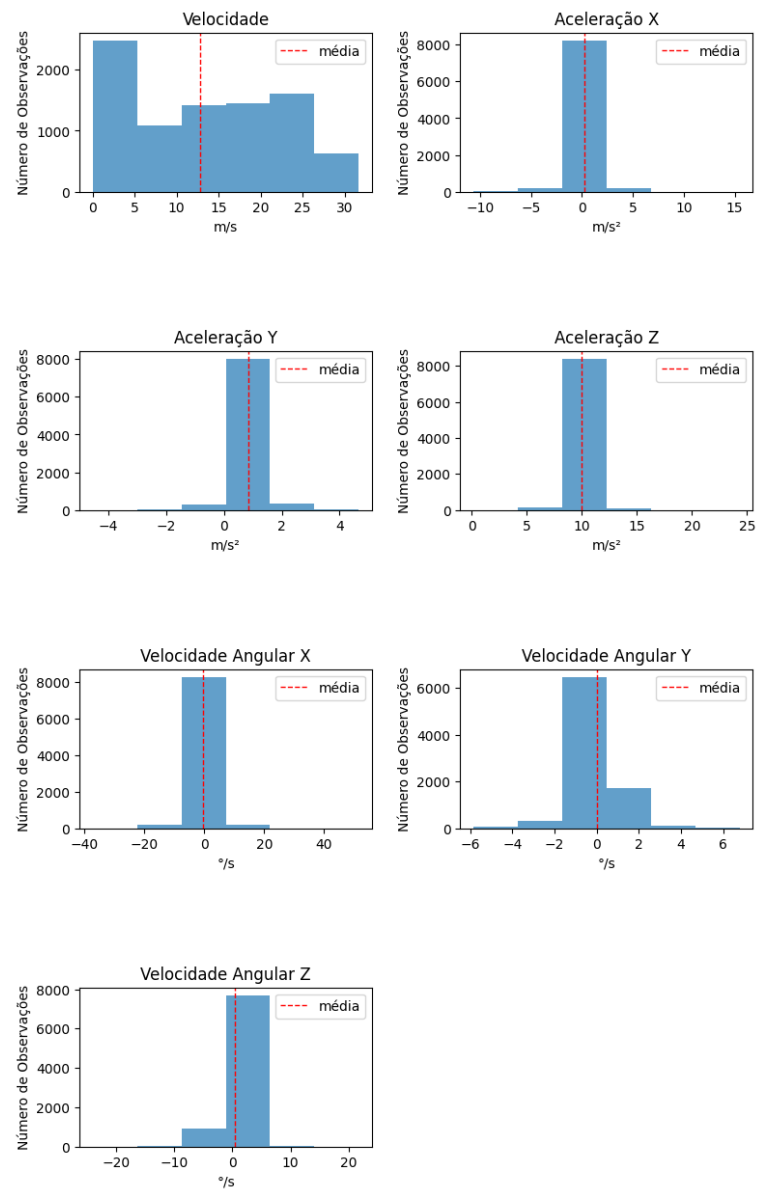


(b) Histograma das Features - Rua da Mooca

Figura 54 – Análise do arquivo 005 - Rua da Mooca. (Fonte: Autor 2024)



(a) Features - Avenida Paes de Barros



(b) Histograma das Features - Avenida Paes de Barros

Figura 55 – Análise do arquivo 005 - Avenida Paes de Barros. (Fonte: Autor 2024)

5.4 Análise no *Dataset* de Teste

Para esta base foram coletadas 285 amostras (vide Apêndice E), no qual um passageiro classificou quatro classes de interesse conforme descrito em Tabela 12, sendo três classes anômalas (buraco, lombada e sarjetão) e uma não anômala (asfalto regular), onde procurou-se trafegar em diversos tipos de pavimentos e em vias que não foram utilizadas para treino, garantindo assim a confiabilidade e imparcialidade do teste.

Tabela 12 – Quantidade de amostras - Dataset de teste. (Fonte: Autor 2024)

Amostra	Quantidade de arquivos	Quantidade de Anomalias/ Não Anomalias
Asfalto regular	123	123
Buraco	56	162
Lombada	52	
Sarjetão	54	

5.5 Módulo de Inteligência Artificial

Com a coleta dos dados realizadas, pode-se então utilizar o algoritmo de UL previamente definido para a construção da etapa de inteligência artificial, onde os resultados aqui apresentados estão divididos em: treino de máquina, limiar do erro e teste controlado.

5.5.1 Treino de Máquina

Na fase de treino além de expor a máquina as informações coletadas, o arquiteto da solução deve assegurar que o processo de aprendizado ocorra com o menor viés possível, ou seja, ele deve ser capaz de impedir que exista um sobre ajuste (do inglês *overfitting*) nos dados.

Desse modo, uma ferramenta bastante utilizada nesse momento é a análise gráfica, explicitando como se dá esse processo, conforme ilustrado na Figura 56, onde pode-se concluir que no treino em questão o erro do conjunto de validação permaneceu ligeiramente maior que o erro do conjunto de treino durante toda a fase de instrução de máquina, o que pode indicar *overfitting*, entretanto vale ressaltar que o nível do erro médio para os dois grupos é relativamente baixo, girando em torno de 0.025 para o conjunto de treino e 0.030 para validação, indicando assim que o sistema está apto a fazer previsões em novas amostras.

Ainda para a fase de treino pode-se elencar que o processo se deu em seis épocas, e após alcançar um valor de erro pré estabelecido o sistema atingiu o *early stop* e se auto finalizou.

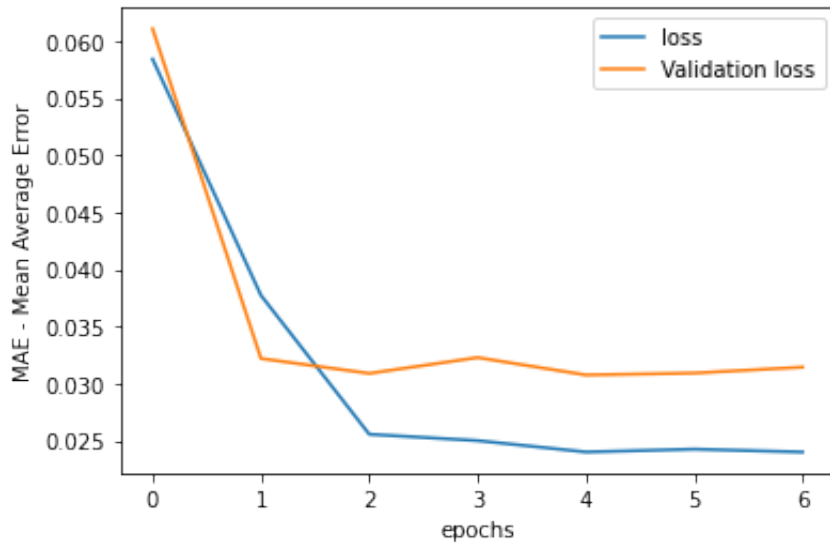


Figura 56 – Histórico do treino de máquina. (Fonte: Autor 2024)

5.5.2 Definição do Limiar de Erro

Após o treino de máquina deve-se definir o limiar de erro, onde realiza-se a inferência a partir dos dados de treino, contabilizando assim o erro de cada amostra, gerando um histograma como o evidenciado na Figura 57, onde pode-se observar que o limiar para se considerar um valor anômalo está definido como sendo 0.15, dado que não há amostras com erro maior que esse valor.

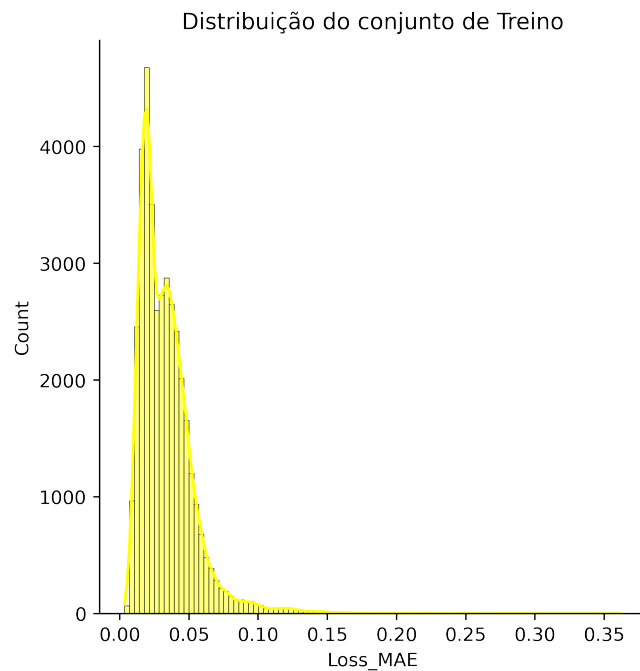


Figura 57 – Histograma para determinação do limiar do erro. (Fonte: Autor 2024)

5.5.3 Teste Controlado

Com o sistema de inteligência artificial devidamente treinado e o limiar do erro definido é realizado então o teste controlado, onde cada amostra do *dataset* de teste é pré processada conforme descrito anteriormente e inserida na RNA para que haja uma predição de sua condição, onde a máquina rotula se tal janela é ou não anômala. Por fim a conferência do resultado se dá confrontando a predição entregue pela máquina com o real valor encontrado em campo.

Logo, para a realização do teste são utilizados quatro critérios de avaliação:

$$precision = \frac{TP}{TP + FP} \quad (5.2)$$

Quantificando o acerto do sistema em relação a todos os resultados que foram classificados como anomalias.

$$recall = \frac{TP}{TP + FN} \quad (5.3)$$

Quantificando o acerto do sistema em relação a todos as anomalias que realmente existiam.

$$F1_{score} = \frac{2 \times precision \times recall}{precision + recall} \quad (5.4)$$

Sendo este a média harmônica entre o *precision* e *recall*.

$$acurácia = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.5)$$

Quantificando o acerto do sistema de forma global desde que não haja desequilíbrio entre as classes.

Desse modo, a Figura 58 traz uma matriz de confusão, evidenciando as quatro condições possíveis e necessárias (FP, FN, TP, TN) para se calcular as métricas acima expostas, onde pode-se atentar que os maiores valores nesse diagrama se concentram na diagonal principal, significando uma alta assertividade nos resultados preditos, sendo a classe "0" designada para representar os resultados de não anomalia e "1" para anomalia. No tocante a diagonal oposta é possível verificar que esse modelo possui um baixo índice de FP (9 predições), o que é bom em situações onde o sistema conta apenas com um estágio, entretanto quando adicionado o módulo de classificação um maior valor de FP na detecção pode ser tolerado, visto que haverá uma segunda análise.

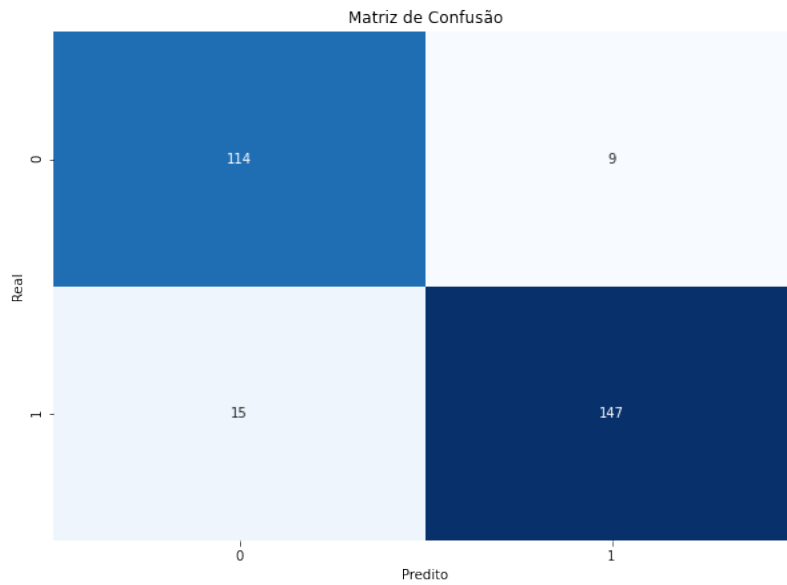


Figura 58 – Matriz de confusão. (Fonte: Autor 2024)

Portanto a Tabela 13 sumariza os resultados do módulo de detecção obtidos em teste, no qual pode-se observar um índice de *acurácia* e $F1_{score}$ muito próximo aos de trabalhos de detecção de anomalias no pavimento utilizando algoritmos de SL [82, 106], o que torna a utilização dos algoritmos de UL ideais para este tipo de problema, dado todas as vantagens já elencadas.

Tabela 13 – Métricas do sistema de inteligência artificial - Resultado. (Fonte: Autor 2024)

<i>Precision</i>	0.94
<i>Recall</i>	0.91
$F1_{score}$	0.92
<i>Acurácia</i>	0.92

6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

O trabalho aqui realizado teve como tema principal a utilização de sensores inerciais em carros para a detecção de anomalias no pavimento, de modo que ao fim da pesquisa possa se obter um sistema automático que possibilite ações proativas dos gestores de centros urbanos, como a localização e manutenção corretiva de buracos, a engenharia de tráfego através da aquisição da geolocalização de lombadas, a geração do inventário de ativos urbanos como sarjetões, etc. Neste contexto grandes dúvidas se levantaram em relação a arquitetura desse sistema, uma vez que na literatura há diferentes modos de operações com distintas tecnologias, entretanto a partir de estudos a artigos publicados verificou-se que um sistema de detecção seguido de um fase de classificação poderia mitigar os FP, uma vez que este seria um sistema de dupla verificação.

Assim, optou-se por focar no módulo de detecção, uma vez que a fase secundária esta amplamente difundida e os algoritmos que a compõe são de fácil acesso a engenheiros da computação. Desse modo, ao se pesquisar sobre métodos de detecção de anomalias percebeu-se que algoritmos de inteligência artificial de UL possuem diversas aplicações para problemas semelhantes em outras áreas do conhecimento com uma alta taxa de acerto, se mostrando assim promissor para a aplicação em pavimentos.

Outro problema em aberto na literatura está na frequência de amostragem desses sensores, uma vez que a escolha dessa grandeza é feita por tentativa e erro e, por inúmeras vezes é utilizado um valor maior que o necessário para a captura do sinal, sendo um problema para sistemas que utilizam dados de inúmeros carros, uma vez que a economia de armazenamento e transmissão dessas informações dependem diretamente da taxa de amostragem.

A pesquisa dividiu-se então em dois objetivos específicos, sendo o primeiro a criação de um método para a definição da frequência de amostragem para sensores inerciais e o segundo a validação de algoritmos de UL para a detecção de anomalias no pavimento, de modo que no final com a junção dessas partes têm-se o sistema de detecção completo.

O método para a definição da frequência de amostragem sugeriu o uso de um modelo matemático de 2DOF que descreve a região de acoplamento do sensor, permitindo simular as respostas do sistema, de modo que no final foram encontrados três regiões de frequência possíveis de atuação para se realizar a coleta dos dados, abrindo a possibilidade de se trabalhar com frequências de captação menores que a frequência de Nyquist.

Para o módulo de detecção das anomalias, optou-se por utilizar uma RNA *autoencoder* LSTM devido a sua topologia que permite a manipulação de dados temporais, onde foi obtido uma taxa de acerto muito próxima a de sistemas que fazem uso de algoritmos

de SL, viabilizando assim a sua aplicação.

Por fim, vale ressaltar que o sistema aqui descrito poder ter uma alta taxa de FN associado aos seus resultados, pois a coleta de um ponto anômalo depende exclusivamente do motorista passar o veículo sobre o elemento em questão, no qual esse déficit pode ser suplantado a medida em que se aumenta a frota de carros que dispõem do sensor inercial, elevando a probabilidade de detecção de um ponto na via pelo simples fato do aumento de tráfego, ou integrando diferentes tecnologias, transformando-o num sistema híbrido, como por exemplo a adição de uma câmera para a utilização de algoritmos de visão computacional.

6.1 Sugestões para Trabalhos Futuros

Ainda que atingido os objetivos da pesquisa, entende-se que os esforços nessa área devem continuar para que haja um sistema íntegro e a prova de falhas, que consiga trafegar o menor volume de dados possível e não apenas detectar as anomalias mas também as classificar. Desse modo é elencado algumas sugestões e observações para trabalhos futuros, são elas:

- Análise de outros algoritmos de UL para o módulo de detecção, explicitando as métricas para avaliação dos modelos e o desempenho obtido em cada caso.
- Incorporar o módulo de classificação das anomalias a arquitetura proposta, estudando o desempenho geral e comparando com outros sistemas de SL.
- Adicionar outros modelos com maiores DOF e não linearidade ao método proposto, os estudando sob a ótica de sistemas caóticos, a fim de se observar se há uma mudança significativa nas regiões de amostragem obtidas.
- Análise do impacto da inclusão de diferentes veículos no sistema, estudando qual a quantidade de passadas mínima em um elemento para se detectar uma anomalia, e se a mudança dos parâmetros de modelo e simulação afetam diretamente o método sugerido para a definição da frequência de amostragem.

REFERÊNCIAS

- [1] FESZCZUK, E. *Pothole city: How wild weather has blown out road repair budget*. 2022. Disponível em: <<http://web.archive.org/web/20230113022344/https://westernweekender.com.au/2022/12/pothole-city-how-wild-weather-has-blown-out-road-repair-budget/>>. Acesso em: 13.01.2023.
- [2] ALLIANCE, A. I. *Annual Local Authority Road Maintenance Survey*. 2022. Disponível em: <<http://web.archive.org/web/20230113032436/https://www.asphaltuk.org/wp-content/uploads/ALARM-survey-2022-FINAL.pdf>>. Acesso em: 13.01.2023.
- [3] PREZ, M. de. *Pothole repairs to cost £12bn and take 9 years, report suggests*. 2022. Disponível em: <<https://www.fleetnews.co.uk/news/car-industry-news/2022/03/22/pothole-repairs-to-cost-12bn-and-take-9-years-report-suggests>>. Acesso em: 13.01.2023.
- [4] MARTINEZ-RÍOS, E. A.; BUSTAMANTE-BELLO, M. R.; ARCE-SÁENZ, L. A. A review of road surface anomaly detection and classification systems based on vibration-based techniques. *Applied Sciences*, v. 12, n. 19, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/19/9413>>.
- [5] ŽALIMIENĖ, L.; LAURINAVIČIUS, A.; VAITKUS, A. The defects of asphalt pavements and their impact to driving quality in vilnius city streets. *7th International Conference on Environmental Engineering, ICEE 2008 - Conference Proceedings*, p. 1247–1252, 01 2008.
- [6] KIM, Y.-M. et al. Review of recent automated pothole-detection methods. *Applied Sciences*, v. 12, n. 11, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/11/5320>>.
- [7] MARTINELLI, A. et al. Road surface anomaly assessment using low-cost accelerometers: A machine learning approach. *Sensors*, v. 22, n. 10, 2022. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/22/10/3788>>.
- [8] PHAM, D.; SAGIROGLU, S. Processing signals from an inertial sensor using neural networks. *International Journal of Machine Tools and Manufacture*, v. 36, n. 11, p. 1291–1306, 1996. ISSN 0890-6955. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0890695595000879>>.
- [9] LEE, T.; CHUN, C.; RYU, S. Detection of road-surface anomalies using a smartphone camera and accelerometer. *Sensors*, v. 21, p. 01–17, 2021.
- [10] *A smartphone based method to enhance road pavement anomaly detection by analyzing the driver behavior*. Osaka, Japan: Association for Computing Machinery, 2015. 1169–1177 p.
- [11] LI, X. et al. Embracing crowdsensing: An enhanced mobile sensing solution for road anomaly detection. *ISPRS International Journal of Geo-Information*, v. 8, n. 9, 2019. ISSN 2220-9964. Disponível em: <<https://www.mdpi.com/2220-9964/8/9/412>>.

- [12] NGUYEN, T.; WONG, Y.; LECHNER, B. Response-based methods to measure road surface irregularity: a state-of-the-art review. *European Transport Research Review*, v. 11, 09 2019.
- [13] VILLA-PÉREZ, M. E. et al. Semi-supervised anomaly detection algorithms: A comparative summary and future research directions. *Knowledge-Based Systems*, v. 218, p. 106878, 2021. ISSN 0950-7051. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0950705121001416>>.
- [14] EGAJI, O. A. et al. Real-time machine learning-based approach for pothole detection. *Expert Systems with Applications*, v. 184, p. 115562, 2021. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417421009684>>.
- [15] THILL, M. et al. Temporal convolutional autoencoder for unsupervised anomaly detection in time series. *Applied Soft Computing*, v. 112, p. 107751, 2021. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494621006724>>.
- [16] KANARACHOS, S. et al. Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and hilbert transform. *Expert Systems with Applications*, v. 85, p. 292–304, 2017. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417417302737>>.
- [17] LIU, P. et al. Arrhythmia classification of lstm autoencoder based on time series anomaly detection. *Biomedical Signal Processing and Control*, v. 71, p. 103228, 2022. ISSN 1746-8094. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1746809421008259>>.
- [18] BAUR, C. et al. *Autoencoders for Unsupervised Anomaly Segmentation in Brain MR Images: A Comparative Study*. arXiv, 2020. Disponível em: <<https://arxiv.org/abs/2004.03271>>.
- [19] K, N. et al. Detection of overdose and underdose prescriptions—an unsupervised machine learning approach. *PLOS ONE*, v. 16(11): e0260315, p. 1–14, 2021. Disponível em: <<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0260315>>.
- [20] ZAMINI, M.; MONTAZER, G. Credit card fraud detection using autoencoder based clustering. In: *2018 9th International Symposium on Telecommunications (IST)*. [S.l.: s.n.], 2018. p. 486–491.
- [21] KARIMIPOUR, H. et al. A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids. *IEEE Access*, v. 7, p. 80778–80788, 2019.
- [22] DONNER, P.; LEGER, A. S.; BLAINE, R. Unsupervised machine learning for anomaly detection in synchrophasor network traffic. In: *2019 North American Power Symposium (NAPS)*. [S.l.: s.n.], 2019. p. 1–6.
- [23] NGUYEN, H. et al. Forecasting and anomaly detection approaches using lstm and lstm autoencoder techniques with the applications in supply chain management. *International Journal of Information Management*, v. 57, p. 102282, 2021. ISSN

- 0268-4012. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S026840122031481X>>.
- [24] LÓPEZ, D. et al. Fusing anomaly detection with false positive mitigation methodology for predictive maintenance under multivariate time series. *Information Fusion*, v. 100, p. 101957, 2023. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253523002737>>.
- [25] SOUZA, V. M. Asphalt pavement classification using smartphone accelerometer and complexity invariant distance. *Engineering Applications of Artificial Intelligence*, v. 74, p. 198–211, 2018. ISSN 0952-1976. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0952197618301349>>.
- [26] PAIXÃO, M. J. S.; MARANHÃO, F. L. Geovista – gestão descentralizada das condições do pavimento em ambiente urbano / geovista - decentralized management of pavement conditions in an urban environment. *Brazilian Journal of Development*, v. 6, n. 7, p. 53235–53244, Jul. 2020. Disponível em: <<https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/14158>>.
- [27] CHEN, S.; GUO, W. Auto-encoders in deep learning mdash a review with new perspectives. *Mathematics*, v. 11, n. 8, 2023. ISSN 2227-7390. Disponível em: <<https://www.mdpi.com/2227-7390/11/8/1777>>.
- [28] HASSAAN, G. A. Car dynamics using quarter model and passive suspension, part i: Effect of suspension damping and car speed. In: . [s.n.], 2014. Disponível em: <<https://api.semanticscholar.org/CorpusID:16786012>>.
- [29] Zhirui Qin, Z. Z.; OHORE, O. E. Research trends and hotspots of aquatic biofilms in freshwater environment during the last three decades: a critical review and bibliometric analysis. *Environmental Science and Pollution Research*, v. 29, n. 1, p. 47915–47930, 2022.
- [30] Çağlayan Akay E. YILMAZ SOYDAN, N.; GACAR, K. Bibliometric analysis of the published literature on machine learning in economics and econometrics. *Social Network Analysis and Mining*, v. 109, n. 12, p. 01–20, 2022.
- [31] Huang J, W. X.; S, L. A bibliometric and content analysis of research trends on gis-based landslide susceptibility from 2001 to 2020. *Environmental Science and Pollution Research*, v. 29, n. 1, p. 86954–86993, 2022.
- [32] *The pothole patrol: using a mobile sensor network for road surface monitoring*. Breckenridge CO, USA: Association for Computing Machinery, 2008. 29–39 p.
- [33] MOHAN, P.; PADMANABHAN, V.; RAMJEE, R. Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In: . [S.l.: s.n.], 2008. p. 323–336.
- [34] SUN, W.; ZHAO, Z.; GAO, H. Saturated adaptive robust control for active suspension systems. *IEEE Transactions on Industrial Electronics*, v. 60, n. 9, p. 3889–3896, 2013.
- [35] NGUYEN, Q.-H.; CHOI, S.-B. Optimal design of mr shock absorber and application to vehicle suspension. *Smart Materials and Structures*, v. 18, n. 3, p. 035012, feb 2009. Disponível em: <<https://dx.doi.org/10.1088/0964-1726/18/3/035012>>.

- [36] DU, H.; LI, W.; ZHANG, N. Semi-active variable stiffness vibration control of vehicle seat suspension using an mr elastomer isolator. *Smart Materials and Structures*, v. 20, n. 10, p. 105003, aug 2011. Disponível em: <<https://dx.doi.org/10.1088/0964-1726/20/10/105003>>.
- [37] BHORASKAR, R. et al. Wolverine: Traffic and road condition estimation using smartphone sensors. In: *2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012)*. [S.l.: s.n.], 2012. p. 1–6.
- [38] PANG, G. et al. Deep learning for anomaly detection: A review. *CoRR*, abs/2007.02500, 2020. Disponível em: <<https://arxiv.org/abs/2007.02500>>.
- [39] LIU, F. T.; TING, K. M.; ZHOU, Z.-H. Isolation-based anomaly detection. Association for Computing Machinery, New York, NY, USA, v. 6, n. 1, mar 2012. ISSN 1556-4681. Disponível em: <<https://doi.org/10.1145/2133360.2133363>>.
- [40] MEHROTRA, K. G.; MOHAN, C. K.; HUANG, H. *Anomaly Detection Principles and Algorithms*. [S.l.]: Springer Cham, 2017.
- [41] FU, C.; YANG, J. Granular classification for imbalanced datasets: A minkowski distance-based method. *Algorithms*, v. 14, p. 54, 02 2021.
- [42] YEPMO, V.; SMITS, G.; PIVERT, O. Anomaly explanation: A review. *Data and Knowledge Engineering*, v. 137, p. 101946, 2022. ISSN 0169-023X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169023X21000720>>.
- [43] BREUNIG, M. M. et al. Lof: Identifying density-based local outliers. *SIGMOD Rec.*, Association for Computing Machinery, New York, NY, USA, v. 29, n. 2, p. 93–104, may 2000. ISSN 0163-5808. Disponível em: <<https://doi.org/10.1145/335191.335388>>.
- [44] TANG, J. et al. Enhancing effectiveness of outlier detections for low density patterns. In: . Berlin, Heidelberg: Springer-Verlag, 2002. (PAKDD '02). ISBN 3540437045.
- [45] LLOYD, S. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, v. 28, n. 2, p. 129–137, 1982.
- [46] CHENG, Z. et al. An improved knn classification algorithm based on sampling. In: *Proceedings of the Advances in Materials, Machinery, Electrical Engineering (AMMEE 2017)*. Atlantis Press, 2017/06. p. 220–225. ISBN 978-94-6252-350-0. ISSN 2352-5401. Disponível em: <<https://doi.org/10.2991/ammee-17.2017.45>>.
- [47] CORTES, C.; VAPNIK, V. Support vector networks. *Machine Learning*, v. 20, p. 273–297, 1995.
- [48] SCHÖLKOPF, B. et al. Estimating the support of a high-dimensional distribution. *Neural Comput.*, MIT Press, Cambridge, MA, USA, v. 13, n. 7, p. 1443–1471, jul 2001. ISSN 0899-7667. Disponível em: <<https://doi.org/10.1162/089976601750264965>>.
- [49] FILHO, J. E. D. A. et al. A review of neural networks for anomaly detection. *IEEE Access*, v. 10, p. 112342–112367, 2022.

- [50] ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, v. 65, n. 6, p. 386–408, 1958. ISSN 0033-295X. Disponible em: <<http://dx.doi.org/10.1037/h0042519>>.
- [51] HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, v. 2, n. 5, p. 359–366, 1989. ISSN 0893-6080. Disponible em: <<https://www.sciencedirect.com/science/article/pii/0893608089900208>>.
- [52] LECUN, Y.; BENGIO, Y. Convolutional networks for images, speech and time series. In: _____. *The Handbook of Brain Theory and Neural Networks*. [S.l.: s.n.], 1995.
- [53] CHEN, Y.; LI, J. Recurrent neural networks algorithms and applications. In: *2021 2nd International Conference on Big Data and Artificial Intelligence and Software Engineering (ICBASE)*. [S.l.: s.n.], 2021. p. 38–43.
- [54] BERGMAN, L.; HOSHEN, Y. *Classification-Based Anomaly Detection for General Data*. arXiv, 2020. Disponible em: <<https://arxiv.org/abs/2005.02359>>.
- [55] HAGEMANN, T.; KATSAROU, K. Reconstruction-based anomaly detection for the cloud: A comparison on the yahoo! webscope s5 dataset. In: . New York, NY, USA: Association for Computing Machinery, 2020. (ICCBDC '20), p. 68–75. ISBN 9781450375382. Disponible em: <<https://doi.org/10.1145/3416921.3416934>>.
- [56] CHEN, Z. et al. Autoencoder-based network anomaly detection. In: *2018 Wireless Telecommunications Symposium (WTS)*. [S.l.: s.n.], 2018. p. 1–5.
- [57] HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. *Science*, v. 313, n. 5786, p. 504–507, 2006. Disponible em: <<https://www.science.org/doi/abs/10.1126/science.1127647>>.
- [58] SRIVASTAVA, N.; MANSIMOV, E.; SALAKHUTDINOV, R. *Unsupervised Learning of Video Representations using LSTMs*. arXiv, 2015. Disponible em: <<https://arxiv.org/abs/1502.04681>>.
- [59] MA, N. et al. Computer vision for road imaging and pothole detection: a state-of-the-art review of systems and algorithms. *Transportation Safety and Environment*, v. 4, n. 4, 11 2022. ISSN 2631-4428. Tdac026. Disponible em: <<https://doi.org/10.1093/tse/tdac026>>.
- [60] AKAGIC, A.; BUZA, E.; OMANOVIC, S. Pothole detection: An efficient vision based method using rgb color space image segmentation. In: *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. [S.l.: s.n.], 2017. p. 1104–1109.
- [61] BUZA, E.; OMANOVIC, S.; HUSEINOVIĆ, A. Pothole detection with image processing and spectral clustering. In: . [S.l.: s.n.], 2013.
- [62] JAKŠTYS, V. et al. Detection of the road pothole contour in raster images. *Information Technology and Control*, v. 45, n. 3, p. 300–307, 2016.

- [63] KOCH, C.; BRILAKIS, I. Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, v. 25, n. 3, p. 507–515, 2011. ISSN 1474-0346. Special Section: Engineering informatics in port operations and logistics. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1474034611000036>>.
- [64] PEREIRA, V. et al. Semantic segmentation of paved road and pothole image using u-net architecture. In: *2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*. [S.l.: s.n.], 2019. p. 1–4.
- [65] J, D. et al. Deep learning based detection of potholes in indian roads using yolo. In: *2020 International Conference on Inventive Computation Technologies (ICICT)*. [S.l.: s.n.], 2020. p. 381–385.
- [66] PAN, Y. et al. Object-based and supervised detection of potholes and cracks from the pavement images acquired by uav. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W4, p. 209–217, 2017. Disponível em: <<https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-4-W4/209/2017/>>.
- [67] FAN, R.; AI, X.; DAHNOUN, N. Road surface 3d reconstruction based on dense subpixel disparity map estimation. *IEEE Transactions on Image Processing*, v. 27, n. 6, p. 3025–3035, 2018.
- [68] FAN, R. et al. Real-time stereo vision for road surface 3-d reconstruction. In: *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*. [S.l.: s.n.], 2018. p. 1–6.
- [69] SCHNEBELE, E. et al. Review of remote sensing methodologies for pavement management and assessment. *European Transport Research Review*, v. 7, p. 1866–8887, 2015. Disponível em: <<https://doi.org/10.1007/s12544-015-0156-6>>.
- [70] ZHANG, D. et al. Automatic pavement defect detection using 3d laser profiling technology. *Automation in Construction*, v. 96, p. 350–365, 2018. ISSN 0926-5805. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0926580518304813>>.
- [71] ALI, A. et al. Drivable area segmentation in deteriorating road regions for autonomous vehicles using 3d lidar sensor. In: *2021 IEEE Intelligent Vehicles Symposium (IV)*. [S.l.: s.n.], 2021. p. 845–852.
- [72] LI, Y.; PAPACHRISTOU, C.; WEYER, D. Road pothole detection system based on stereo vision. In: *NAECON 2018 - IEEE National Aerospace and Electronics Conference*. [S.l.: s.n.], 2018. p. 292–297.
- [73] DU, Y. et al. A pothole detection method based on 3D point cloud segmentation. In: JIANG, X.; FUJITA, H. (Ed.). *Twelfth International Conference on Digital Image Processing (ICDIP 2020)*. SPIE, 2020. v. 11519, p. 1151909. Disponível em: <<https://doi.org/10.1117/12.2573124>>.
- [74] MOAZZAM, I. et al. Metrology and visualization of potholes using the microsoft kinect sensor. In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. [S.l.: s.n.], 2013. p. 1284–1291.

- [75] LION, K. M.; KWONG, K. H.; LAI, W. K. Smart speed bump detection and estimation with kinect. In: *2018 4th International Conference on Control, Automation and Robotics (ICCAR)*. [S.l.: s.n.], 2018. p. 465–469.
- [76] S, Y.; J, C. Convergence of stereo vision-based multimodal yolos for faster detection of potholes. *Computers, Materials and Continua*, v. 73, p. 2821–2834, 2022. Disponível em: <<https://doi.org/10.32604/cmc.2022.027840>>.
- [77] BASAVARAJU, A. et al. A machine learning approach to road surface anomaly assessment using smartphone sensors. *IEEE Sensors Journal*, v. 20, n. 5, p. 2635–2647, 2020.
- [78] FOX, A. et al. Multi-lane pothole detection from crowdsourced undersampled vehicle sensor data. *IEEE Transactions on Mobile Computing*, v. 16, n. 12, p. 3417–3430, 2017.
- [79] BADUROWICZ, M.; KARZMAREK, P.; MONTUSIEWICZ, J. Fuzzy extensions of isolation forests for road anomaly detection. In: *2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. [S.l.: s.n.], 2021. p. 1–6.
- [80] BANSAL, K. et al. Deepbus: Machine learning based real time pothole detection system for smart transportation using iot. *Internet Technology Letters*, v. 3, n. 3, p. e156. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/itl2.156>>.
- [81] SINGH, G. et al. Smart patrolling: An efficient road surface monitoring using smartphone sensors and crowdsourcing. *Pervasive and Mobile Computing*, v. 40, p. 71–88, 2017. ISSN 1574-1192. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1574119216301262>>.
- [82] SERAJ, F. et al. Roads: A road pavement monitoring system for anomaly detection using smart phones. In: *International Workshop on Modeling Social Media International Workshop on Mining Ubiquitous and Social Environments International Workshop on Machine Learning for Urban Sensor Data*. Switzerland: Springer, 2016. p. 128–146.
- [83] VARONA, B.; MONTESERIN, A.; TEYSEYRE, A. A deep learning approach to automatic road surface monitoring and pothole detection. *Personal and Ubiquitous Computing*, v. 24, n. 4, p. 519–534, 2020.
- [84] PANDEY, A. K. et al. Deep neural networks based approach for pothole detection. In: *2021 4th International Conference on Signal Processing and Information Security (ICSPIS)*. [S.l.: s.n.], 2021. p. 1–4.
- [85] MEDNIS, A. et al. Real time pothole detection using android smartphones with accelerometers. In: *2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*. [S.l.: s.n.], 2011. p. 1–6.
- [86] XUE, G. et al. Pothole in the dark: Perceiving pothole profiles with participatory urban vehicles. *IEEE Transactions on Mobile Computing*, v. 16, n. 5, p. 1408–1419, 2017.
- [87] WANG, H. et al. A real-time pothole detection approach for intelligent transportation system. *Mathematical Problems in Engineering*, v. 2015, p. 1–7, 2015.

- [88] ANAISSI, A. et al. Smart pothole detection system using vehicle-mounted sensors and machine learning. *Journal of Civil Structural Health Monitoring*, v. 09, p. 91–102, 2019.
- [89] HEGAZY, S.; SHARAF, A. Ride comfort analysis using quarter car model. In: . [S.l.: s.n.], 2013.
- [90] NGUYEN, L. H.; HONG, K.-S.; PARK, S. Road-frequency adaptive control for semi-active suspension systems. *International Journal of Control Automation and Systems*, v. 8, p. 1029–1038, 10 2010.
- [91] MAHALA, M.; GADKARI, P.; DEB, A. Mathematical models for designing vehicles for ride comfort. In: . [S.l.: s.n.], 2009.
- [92] THITE S. BANVIDI, T. I. A. N.; BENNETT, L. Suspension parameter estimation in the frequency domain using a matrix inversion approach. *Vehicle System Dynamics*, Taylor and Francis, v. 49, n. 12, p. 1803–1822, 2011.
- [93] SOONG, M. F.; RAMLI, R.; SAIFIZUL, A. Between simplicity and accuracy: Effect of adding modeling details on quarter vehicle model accuracy. *PLOS ONE*, Public Library of Science, v. 12, p. 1–23, 06 2017. Disponível em: <<https://doi.org/10.1371/journal.pone.0179485>>.
- [94] HYUN, Y. *The effect of suspension damping on vehicle response to transient road inputs*. Tese (Doutorado) — University of Southampton, 2005. Disponível em: <<https://eprints.soton.ac.uk/465852/>>.
- [95] TALUKDAR, S. et al. Mathematical modeling in vehicle ride dynamics. In: . [S.l.: s.n.], 2012.
- [96] LI, K.; NAGAI, M. Control and evaluation of active suspension for mdof vehicle model. *JSAE Review*, v. 20, n. 3, p. 343–348, 1999. ISSN 0389-4304. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0389430499000193>>.
- [97] KIEU, T.; YANG, B.; JENSEN, C. S. Outlier detection for multidimensional time series using deep neural networks. In: *2018 19th IEEE International Conference on Mobile Data Management (MDM)*. [S.l.: s.n.], 2018. p. 125–134.
- [98] DESAI, R.; GUHA, A.; SESHU, P. A comparison of quarter, half and full car models for predicting vibration attenuation of an occupant in a vehicle. *Journal of Vibration Engineering and Technologies*, v. 9, 02 2021.
- [99] MANUAL do Proprietário Honda City 2013. [S.l.], 2013. Disponível em: <<https://web.archive.org/web/20230218022102/https://www.honda.com.br/sites/cbw/files/2016-08/City%202013.pdf>>. Acesso em: 17.2.2023.
- [100] GILLESPIE, T. D. *Fundamentals of Vehicle Dynamics*. [S.l.]: Society of Automotive Engineers, Inc., 1992.
- [101] RILL, G. *Road Vehicle Dynamics Fundamentals and Modeling*. [S.l.]: CRC Press, 2011.

- [102]OGATA, K. *Engenharia de Controle Moderno*. Pearson Universidades, 2010. ISBN 9788576058106. Disponível em: <<https://books.google.com.br/books?id=iL3FYgEACAAJ>>.
- [103]KUMAR, G. A. et al. Road quality management system using mobile sensors. In: *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*. [S.l.: s.n.], 2017. p. 1–6.
- [104]JR., L. M. P. F. *Estudo da dinâmica vertical de uma suspensão veicular do tipo Macpherson*. Dissertação (Mestrado) — Universidade de São Paulo, 2006.
- [105]KAWAGUCHI, Y. Anomaly detection based on feature reconstruction from subsampled audio signals. In: *2018 26th European Signal Processing Conference (EUSIPCO)*. [S.l.: s.n.], 2018. p. 2524–2528.
- [106]KYRIAKOU, C.; CHRISTODOULOU, S. E.; DIMITRIOU, L. Smartphone-based pothole detection utilizing artificial neural networks. *Journal of Infrastructure Systems*, v. 25, n. 3, p. 04019019, 2019. Disponível em: <<https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29IS.1943-555X.0000489>>.

Apêndices

APÊNDICE A – REVISÃO BIBLIOMÉTRICA - QUERIES

Fase 1 - Pesquisa para entender o cenário de detecção de anomalias

Query (WoS) TS=(ROAD AND ANOMALY AND DETECTION) OR TS=(ROAD AND ANOMALY AND CLASSIFICATION) OR TS=(PAVEMENT AND ANOMALY AND DETECTION) OR TS=(PAVEMENT AND ANOMALY AND CLASSIFICATION) OR TS=(STREET AND ANOMALY AND DETECTION) OR TS=(STREET AND ANOMALY AND CLASSIFICATION) OR TS = (ROAD AND POTHOLE AND DETECTION) OR TS = (STREET AND POTHOLE AND DETECTION) OR TS = (PAVEMENT AND POTHOLE AND DETECTION) OR TS = (ROAD AND BUMP AND DETECTION) OR TS = (STREET AND BUMP AND DETECTION) OR TS = (PAVEMENT AND BUMP AND DETECTION) OR TS = (ROAD AND GUTTER AND DETECTION) OR TS = (STREET AND GUTTER AND DETECTION) OR TS = (PAVEMENT AND GUTTER AND DETECTION)

Query (Scopus) ABS(road AND anomaly AND detection) OR ABS(road AND anomaly AND classification) OR ABS(pavement AND anomaly AND detection) OR ABS(pavement AND anomaly AND classification) OR ABS(street AND anomaly AND detection) OR ABS(street AND anomaly AND classification) OR ABS(road AND pothole AND detection) OR ABS(street AND pothole AND detection) OR ABS(pavement AND pothole AND detection) OR ABS(road AND bump AND detection) OR ABS(street AND bump AND detection) OR ABS(pavement AND bump AND detection) OR ABS(road AND gutter AND detection) OR ABS(street AND gutter AND detection) OR ABS(pavement AND gutter AND detection)

Fase 2 - Pesquisa para entender os principais artigos e autores de anomalias com sensores inerciais

Query (WoS) TS=(ROAD AND ANOMALY AND VIBRATION) OR TS=(PAVEMENT AND ANOMALY AND VIBRATION) OR TS=(STREET AND ANOMALY AND VIBRATION) OR TS=(ROAD AND POTHOLE AND VIBRATION) OR TS=(PAVEMENT AND POTHOLE AND VIBRATION) OR TS=(STREET AND POTHOLE AND VIBRATION) OR TS=(ROAD AND BUMP AND VIBRATION) OR TS=(PAVEMENT AND BUMP AND VIBRATION) OR TS=(STREET AND BUMP AND VIBRATION) OR TS=(ROAD AND GUTTER AND VIBRATION) OR TS=(PAVEMENT AND GUTTER AND VIBRATION) OR TS=(STREET AND GUTTER AND VIBRATION) OR TS=(ROAD AND ANOMALY AND ACCELEROMETER) OR TS=(PAVEMENT AND ANOMALY AND ACCELEROMETER) OR TS=(STREET AND ANOMALY AND ACCELEROMETER) OR TS=(ROAD AND POTHOLE AND ACCELEROMETER) OR TS=(PAVEMENT AND POTHOLE AND ACCELEROMETER) OR TS=(STREET AND POTHOLE AND ACCELEROMETER) OR TS=(ROAD AND BUMP AND ACCELEROMETER) OR TS=(PAVEMENT AND BUMP AND ACCELEROMETER) OR TS=(STREET AND BUMP AND ACCELEROMETER) OR TS=(ROAD AND GUTTER AND ACCELEROMETER) OR TS=(PAVEMENT AND GUTTER AND ACCELEROMETER) OR TS=(STREET AND GUTTER AND ACCELEROMETER)

Query (Scopus) ABS (road AND anomaly AND vibration) OR ABS (pavement AND anomaly AND vibration) OR ABS (street AND anomaly AND vibration) OR ABS (road AND pothole AND vibration) OR ABS (pavement AND pothole AND vibration) OR ABS (street AND pothole AND vibration) OR ABS (road AND bump AND vibration) OR ABS (pavement AND bump AND vibration) OR ABS (street AND bump AND vibration) OR ABS (road AND gutter AND vibration) OR ABS (pavement AND gutter AND vibration) OR ABS (street AND gutter AND vibration) OR ABS (road AND anomaly AND accelerometer) OR ABS (pavement AND anomaly AND accelerometer) OR ABS (street AND anomaly AND accelerometer) OR ABS (

APÊNDICE B – REVISÃO BIBLIOMÉTRICA - GRUPO DE PALAVRAS RELACIONADAS

Grupos WoS

Cluster 1

3d
3d asphalt surfaces
artificial intelligence
asphalt
asphalt pavement
automatic crack detection
blockchain
convolutional neural network
crack
crack detection
cracks
damage detection
data augmentation
deep convolutional neural network
deep learning
defect detection
depth
distress detection
generative adversarial network
gpr
ground penetrating radar
ground-penetrating radar
identification
image classification
image segmentation
images
inspection
neural network
neural-network
neural-networks
object detection
pavement assessment
pavement distress
pavement distress detection
pavement inspection
pavement management
pavement monitoring
pothole
prediction
principal component analysis
remote sensing
road crack detection
robustness
structural health monitoring
temperature
uav

vehicle detection
wireless sensor networks
yolov3

Cluster 2

accident detection
accidents
anomaly detection
big data
cloud computing
clustering
crowdsensing
cybersecurity
data fusion
data models
edge computing
event detection
gaussian mixture model
incident detection
intelligent transportation systems
internet of things
iot
model
models
networks
objects
predictive maintenance
regression
road anomaly detection
road surface monitoring
road traffic
smart cities
smart city
smart transportation
smartphone sensors
social media
state estimation
systems
things
time series
time-series
traffic congestion
unsupervised learning

Cluster 3

algorithm
automated vehicles
autonomous driving
autonomous vehicle
autonomous vehicles

cameras
estimation
extraction
feature extraction
lane detection
lidar
localization
monitoring
navigation
network
obstacle avoidance
obstacle detection
pedestrian detection
road vehicles
roads
safety
semantic segmentation
semantics
sensor fusion
speed bump
speed bump detection
stereo vision
surveillance
task analysis
three-dimensional displays
time series analysis
tracking
training
unmanned aerial vehicles
vehicle
video anomaly detection
visualization

Cluster 4

3d reconstruction
accelerometer
android
anomaly
automated detection
classification
crowdsourcing
detection
distress
flow
image
inertial sensors
intelligent transportation systems
machine learning
management

pattern recognition
pavement defect
recognition
reconstruction
road
road monitoring
road surface
segmentation
sensor
simulation
smartphone sensors
system
traffic
video
vision
wireless sensor network

Cluster 5

acceleration
attacks
authentication
automobiles
challenges
computational modeling
connected vehicles
convolutional neural network
dataset
framework
intelligent transportation system
internet
internet of vehicles
intrusion detection
intrusion detection system
lstm
misbehavior detection
optimization
random forest
real-time systems
scheme
sdn
secure
security
support vector machine
vanet
vanets
vehicle dynamics
vehicular ad hoc networks

Cluster 6

abnormalevent detection

adas
attention mechanism
camera
cnn
computer vision
convolutional neural network (cnn)
detection system
gps
image processing
neural networks
object tracking
pavement
pavement crack
pothole detection
potholes
raspberry pi
real-time
road damage
road damage detection
road maintenance
ultrasonic sensor
video surveillance
yolo

Cluster 7

accelerometers
algorithms
convolutional neural networks
design
features
hidden markov models
intelligent vehicles
mobile sensing
one-class svm
participatory sensing
road anomalies
road condition assessment
sensors
signal processing
smartphones
support vector machine (svm)
traffic anomaly detection
transfer learning
transportation
vehicles
video analysis

Cluster 8

anomalous trajectory detection
behavior

data mining
global positioning system
movement
outlier detection
patterns
sliding window
space
trajectory
trajectory analysis
trajectory data
visual analytics

Cluster 9

accuracy
anomalies
diagnosis
energy
fusion
performance
road anomaly detection
road safety
validation
vibrations

Grupos Scopus

Cluster 1

algorithm
algorithms
anomaly detection
anomlay detection methods
anomaly detection systems
anomaly-detection algorithms
article
artificial intelligence
auto encoders
automation
behavioral research
big data
clustering algorithms
data handling
data mining
data process
data set
decision making
detection method
driving behavior
errors
forecasting
geographic information systems
hidden markov model
highway accidents
highway administration
highway traffic control
human
humans
information management
information systems
information use
intelligent systems
intelligent transport systems
intelligent transportation systems
intelligent vehicle highway systems
intrusion detection
motor transportation
network security
optimization
pedestrian safety
real time systems
real-world
regression analysis
risk assessment
road network

road safety
road traffic
robotics
safety engineering
sensor
sensor networks
signal detection
signal encoding
smart city
statistical tests
statistics
street traffic control
taxicabs
time series analysis
traffic anomalies
traffic anomaly detections
traffic conditions
traffic congestion
traffic control
traffic flow
traffic management
traffic surveys
trajectories
transportation
transportation system
travel time
urban area
urban transportation
vehicle to vehicle communications
vehicle trajectories
vehicles
vehicular ad hoc networks
visualization

Cluster 2

aircraft detection
antennas
asphalt pavements
automatic detection
cameras
complex networks
condition
costs
crack detection
cracks
damage detection
data acquisition
data collection
defects

detection accuracy
detection methods
detection performance
deterioration
edge detection
extraction
feature extraction
features extraction
geological surveys
graphic methods
ground penetrating radar
ground penetrating radar systems
highway planning
image analysis
image enhancement
image processing
image reconstruction
image segmation
inspection
k-means clustering
landforms
mapping
neural networks
nondestructive examination
obstacle detectors
optical radar
pavement
pavement condition
pavement distress
pavement distress detection
pavement surface
pavements
pixels
point-clouds
remote sensing
road
road pavents
semantic segmentation
semantics
state of the art
stereo image processing
stereo vision
surface defects
surveys
textures
three dimensional computer graphics
unmanned aerial vehicles (uav)

Cluster 3

acceleration data
accelerometer
accelerometers
accidents
condition monitoring
cost effectiveness
crowdsourcing
decision trees
detection system
developing countries
digital storage
embedded systems
global positioning system
gps
internet of things
iot
maintenance
mobile devices
pothole
potholes
rasberry pi
repair
road anomaly
road anomaly detection
road condition
road condition monitoring
road infrastructures
road maintenance
road surface condition
road surfaces
roads
roads and streets
sensors
signal processing
smart phones
smartphone
smartphones
support vector machine
support vector machines
ubiquitous computing
ultrasonic applications
ultrasonic sensor
ultrasonic sensors

Cluster 4

classification
classification (of information)
cnn
computer vision

convolution
convolution neural network
convolutional neural network
convolutional neural networks
crime
deep learning
deep neural networks
detection models
image classification
images processing
large dataset
learning algorithms
learning approach
learning systems
learning technincs
long short-term memory
machine learning
machine learning techniques
machine-learning
monitoring
object detection
object recognition
performance
pothole detection
performance
pothole detection
real time systems
real-time
road cracks
road damage
security systems
survaillance video
transfer learning
video surveillance
yolo

Cluster 5

accident prevention
advanced driver assistance systems
automobile drivers
automobile suspension
autonomous driving
autonomous vehicles
bump detection
detection algorithm
fault detection
hazards
pattern recognition
road and street markings

road quality
road vehicles
speed
speed bumps
structural health monitoring
suspensions (components)
traffic signs
vibrations (mechanical)
wavel trasforms

APÊNDICE C – CÓDIGO - SIMULAÇÃO DA RESPOSTA ESTACIONÁRIA

```

%% Modelo Quarto de carro
%%
clear ; close all ; clc
%% Parâmetros
% Veículo
M = 330; % Sprung mass [kg]
m = 62; % Unsprung mass [kg]
k = 20e3; % Spring constant suspension [N/m]
kt = 200e3; % Spring constant tire [N/m]
c = 2.00e3; % Damping constant suspension [N.s/m]

%% Função de transferencia massa não amortecida
alphaunsp = M*m;
betaunsp = c*(M+m);
gamaunsp = (M*(k+kt))+ (k*m) ;
thetaunsp = c*kt;
episolonunsp = k*kt;
numunsp = [M, c, k]*kt;
denunsp = [alphaunsp, betaunsp, gamaunsp, thetaunsp, episolonunsp];
TransferFunctionUnsprung = tf(numunsp,denunsp);

%% Parâmetros
lb = 0.6; % [m]
hb = 0.1; % [m]
v = 13.89; % [m/s]
tp = lb/v;
w = (2*pi)/tp;
ds = 1/3000;
t = 0:ds:10*tp;
zr = (hb/2)*(1-cos(w*t));
zr(round(tp/ds)+1:end) = 0;
omeganw = sqrt((k+kt)/m);
ar = (((hb*(w^2))/2)/(omeganw *hb))*cos(w*t);
ar(round(tp/ds)+1:end) = 0;

```

```
[resp , tresp] = lsim(TransferFunctionUnsprung , ar , t);
```

```
%% Diagrama de bode
```

```
figure(2)
```

```
opts = bodeoptions;
```

```
opts.FreqUnits = 'Hz';
```

```
opts.PhaseVisible = 'off';
```

```
opts.MagUnits = 'db';
```

```
bode(TransferFunctionUnsprung , opts);
```

```
grid;
```

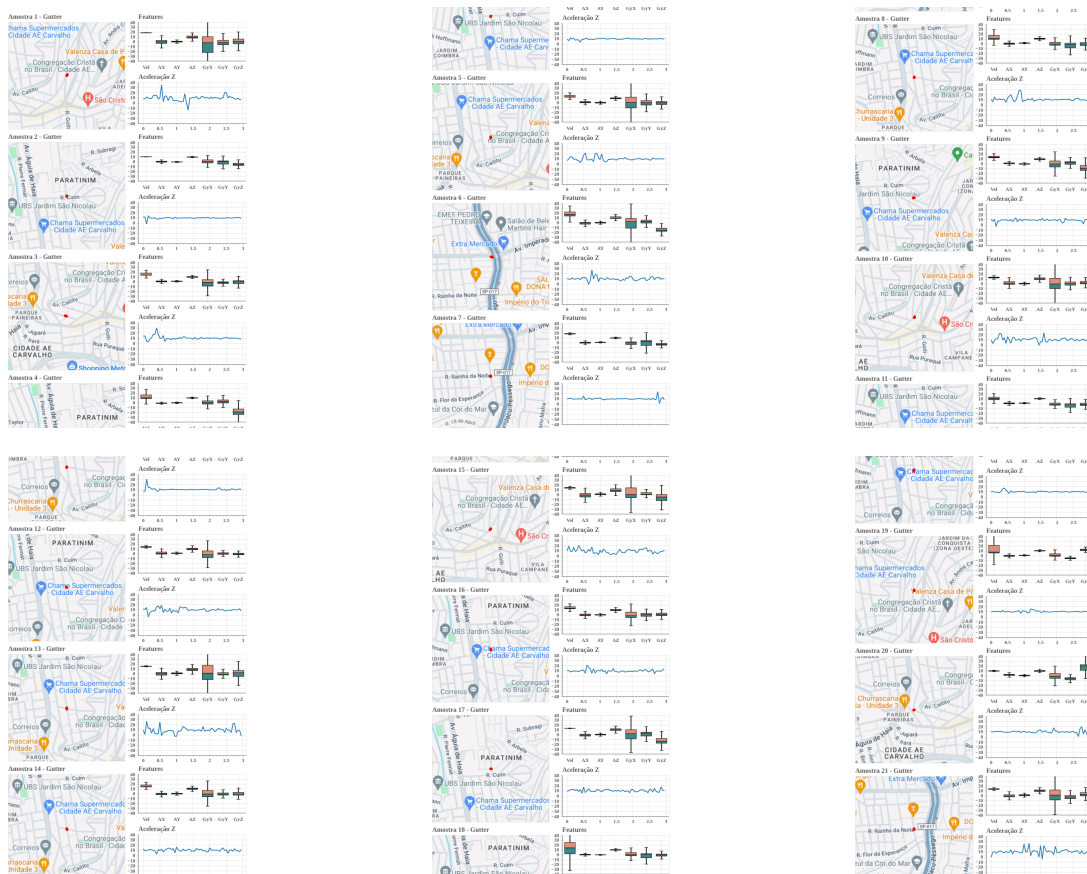
APÊNDICE D – CÓDIGO - SIMULAÇÃO DA RESPOSTA TRANSITÓRIA

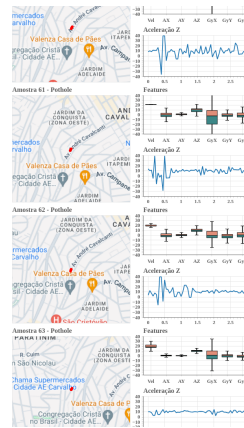
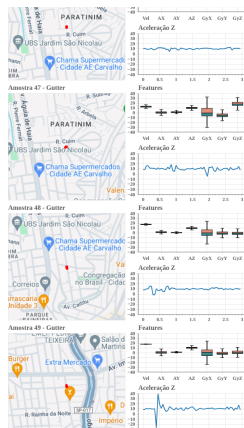
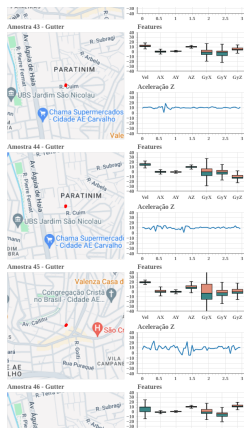
```

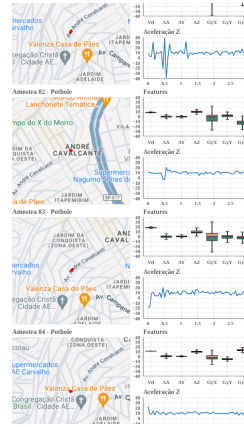
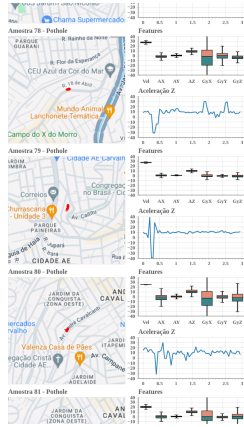
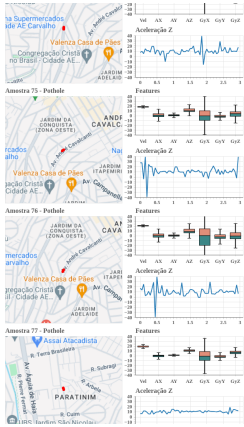
%% FFT
y = fft(resp);
z = fftshift(2*y/length(t));
Fs = 1/ds;
ly = length(y);
f = (-ly/2:ly/2-1)/ly*Fs;
figure(1);
stem(f,abs(z))
str = "Espectro de Frequência – Resposta Transitoria ";
title(append(str,num2str(round(3.6*v)), 'km/h'))
xlabel("Frequencia – Hz")
xlim([0,100])
ylabel("\omega_{n}^2/_{h_b}")
r = rectangle('Position',[0 0 10 8], 'EdgeColor','b', 'LineWidth',3);
txt = '\leftarrow faixa de Amostragem';
text(15,7,txt, 'FontSize',10)
grid on;

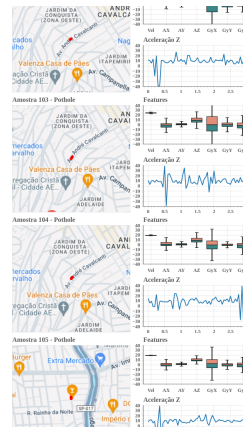
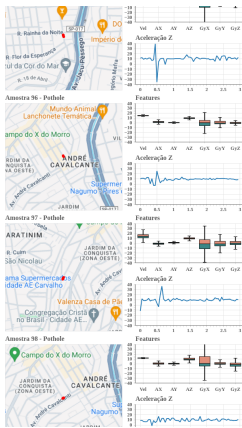
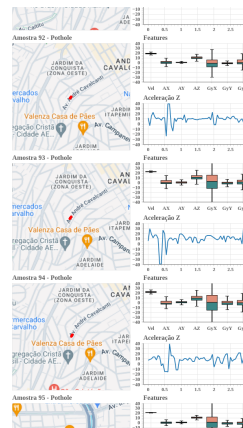
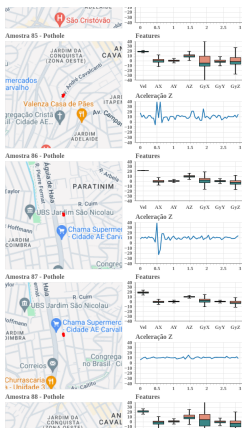
```

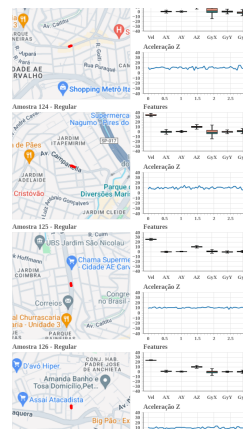
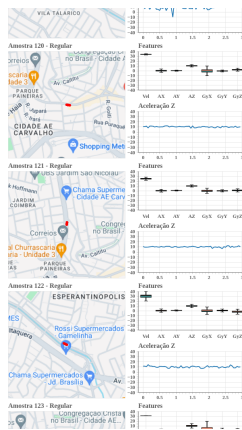
APÊNDICE E – ANÁLISE - DATASET DE TESTE

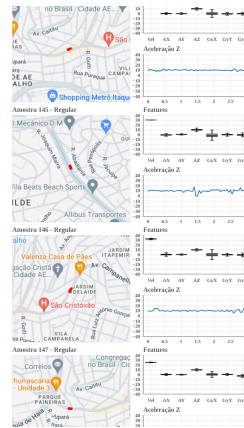
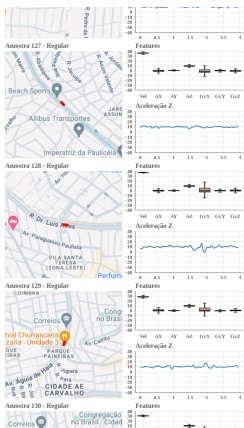


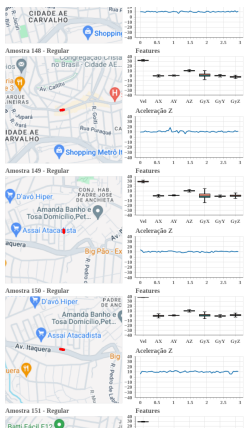


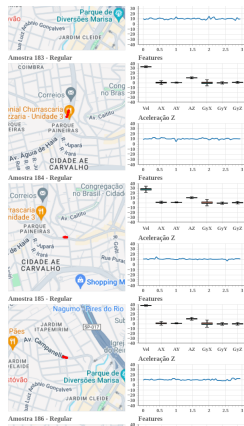
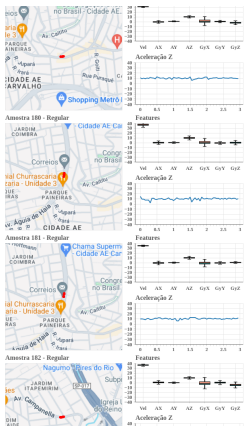
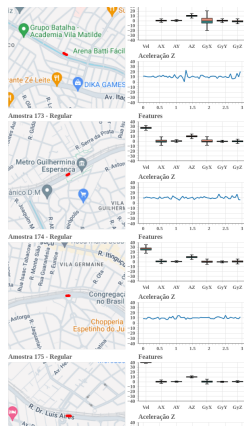
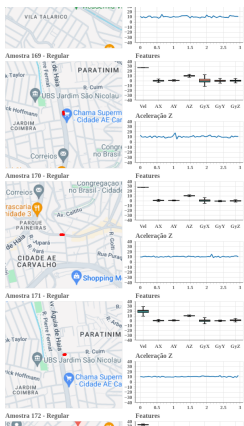


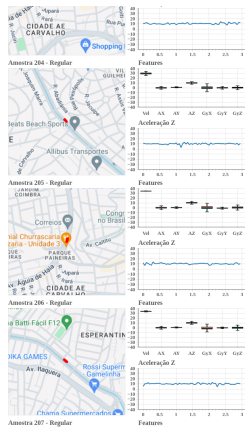


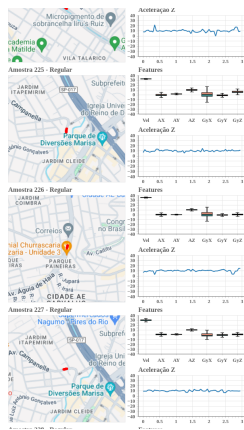
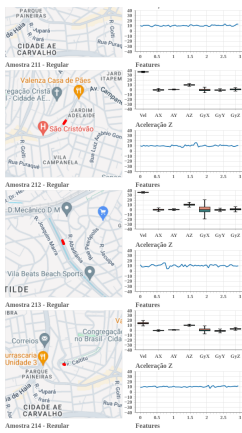


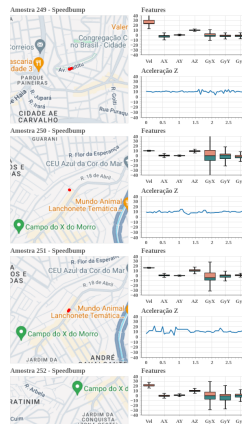
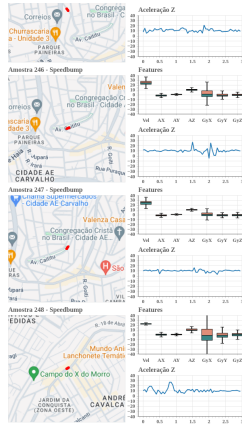
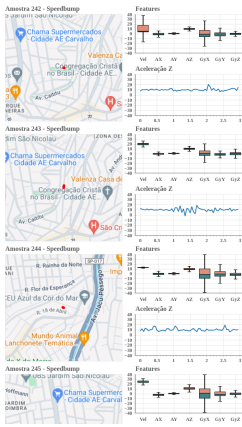
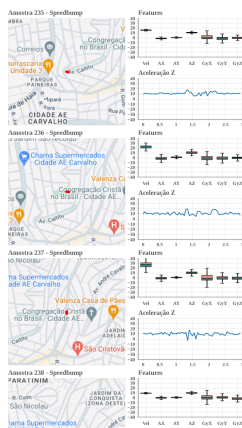
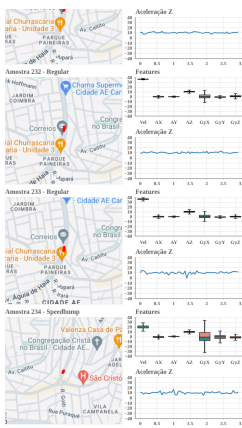


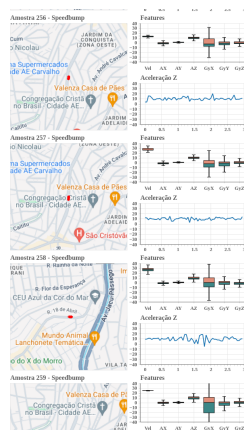
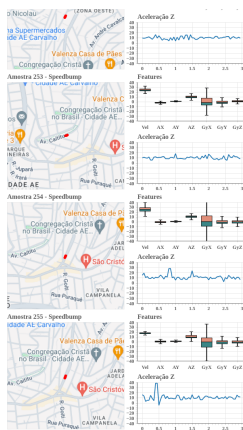


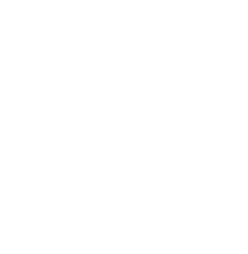
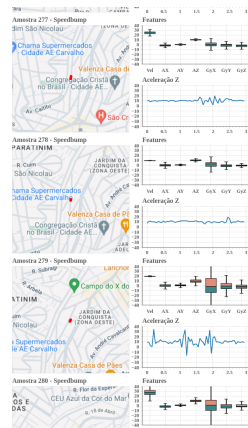
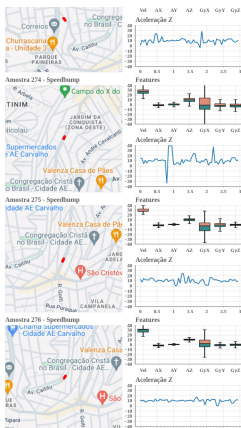












Anexos

ANEXO A – ESPECIFICAÇÕES DO VEÍCULO UTILIZADO

Especificações

■ Especificações do Veículo

Modelo	CITY	
Peso em Ordem de Marcha	TM	1.134 - 1.157 kg
	TA	1.172 - 1.192 kg
Peso Máximo Permitido	TM	1.509 kg ^{*1} - 1.532 kg ^{*2}
	TA	1.546 kg ^{*1} - 1.569 kg ^{*2}

■ Especificações do Motor

Tipo	SOHC i-VTEC FLEX	
Diâmetro Interno x Curso	73,0 x 89,4 mm	
	Cilindrada 1.496 cc	
Velas de Ignição	NGK	IZFR6K13DN
	DENSO	SKJ20DR-M13
Potência	Etanol	116 cv x 6.000 rpm
	Gasolina	115 cv x 6.000 rpm
Torque	Etanol	14,8 kgf.m x 4.800 rpm
	Gasolina	14,8 kgf.m x 4.800 rpm
Taxa de Compressão	10,4 : 1	
Rotação de	T/M	700 ± 50 rpm
Marcha Lenta	T/A	700 ± 50 rpm
Avanço da Ignição	4° ± 2° APMS	
Folga das Válvulas	Admissão	0,15 - 0,19
	Escape	0,26 - 0,30

■ Combustível

Tipo	Tanque Principal	Etanol e/ou gasolina ¹ comum, comercialmente disponível ou aditivado em qualquer proporção.
	Reservatório de Gasolina	Gasolina ¹ comum, comercialmente disponível ou aditivada.
Capacidade do Tanque de Combustível	Tanque Principal	47 ℓ
	Reservatório de Gasolina	0,7 ℓ

*1 A gasolina aditivada distingue-se da comum tipo C por receber aditivos com propriedades detergentes e dispersantes. Dessa forma, evita-se a formação de corrosão e depósitos no sistema de alimentação do veículo, permitindo sempre um bom fluxo de combustível ao longo do tempo. A decisão pelo uso de gasolina aditivada cabe ao cliente, porém ressalta-se primordialmente o cuidado com sua procedência, seja comum ou aditivada.

Ao trafegar por outros países, deve-se sempre utilizar gasolina premium com octanagem 95 ou superior sem chumbo, ou seja, verifique a seguinte inscrição na bomba do posto de abastecimento: "sin plomo", em países de língua espanhola.

■ Bateria

Capacidade/Tipo	12 V - 27 Ah (5h) / 30 Ah (20h)
-----------------	---------------------------------

■ Fluido do Lavador

Capacidade do Reservatório	2,5 ℓ
----------------------------	-------

■ Lâmpadas

Faróis (Farol Baixo)	55 W
Faróis (Farol Alto)	60 W
Faróis de Nebolina*	55 W
Lanternas Dianteiras	5 W
Luzes dos Sinalizadores de Direção Dianteiros	21 W (Âmbar)
Sinalizadores de Direção Laterais (Nos Retrovisores Externos)*	LED
Luzes das Lanternas	5 W
Luzes de Freio	21 W
Luzes dos Sinalizadores de Direção Traseiros	21 W (Âmbar)
Luzes de Ré	21 W
Luzes da Placa de Licença Traseira	5 W
Brake Light	21 W
Luzes Internas	
Luzes de Leitura	8 W
Luz de Cortesia	8 W
Luz do Porta-malas	5 W
Luz de Posição	5 W

^{*1}: Modelos sem Controle de Velocidade de Cruzeiro

^{*2}: Modelos com Controle de Velocidade de Cruzeiro

250 * Não disponível em todos os modelos

■ Fluido de Freio/Embreagem

Especificado	Fluido de Freio DOT 3 ou DOT 4
--------------	--------------------------------

■ Fluido da Transmissão Automática

Especificado	Honda ATF 11 (fluido para transmissão automática)
Capacidade	Troca 2,5 ℓ

■ Fluido da Transmissão Manual

Especificado	Fluido para Transmissão Manual Honda
Capacidade	Troca 1,5 ℓ

■ Óleo do Motor

Recomendado	Óleo Genuino Honda SAE 0W-20 API SL ou superior
Capacidade	Troca 3,4 ℓ
	Troca incluindo o filtro 3,6 ℓ

■ Líquido de Arrefecimento do Motor

Especificado	Líquido de Arrefecimento/ Anticongelante Honda para Todas as Estações, Tipo 2
Proporção	50/50 com água destilada
Capacidade	TM 4,78 ℓ (troca incluindo o 0,44 ℓ no reservatório)
	TA 4,71 ℓ (troca incluindo de 0,44 ℓ no reservatório)

■ Pneu

Normal	Tamanho ¹	175/65R15 84T 185/55R16 83V
	Pressão	Dianteiro 220 kPa, 2,2 kgf/cm ² , 32 psi
		Traseiro 210 kPa, 2,1 kgf/cm ² , 30 psi
Estepe	Tamanho	175/65R15 84T
	Pressão	240 kPa, 2,4 kgf/cm ² , 35 psi
Tamanho da Roda	Liga leve* ou aço*	15 x 5,5 J
	Liga leve*	16 x 6 J*
	Estepe	15 x 5,5 J ou
		175/65 R15 84T

■ Suspensão

Tipo	Dianteira	McPherson
	Traseira	Barra de torção

■ Alinhamento

Convergência	Dianteira	0,0 mm
	Traseira	1,0 mm
Câmbor	Dianteira	0°9'±2 0°10'±1
	Traseira	- 1°
Cáster	Dianteira	3°9'

¹ O tamanho original do pneu encontra-se na etiqueta de informações do pneu, localizada no batente da porta do motorista.

■ Direção

Tipo	Pinhão e cremalheira, com assistência elétrica
------	--

■ Freio

Tipo	Com Assistência
Dianteira	Disco Ventilado
Traseira	Disco rígido ou tambor
Estacionamento	Mecânico

■ Dimensões

Comprimento	4.400 mm	
Largura	1.695 mm	
Altura	1.480 mm	
Distância Entre Eixos	2.551 mm	
Distância entre Rodas ¹	Dianteiras	1.488 mm
	Traseiras	1.475 mm
Distância entre Rodas ²	Dianteiras	1.472 mm
	Traseiras	1.459 mm

¹ Modelos com pneus 175/65R15

² Modelos com pneus 185/55R16

* Não disponível em todos os modelos 251

Índice capítulo

Índice Geral