

DOI: 10.5748/20CONTECSI/PSE/CID/7295

eLocator: e207295

**IDENTIFICAÇÃO DE BURACOS EM VIAS PAVIMENTADAS UTILIZANDO
ACELERÔMETRO: UMA REVISÃO SISTEMÁTICA**

Ícaro Gonçalves – <https://orcid.org/0000-0002-2735-2642>

Ipt - Instituto De Pesquisas Tecnológicas

Vagner Luiz Gava – <https://orcid.org/0000-0001-5965-957X>

Douglas Bellomo Cavalcante – <https://orcid.org/0000-0002-0209-1562>

IDENTIFICATION OF HOLES IN PAVED ROADS USING ACCELEROMETER: A SYSTEMATIC REVIEW

ABSTRACT

The excessive presence of potholes on the paved roads in the state of São Paulo results in a series of problems, including traffic accidents, financial impacts, and even environmental damage. There are various solutions addressing this issue, such as visual processing, distance sensors (such as laser or ultrasonic), or response analysis using vibration. This article seeks to answer the following research question: "how to develop a low-cost embedded system based on machine learning using accelerometer and GPS capable of identifying irregularities (potholes) in road pavements?". To achieve this, a systematic literature review is conducted, with the main result being a process based on the use of accelerometers to collect vibration and the application of artificial intelligence to identify and classify potholes on paved roads, indicating the feasibility of the process in terms of installation practicality, cost, and comparative results found in the literature.

Keywords: Accelerometer, Pothole, Irregularity, SVM (Support Vector Machine), Edge Computing, ThingsBoard, Internet of Things, Automatic Pothole Detection.

IDENTIFICAÇÃO DE BURACOS EM VIAS PAVIMENTADAS UTILIZANDO ACELERÔMETRO: UMA REVISÃO SISTEMÁTICA

RESUMO

A presença excessiva de buracos nas vias pavimentadas do estado de São Paulo resulta em uma série de problemas, incluindo acidentes de trânsito, impactos financeiros e até danos ambientais. Existem diversas soluções abordando essa questão, como por exemplo processamento visual, sensores de distância (como laser ou ultrassônico) ou por análise de resposta, utilizando a vibração. Este artigo procura responder à seguinte pergunta de pesquisa: “como desenvolver um sistema embarcado, de baixo custo, baseado em *machinelearning* utilizando acelerômetro e GPS capaz de identificar irregularidades (buracos) em pavimentos viários?”. Para tal, é desenvolvida uma revisão sistemática da literatura que tem como principal resultado um processo baseado na utilização de acelerômetros para coletar a vibração, e da aplicação de inteligência artificial para identificar e classificar buracos nas vias pavimentadas, indicando a viabilidade do processo em função da praticidade da instalação, do custo, e dos comparativos dos resultados levantados na literatura.

Palavras-chave: Acelerômetro, Buraco, Irregularidade, SVM, Computação de Borda, Thingsboard, Internet das Coisas, Detecção Automática de Buracos.

1 INTRODUÇÃO

Nesta seção são descritos o problema e sua contextualização, os objetivos a serem alcançados, a questão de pesquisa, a metodologia adotada, as contribuições à comunidade científica e da indústria de software e a organização das demais seções deste trabalho.

1.1 Apresentação e Contextualização do Problema

A qualidade das vias pavimentadas implica diretamente em questões financeiras. Cerca de 95% dos passageiros e mais de 65% dos veículos de carga dependem unicamente do transporte rodoviário, e por conta disso, buracos e ondulações presentes no pavimento geram custos operacionais de transporte, elevando os níveis de acidente para os passageiros e veículos de cargas, acarretando também em prejuízos ambientais (CNT, 2021). Um exemplo disso é uma avenida que é uma das principais da cidade de Osasco, situada no estado de São Paulo, a Avenida dos Autonomistas, Figura 1, demonstrando a situação precária, em termos de buracos.



Figura 1 - Buraco presente na Avenida dos Autonomistas

Fonte: Elaborado pelo autor

Conforme pesquisa realizada pela Confederação Nacional do Transporte (CNT, 2021), o transporte rodoviário tem um papel fundamental para a sociedade brasileira, sendo que é o principal método de locomoção de cargas no país. Cerca de 52,2% (56.970 quilômetros) da extensão total das rodovias avaliadas apontam algum defeito no pavimento. Dos custos operacionais do transporte, a porcentagem dos valores relacionados a qualidade do asfalto está estimada cerca de 30,9% devido a esses diversos defeitos nas vias pavimentadas, além de causar a diminuição da velocidade do veículo, aumentando o tempo de viagem das cargas e diminuindo a vida útil dos veículos, principalmente em falhas ligadas ao amortecedor, acarretando posteriormente em outras falhas graves no sistema mecânico.

Um buraco em vias pavimentadas é uma abertura ou depressão que ocorre na superfície da estrada asfaltada. Esses buracos são formados devido a uma combinação de fatores, como o desgaste natural do pavimento, ação do clima, tráfego intenso e falta de manutenção adequada (Costa, 2020).

Segundo CNT (2021) somente em custos excedentes anuais de combustível, causado pela má conservação dos pavimentos, ocasionam um valor estimado em 955,99 milhões de litros de diesel, o que gera valores adicionais para o transportador de R\$ 4,21 bilhões.

Essas falhas nas vias pavimentadas são ocasionadas ou pela falta da execução do projeto adequado na elaboração do asfalto, ou pela falta de manutenção preventiva. Existem diversas técnicas para aplicar a pavimentação asfáltica, sendo uma das mais utilizadas nas diversas cidades de São Paulo pavimento flexível. Essa técnica é chamada de

"flexível" porque a estrutura do pavimento é projetada para flexionar sob o peso dos veículos que passam sobre ele (DNT, 2022).

Para realizar a manutenção em buracos nas ruas da cidade de São Paulo, é necessária a indicação da população para a prefeitura de São Paulo, por meio de canais eletrônicos (Aplicativo SP156 e Portal de Atendimento SP156), presencialmente nos atendimentos das Subprefeituras ou pela ligação telefônica na central SP156, informando a localização dos buracos. (Prefeitura de São Paulo, 2023).

Comumente conhecido como operações "Tapa Buraco", Figura 2, é passível de falhas por depender da ação humana. Como o processo é somente em identificar os buracos, não sendo possível qualificar os mesmos, perde o conceito de prioridade, acarretando possíveis transtornos, recapando vias que poderiam aguardar mais tempo, em comparação a outra situação de emergência (Prefeitura de São Paulo, 2023).



Figura 2 - Operações "tapa buraco"
Fonte: Adaptado de Prefeitura de São Paulo (2023)

Existem diversos métodos para automatizar o processo de localização e identificação das falhas nas vias pavimentadas, sendo grande parte utilizando processamento visual, sensores ultrassônicos e de vibração.

Em 2022, a prefeitura de São Paulo investiu 1 bilhão de reais no programa de recapeamento, chamado de Recape. Esta iniciativa trouxe um investimento na área de automatização na identificação de falhas nas vias pavimentadas chamado de Sistema Gaia. Consiste em 108 veículos, entre táxis e carros de aplicativos, contendo um sensor instalado na suspensão desses veículos, cuja finalidade é na identificação das mudanças que ocorreram ao longo do tempo que podem ter sido causadas pelo desgaste do asfalto, ou por obras realizadas por concessionárias, e em algum deles possuem também uma câmera no retrovisor identificando o local percorrido pelo veículo (Prefeitura de São Paulo, 2023).

Segundo Kandoi (2021) os buracos são uma das principais causas de acidentes rodoviários. O autor desenvolveu uma solução capaz de identificar buracos e lombadas a partir do uso de acelerômetros embarcados do smartphone, realizando um comparativo com uma outra solução baseada em processamento visual. Carlos (2017) realizou um comparativo de técnicas utilizando *machinelearning* e algoritmos probabilísticos com a finalidade de identificar a técnica com maior eficiência no ambiente de identificação de buracos e lombadas em vias as pavimentadas.

Cavalcante (2019) propôs uma solução baseada na coleta de vibração de um acelerômetro instalado na bandeja de suspensão do veículo. Quando as rodas do veículo passam por algum buraco, o sensor capta essa vibração, e aplicando o algoritmo de desvio padrão no eixo Z consegue identificar um buraco. Via bluetooth, o dispositivo transmite as informações para um celular, que por sua vez transmite para uma plataforma de IoT, em

conjunto com a localização do buraco, podendo geolocalizar essas informações em tempo real.

O artigo está dividido em quatro seções. A segunda traz a fundamentação teórica com uma visão geral sobre dívida técnica e métodos ágeis. A terceira apresenta o método de revisão sistemática da literatura adotado neste trabalho. A quarta seção fornece a análise e os resultados obtidos na revisão sistemática e a quinta seção traz as conclusões, limitações e sugestões para futuros trabalhos de pesquisa e aplicação prática dos conceitos nas organizações.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão tratadas as principais referências já publicadas que formam a base de conhecimento para estruturação desta pesquisa, compostas pela fundamentação teórica e a revisão sistemática da literatura.

2.1 Conceitos Gerais

Nesta subseção é apresentado uma visão geral dos termos utilizados para essa pesquisa, iniciando pela definição do microcontrolador escolhido no projeto, o ESP32 HELTEC v2, do sensor acelerômetro, modulo wireless GPS, os protocolos de comunicação utilizados no hardware, I²C, o SPI e o UART assim como os demais conceitos necessários para o entendimento do trabalho.

2.1.1 Microcontrolador Esp32 Heltec

Utilizou-se nesse trabalho o microcontrolador ESP32 Heltec V2, por já possuir uma plataforma vasta de funcionalidades embutidas em seu hardware, sendo essenciais para o andamento do estudo, facilitando o desenvolvimento do hardware final. Utilizando o chip ESP32-D0WDQ6, possuindo 2 núcleos de processamento, com acesso a 32 bits de barramento, com um clock máximo de 240 MHz, memória não volátil (Flash) de 8MB e 520kB de memória volátil (SRAM) (Espressif Systems, 2023).

O Chip ESP32-D0WDQ6 possui pinos de Universal Asynchronous Receiver / Transmitter (UART) x 3, Serial Peripheral Interface (SPI) x 2, Inter Integrated Circuit (I²C) x 2, Inter-ICSound (I²S) x 1, 12 - bits ADC (Analogue Digital Converter) de entrada x 18, 8 bits de saída para um conversor de analógico / digital x 2, GPIO x 22 e GPI x 6 (Espressif Systems, 2023).

Essa plataforma possui 3 módulos de wireless, sendo: o de WiFi 802.11 b/g/n (802.11n até 150Mbps), duas variantes de bluetooth, o V4.2 e o LE (Low Energy) e o de LoRa, funcionando na frequência aceita pela Anatel no Brasil de 915MHz ~ 928MHz, usando o chip SX127x (Heltec Automation, 2023).

Possui o chip CP2102, convertendo o protocolo de comunicação UART, fornecido pelo chip ESP32-D0WDQ6 para o USB, facilitando a compilação dos códigos. Tem um conversor de bateria de Lithium de 3.7V, identificando quando se tem energia na entrada do microcontrolador, carregando a bateria, ou se não tiver, chaveia automaticamente para usar a energia contida na bateria. Possui um display OLED de 0.96 polegadas. Na Figura 3 é mostrado as funções dos pinos do ESP32 Heltec V2, com suas respectivas entradas e saídas (Heltec Automation, 2023).

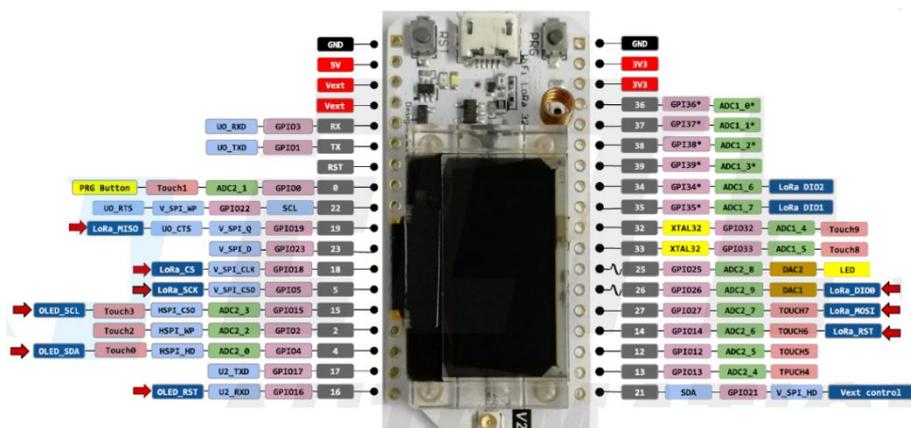


Figura 3 - Pinout do Microcontrolador ESP32 Heltec v2

Fonte: Adaptado de Heltec Automation (2023)

2.1.2 Acelerômetro

Os acelerômetros Micro-ElectroMechanical Systems (MEMS) têm sido amplamente adotados em muitas áreas de aplicações eletrônicas, sendo eles em sistemas automotivos, industriais, propiciando a identificação e monitoramento de vibrações em sistemas mecânicos, tais como dobradiças de elementos rotativos, caixas de engrenagem, componentes estruturais em geral e em muitas outras áreas, devido ao seu pequeno tamanho, baixo consumo de energia e integração de baixo custo.

O acelerômetro é um sensor capaz de medir a aceleração própria de um sistema. A ideia de aceleração originou-se nos séculos XVII e XVIII com os trabalhos científicos de Galileu, estabelecendo a relação entre força e aceleração, e com Newton, que desenvolveu suas leis do movimento, incluindo a segunda lei de Newton, definindo a relação entre força e aceleração. A aceleração unitária de um sistema é mensurada em relação a outro sistema em queda livre, de modo que está ligada à sua percepção de peso. Portanto, um acelerômetro em queda livre na Terra resultará em uma aceleração nula, entretanto se o acelerômetro estiver em repouso sobre a superfície da Terra, indicará uma aceleração g ($9,8 \text{ m/s}^2$) para cima, pois, em relação a um objeto em queda livre, o acelerômetro está acelerado a $9,8 \text{ m/s}^2$ para cima. (Hewitt, 2011).

Na Figura 4 é mostrado o funcionamento de um acelerômetro do tipo MEMS (Micro-Electro-Mechanical System), tendo placas fixas captando a capacitância emitida por uma peça central, que é movida de acordo com a aceleração externa.

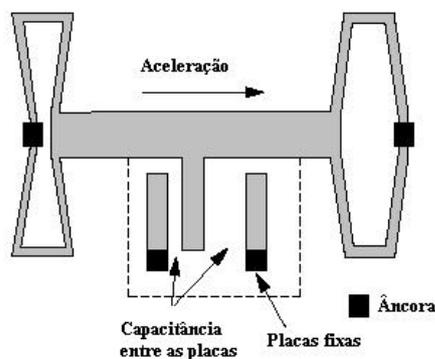


Figura 4 - Sistema de massa e mola utilizado em muitos acelerômetros

Fonte: Adaptado de Analog Devices (1998)

Para esse trabalho foi escolhido o chip MPU-9250, sendo um acelerômetro de três eixos, giroscópio de três eixos e magnetômetro de três eixos, usando a tecnologia MEMS.

Possuindo um ADC interno de 16-bits de resolução, convertendo para sinal analógico, facilitando a implementação deste sensor. O fundo de escala do acelerômetro pode ser ajustado para $\pm 2g$, $\pm 4g$, $\pm 8g$ ou $\pm 16g$. Possui pinos de comunicação I²C, dois possíveis endereços, o x68 e o x69, de acordo com o nível lógico aplicado no pino ADO. Na Figura 5 é mostrado o módulo que utiliza o chip MPU-9250, utilizado nesse trabalho (InvenSense, 2014).



Figura 5 - Módulo Acelerômetro e Giroscópio 9 Eixos MPU-9250 GY-9250
Fonte: Adaptado de Saravati (2023)

2.1.3 GPS

O GPS (Global Positioning System) é um sistema que utiliza mais de 30 satélites de navegação que circulam próximos a órbita do planeta Terra. Devido às constantes transmissões de dados desses satélites, são calculadas suas localizações aéreas, determinado com precisão sua localização. Um receptor GPS no telefone celular, por exemplo, recebe esses sinais. Uma vez que o receptor calcula sua distância de quatro ou mais satélites GPS, pode identificar a localização, convertidos em latitude e longitude (NASA, 2019).

GPS é um sistema composto por três partes: satélites, estações terrestres e receptores. Os satélites agem como as estrelas nas constelações, sabe-se onde devem estar a qualquer momento, a partir de cálculos para identificar suas trajetórias. As estações terrestres usam radar para se certificar de que estão realmente onde pensamos que estão. Um receptor está constantemente recebendo um sinal desses satélites, descobre a que distância está de alguns deles. Uma vez que o receptor calcula sua distância de quatro ou mais satélites, é possível obter a localização exata desse receptor, com um erro de poucos centímetros (NASA, 2019).

Na Figura 6 é demonstrado o funcionamento do GPS, ilustrando o satélite, o receptor e a estação terrestre.

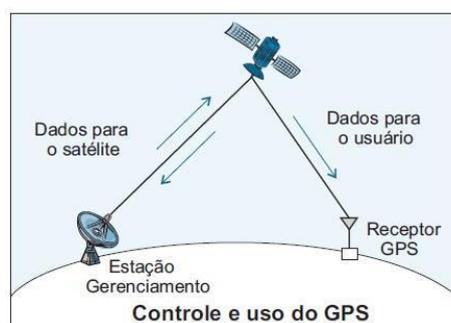


Figura 6 - Sistema de Posicionamento Global
Fonte: Adaptado de Alves (2023)

Outra característica do GPS é a possibilidade de atualizar o Real Time Clock (RTC) local, graças ao relógio atômico de Rubídio e Césio a bordo nos satélites, possibilitando

uma precisão na ordem de nano a pico segundos. Isso é vantajoso pelo fato de que muitos dispositivos utilizam como referência de *clock* os osciladores de cristais, como por exemplo o de quartzo, que tem como referência valores na ordem de microssegundos (BIDIKAR, 2014)

Utiliza-se nesse trabalho o chip de GPS GY-NEO7MV2-NEO 7M-0-000, que possui uma antena de cerâmica. Além do convencional GPS, esse módulo tem um suporte adicional para o GNSS, podendo funcionar com satélites adicionais ao sistema do GPS, obtendo valores mais precisos. Apresenta uma baixa corrente no modo convencional de operação de 17 mA, e 5 mA no modo de economia. Possui saída digital com o protocolo UART, podendo facilmente ser integrado a um microcontrolador. Na Figura 7 é mostrado o módulo que utiliza o chip comentado, além de possuir um oscilador de Crystal de quartzo com uma bateria local, para armazenar o RTC.

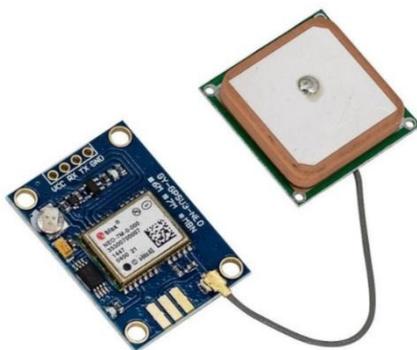


Figura 7 - Módulo GPS GY-NEO7MV2-NEO 7M-0-000
Fonte: Adaptado de Saravati (2023)

2.1.4 LoRa

Com a crescente aplicação de IoT em ambientes diversos, como em cidades e agricultura inteligentes ou monitoramento de mercadorias, entre muitos outros, exige-se fortes requisitos para a Rede de Acesso por Rádio (RAN) em termos de cobertura, consumo de energia do módulo final, e escalabilidade. Por causa dessas necessidades, foi desenvolvido o conceito de *Low Power - Wide Area Network* (LP-WAN) surgindo como uma alternativa para poder solucionar essas três demandas. Dentre vários protocolos desse agrupamento, uma tecnologia destaca-se pela sua versatilidade em ambientes de IoT, que é o *Long-Range Wide Area Network* (LoRaWAN) (Iborra, 2018).

Esse protocolo de comunicação consegue abranger uma área por volta de 10 km com visada, com uma largura de banda de 500 kHz dependendo de sua configuração, funcionando em uma topologia de estrela, sendo um coordenar para diversos dispositivos finais, como é mostrado na Figura 8. O LoRa apresenta três parâmetros configurados, que de acordo com essas definições, consegue alcançar resultados específicos de taxa de transmissão dos dados, consumo energético, alcance de comunicação e validade dos dados recebidos.

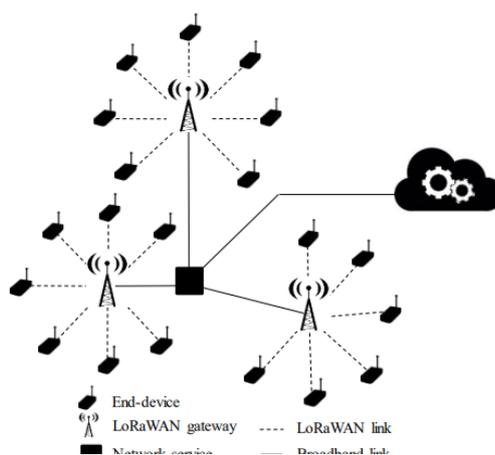


Figura 8 - Topologia em estrela do LoRa

Fonte: Adaptado de Iborra (2018)

Os parâmetros de configuração são o *Spreading Factor* (SF), o *Coding Rate* (CR) e o *Bandwidth* (BW) (Iborra, 2018):

- SF é a taxa de espalhamento, podendo variar SF de 7 a 12, de modo que valores menores resultam em uma rede com maior taxa de transmissão, maior gasto energético e o menor alcance dos dados. Já os valores maiores focam em uma rede com um maior alcance, com um menor gasto energético, e com uma baixa taxa de transmissão. Esse conceito em LoRa também é conhecido como *Data Rate* (DR), sendo uma escala invertida, que no caso o SF12 é proporcional ao DR0, e o DR5 é equivalente ao SF7.

- CR é definido pelo código de erro. Com o objetivo de melhorar a robustez do link, o LoRa emprega codificação de erro cíclica para realizar a detecção e correção de erros diretos. Essa codificação de erro ocorre em uma sobrecarga de transmissão dos dados, que é determinada pelo parâmetro CR.

- BW é a definição da largura de banda, sendo comumente aplicado o valor de 125 kHz, variando em até 500 kHz.

Nesse trabalho será utilizado o serviço da empresa IoTLabs, que disponibiliza uma rede de LoRaWAN em diversas regiões, como é mostrado na Figura 9. O IoT Labs é a PLATAFORMA onde empresas parceiras, clientes, desenvolvedores e estudantes experimentam, interagem, testam e entendem fim a fim as aplicações em uma rede real, a Rede Neutra IoT LoRaWAN da American Tower (IoTLabs, 2023).

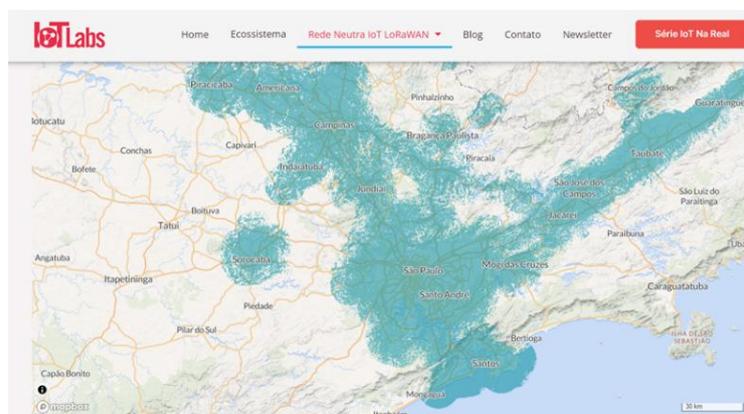


Figura 9 - Mapa de cobertura de LoRaWAN oferecido pela empresa IoTLabs

Fonte: Adaptado de IoTLabs (2023)

2.1.5 Inteligência Artificial

O termo inteligência artificial foi criado por John McCarthy no ano de 1956, com o artigo: “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence”. No artigo é descrito que a inteligência artificial consiste em simular funções do cérebro humano, porém descreve que o grande limitante é a capacidade humana em escrever programas eficientes para lidar com situações complexas (McCarthy, 1955).

Desde então, a ideia de inteligência foi se desenvolvendo, porém de uma forma diferente dos dias atuais e que consistia em um conjunto de regras que ao receber os dados, fornecia a resposta, dependendo de instruções prévias.

Com o avanço da tecnologia, veio o conceito de *Machine Learning*, que ao invés de o algoritmo fornecer as respostas, tendo como entradas os dados e as regras, nesse caso são fornecidas como entradas os dados e as respostas, chamadas de características e rotulagem, obtendo como saídas o *machinelearning*. Na Figura 10 é mostrado a diferença de inteligência artificial para aprendizado de máquina (Pacheco, 2021).

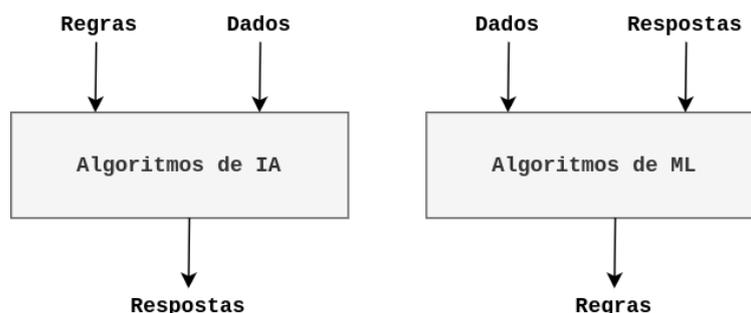


Figura 10 - Diferenças dos modelos de Inteligência Artificial e Machine Learning
Fonte: Adaptado de Pacheco (2021)

2.1.6 Machine Learning - SVM

O conceito de *machinelearning* vem com a ideia de que a própria máquina encontrará as regras eficientes se tiver as respostas pré-definidas e com a entrada de dados. Nunca encontrará o melhor modelo possível, conhecida como solução forte, o *machinelearning* encontrará uma solução viável para solucionar o problema (Leilani, 2018).

Existem diversos algoritmos que aplicam *machinelearning*, cada um com suas vantagens e desvantagens, sendo divididos em três classes, os supervisionados, os não supervisionados e os de aprendizagem por reforço. Supervisionados são os algoritmos que precisam de ensinamento prévio, como as características e rótulos, para encontrarem uma solução rápida e eficiente. São alguns exemplos o K-NN (k-nearestneighbors), o SVM (support vector machine), as árvores de decisão, as redes neurais, dentre outros (Singh, 2012).

Os algoritmos não supervisionados são os que conseguem obter as respostas sem ter as regras de forma de rótulos e características, nesse caso, analisam semelhanças das características dos dados, e conseguem agrupar em definições próprias de classes. São alguns exemplos o clusterização, o Análise de Componentes Principais (PCA), entre outros (Alloghani, 2020).

Os algoritmos de aprendizagem por reforço, na fase de treinamento, recebem uma recompensa pelas suas ações, e ao executarem a ação treinada, sempre vai pelo caminho que mais recebeu incentivo. São alguns exemplos o Q-learning e o Sarsa (Bellemare, 2017).

Para a aplicação de *machinelearning* supervisionado utilizado nessa pesquisa, é importante apresentar os seguintes termos que serão utilizados na seção 4:

Features: As *features* são as características ou variáveis de entrada usadas pelo modelo para realizar previsões ou classificações, sendo fundamentais para que o modelo possa identificar e capturar os padrões relevantes nos dados. A escolha adequada das *features* é crucial para otimizar o desempenho do modelo, permitindo uma melhor compreensão e previsão do fenômeno em questão (Kumar, 2022),

Underfitting: Ocorre quando um modelo não consegue identificar os padrões do fenômeno que deseja identificar, por ser um modelo simples ou não conseguir se ajustar aos dados de treinamento, não identificando com precisão a relação das *features* com os rótulos (IBM, 2020),

Overfitting: Ocorre quando um modelo é treinado em um único conjunto de dados, não tendo a capacidade de generalização. Dessa forma o modelo se adequa exclusivamente aos dados de treinamento, aprendendo até mesmo os ruídos e as peculiaridades desses dados, não existindo necessariamente em outros conjuntos de dados (IBM, 2020) e

Data Augmentation: É uma técnica utilizada em *machinelearning* para aumentar a quantidade de dados e a sua variedade para treinar um modelo, melhorando a capacidade de generalizar e realizar previsões precisas em dados não aplicados anteriormente (Perez, 2017).

Matriz de confusão: Uma técnica utilizada para avaliar o desempenho de um modelo de classificação, sendo contruída em quatro elementos principais: verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos, e partir dessas características, pode se concluir a acurácia, precisão, sensibilidade e F1-score (medida que combina precisão e sensibilidade em uma única métrica, fornecendo uma visão mais ampla do desempenho) (Düntsche, 2019).

Hiperparâmetros do SVM: São parâmetros que definem o comportamento do algoritmo de *machinelearning* antes de aplicar o treinamento. Precisam ser definidos previamente, e a sua escolha otimizará o desempenho do modelo, afetando a maneira como o modelo é treinado e como as previsões são realizadas. Alguns dos hiperparâmetros utilizados no SVM (Padierna, 2017):

C: Controla a penalidade por violações das margens no SVM de classificação. Um valor alto permite ao modelo ter uma classificação mais precisa dos dados de treinamento, porém não tendo um modelo tão genérico e passivo de ruídos. Com um valor baixo suaviza a classificação, possuindo uma margem mais ampla e aceitando mais erros na classificação (Yıldırım, 2020);

Kernel: Em um algoritmo de SVM, *kernel* é uma função matemática que mapeia a entrada dos dados para um conjunto de características pré-definidas, podendo ser *Kernel Linear*, *Kernel Polinomial*, *Kernel RBF (Radial BasisFunction)*, *kernel sigmoidal* e o *kernel laplaciano* (Padierna, 2017);

Gamma: Afeta a forma como o SVM lida com a influência de um único exemplo de treinamento. Valores baixos de gamma ocasionam em uma influência mais abrangente e suave de cada exemplo, enquanto valores altos de gamma dão maior peso aos exemplos mais próximos ao vetor de suporte. A escolha do valor de gamma pode influenciar a flexibilidade do modelo e sua capacidade de se ajustar a diferentes formas de distribuição de dados (Yıldırım, 2020) e

Tolerância à parada (tol): Parâmetro que determina a precisão do critério de parada no processo de otimização. O treinamento do SVM implica em um procedimento iterativo de otimização, cujo objetivo é encontrar os vetores de suporte e os coeficientes de margem ótimos. A tolerância de parada estabelece o limite máximo permitido para a diferença entre

iterações consecutivas, antes que o processo de otimização seja considerado concluído (Loosli, 2007).

Nesse trabalho será utilizado o algoritmo supervisionado SVM (3.2) como um modelo de classificação, solucionando problemas e separando as classes, utilizando um hiperplano. A técnica consiste em aplicar uma linha que possa separar os grupos que possuem padrões definidos, de forma que essa linha se situe centralizada em relação aos dois grupos. Na Figura 11 é mostrado dois grupos, estrelas e círculos. Nesse caso foi definido um vetor de suporte no limite de cada grupo, calculando a mediana, possuindo duas margens equidistantes (Cervantes, 2020).

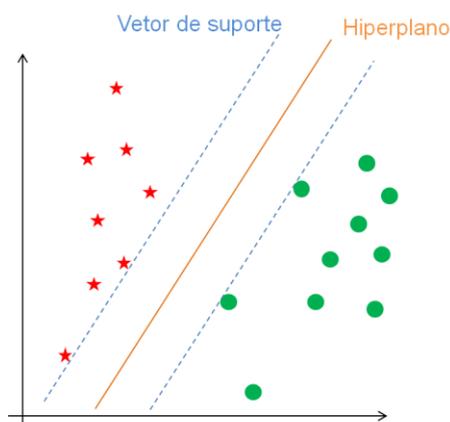


Figura 11 - Hiperplano e vetor de suporte
Fonte: Elaborado pelo autor

O SVM foi criado inicialmente para aplicar em grupos que possam ser linearmente separados, porém com o avanço da tecnologia, hoje existem ferramentas que aplicam SVM em situações mais complexas. Por exemplo, a biblioteca *pythonscikit-learn* possui várias algoritmos e modelos prontos para facilitar o treinamento, tendo potencial para calcular formas lineares, polinômios, função de base e sigmoide, se adequar em diversas situações complexas. Na Figura 12 mostra alguma das possibilidades da biblioteca *scikit-learn* para separar diversos padrões, utilizando algoritmos de *machinelearning* (scikit-learn, 2023).

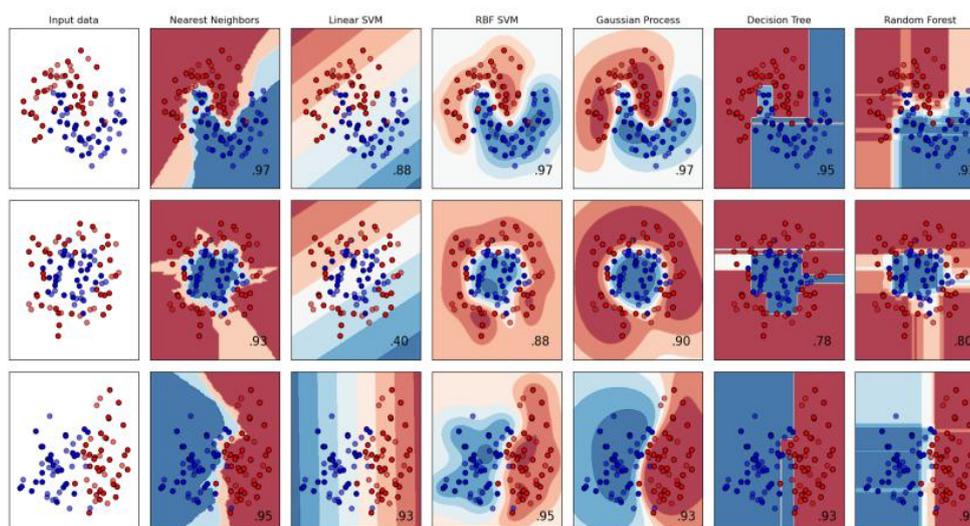


Figura 12 - Pontos de treinamento em cores sólidas e os pontos de teste semitransparentes
Fonte: Adaptado de scikit-learn (2023)

2.1.7 MQTT - Message Queuing Telemetry Transport

O MQTT é um protocolo de comunicação que usa princípios de pilha TCP/IP para transferir dados como formato de Cliente/Servidor, necessitando de um gerenciador de mensagens, chamado de *Broker*. Foi desenvolvido na década de 90 pela IBM e Eurotech com a finalidade de otimizar o protocolo de comunicação HTTP (Hypertext Transfer Protocol) para o ambiente de *MachinetoMachine* (M2M) e posteriormente aplicado também no âmbito de IoT. Tem a vantagem de ser um protocolo mais simples, com um baixo consumo de banda, sendo seguro e confiável (Santos, 2023).

Para o funcionamento do projeto, será utilizado um gateway para receber as informações pelo módulo, via LoRa. Um gateway é responsável por atuar como uma ponte entre protocolos diferentes, que no caso do projeto converterá os pacotes recebidos para um protocolo baseado em TCP/IP (Transmission Control Protocol/Internet Protocol), que nesse caso é o MQTT (IoT Labs, 2023).

ThingsBoard

É uma plataforma de IoT com a finalidade de auxiliar a análise e gerenciamento de dados, utilizando alguns dos protocolos mais comuns de rede, como o MQTT e HTTP. Possui diversas ferramentas de análise de dados em tempo real, podendo desenvolver dashboards iterativos e customizáveis de acordo com a necessidade. Utilizando o protocolo MQTT, o *thingsboard* pode atuar como servidor, executando a função de um Broker (local ou na nuvem), apresentando também, assim como a função cliente, executando comandos como *publish* ou *subscriber* (Thingsboard, 2023).

Na Figura 13 é mostrado uma de suas possibilidades de criação de mapas, atualizados em tempo real.

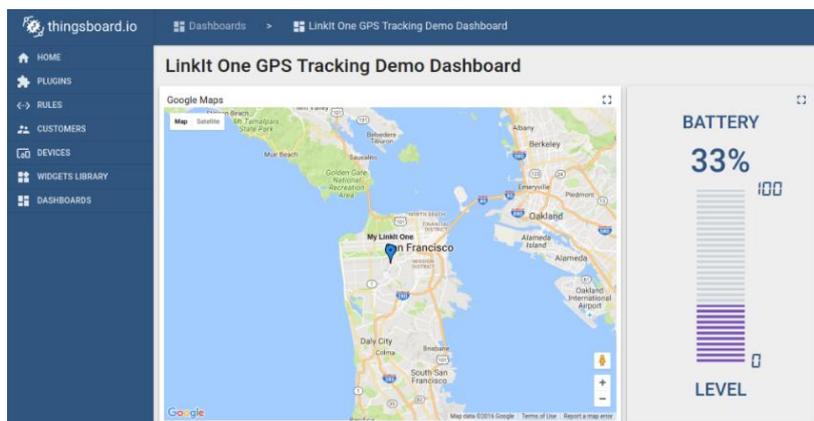


Figura 13 - Plataforma Thingsboard demonstrando a funcionalidade de geolocalização
Fonte: Adaptado de Thingsboard (2023)

2.1.8 Edge Computing e Fog Computing

São consideradas arquiteturas de processamento de dados, sendo melhorias em relação à arquitetura clássica do *Cloud Computing*, que consiste na transmissão total dos dados de um dispositivo para a nuvem. O *Fog Computing* traz a melhoria de filtrar os dados de um conjunto de dispositivos, sendo comumente aplicado no *gateway* ou em algum dispositivo intermediário. Segundo Tedeschi (2019), essa arquitetura traz o benefício da diminuição da latência e do consumo da banda de dados, permitindo o gerenciamento de diversos dispositivos e transmitindo somente o que for necessário.

A arquitetura do *Edge Computing* consiste no poder de processamento localizado próximo ao sensor. Por essa razão, o dispositivo consegue tomar decisões e, caso interprete como relevante para a situação, transmite somente o necessário. Isso traz as vantagens mencionadas no *Fog Computing* e também melhora o consumo energético do dispositivo,

uma vez que a transmissão de dados é uma ação que consome energia adicional em relação ao processamento e coleta dos dados. Essa característica é especialmente importante para dispositivos alimentados por bateria. Na Figura 14 é mostrada as regiões de aplicação destes conceitos.

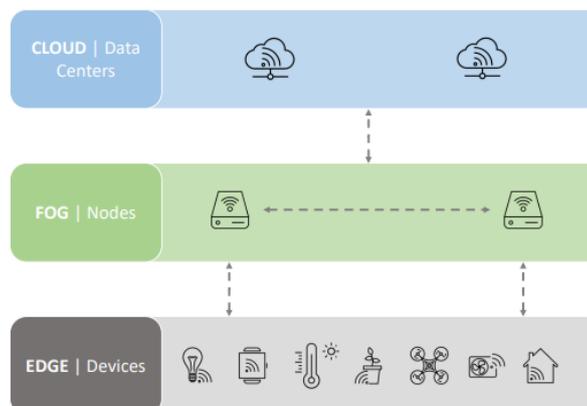


Figura 14 - Camadas de processamento - Cloud, Fog e Edge
Fonte: Adaptado de Tedeschi (2019)

Ambas as arquiteturas citadas trazem benefícios distintos para determinadas situações. No caso do *Fog Computing*, é ideal para dispositivos mais simples, em termos de poder e lógica de processamento, que sejam energizados de forma constante. A arquitetura de *Edge Computing* necessita de um hardware e uma lógica mais robusta, que seja capaz de processar os dados, porém sendo melhor adaptados para dispositivos alimentados por bateria, uma vez que é mais eficiente em termos energéticos.

3 MÉTODO

Esta etapa tem como o objetivo de identificar e analisar métodos, técnicas, algoritmos existentes na literatura na identificação e classificação de buracos em vias pavimentadas.

A revisão sistemática busca detectar estudos na qual possam contribuir na questão de pesquisa:

Objetivo: Desenvolver um dispositivo capaz de identificar os buracos em vias pavimentadas, utilizando acelerômetro instalado na bandeja de suspensão de um carro para medir as vibrações. Será utilizado *Machine Learning* para identificar os buracos, os dados serão processados no dispositivo e posteriormente transmitidos para uma plataforma de IoT.

Questão principal: Como desenvolver um sistema embarcado, de baixo custo, baseado em machine learning utilizando acelerômetro e GPS capaz de identificar irregularidades (buracos) em pavimentos viários?

Critérios PICO: A questão de pesquisa será analisada e estruturada aplicando a estratégia PICO de Petticrew e Roberts (2005). PICO representa um acrônimo para Population, Intervention, Comparison e Outcome.

População: Identificação de buracos no asfalto.

Intervenção: Uso de acelerômetros e GPS para a identificação dos buracos.

Controle: Artigos da revisão exploratória.

Resultado: Dispositivo capaz de identificar buracos no asfalto, utilizando machine Learning.

Palavras-chave: Foi utilizado as seguintes palavras chaves em inglês - Accelerometer, Detect, Find, Identify, Observe, Pothole, Pavement, Irregularity.

Aplicou-se os operadores lógicos “OR” e “AND” na aplicação das palavras chaves: ("detect" OR "find" OR "identify" OR "observe") AND ("pothole" OR "pavement irregularity") AND ("accelerometer").

Crítérios de Seleção das Bases de Busca:

A escolha das fontes de busca foi realizada por meio dos seguintes critérios: Bibliotecas digitais on-line, bases eletrônicas indexadas, anais de eventos das áreas, consultas a especialistas e patentes. Outras fontes: teses e dissertações (USP, IFSP mestrado do IPT) e bibliografias da análise exploratória.

Baseado nos critérios citados, foram selecionadas as fontes do Quadro 1, por terem grande relevância mundial pela produção literária de alta qualidade e por incluírem as principais revistas literárias e eventos científicos no âmbito da engenharia e de tecnologia:

Nome	URL
ACM	https://www.acm.org/
IEEE Digital Library	https://ieeexplore.ieee.org/
ISI Web of Science	http://isiknowledge.com
Scopus	https://www.scopus.com/

Quadro 1 - Bases de pesquisa utilizadas na revisão sistemática

Fonte: Elaborado pelo autor

Idioma dos Trabalhos: Tratando-se da maior aceitação internacional das pesquisas científicas na área, o inglês foi escolhido para ser o idioma utilizado nas buscas das bases científicas selecionadas, com exceção da dissertação referência, que está em português.

Crítérios de Inclusão e Exclusão dos Trabalhos:

Nessa revisão sistemática, foram aplicados critérios de inclusão e exclusão, com a finalidade de selecionar os trabalhos científicos que teriam uma relevância para solucionar o problema de pesquisa na qual esse artigo está propondo, sendo eles:

Crítérios de inclusão:

- (1) Celular e Machine Learning;
- (2) Celular e Algoritmos Determinísticos;
- (3) Acelerômetro e Machine Learning;
- (4) Acelerômetro e Algoritmos Determinísticos;
- (5) Otimiza falhas de GPS;
- (6) Trabalho científico teórico sobre o tema.

Crítérios de exclusão:

- (6) O estudo não abrange a área de identificação de buracos no asfalto;
- (7) O estudo não utiliza acelerômetros para medir vibração;
- (8) O estudo já encontrado em outras buscas, duplicidade;
- (9) O estudo completo não disponível para acesso por meio digital;
- (10) O estudo não coleta as informações de forma a auxiliar o problema proposto.

Estratégia para seleção dos estudos:

Nos artigos retornados após a aplicação da string de busca, foi aplicado o seguinte procedimento para encontrar os artigos mais relevantes para a pesquisa.

Realizar a leitura do título e resumo (abstract)
Aplicar critérios de inclusão e critérios e exclusão;
Realizar a leitura da introdução e conclusão;
Reaplicar critérios de inclusão e critérios e exclusão;
Realizar a leitura do texto completo;
Reaplicar critérios de inclusão e critérios e exclusão;
Criação e tabulação da lista refinada final.

Estratégia de extração e síntese dos dados:

Para todos os estudos da lista final de aceite foi realizada a leitura integral deles para extração das seguintes informações

Calibração dos sensores;
Coleta de dados;
Análise dos dados e rotulagem dos buracos;
Implementação do *machinelearning*;
Aplicar a solução in loco para identificar os buracos;
Transmissão e geolocalização dos dados.

Estratégia de sumarização dos dados:

- Leitura Integral;
- Análise, contextualização e entendimento dos resultados;
- Comparação entre as obras;
- Contribuições para esta pesquisa;

Condução

Para a etapa da condução da revisão sistemática deste estudo foram aplicadas as etapas propostas por Felizardo et al (2017), as quais estão detalhadas em cada uma das subseções a seguir, sendo identificação, seleção e extração/síntese dos dados dos estudos.

A busca, com a string ajustada, foi aplicada nos motores de busca contendo os seguintes dados:

Data de busca:	07/02/2022
Amplitude da busca:	Scopus, Web of Science, ACM e IEEE.
Idioma:	Inglês
Campos pesquisados:	Abstract, Title and Key Words
Período considerado:	Artigos publicados entre 2017 e 2022
Palavras-chave:	Detect, find, identify, observe, pothole, pavement, irregularity, accelerometer.
String executada:	("detect" OR "find" OR "identify" OR "observe") AND ("pothole" OR "pavement irregularity") AND ("accelerometer")
	Retornou 91 documentos

3.1 Seleção final dos estudos

No Quadro 2 são mostrados os trabalhos que foram selecionados por essa revisão sistemática

ID	Nome dos artigos
1	Evaluation of Detection Approaches for Road Anomalies Based on Accelerometer Readings- Addressing Who's Who
2	Mobile Crowdsensing Based Road Surface Monitoring Using Smartphone Vibration Sensor and Lorawan
3	A Machine Learning Approach to Road Surface Anomaly Assessment Using Smartphone Sensors
4	Vibration-based damaged road classification using artificial neural network
5	Smart patrolling: An efficient road surface monitoring using smartphone sensors and crowdsourcing
6	A deep learning approach to automatic road surface monitoring and pothole detection
7	Pothole detection using accelerometer and computer vision with automated complaint redressal
8	MagTrack: Detecting Road Surface Condition using Smartphone Sensors and Machine Learning
9	An early warning system for traffic and road safety hazards using collaborative crowd sourcing
10	Desenvolvimento de um sistema automático para detectar e avaliar buracos em pavimentos viários, utilizando acelerômetros MEMS

Quadro 2 - Artigos selecionados após a conclusão da revisão sistemática

Fonte: Elaborado pelo autor

O processo realizado é resumido na Figura 15:



Figura 15 - Diagrama das etapas para a seleção dos artigos finais
Fonte: Elaborado pelo autor

3.2 Extração e síntese dos dados

Para esta seção foram utilizados os campos de extração definidos no protocolo de pesquisa nesta pesquisa.

Cavalcante (2019) desenvolveu um dispositivo na qual é capaz de coletar informações como Acelerômetro, giroscópio, latitude, longitude e velocidade. Esse módulo foi implementado na bandeja do veículo, com a finalidade de coletar dados mais realistas, não estando passivo pelo filtro mecânico do amortecedor. Utilizando conceitos de algoritmos determinísticos, o STDEV(Z) modificado, foi capaz de identificar buracos com uma taxa de acerto de 70% dos buracos um raio de 10 metros, e 100% em um raio de 20 metros. Utilizando um smartphone para receber os dados primordiais vindo desse módulo, via bluetooth, e transmitindo-os via 4G, utilizando protocolo MQTT para uma plataforma de IoT, o ThingsBoard, plotando essas informações geolocalizadas em um mapa.

Cavalcante (2019) contribuiu para essa dissertação com relação à calibração do acelerômetro de baixo custo, demonstrando a calibração por estática e por aceleração centrípeta. Realizou um comparativo no domínio das frequências, do acelerômetro de baixo custo em relação a um acelerômetro industrial, validando esse tipo de sensor, se for aplicadas técnicas como filtro passa baixo e média móvel. Descreveu a técnica de reorientação da aceleração, utilizando ângulos de Euler, sendo essencial para o bom andamento da pesquisa.

O trabalho do Cavalcante apresentou hardware dedicado e instalado na bandeja do amortecedor do veículo, captando as vibrações dos buracos diretamente, sem passar pelo filtro do amortecedor. Demonstrou a dificuldade em lidar com a imprecisão do módulo de GPS, e como lidar com esses erros, utilizando o algoritmo DBSCAN para agrupar os valores.

Varona (2018) propôs identificar automaticamente os diferentes tipos de superfície da estrada e distinguir automaticamente os buracos e as lombadas, utilizando aprendizagem profunda. Foi aplicado diferentes modelos de deeplearning: redes neurais convolucionais, redes LSTM e modelos de computação de reservatórios. Os experimentos foram realizados com informações do mundo real, e os resultados mostraram uma precisão promissora na resolução de ambos os problemas. Foi importante para descrever a técnica de slidingwindow com a objetivo de segmentar os dados e os métodos de expansão de dados (data augmentation), com a finalidade de evitar os overfitting.

Basavaraju (2019) analisou diferentes técnicas de aprendizagem de máquina supervisionadas para rotular efetivamente as condições da superfície rodoviária usando dados de acelerômetro e GPS, coletados por um celular. O trabalho concentrou-se na classificação de três rótulos de classes principais: estrada lisa, buracos e rachaduras transversais profundas. Foi importante para descrever a extração das features no domínio do tempo, frequência e wavelet, demonstrando também a importância de trabalhar com os 3 eixos do acelerômetro. Aplicou o algoritmo de machinelearning SVM, descrevendo todas as etapas até a conclusão do modelo final.

Pooja (2017) apresenta um modelo analítico de um sistema que detecta deformidades no pavimento. O sistema apresenta um aplicativo móvel que captura o perfil das vibrações, utilizando um acelerômetro embutido no smartphone e também o GPS interno, para obter a localização das deformidades no pavimento. Foi importante para descrever o filtro passa baixo gaussiano para tratar o dataset antes do treinamento. Foi descrito também as etapas de treinamento utilizando o Sequentialminimaloptimization (SMO), que é um classificador binário baseado no SVM

Carlos (2017) elaborou um software de simulação de gerador de dados, chamado de PhotoleLab, que tem a finalidade de gerar dados dos 3 eixos de aceleração, velocidade, veículo utilizado e por fim alguns rótulos como lombadas e buracos. Esse artigo foi importante por realizar um comparativo de diversas técnicas de identificação de buracos, desde algoritmos estatísticos até machinelearning, alcançando as seguintes médias da taxa de acerto: Z-THRESH (45%), Z-DIFF (51,6%), STDEV(Z) (74,4%), GZERO (7,4%), NERICELL (33,6%), POTHOLE PATROL (8,6%) e SVM(Z) (87,7%). Descreveu também as seguintes técnicas de extração das features: Decomposição estacionária wavelettransform (SWT), média dos valores do acelerômetro, Desvio Padrão, Coeficiente de Variação e a diferença entre os valores Max e Min.

Para evitar acidentes e melhorar a segurança da estrada, Dey (2019) propôs um modelo baseado em machinelearning, o MagTrack, para detectar buracos e ranhuras nas estradas. Coletou-se várias condições da superfície da estrada, como estradas lisas, estradas irregulares, buracos e quebra-velocidade, utilizando acelerômetro embutido no Smartphone, e analisando vários algoritmos de classificação como Random Forest (RF), Random Tree (RT) e Support Vector Machine (SVM). Este artigo foi importante por ter descrito todas as camadas de treinamento do SVM, além de ter descrito sobre os algoritmos BestFit, Ranker e GreedyStepp, para a extração das features.

Seid (2020) apresentou um monitoramento de superfície rodoviária baseado em coletas de vibrações em dispositivos móveis usando sensores de Smartphone e uma rede LoRaWAN. Este artigo foi importante por ter descrito todas as etapas de um treinamento

SVM, realizando uma prova de conceito em campo. Utilizando o protocolo de comunicação LoRa, para transmitir os dados até um gateway, utilizando o conceito de fog-computing. Extraíu as seguintes features no domínio do tempo: Média, min, max, desvio padrão (STD), inclinação, curtose, intervalo interquartil (IQR), desvio absoluto mediano (MAD), área sob a curva e área sob a curva ao quadrado e no domínio da frequência, usando a Transformada rápida de Fourier (FFT), extraíndo as características de energia, de entropia e da frequência média.

Kandoi (2021) propôs uma solução que faz o uso de sensores contidos em um smartphone, juntamente com um comparativo de técnicas baseadas em processamento visual para detectar buracos em tempo real, e o resultado foi, respectivamente, a acurácia da técnica Random Forest de 89,63%, enquanto o processamento visual foi de 83%. Além do Random Forest, foi utilizado também o DecisionTree (83,95%), NaiveBayes (84,52%), LogicRegression (82,61%) e o K — NearestNeighbors (KNN) (85,86%). Esse artigo foi importante por demonstrar que o uso de técnicas que usam a aceleração foi mais assertivo do que técnicas que usam processamento visual.

Singh (2017) apresenta uma técnica de sensoriamento baseada em smartphones para detectar as condições da superfície da estrada. Os sensores embutidos no smartphone, como acelerômetro e GPS, têm sido usados para observar as condições da estrada. O artigo focou na otimização da precisão no processo de classificação e identificação das condições da superfície da estrada usando a técnica Dynamic Time Warping (DTW). Esse artigo foi importante para demonstrar algumas das técnicas de tratamento dos dados, como Média Móvel Simples (SMA), Filtro passa banda.

Purnama (2018) explora o uso da Rede Neural Artificial (ANN) para detectar anomalias rodoviárias com base em dados do acelerômetro de veículos. O veículo é equipado com um smartphone que possui um acelerômetro de 3 eixos e sensores de geolocalização. Em seguida, o veículo é usado para escanear a rede rodoviária com várias anomalias rodoviárias, como buracos, lombada e juntas de expansão. Este artigo foi importante para identificar técnicas de filtragem, como o Filtro Savitzky-Golay removendo o ruído desses dados de aceleração.

3.3 Análise dos dados

Durante a análise da síntese dos artigos, foi possível identificar similaridades entre alguns estudos. Por contribuírem com a pergunta de pesquisa e pelo destaque dado a estes temas, no Quadro 3 é demonstrado as principais contribuições para esta pesquisa, considerando as 5 etapas descritas nesse trabalho.

ID	Artigos	Etapas de contribuição					
		1ª Planejamento do hardware e Calibração dos sensores	2ª Dataset Coleta de dados e Análise dos dados e rotulagem dos buracos		3ª Implementação do machine learning	4ª Aplicação da solução in loco na identificação dos buracos e lombadas	5ª Transmissão e geolocalização dos dados
1	A deep learning approach to automatic road surface monitoring and pothole detection		X	X			
2	A Machine Learning Approach to Road Surface Anomaly Assessment Using Smartphone Sensors		X	X	X		
3	An early warning system for traffic and road safety hazards using collaborative crowd sourcing		X	X			
4	Desenvolvimento de um sistema automático para detectar e avaliar buracos em pavimentos viários, utilizando acelerômetros MEMS	X		X		X	X
5	Evaluation of Detection Approaches for Road Anomalies Based on Accelerometer Readings- Addressing Who's Who			X	X		
6	MagTrack: Detecting Road Surface Condition using Smartphone Sensors and Machine Learning			X	X		
7	Mobile Crowdsensing Based Road Surface Monitoring Using Smartphone Vibration Sensor and LoraWAN				X	X	X
8	Pothole detection using accelerometer and computer vision with automated complaint redressal			X			
9	Smart patrolling: An efficient road surface monitoring using smartphone sensors and crowdsourcing			X			
10	Vibration-based damaged road classification using artificial neural network			X		X	

Quadro 3 - Contribuições dos artigos selecionados para as etapas da solução proposta
Fonte: Elaborado pelo autor

3.4 Considerações Finais da Seção

Na presente seção foram abordados a teoria que sustenta e delimita o problema de pesquisa, bem como os artigos selecionados pela Revisão Sistemática e suas respectivas contribuições para a pesquisa. Na próxima seção será fornecida uma definição do experimento com o objetivo de destacar a relevância dos artigos identificados na seção anterior, na solução do problema proposto pela presente pesquisa.

4 RESULTADOS

Essa seção consiste na apresentação teórica do experimento, aplicando os conceitos descritos, tanto na parte teórica, como na revisão sistemática da literatura apresentados na seção anterior.

4.1 Fases do experimento

Esse experimento consiste em 5 etapas: 1ª - Planejamento do hardware e Calibração dos sensores, 2ª - Dataset, coleta e análise dos dados, com a rotulagem de buracos, 3ª - Implementação do *machinelearning*, 4ª - Aplicação da solução in loco na identificação dos

buracos e a 5ª - Transmissão e geolocalização dos dados. Essas etapas são demonstradas na Figura 16.

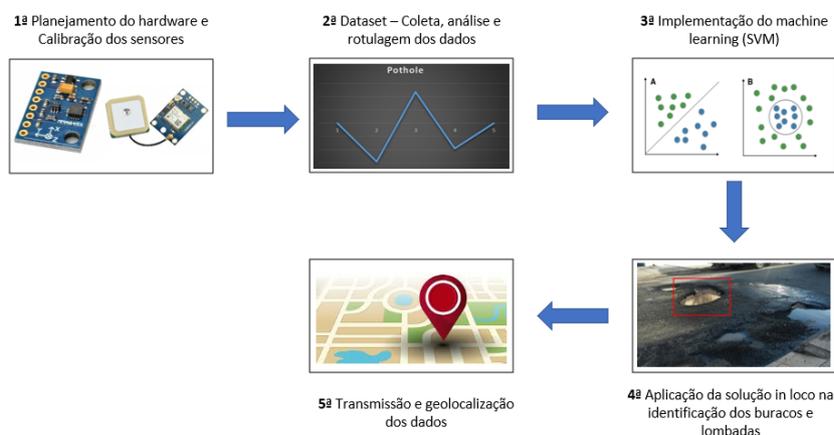


Figura 16 - Etapas do andamento do projeto

Fonte: Elaborado pelo autor

4.1.1 1ª - Planejamento do hardware e Calibração dos sensores:

As fases do processo são baseadas na análise dos artigos descritos em 3.1.A primeira fase é relacionada ao Planejamento do hardware e Calibração dos sensores. A arquitetura do hardware foi inspirada e evoluída a partir do projeto do Cavalcante (2019), baseados no microcontrolador ESP32-WROOM-32, o acelerômetro MPU 6050, GPS e a memória de armazenamento (esses dois últimos fornecidos externamente por meio de celular comercial). Uma das evoluções propostas no presente trabalho foi juntar essas duas funcionalidades em uma única placa, sem a necessidade de um celular intermediário ao sistema, transmitindo as informações diretamente para um gateway, via LoRa.

Diferente dos demais artigos abordados, por desenvolver o próprio hardware de coleta dos valores de vibração, o trabalho do Cavalcante teve a flexibilidade de instalar o módulo em um local mais estratégico, coletando os dados de vibração na bandeja do amortecedor do carro, não sendo afetado pelo filtro do amortecedor do veículo, extraindo valores mais precisos do fenômeno que será analisado. Na Figura 17 é mostrado o posicionamento de seu hardware contendo o acelerômetro.



Figura 17 - Posicionamento dos módulos de acelerômetro

Fonte: Adaptado de Cavalcante (2019)

Além dessa nova proposta de coleta de dados, Cavalcante (2019) mostrou como calibrar os acelerômetros. Utilizando a gravidade como referência, assumindo $9,81 \text{ m/s}^2$ no eixo Z, e 0 m/s^2 nos eixos X e Y, realizando a média de diversas coletas. O procedimento descrito foi:

1. inicia a coleta de 1024 amostras de aceleração dos eixos X, Y e Z;
2. calcula a média para cada eixo;
3. repete os itens 1 e 2 por 10 vezes;

4. calcula a média das 10 repetições;
5. armazena os resultados nos registradores de offset de cada eixo;
6. armazena-os em memória permanente do ESP32, para recuperar a cada reinicialização dos acelerômetros.

Na Figura 18 é mostrado o gráfico com os valores antes e depois dessa calibração.

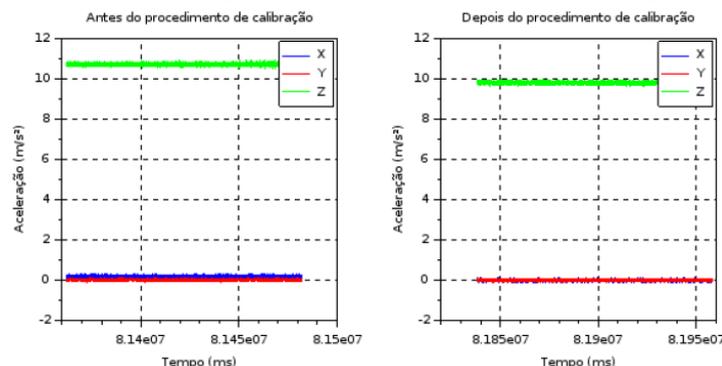


Figura 18 - Calibração por estática
Fonte: Adaptado de Cavalcante (2019)

O Cavalcante (2019) também demonstrou um comparativo do acelerômetro de baixo custo, o MPU-6050, com um acelerômetro com o preço mais elevado, fabricado pela Silicon Designs, modelo 2210 002. Realizou o ensaio coletando a resposta de atuação no âmbito das frequências, em uma mesa de vibração fabricada pela empresa VEB RFT MESSELEKTRONIK, modelo 11076, na qual é possível controlar a altura e a velocidade na qual se movimenta, obtendo os seguintes gráficos comparativos demonstrados na Figura 19, realizando o ensaio a 2 Hz.

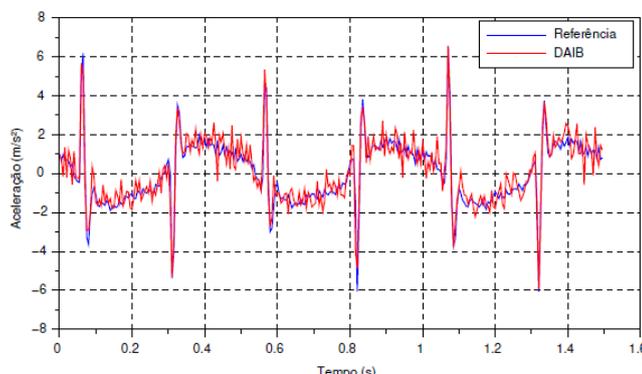
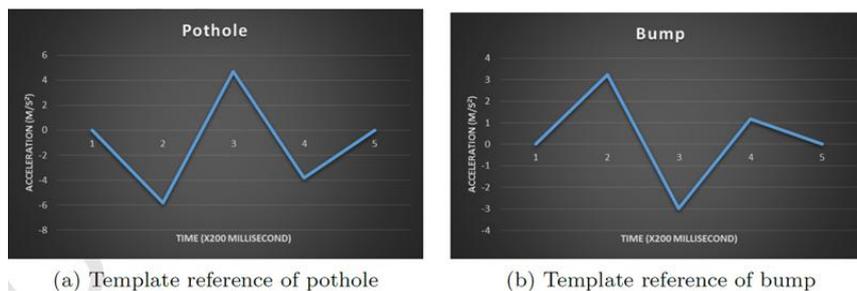


Figura 19 - Resultado do comparativo em frequência do eixo Z do para 2 Hz
Fonte: Adaptado de Cavalcante (2019)

4.1.2 2ª - Dataset, coleta e análise dos dados, com a rotulagem de buracos

Nessa fase, são abordados os artigos que contribuíram para a segunda etapa do projeto, que consiste na construção do dataset, incluindo a parte de coleta, a análise dos dados até a rotulagem dos buracos.

BASAVARAJU (2020) descreve a importância de coletar as vibrações em conjunto com gravações visuais, para o auxílio no processo de rotulagem dos dados. No processo de rotulagem, o momento de cada ação foi identificado pela câmera, e de acordo com os padrões descrito por SINGH (2017), é possível uma identificação do fenômeno, como é demonstrada na Figura 20, com padrões da resposta de aceleração do eixo Z para buracos e lombadas.



(a) Template reference of pothole

(b) Template reference of bump

Figura 20 - Assinatura no sinal para buracos e lombadas

Fonte: Adaptado de Singh (2017)

Caso os dados coletados apresentem ruídos que estejam atrapalhando sua interpretação, visando a facilitar a identificação dos fenômenos, Carlos (2018) utiliza técnicas de média móvel para eliminar possíveis ruídos, podendo analisar de forma mais clara o fenômeno.

No treinamento utilizado em *machinelearning*, é necessário um dataset com dados suficientes, heterogêneo, pois se possuir poucos dados, há a possibilidade de causar um *underfitting* que no caso é elaborar um modelo de um treinamento que não consegue traduzir a aplicação, ou um *overfitting*, que gera um modelo a partir de um treinamento que não seja genérico, funcionando para uma situação específica.

Varona (2019) demonstrou uma técnica matemática que é possível a expansão dos dados (*Data Augmentation*), a partir de um dataset inicial, modificando as amostras existentes para criar novas amostras para que seja possível a obtenção de mais dados na realização do treinamento, modificando sua frequência base como é demonstrado na Figura 21. Ao introduzir variações nos dados de treinamento, o modelo evita o *overfitting*, generalizando o aprendizado do modelo, se tornando mais flexível em novos ambientes de coletas.

Um exemplo de sua aplicação é nas possíveis variações de velocidade do veículo. Independente dessa variação, o padrão do sinal continua o mesmo, variando algumas características como a amplitude e a frequência. Com o *Data Augmentation* é possível simular velocidades diferentes coletando o mesmo buraco, melhorando o aprendizado do modelo, tornando-o mais heterogêneo.

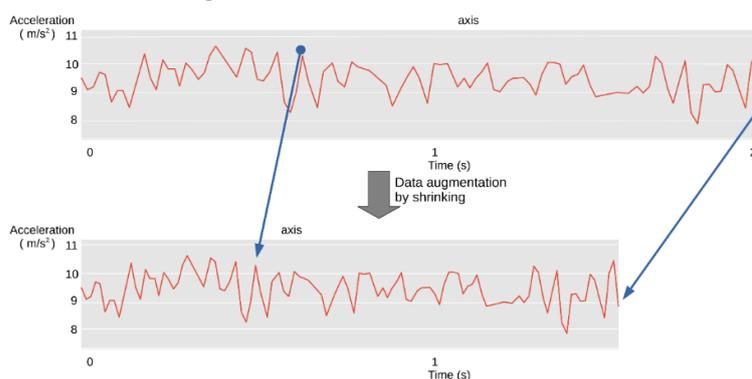


Figura 21 - Técnica para expandir os dados, base de 20 ms para 16 a 24 ms

Fonte: Adaptado de Varona (2019)

4.1.3 3ª - Implementação do *machinelearning*:

Nessa etapa serão demonstrados os métodos necessários para realizar o treinamento de SVM, e a justificativa do uso dessa técnica.

Para se realizar um treinamento utilizando SVM, é necessário a extração das *features*, evidenciando o fenômeno na qual se quer obter para o treinamento supervisionado

de *machinelearning*. Dos trabalhos científicos abordados pela RSL, foram obtidas as seguintes *features* apresentadas no Quadro 4.

Artigos	Extração das features
A Machine Learning Approach to Road Surface Anomaly Assessment Using Smartphone Sensors	Domínio do tempo: Valor Máximo, Valor Mínimo, Valor Médio, Valor RMS, Valor de Pico para Pico e Valor Médio de Dez Domínio de frequência: densidade espectral de energia de sinais de vibração Domínio de wavelet: as ondas mortlet e as ondas Daubechies 6 e 10
An early warning system for traffic and road safety hazards using collaborative crowd sourcing 4.0	Aceleração dos 3 eixos
Evaluation of Detection Approaches for Road Anomalies Based on Accelerometer Readings- Addressing Who's Who 8.0	Decomposição estacionária wavelet transform (SWT), Média dos valores do acelerômetro, Desvio Padrão, Coeficiente de Variação e a diferença entre os valores Max e Min.
MagTrack: Detecting Road Surface Condition using Smartphone Sensors and Machine Learning 4.0	Algoritmos: BestFit, Ranker e GreedyStepp
Mobile Crowdsensing Based Road Surface Monitoring Using Smartphone Vibration Sensor and Lorawan	Domínio do tempo: Média, min, max, desvio padrão (STD), inclinação, curtose, intervalo interquartil (IQR), desvio absoluto mediano (MAD), área sob a curva e área sob a curva ao quadrado. Domínio da frequência: Transformada rápida de Fourier (FFT), extraindo as seguintes características: energia, entropia e frequência média.
Pothole detection using accelerometer and computer vision with automated complaint redressal	Aceleração e velocidade angular dos 3 eixos

Quadro 4 - *Features* extraídas pelos artigos encontrados na revisão sistemática
Fonte: Elaborado pelo autor

Carlos (2018) realizou um comparativo das técnicas sendo possível analisar a taxa de acerto no mesmo ambiente, justificando a técnica mais eficiente para identificar buracos. Na Tabela 1 é mostrado os algoritmos aplicados, na qual pode-se verificar que o SVM(Z) – Suporte Vector Machine aplicado no eixo Z, foi a técnica com a maior taxa de acerto, em comparação com as demais técnicas aplicadas. Para coletar os dados, utilizou-se o smartphones Moto G Android, com uma frequência de amostragem de 50 Hz.

Algoritmo	Acurácia
Z-THRESH	0,450
Z-DIFF	0,516
STDEV (Z)	0,744
GZERO	0,074
NERICELL	0,336
POTHOLE PATROL	0,086
SVM (Z)	0,877

Tabela 1 - Acurácia dos algoritmos de Machine Learning
Fonte: Adaptado de Carlos (2018)

Seid (2020) também fez um comparativo de resultados entre técnicas, e o resultado também foi que o SVM se demonstrou uma técnica eficiente para identificar buracos, conforme a Tabela 2. Os dados do acelerômetro foram coletados por um smartphone, com uma frequência de amostragem de 16 Hz.

Algoritmo	Acurácia
Support Vector Machine (SVM)	0,98
Multi-class Logistic Regression	0,68
Convolutional Neural Network (CNN)	0,96

Tabela 2 - Acurácia dos algoritmos
Fonte: Adaptado de Seid (2020)

4.1.4 4ª - Aplicação da solução *in loco* na identificação dos buracos

Varona (2019) descreveu a aplicação da técnica *slidingwindow*, na qual consiste em processar os dados em uma determinada janela de tempo. O autor identificou que de acordo com o tempo da janela, sendo sobreposta ou não, e pelo ajuste de amostras, obteve resultados variados. Varona embarcou um modelo gerado por uma rede CNN (Convolutional Neural Network), realizando testes sem sobreposição, sendo nomeado como janela de passo 1, e com sobreposição, sendo eles o $\frac{1}{4}$ e $\frac{1}{2}$ de sobreposição entre as janelas e com 5 amostras de sobreposição. Na Figura 22 é mostrado a acurácia obtida em relação a quantidade de amostras utilizadas.

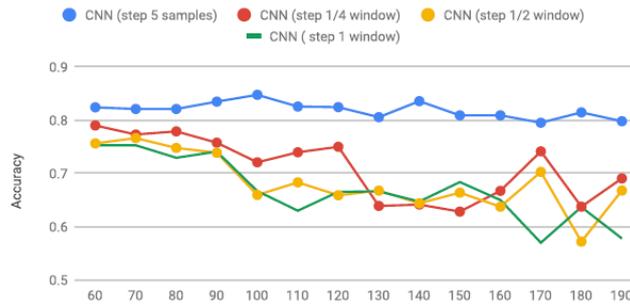


Figura 22 - Aplicação das técnicas baseadas em *slidingwindow*

Fonte: Adaptado de Varona (2019)

Basavaraju (2019) aplicou o conceito de *fog computing*, na qual o celular processava os dados gerados pelo acelerômetro, e com o treinamento embarcado foi capaz de identificar os buracos nas vias pavimentadas, em tempo real. Pelo fato de o celular possuir também um GPS, transmitiu somente os dados válidos geolocalizados, a classificação do que identificou, com a latitude e longitude da informação.

Seid (2020) propôs o uso da tecnologia LoRa (Long Range) para transmitir essas ocorrências em grandes distâncias para um *Gateway*, e posteriormente para um *Cloud server* (serviço disponível na nuvem) possibilitando uma rápida visualização em dashboards e em mapas georreferenciados. Na Figura 23 é mostrado a arquitetura da solução proposta pelo artigo.

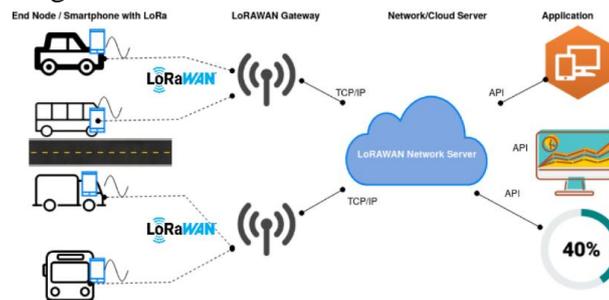


Figura 23 - Arquitetura da solução para identificar buracos

Fonte: Adaptado de Seid (2020)

4.1.5 5ª - Transmissão e geolocalização dos dados

Cavalcante (2019) pontuou em sua pesquisa que ao geolocalizar os dados em um mapa, havia uma elevada concentração de dados em um raio próximo, significando que aquelas informações representavam um único buraco, poluindo a visualização dos dados, conforme é demonstrado na Figura 24.

4.2 Considerações Finais da Seção

Nessa seção foram descritas as contribuições teóricas da revisão sistemática para as 5 etapas do experimento, desde contribuição para a área de hardware e calibração dos sensores, geração do dataset, rotulagem dos dados, transmissão da informação, e a geolocalização dessa ocorrência.

5 CONCLUSÃO

No início desse trabalho foram apresentadas diversas soluções pertinentes à identificação e classificação de falhas em vias pavimentadas, sendo utilizadas como base para a estruturação da revisão sistemática da literatura, cuja resposta à pergunta de pesquisa “como desenvolver um sistema embarcado, de baixo custo, baseado em *machinelearning* utilizando acelerômetro e GPS capaz de identificar irregularidades (buracos) em pavimentos viários?” indica que os artigos extraídos demonstraram a viabilidade na utilização de técnicas de calibração e filtros, assim como de acelerômetros do tipo MEMs para coletar a vibração. Técnicas de expansão de dados e de rotulagem mostraram-se eficientes para aumentar a base do dataset e identificar assinaturas dos buracos, necessários para realizar a rotulagem. Na parte do treinamento, SVM foi a técnica que teve a maior taxa de acerto dentre as demais técnicas e utilizando LoRa como protocolo de comunicação, para identificar e classificar os buracos em campo e em tempo real, transmitindo-os para uma plataforma de IoT.

Este trabalho contribui para a academia com uma revisão sistemática da literatura, mostrando os principais trabalhos relacionados ao tema específico e para a indústria, com o resultado da análise dos artigos, resultando no processo indicado.

A principal limitação desta pesquisa corresponde ao fato que o processo, embora estabelecido em termos teóricos, encontra-se em fase de implementação, de modo que não foi possível apresentar resultados empíricos da proposta, indicando como trabalho futuro, a colocação em prática do processo proposto, desenvolvendo um módulo capaz de coletar informações de vibração, velocidade e GPS, gerar um *dataset*, rotulando os dados, e por fim implementar identificação e classificação das informações dos buracos. Espera-se que com essas técnicas aplicadas, seja possível desenvolver um dispositivo capaz de identificar e classificar buracos nas vias pavimentadas, geolocalizando as ocorrências, em tempo real.

REFERÊNCIA BIBLIOGRÁFICA

- ALLOGHANI, M. . A.-J. D. . M. J. . H. A. . A. A. J. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. **Springer, Cham**, p. 3-21, 2020.
- ALVES, S. O que é o GPS e como funciona? **IME _ USP**, 26 abr. 2023. Disponível em: <<https://www.rpm.org.br/cdrpm/59/5.htm>>.
- ANALOG DEVICES. **ADXL150/ADXL250**. Norwood, p. 15. 1988.
- BASAVARAJU, A. et al. A Machine Learning Approach to Road Surface Anomaly Assessment Using Smartphone Sensors. **IEEE Sensors Journal**, 5 março 2020. 2635–2647.
- BELLEMARE, M. ; DABNEY, W.; MUNOS, R. A Distributional Perspective on Reinforcement Learning. **arxiv**, p. 19, 2017.

BERNUCCI, B. et al. **Pavimentação asfáltica Formação básica para engenheiros**. Rio de Janeiro: Associação Brasileira das Empresas Distribuidoras de Asfaltos – ABEDA, 2010.

BIDIKAR, B. Satellite Clock Error and Orbital Solution Error Estimation for Precise Navigation Applications. **Scientific Research**, Curitiba, 01 Fevereiro 2014. 22-26.

CARLOS, ; EZRA ARAGÓN, M.; GONZÁLEZ, L. C. Evaluation of Detection Approaches for Road Anomalies Based on Accelerometer Readings—Addressing Who’s Who. **IEEE Transactions on Intelligent Transportation Systems**, 10 outubro 2018. 3334–3343.

CAVALCANTE, D. B. **Desenvolvimento de um sistema automático para detectar e avaliar buracos em pavimentos viários, utilizando acelerômetros MEMS**. Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP). São Paulo, p. 141. 2019. (Mestrado Profissional em Automação e Controle).

CERVANTES, J. et al. A comprehensive survey on support vector machine classification: Applications, challenges and trends. **Neurocomputing**, p. 189-215, 2020.

CNT - CONFEDERAÇÃO NACIONAL DO TRANSPORTE. **Pesquisa CNT de rodovias 2021**. Brasília, p. 231. 2021. (CDU 656.11(81)(047)).

COSTA, F.; GONÇALVES, C. V. R.; BRITO, P. Patologias em vias urbanas: Estudo de caso na Cidade de Vitória da Conquista – Bahia. **Multidisciplinar Núcleo do Conhecimento**, 20 maio 2020.

DNIT - DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES. **Manual de Pavimentação**. Rio de Janeiro. 2022.

DÜNTSCH, ; GEDIGA,. Confusion Matrices and Rough Set Data Analysis. **Journal of Physics: Conference Series**, 4 Fevereiro 2019.

ESPRESSIF SYSTEMS. **ESP32 Series Datasheet**. Shanghai, p. 70. 2023.

FELIZARDO, R. et al. **Revisão sistemática da literatura em engenharia de software**. São Paulo: Grupo GEN, 2017.

HELTEC AUTOMATION. WiFi LoRa 32 (V2.1), 17 Abil 2023. Disponível em: <<https://heltec.org/project/wifi-lora-32/>>.

IBM. Learn how to avoid underfitting, so that you can generalize data outside of your model accurately. **IBM**, 2020. Disponível em: <<https://www.ibm.com/topics/underfitting#:~:text=Underfitting%20is%20a%20scenario%20in,training%20set%20and%20unseen%20data.>>.

IBORRA, S. et al. Performance Evaluation of LoRa Considering Scenario Conditions. **Journal**, 3 Março 2018. 19.

INVENSENSE. **MPU-9250 Product Specification Revision 1.1**. San Jose, p. 42. 2016.

IOTLABS. Mapa a presença da Rede Neutra IoT LoRaWAN® American Tower. **iot-labs.io**, 05 abr. 2023. Disponível em: <<https://iot-labs.io/rede-neutra-iot-lorawan-da-american-tower/#cobertura>>.

IRAZABAL, J.-M.; BLOZIS,. NXP. **Application Note - I²C Manual (AN10216-01)**, 24 março 2003. Disponível em: <<https://www.nxp.com/docs/en/application-note/AN10216.pdf>>. Acesso em: 11 Abril 2023.

KANDOI, et al. Pothole Detection using Accelerometer and Computer Vision with Automated Complaint Redressal. **2021 3rd International Conference on Signal Processing and Communication (ICPSC)**, , 15 junho 2021. 257–264.

KUMAR , A. What are Features in Machine Learning? **Data Analytics - AI, Data, Data Science, Machine Learning, Blockchain, Digital**, 29 Outubro 2022. Disponível em: <<https://vitalflux.com/what-are-features-in-machine-learning/>>.

LEILANI H. GILPIN, D. B. B. Z. Y. A. B. M. S. L. K. Explaining Explanations: An Overview of Interpretability of Machine Learning. **The 5th IEEE International Conference on Data Science and Advanced Analytics (DSAA 2018)**, p. 10, 2019.

LOOSLI,. Comments on the “Core Vector Machines: Fast SVM Training on Very Large Data Sets”. **Journal of Machine Learning Research**, 2 Julho 2007. 291- 301.

MCCARTHY,. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. **AI Magazine**, 31 Agosto 1955. 13.

NASA. NASA Science Space Place Explore Earth and Space! **How Does GPS Work?**, 27 Junho 2019. Disponível em: <<https://spaceplace.nasa.gov/gps/en/>>.

NXP. **I²C-bus specification and user manual**. [S.l.], p. 62. 2021.

PACHECO,. O que é Machine Learning? **Computação Inteligente**, 11 Março 2021. Disponível em: <<http://computacaointeligente.com.br/conceitos/o-que-e-machine-learning/>>.

PADIERNA,. Hyper-Parameter Tuning for Support Vector Machines by Estimation of Distribution Algorithms, Dezembro 2017. 787-800.

PEREZ, ; WANG, J. The Effectiveness of Data Augmentation in Image Classification using Deep, 13 Dezembro 2017.

POOJA, P. R.; HARIHARAN, B. An early warning system for traffic and road safety hazards using collaborative crowd sourcing. **2017 International Conference on Communication and Signal Processing (ICCSP)**, 08 fevereiro 2017. 1203–1206.

PREFEITURA DE SÃO PAULO. Programa Recape, 25 Abril 2023. Disponível em: <https://recape.prefeitura.sp.gov.br/recape_/>.

PREFEITURA DE SÃO PAULO. SP156. **Portal de Atendimento Prefeitura de São Paulo**, 23 Abril 2023. Disponível em: <<https://sp156.prefeitura.sp.gov.br/portal/servicos/informacao?servico=952>>.

PURNAMA, ; FERGYANTO E. , G. Vibration-Based Damaged Road Classification Using Artificial Neural Network. **TELKOMNIKA (Telecommunication Computing Electronics and Control)**, outubro 2018. 2179-2190.

RANI DEY, et al. MagTrack: Detecting Road Surface Condition using Smartphone Sensors and Machine Learning. **2019 IEEE Region 10 Conference (TENCON)**, 12 dezembro 2019. 2485–2489.

SANTOS,. Protocolo MQTT: O Que é, Como Funciona e Vantagens. **Automação Industrial**, 15 Março 2023. Disponível em: <<https://www.automacaoindustrial.info/mqtt/>>.

SARAVATI, 17 Abril 2023. Disponível em: <<https://www.saravati.com.br/>>.

SAYERS, M. W.; GILLESPIE, T. D.; PATERSON, D. O. **Guidelines for Conducting and Calibrating**. Washington - USA: The International Bank for Reconstruction/ The World Bank, 1986.

SCIKIT-LEARN. Machine Learning in Python, 20 Abril 2023. Disponível em: <<https://scikit-learn.org/stable/index.html>>.

SEID, et al. Mobile Crowdsensing Based Road Surface Monitoring Using Smartphone Vibration Sensor and Lorawan. **Proceedings of the 1st Workshop on Experiences with the Design and Implementation of Frugal Smart Objects**, 21 setembro 2020. 36–41.

SINGH, ; THAKUR, ; SHARMA,. A review of supervised machine learning algorithms. **2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)**, p. 1310-1315, 2016.

SINGH, et al. Smart Patrolling: An Efficient Road Surface Monitoring Using Smartphone Sensors and Crowdsourcing. **Pervasive and Mobile Computing**, setembro 2017. 71–88.

TENSORFLOW. <https://www.tensorflow.org/lite/microcontrollers?hl=pt-br>. **For Mobile & Edge**, 24 dez. 2021. Disponível em: <<https://www.tensorflow.org/lite/tutorials?hl=pt-br>>.

THINGSBOARD. ThingsBoard Open-source IoT Platform, 05 abr. 2023. Disponível em: <<https://thingsboard.io/docs>>.

TURING,. Computing Machinery and Intelligence. **Mind**, Outubro 1950. 433-460.

VARONA, B.; MONTESERIN, ; TEYSEYRE,. A Deep Learning Approach to Automatic Road Surface Monitoring and Pothole Detection. **Personal and Ubiquitous Computing**, 27 maio 2019. 519–534.

YİLDİRİM, S. Hyperparameter Tuning for Support Vector Machines — C and Gamma Parameters. **Towardsdatascience**, 31 Maio 2020. Disponível em: <<https://towardsdatascience.com/hyperparameter-tuning-for-support-vector-machines-c-and-gamma-parameters-6a5097416167>>.