

# Relatório Técnico-Científico

Felipe Takasawa Paiva

## 1 Identificação

Referente ao Período de Novembro/2022 a Julho/2023.

Título: Detecção de Buracos em Vias com Aprendizado Federado

Bolsista: Felipe Takasawa Paiva

Orientadores: Luís Henrique Maciel Kosmalski Costa e Rodrigo de Souza Couto

Departamento: Programa de Engenharia Elétrica

Unidade: COPPE

Centro: Centro de Tecnologia

## 2 Resumo

O trabalho propõe a utilização de um modelo de aprendizado de máquina para detecção de buracos nas ruas através de imagens capturadas por uma câmera acoplada ao painel de um automóvel. A existência do modelo permite um futuro mapeamento dessas irregularidades nas vias e a utilização dessas informações de localização por autoridades responsáveis pela manutenção das ruas. Este trabalho abrange temas como visão computacional, aprendizado federado e transferência de aprendizado.

## 3 Introdução

Com a popularização da inteligência artificial, áreas como o aprendizado de máquina (*Machine Learning*) encontraram espaço em diversos campos e o uso de suas técnicas está presente em uma ampla diversidade de aplicações. O aprendizado de máquina se tornou uma ferramenta útil para a resolução de uma grande variedade de problemas de forma adaptativa.

Algumas aplicações, no entanto, tratam de dados sensíveis, cuja privacidade deve ser preservada. Além disso, o poder computacional de dispositivos como veículos autônomos e smartphones estimula o armazenamento de dados e aprendizado de forma local. Nesse contexto, o aprendizado federado oferece uma solução para o treinamento de modelos estatísticos de forma descentralizada. Os dispositivos de borda treinam o modelo localmente com seus próprios dados e compartilham os resultados do aprendizado local à entidade central, que digere-os para gerar um modelo global que envolve dados de outros dispositivos de borda. Ao final, o resultado do aprendizado global é enviado para os dispositivos de borda a fim de retroalimentar o modelo local, tornando este um processo distribuído.

Considerando o contexto de veículos autônomos, é essencial a prevenção de acidentes para a garantia da segurança no trânsito. Ademais, a presença de buracos gera um custo relacionado aos danos a longo prazo causados nos carros que passam pelas vias. Nesse sentido, a detecção de buracos na via é mister para permitir que os veículos evitem rotas perigosas para assegurar viagens livres de contratemplos e promover a preservação da longevidade do carro. Sendo assim, esse trabalho visa a criação de um modelo de aprendizado de máquina para a detecção de buracos nas vias com a utilização de aprendizado federado.

## 4 Objetivos

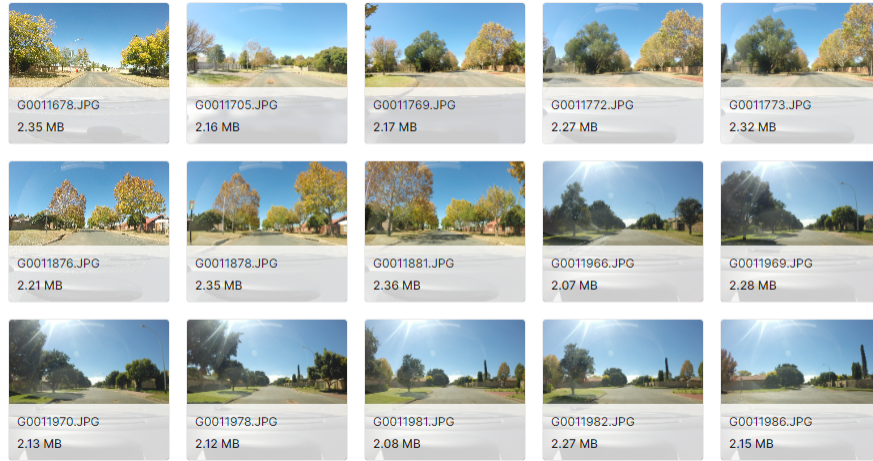
Assim sendo, O trabalho foi desenvolvido com o objetivo de obter um modelo capaz de detectar a presença de buracos em imagens de ruas de forma suficientemente rápida e precisa, de acordo com requisitos da aplicação, assim como próprio para futuro treinamento com novos dados em contexto de aprendizado federado.

Outro objetivo era encontrar os modelos de rede neural convolucional que melhor se adaptassem aos dados do contexto de detecção de buracos nas vias. Além disso, os modelos obtidos com aprendizado federado tiveram seu desempenho comparado com os modelos obtidos através de aprendizado clássico em respeito a métricas como acurácia e precisão.

## 5 Metodologia

### 5.1 Escolha da base de dados

Figura 1: Imagens do *dataset Road Pothole Images for Pothole detection*.



Primeiramente, foram reunidos os dados para o treinamento do modelo. Foi selecionada uma base de dados chamada *Road Pothole Images for Pothole detection*, disponível no site [www.kaggle.com](http://www.kaggle.com), um repositório de bases de dados públicas usadas para desenvolvimento de modelos de aprendizado de máquina. A base de dados escolhida possui 3777 imagens, sendo 1119 imagens com buracos e 2658 sem buracos, todas captadas por uma

câmera veicular, no estilo *dashcam* que é um modelo de câmera que acopla-se ao painel do automóvel, registrando a visão do motorista. Vale notar que todas as imagens foram capturadas na África do Sul e possuem características específicas da região, como coloração da terra, tipo de vegetação, condições de iluminação, tipo de calçada, tamanho e abundância de prédios, etc.

## 5.2 Transferência de aprendizado

Figura 2: Exemplos de imagens do *dataset ImageNet*.



A transferência de aprendizado consiste no aproveitamento de parte das camadas e pesos de uma rede pré-treinada para uma aplicação similar afim de acelerar radicalmente o processo de aprendizado de um novo modelo. Essa técnica possibilita melhor rendimento no treinamento de um modelo de aprendizado com um dataset pequeno.

Foi realizado o treinamento e teste do modelo utilizando a técnica de *transfer learning* com diversas redes neurais convolucionais pré-treinadas para identificar qual melhor se adequa ao contexto. As redes utilizadas foram **VGG16** (138,4M parâmetros), **ResNet50V2** (25,6M parâmetros),

**Xception** (22,9M parâmetros) e **MobileNetV3** (5,4M parâmetros, voltada para CPUs de dispositivos móveis), todas pré-treinadas utilizando a base de dados **ImageNet**, que contém mais de 14 milhões de imagens, separadas em mais de 20 mil classes.

### 5.3 Modelo federado

Posteriormente, foi realizado o treinamento e teste do modelo com a utilização das técnicas do aprendizado federado com assistência do framework Flower (<https://flower.dev/>), visto o caráter dos dados, que adequam-se ao chamado aprendizado horizontal, onde os tipos de dados coletados pelos diferentes nós contêm o mesmo conjunto de características.

A estratégia de agregação de pesos escolhida foi a FedAvg, em que o servidor realiza uma média ponderada dos pesos obtidos pelos clientes. As diferentes redes utilizadas na etapa anterior do trabalho foram também comparadas nesta etapa afim de visualizar novamente seu desempenho, agora no contexto de aprendizado federado.

Nos testes, havia 5 clientes, cada um com uma parte do conjunto de dados (sem sobreposição e todas com aproximadamente o mesmo tamanho) e um servidor. O servidor e cada cliente eram processos no mesmo computador e a latência de rede não foi simulada.

### 5.4 Ambiente de desenvolvimento

Todos os testes e treinamentos foram realizados localmente, em um computador com um processador AMD Ryzen 5 5600 3.5GHz (6 núcleos) e memória RAM de 32GB (2x16GB), 3200MHz. Não foi utilizada GPU, afim de melhor simular o ambiente de dispositivos móveis.

A linguagem de programação utilizada foi Python e o Framework utilizado para as redes neurais foi o Keras com o TensorFlow. O Framework usado para aprendizado federado foi o Flower. Foram também utilizadas as bibliotecas Numpy e Matplotlib para cálculos e visualização dos dados, respectivamente.

## 6 Resultados

### 6.1 Transferência de aprendizado

Figura 3: Métricas de treinamento dos modelos.

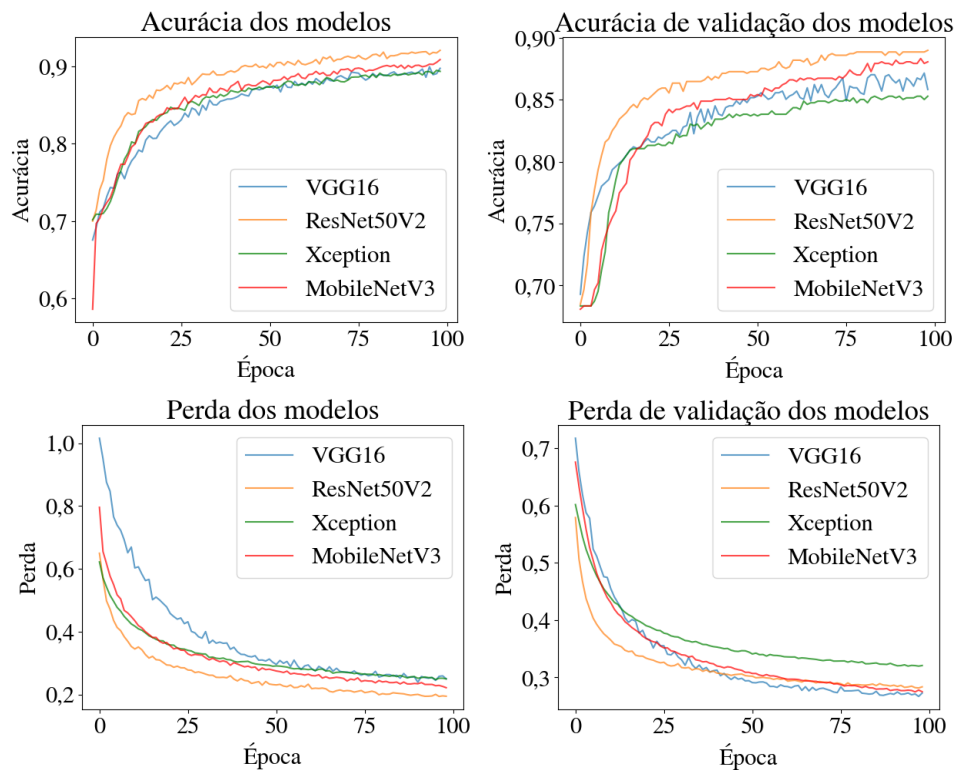


Tabela 1: Tempos de treinamento, transferência de aprendizado.

Rede neural	Tempo de treinamento (minutos)
VGG16	365
ResNet50V2	148
Xception	176
MobileNetV3	54

As métricas de acurácia e perda foram calculadas a cada época de treinamento. A acurácia é dada pela divisão do número de previsões corretas

pelo número total de previsões. A função de perda utilizada foi a *binary cross-entropy*.

## 6.2 Modelo federado

Figura 4: Métricas de treinamento dos modelos federados.

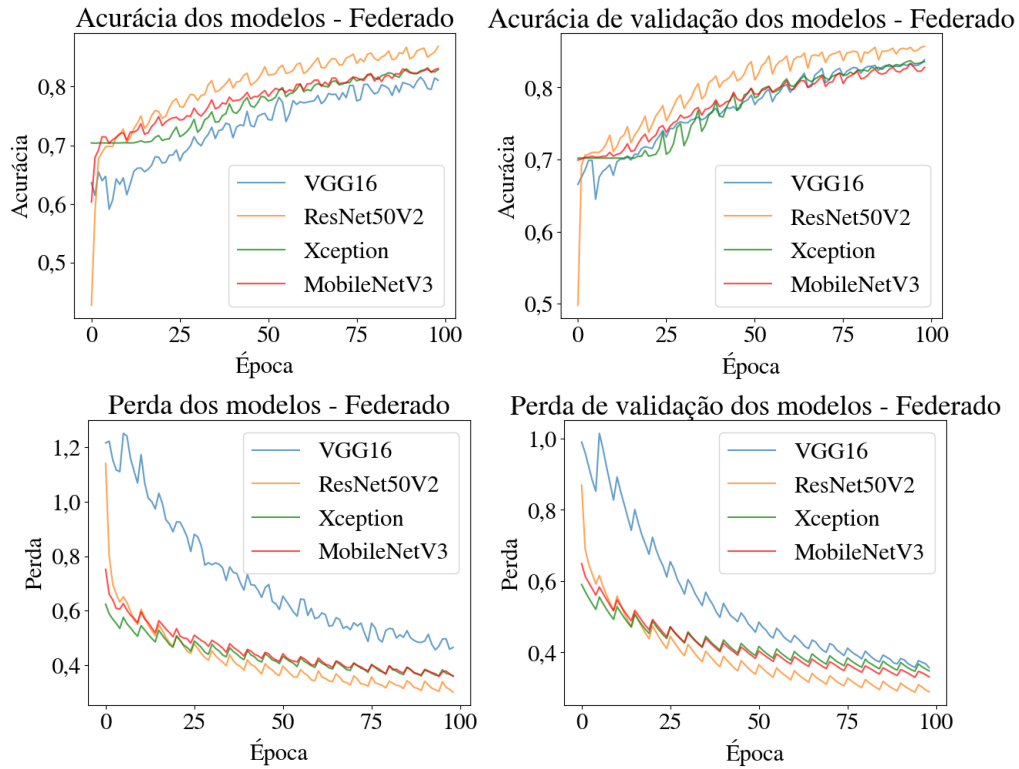


Tabela 2: Tempos de treinamento, aprendizado federado.

Rede neural	Tempo de treinamento (minutos)
VGG16	356
ResNet50V2	138
Xception	165
MobileNetV3	52

As métricas para os modelos federados foram calculadas a partir da média de cada métrica entre os clientes.

## 7 Conclusão

A partir dos resultados obtidos, é possível observar que duas redes se destacam tanto no contexto de aprendizado de máquina tradicional quanto no aprendizado federado: A ResNet50 e a MobileNetV3. Enquanto a ResNet50 apresenta maior acurácia, o tempo de treinamento da MobileNetV3 é o menor de todos, além de manter um nível de acurácia semelhante (até melhor em alguns casos) às outras redes utilizadas. Além disso, é possível notar que ao introduzir o aprendizado federado, o desempenho das redes se manteve relativamente semelhante, indicando possível viabilidade desse método para o contexto estudado.

## 8 Próximos passos

Para dar continuidade ao projeto, é interessante utilizar um conjunto de dados de imagens capturadas nas vias brasileiras, com o intuito de se adequar melhor às peculiaridades do nosso ambiente, como condições de luminosidade, tipo de asfalto, tipo de vegetação, cor do solo, etc. Ademais, poderia ser realizada uma comparação mais profunda entre as diferentes redes neurais utilizadas, empregando técnicas como matrizes de confusão e visualizando as diferenças entre os modelos sob a luz de mais métricas. Adicionalmente, para reduzir o impacto da aleatoriedade da divisão dos conjuntos de treino e validação, poderia ser realizada a técnica de validação cruzada.

Além disso, o estudo se beneficiaria com realização de testes de treinamento em dispositivos móveis, simulando melhor uma situação real do aprendizado federado, com poder de processamento menor e separação física entre os clientes, fazendo com que o atraso da rede e a possível perda de pacotes seja levada em conta.

## Referências

- [1] Peter Kairouz et al. “Advances and Open Problems in Federated Learning”. Em: *Found. Trends Mach. Learn.* 14.1-2 (2021), pp. 1–210.
- [2] Fernando Molano Ortiz et al. “Applications and Services Using Vehicular Exteroceptive Sensors: a Survey”. Em: *IEEE Trans. Intell. Veh.* (2022), pp. 1–20.



- [3] Qiang Yang et al. “Federated Machine Learning: Concept and Applications”. Em: *ACM Trans. Intell. Syst. Technol.* 10.2 (jan. de 2019).