Treinamento Federado Aplicado à Segmentação do Ventrículo Esquerdo

Vinicios B. da Silva[1], Renan R. de Oliveira[1,2], Antonio Oliveira-Jr[1,3] e Ronaldo M. da Costa[1]

[1]Instituto de Informática – Universidade Federal de Goiás (UFG), GO, Brasil

[2]Instituto Federal de Goiás (IFG), GO, Brasil

[3]Fraunhofer Portugal AICOS, Porto, Portugal

viniciosbarbosa@discente.ufg.br, renan.rodrigues@ifg.edu.br,

{antoniojr,ronaldocosta}@ufg.br

Abstract. *The sensitive nature of medical data is a challenge for the use of centralized Machine Learning models. In contrast to traditional ML, Federated Learning allows models to be trained across institutions without sharing data. Therefore, this article presents a comparative analysis of the use of a traditional ML model for the segmentation of medical images compared to the FL paradigm, highlighting its benefits in the development of collaborative models.*

Resumo. *A natureza sensível dos dados médicos é um desafio para a utilização de modelo centralizados de Aprendizado de Máquina (Machine Learning - ML). Em contraste com o ML tradicional, o Aprendizado Federado (Federated Learning - FL) permite que modelos sejam treinados entre instituições sem o compartilhamento de dados. Dessa forma, este artigo apresenta uma análise comparativa do uso de um modelo tradicional de ML para a segmentação de imagens médicas em comparação com o paradigma de FL, destacando seus benefícios no desenvolvimento de modelos colaborativos.*

1. Introdução

Nos últimos anos, os dados digitais de saúde cresceram significativamente. Ao mesmo tempo, avanços recentes no campo do Aprendizado de Máquina (*Machine Learning* - ML) têm sido usados em uma variedade de processos que utilizam dados médicos, incluindo diagnóstico automático de doenças, classificação, análise de dados e segmentação de imagens médicas [Joshi et al. 2022].

No entanto, à medida que a demanda por modelos mais complexos e precisos cresce, surgem desafios significativos relacionados à privacidade, segurança e eficiência no treinamento desses modelos. O Aprendizado Federado (*Federated Learning* - FL) é um paradigma que permite o treinamento de modelos de ML em um conjunto de dados distribuído em *data centers*, hospitais, laboratórios de pesquisa clínica e dispositivos móveis sem compartilhamento de dados entre estas entidades. Nesta abordagem, somente os parâmetros dos modelos treinados localmente são compartilhados com o servidor agregador [McMahan et al. 2023].

O FL têm o potencial de revolucionar a forma como os modelos de ML são treinados, sendo especialmente benéfico na área da medicina. O FL é composto por duas entidades principais: os clientes (participantes) que são os proprietários dos dados e um servidor agregador. Inicialmente, um modelo global é compartilhado para um conjunto de clientes. Cada participante treina um modelo local usando seu subconjunto de dados locais e, em seguida, envia apenas os parâmetros do modelo local (ou seja, os pesos do modelo) para o servidor agregador. Os modelos locais recebidos são agregados para criar um modelo global que é transmitido para os clientes [Li et al. 2021].

Conforme apresentado por [Joshi et al. 2022], espera-se que os serviços de saúde sejam aprimorados pelas capacidades de ML e FL, melhorando a qualidade de vida dos pacientes e reduzindo as hospitalizações. Neste contexto, este trabalho realiza uma análise comparativa do uso de um modelo tradicional de ML para a segmentação de imagens médicas em comparação com o paradigma de FL, apresentando as curvas de aprendizagem, desafios de treinamento, bem como, as vantagens e desvantagens inerentes ao treinamento distribuído. O principal contribuição deste trabalho é demonstrar a viabilidade de treinar um modelo de ML no contexto da área médica, usando um método de preservação de privacidade que não requer troca de dados entre centros nem armazenamento centralizado de dados.

Para além desta seção introdutória, o restante deste artigo está organizado em seções, conforme descrito a seguir. A Seção 2 apresenta os trabalhos relacionados. A Seção 3 aborda as diferenças entre o ML centralizado e o FL. A Seção 4 descreve os elementos que compõem o cenário experimental. A Seção 5 apresenta a análise dos resultados obtidos. Por fim, a Seção 6 apresenta as considerações finais.

2. Trabalhos Relacionados

A aplicação de ML de imagens biomédicas tem visto avanços recentes significativos. Em [Ouyang et al. 2019], os autores apresentamos o conjunto de dados *EchoNet-Dynamic* de 10.036 vídeos de ecocardiografia, abrangendo a gama de condições típicas de imagem de laboratório de ecocardiografia. Além disso, os autores apresentam o desempenho de três arquiteturas convolucionais 3D para avaliar a função cardíaca.

Os trabalhos que utilizam o FL no setor saúde são diversos, abordando o desenvolvimento de aplicações de prognóstico, diagnóstico e fluxo de trabalho clínico. Em [Joshi et al. 2022], os autores apresentam estudos sobre FL no setor de saúde em uma variedade de casos de uso e aplicações. Além do mais, os autores apresentam os desafios, métodos e aplicações do FL para os setores de saúde.

A utilização de FL têm sido utilizada para detectar anormalidades tomográficas relacionadas à COVID-19. Em [Dou et al. 2021], os autores exploram algoritmos de FL para desenvolver um modelo de ML que preserva a privacidade para diagnóstico de imagens médicas. Os autores afirmam que o FL pode fornecer um mecanismo eficaz durante pandemias para desenvolver rapidamente modelo de ML clinicamente útil, superando o fardo da agregação central de grandes quantidades de dados sensíveis. Do mesmo modo, em [Flores et al. 2021] os autores aplicaram o paradigma FL para facilitar uma colaboração entre centros de pesquisa sem troca de dados, resultando em um modelo de ML que respondeu aos desafios da COVID-19.

O FL também tem sido utilizado devido às regulamentações de privacidade de

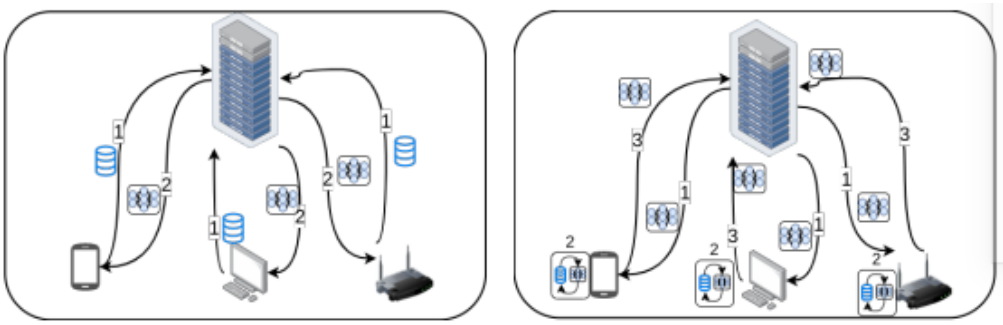
dados médicos. Em [Li et al. 2019], os autores utilizam o FL para a segmentação federada de tumor cerebral com preservação de privacidade, aplicando técnicas de privacidade diferencial para proteger os dados do paciente. De acordo com os autores, os resultados mostram que existe um compromisso entre o desempenho do modelo e os custos de proteção da privacidade.

A avaliacão dos trabalhos relacionados permitiu compreender os desafios para a implementação de métodos segmentação de imagens médicas para a implementação da proposta deste artigo.

3. Aprendizado Centralizado vs Aprendizado Federado

Conforme apresentado por [Bochie et al. 2021], a Figura 1 comparara o ML tradicional (ou centralizado) e o modelo de FL. Na etapa 1(a), observa-se o processo de coleta e envio de diversos tipos de dados, tais como imagens, vídeos e parâmetros de funcionamento, provenientes dos dispositivos de borda. Esses dados são então agregados e direcionados a uma máquina centralizada.

Na etapa seguinte, a máquina centralizada realiza a agregação e processamento dos dados recebidos, com o intuito de desenvolver um modelo de aprendizado. Esse modelo é treinado com base nas informações compiladas e, na etapa (2a), é enviado de volta aos dispositivos de borda. Essa abordagem tradicional oferece controle e coordenação centralizados, porém pode ser limitada em cenários de escalabilidade e privacidade.



(a) Cenário de aprendizado tradicional ou centralizado.

(b) Cenário de aprendizado federado.

Figura 1. Diferentes cenários de treinamento [Bochie et al. 2021]. (a) Na etapa 1 os clientes enviam os dados ao servidor e na etapa 2 o servidor devolve os modelos treinados para os clientes. (b) Na etapa 1 o servidor inicializa e distribui um modelo de aprendizado, na etapa 2 os clientes treinam o modelo com seus dados locais e na etapa 3 os modelos são devolvidos ao servidor para agregação.

Em contrapartida, o modelo de FL ilustrado na Figura 1(b) apresenta uma abordagem cliente servidor distribuída e colaborativa. Nesse caso, os dispositivos de borda continuam desempenhando um papel fundamental na geração de dados, porém, em vez de enviar todos os dados e informações a uma máquina centralizada, ocorre um processo descentralizado.

No estágio inicial (1b), o modelo global é enviado pela máquina servidora aos dispositivos de borda, que são comumente referidos como clientes. Nessa etapa, é estabelecida uma conexão colaborativa entre o servidor e os dispositivos de borda. Ao receber

o modelo global, os dispositivos de borda se tornam participantes ativos no processo de treinamento.

Na sequência, na etapa (2b), cada dispositivo de borda utiliza o modelo global recebido para realizar treinamento local. Essa abordagem aproveita os dados armazenados nos dispositivos, eliminando a necessidade de enviar esses dados para a máquina servidora central. Isso não apenas reduz a sobrecarga na comunicação, mas também endereça preocupações relacionadas à segurança e privacidade dos dados pessoais, uma vez que os dados permanecem localizados nos dispositivos de origem.

Após o treinamento local, na etapa (3b), o modelo aprimorado é retransmitido para a máquina servidora. Nesse ponto, ocorre um processo de agregação do modelo, onde as contribuições dos diversos dispositivos de borda são combinadas de forma inteligente para criar uma versão atualizada e refinada do modelo global. Esse processo de agregação incorpora as nuances aprendidas individualmente por cada dispositivo, enriquecendo assim o conhecimento geral do modelo. Essa abordagem de treinamento federado oferece vantagens significativas em termos de privacidade e segurança, bem como, reduz o consumo de largura de banda e recursos computacionais, resultando em um sistema mais eficiente e escalável.

É fundamental destacar os requisitos e problemas inerentes aos modelos, como exemplificado pelo modelo de aprendizado centralizado. Dependendo do volume de dados, a máquina centralizada, de acordo com sua configuração, pode enfrentar um período substancial de recebimento e processamento de dados, ocasionando atrasos na preparação para o início do treinamento. Nesse intervalo, é ainda mais provável ocorrerem falhas de rede, as quais podem comprometer a continuidade do processo.

Por outro lado, no contexto do modelo federado, é necessário que os dispositivos clientes possuam poder computacional adequado para realizar o treinamento dentro do prazo estabelecido. Além disso, o treinamento federado pode ser afetado por falhas de rede, latência ou sobrecarga. Esses obstáculos têm o potencial de impactar negativamente o desenvolvimento do modelo, tanto em termos de tempo quanto de acurácia.

4. Configuração dos Experimentos

Esta seção apresenta o ambiente de simulação utilizado para a avaliação da convergência do treinamento federado com o objetivo de realizar a segmentação do ventrículo esquerdo utilizando um conjunto de dados com imagens de ecocardiografia.

4.1. Modelo de ML

Este trabalho utilizou uma arquitetura de rede neural para o treinamento centralizado com o objetivo de estabelecer resultados base para a comparação com o cenário federado. O resultado base foi obtido com um mecanismo de parada antecipada para calcular o valor ótimo de classificação das amostras centralizadas.

O modelo empregado neste estudo é uma arquitetura UNet, uma rede neural convolucional popularmente utilizada em tarefas de segmentação de imagens. A UNet se destaca por sua capacidade de capturar informações contextuais enquanto mantém a precisão nos detalhes. Consiste em um encoder (parte descendente) e um decoder (parte ascendente) conectados por um caminho de atalho, permitindo a preservação da resolução espacial.

A função de perda adotada para guiar o treinamento é a `SparseCategoricalCrossentropy`. Esta é uma escolha adequada para tarefas de segmentação, onde cada pixel é atribuído a uma classe específica. O parâmetro `from_logits=True` indica que o modelo está gerando logits, ou seja, previsões não normalizadas antes da aplicação da função softmax, o que pode melhorar a estabilidade numérica durante o treinamento. A combinação do modelo UNet com a função de perda `SparseCategoricalCrossentropy` proporciona uma abordagem robusta e eficaz para a segmentação de imagens, permitindo a criação de representações detalhadas e precisas das estruturas de interesse.

Os experimentos utilizaram diferentes tamanhos de lote, especificamente 64 e 256, para avaliar o processo de aprendizado no modelo federado. Foram realizadas 100 rodadas de treinamento, cada uma composta por 2 épocas locais. No contexto do treinamento centralizado, manteremos um tamanho de lote fixo em 256, enquanto o número de épocas de treinamento foram fixadas em 200.

Adicionalmente, foram exploradas diversas taxas de aprendizado tanto no modelo federado quanto no centralizado, a fim de avaliar seu impacto na velocidade de convergência e no desempenho final do modelo. Essa avaliação abrangente forneceram *insights* valiosos sobre os efeitos do tamanho do lote, das rodadas de treinamento e das épocas na eficiência e eficácia de ambos os paradigmas de aprendizado. Ao investigar esses fatores em diferentes cenários, foram reveladas as dinâmicas sutis do FL e centralizado, lançando luz sobre suas forças e limitações relativas em diferentes contextos.

4.2. Conjunto de Dados

Este trabalho implementou uma aplicação de segmentação do ventrículo esquerdo em imagens de ecocardiografia utilizando o conjunto de dados *EchoNet-Dynamic* da *Stanford University*. Este conjunto de dados é composto por mais de 10.000 imagens abrangendo diversas variações anatômicas e patológicas de pacientes cardíacos [Ouyang et al. 2019]. A seguir, na Figura 2 é apresentado um exemplo de amostra do conjunto de dados e a segmentação do ventrículo esquerdo.

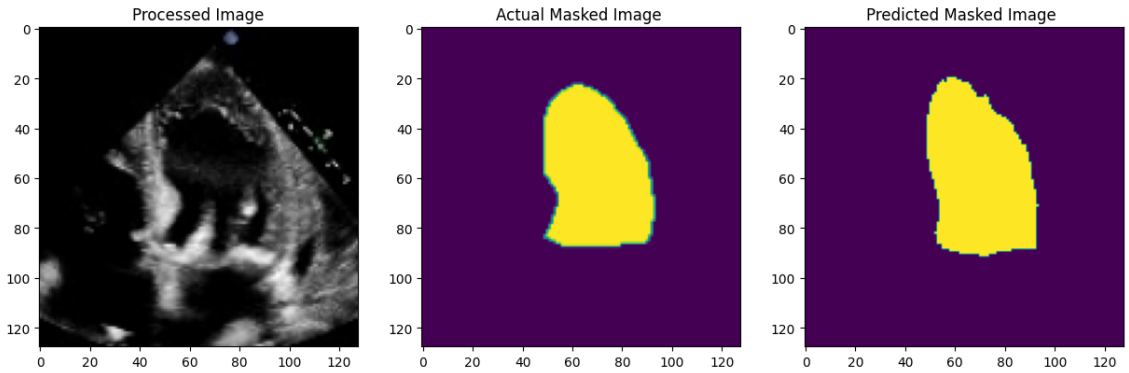


Figura 2. Exemplo de amostra do conjunto de dados e a segmentação do ventrículo esquerdo.

Para otimizar o treinamento, foram sub-amostrados $1/5$ das imagens do conjunto de dados. Essa abordagem visa garantir eficiência computacional, sem comprometer a

amplitude representativa da coleção. Posteriormente, realizou-se a seguinte estratégia de divisão da sub-amostra: 70% das imagens foram alocadas para o conjunto de treinamento, enquanto os 30% restantes compuseram o conjunto de teste. Essa divisão foi realizada de forma aleatória, assegurando que ambas as porções capturassem adequadamente a diversidade anatômica e patológica presente no conjunto completo.

Dessa forma, a metodologia não apenas equilibrou a eficiência computacional, mas também preservou a robustez e generalização da modelagem, possibilitando uma segmentação precisa e confiável do ventrículo esquerdo em imagens de ecocardiografia.

4.3. Ambiente de Simulação

Para a realização dos experimentos, utilizou-se um computador com especificações notáveis: um processador Intel Core i9-12900F operando a 2.4GHz, equipado com 16 núcleos, sendo 8 núcleos de alta performance e 8 núcleos de eficiência. Além disso, a máquina dispõe de 64 GB de RAM e uma GPU RTX 3060 com 12GB de memória.

Devido à arquitetura escolhida, a rede neural resultante abrange um total de 8643779 parâmetros. O arquivo que armazena essa rede, formatado como *Hierarchical Data Format* (HDF5), ocupou um espaço de 32.97 MB em armazenamento. Esses aspectos de configuração e arquitetura desempenham papéis cruciais nos resultados alcançados, impactando diretamente o desempenho e a eficácia das análises realizadas.

A Figura 3 apresenta o ambiente de simulação de FL deste trabalho. A orquestração do processo de FL é implementada utilizando o *framework* Flower [Beutel et al. 2020] em um ambiente baseado em contêineres Docker.

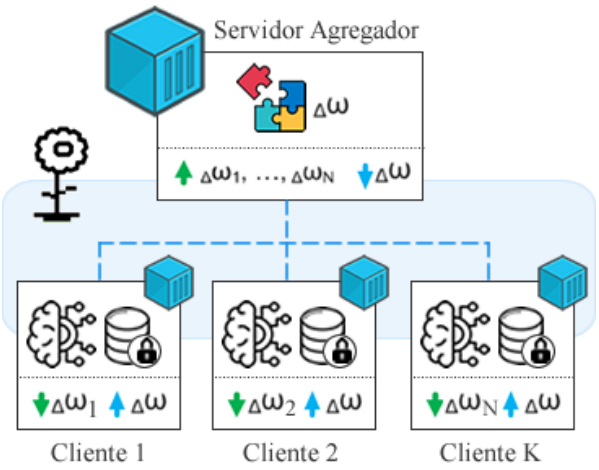


Figura 3. Ambiente de simulação de FL.

O Flower oferece uma implementação estável dos principais componentes de um sistema de FL, independente de linguagem de programação e de estrutura de ML. A utilização de contêineres permite o isolamento do servidor agregador e dos clientes participantes do processo de FL.

5. Resultados

Esta seção apresenta os resultados das tarefas de segmentação do ventrículo esquerdo utilizando um conjunto de dados com imagens de ecocardiografia. Para avaliar a con-

vergência do algoritmo de FL, monitorou-se cuidadosamente o progresso das métricas de classificação durante as iterações. A partir da análise das curvas de aprendizado, buscou-se sinais de estabilidade e consistência ao longo do treinamento. Ao comparar o desempenho do algoritmo distribuído com o desempenho do algoritmo centralizado, foi possível determinar se o desempenho do treinamento federado atingiu resultados satisfatórios.

5.1. Treinamento Centralizado

Esta seção apresenta o treinamento centralizado como forma de estabelecer um ponto de referência para avaliar qualitativamente a convergência da tarefa de FL para segmentação do ventrículo esquerdo. Após a calibração dos hiperparâmetros, o algoritmo centralizado foi treinado por 200 épocas de treinamento. Conforme pode ser observado na Figura 4, o treinamento centralizado atingiu uma acurácia de 98%.

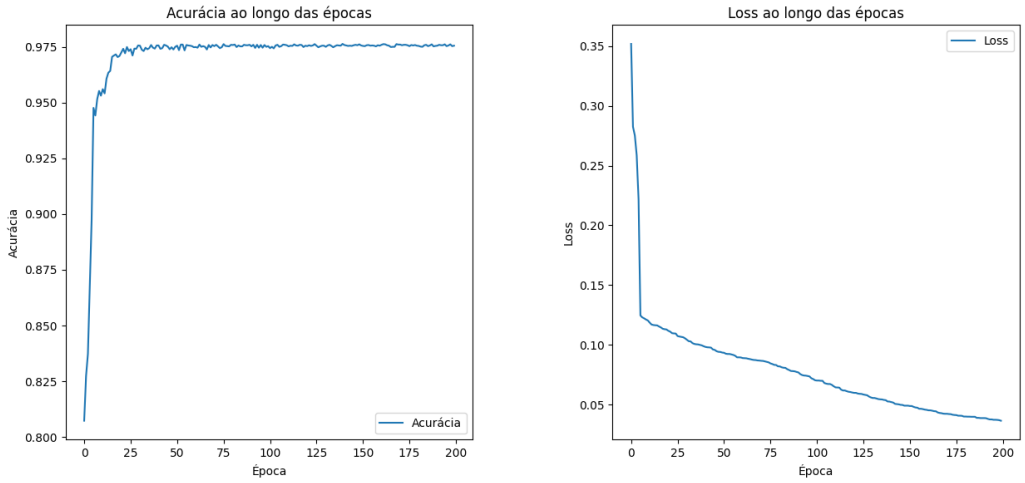


Figura 4. Curvas da acurácia e função de perda do treinamento do modelo centralizado.

Analisando a evolução da acurácia em cada época de treinamento centralizada, nota-se que por volta da época 25, o gráfico da acurácia já convergiu para aproximadamente 98%. No entanto, é observa-se que a função de perda continua a decrescer a cada nova época. Este comportamento pode ser um indicativo de que o modelo ainda está refinando seus parâmetros, buscando otimizar a representação dos dados de treinamento. Portanto, apesar da acurácia estar próxima de seu ponto de saturação, a continuidade na redução da função de perda sugere que o modelo tem potencial para se aprimorar ainda mais em termos de generalização e precisão.

Esse fenômeno pode ser atribuído a vários fatores, como a complexidade do modelo, o tamanho do conjunto de dados ou a escolha da função de perda. Considerando tal cenário, é essencial monitorar o desempenho do modelo em um conjunto de validação para evitar possíveis problemas de sobreajuste (*overfitting*) e garantir que ele esteja capturando corretamente os padrões do dado.

5.2. Treinamento Federado

Esta seção apresenta a utilização do paradigma FL para a tarefa de segmentação do ventrículo esquerdo considerando um cenário de rede ideal. Este cenário considera que

não haverá atraso na transmissão, nem probabilidade de desconexão entre os clientes e o servidor agregador. Além disso, os clientes sempre enviarão os pesos corretamente ajustados para o servidor. Neste contexto, os clientes podem ser entendidos como hospitais ou laboratórios de pesquisa clínica.

Além disso, é importante mencionar que esse cenário ideal representa uma simplificação da realidade, uma vez que ambientes reais muitas vezes apresentam variações nas condições da rede, atrasos na transmissão e possibilidades de desconexão. Ao considerar esses fatores, os algoritmos e métodos de aprendizado precisam ser robustos o suficiente para lidar com tais desafios e garantir um desempenho eficaz mesmo em cenários menos ideais.

Conforme apresentado na Figura 5, a configuração do cenário federado deste trabalho utilizou um número de clientes igual a 5. Quanto ao tamanho dos lotes (*batches*), foram utilizados os valores de 64 e 256. O numero de rodadas de comunicação foi fixado em 100 com duas épocas de treinamento local no dispositivo a cada rodada.

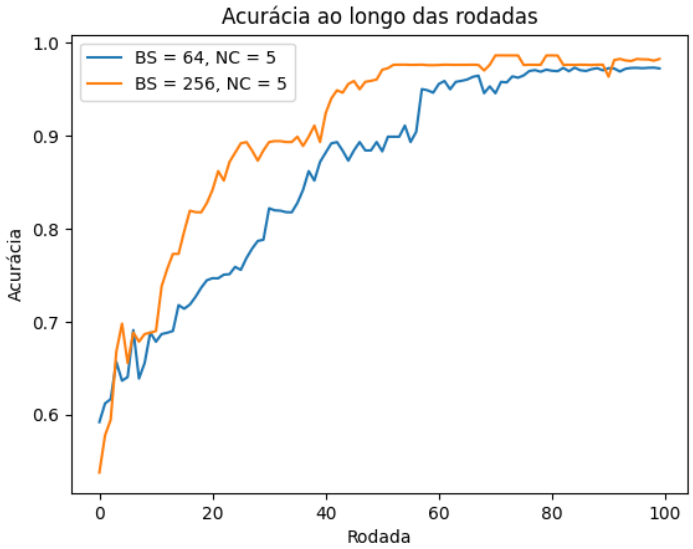


Figura 5. Acurácia do FL variando ao longo das rodadas de comunicação, onde BS representa o tamanho do lote e NC é o número de clientes.

Analisando o gráfico da Figura 5, é possível observar que à medida que aumenta o tamanho do lote (*batch*), a curva de crescimento também aumenta em comparação com os treinamentos utilizando um tamanho de lote menor. Esse fenômeno indica que o aumento do tamanho do lote está correlacionado a um aumento mais acentuado na curva de crescimento. Isso sugere que o uso de lotes maiores pode levar a um aprendizado mais eficaz e rápido nos modelos de treinamento. No entanto, é importante considerar que o aumento do tamanho do lote também pode implicar em maior consumo de recursos computacionais e possíveis desafios de convergência.

Portanto, a escolha do tamanho do lote deve ser cuidadosamente ponderada para otimizar o desempenho do treinamento e a utilização de recursos. No entanto, como demonstrado em outros artigos do mesmo gênero, nem sempre adotar um tamanho de

lote grande resulta em uma convergência eficaz devido a possível heterogeneidade dos dados dos clientes. Dessa forma, torna-se necessário encontrar um limiar adequado entre o número de usuários participantes e o tamanho do lote utilizado para alcançar uma convergência otimizada.

Essa busca pelo equilíbrio ideal entre esses dois parâmetros é crucial para garantir uma convergência eficiente do processo de aprendizado federado. Conforme apontado por [Bochie et al. 2021], diversos fatores podem influenciar essa relação, incluindo a heterogeneidade dos dados dos clientes, a capacidade computacional disponível e a dinâmica da rede. Portanto, ao projetar uma configuração de aprendizado federado, é essencial considerar cuidadosamente esses elementos a fim de atingir os melhores resultados de convergência.

Ressalta-se que a capacidade de treinar modelos diretamente nos dispositivos, sem a necessidade de enviar dados sensíveis para um servidor centralizado, representa um avanço significativo para a segurança e a proteção dos dados dos usuários. No entanto, a implementação bem-sucedida do aprendizado federado em redes móveis não é isenta de desafios. A otimização desses parâmetros requer uma abordagem adaptativa, levando em consideração a heterogeneidade dos dispositivos, a disponibilidade de recursos e a qualidade da conexão de rede.

6. Considerações Finais

Este artigo apresenta explorou o potencial do FL para a realização de tarefas segmentação de imagens médicas, destacando seus benefícios no desenvolvimento de modelos colaborativos. Os experimentos foram realizado utilizando o *framework* Flower para a orquestração do processo de FL em um ambiente baseado em contêineres Docker. Dessa forma, este estudo demonstrou o potencial do FL no setor da saúde, onde um modelo de ML pode ser treinado de forma distribuída sem o compartilhamento de dados entre hospitais ou laboratórios de pesquisa clínica.

Referências

Beutel, D. J., Topal, T., Mathur, A., Qiu, X., Parcollet, T., and Lane, N. D. (2020). Flower: A Friendly Federated Learning Research Framework. *arXiv*, abs/2007.14390.

Bochie, K., Sammarco, M., Detyniecki, M., and Campista, M. (2021). Análise do Aprendizado Federado em Redes Móveis. In *Anais do XXXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 71–84, Porto Alegre, RS, Brasil. SBC.

Dou, Q., So, T. Y., Jiang, M., Liu, Q., Vardhanabhuti, V., Kaissis, G., Li, Z., Si, W., Lee, H. H. C., Yu, K., Feng, Z., Dong, L., Burian, E., Jungmann, F., Braren, R., Makowski, M., Kainz, B., Rueckert, D., Glocker, B., Yu, S. C. H., and Heng, P. A. (2021). Federated Deep Learning for Detecting covid-19 Lung Abnormalities in CT: a Privacy-Preserving Multinational Validation Study. *Digital Medicine*.

Flores, M., Dayan, I., Roth, H., Zhong, A., Harouni, A., Gentili, A., Abidin, A., Liu, A., Costa, A., Wood, B., Tsai, C., CH Wang, C. H., and Wen, Y. (2021). Federated Learning used for Predicting Outcomes in SARS-COV-2 Patients. *Nat Med*.

Joshi, M., Pal, A., and Sankarasubbu, M. (2022). Federated Learning for Healthcare Domain - Pipeline, Applications and Challenges. *ACM Trans. Comput. Healthc.*

Li, M., Soltanolkotabi, M., and Oymak, S. (2020). Gradient Descent with Early Stopping is Provably Robust to Label Noise for Overparameterized Neural Networks. In *International Conference on Artificial Intelligence and Statistics*, pages 4313–4324. PMLR.

Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., Liu, X., and He, B. (2021). A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection. *IEEE Transactions on Knowledge and Data Engineering*.

Li, W., Milletarì, F., Xu, D., Rieke, N., Hancox, J., Zhu, W., Baust, M., Cheng, Y., Ourselin, S., Cardoso, M. J., et al. (2019). Privacy-preserving Federated Brain Tumour Segmentation. In *Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 10*, pages 133–141. Springer.

Lim, W. Y. B., Luong, N. C., Hoang, D. T., Jiao, Y., Liang, Y.-C., Yang, Q., Niyato, D., and Miao, C. (2020). Federated Learning in Mobile Edge Networks: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, 22(3):2031–2063.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. (2023). Communication-Efficient Learning of Deep Networks from Decentralized Data. *arXiv*.

Ouyang, D., He, B., Ghorbani, A., Lungren, M. P., Ashley, E. A., Liang, D. H., and Zou, J. Y. (2019). EchoNet-Dynamic: a Large New Cardiac Motion Video Data Resource for Medical Machine Learning. Computer Science, Medicine.