International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 2, ISSN No. 2455-2143, Pages 189-193
Published Online June 2022 in IJEAST (http://www.ijeast.com)

UNDERSTANDING FEDERATED LEARNING AND ITS APPLICATION IN VIDEO ANOMALY DETECTION

Aniket Singh, Zaifa Khan, Sathvik KP, Srinidhi H
Department of CSE
Ramaiah Institute of Technology (RIT)
Bengaluru

***Abstract*—Federated Learning is a relatively newer approach, and it tries to solve problems such as reluctance of clients to share data. Instead, a model is sent from server to client(s) where training happens on various edge clients who later update the central model. This is a redesign of the traditional approach of learning from data where all data had to be uploaded to a central server. In contrast, federated learning promotes the idea of bringing code to data which can help overcome latency in time-critical operations. Privacy gives every individual a right to prevent release of data, and several machine learning use-cases deal with such sensitive information for e.g., disease prediction (patient data), anomaly detection (security footage) etc. By introducing the use of federated learning, we can overcome the issue of data sharing and all data will be retained at the edge nodes where the training happens. Through this work we walk through a system design of applying federated learning in a real-time use case of video anomaly detection and present our preliminary results where we achieved an accuracy of 83.34% using the federated paradigm, higher than the centralized accuracy of 81.29%.**

***Keywords*—Federated Learning, Federated Averaging, Anomaly Detection, Autoencoders, Convolutional Neural Networks, Edge Computing.**

I. INTRODUCTION

Federated Learning is a distributed Machine Learning (ML) approach for training models on large, decentralized data. The idea behind FL is to leverage the compute power becoming increasingly ubiquitous at the edge. Traditionally, in distributed scenarios of ML we utilize distributed storage of data using distributed file systems and big data stores and for the purpose of training load the data centrally on a computationally powerful server, where the training proceeds. However, in this approach there is a bulky movement of data across devices which increases the overall runtime and latency along with limiting the applications of ML as several use-cases deal with sensitive personal information (PI) which cannot be openly shared. To overcome these challenges from a security, latency, and use-case point of view, we explore the utilization of federated learning approaches in our real-time ML solvable problems.

Federated Learning expects no data sharing between the edge clients and the central server. While the agents in the system remain the same including edge devices like pi devices, personal computers, mobile phones, fog devices, and central servers, the approach for applying ML is reengineered. In FL approach the edge devices retain the right to hold and protect their data and train models locally only communicating with the central server the updates generated to the model weights per iteration. These weights are combined at the server with different algorithms and recirculated in the system. We will take a deeper dive into the system design in upcoming sections. Since FL provides privacy, we can use it to solve the problem of video anomaly detection which consists of sensitive data. While referencing existing work and literature we saw that FL being up and coming, has not been applied to any video use-cases.

Video Anomaly Detection is the problem of identifying any out of the ordinary behaviour in surveillance footage to raise alarms and provide strong security systems. Hence, we apply FL to train our deep learning model which utilizes auto encoders to perform anomaly detection and train it is using FL to measure its effects on model accuracy. Our aim is to fulfil the following objectives:

- Understand the paradigm of FL and provide a system design which can be utilized for real-time ML and DL applications
- To test the performance of video anomaly detection models in the FL paradigm and compare metrics with centralized training.
- To boost the use of FL in real-time scenarios where privacy, latency and security are major concerns.

The rest of this paper is organized as follows; Section II explores the related works in the field, followed by Methodology in Section III, Results in Section IV, and Conclusion in Section V.

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 2, ISSN No. 2455-2143, Pages 189-193
Published Online June 2022 in IJEAST (http://www.ijeast.com)

II. RELATED WORKS

[1] The primary work of focus which we surveyed, giving insight about the design aspect of federation algorithms, design aspect of federated systems and about the drawbacks and limitations. They have built a scalable and performant Federated Learning system for mobile devices using Tensor Flow to solve problems like keyword-based search, next-word predictions etc. They also presented a High-level design for FL on mobile phones based on device configuration along with server configuration. The main drawbacks and future scope were stated which included changing the effect of bias during selection, finding the optimal convergence time between training and higher accuracy, introducing newer scheduling algorithms, and optimizing bandwidth consumption.

[2] Introduced Federated Learning as a method to perform decentralized machine learning while allowing sensitive client data to reside on the client device itself and perform local model training on these client devices. These local updates can be averaged over several iterations to finally aggregate the global model for all devices. Thus, this is the main federation algorithm. They also presented a Federated Averaging algorithm which can handle deep networks and performs weighted averaging based on the number of examples available on the client w.r.t. the overall data points. A 10-100 times reduction in communication rounds is shown as compared to centralize SGD.

[3] This paper provided a new framework called flower for user friendly development of federated systems. To ease research and real-time deployment of federated learning containing a heterogeneous setup of clients, authors proposed the development of a new framework Flower. The purpose of the work was to allow easy changing of algorithms which perform federated aggregation along with allowing developers to scale to many clients without the challenges involved in underlying communication protocols. The framework had five state-of-the-art federated algorithms implemented along with providing the ability to fit in custom algorithms. They also tested it on various devices of various capacities. The framework is tested on several sets of devices like CPU, GPU, Raspberry Pi and NVIDIA Jetson and is performant in all cases. The drawbacks were not testing on sufficient data. There are limited datasets which have been distributed over clients and each distribution can cause a difference in the overall outcome of training hence more such datasets must be generated or tested. Testing in newer areas remains as a future scope for ex in the case of multi-modal data such as audio, video etc.

[4] The problem of limited abilities of humans to be at par with the high influx of video data is recognized. And the excess of labour invested in monitoring is mostly idle and unused due to the rare occurrences of any unusual pattern or behaviour. Hence a deep learning approach to detect anomalies in real-time with the intent of public safety was developed. The data used for training was weakly labelled i.e., only the video-level classification was done to indicate anomalous or not, and the attempt is to generate a clip or segment from the video which holds the actual occurrence of anomaly. Besides working on the deep learning architecture and extensive video surveillance dataset was introduced containing 128 hours of videos.

[5] Identifying the most important aspects of privacy and implications of using federated learning to solve the problem of data exchange, the authors have laid down the definitions for federated learning and provided analysis for the different components existing in the system. Authors have identified that FL approaches can be categorized using underlying data distribution, machine learning, scale of design, privacy, communication, and inspiration for federation. Three main system components identified were parties, manager, and communication-computation framework. Fate, TFF, PySyft, PaddleFL and FedML were reviewed as FL frameworks. Possible fields of exploration were identified as dynamic scheduling, IoT, Healthcare etc. where use of FL could create significant impact.

[6] The variety of data available in video format often comes with the challenges of being unlabelled. Identifying this the authors compiled the deep learning models available for anomaly detection, with classification based on type of model. Both Convolutional models for feature extraction and prediction based generative models were studied. The datasets that were noted to be of importance included UCSD Dataset of pedestrian footage, UMN Dataset showing out of the normal crowded scenes, and the CUHK Avenue Dataset related to strange activities like throwing bags etc.

III. METHODOLOGY

The work was carried out in keeping with the three important aspects of study which were: understanding the system implications of federated learning, performing preliminary experiments on image datasets to understand and generate baseline results, and finally performing a federated training of a deep learning model which is built for video anomaly detection. The first aspect involves the understanding of the federated learning system and implementation techniques utilized for the purpose of this study. The system design of federated learning can be broken down into three major layers of operation: Edge layer, Network layer, and Server layer. The edge layer comprises the client devices which are responsible for data collection and perform training on the collected data. These devices can be personal mobiles, personal computers, embedded devices, pi devices etc. For the use-case of video anomaly detection the ideal edge device will be an IoT enabled CCTV camera with compute abilities. The second layer involves the network layer over which communication of model training is carried out. Network layer facilitates transfer of updated model weights between client devices and server. The last layer is the server where the server resides, responsible for coordinating all the connections

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 2, ISSN No. 2455-2143, Pages 189-193
Published Online June 2022 in IJEAST (http://www.ijeast.com)

and communications and performing federated combinations of distributed weights received.

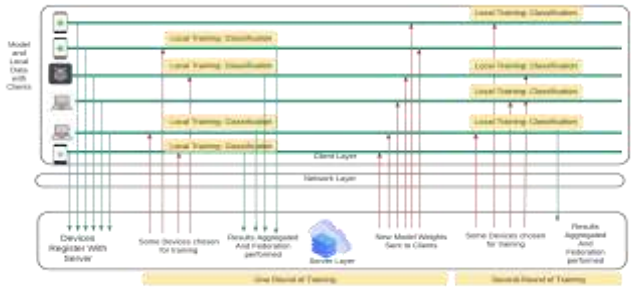

Fig. 1. Federated Learning Setup

Figure 1 shows the major system components and overview of the training process, where federated learning proceeds in the form of rounds. After performing a thorough review of the existing literature and framework available to perform experiments, we proceeded with the flower framework.

The second aspect was to perform a preliminary set of experiments using the mentioned federated setup and measure different aspects of training such as accuracy and time. These experiments dealt with image dataset CIFAR-10 which is a standard dataset frequently used for benchmarking performance. Since the final goal was to perform federation on video data, image data was a first step in achieving the same. The dataset comprises 10 classes of images with 50,000 training and 10,000 testing images. A neural network with 3 Conv2D layers, 2 Max Pooling layers, 2 Dense layers and 1 Flatten layer was trained over a varying set of configurations. The training was performed for up to 5 clients.

The final aspect was to apply federation to the video anomaly detection model. The anomaly detection was performed using a C3D based feature extractor and an autoencoder classifying model. The C3D feature extractor provided a 4096-dimension output vector as a set of spatio-temporal features captured from the video frames. The input videos were split into 32 clips at the rate of 16fps. The feature vector was then passed to the three-layer autoencoder which consisted of 512, 32 and 1 neurons, respectively. The dataset used was UCF Crime Dataset. In our experiments we have used 1610 videos (800 normal class, 810 anomalous class). Figure 2 covers the flow of proposed work.

Fig. 2. System and Model Development Methodology

From the above set of experiments, we established that federation can successfully be applied to multi-modal data including images and videos. Video anomaly detection is a crucial use-case. Federation solves the problem of data sharing and privacy.

IV. RESULTS

Preliminary results from the image classification dataset were carried out over 2-5 client setups. For comparison and analysis centralized training was also performed. The training data images (50,000) and testing data images (10,000) were split across all the clients on an almost equivalent basis. The neural network was trained, and the timing and accuracy plots were generated. From the timing plot (Figure 3) it was observed, increasing number of clients reduces training time. Figure 4 shows comparable accuracies reported by each client. The configurations for training of video anomaly detection with federated setup have been enlisted in Table 1. Data Distribution for video anomaly detection was done in a non-uniform way unlike the image dataset to simulate closeness to real-world scenarios.The distribution has been plotted in Figure 5 (roughly 1:9 distribution ratio) and Figure 6, reports the accuracy obtained which indicates that overcoming the data disparity the accuracies obtained are comparable and even better than centralized training. These results have been summarized in Table 2.

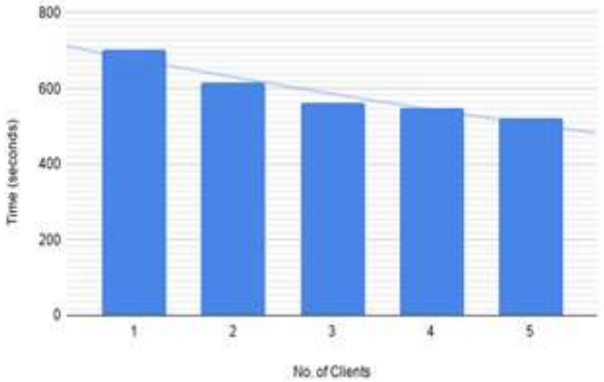

Fig. 3. Timing Plot (Image Dataset)

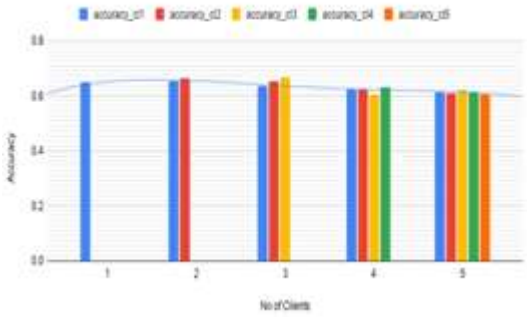

Fig. 4. Accuracy Plot (Image Dataset)

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 2, ISSN No. 2455-2143, Pages 189-193
Published Online June 2022 in IJEAST (http://www.ijeast.com)

Table-1 Experiment Setup for Video Anomaly Detection with Federation

Feature	Value
No of Anomaly Classes	4 (Assault, Accident, Arson, Burglary)
Optimizer	Adagrad
Optimizer Learning Rate	0.001
Dropout Regularization	60%
No. of Clients	Up to 2
Centralized Epochs	1000
Federated Epochs	5 Rounds, 100 epochs

IV. CONCLUSION

The work aimed at providing a stronger understanding of the system agents and their communications in the process of federated learning and establishes that federation can be used successfully in the field of video anomaly detection or other such real-time use-cases. As a part of this work, we achieved successful classification of normal and anomalous surveillance footage while applying a federated deep learning model and reported accuracies consistent with the centralized training agents.

Fig. 5. Data Distribution (Video Dataset)

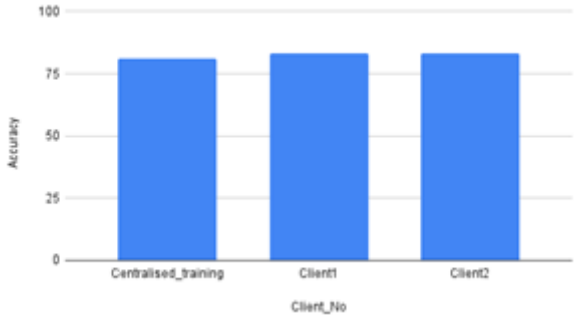

Fig. 6. Accuracy Pot (Video Dataset)

Additionally, the issue of privacy and uneven data distribution of clients was successfully tackled. As an extension of this work, we wish to explore different configurations of the federated setup which involves introducing new client device types such as embedded devices and IoT enabled devices. Also, newer strategies for federated combination of model weights, more settings for rounds and epochs, and focus on encryption of transmitted model weights can be future points of focus.

Table-2 Result Summary

Parameters	Baseline Model	Proposed Model
Configuration	Centralized Deep Learning model	Deep Learning Model + Federated Learning Model
Accuracy	81.29%	83.34%
Time taken	5-6 mins per video	5-7 mins per video
No. of videos	Complete dataset	Distribution of data in 1:9 but accuracy remains unaffected

V. REFERENCE

[1] Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., Kiddon, C., Konečný, J., Mazzocchi, S., McMahan, B. and Van Overveldt, T., 2019. Towards federated learning at scale: System design. Proceedings of Machine Learning and Systems, 1, pp.374-388.

[2] McMahan, B., Moore, E., Ramage, D., Hampson, S. and y Arcas, B.A., 2017, April. Communication-efficient learning of deep networks from decentralized data. In Artificial intelligence and statistics (pp. 1273-1282). PMLR.

[3] Beutel, D.J., Topal, T., Mathur, A., Qiu, X., Parcollet, T., de Gusmão, P.P. and Lane, N.D., 2020. Flower: A friendly federated learning research framework. arXiv preprint arXiv:2007.14390.

[4] Sultani, W., Chen, C. and Shah, M., 2018. Real-world anomaly detection in surveillance videos. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6479-6488).

[5] Yang, Q., Liu, Y., Chen, T. and Tong, Y., 2019. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), pp.1-19.

[6] Kiran, B.R., Thomas, D.M. and Parakkal, R., 2018. An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. Journal of Imaging, 4(2), p.36.

[7] Li, Qinbin, Zeyi Wen, and Bingsheng He. "Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection." (2019).

[8] Li, Q., Wen, Z. and He, B., 2019. Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection.

[9] Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R. and D'Oliveira, R.G., 2021. Advances and open problems in federated learning. Foundations and Trends® in Machine Learning, 14(1–2), pp.1-210.

[10] Li, T., Sahu, A.K., Zaheer, M., Sanjabi, M., Talwalkar, A. and Smith, V., 2020. Federated optimization in heterogeneous networks. Proceedings of Machine Learning and Systems, 2, pp.429-450.

[11] Chai, Z., Ali, A., Zawad, S., Truex, S., Anwar, A., Baracaldo, N., Zhou, Y., Ludwig, H., Yan, F. and Cheng, Y., 2020, June. Tifl: A tier-based federated learning system. In Proceedings of the 29th International Symposium on High-Performance Parallel and Distributed Computing (pp. 125-136).

[12] Kulkarni, V., Kulkarni, M. and Pant, A., 2020, July. Survey of personalization techniques for federated learning. In 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4) (pp. 794-797). IEEE.

[13] Duman, E. and Erdem, O.A., 2019. Anomaly detection in videos using optical flow and convolutional autoencoder. IEEE Access, 7, pp.183914-183923.

[14] Li, H., Achim, A. and Bull, D., 2012. Unsupervised video anomaly detection using feature clustering. IET signal processing, 6(5), pp.521-533.

[15] Blair, C.G. and Robertson, N.M., 2015. Video anomaly detection in real time on a power-aware heterogeneous platform. IEEE Transactions on Circuits and Systems for Video Technology, 26(11), pp.2109-2122.