Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter and M. D. Rossetti, eds.

SIMULATING FEDERATED LEARNING WITH DATA AUGMENTATION FOR CULVERT CONDITION PREDICTION IN UTAH: A CASE STUDY

Pouria Mohammadi[1], Abbas Rashidi[1], and Sadegh Asgari[2]

[1]Dept. of Civil and Environmental Eng., University of Utah, Salt Lake City, UT, USA
[2]Dept. of Civil and Environmental Eng., Merrimack College, North Andover, MA, USA

ABSTRACT

Transportation agencies are increasingly adopting cutting-edge solutions to enhance infrastructure management strategies. Infrastructure condition prediction models, particularly important for optimizing inspection resources, are at the forefront of this trend. Machine learning (ML) algorithms play a crucial role in these models, but traditional centralized ML models often struggle with data scarcity, privacy, and transferability. This paper introduces a decentralized approach using Federated Learning (FL) for infrastructure condition prediction, simulating its effectiveness with culvert inventories in Utah and five other states in the United States. We analyzed two FL models—one with data augmentation and one without—against a traditional centralized model. Our results demonstrated that FL-based models improved prediction accuracy by 30%, ensured data privacy, and reduced data transmission overheads. Moreover, due to its limited data, Utah benefited from federated insights, illustrating FL's potential to effectively enhance infrastructure management. The simulation highlights the advantages of utilizing FL in real-world scenarios.

1 INTRODUCTION

In recent years, the convergence of artificial intelligence (AI) and infrastructure asset management has led to a transformative approach to maintaining and inspecting critical infrastructures. Transportation agencies are adopting machine learning (ML), a branch of AI, at an increasing rate because it empowers these agencies to make smarter and more effective decisions regarding the maintenance and operation of infrastructure systems (Iyer 2021). One of the key applications of ML in this sector is predictive modeling, which allows agencies to predict and address potential infrastructure issues before they lead to system failures. This proactive strategy minimizes the costs associated with repairs and downtime and significantly reduces the risk of catastrophic failures. Consequently, this proactive strategy contributes to extending the lifespan of transportation infrastructures, ensuring that the components of the transportation systems are more durable and reliable (Mohammadi et al. 2023a). Despite the potential of ML for enhancing transportation infrastructure management, several significant obstacles prevent its widespread implementation. One of the primary challenges agencies face is the scarcity of comprehensive data inventory. The availability of robust and detailed local datasets is crucial for effectively applying ML techniques (Sinha et al. 2017).

The lack of comprehensive datasets poses several challenges for ML models in infrastructure management. First, it can lead to overfitting. In this scenario, the model memorizes the training data too well, leading to poor performance on unseen data. Second, limited data can restrict the diversity of the dataset. This can introduce biases into the model's outputs, making it unfair and less effective in handling the full range of potential infrastructure conditions. Finally, leveraging complex ML algorithms with strong prediction performance, which often require a substantial amount of data for training, becomes impractical when data is scarce (Theodorou et al. 2023). Several factors impede the accessibility of comprehensive local infrastructure datasets. These limitations are due to resource constraints, such as limited budgets and a shortage of inspectors, which can restrict data collection efforts. Additionally, the absence of a centralized management system hinders systematic data collection and consolidation. Furthermore, many transportation systems were not originally engineered to incorporate modern data collection technologies, leading to significant gaps in historical data. Several factors impede the accessibility of comprehensive local infrastructure datasets. These limitations are due to resource constraints, such as limited budgets and a shortage of inspectors, which can restrict data collection efforts. Additionally, the absence of a centralized management system hinders systematic data collection and consolidation (Sinha et al. 2017). Utah's culvert inventory is an example of this data scarcity. The absence of a comprehensive culvert management system has resulted in a limited number of inspection records for these critical infrastructure assets (McGrath and Beaver 2004).

To effectively utilize ML in transportation infrastructure management, tackling the issue of data scarcity is significant. There are a variety of strategic approaches, including improving data collection approaches, applying advanced analytical methods to available data, and promoting cooperation between transportation agencies to facilitate the sharing of data. A useful and commonly adopted strategy among transportation agencies involves acquiring and merging inventory data from other agencies into their own datasets (Mohammadi et al. 2024a). Such collaborative data sharing can substantially enrich the datasets available for ML models, offering more details on transportation infrastructures' needs and usage patterns. However, the sharing of data between transportation agencies comes with several challenges. First, standardized data formats and protocols are necessary to ensure seamless integration and interpretation of shared data. Second, a reluctance to share data among transportation agencies (Figueiredo 2017), which can be attributed to competitive interests or bureaucratic barriers. Last and foremost, there are legal concerns related to privacy regulations. General Data Protection Regulation (GDPR 2018) and the California Consumer Privacy Act (CCPA 2019) are two regulations that enforce strict controls and limitations on data sharing. This regulatory scheme leads to a cautious approach to raw data sharing, often resulting in the privatizing of raw data and limiting its availability.

Researchers have explored creative approaches to overcome the challenges associated with infrastructure data scarcity and barriers to data sharing. A notable advancement in this field was the introduction of federated learning (FL) by Google in 2016 (Konečný et al. 2016), a cutting-edge ML approach that simultaneously safeguards data privacy and mitigates issues related to data scarcity. At first, developed to boost text prediction on a wide range of Android devices, while also adhering to strict privacy regulations like the GDPR and CCPA. In FL, a network of decentralized entities—which can be mobile devices or local servers belonging to various organizations—is managed by a central server. These entities work together in the training phase. This methodology allows for the collaborative training of a model without the need to share raw data among participants. The central server provides the initial model for each entity, which is subsequently trained locally. Only the model updates are then sent back to the server. By aggregating these updates, the central server improves the overall efficacy of the model. Therefore, this strategy successfully preserves data privacy and utilizes the collective knowledge of all involved parties (McMahan et al. 2017).

Previous research has underscored the effectiveness of FL across various domains, highlighting its robust performance. However, its implementation in environments characterized by scarce transportation infrastructure data is still largely untapped. Implementing the FL model in a real-world scenario requires significant investment in money and resources. Therefore, we first need to simulate it on a single server to demonstrate its advantages. Once validated, we can expand by connecting several states and applying it to address the data limitations faced by DOTs. Our study focuses on exploring the utility of FL in predictive analytics, especially in contexts where infrastructure historical data are sparse. We selected the culvert inventory of Utah, which comprises only 272 records, as our case study. This choice allows us to evaluate FL's capability to overcome challenges faced by the Utah Department of Transportation (UDOT) and other similar transportation agencies. In our research, we proposed an FL framework integrated with data augmentation techniques that enables UDOT to collaboratively use data from other state DOTs while ensuring the privacy and security of the data. To accomplish this, our study consisted of a thorough literature review, collecting and preprocessing of historical culvert inspection data, and simulating two developed FL models - one with and one without data augmentation. We subsequently assessed the performance of these decentralized FL models in comparison to a traditional local model that relies solely on Utah's limited dataset of 272 culvert records. Through FL simulation, we aimed to offer UDOT a more efficient model for evaluating culvert conditions and organizing maintenance tasks. This will ultimately improve the accuracy and effectiveness of infrastructure management.

2 LITERATURE REVIEW

FL has rapidly gained popularity among researchers across various disciplines seeking to address data privacy, data scarcity, and communication overhead challenges. Civil engineering is one of the many fields that are starting to understand the advantages of FL. In this section, we begin with reviewing the research conducted on the implementation of FL in civil engineering. Next, we explore the latest predictive modeling techniques focused on enhancing the effectiveness of culvert inspection programs.

2.1 FL in Civil Engineering

FL has the potential to bring about major improvements in the field of civil engineering. It can develop ML models that are not only more effective and accurate, but also prioritize data confidentiality. For instance, Saputra et al. (2019)introduced an FL method for forecasting energy demand within electric vehicle networks. Their approach, Federated Energy Demand Learning (FEDL), significantly improved the accuracy of energy demand forecasts in a region. The algorithm, implemented by a charging station provider (CSP), demonstrated high precision in estimating energy demand across multiple charging stations. Notably, FEDL successfully reduced communication overhead and secured the privacy of electric vehicle users by enabling the sharing of model updates rather than raw data with the CSP. This innovative method outperformed traditional ML algorithms in efficacy. In another study, Li et al. (2021) utilized FL to tackle privacy and security concerns related to gathering personal image data of construction workers for occupational health and safety (OHS) status monitoring. They presented a new framework called federated smart work packaging (FedSWP), which was developed based on federated transfer learning to protect the personal image data of construction workers. This framework was built using a hybrid deep neural network that included Multi-Task Cascaded Convolutional Neural Networks, MobileNet, and Long Short-Term Memory. The test of FedSWP in monitoring facial fatigue among crane operators showcased its effectiveness in providing personalized safety alerts and healthcare advice. According to the authors, FedSWP could be applied in a variety of construction OHS monitoring scenarios.

Expanding the application of FL, Khalil et al. (2022) explored FL's utility in preserving privacy within ML operations in industrial Internet of Things (IoT) settings. Their FL algorithm, Fed-NN, which was based on a neural network, delivered accurate predictions of thermal comfort for workers while ensuring their privacy. Upon testing with a real dataset, Fed-NN achieved an 80% accuracy rate, surpassing traditional models like support vector machines and multiple linear regression. This demonstrated Fed-NN's capability to provide reliable and privacy-preserving predictions of thermal comfort in industrial IoT environments. Also, Moretti et al. (2023) created a federated open data model for digital twins (DTs) in the built environment, aimed at addressing interoperability challenges in DT applications. They built a process modeling approach with three tiers that ensures asset information standards are aligned with service expectations throughout the lifecycle of the assets. The usability and effectiveness of this model were confirmed through its implementation in developing a building-level DT data model for the West Cambridge Campus. This federated data model demonstrated its capability to enhance DT-based asset management across both individual building and built environment levels. These studies indicate the transformative potential of FL, suggesting that further advancements in this area could significantly digitalize and enhance operational efficiencies in the civil industry. Such progress could revolutionize how data-driven decisions are made, maintaining high standards of privacy and reducing operational costs across various sectors.

2.2 Predictive Modeling for Culverts

For enhancing the effectiveness of culvert inspection programs managed by transportation agencies, researchers created a number of predictive analytics models. One notable effort by Tatari et al. (2013) involved the development of a regression model using Artificial Neural Networks (ANN) to forecast the condition of culverts in Ohio. However, the model's effectiveness was limited by its dependence on a relatively small dataset from the Ohio DOT, which comprised only 39 culvert records and nine different features, resulting in a disproportionately high feature-to-data ratio. Similarly, Stoner et al. (2019) utilized ANNs and logistic regression to construct models that predict the conditions of culverts with a significantly larger dataset from the South Carolina DOT, encompassing 8,000 culvert entries. These models, which covered various types of culverts and ten distinct categories of defects, proved proficiency at multiclass classification, even without taking the age of the culverts into account. As these studies were based on Ohio state's culvert data, their applicability is somewhat limited.

Additionally, Gao and Elzarka (2021) crated a decision tree-based binary classification model focused on culvert condition assessment. This model, which incorporated 11 different features including material, shape, and span, was applied to an extensive dataset of 12,400 culverts from Ohio, achieving a classification accuracy of 75%. Nevertheless, the model's binary nature and reliance only on Ohio data were viewed as significant limitations. As part of another endeavor, Mohammadi et al. (2023) assessed the efficacy of five recognized multiclass classification ML algorithms—including random forest, decision tree, support vector machine, k-nearest neighbor, and ANN—in predicting the condition of Utah's culverts. Their research, which analyzed a dataset of 2,555 culvert records from four states, highlighted the random forest algorithm as the most effective across all evaluation metrics. However, this study also faced challenges related to data privacy due to merging data from multiple state inventories.

While these previous efforts have advanced the field of culvert condition prediction, they fell short in offering practical solutions for states like Utah, which lack robust predictive models due to limited datasets. The previous models typically relied on binary classification and were restricted to data from specific state inventories. Given the proven capabilities of FL in civil engineering, our research introduces a more advanced and generalized model employing FL. This model is designed to overcome the limitations observed in earlier studies, especially those related to data scarcity and diversity, while maintaining strict regulations for data confidentiality and privacy.

3 METHODOLOGY

Our research's main objective is to address data scarcity in the inventories of transportation agencies' infrastructure assets. We used Utah's 272 data row culvert inventory for our case study. The limited culvert inspection records within Utah's inventory represent a significant barrier to effective decision-making regarding culvert inspection and maintenance (McGrath and Beaver 2004; Mohammadi et al. 2023b). To bridge this gap, our strategy involves aggregating culvert data from the inventories of other states across the US, thereby enriching the dataset available for analysis. Data was collected from the state culvert inventories of Vermont, Ohio, New York, Massachusetts, and Colorado. Following the data collection phase, we performed a series of preprocessing steps to refine the dataset. This included outlier removal, data sampling, normalization, and feature combination to enhance data quality and consistency. Instead of merging these datasets, we retained them as distinct entities to facilitate the simulation of FL. During FL simulation, we trained and tested base ML models on these separate preprocessed datasets as a preliminary step in developing our FL models. These models must be validated and tested to ensure their reliability. Figure 1 illustrates the FL simulation we conducted in this study.

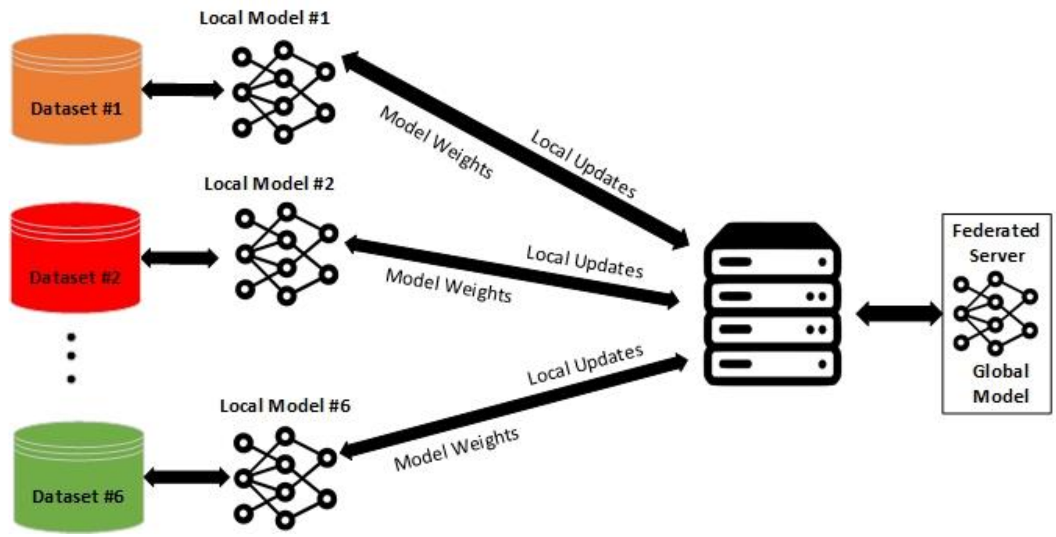


Figure 1: The FL framework for culvert condition prediction modeling.

3.1 Data Collection and Preprocessing

We collected data from five additional states to expand Utah's culvert dataset. The states included New York with 1050 data rows, Massachusetts with 417 data rows, Colorado with 766 data rows, Vermont with 3884 data rows, and Ohio with 1851 data rows. This choice was made due to the ease of accessing data and the similarities between these states' culvert inventory features and Utah's. Due to the variations in inspection criteria used in different states, it was necessary to perform a preprocessing step on the features in the datasets. In addition, the target labels were adjusted to align with UDOT's culvert rating criteria. The rating scales were adjusted to match UDOT's 5-point rating scale in these additional states. Table 1 illustrates the mapping of each state's rating scale onto Utah's 5-point scale. To ensure the conversions were accurate and reflective of comparable assessment standards, we thoroughly reviewed the inspection criteria used by these five DOTs compared with UDOT's criteria.

Table 1: Rating scales conversion to Utah's.

Utah	**5**	**4**	**3**	**2**	**1**
Colorado & Ohio	9 & 8	7	6	4 & 5	1 & 2 & 3
Vermont	7	6 & 5	4	3	2 & 1
New York	7	6 & 5	4	3	0 & 1 & 2
Massachusetts	9 & 8	7	6	5	0 & 1 & 2 & 3 & 4

To develop a more robust and generalizable ML model for application across various states, we enriched our six datasets with environmental features such as soil pH and soil moisture, obtained from the Web Soil Survey website (Web Soil Survey). As shown in Table 2, our final dataset includes ten features classified as either physical or environmental.

Table 2: Final dataset's labels and features.

Label	Culvert condition rating				
Physical Features	Culvert installation year	Culvert structure	Culvert shape	Culvert length	Culvert inspection date
Environmental Features	Soil electrical conductivity	Soil drainage class	Soil pH	Soil Moisture	Flooding frequency

3.2 Model Development

3.2.1 Artificial Neural Network

In this research, we employed ANNs to develop our predictive models. ANNs are complex ML algorithms that mimic the structure of the human brain. They consist of interconnected nodes—often referred to as "artificial neurons"—arranged in several layers, including an input layer, multiple hidden layers, and an output layer (Tatari et al. 2013). These neurons are interconnected through weighted connections that signify the influence each neuron has on the output. The training process involves adjusting these weights to minimize the difference between the predicted output and the actual data. The fundamental operation within a neuron is the aggregation of weighted inputs, which are then processed through an activation function. Using non-linear activation functions such as sigmoid, tanh, and ReLU is crucial because it allows the ANN to capture and model complex patterns in the data (Hassandokht Mashhadi et al. 2024). In our specific application, we implemented an ANN architecture that utilizes the ReLU activation and cross-entropy loss functions. This configuration was chosen because it effectively facilitates the learning of complex, non-linear relationships inherent in the dataset, enhancing the model's predictive capability.

3.2.2 Synthetic Minority Over-sampling Technique (SMOTE)

A common data augmentation technique in ML is SMOTE. It can address challenges posed by data scarcity and imbalanced classes. When datasets exhibit a significant disparity in the number of samples between classes—where the minority class has far fewer samples compared to the majority class—machine learning algorithms tend to develop a bias toward the majority class (Wijs et al. 2020). To address this, Chawla et al. (2002) introduced an approach that involves augmenting the minority class by generating "synthetic" data points, rather than just duplicating the existing ones. SMOTE generates these synthetic data points by interpolating between several instances of the minority class. Specifically, the algorithm selects k nearest neighbors in the feature space for a given minority class sample, draws lines between these neighbors and the sample, and then generates new samples along these lines. This technique not only achieves a more balanced class distribution in the training data but also enhances the diversity within the dataset. By doing so, it allows ML models to capture a more comprehensive representation of the minority class characteristics.

3.2.3 Federated Learning

FL represents a cutting-edge method in the field of distributed learning, enabling the creation of a global model by combining locally trained models from various decentralized entities. Of significance, the aggregation process does not require the direct exchange of the data used in these models. Instead, it generally aggregates the local models by averaging their parameters or gradients, which helps protect data privacy and enhances computational efficiency (Li et al. 2020b). For simulating FL, we utilized Flower (Beutel et al. 2020), an advanced open-source framework specifically developed for the efficient development and simulation of FL models. Flower's design provides substantial flexibility, allowing researchers to select from a wide range of ML algorithms and optimization techniques to best fit the specific needs of their research. For our purposes, we chose ANN as our base ML algorithm. This decision was informed by the ANN's proven capability in handling complex predictive tasks, its expertise at multiclass classification challenges, and its seamless compatibility with the Flower framework. This integration effectively leverages ANN's strengths in robust prediction and classification within the FL environment facilitated by Flower.

FL can be classified into three distinct types: horizontal, vertical, and federated transfer learning. Horizontal FL, often referred to as data parallelism, is utilized when various agents possess the same features but differ in their data samples. This type enables the agents to collaboratively train a global model while preserving the privacy of their respective data. Vertical FL, known as feature parallelism, is appropriate when agents have different features but share the same data samples. In this approach, agents exchange intermediate results to facilitate collective learning. Federated transfer learning is applied in scenarios where agents have different features and data samples. This method leverages a pre-trained model to transfer knowledge, enhancing learning across different agents (Li et al. 2020a). In our research, we opted for horizontal FL due to the similarity in features across the DOT inventories, despite the variations in their data samples.

Accordingly, we developed two FL models: one with integrated data augmentation, SMOTE, and one without. Our FL framework incorporates the Federated Averaging algorithm (FedAvg), which is widely recognized for its simplicity and effectiveness in aggregating updates from multiple models derived from different agents, as outlined in Algorithm 1. To evaluate the performance of our ML models on new, unseen data, we implemented the hold-out cross-validation technique. Typically, this method randomly splits the dataset into training and testing subsets (Gondia et al. 2020), with 80% of the data allocated for training and 20% reserved for testing. However, within the FL context, we adapted this approach: each agent independently tests the global model using their specific hold-out data. This enables the aggregation of key performance metrics, such as accuracy and F1 score, from each agent to comprehensively evaluate the global model's overall efficacy. Following is the pseudo-code of the FL framework developed for further understanding.

Algorithm 1: Federated Averaging Algorithm (FedAvg)

Input: Number of agents (N), Local datasets (D_n), the strategy of the framework, including the number of communication rounds (m), the fraction of agents on each round (s), the number of local epochs (E), learning rate (η), optimization function ($Adam$), and local mini-batch size (B).

Output: updated parameters of global ANN model (ω), ANN model's performance metrics (p)

Server Execution:

Initialize global neural network model (ω_0)

for round t = 1, 2, …, m **do**

$S_t \leftarrow$ random subset of $max(\mathrm{s} \times \mathrm{N}, 1)$ branches to participate in the training

Distribute ω_0 to all agents in S_t

for each agent $D_n \in S_t$ **in parallel do**

Initialize $\omega_t^n = \omega_0$

$\omega_{t+1}^n, p_t^n \leftarrow$ Agent_Update ($D_n, \omega_t^n, \eta, E, B, Adam$)

end

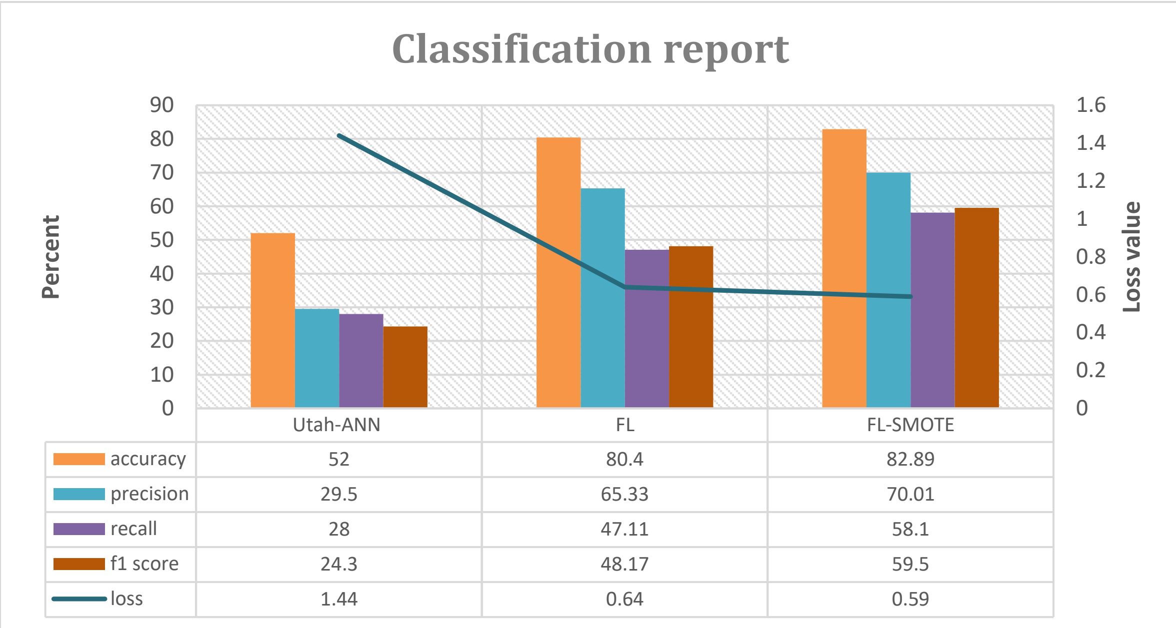
$\omega_{t+1} \leftarrow \frac{1}{|\{S_t\}|} \sum_{D_n \in S_t} \omega_{t+1}^n$

$p_t \leftarrow \sum_{D_n \in S_t} \frac{|D_n|}{\sum_{D_n \in S_t} |D_n|} \times p_t^n$

end

4 RESULTS AND DISCUSION

Challenges such as data scarcity and privacy have significantly impeded the widespread use of ML in infrastructure management. To address these challenges, we proposed the adoption of FL over traditional centralized learning methods. As a practical application, we simulated the effectiveness of FL in developing predictive models for culvert conditions, specifically focusing on the state of Utah. Given the limited size of Utah's culvert dataset, we leveraged data from five additional states using the FL approach. Accordingly, we developed and evaluated three different models: two FL models (one with SMOTE and one without SMOTE) utilizing data from six states and one centralized ANN model (Utah-ANN) based only on Utah's small dataset. The performance of these models was assessed using various metrics, including accuracy, precision, recall, F1 score, and total loss. As depicted in Figure 2, the FL-based models, FL-SMOTE and FL, demonstrated superior performance to the Utah-ANN model, exhibiting notable enhancements in precision and accuracy. The Utah-ANN model's accuracy improvement by 30 percent suggests that integrating local parameters from multiple entities into a global model significantly enhances the model's generalization and robustness.



	Utah-ANN	FL	FL-SMOTE
accuracy	52	80.4	82.89
precision	29.5	65.33	70.01
recall	28	47.11	58.1
f1 score	24.3	48.17	59.5
loss	1.44	0.64	0.59

Figure 2: Analyzing classification reports of FL-SMOTE, FL, and Utah-ANN models.

After simulating FL for the model without using SMOTE, its results showed that there were significant improvements compared to the centralized model. The limitations of the centralized model were mainly due to its smaller dataset. However, FL-SMOTE's results emphasized the importance of incorporating synthetic data to improve the balance of the dataset, thereby increasing its reliability. Beyond performance improvements, the FL and FL-SMOTE models also offer significant benefits regarding data privacy, which is critical in sensitive infrastructure data management. This reinforces the value of FL in building more reliable, robust, and private predictive analytics frameworks in infrastructure management.

The FL-SMOTE model exhibited superior performance compared to FL when applied to the six states' datasets due to its balanced dataset and avoiding its bias toward a specific target value. The standard FL setup has the potential to sustain or worsen class imbalances. This is because it typically aggregates updates from locally trained models without specifically addressing imbalances in class distribution across different entities. This can result in models that exhibit bias, performing favorably on majority classes while struggling with minority classes. Consequently, the model's overall accuracy and fairness can be significantly impacted. FL-SMOTE, on the other hand, allows each entity to generate synthetic samples for minority classes before training begins. This leads to a more balanced class distribution, which helps in training more generalizable models. Also, diversifying training examples can help to mitigate overfitting and improve the robustness of the model against unseen data.

Although the FL-SMOTE model demonstrated notable effectiveness in culvert condition prediction, it fell short of matching the performance levels achieved by the centralized ML model proposed by Mohammadi et al. (2024b). They leveraged a combined dataset from four different culvert inventories to enhance the predictive accuracy of their ML models. Centralized models typically exhibit superior performance due to the combination of diverse datasets into a unified database, though at the expense of compromising data privacy. However, enhancements in the predictive capabilities of FL-based models offer a dual benefit by eliminating the need to give up privacy for accuracy. Consequently, organizations stand to benefit from the precision of FL-based models while concurrently upholding the confidentiality of their data.

5 CONCLUSIONS

In this study, we simulated the application of a new ML paradigm, FL, as a solution to the challenge of data scarcity when developing efficient predictive models for transportation infrastructure management. Using Utah as a case study, we explored predictive modeling for culvert conditions, demonstrating the feasibility of achieving accurate predictions for UDOT while safeguarding the data privacy, security, and access rights of other DOTs. Utilizing an FL framework, we developed two FL-based models—one with SMOTE and one without—that surpassed the performance of a local ANN model built using Utah's limited culvert dataset. Notably, the FL-SMOTE model achieved more than 30% accuracy than the Utah-ANN model. According to the results of FL simulations, while standard FL offers significant privacy and high-performance benefits, FL-SMOTE further enhances these advantages by addressing the critical issue of class imbalance. This adaptation makes FL-SMOTE particularly valuable in fields like infrastructure management, where there is a scarcity of data and potential data imbalance issues. This advantage is particularly evident in the precision and recall metrics, where FL-SMOTE consistently outperforms FL. FL's methodology enables each participating state to contribute to and benefit from a collective global model while maintaining the confidentiality of its specific culvert data. This methodology is becoming more and more important due to the increasing focus on data privacy and security. While the centralized model, which combines data from multiple states, showed superior performance to the FL-based models, its compromise on data privacy poses significant limitations. Our findings suggest that improving the predictive capabilities of FL-based models can allow agencies like UDOT to utilize highly accurate models without sacrificing data confidentiality. Therefore, we suggest using optimization algorithms to further improve FL-based models or the inclusion of more state culvert datasets to enhance generalizability in future research. This study not only underscores the efficacy of FL in contexts marked by data scarcity and privacy concerns but also sets the stage for expanded investigations into the potential of FL in similar scenarios.

ACKNOWLEDGMENTS

This research project has been funded by the Utah Department of Transportation (UDOT) and the Mountain-Plains Consortium (MPC). The authors gratefully acknowledge the UDOT and MPC’s support and help. Any opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not reflect the views of the funding agencies.

REFERENCES

Beutel, D. J., T. Topal, A. Mathur, X. Qiu, J. Fernandez-Marques, Y. Gao, et al. 2020. “Flower: A Friendly Federated Learning Research Framework.”

Bukaty, P. 2019. *The california consumer privacy act (ccpa): An implementation guide*. IT Governance Ltd.

Chawla, N. V, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. "SMOTE: synthetic minority over-sampling technique." *J. Artif. Intell. Res.*, 16: 321–357.

Figueiredo, A. S. 2017. "Data sharing: convert challenges into opportunities." *Front. public Heal.*, 5: 327. Frontiers Media SA.

Gao, C., and H. Elzarka. 2021. "The use of decision tree based predictive models for improving the culvert inspection process." *Adv. Eng. Informatics*, 47: 101203. Elsevier.

Gondia, A., A. Siam, W. El-Dakhakhni, and A. H. Nassar. 2020. "Machine learning algorithms for construction projects delay risk prediction." *J. Constr. Eng. Manag.*, 146 (1): 4019085. American Society of Civil Engineers.

Hassandokht Mashhadi, A., A. Rashidi, and N. Marković. 2024. "A GAN-Augmented CNN Approach for Automated Roadside Safety Assessment of Rural Roadways." *J. Comput. Civ. Eng.*, 38 (2): 4023043. American Society of Civil Engineers.

Iyer, L. S. 2021. "AI enabled applications towards intelligent transportation." *Transp. Eng.*, 5: 100083. Elsevier.

Khalil, M., M. Esseghir, and L. Merghem-Boulahia. 2022. "A federated learning approach for thermal comfort management." *Adv. Eng. Informatics*, 52 (February): 101526. Elsevier Ltd. https://doi.org/10.1016/j.aei.2022.101526.

Konečný, J., H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon. 2016. "Federated learning: Strategies for improving communication efficiency." *arXiv Prepr. arXiv1610.05492*.

Li, L., Y. Fan, M. Tse, and K.-Y. Lin. 2020a. "A review of applications in federated learning." *Comput. Ind. Eng.*, 149: 106854. Elsevier.

Li, T., A. K. Sahu, A. Talwalkar, and V. Smith. 2020b. "Federated Learning: Challenges, Methods, and Future Directions." *IEEE Signal Process. Mag.*, 37 (3): 50–60. https://doi.org/10.1109/MSP.2020.2975749.

Li, X., H. lin Chi, W. Lu, F. Xue, J. Zeng, and C. Z. Li. 2021. "Federated transfer learning enabled smart work packaging for preserving personal image information of construction worker." *Autom. Constr.*, 128 (March): 103738. Elsevier B.V. https://doi.org/10.1016/j.autcon.2021.103738.

McGrath, T. J., and J. L. Beaver. 2004. *Condition Assessment of Utah Highway Culverts and Determination of Culvert Performance Measures. Utah Dep. Transp. Rep. UT-81FR0132, Salt Lake City, UT*.

McMahan, B., E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas. 2017. "Communication-efficient learning of deep networks from decentralized data." *Artif. Intell. Stat.*, 1273–1282. PMLR.

Mohammadi, P., A. Rashidi, and S. Asgari. 2024a. "Improving Culvert Condition Prediction Models Using Federated Learning: The Case Study of Utah." *Constr. Res. Congr. 2024*.

Mohammadi, P., A. Rashidi, and S. Asgari. 2024b. "An Improved Hybrid XGBoost Model for Culvert Inspection Using Swarm Intelligence Algorithms." *Comput. Civ. Eng. 2023*, Proceedings.

Mohammadi, P., A. Rashidi, M. Malekzadeh, and S. Tiwari. 2023a. "Evaluating various machine learning algorithms for automated inspection of culverts." *Eng. Anal. Bound. Elem.*, 148: 366–375. https://doi.org/10.1016/j.enganabound.2023.01.007.

Mohammadi, P., B. Sherafat, and A. Rashidi. 2023b. *Developing a Culvert Inspection Manual and Estimating Culverts' Deterioration Curve, Inspection Frequency and Service Life for UDOT*. Utah. Department of Transportation.

Moretti, N., D. Ph, X. Xie, D. Ph, J. M. Garcia, D. Ph, J. Chang, A. K. Parlikad, and D. Ph. 2023. "Federated Data Modeling for Built Environment Digital Twins." 37 (4): 1–15. https://doi.org/10.1061/JCCEE5.CPENG-4859.

Regulation of the European Parliament and of the council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive (general data protection regulation). 2018. "EU." Accessed April 22, 2023. https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en.

Saputra, Y. M., D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, M. D. Mueck, and S. Srikanteswara. 2019. "Energy demand prediction with federated learning for electric vehicle networks." *2019 IEEE Glob. Commun. Conf.*, 1–6. IEEE.

Sinha, K. C., S. Labi, and B. R. D. K. Agbelie. 2017. "Transportation infrastructure asset management in the new millennium: continuing issues, and emerging challenges and opportunities." *Transp. A Transp. Sci.*, 13 (7): 591–606. Taylor & Francis.

Stoner, M., W. Pang, and K. Piratla. 2019. "Predicting Culvert Deterioration Using Physical and Environmental Time-Independent Variables." *J. Pipeline Syst. Eng. Pract.*, 10 (4): 04019035. https://doi.org/10.1061/(asce)ps.1949-1204.0000409.

Tatari, O., S. M. Sargand, T. Masada, and B. Tarawneh. 2013. "Neural Network Approach to Condition Assessment of Highway Culverts: Case Study in Ohio." *J. Infrastruct. Syst.*, 19 (4): 409–414. https://doi.org/10.1061/(asce)is.1943-555x.0000139.

Theodorou, E., E. Spiliotis, and V. Assimakopoulos. 2023. "Optimizing inventory control through a data-driven and model-independent framework." *EURO J. Transp. Logist.*, 12: 100103. Elsevier.

Web Soil Survey. n.d. Accessed May 18, 2023. https://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx

Wijs, R. J. A., G. F. Nane, G. Leontaris, T. R. W. Van Manen, and A. R. M. Wolfert. 2020. "Improving subsurface asset failure predictions for utility operators: A unique case study on cable and pipe failures resulting from excavation work." *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.*, 6 (2): 5020002. American Society of Civil Engineers.

AUTHOR BIOGRAPHIES

POURIA MOHAMMADI is a Ph.D. student in Construction Management program at the Department of Civil and Environmental Engineering of the University of Utah. He received his B.Sc. degree in Civil Engineering from University of Tehran and M.Sc. degree in Construction Engineering and Management from University of Tehran. He has done researches about Building Information Modeling (BIM), Construction Simulation, and Sustainability. He is currently focusing on enahnacing transportation infrastructure management using Machine Learning, Computer Vision, and Artificial Intelligence algorithms. His email address is pouria.mohammadi@utah.edu and his website is https://faculty.utah.edu/u1343282-POURIA_MOHAMMADI/hm/index.hml.

ABBAS RASHIDI is an Associate Professor in the Civil and Environmental Engineering Department at the University of Utah. He earned his Ph.D. in Civil Engineering from the Georgia Institute of Technology. His research interests include information and sensing technologies for construction engineering and management, audio signal processing for civil infrastructure systems, and video/image processing for 3D reconstruction of civil infrastructure systems. . He currently serves as a member for several professional committees, including the Signal Processing in Acoustics Committee of the Acoustical Society of America and the Data Sensing and Analysis (DSA) Committee of the American Society of Civil Engineers (ASCE). He is an Associate Editor and a member of the Editorial Board of two ASCE journals, such as Journal of Construction Engineering and Management (ASCE) and Journal of Performance of Constructed Facilities (ASCE). His email address is abbas.rashidi@utah.edu. His webpage is https://faculty.utah.edu/u6013686-Abbas_Rashidi/hm/index.hml.

SADEGH ASGARI is an Associate Professor of Civil Engineering at Merrimack College and an Adjunct Professor at Columbia University. His research focuses on the application of artificial intelligence developing new technologies and decision-making tools to advance analysis and design of complex systems. He is the recipient of the Zampell family faculty fellowship and Edward G. Roddy outstanding teacher of the year award. He holds a doctoral degree in Civil Engineering from Columbia University. His e-mail address is asgaris@merrimack.edu. His website is https://www.merrimack.edu/profiles/sadegh-asgari/