Detection of the Pipeline Pig using Percussion and Machine Learning

Shengjun Lu*, and Gangbing Song†

Department of Mechanical Engineering, Cullen College of Engineering

*Email address: slu28@uh.edu, †Email address: gsong@uh.edu

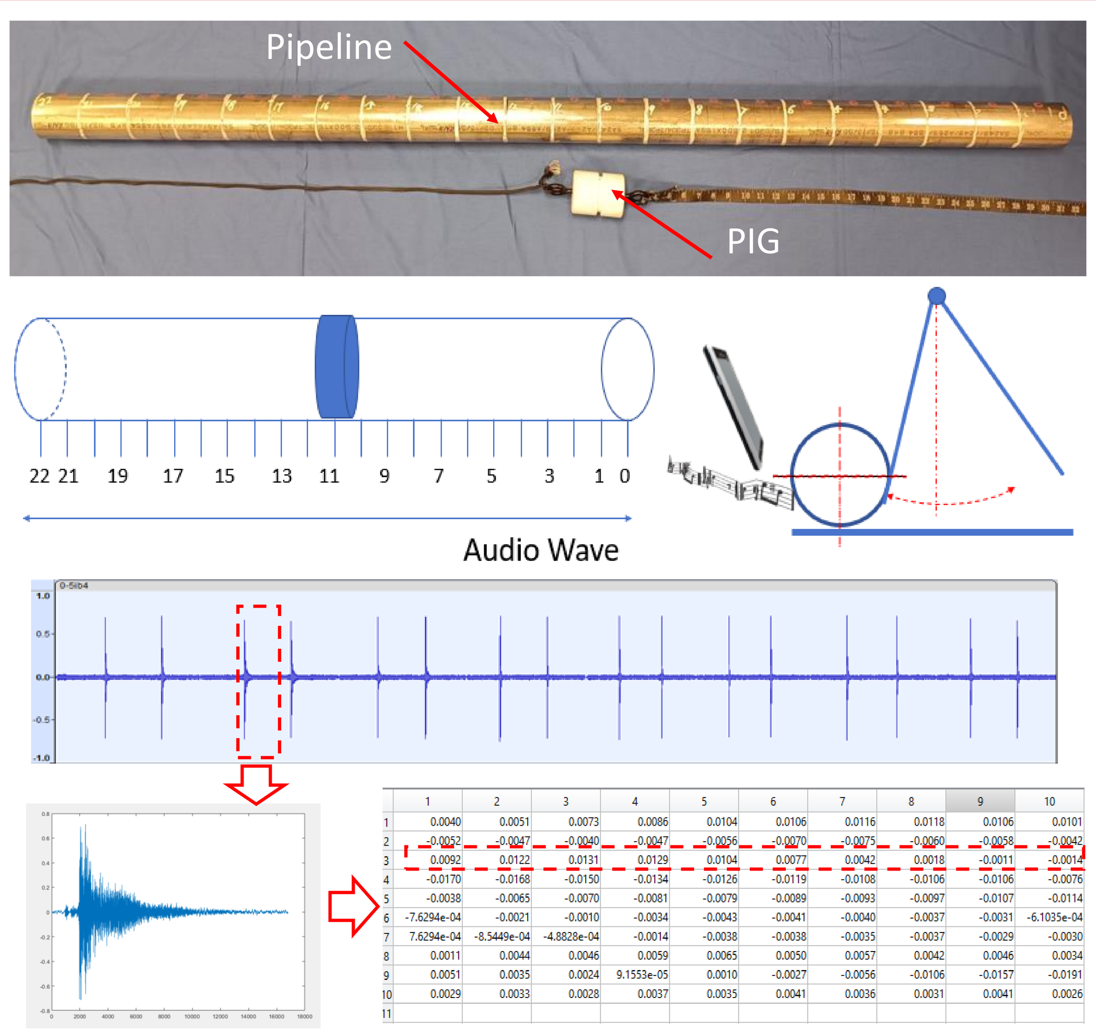
Problem Statement

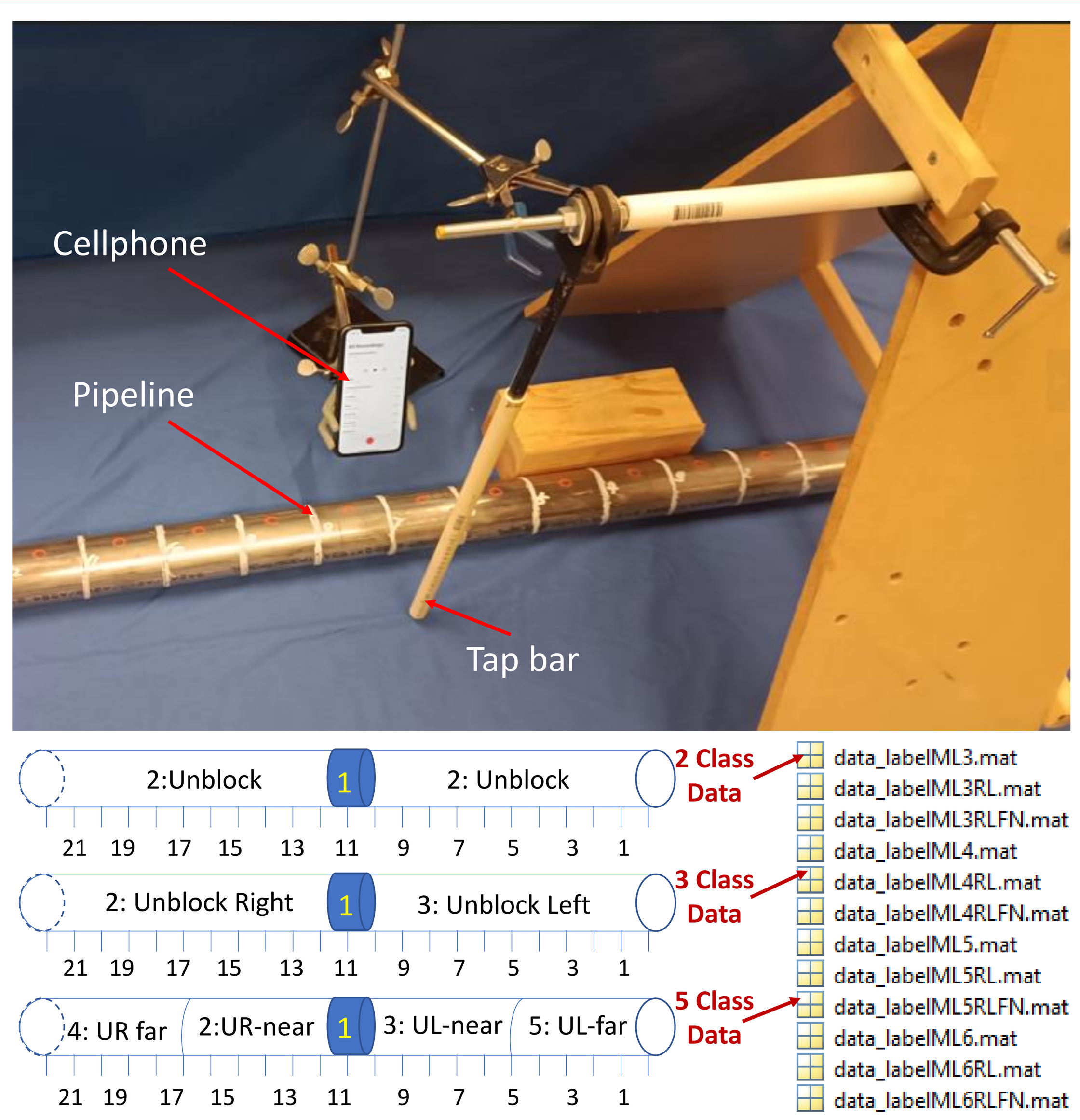
- Ambient noise, varying pipeline configurations, and different types of pigs further complicate the detection process.
- Machine learning algorithms have played a significant role in enhancing the accuracy and reliability of pipeline pig detection systems.

Brief Literature Review

- Static networks and recurrent networks (LSTM) were built, to collect data for the model training. The implementation of the supervised neural networks used the Python library TensorFlow package.
- Compared with the support vector machine (SVM) model and the decision tree model, the CNN model has better recognition performance.

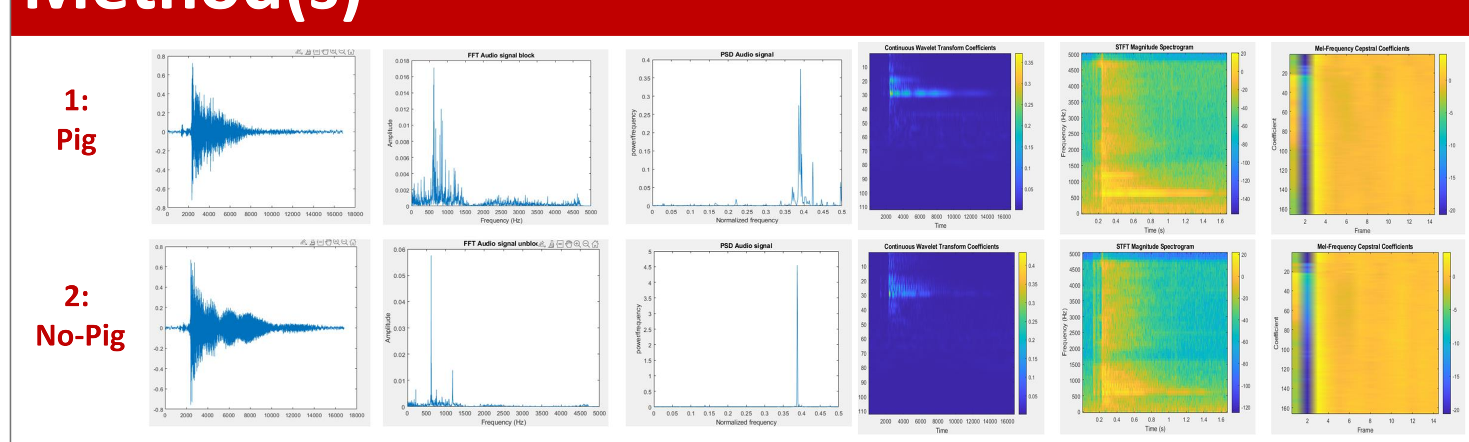
Experimental Setup and Collection of Data

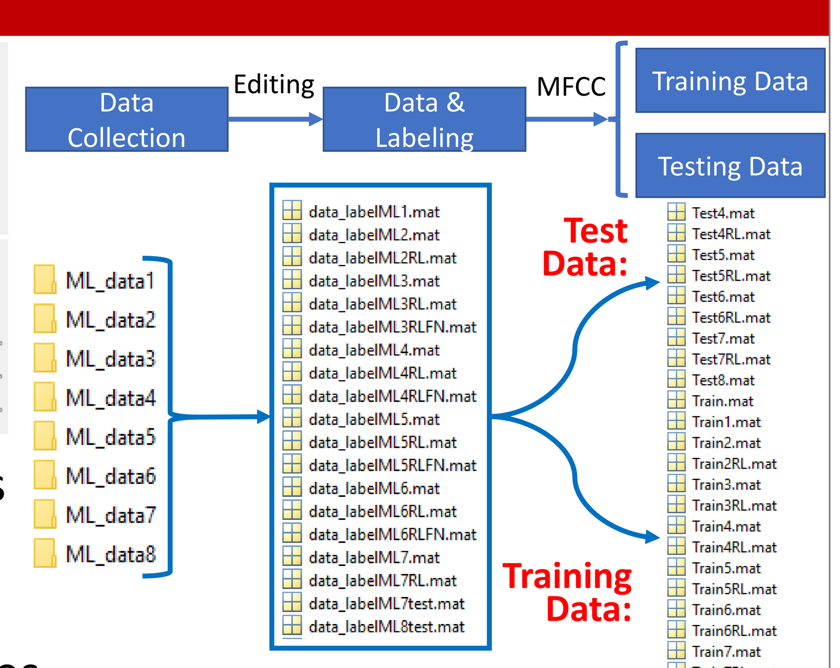




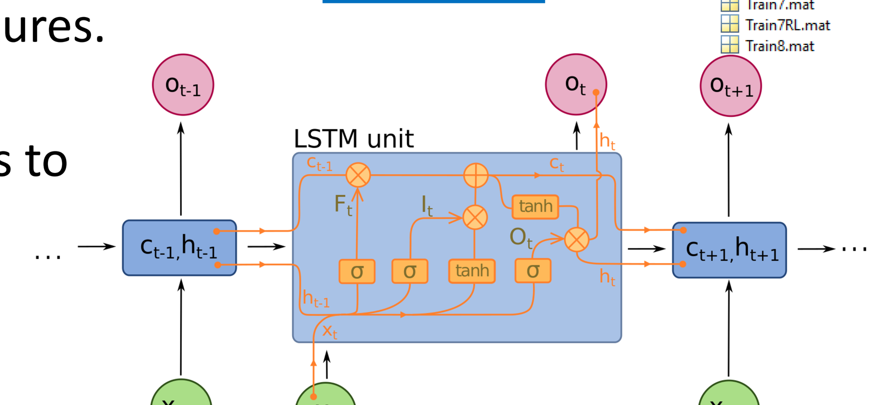
- This produces a total of 440 audio signals for each set of data, 220 at the Pig position were labeled as blocked and 220 at the no Pig position were labeled as unblocked.
- Collected 8 groups of data (5500 signals) and processed them into various data sets two-class, three-class(RL), and five-class(RLFN) to train and test machine learning models effectively.

Method(s)





- There are various techniques for processing audio signals such as FFT, PSD, CWT, STFT, and MFCC.
- The audio signals through MFCC signal processing and formed multiple training and testing datasets using the processed features.
- Using the dependent data to begin the process, we utilize standard machine learning such as KNN, DT, and SVM practices to classify the data.
- Recurrent neural networks (RNN) are theoretically Turing complete and can run arbitrary programs to process arbitrary sequences of inputs.

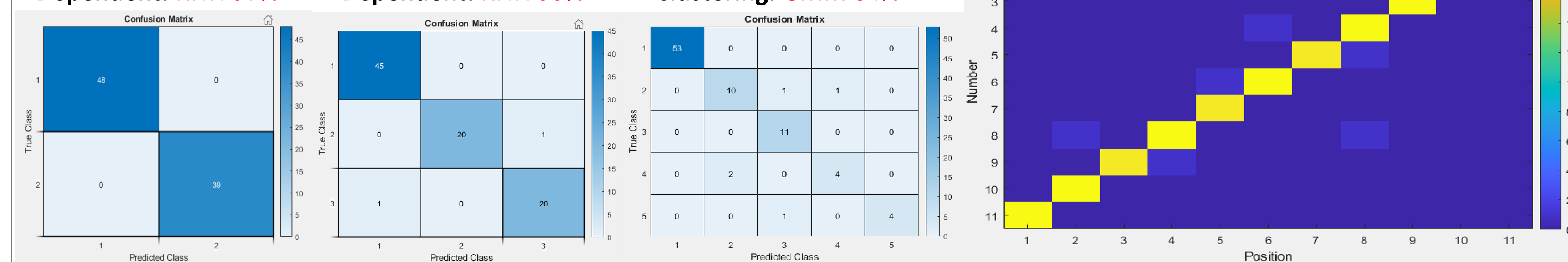


Results, Analysis and Discussion

Dependent: SVM 100%
Dependent: KNN 100%
Dependent: RNN 97%

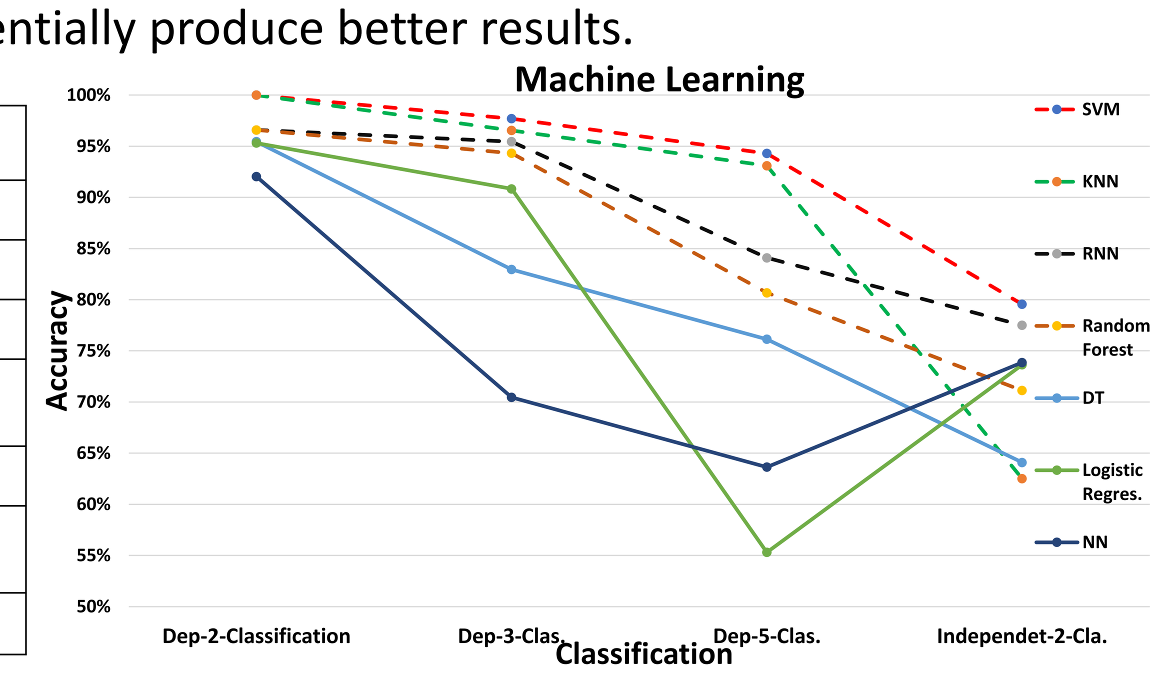
Dependent: SVM 98%
Dependent: KNN 97%
Dependent: RNN 95%

Independent: SVM 80%
Independent: RNN 78%
Clustering: GMM 84%



- After trying out several machine learning methods and have realized that to achieve high accuracy with multiple classifications, we need to use more training data. SVM and RNN have demonstrated advantages in classification for this project.
- Deep learning methods with more training and test data, we may be able to achieve higher accuracy by adjusting the parameters of deep learning for independent classifications.
- I have used independent training and test data for 2 classifications, we can extend that to 3 and 5 classifications, or even more fine-grained classifications. The accuracy of clustering is lower, need a better clustering method for classification.
- In the future, I believe that reinforcement learning classification methods could prove useful, and it could lead to more accurate data classification. Automating this process could generate a massive amount of data and not only use MFCC for processing both the training and test data but also find other signal processing methods that could potentially produce better results.

	Dependent-2-Classification	Dependent-3-Classification	Dependent-5-Classification	Independent-2-Classification
SVM	100%	98%	94%	80%
RNN	97%	95%	84%	78%
KNN	100%	97%	93%	63%
Random Forest	97%	94%	81%	71%
DT	95%	83%	76%	64%
Logistic Regression	95%	91%	55%	74%
NN	92%	70%	64%	74%



Conclusion

- It is observed that dependent data can lead to higher accuracy, especially when there are only two classifications, although the accuracy of higher classifications may be lower.
- Independent data can be more challenging to obtain accurate results than dependent data, resulting in lower accuracy.
- Support Vector Machine (SVM) has proven to be a reliable method with better results compared to deep learning methods such as Neural Networks (NN) and even better than Recurrent Neural Networks (RNN). This highlights the importance of selecting the appropriate method for a particular task.
- While the clustering method has the lowest accuracy compared to other methods mentioned, it still has its advantages without labels and therefore can be useful in certain scenarios.

Acknowledgements

The financial support from Midstream Integrity Services (MIS) and technical support from Smart Materials & Structures Lab (SMSL) and Artificial Intelligence Lab for Monitoring & Inspection (AILMI) at UH.

References (brief)

1. Victor Carvalho Galvão De Freitas, Valbério Gonzaga De Araujo, Daniel Carlos de Carvalho Crisóstomo, Gustavo Fernandes De Lima, Adrião Duarte Dória Neto, Andrés Ortiz Salazar, Velocity Prediction of a Pipeline Inspection Gauge (PIG) with Machine Learning Sensors 2022, 22(23), 9162; https://doi.org/10.3390/s22239162
2. Dan Yang, Mengzhou Xiong, Tao Wang, Guangtao Lu, Percussion-Based Pipeline Ponding Detection Using a Convolutional Neural Network Appl. Sci. 2022, 12(4),2127; https://doi.org/10.3390/app12042127
3. Jiacheng Wei, Xi Tang, Jinxiu Liu, and Zhiyan Zhang, Detection of Pig Movement and Aggression Using Deep Learning Approaches, Animals 2023, 13(19), 3074; https://doi.org/10.3390/ani13193074
4. How to ensure accurate and reliable pipeline pig tracking, Posted by Nicola Porterhttps://blog.tracerco.com/marketing/accurate-and-reliable-pipeline-pig-tracking3.