THE ANNALS OF "DUNAREA DE JOS" UNIVERSITY OF GALATI
FASCICLE IX. METALLURGY AND MATERIALS SCIENCE
N°. 4 - 2023, ISSN 2668-4748; e-ISSN 2668-4756
Article DOI: https://doi.org/10.35219/mms.2023.4.14

CHAINSAW SOUND DETECTION USING DNN ALGORITHM

Florin Bogdan MARIN, Mihaela MARIN
"Dunarea de Jos" University of Galati, Romania
e-mail: florin.marin@ugal.ro

ABSTRACT

Deforestation and illegal logging stand as important environmental problems. In this paper we propose a DNN architecture for sound recognition of chainsaw detection. Various parameters need to be tuned in order to identify the sound of chainsaw but not to produce too much amounts of false positive detection. The task is challenging as different sound emerge in the forest.

KEYWORDS: sound processing, sound recognition, chainsaw sound detection

1. Introduction

Deforestation, the widespread cutting of forests, and illegal logging, the illicit harvesting of timber, stand as important environmental menaces with far-reaching consequences. These practices, driven by various factors including economic interests, agricultural expansion, and demand for timber, pose significant threats to biodiversity, climate stability, and the livelihoods of local communities. Illegal logging, a subset of deforestation, is perpetuated by the lure of financial gain. The demand for timber, often driven by the construction industry and global markets, creates lucrative opportunities for illegal loggers. Weak law enforcement, corruption, and inadequate governance in some regions exacerbate the problem, allowing illegal logging to persist [1-3].

The impact of deforestation and illegal logging on biodiversity is profound. Forests are home to a staggering array of plant and animal species, many of which are endemic and found nowhere else on Earth [4].

One solution to identify the use of chainsaw is to detect the specific sound using portable devices that are dedicated to this or to use cameras used for animal activity, as presented in Fig.1. The task to recognize the specific sound of chainsaw is quite challenging taking into account vast sound that emerge in forests. There are several algorithms commonly used for sound recognition.

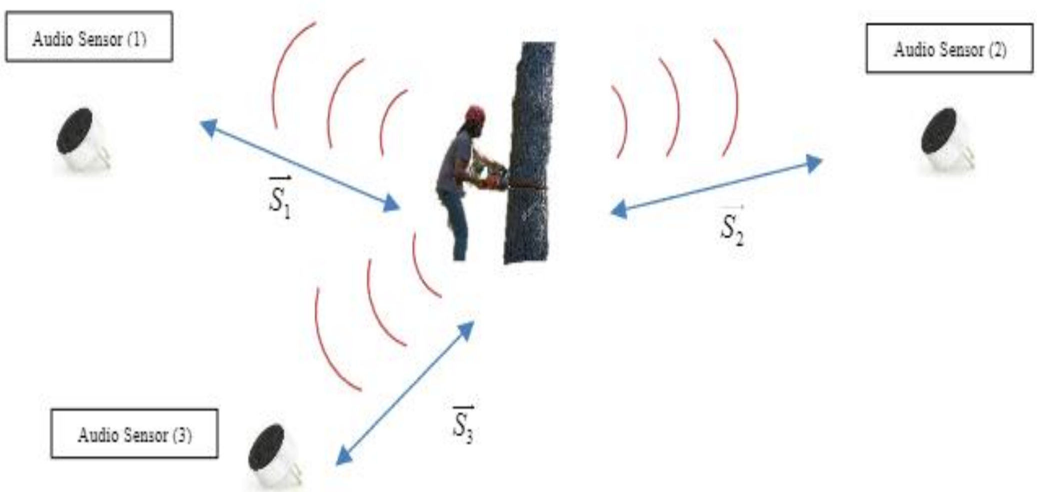

Fig. 1. *The concept of chainsaw sound detection* [1]

THE ANNALS OF "DUNAREA DE JOS" UNIVERSITY OF GALATI
FASCICLE IX. METALLURGY AND MATERIALS SCIENCE
N°. 4 - 2023, ISSN 2668-4748; e-ISSN 2668-4756
Article DOI: https://doi.org/10.35219/mms.2023.4.14

Fast Fourier Transform (FFT) is a widely used algorithm for transforming a signal from the time domain to the frequency domain. It devises a signal into its constituent frequencies, allowing for the analysis of the spectral content of the sound.

Mel-Frequency Cepstral Coefficients (MFCC) is a feature extraction technique commonly used in speech and audio processing. It represents the short-term power spectrum of a sound and is widely used in sound recognition tasks. MFCCs capture the characteristics of the human auditory system.

Hidden Markov Models (HMMs) are used for modeling sequential data, making them suitable for tasks where the temporal dynamics of sound are important. HMMs have been applied in speech recognition and environmental sound classification.

Gaussian Mixture Models (GMMs) are probabilistic models that can be used for sound modeling and recognition. They are often applied to represent the statistical distribution of feature vectors extracted from sound signals.

Convolutional Neural Networks (CNNs), known for their success in image processing, have also been applied to sound recognition tasks. They can learn hierarchical representations of sound features by processing spectrogram or other time-frequency representations.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTMs) are well-suited for tasks involving sequential data, making them applicable to sound recognition. They can capture temporal dependencies in sound sequences.

Support Vector Machines (SVMs) are used for classification tasks, including sound recognition. They work by finding the hyperplane that best separates different classes of sound based on their feature vectors.

Dynamic Time Warping (DTW) is a method for measuring the similarity between two temporal sequences that may vary in speed. It has been used in sound recognition to compare and align time series data.

Nearest Neighbor Algorithms, known as k-Nearest Neighbors (k-NN) and other nearest neighbour algorithms can be used for sound recognition by comparing the input sound features with those of known sound classes.

Ensemble methods, such as Random Forests or Gradient Boosting, combine the predictions of multiple base models. They are effective for improving the robustness and generalization of sound recognition systems.

Transfer learning involves leveraging pre-trained models on large datasets and fine-tuning them for specific sound recognition tasks. This approach is especially useful when limited labelled data is available.

These algorithms can be used individually or in combination, depending on the complexity of the sound recognition task and the characteristics of the data. The choice of algorithm often depends on factors such as the nature of the sound data, the available computational resources, and the specific requirements of the application. Chainsaw sound recognition using deep learning has practical applications in monitoring and combatting illegal logging activities, promoting conservation efforts, It showcases the potential of advanced technologies in addressing environmental challenges and promoting sustainable practices [5-10].

Deep learning algorithms excel at capturing complex patterns and dependencies in data, making them well-suited for tasks where distinguishing between subtle variations in sound is crucial [11, 12].

The models can adapt to variations in environmental conditions, such as changes in background noise, allowing for robust performance in diverse forest settings [13, 14].

2. Technique proposed

To train a deep learning model for chainsaw sound recognition, a dataset needs to be collected. This dataset should include audio recordings of various forest sounds, with a focus on capturing chainsaw sounds in different contexts and environmental conditions. We used own audio data consisting of 2 hours of recording.

The collected audio data is pre-processed to extract relevant features. This means converting the audio signals into a spectrogram, a visual representation of the spectrum of frequencies over time. Spectrograms provide a rich input for deep learning models to learn patterns associated with different sounds.

The deep learning model is designed to process the spectrogram data. Convolutional Neural Networks (CNNs) are used for sound recognition tasks. Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks are other algorithm employed to capture temporal dependencies in the audio signals.

Training the Model is one important step in using CNN.

The model is trained using the pre-processed data, with chainsaw and non-chainsaw sounds labelled appropriately. During training, the deep learning algorithm adjusts its internal parameters based on the input data, learning to recognize features that distinguish chainsaw sounds from other sounds in the forest.

The trained model is evaluated on a separate dataset to assess its performance. Adjustments and fine-tuning are done based on the evaluation results to

N°. 4 - 2023, ISSN 2668-4748; e-ISSN 2668-4756
Article DOI: https://doi.org/10.35219/mms.2023.4.14

enhance the model's accuracy and generalization capabilities.

Once trained, the deep learning model can be used for chainsaw sound recognition in real-time. When deployed in the field, the model processes incoming audio signals and predicts whether the sound corresponds to a chainsaw.

The strengths of deep learning in this context include an important stage in learning, feature learning. Deep learning models automatically learn hierarchical representations of features from the raw audio data, eliminating the need for manual feature engineering.

In this study, we test several parameters for a deep learning algorithm to detect the presence of acoustic data of engine chainsaw. As chainsaw sound are rare events, we aimed to create a network architecture that maximises parameters in order to minimise false negatives and protecting forest ecosystems.

Using a Convolutional Neural Network (CNN) for audio recognition involves experimenting with various architectural elements and hyperparameters to achieve better performance.

Concerning convolutional layers, we experiment with the number of convolutional layers, kernel sizes, and the number of filters in each layer. Increasing the depth and width of the network capture more complex features. We tested the network for two convolutional layers with the size 3X3 and 6 layers with the size 3X3. The difference in accuracy to identify was very small, 0.5 percent, but the learning time increased by 30 percent.

Adjust the pooling layer's size and stride may produce improved results. Pooling helps reduce spatial dimensions and control overfitting. We tested using

Pool Size of (2,2) (4,4) and the identification rate drops from 98.2% to 90%. That is because the pool size defines the spatial extent over which the pooling operation is applied. A larger pool size reduces spatial dimensions and discard fine features.

Adjusting the learning rate (lr) with a too high learning rate can cause the model to converge too quickly or oscillate, while too low learning rate can lead to slow convergence. We experimented with optimizers Adam and SGD. For our own data set Adam predicted with 12% better results than SGD. The batch size can impact training speed and memory usage.

We experiment different size of 32,64,128. The best prediction rate was 94.2%, 95.2%, 92.3 % respectively. However, we note that false pozitive detection are higher in case of 64 dimensions.

3. Conclusions

Acoustic technology offers a multitude of opportunities for monitoring biodiversity, environmental health, and human disturbance such as gun hunting. However, there is currently a mismatch between the speed that affordable hardware is being developed and the ability of ecologists to process the vast amounts of data collected.

We have developed a deep learning technique that is able to classify sound information. Our algorithm achieve > 95% accuracy even we used relatively small training data sample size. Our algorithm is designed to work on low computational resources devices such as remotely processing units.

References

[1]. Wang S. F., Wang K. Y., Wang X. J., Liu Z. Q., *A Novel Illegal Logging Monitoring System Based on WSN*, Advanced Materials Research, p. 1417-142, 2012.

[2]. Tang Y., Han P., Wang Z., Hu L., Gao Y., Li H., *Based on intelligent voice recognition of forest illegal felling of detecting methods*, 2nd International Conference on Cloud Computing and Intelligent Systems, p. 1153-1156, 2012.

[3]. Soisoonthorn T., Rujipattanapong S., *Deforestation detection algorithm for wireless sensor networks*. 10.1109/ISCIT.2007.4392237, 2007.

[4]. Colonna J. G., Gatto B., Santos E. D., Nakamura E. F., *A framework for chainsaw detection using one-class kernel and wireless acoustic sensor networks into the amazon rainforest*, Mobile Data Management (MDM), 17th IEEE International Conference, vol. 2, p. 34-36, 2016.

[5]. Harvanova V., Vojtko M., Babis M., Duricek M., Pohronska M., *Detection of wood logging based on sound recognition using zigbee sensor network*, In Proceedings of the International Conference on Design and Architectures for Signal and Image Processing, Tampere, Finland, 2-4 November 2011.

[6]. Kalhara P. G., Jayasinghearachchd V. D., Dias A. H. A. T., Ratnayake V. C., Jayawardena C., Kuruwitaarachchi N., *TreeSpirit: Illegal logging detection and alerting system using audio identification over an IoT network*, In Proceedings of the 2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Malabe, Sri Lanka, 6–8 December 2017.

[7]. Prasetyo D. C., Mutiara G. A., Handayani R., *Chainsaw sound and vibration detector system for illegal logging*, In Proceedings of the 2018 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC), Bali, Indonesia, 5–7 December 2018.

[8]. Jubjainai P., Pathomwong S., Siripujaka P., Chiengmai N., Chaiboot A., Wardkein P., *Chainsaw location finding based on travelling of sound wave in air and ground*, IOP Conf. Ser. Earth Environ. Sci. 2020.

[9]. Andrei V., Cucu H., Petrică L., *Considerations on developing a chainsaw intrusion detection and localization system for preventing unauthorized logging*, Journal of Electrical and Electronic Engineering, 3(6), p. 202-207, 2015.

[10]. Czuni L., Varga P. Z., *Time Domain Audio Features for Chainsaw Noise Detection Using WSNs*, IEEE Sens. J. 17, 1, 2017.

[11]. Meedeniya D., Ariyarathne I., Bandara M., Jayasundara R., Perera C., *A Survey on Deep Learning Based Forest Environment Sound Classification at the Edge*, ACM Computing Surveys, 56(3), p. 1-36, 2023.

[12]. Somwong B., Kumphet K., Massagram W., *Acoustic Monitoring System with AI Threat Detection System for Forest*

N°. 4 - 2023, ISSN 2668-4748; e-ISSN 2668-4756
Article DOI: https://doi.org/10.35219/mms.2023.4.14

Protection, 20th International Joint Conference on Computer Science and Software Engineering (JCSSE), p. 253-257, IEEE, 2023.

[13]. Stefanakis N., Psaroulakis K., Simou N., Astaras C., *An Open-Access System for Long-Range Chainsaw Sound Detection*, 30th European Signal Processing Conference (EUSIPCO), p. 264-268, IEEE, 2022.

[14]. Prasetyo D. C., Mutiara G. A., Handayani R., *Chainsaw sound and vibration detector system for illegal logging*, International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC), p. 93-98, IEEE, 2018.