Mateusz Erezman 171675

Michał Hajdasz 172156

Dariusz Kobiela 175656

Systematic Literature Review Report

1. Research project

1.1. Title

A vehicle type recognition model based on audio data from its drive-through

1.2. Supervisor

mgr inż. Szymon Zaporowski

1.3. Goals and short description

The goal of the project is to develop a deep neural network model that is capable of recognizing the type of passing vehicle based on the input of an audio file. The approach presented in the literature can be used as a reference for developing researchers own more advanced model using e.g. Transformers.

Main tasks to accomplish are:
1. Literature review
2. Training data analysis
3. Labelling additional data
4. Network architecture selection
5. Model training
6. Model verification
7. Analysis of results

2. Systematic Literature Review plan

2.1. Goals and questions

The goal of the SLR is to get knowledge about models used to recognize vehicles by their sound. Furthermore, we would like to see the advantages and disadvantages of those models in our field of interest and also find out what results we should be expecting from our implementation.
So, the question is:
What machine learning models are used in vehicle type recognition based on the sound?
PL: Jakie modele uczenia maszynowego są używane do rozpoznawania typu pojazdu na podstawie dźwięku?

2.2. Keywords

- machine learning, deep learning, neural network
- vehicle type recognition, classification, identification
- sound, acoustics

2.3. Search strings

- (“machine learning” OR “deep learning” OR “neural network”)

- (“vehicle recognition” OR “vehicle classification” OR “vehicle identification” OR “car recognition”)
- (sound OR acoustic)
- final search string: (“machine learning” OR “deep learning” OR “neural network”) AND (“vehicle recognition” OR “vehicle classification” OR “vehicle identification” OR “car recognition”) AND (sound OR acoustic)

2.4. Literature databases

- Scopus
- SpringerLink
- IEEExplore
- ACM Digital Library – optional – only if there will be not enough articles in three above

2.5. Inclusion criteria

- conference and journal papers published 2016-2022 (article, conference paper, conference review)
- publications in English
- only vehicle recognition (classification)
- vehicle recognition based on sound

2.6. Exclusion criteria

- research related to electric vehicles
- articles closely related to other articles (e.g. extended versions, reference work)
- research not related to machine learning approach
- vehicle recognition not only based on sound
- recognition of different objects (eg. road surface wetness detection, road surface quality detection based on gravel) based on sound

2.7. Quality criteria

- paper reviews or compares machine learning models for sound recognition
- the similarity of the recognized vehicle types to ours (cars and similar urban vehicles)
- the description of the model(s) used (details about the model)

2.8. Data extraction

- models for vehicle sound recognition with its pros and cons with regard to our topic

2.9. SLR process

2.9.1. Steps

1. Plan definition, verification and acceptance
2. Search string construction and validation
3. Databases search with search string and inclusion criteria
4. Selection 1st phase based on paper titles, abstracts and exclusion criteria (scoring from **0 to 2** by each person). Only articles with summary **scoring >=4** are taken into further consideration.
5. Reading paper content and assessment of quality criteria – each article is given score from **1 to 5**
6. Final selection of papers based on **score (>=3)** and quality
7. Snowball sampling – review of references in the selected papers
8. Extraction of the models from each paper
9. Aggregation and compilation of a final list of models that we can use in our project
10. Selecting the model or models that suits our needs best

2.9.2. Assignments

Step	Executioners	Reviewers
1, 2	MH	ME, DK
3, 4	DK(S), ME(SL), MH(IE)	DK, ME, MH
5	DK(A), ME(B), MH(C)	DK, ME, MH
6	DK	MH
7	ME	DK
8	MH, DK	ME
9	ME	MH
10	DK, ME, MH	DK, ME, MH

ME – Mateusz Erezman
MH – Michał Hajdasz
DK – Dariusz Kobiela

S – Scopus
SL – SpringerLink
IE - IEEExplore

A, B, C – groups that each contains 1/3 out of all articles from previous step

2.9.3. Tools to use

- mendeley.com – for easy articles storage
- clickup.com – for task assignment and workflow control
- Discord – for online meetings and problems discussing
- Google Spreedsheets - for statistics, charts, reviewing and scoring process

3. Systematic Literature Review results

3.1. Results in numbers

3.1.1. Step 3 – search with constraints

- Scopus - 48 articles
- SpringerLink - 65 articles
- IEEExplore - 134 article
- All together - 247 articles

3.1.2. Step 4 – Selection 1st Phase

- 21 articles

3.1.3. Step 6 - Final selection

- 16 articles

3.1.4. Step 8 - Snowballing

- 6 additional articles

3.2. Articles selected for data extraction

1D convolutional context-aware architectures for acoustic sensing and recognition of passing vehicle type	2020 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)	A. Kurowski; S. Zaporowski; A. Czyżewski	2020
Acoustic Emission Signal Classification Using Feature Analysis and Deep Learning Neural Network	Fluctuation and Noise Letters	Wu J.-D., Wong Y.-H., Luo W.-J., Yao K.-C.	2021
Acoustic Signal Classification Using Symmetrized Dot Pattern and Convolutional Neural Network	Machines	Wu J.-D., Luo W.-J., Yao K.-C.	2022
Audio Feature Extraction for Vehicle Engine Noise Classification	ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)	L. Becker; A. Nelus; J. Gauer; L. Rudolph; R. Martin	2020
Automatic labeling of traffic sound recordings using autoencoder-derived features	2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)	A. Kurowski; A. Czyżewski; S. Zaporowski	2019
Deep Recurrent Neural Networks for Audio Classification in Construction Sites	2020 28th European Signal Processing Conference (EUSIPCO)	M. Scarpiniti; D. Comminiello; A. Uncini; Y. -C. Lee	2021
Hybrid Neural Network based on Feature Fusion for Vehicle Type Identification	2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)	H. Chen; Z. Zhang; W. Yin; M. Wang; M. Lifan; X. Hao	2020
Recognition of Moving Tracked and Wheeled Vehicles Based on Sound Analysis and Machine Learning Algorithms	International Journal of Automotive and Mechanical Engineering	Jackowski J., Jakubowski J.	2021
Sound-Convolutional Recurrent Neural Networks for Vehicle Classification Based on Vehicle Acoustic Signals	2021 International Conference on Smart City and Green Energy, ICSCGE 2021	Luo Y., Chen L., Wu Q., Zhang X.	2021

Sound-based Transportation Mode Recognition with Smartphones	ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)	L. Wang; D. Roggen	2019
Spectral features for audio based vehicle and engine classification	Journal of Intelligent Information Systems	Alicja Wieczorkowska, Elżbieta Kubera Tomasz Słowik Krzysztof Skrzypiec	2018
Transportation mode detection using cumulative acoustic sensing and analysis	Frontiers of Computer Science	Dinesh VijNaveen Aggarwal	2020
Unsupervised vehicle recognition using incremental reseeding of acoustic signatures	Lecture Notes in Computer Science	Sunu J., Percus A.G., Hunter B.	2018
Vehicle Identification Based on Improved 1/3 Octave and Bark-Scale Wavelet Packet Methods	2021 IEEE 4th International Conference on Electronics Technology (ICET)	Q. Wang; Y. He; Z. Chen; Y. Luo	2021

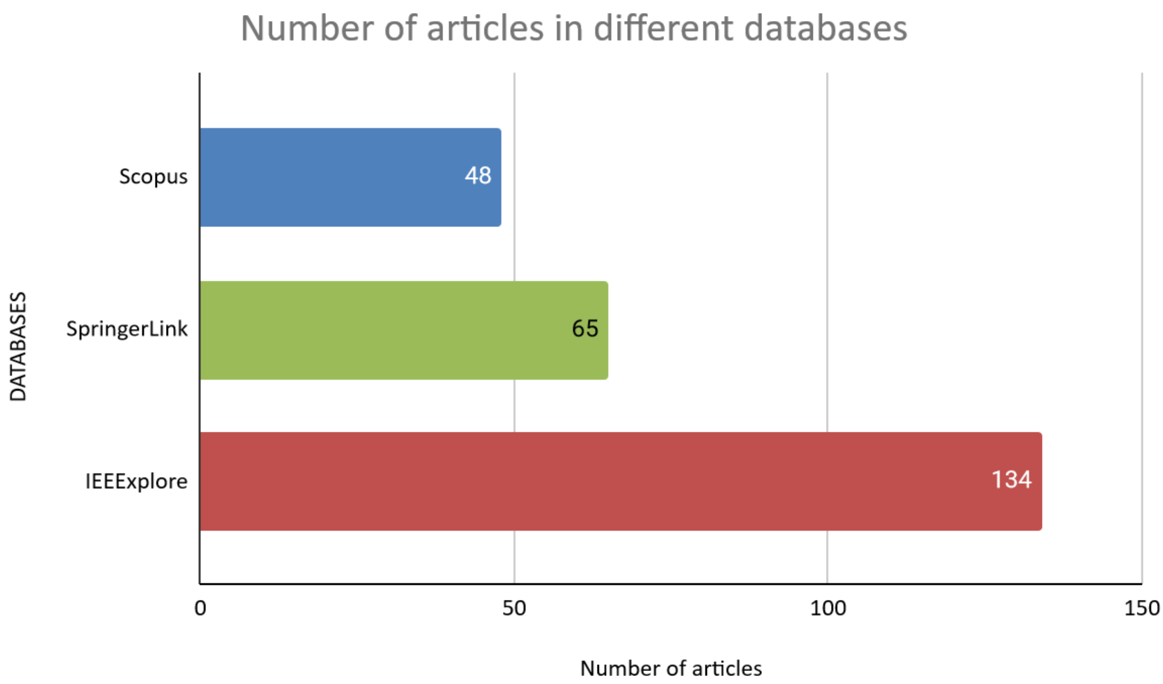
3.3. Snowballed articles

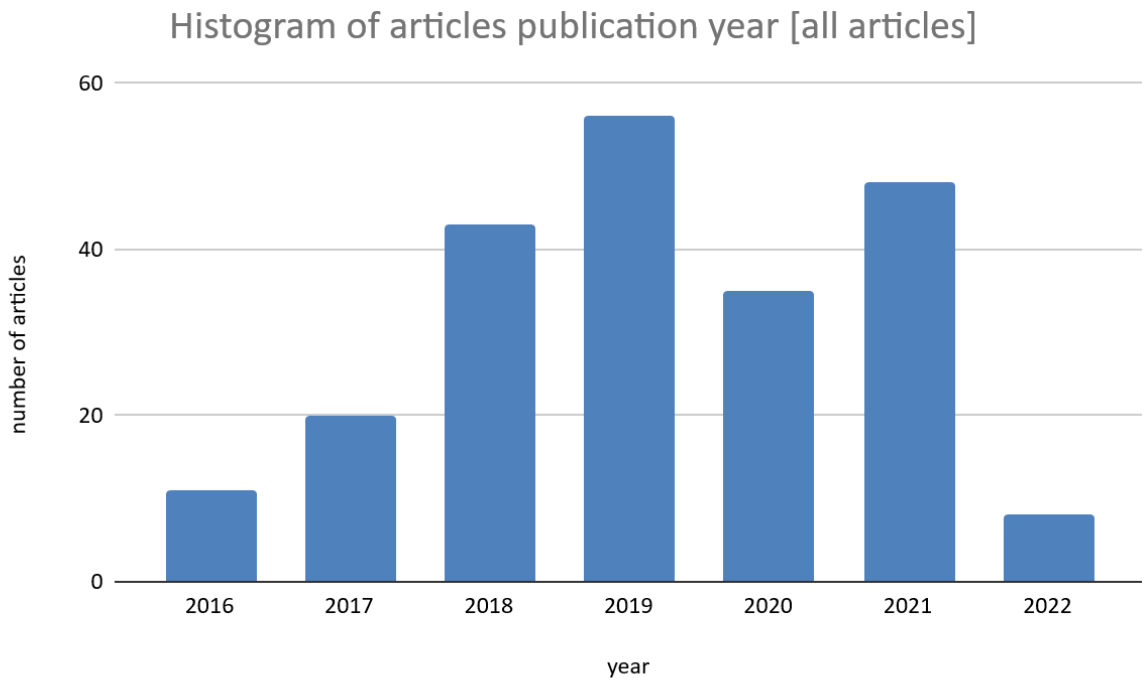
Environmental Intelligence for Embedded Real-time Traffic Sound Classification	P. Montino; D. Pau	2019
Sound event classification using neural networks and feature selection based methods	A. Ahmed; Y. Serrestou; K. Raoof; J. -F. Diouris	2021
Machine Learning Inspired Sound-Based Amateur Drone Detection for Public Safety Applications	M. Z. Anwar; Z. Kaleem; A. Jamalipour	2019
Research on Environmental Sound Classification Algorithm Based on Multi-feature Fusion	Ruixue Li; Bo Yin; Yongchao Cui; Zehua Du; Kexin Li	2020
VehicleSense: A Reliable Sound-based Transportation Mode Recognition System for Smartphones	Sungyong Lee, Jinsung Lee, Kyunghan Lee	2017
Moving Vehicle Noise Classification using Multiple Classifiers	N. Abdul Rahim, Paulraj M P, A. H. Adom, and S. Sathish Kumar	2011

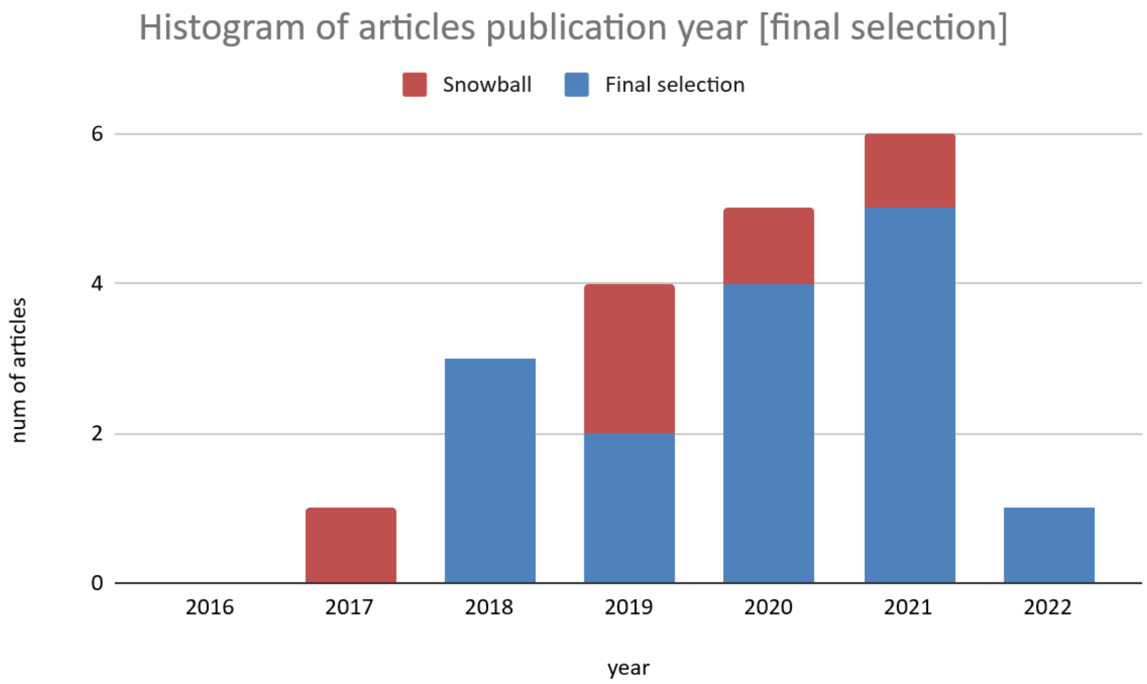
3.4. Article statistics

STATISTICS	
Sum of all articles at the beginning	247
Sum of articles after removing duplicates	241
Sum of articles after exclusion criteria (removing books, articles < 2016 and referenceWorks) [1st selection]	184
Sum of articles with the highest rating (6, 5 and 4 points) [2nd selection]	21
Sum of articles after final scoring (1-5) and quality criteria [articles with mark >= 3]	16
Sum of articles after snowballing	22
FINAL SUM OF ARTICLES	**22**

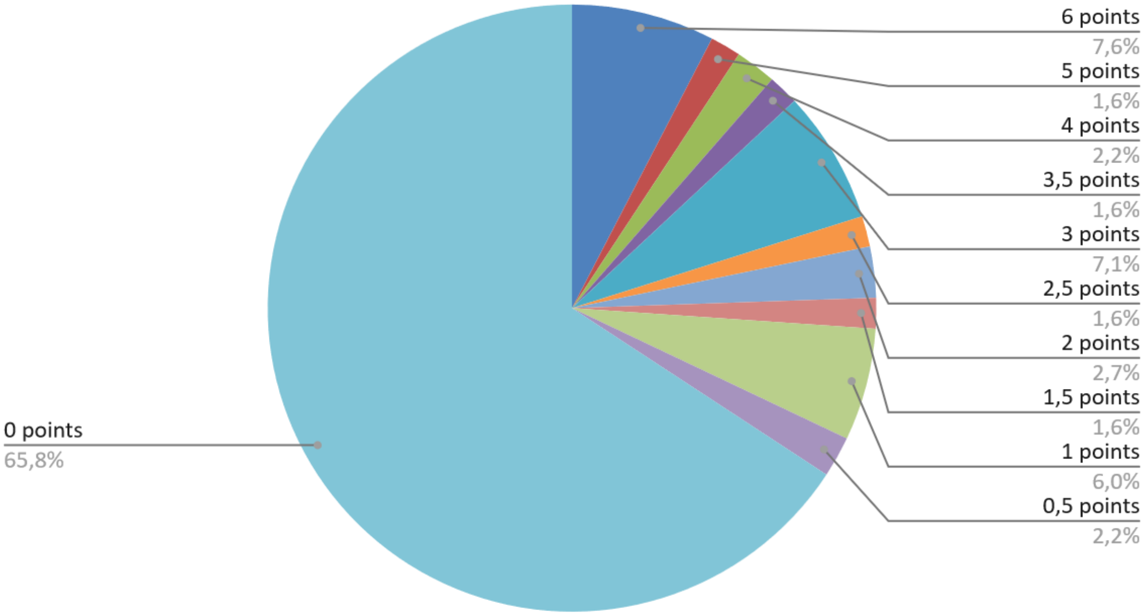
After rating 0-2 by title and abstract	
No. points	**No. articles**
6 points	14
5 points	3
4 points	4
SUM	**21**



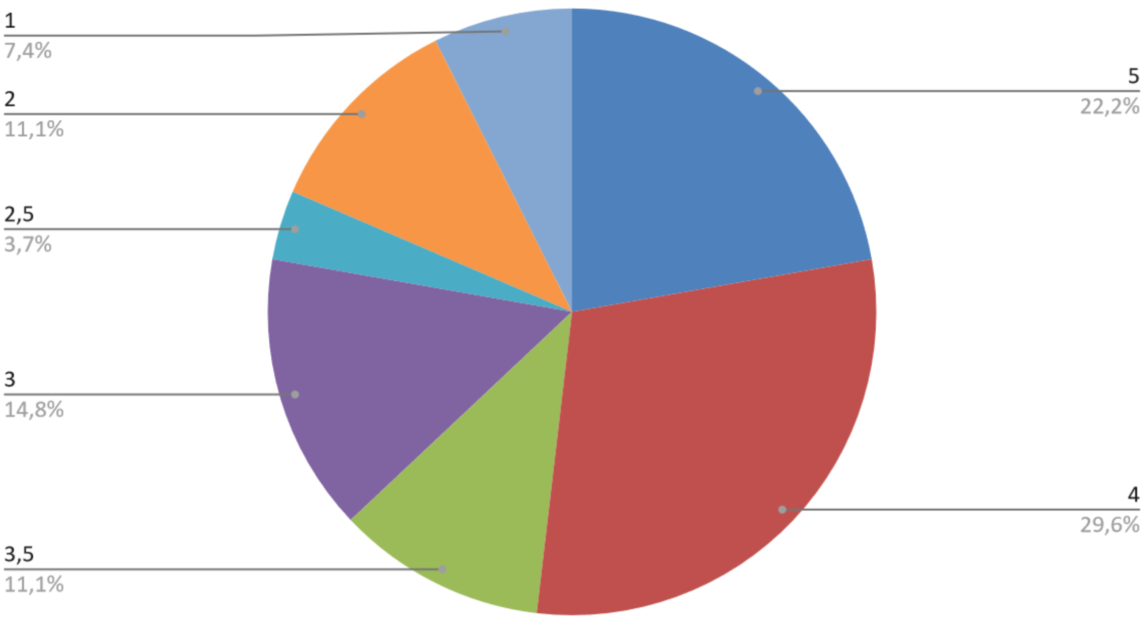

Histogram of articles publication year [all articles]
number of articles
60
40
20
0
2016
2017
2018
2019
2020
2021
2022
year


Histogram of articles publication year [final selection]
Snowball
Final selection
num of articles
6
4
2
0
2016
2017
2018
2019
2020
2021
2022
year

Scores (1st phase)



Scores by content and quality of paper



3.5. Initial extracted data

MODEL	Input data	Output data	Technologies used	Pros (+)	Cons (-)
DNN, classical Deep Neural Network (10 models, the hidden layer is set to three layers, wherein the specified gravity is 1.5, 2.5 and 1.5,)	engine acoustic emission signal: engine speed at 750 rpm (gained by PCB 426E01 microphone)	classified and identified vehicle types	Python, Tensorflow, Anaconda, Numpy, Spyder, scikit-learn, matplotlib, Labview	unsupervised learning method (no need to tag the data)	Focuses more on data preprocessing then modelling, compares 3 methods: time domain, frequency domain and the wavelet transform domain (DWT)
CNN (Convolutional Neural Network): four convolutional layers and four pooling layers, finally leading to a fully connected layer	traditional engine vehicle noise and electric motor noise data presented in the form of visualised sound signals (as a dot pattern = symmetrical point map, snowflakes). The vehicle speed was maintained at 750 rpm.	recognition (classification) of non-electrical and electrical Vehicles (different types of vehicle noise)	Python, Tensorflow	Uses innovative technique - sound as image classification (sound symmetry point pattern)	Only one model, both non and electrical vahicles
Siamese neural network (SNN), used mainly for binary classification tasks (inside 16 convolution layers, so CNN also	classifying vehicles based on audio recordings; features are based on the mel-spectrogram (MFCC);	diesel vs. petrol and heavy goods vehicle vs. personal car classification	-	2 models used together; f1 score highest (>90%) with the usage of HGV/PC dataset	not clear conclusion about models (only claim about the advantage of the proposed approach compared to conventional feature representations and classifiers)

used inside)					
Hybrid - LSTM and CNN, SVM	Features extracted from audio signal (MFCC, PCP, Short-Term Energy)	classified vehicle	-	Results achieved by Hybrid CNN-LSTM model are better than SVM classifier (about 98% vs 92%). Also, only CNN perform worse than CNN-LSTM hybrid (3,39% worse).	-
SVM (Support Vector Machine) and k-NN (k nearest neighbours)	audio data preprocessed using PSD (power spectral density - Fourier transform) and LPC (linear predictive coding)	vehicle classification as a wheeled or tracked one (binary classification)	Python	linear kernel used for SVM, very simple methods. Very good results for SVM (99.5% accuracy) and k-NN (96,4% accuracy)	Method was checked only on the specified subset. After aplying sequential feature selection, the accuracy for both models took 100%. Of course, this result will probably be lower on different data set.
SVM (Support Vector Machine), DNN (Deep Neural Network)	Acoustic data of airplane, on-road(buses, cars, auto-rickshaw and pedestrian) and train. Then windowed and feature extracted using MFCC, wavelet transform (WPT)	Classification into classes: airplane, on-road(buses, cars, auto-rickshaw and pedestrian) and train	-	Usage of normal and noise dataset (better to generalize network). Accuracy up to 97,62%.	Fails to detect then heavy music is played in the vehicle. Classification of also airplanes and trains
CNN with triplet consecutive frames	classified vehicle audio frames 1st model: single frame 2nd model: three consecutive	classified vehicles (heavy car, light car) or silence	Python, Pandas, SciPy, Tensorflow, Scikit-posthocs	Results for triplet input model are slightly better then single input. Easy to implement.	Still not very impresive results (about 80% correctness)

	frames Three different frame lengths: 50, 100, 200, 400ms				
DRNN + LSTM (two recurrent layers with LSTM cells plus a dense layer and a softmax decision)	Concatenation of several spectral features, like MFCCs, mel-scaled spectrogram, chroma and spectral contrast for construction site recordings. Two different frame lengths: 30, 50ms	classified construction site machines (vehicles and tools, 5 classes)	Python, Keras, TensorFlow, CuDNNLSTM, cuDNN	Great result - 97% accuracy (higher then DBN). LTSM resolves vanishing/exploding gradient problem. Exploit the typical time correlations of audio data.	Accuracy not higher then DCNN. Construction site machines may be easier to classify then road vehicles, so the accuracy may be worse.
S-CRNN (CNN + RNN) + LSTM/GRU (3 CNN layers and on RNN layer)	Mel spectrograms out of road vehicle sound recordings	classified vehicles (car, bus motocycle) and components (engine, wheels, compressor, brakes)	Python, librosa, Keras, TensorFlow	It can efficiently pick up local acoustic features. It provides translation invariant convolution in time and space. And it can well extract high-level acoustic features from low-level acoustic features. Enables to mine temporal and semantic information in acoustic data. GRU makes training faster and it's just a little bit worse then LSTM in terms of	High complexity.

				classification.	
PNN (probabilistic neural network)	1/3 Octave and Bark-Scale Wavelet out of road vehicle sound recordings	classified vehicles (truck, bus, motocycle, small vehicle, medium vehicle)	MATLAB, GoldWave	Despite of simplicity of network results are pretty good (90% accuracy) thanks to use of nowel sound extraction methods. Simple model.	Too simple network for our project.
CNN, ConvNet	Denoised MFCC (using Dual-Tree Complex Discrete Wavelet Transform) out of road vehicle sound recordings	classified vehicles (truck, bus, motocycle, car)	-	Good results on simple CNN network (86% accuracy). Can be combined with more advanced neural network. De-noise process made great improvement on the results. Well discribed data preparing process.	Long and complicated Denoising process.
DNN (fully-connected neural network)	original spectrogram image	Transport mode (Still, Walk, Run, Bike, Car, Bus, Train and Subway)	Matlab	Good results 85% accuracy, recognition performance using sound can be improved by optimizing the CNN architecture and by combining with other modalities such as motion and GPS sensors	High complexity
DL neural network	the obtained audio feature vector	Transport mode (bicycles, motorcycles (including scooters), cars (including minibuses), vans (light trucks, up to 3.5 t), small	R (h2o, randomForest, and e1071)	Good in multi-class classi-fication results for the multi-class classifiers for big truck, car, and motorcycle 97%	High complexity

		trucks (above 3.5 t), big trucks (above 3.5 t with trailers or semi-trailers), buses, and tractors (including rollers, excavators etc.))			
SVM, KNN, DT, RF, NB	feature vector (zero-crossing rate and mel-frequency cepstral coefficients (MFCC)	Transport mode (Still, Walk, Run, Bike, Car, Bus, Train and Subway)	Matlab	Classical machine learning, very simple methods	poor accuracy about 62%
-	Audio data preproeceeded using short-time Fourier transform	classified and identified vehicles	Spectral clustering, Incremental Reseeding (INCRES) Algorithm	unsupervised learning method (no need to tag the data)	Identification of similar vehicles, Overtraining possible in our case

4. Conclusions

4.1. SLR process

Systematic Literature Review is a time consuming process, but it is worth the effort if we want to discover what is the current research state in our area of interest and gain valuable answers to our research question.

It is important to formulate the right questions of the SLR and to overthink the goals before proceeding to the main SLR process.

Through reading articles we can also gain more comprehensive knowledge of the topic and find out problems that we hadn't considered before.

4.2. SLR results

MODEL	COUNT
CNN (Convolutional Neural Network)	9
SVM (Support Vector Machine)	7
DNN (Deep Neural Network)	7
LSTM (Long-Short Term Memory Neural Network)	4
RNN (Recurrent Neural Network)	2
k-NN (k Nearest Neighbours)	2
SNN (Siamese Neural Network)	1
PNN (Probabilistic Neural Network)	1
Hybrid: S-CRNN (CNN + RNN)	1
Hybrid: LSTM+CNN	1
Random Forest	1
Decision tree	1
Naive Bayes	1
GRU (Gated Recurrent Unit Neural Network)	1

We found out that the most common models used in vehicle type recognition based on sound are CNN, SVM, DNN and LSTM.
Also, these models are frequently combined with each other to create more advanced network models and achieve better results.
Thanks to our SLR, we also discovered that sound features extraction that goes into network input is as important, or even more important, than the model used.
Some articles focus on creating a complex network and while others focus on best extraction features out of recorded vehicle sound.
In our approach we would like to combine well extracted and denoised feature vector with advanced model to outperform existing networks for vehicle sound classification.

Whole results of our SLR process are included in Google Spreadsheets document:
https://docs.google.com/spreadsheets/d/1187Cnw3jHyhH_7k5SF4xBH34L8GkEzaF/

5. Literature

- Systematic Literature Review, Research methods in informatics lecture, Jakub Miler
- Systematic Literature Review Plan Example, J. Miler, O. Springer