Meeting of the Acoustical Society of America, San Diego, USA, 2-6 December 2019

3pSP2

Czyżewski A.[1], Kurowski A. [1,2], Zaporowski S. [1,2]

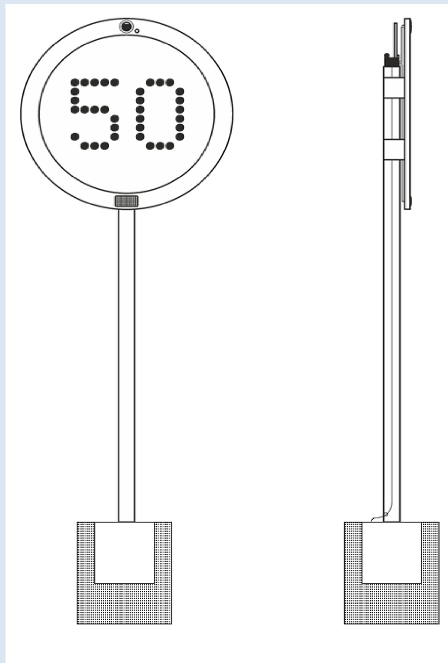
[1]Gdansk University of Technology, Multimedia Systems Department
[2]Gdansk University of Technology, Audio Acoustics Laboratory

Application of autoencoder to traffic noise analysis

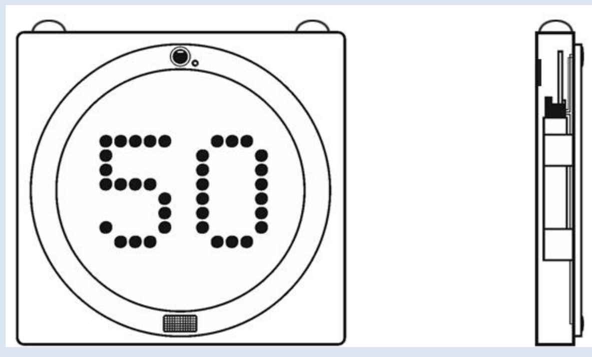
INTELGIENT ROAD SIGN

- The goal of the project is the development, construction and testing of a new type of intelligent road signs that will allow the prevention of the most common collision types on high speed roads resulting from e.g. the rapid accumulation of vehicles.
- As part of the project, a series of products is being developed, including intelligent signs, which can be placed on a mobile stand or can be hung above the road. They display the dynamically updated recommended speed, determined automatically, by the electronic module mounted inside the road sign, enabling multi-modal measurement of traffic intensity employing video and acoustic sensors and weatherstation).

Communication between inteligent road signs and vehicles using V2X Technology



„Standing smart road sign" Patent W.125160

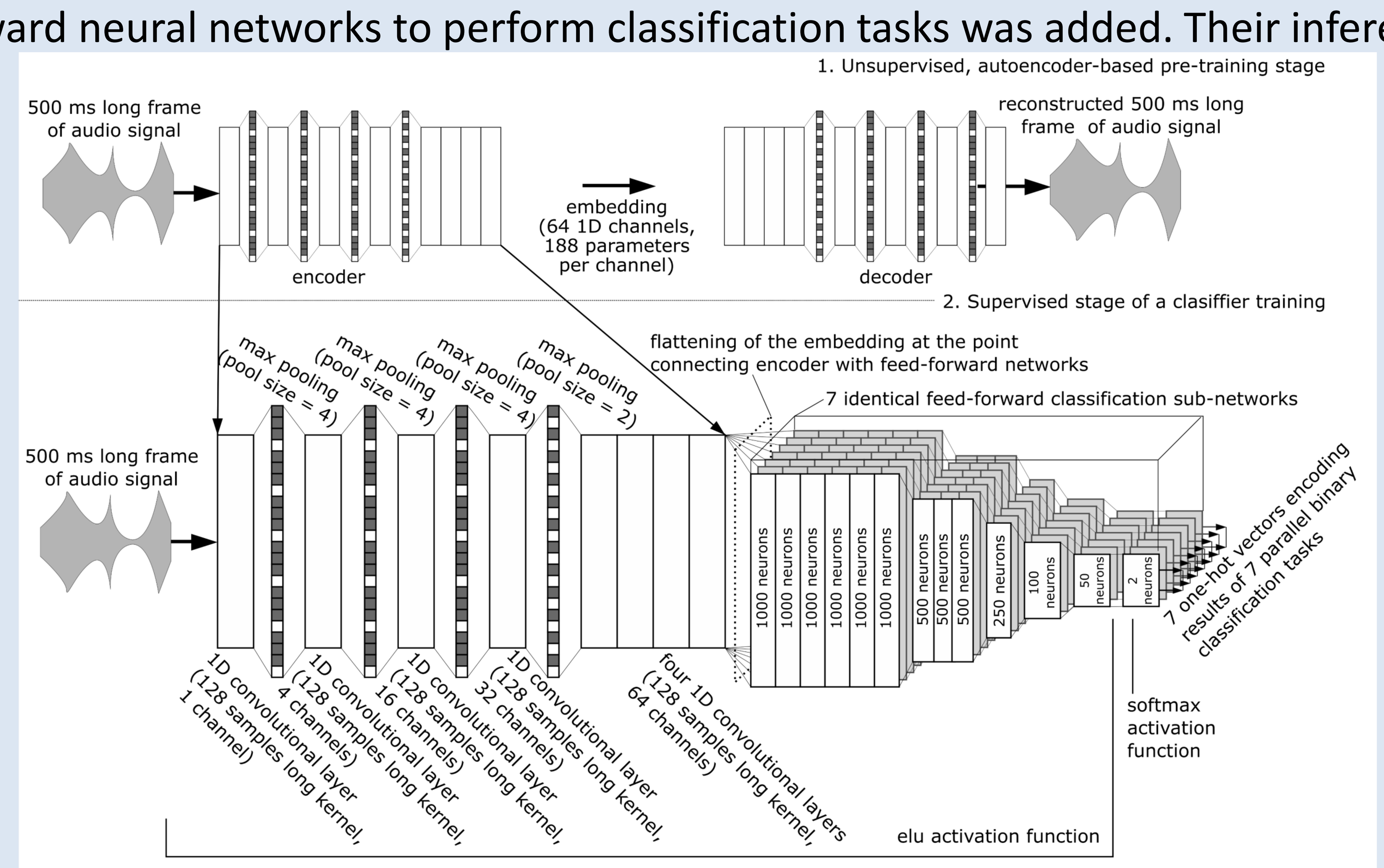


„Hanging smart road sign" Patent W.125159

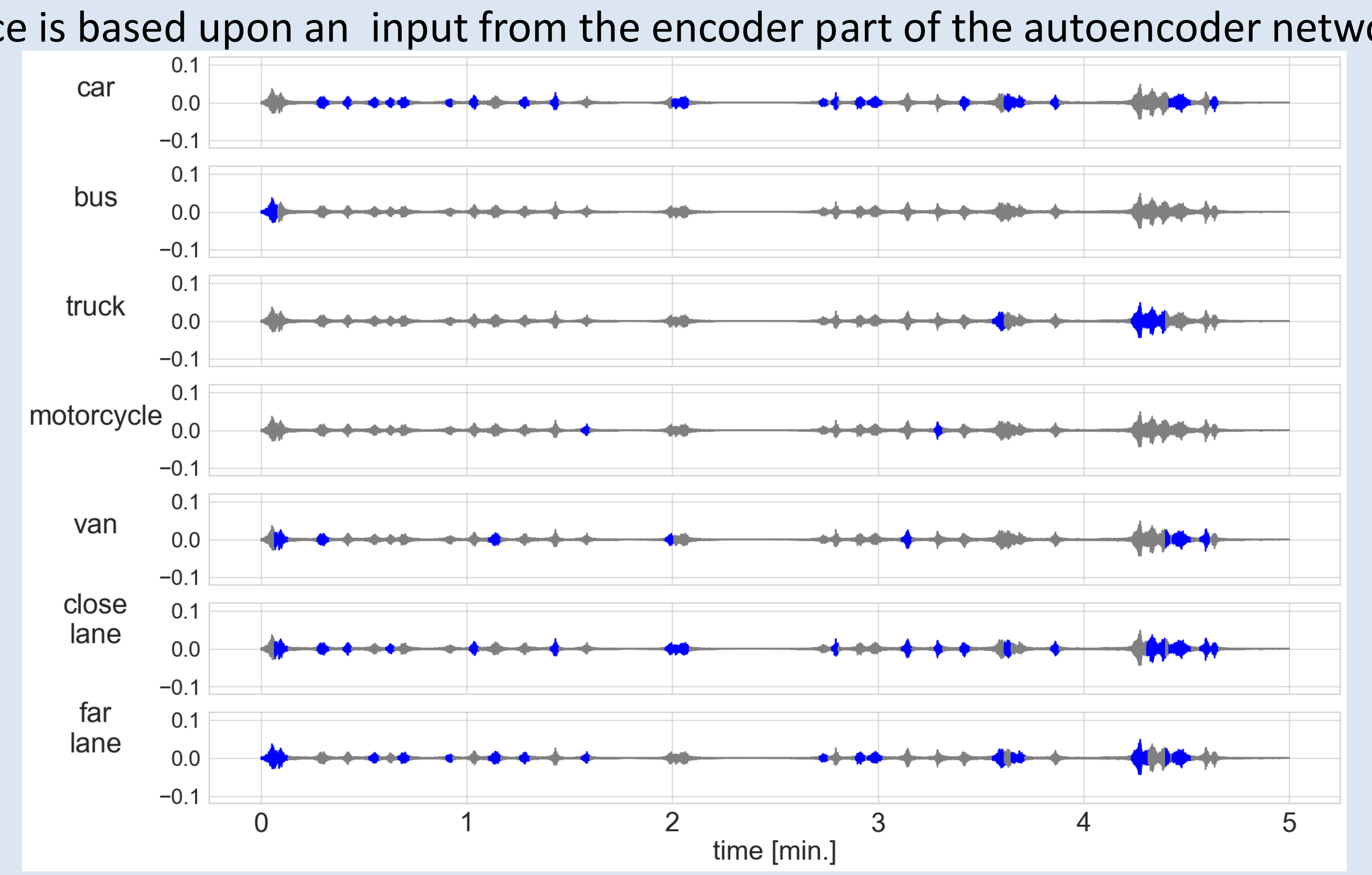
Hanging measurement module of inteligent road sign in Leźno, Poland

DATA ACQUSITION AND PROCESSING

Acoustic signal is gathered by MEMS microphone-based acoustic probe. Moreover, a video signal is captured for the purpose of its use in manual labeling of acoustic data. Seven types of labels are taken into consideration which describe both: the type of the vehicle passing by at the given moment of time and the lanes being occupied at that moment. Eight hours of unlabeled audio was used to train an autoencoder-type 1D convolutional neural network. The encoding part of this network was used as the first part of 1D convolutional network performing classification. Two hours of audio recording were used for the training of the neural network. Frames of audio signals were 500 ms long. To interpret the output of the encoder, an additional set of sub-networks consisting of 7 feed-forward neural networks to perform classification tasks was added. Their inference is based upon an input from the encoder part of the autoencoder network.



Architecture of used neural networks and two stages of training – first one based on use of autoencoder neural network, and second one based on ancoding part of autoencoder and 7 feed-forward classifier networks.



Visual depiction of labels assigned to audio segments from manually anotated dataset with fragments marked in blue. Each label represents a given vehicle type presence or presence of a vehicle on a given road lane.

RESULTS

The result of classification using the presented approach is shown in Table 1. As can be seen, in the table there are 4 classes of 7 discussed before. Since precission and recall score for 3 remaining classes were close to zero, we decided to not include results for them, because such a result indicates poor accuracy. However, in case of detection of cars and trucks, and a vehicle presence on both: a lane which is closer or a lane more distant to the position of a camera, reasonable results of classification were achieved. An autoencoder output allowed training the neural network on both: labelled and unlabelled data.

Table 1. Results of classification using architecture presented earlier

	Accuracy	Precission	Recall	F1 score
car	0.870	0.706	0.703	0.704
truck	0.987	0.136	0.375	0.200
close lane	0.852	0.357	0.405	0.380
far lane	0.852	0.587	0.459	0.515

CONCULSIONS

According to presented results, with our approach it is possible to classify cars with ~ 87% accuracy, satisfying precision and recall, as well with ~70% F1 score. The remaining results are slightly worse, probably due to the fact that classes were unbalanced and most of data was acquired for cars instead of trucks. Despite above difficulties, a distinction between two classes of vehicles, namely: cars and trucks, was possible to make on the basis of acoustical data. The advantage of the proposed approach is the fact, that the use of autoencoder for unsupervised pre-training allowed to utilize unlabelled audio data. Another advantage of the proposed classifier architecture lays in its modularity. The autoencoder-derived part of the network may be reused as a feature extraction module for another neural network performing the classification of vehicles. The frames of raw audio signals acquired in the proximity of a road can be used for both: supervised and unsupervised training of neural networks.

The project entitled: "INZNAK: Intelligent Road Signs with V2X Interface for Adaptive Traffic Controlling„ (No. POIR.04.01.04-00-0089/16) is subsidized from the European Regional Development Fund by the Polish National Centre for Research and Development (NCBR)