Investigations on the implementation of an acoustic rain sensor system

Kevin Hock[1], Julian Götz[1], Mario Seideneck[1], Christoph Sladeczek[1]
[1] *Fraunhofer Institute for Digital Media Technology IDMT, 98693 Ilmenau, Germany, kevin.hock@idmt.fraunhofer.de*

Abstract
The most common method for measuring precipitation is based on the collection of rain in a defined container, where the amount of water accumulated is read off after a time interval. A major disadvantage of this method is that the rainfall amounts measured in this way always represent an average over time, and do not provide data on the course of precipitation with high temporal resolution. Since time immemorial, meteorological elements have been described colloquially by their generated sounds. Following on from this, the aim of this study was the implementation of an acoustic rain sensor. For this purpose, acoustic data were generated in a laboratory setup by a simplified sprinkler system, with which a machine learning algorithm was trained.

Introduction
According to a recent report of the United Nations (UN) the climate change will cause a significant increase in natural disasters and extreme weather events [1]. Considering the prognosticated global warming, in the year 2100 we would face a four to five-fold increase in extreme weather events compared to 2022. Because of the drastic effects of such events on both humanity and infrastructure, the UN underlines the importance of the expansion of early warning systems [2]. Besides this, in virtue of the Deutscher Wetterdienst (DWD) 80 % of the variance of crop yields in Germany are referable to the weather. Hence the prevention of damages due to the weather, as well as the management adopted to it, is essential for agriculture and forestry [3]. A challenge in weather forecasting today is accurately determining local weather impact. One missing element is accurate spatially and temporally highly resolved local precipitation data. These data cannot be determined by systems in use today, such as precipitation radars and ombrometers, for a variety of reasons [4].

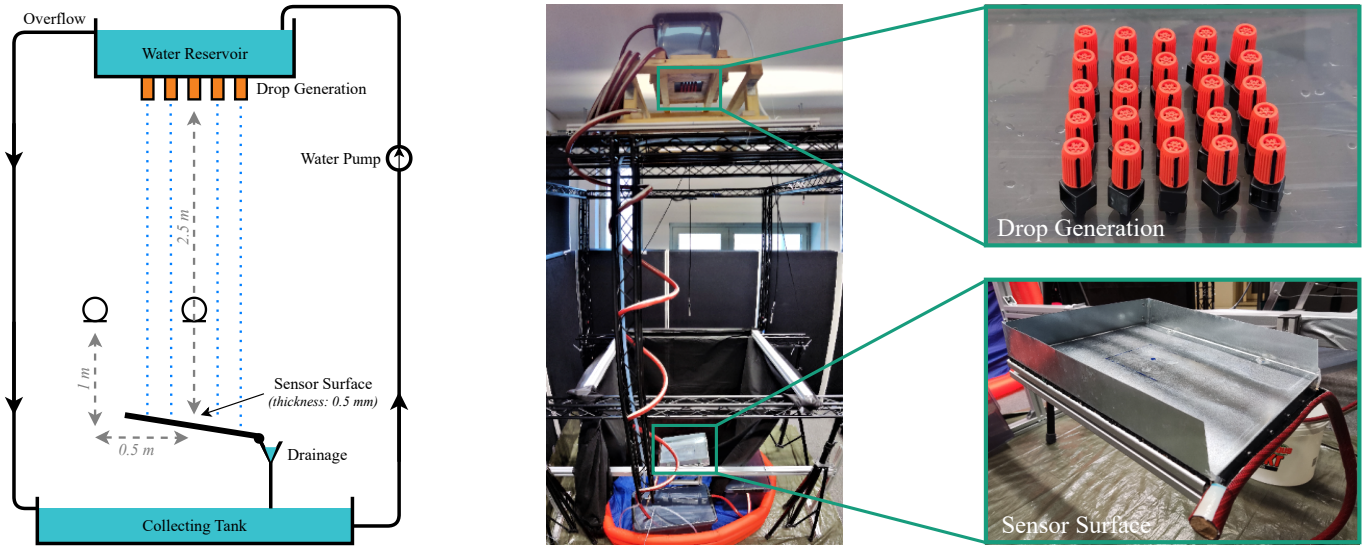
Since rain causes distinctive sounds on different surfaces, the idea is to use acoustic sensor technology to measure precipitation. Acoustic sensor technology could bring the advantage of being relatively energy efficient to operate and inexpensive to implement. Such rain sensor could be used for the early detection of heavy rain in the sense of an expansion of early warning systems, or as an exploration and real-time display of precipitation. In this paper, a proof of concept experiment is described for the use of an acoustic sensor for rain measurement.

Rain simulation
The purpose of the experimental setup, depicted in Figure 1, was to generate acoustic data for a machine learning algorithm, classifying different precipitation intensities. For this a self constructed sprinkler system produces water droplets, falling on the rain sensors surface, where the impact sounds were recorded.

Acoustic rainfall simulator
The sprinkler system consists of a water reservoir connected to a drop generator. The generator consists of 25 drippers (arranged in squares, 2 cm distance), ensuring a stochastic behaviour in dripping. Each dripper can be infinitely adjusted in its flow rate, producing droplets with a constant diameter of round about 5.75 mm. This diameter is represented in precipitation intensities from 25 mm/hr up to at least 150 mm/hr [5]. The generated water droplets are falling on the sensor surface made out of sheet steel (length 37 cm; width 30 cm; thickness 0.5 mm). The area on which the water drops appear is approximately 100 cm^2 large. The distance between the sensor surface and the drop generation is 2.5 m. This is the travel at which the average raindrop (> 0.5 mm diameter) reaches nearly its maximum speed of round about 5.8 m/s [6][7]. The sensor surface is tilted with an angle of 10 degree. This inclination was seen as a compromise between drainage speed of the water and non adulteration of the impact sound. To maintain a constant water level in the water reservoir, water is steadily pumped up from the collecting tank, creating an excessive water supply. This is necessary to guarantee an even water pressure on the drop generation, resulting in a time consistent water drop speed. At the lower edge of the sensor surface, there is a drain back to the collecting vessel to avoid dripping noises. Fabric spanned around the sensor surface prevents noise from water splashes.

Recording setup
The sound of the rain drops impacting on the sensor surface is recorded by two Microtech Gefell MK 221 microphones with pre-amplifier MV 212. The microphones are arranged on the back and the left side of the surface. They are positioned at a distance of 0.5 m from the middle of the surface in a height of 1 m. This was necessary to prevent water splashes from reaching the microphones. The remaining setup consists of an G.R.A.S. Power Module Type 12AQ, a measurement interface from HEIM (PWAC, DIC6B and LMF2FE) and the recording software siRecord from Soundtec. The microphones were calibrated at 1 kHz, 94 dB(SPL) and samplerate of 96 kHz was used.

System properties
The prototype implementation of the rain simulator comes with limitations. However, as the goal was to create different rain intensities, a reference for categorization was needed. Therefore, the categories *light rain*, *moderate rain*, *heavy rain* from the classification by the American Meteorological Society was used [8]. A fourth category (*extreme rain*) was added to increase the vari-



(a) Schematic representation of the experimental rainfall simulator.

(b) Complete view of the experimental setup on the left. Detailed view of the drop generation and the sensor surface on the right.

Figure 1: Overview of the experimental setup, including the sprinkler system, the rain sensor surface and part of the measurement setup.

ability of the dataset and because of the limited adjustment range of the dripper. The range covered by each category is depicted in Figure 2. The mentioned limitations are mainly caused by two problems. Due to manufacturing inaccuracies of the dripper, the drop frequency is subject to certain error tolerances when setting the dropper manually. This error is disproportionately high at low drop frequencies, while high flow rates can be set with a relatively low error tolerance. Considering this, for each precipitation category one similar named rain setting with a certain rounded precipitation rate (25, 43, 82 and 376 mm/h) is derived.

Rain intensity classification

For the classification of rain intensities, it is necessary to process the acoustic records in order to create datasets for the usage of the machine learning algorithm.

Data preparation

Acoustic recordings of 11 hours were made for each rain setting and subsequently cut into segments with a length of 10 seconds. Equivalent to this, a 10 seconds record of the trial-related noise was made before each rain recording. This noise was caused for example by the water pump. For the segments of the rain and noise recordings, a multiband onset detection based on the spectral flux was performed. After subtracting the detected onsets in the noise signal from those of the respective rain segment, the contained raindrops could be quantified. When at least two drops were contained in a segment, it was included in the source dataset (`SD`). Figure 2 shows the distribution within the resulting `SD`, consisting of 3,750 segments per rain setting and 15,000 segments in total.

FSD50K

The FSD50K is a human labeled open dataset, consisting of 51,197 audio clips from freesound.org of common physical sources from everyday situations (*Human sounds, Sounds of things, Animals, Natural sounds and Music*). These are represented by 200 classes, drawn from the AudioSet-Ontology. The recordings are provided as uncompressed PCM 16-bit 44.1 kHz mono audio files [9]. For testing the real world noise resistance of our classification system, 1,918 clips that had at least a duration of 10 seconds, were extracted from the FSD50K. All labels of the selected clips are part of 43 classes, that were chosen from the FSD50K-200-class-ontology. These classes were meant to represent the environmental noise of a rural village. The 43 classes can be divided into four main classes (*Human sounds, Sounds of things, Animals and Natural sounds*). *Human sounds* represents the speech of single persons and small groups. *Sounds of things* contains the noise of means of transportation and travel. *Animals* represents the noise of domestic, farm and wild animals. *Natural sounds* covers wind and sea noise.

Dataset generation

Two datasets, named `A` and `B`, were created (see Figure 2). For `A` the obtained `SD` was split 60 % into a training subset $\mathtt{A}_{\mathtt{Training}}$ and 20 % each into a validation and test subset, $\mathtt{A}_{\mathtt{Validation}}$ and $\mathtt{A}_{\mathtt{Test}}$. The subset $\mathtt{A}_{\mathtt{Test}}$ was overlaid with the extracted clips from the FSD50K dataset at SNR intervals of 0 to 10 dB and 10 to 20 dB, leading to $\mathtt{A}_{\mathtt{Test,SNR\_10\text{-}20dB}}$ and $\mathtt{A}_{\mathtt{Test,SNR\_0\text{-}10dB}}$. This generalized test data was used to consider the classification performance under the influence of simulated environmental noise.

As it was not possible to completely avoid interfering noise during the rain recordings, the `SD` was filtered with the use of spectral gating, resulting in dataset `B`. This denoising method estimates a threshold based on the noise spectrogram to determine a spectral mask. Due to the roughly steady background noise, an stationary approach was chosen. This stationary spectral gating maintains a constant threshold over the entire audio signal [10]. As

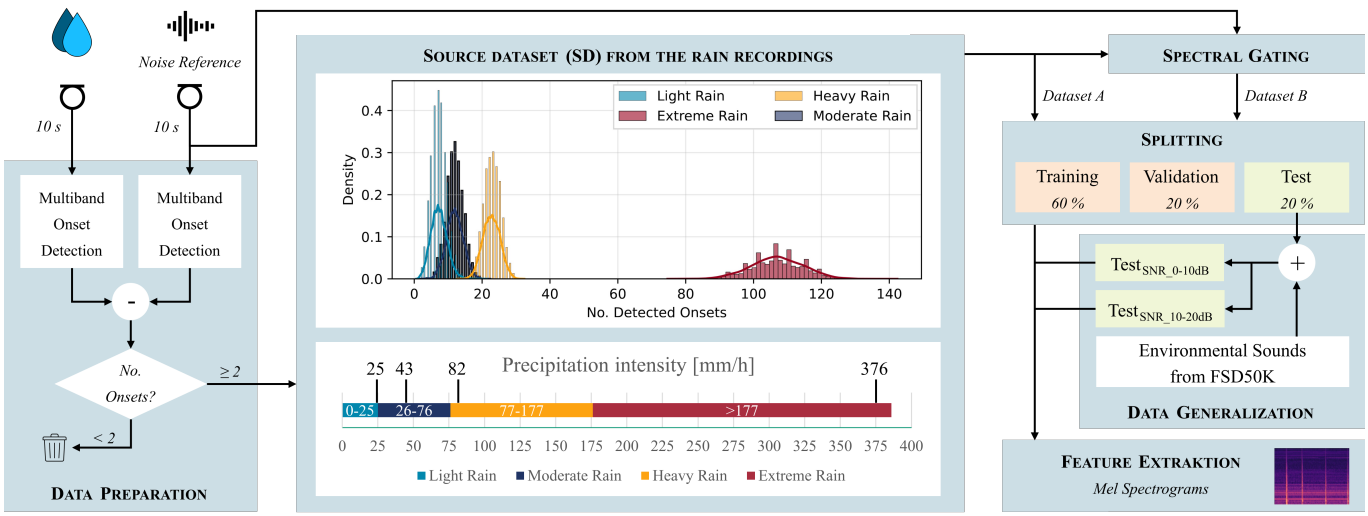


Figure 2: The recordings of the rain settings were cut into 10 second segments and processed by an multiband onset detection. If at least two onsets resp. drops are in a segment, it is inserted to the source dataset (SD). After splitting, the test subsets were generalized by environmental noise. Finally, mel spectrograms of all subsets were extracted as features for the classifier.

noise reference, the same noise recording, as the one for filtering in the onset detection, was used. According to dataset `A`, dataset `B` was split into $\mathtt{B}_{\mathtt{Training}}$, $\mathtt{B}_{\mathtt{Validation}}$ and $\mathtt{B}_{\mathtt{Test}}$ as well as $\mathtt{B}_{\mathtt{Test,SNR\_10\text{-}20dB}}$ and $\mathtt{B}_{\mathtt{Test,SNR\_0\text{-}10dB}}$.

Feature extraction and classification method

To utilize the `SD`, mel spectrograms in the frequency range of 200 to 20,000 Hz were extracted from the audio segments of dataset `A` and `B` . The motivation for this lower cutoff frequency is that in a real application of such precipitation sensors, wind and gusts would have an energetic effect on the low-frequency range.

A convolutional neural network (`CNN`) was trained, using the extracted spectrograms. The network consists of three two-dimensional convolutional layers, each with 2D max-pooling, followed by a flatten layer and three dense layers (see Figure 3a). Dropouts of 0.25 were inserted after the first and second dense layers to reduce potential over fitting. ReLU along with SoftMax activation was chosen to classify the four rain intensities. The four rain intensities were classified using ReLU activation along with SoftMax, except for the last output dense layer where SoftMax was not used. Early stopping was used to determine the training epochs, with training terminated after minimizing the validation loss.

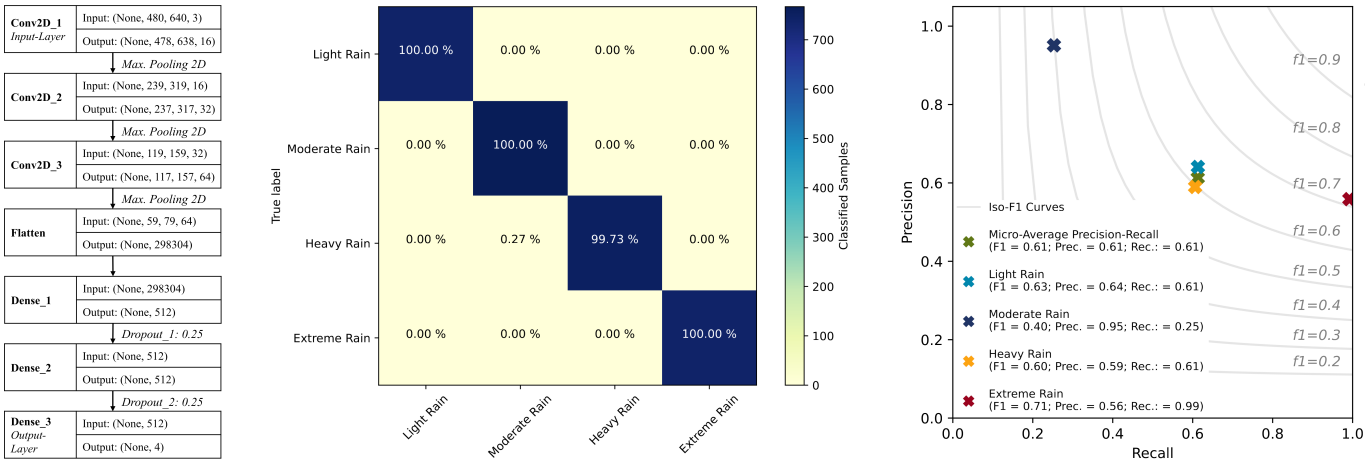
Evaluation

Depending on the intended application of a rain sensor, various metrics can be used for evaluation. Recall is relevant for monitoring the rainfall, since it describes the percentage of a class that was recognized, while precision is more appropriate for focusing on the share of correct classifications within a class. For the following evaluation, the harmonic mean of these two metrics, the F_1-score is considered. First, we consider the classifier $\mathtt{CNN}_{\mathtt{A}}$, which was trained on dataset `A`. It can be seen that the classification on the cleaned test subset achieves approximately 100 % correct classifications across all classes (see Figure 3b). From the confusion matrix it can be derived, that precision, recall and the F_1-score rounded are equal to 1.00 for all perception intensities.

Applying the classifier on the generalized test subset $\mathtt{Test}_{\mathtt{A}}$ showed a degradation of the classification performance. While the classes *light rain*, *heavy rain* and *extreme rain* continue to achieve over 80 % of correct classifications, the performance of *moderate rain* drops to 54.37 %. Table 1 illustrates the impact by a recall of only 0.54 for this class. If the SNR decreases to an interval of 0 to 10 dB, a reliable classification of *moderate rain* is no longer possible. The corresponding recordings were attributed to all classes by the classifier with about 25 % each, which can also be seen in the Precision-Recall-Curve (see Figure 3c) by a recall of 0.25 and an F_1-score of 0.4. The reason for the above-average impairment of this class, compared to the other precipitation intensities could be the overlapping of *moderate rain* with *light rain* and *heavy rain* with respect to the drop frequency and number of drops. *Light rain* and *heavy rain* show a shift in classification towards higher rain intensities with lower SNR. For instance, 61.30 % of *light rain* is classified as *light rain*, 0.27 % as *moderate rain*, 15.56 % as *heavy rain* and 22.87 % as *extreme* rain with a SNR between 0 and 10 dB. This can be explained by percussive signal components contained in the added environmental noise.

Class	$\mathbf{CNN_A}$ SNR: [10; 20] dB			
	TP	Precision	Recall	F_1
Light Rain	85.52 %	0.77	0.86	0.81
Moderate Rain	54.37 %	0.97	0.54	0.70
Heavy Rain	82.10 %	0.76	0.82	0.79
Extreme Rain	100.00 %	0.79	1.00	0.88
Micro-Average		*0.80*	*0.80*	*0.80*

Table 1: Classification results on the generalized test data of the classifier, which was trained on dataset A.



(a) Neural network architecture.

(b) Confusion matrix of $\texttt{CNN}_\texttt{A}$ applied on subset A_{Test} without ambient noise.

(c) Classification results of $\texttt{CNN}_\texttt{A}$ evaluated with subset $A_{Test,SNR\_0\text{-}10dB}$.

Figure 3: Architecture of the used neural network (a) and the resulting classification performance without (b) and with added environmental sounds from the FSD50K dataset with SNR from 0 to 10 dB (c).

Comparing the performance of the classifier ($\texttt{CNN}_\texttt{B}$), which was trained on dataset B, a similarly high classification is achieved on the clean test data. However, if data generalization is used, this is reflected in lower F_1-scores, independent of SNR interval, compared to those of the $\texttt{CNN}_\texttt{A}$. The only exception is the *moderate rain* class (see Tabular 2).In summary, stationary spectral gating worsens this application.

Class	F_1-Scores			
	SNR: [10; 20] dB		SNR: [0; 10] dB	
	CNN_A	CNN_B	CNN_A	CNN_B
Light Rain	0.81	0.88	0.63	0.78
Moderate Rain	0.70	0.47	0.40	0.37
Heavy Rain	0.79	0.59	0.60	0.43
Extreme Rain	0.88	0.72	0.71	0.60
Micro-Average	*0.80*	*0.68*	*0.61*	*0.57*

Table 2: Classification results on the generalized test subsets of both classifier.

Conclusions

The implementation of the prototypical rain sensor was done using a dataset generated in a sprinkler system. Their recordings were passed to the feature extraction and finally to a CNN classifier. The results on the clean test data exhibit a high precision, recall and F_1-score of rounded 1.00 across all intensity classes. Once environmental noise was overlaid on the test data, performance decreased, especially for the *moderate rain* class. The other rain intensities were still correctly classified at least with over 60 %. In an additional experiment spectral gating was used for denoising, which led to a decline of the classification performance. This proof of concept indicates that a good differentiation of rain intensities can be achieved. Nevertheless, these are only tendencies, since e.g. only one raindrop diameter could be realized with the sprinkler system. Therefore, further investigations with a more realistic sprinkler system are necessary.

References

[1] United Nations, "The sustainable development goals report-2022," 2022.

[2] UNDRR and WMO, "Global status of multi-hazard early warning systems: Target G," 2022.

[3] Deutscher Wetterdienst. Land- und Forstwirtschaft. [Online]. Available: https://www.dwd.de/DE/fachnutzer/landwirtschaft/landwirtschaft_node.html

[4] M. Graf, J. Polz, and C. Chwala, "Regenmessung im Mobilfunknetz," *Physik in unserer Zeit*, vol. 52, no. 2, pp. 88–93, 2021.

[5] J. O. Laws and D. A. Parsons, "The relation of raindrop-size to intensity," *Eos, Transactions American Geophysical Union*, vol. 24, no. 2, pp. 452–460, 1943.

[6] W. Kassera, *Motorflug kompakt: das Grundwissen zur Privatpilotenlizenz*. Motorbuch-Verlag, 2012.

[7] P. K. Wang and H. R. Pruppacher, "Acceleration to terminal velocity of cloud and raindrops," *Journal of Applied Meteorology (1962-1982)*, vol. 16, no. 3, pp. 275–280, 1977.

[8] American Meteorological Society. Rain, glossary of meteorology. [Online]. Available: https://glossary.ametsoc.org/wiki/Rain

[9] E. Fonseca, X. Favory, J. Pons, F. Font, and X. Serra, "FSD50K: an open dataset of human-labeled sound events," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 829–852, 2022.

[10] T. Sainburg, M. Thielk, and T. Q. Gentner, "Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires," *PLoS computational biology*, vol. 16, no. 10, p. e1008228, 2020.