Computer vision applied to detect normal or abnormal sound-based uroflowmetry signals

Mario Jojoa [1] , Laura Arjona [2] , Marcos L. Alvarez [3] and Alfonso Bahillo [4]

1,4 University of Valladolid, Valladolid Spain. mariofernando.jojoa@uva.es, alfonso.bahillo@uva.es
2,3 University of Deusto. Bilbao Spain. laura.arjona@deusto.es, alvarez.marcoslazaro@deusto.es

Aim of the Study/Purpose

The aim of our study is to demonstrate the higher capability of deep learning algorithms to classify sound-based uroflowmetry signals. We have used inception v3 algorithm mixed with the representation of the sound signals as images using the Fourier spectrogram technique [1], opening the possibilities to apply complex computer vision algorithms for urologic diseases detection.

Sounds Adquisition using a smart watch → Records organization and data storage → Sounds classification using a trained inception v3 deep learning model → Performance evaluation of the trained model

Study Design

Leavening consumer technology such as smartwatches to objectively and remotely asses people with voiding dysfunction could allow healthcare professionals to deliver more personalised and effective care at home with less waste of time and resources. To face this challenge, we have developed an artificial intelligence algorithm that uses the recorded sounds of male human urination using a smartwatch at home. The dataset is composed by 153 audio signals, where 106 correspond to normal flow and 47 to abnormal flow. The sounds were labeled by one human expert, who identified the differences in the flows using a flowmetric exam as a pattern. The results obtained in the present work are compared with the results published in [2]. Using our proposed approach, the metric accuracy (ACC) was improved in 3.65 percentage points, and the metric Area Under the Curve (AUC) was improved in 3.46 percentage points. The next sections describe in detail the materials and methods used in our study. We have decided to select those two metrics since the ACC gives information about the performance of the inception model in terms of the number of sounds classified correctly, without considering the distribution of the inputs as normal or abnormal; and the AUC provides a performance metric across all possible classification thresholds (normal or abnormal classes)

Materials

The dataset is composed of 153 voiding audios from Oppo smartwatches, across 14 study participants. The labelling process was performed according to the urologist author of [3]. The labels were assigned as it is described in [2]. As a result, 47 sounds were labeled as abnormal flows, while the remaining 106 are labeled as normal flows. Abnormal flows account for approximately 30% of all recordings. This information is outlined in Table 1, where 0 represents normal flows, and 1 represents abnormal flows. This table 1 also presents the distribution of trials (number of audio recordings) for each participant.

Table 1 Distribution of urologic sounds per participant and their labels [2]

User	Trials	0 #	1 #	User	Trials	0 #	1 #
A2	12	9	3	B5	12	7	5
A3	14	5	9	B7	14	9	5
A4	15	8	7	B8	7	3	4
A5	9	8	1	B9	8	4	4
A6	12	12	0	B10	3	2	1
A7	15	11	4	B11	9	9	0
B3	16	12	4	B12	7	7	0

The complete used dataset is available in https://github.com/DeustoTech/UroSound

Methods

To perform the automatic classification of the sound-based uroflowmetries, we have implemented a computer vision model called inception v3. This model is a deep neural network that was created and used for general purpose image classification. To give the mandatory inputs attributes needed for this algorithm, such as 2D shape, number of channels and tensor form, we have decided to calculate per each sound-based signal their respective spectrogram using the Fast Fourier Transform algorithm (FFT) with 512 points. The number of points was selected to obtain a good resolution in the frequency domain, that is needed to build the spectrograms and to avoid losing information in this step. The inception model was configured as a classifier using a pretrained structure, which was trained using the Imagenet repository. The hyper parameters were set up using the grid search method, where different values were searched in an arbitrary set of possible values. Finally, based on the number of samples available for training and validation (153 values in total) we have used a three-fold cross validation technique to eliminate the subjective behavior of the training task, such as the overfitting phenomena. The results presented in the next section are the average values obtained in each fold.

Results

In Table 2 we present a comparative analysis of the proposed model performance in terms of AUC and accuracy and the results presented in [2]. We found that our algorithm improved the accuracy and AUC metrics by 3,38 and 3,81 percentage points, respectively. Besides, our model does not use a preprocessing stage since it uses the spectrograms of raw sound-based uroflowmetries as inputs, without a stage of envelope extraction or another feature extraction technique.

Table 2 Performance comparison between models in classification task

Model	AUC	Accuracy
Inception V3 + Spectrogram	0.9300	0.8954
Ensemble method	0.8919	0.8616
Regression Forest	0.8754	0.8562
Random Forest	0.8476	0.8403

Interpretation of results

The results show a novel milestone for the automatic sound-based uroflowmetries classification. The computer vision algorithm, inception v3, mixed with the representa-tion of the sound-based signals as Fourier spectrograms performs better compared to the performance of the classical machine learning algorithms presented in [2], for the same uroflowmetry classification task. The improvement in the identification process per each sound-based signal (normal or abnormal classes) is represented for 3.65 per-centage points higher accuracy using our approach, compared to the best value ob-tained for the ensembled method. On the other hand, the metric AUC was improved using our approach in 3.46 percentage points, compared to the best configuration ob-tained for the ensembled method as well. This improvement was reached using raw sound-based uroflowmetries without any preprocessing stage, enabling the use of more complex deep learning structures in the future that use the same approach for this task.

Concluding message

The proposed approach opens the possibility to use computer vision algorithms for sound-based uroflowmetries classification without using a preprocessing stage. This contribution opens the doors to use advanced (state of the art) computer vision algorithms without complex and subjective feature extraction process, to perform an automatic diagnostic task, helping health care professionals to identify possible abnormalities in the human urination process. Besides, these algorithms could be used to execute massive screening programs, only using a low-cost smartwatch, without the need to require the people to go to the hospital. Our approach could be a solution of e-health and tele diagnostic current challenges.

Institutional *Review* Board Statement

The study was conducted according to the guidelines of the Declaration of Helsinki.

Informed Consent Statement

Informed consent was obtained from all subjects and family members involved in the study.

References

[1] GELPUD, John, et al. Deep Learning for Heart Sounds Classification Using Scalograms and Automatic Segmentation of PCG Signals. En Advances in Computational Intelligence: 16th International Work-Conference on Artificial Neural Networks, IWANN 2021, Virtual Event, June 16–18, 2021, Proceedings, Part I 16. Springer International Publishing, 2021. p. 583-596.

[2] NARAYANSWAMY, Girish, et al. Automatic classification of audio uroflowmetry with a smartwatch. En 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2022. p. 4325-4329.

[3] L. Arjona, L. Enrique Diez, A. Bahillo Martinez, and A. Arruza Echevarria, “Urosound: A smartwatch-based platform to perform non-intrusive sound-based uroflowmetry,” IEEE Journal of Biomedical and Health Informatics, pp. 1–1, 2022.