CS442 Mobile Computing, Networking & Applications
Essay #3
PDVocal: Towards Privacy-preserving Parkinson's Disease Detection using Non-speech Body Sounds

20160042 Inyong Koo

The paper presents *PDVocal*, a private-preserving Parkinson's disease (PD) digital biomarker via non-speech body sounds passive sensing. Researches about mobile health are always inspiring, and I feel grateful that people are working on PD detection since the early diagnosis of PD is crucial. Though I believe there are other ways to cope with privacy impairment, I liked how they approached using non-speech sound. The authors have reasonably inferred that the non-speech body sounds can be a PD biomarker with remarkable intuition. This research expands a whole new domain and possibilities of mobile PD detection.

Another thing I think highly of this paper is their consideration of privacy. Existing mobile PD detectors use voice monitoring, which involves pressing concerns about user's privacy. The authors acknowledged this issue and came up with a solution. Their motivation is faultless; I believe this research is what people need and deserve.

However, I am a little doubtful regarding the reliability. The experiment results were impressive, showing comparable accuracy with the state-of-the-art methods. But to me, the dataset seemed so small and biased. The fact that most of the participants being white, well-educated U.S. citizens was rather disappointing. [1] The authors said they addressed the small dataset problem by using residual architecture (ResNet[2]). However, to my best knowledge, a network can still overfit while using residual architecture.

Speaking of network architecture, I was excited to learn about MobileNets[3] and would like to read more about its usage. I'd also like to see the result for utilizing DenseNet[4] instead of ResNet since DenseNet is better at reducing overfitting and preserving original features compared to ResNet.

Also, I believe the authors missed an important criterium on evaluation - Power consumption. One of the key aspects of this research is passive sensing. *PDvocal* continuously monitors user's phone usage and extracts passive non-speech body sounds. Mobile devices have limited battery, and thus we should consider how much power the module would consume. Similarly, they've commented about extra memory caused by their module but did not mention the data usage amount through the wireless network. If the application runs in the background and spends resources, we need to learn how costly it is.

In summary, I'm interested in audio processing and machine learning, and I enjoyed reading this paper. It was thrilling to see my field of interest applied to mobile health with great intuition and cause. However, I had some concerns regarding reliability and efficiency. With all respect, I hope the future works would supplement those issues.

---

[1]90% of participants were white race, 93.8% of participants received higher education. All participants come from the U.S.

[2]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[3]A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

[4]Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.