Evidential Deep Learning for Uncertainty-Aware Mobile Health

TONG XIA, JING HAN, LORENA QENDRO, CECILIA MASCOLO.
UNIVERSITY OF CAMBRIDGE.

Background. Ubiquitous smartphones and massive mobile sensors, with the advance of artificial intelligence (AI), are reshaping public health. Through deep learning, preferable remote disease diagnosis and mobile health monitoring performance have been achieved in the literature. However, despite the great potential and extensive benefits of AI-enabled healthcare systems, there are still concerns about their performance in real-world settings [3]. This is largely because traditional deep learning can be poorly-calibrated, yielding overconfident predictive probabilities which cannot reflect the true confidence.

Goals and Challenges. To enable reliable remote diagnosis and to mitigate the risk of misdiagnosis in mobile health scenarios, incorporating well-calibrated predictive uncertainty which allows the models to admit what is unknown is of crucial importance. Yet, it is challenging due to many reasons. First, health datasets are usually small-scaled with skewed class distribution, i.e., with abnormal conditions having fewer training samples. Such data imbalance increases the difficulty of training a fair and reliable diagnostic model. In addition, due to the health data collection hurdles including annotation costs, datasets for many tasks hardly ever get complete coverage over a domain. Previous work has found that deep learning behaves unpredictably on unfamiliar data [4]. This could have profound effects in the mobile health context: users may upload arbitrary and unqualified data to the diagnostic model through their personal mobile devices. To avoid misdiagnosis from unexpected inputs, it is required that the predictive uncertainty should be out-of-distribution (OOD) aware. Among massive uncertainty quantification methods [1], evidential deep learning (EDL) [2] has been recognised as the new state-of-the-art in many applications. Differing from traditional Softmax-based deep neural networks, EDL learns a distribution over the class probabilities and quantifies classification evidence in the latent feature space, which allows the models to capture what is unknown from training. EDL is also efficient for mobile health: it only needs a single model and a single forward pass to quantify the uncertainty. Nevertheless, EDL implicitly assumes an equal occurrence of classes, while the effectiveness of the estimated uncertainty for health diagnosis in the presence of sparse and imbalanced medical data remains under-verified.

Our Solution and Contributions. In this abstract, we present our recent study exploring EDL in the context of healthcare, and the proposed *DirichNet*, a framework that enables health-focused deep learning models to output uncertainty estimations alongside the predictions. Specifically, in this framework, existing advanced health models such as DenseNet-based clinical image classifier, with customised data processing pipelines, can be leveraged as feature extractors which are further enhanced by a density estimation module to map the provided features into classification evidence. The evidence will then be translated into a Dirichlet distribution representing the model output. To tackle the data sparsity and imbalance, during training the class distribution is balanced through data augmentation. Additionally, a novel re-calibration method is applied to the held-out validation set to further boost the performance of minority classes, which are usually abnormal classes in health applications.

We evaluate *DirichNet* on three different mobile health tasks including sound-based respiratory condition screening, image-based skin lesion detection, and psychological signal-based heart attack prediction. Through extensive experiments, we find that the original EDL cannot achieve accurate health diagnosis due to the imbalanced data, but data augmentation can help to improve the performance. Furthermore, with the proposed re-calibration method, *DirichNet* surpasses existing uncertainty baselines by 1.8% and 8.2% in terms of diagnosis performance and confidence estimation for in-distribution data. It also improves the near and far OOD detection performance in terms of AUROC by around 7.6% and 10.5% against the state-of-the-art methods, respectively. Our insightful analysis paves the way to trustworthy health deep learning deployment in the real world.

REFERENCES

[1] J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, et al. A survey of uncertainty in deep neural networks. *arXiv preprint arXiv:2107.03342*, 2021.

[2] A.-K. Kopetzki, B. Charpentier, D. Zügner, S. Giri, and S. Günnemann. Evaluating robustness of predictive uncertainty estimation: Are dirichlet-based models reliable? In *Proceedings of the 33th International Conference on Machine Learning (ICML)*, pages 5707–5718, 2021.

[3] C. Park, A. Awadalla, T. Kohno, and S. Patel. Reliable and trustworthy machine learning for health using dataset shift detection. In *Proceedings of the 35th International Conference on Neural Information Processing Systems (NeurIPS)*, pages 1–13, 2021.

[4] T. Xia, J. Han, and C. Mascolo. Benchmarking uncertainty quantification on biosignal classification tasks under dataset shift. *In International Workshop on Health Intelligence W3PHIAI-22 (colocated with AAAI22)*, 2022.