MASCOT-NUM 2022
Clermont Ferrand, France
7-9 June 2022

Dealing with confusing samples into learning based model applied to image classification

J. Xie
Université Clermont-Auvergne

Supervisor(s): T. Chateau (Université Clermont-Auvergne), V. Antoine (Université Clermont-Auvergne)

PhD expected duration: Sep. 2019 - Sep. 2022

Address: 1 rue de la Chebarde 63178 Aubiere France

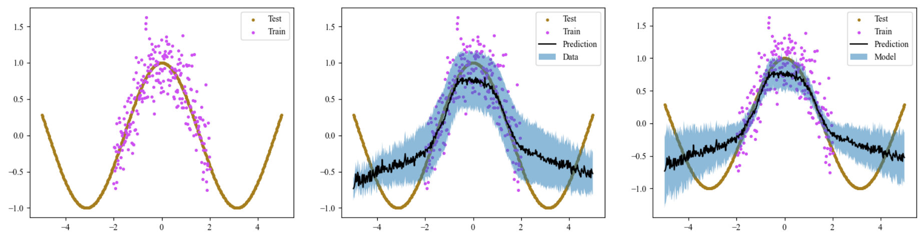
E-mail: jia_rui.xie@limos.fr

Abstract:

The open world is composed by a continuous and evolving environment which requires an associated learning based model that can continuously be updated and robust to confusing samples. It is significant for the model to detect such samples instead of wrongly classifying them to one of the training classes. The first autonomous driving accident in human history occurred in 2016 when a Tesla vehicle hit a white truck turning left at full speed [2]. The post-incident investigation revealed several reasons, one of which was that the autopilot system recognized the white truck as a cloud in the sky [2]. Similar to the Tesla accident, a Uber vehicle collision happened because the autopilot system failed to recognize a walking pedestrian [1]. In these two examples, both the white truck and the walking pedestrian were confusing samples. The autopilot system did not recognize the potential risk, which in turn led to the accidents.

A confusing sample can be divided into two branches, i.e., the imprecise and out-of-domain sample. The imprecise sample located at the intersection of several classes occupies a high risk of misclassification due to the equal probabilities for several classes. The out-of-domain sample represents the samples that are different from the training distribution, which have two typical sources. (1) An outlier situated at a large distance from each of the training classes. (2) The out-of-domain sample might also come from samples that are not represented in the training dataset.

The uncertainty caused by confusing samples is the root of these problems. The predictive uncertainty is in general separated into data uncertainty and model uncertainty as shown in Fig.1. The data uncertainty is inherent in the training dataset and describes the confidence in the training dataset. The noise in the dataset or the overlapping among classes can lead to data uncertainty. It cannot be reduced by adding more data. The model uncertainty rises due to the model itself, which describes the confidence of the prediction. The model structure, overfitting, or underfitting is the reason for this type of uncertainty. The model uncertainty can be reduced by adding more training data and optimizing the model sufficiently by regulating or changing the complexity.



(a) The dataset. (b) The data uncertainty. (c) The model uncertainty.

Figure 1: Visualization of the data and model uncertainties based on regression. The dataset generator is cos function, and the noisy generator is $0.2 \times (x + 1)$.

Overall, we consider two categories to detect confusing samples as shown in Fig.2.

1. **Uncertainty estimation** can provide a single value representing uncertainty and decide whether to reject the input sample. Based on the evidential deep learning method [3] and take the base rates explicitly into account, we proposed the SLUE method [5]. The initialization of the base rate is evaluated. Since the update of base rates is after each batch, a comprehensive analysis with experiments under the batch size setting was carried out.

2. **Partial classification** is another perspective to deal with confusing samples, which can classify the confusing samples into subsets. We fulfilled Partial Classification only based on pre-trained Model Outputs (PCMO) [4], by transforming the model outputs to beliefs for predicted sets under the Dempster-Shafer theory. The most striking achievement is that the proposed method is fulfilled only based on model outputs that can be applied to any model without any demand to retrain the model or conduct any further modifications.

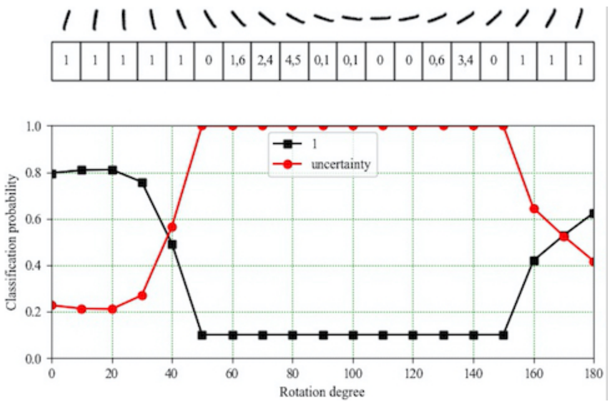


Figure 2: An example of uncertainty estimation and partial classification based on the rotated digit. The red line represents the uncertainty value and the value in the table cell of the second line represents the predicted subset.

References

[1] The economist. Why uber's self-driving car killed a pedestrian, 2018. Available online: https://www.economist.com/the-economist-explains/2018/05/29/why-ubers-self-driving-car-killed-a-pedestrian.

[2] Electrek. Understanding the fatal tesla accident on autopilot and the nhtsa probe, 2016. Available online: https://electrek.co/2016/07/01/understanding-fatal-tesla-accident-autopilot-nhtsa-probe/.

[3] Murat Sensoy, Lance Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pages 3183–3193, 2018.

[4] Jiarui Xie, Violaine Antoine, and Thierry Chateau. Pcmo: Partial classification from cnn-based model outputs. In *Neural Information Processing*, pages 150–163, 2021.

[5] Jiarui Xie, Thierry Chateau, and Violaine Antoine. A subjective-logic-based model uncertainty estimation mechanism for out-of-domain detection. In *2021 International Joint Conference on Neural Networks*, pages 1–6, 2021.

Short biography – I am currently pursuing a Ph.D. degree at the Université Clermont-Auvergne. My research interests include uncertainty estimation and partial classification.