SUMMARY

Ensemble Augmented-Shot-Y-Shaped Learning: State-of-the-Art Few-Shot Classification with Simple Components – EASY

Why it's interesting to me:

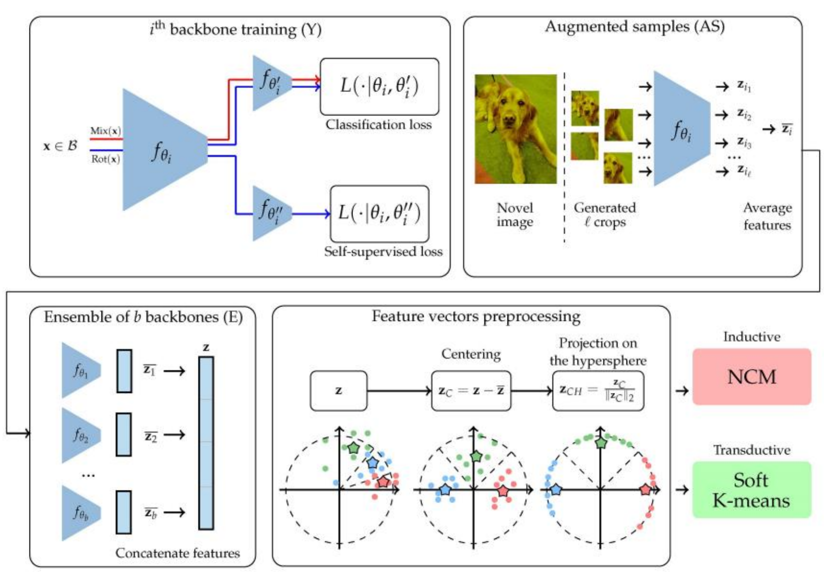
- Few-shot learning is a hot topic in machine learning where the model makes predictions based on a few training examples. It pushes data efficiency to its limits and has potential to reduce collecting large amounts of data.
- It's practical. With real-world situations, it is often the dataset is small. In this paper, they propose a very simple method combining components commonly found in literature and yet achieving competitive performance. They also aim to show that a simple approach reaches higher performance than increasingly complex methods proposed in the recent few-shot literature. With a clear and simple pipeline, this method can be easier to apply for common tasks.
- It's relevant to the subject of this technical test, in the sense of dealing with small dataset. In our case the amount of data is still sufficient for traditional transfer learning techniques, but we can use this method to develop a model with far less data (1-5 samples per label)

Paper summary:

The classical few-shot setup involves a base dataset (ImageNet, CIFAR …) for training the features extractor and a separate novel dataset to fine-turn. These datasets have distinct classes, challenging model generalization on new classes. The novel dataset contains only a few labeled examples (1-5) per class called support set and unlabeled examples called query set.

<u>Results</u>: state-of-the-art accuracy for 5 benchmarks dataset (MiniImagenet, TieredImagenet…) with a significant margin (>1%), using only popular methods at each step and well combine them

<u>*Methods:*</u>



- **Y**: train popular architecture (ex: Resnet) on base dataset (ex: ImageNet…), with 2 cross-entropy losses in parallel: one for the classification of base classes and the other for the self-supervised targets (rotations) – hence called Y-shape training. Drop the last classification layer to use these backbone and feature extraction.
- **AS**: doing data argumentation on novel dataset (flipping, cropping…) and extract their average features using multiple extractors – this approach is the most valuable for this pipeline and consistently improve performance for all cases.
- **E**: using ensemble of extractors to concatenate feature vectors of one image into a single, longer vector. This concatenated vector now represents the input data in a potentially richer and more diverse way, as it incorporates the distinct perspectives learned by each backbone.
- **Classification**: centering and project feature vectors into a hyperspace, then use suitable clustering methods (K-means…) to find the most likely group that new images (query set) belong to.