BC-Learning+ProtoNet Thought 1

Rick Liao

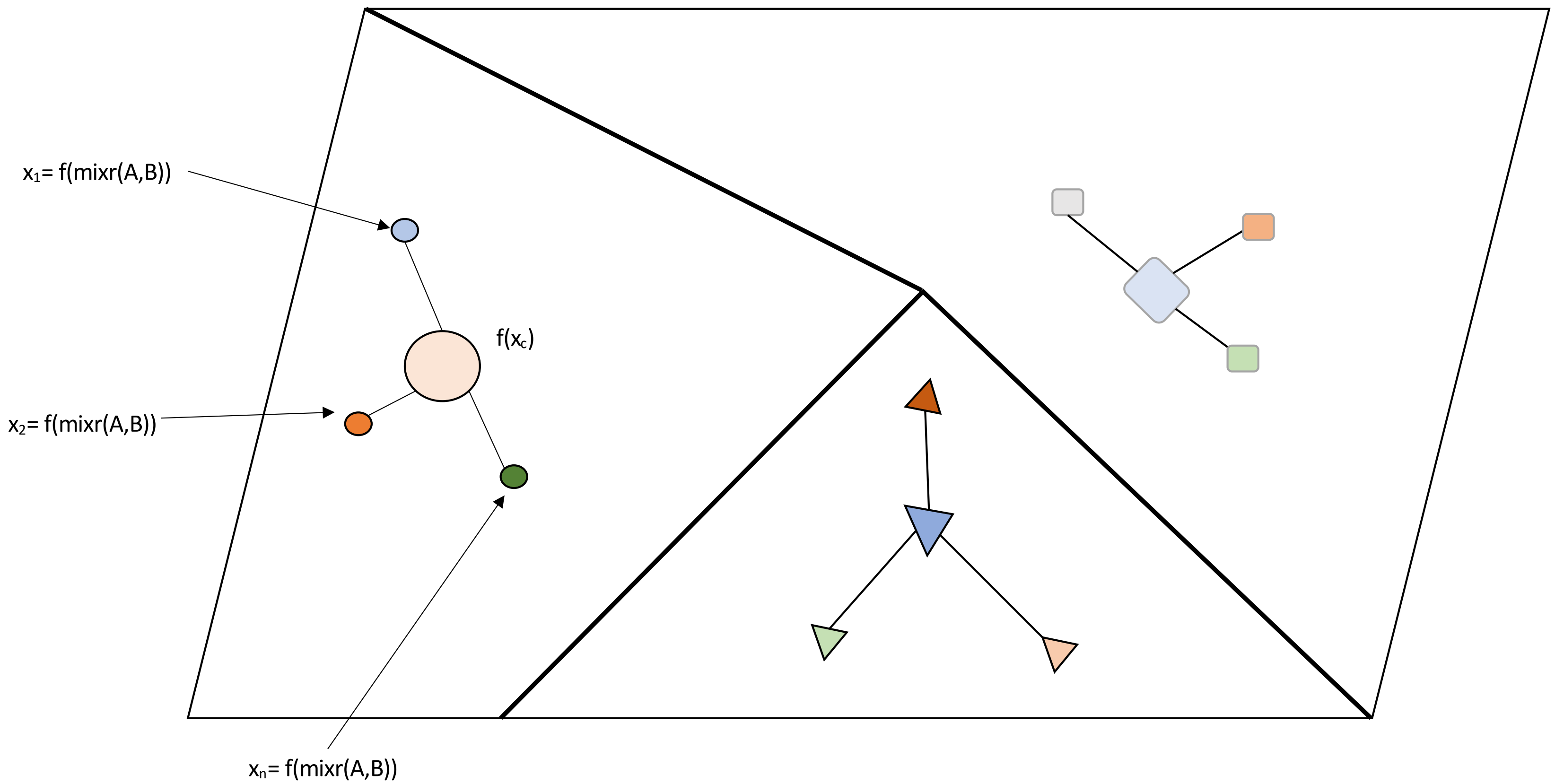
研究目的

- 採用少量的數據產生好的聲音分類模型
- 好的模型代表:
  - 泛化能力好。
  - 小模型一樣達成接近較大模型的分類能力。
  - 準確率高。

BC-Learning+ProtoNet想法

- BC-learning簡單來說:
  - 有Data-Augmentation的功能。
  - 強化Fisher's Criterion以達成更好的分類。
- Proto-Net簡單來說，做二件事:
  - 學習input→embedding的非線性映射。
  - 計算類別的protype representation。
- 結合BC-Learning及ProtoNet的目的:
  - 以更少的樣本得到泛化更好的訓練模型。

BC-Learning+ProtoNet概念圖



BC-Learning+ProtoNet想法

- BC-learning實作上的重點在一個有N個類別的訓練集中，隨機選取個M個類別(N >M)，在從這M個類別中各隨機選取1個樣本進行混合，而這個混合的樣本的標籤，也是依照樣本混合比例進行產生。這是BC-Learning不嚴謹的白話文描述。
- Proto-Net簡單來說，做二件事:
    - 學習input→embedding的非線性映射。
    - 計算類別的protype representation。
- 結合BC-Learning及ProtoNet的目的:
    - 以更少的樣本得到泛化更好的訓練模型。

BC-Learning+ProtoNet想法(Cont.)

- 問題點:
  - BC-Learning中樣本的產生方式修改。
  - 由BC-Learning產生出來的樣本，對應的 Label如何產生，才能符合ProtoNet的需求。
  - 類別的protype representation與BC-Learning產生的訓練集如何…
  - Loss-Function的定義。

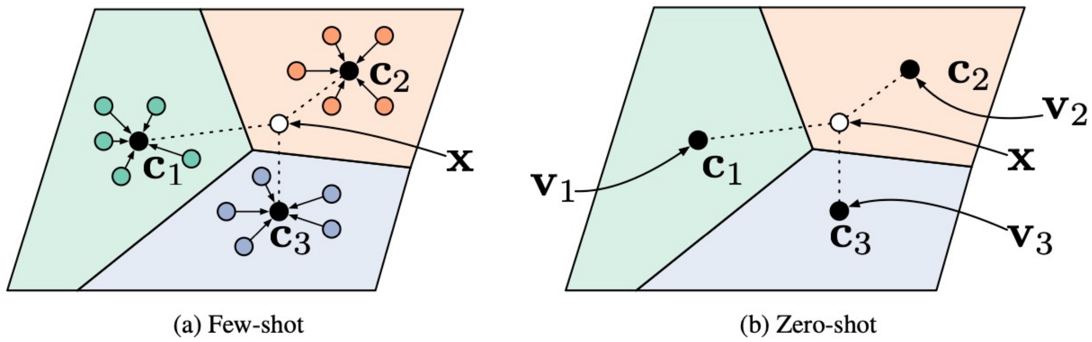


Figure 1: Prototypical networks in the few-shot and zero-shot scenarios. **Left**: Few-shot prototypes $\mathbf{c}_k$ are computed as the mean of embedded support examples for each class. **Right**: Zero-shot prototypes $\mathbf{c}_k$ are produced by embedding class meta-data $\mathbf{v}_k$. In either case, embedded query points are classified via a softmax over distances to class prototypes: $p_{\boldsymbol{\phi}}(y = k|\mathbf{x}) \propto \exp(-d(f_{\boldsymbol{\phi}}(\mathbf{x}), \mathbf{c}_k))$.

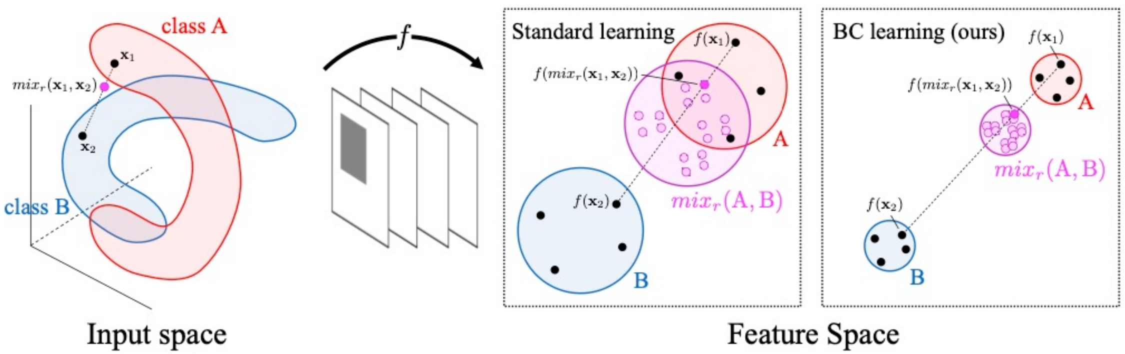


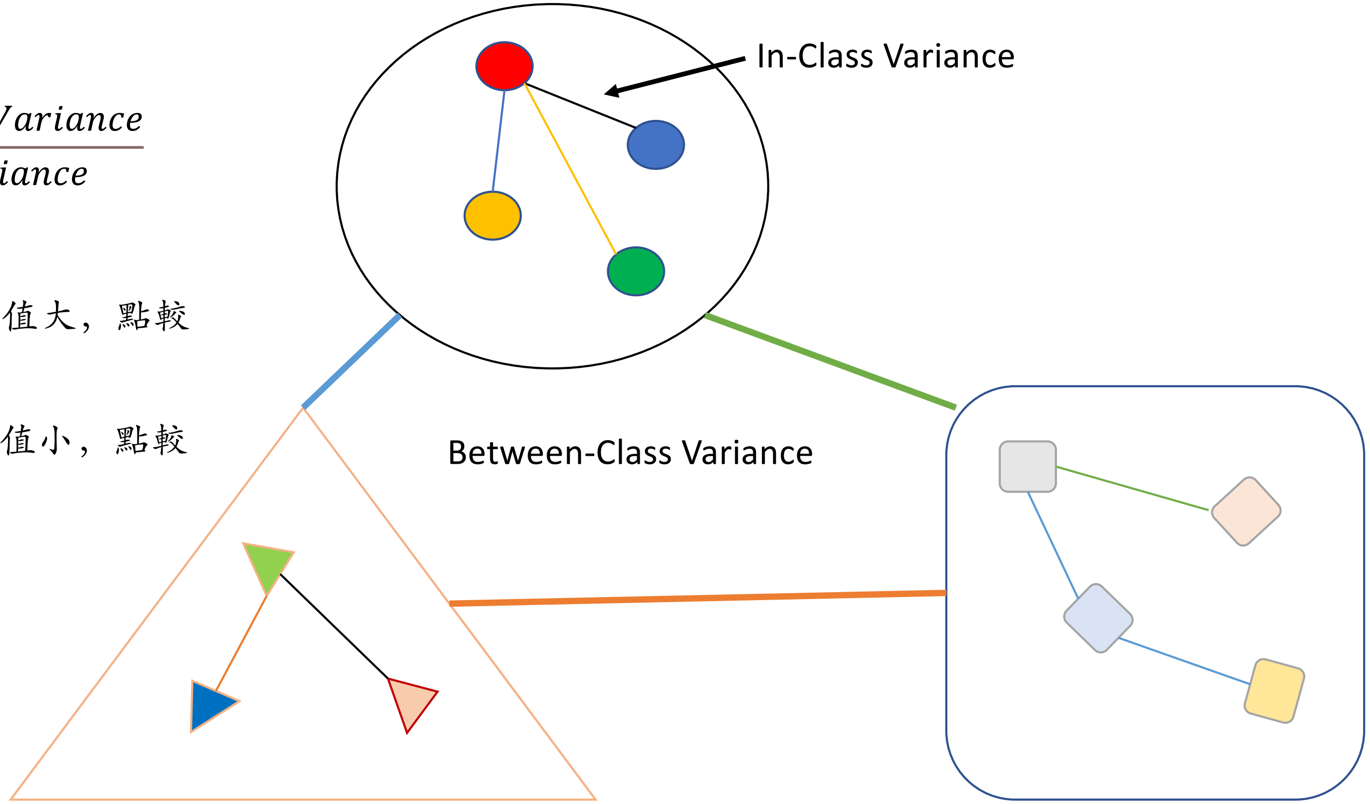
Figure 2: BC learning enlarges Fisher's criterion in the feature space, by training the model to output the mixing ratio between two classes. We hypothesize that a mixed sound $mix_r(\mathbf{x}_1, \mathbf{x}_2)$ is projected into the point near the internally dividing point of $f(\mathbf{x}_1)$ and $f(\mathbf{x}_2)$, considering the characteristic of sounds. **Middle**: When Fisher's criterion is small, some mixed examples are projected into one of the classes, and BC learning gives a large penalty. **Right**: When Fisher's criterion is large, most of the mixed examples are projected into between-class points, and BC learning gives a small penalty. Therefore, BC learning leads to such a feature space.

Fisher's Criteria

$$Ratio = \frac{Between - Class\ Variance}{In - Class\ Variance}$$

Ratio值小表示in-class variance值大，點較分散。

Ratio值大表示in-class variance值小，點較集中。



Pipe-Line of BC-Learning

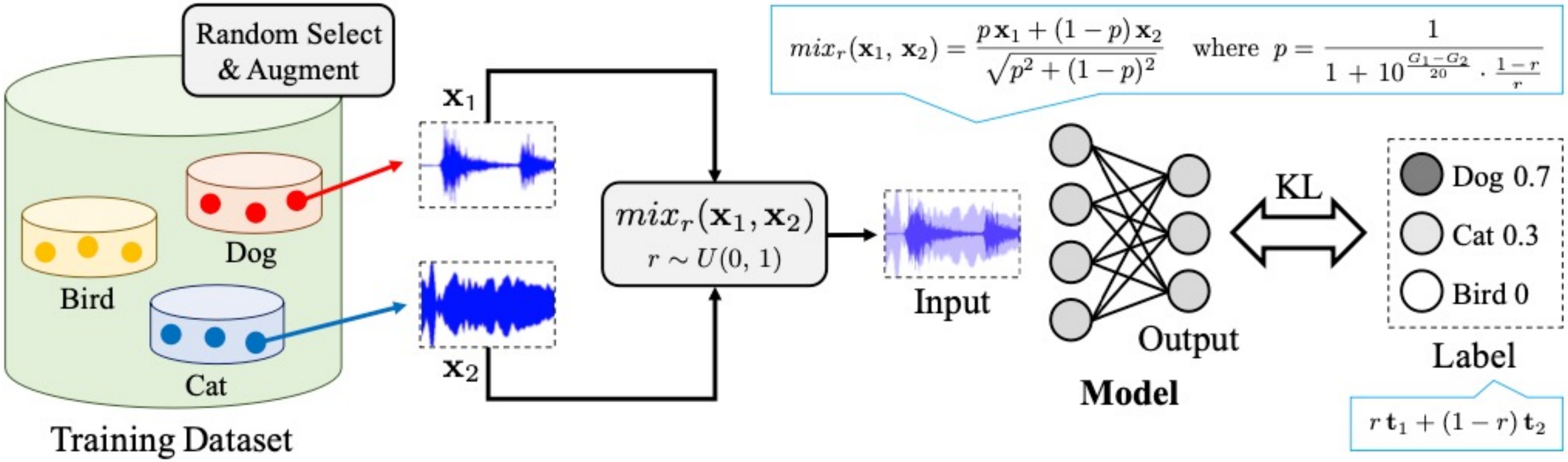
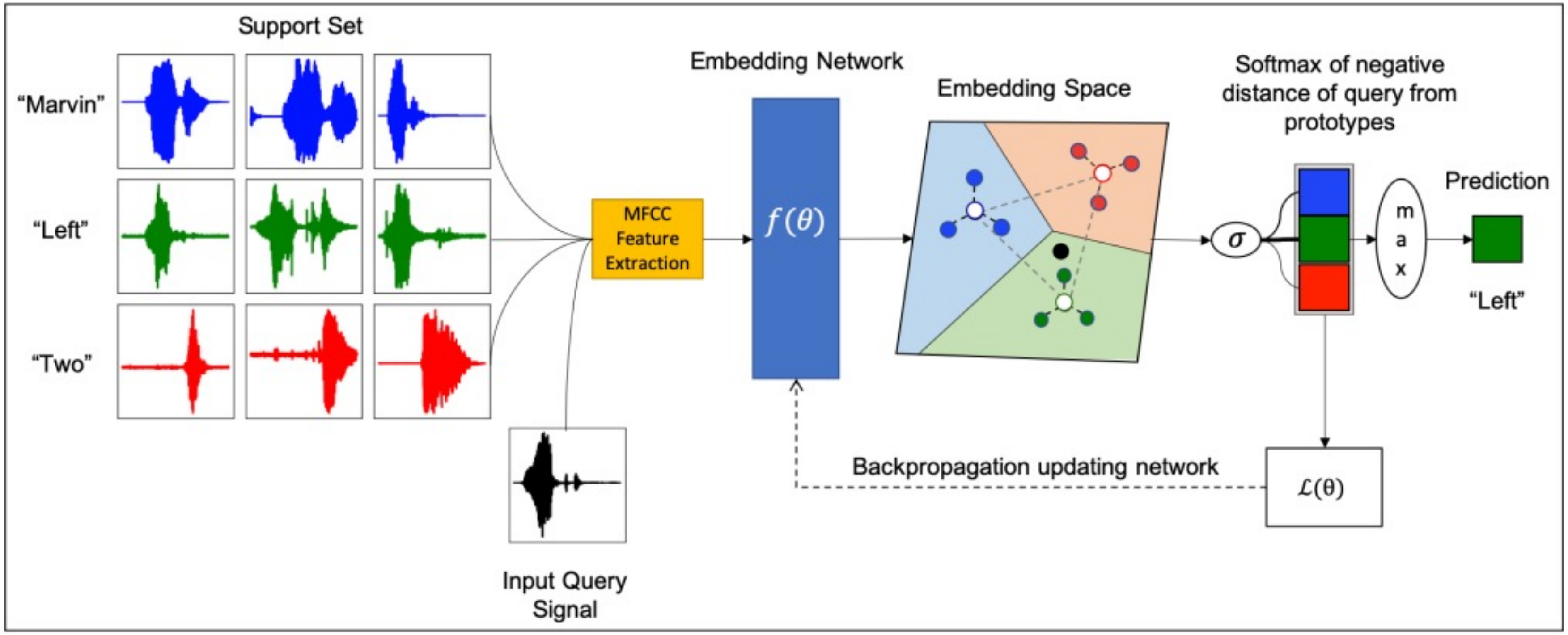


Figure 1: Pipeline of BC learning. We create each training example by mixing two sounds belonging to different classes with a random ratio. We input the mixed sound to the model and train the model to output the mixing ratio using the KL loss.

Pipeline example of ProtoNet for KWS



Prototypical Network Review

- Prototypical Network主要在解決few-shot even one-shot learning中overfit的問題以提高分類的準確率。
- Prototypical Network應用在few-shot and zero-shot learning.
- 其假設為：
  - 每個類別，存在一個Embedding，在這個Embedding中，Point Cluster(點簇)是圍繞在這個類別的單一原型(a single prototype representation)表示。
- 為達成上述假設，針對few-shot learning and zero-shot learning, we do：
  - For few-shot learning
    - 使用NN學習一個input→embedding的非線性映射。(Learn a mapping of input into embedding)
    - 每個類別，以類別中的所有點的mean做為其protype representation。

Prototypical Network Review (Cont.)

- For zero-shot learning
  - 每個類別以meta-data來表示。
  - 針對每個類別，以meta-data學習一個此類別的embedding到一共享的空間，並以這個embedding做為此類別的protype-representation。(learn an embedding of the meta-data into a shared space to serve as the prototype for each class)
- 分類方法則是找出embedded query input距離哪一個類別的protype representation最近。