

野鳥の鳴き声識別における特徴量類似性と認識精度に関する検討

古江 智瑛† 藤岡 優也† 三好 力†

龍谷大学†

1. はじめに

機械学習における教師あり学習では人手によるラベル付きデータを多数学習に用いるほど識別率が高くなることが知られている。しかし、ラベル付きデータは専門家の知識や人の手間などがかかり、高コストである。このコストを削減し、識別器の性能を向上させることは機械学習において重大な課題の一つである。ラベル付きデータが高コストである一方で、ラベルなしデータの場合は低コストで大量に獲得できる場合が多い。たとえば鳥の鳴き声の音声データであれば、森の中に録音機を設置するだけで容易に獲得できる。そこで本研究では特徴ベクトル間の類似性に着目し、少数のラベル付きデータから多数のラベルなしデータのクラスを特徴ベクトル間距離によって決定して訓練データとして用いる手法を検討する。特徴量間の距離尺度についてバタチャリヤ距離の平方根など多くの距離が提案されているが、本研究では最も広く使われており、直感的に理解がしやすいユークリッド距離を距離尺度として用いる。ニュージーランドに生息する野鳥の鳴き声のデータを例に、SVM において訓練データの数と識別率の推移の関係を検討するための実験を行った。

2. 従来手法

藤岡らの先行研究[1]では、SVM において類似性を特徴ベクトル間のユークリッド距離によって決定している。ラベル付きデータからランダムで少数のデータを選び取り、その選んだ各々のデータの平均ベクトルを中心として、そこからのユークリッド距離を全てのデータについて測定し、定めた距離内のもの全てを少数のラベル付きデータと同じラベルを付けて訓練データとする。この手法を用いた実験では、鳴き声の種類が少ない鳥は選んだデータ数が少なく距離が大きいと識別率が高くなり、鳴き声の種類が多い鳥は識別率が安定しない、と報告している。

3. 提案手法

藤岡らの手法は、少数のラベル付きデータの平均ベクトルを中心点として一定距離以内のラベルなしデータを訓練データとして取得している。これは、特徴空間内でラベル付きデータと同種の鳥の鳴き声データがラベルなしデータの中に等方的に分布していることを暗に仮定している。筆者らは、鳴き声の種類が少ない場合にはこの仮定は妥当だが、鳴き声の種類が多くなると鳴き声データは特徴空間内に鳴き声ごとに偏在し、等方性が減少して中心から等距離では不適切なデータが含まれやすいのではないかと考えた。そこで、これに対応する手法として、少数のラベル付きデータの平均ベクトルを中心点として1回だけラベルなしデータを取得するのではなく、少数のラベル付きデータのそれぞれを中心点として中心点の個数回ラベルなしデータの取得を行う方法と提案する。これにより、ランダムに選択されたラベル付きデータの分布を利用してラベルなしデータの偏在にある程度対応できるのではないかと考えた。

4. 実験

提案手法によって決定した訓練データを用いて検証実験を行った。テストデータは訓練データとの重複を避けるためあらかじめデータセットからランダムで20個抜き出して用いた。距離には5段階の基準を設け、基準値以下であれば正例、以上であれば負例のラベルをつけて訓練データとして用いる。識別率の測定は各距離ごと、各鳥ごと、中心データの数ごとに行い、30回繰り返した場合の識別率の平均値を測定値とする。使用するデータセットは野鳥の鳴き声データで、ミヤマオウム、キジカッコー、ニュージーランドアオバズクの3種類それぞれ60セグメントのデータセットをした。なお、ミヤマオウムについては3種類以上の鳴き声、キジカッコーについては3種類程度の鳴き声、ニュージーランドアオバズクについては1種類の鳴き声のデータで構成されている。これらの鳴き声データをMFCCによる特徴抽出を行い、910次元ベクトルを作成し、特徴ベクトルとした。また、機械学習と認識精度の測定にはサポートベクターマシン(SVM)を用いる。

Study on feature similarity and accuracy in bird call recognition

†Furue Tomoaki, Fujioka Yuya, Miyoshi Tsutomu
Ryukoku University

5. 実験結果と考察

実験で得られた正例ニュージーランドアオバズクの場合の結果を図1、図2に、正例キジカッコウの場合の結果を図3、図4に示す。横軸は訓練データ数、縦軸は識別率を表している。

ニュージーランドアオバズクの場合、1種類のみで構成されており、中心点の数、距離を変化させても比較的良い結果が得られている。一方、キジカッコウの場合複数の鳴き声で構成されており、中心点の数が多く、距離が大きい場合に良い結果が得られている。いずれも場合も、従来手法より識別率が安定して近い値が得られた。

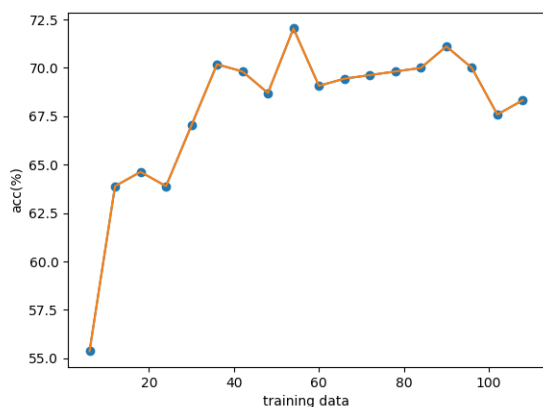


図1: アオバズク 中心点3 距離1500

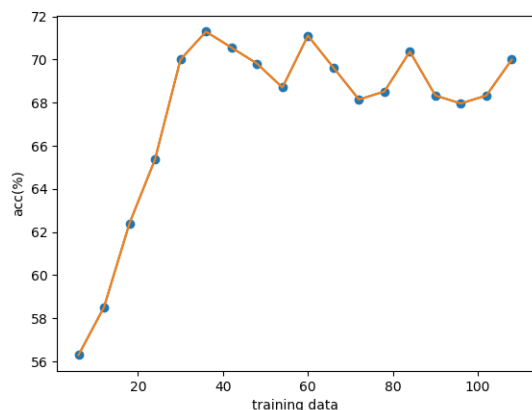


図2: アオバズク 中心点5 距離1600

6. おわりに

教師あり学習において、学習に多数の訓練データを用いるほど識別率が高くなることが知られていることから、少数のラベル付きデータから

の距離が近いものをラベル付きデータと類似したデータとみなして訓練データとして用いる手法の改良手法を提案し、実験を行った。実験結果より、従来手法では識別率が不安定であった鳴き声の種類が多い鳥に対しても、提案手法では比較的安定した識別率が得られることが分かった。

今後の課題として、中心点の数と有効なデータが含まれる距離との相関関係を実験により確認する。

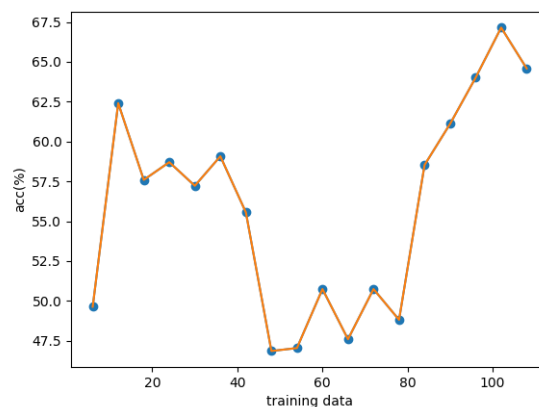


図3: キジカッコウ 中心点3 距離1500

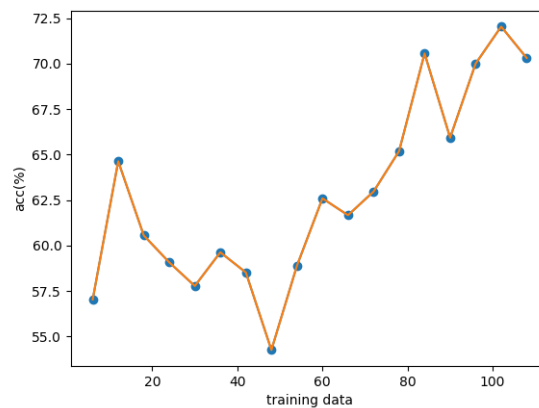


図4: キジカッコウ 中心点5 距離1600

参考文献

[1] 藤岡 優也, 三好 力: “機械学習における特徴量類似性と認識精度に関する検討,” 第17回情報科学技術フォーラム論文集, F-038, 福岡県, 福岡工業大学 (2018.09).