

機械学習に基づく画像認識システムの開発環境構築と画像アノテーションへの応用

担当者：寺本 隆磨

指導教員：長田 茂美 教授

1. まえがき

従来のテキストベースの検索方法では、インターネットに存在する膨大なマルチメディア情報の中から目的の情報を探し出すことは困難である。そのため、ファイル名とは別にそのマルチメディア情報の内容を記したタグ等のメタデータを付与すること、すなわち、アノテーションを行うことで、大容量のマルチメディアの検索を効率的に実行する手法が用いられている。

しかし、ファイル作成時にユーザーが手でアノテーションを行うことは、手間がかかるため現実的ではない。また、既存のファイルへのアノテーションも行わなければ、そのファイルは検索の対象外として扱われてしまう。

本研究では、サポートベクターマシン (Support Vector Machine, SVM) を用いて画像内の物体を認識することにより、自動でメタデータを作成するシステムを開発する。また、大量の画像データへの自動アノテーションに適用し、その有効性を示す。

2. SVM および BoF (Bag of Features) とは？

SVM とは、教師あり学習を用いて 2 クラスのパターン識別器を構成する手法であり、高い汎化能力を有するため、優れた認識性能を発揮できる。

また、BoF (Bag of Features) とは、文書分類などで使われている Bag of Words を画像に応用したもので、画像を visual word と呼ばれる局所特徴量の集まりと見なして、画像認識を行う手法である。一般に、局所特徴量として SIFT (Scale Invariant Feature Transform) 特徴量を用いることで、高精度な画像認識が可能となる。

3. システム概要

SVM を用いた画像認識システムの開発環境を構築するために、SVM 学習システム、アノテーション・検索実行システムを開発した。SVM 学習システムでは、物体の画像を SVM に学習させ、XML 形式で保存する。画像の SVM への入力の際には、SIFT 特徴量と BoF を使い、高精度な学習・認識が可能である。アノテーション・検索実行システムでは、画像内の物体認識を実行すると同時に、SVM 名と画像パスを XML ファイルに書き込む。検索時には、上記の XML ファイルを読み込み、SVM 名を入力すると画像パスを返すという方法で、タグによる検索を実現している。以下に、今回開発したシステムの概要を、また、図 1 に、システム概要図を示す。

➤ SVM 学習システム

- ・正例画像と負例画像を用いて SVM を学習させる。
 - ・SIFT 特徴量と BoF を利用する。
- ⇒ SVM の学習を実行できる。
- ⇒ テスト画像を用いて精度評価を実行できる。

➤ アノテーション・検索実行システム

- ・学習済み SVM を用いて画像内の物体を認識する。
 - ・SIFT 特徴量と BoF を利用する。
 - ・認識結果を基にタグ付けを行う。
 - ・タグを読み込み、SVM 名で検索する。
 - ・SVM 名ノードに記されている画像パスを出力する
- ⇒ 自動的なタグ付け、メタデータによる画像検索を実現できる。

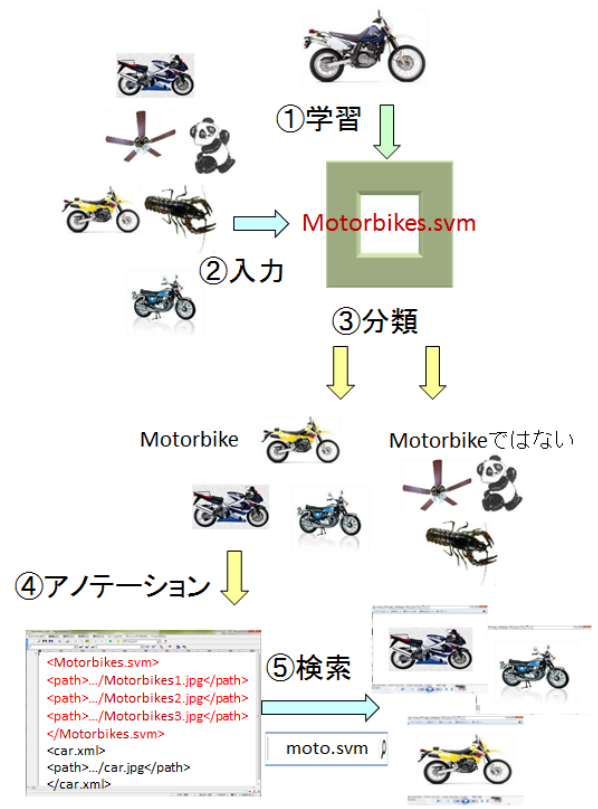


図 1 システム概要図

4. システム評価

Caltech101 データセットを用い、SVM の認識精度を評価した。その結果、平均 53.8%と良好な認識精度が得られた。一方、最高の認識精度が 98.5%であるのに対し、最低では 11.6%と、かなり大きな差がみられた。アノテーションの精度は SVM の認識精度に依存するため、精度が低いクラスの物体については、アノテーションに用いることができない。より信頼性の高いアノテーションを実現するためには、認識が難しいクラスに対する認識精度を向上させる必要がある。

5. むすび

本研究では、SVM を用いた画像認識システムの開発環境を構築するとともに、自動アノテーションシステムを開発した。これにより、大量の画像に対する高精度なアノテーションが可能となった。

しかし、精度が低いクラスの物体も存在することが判明した上、処理時間も短いとはいえず、開発環境を活用するためには、さらなる改良が必要である。

今後の展開として、SVM がより高精度な認識を実現できるようにチューニングする、SVM の他にも Adaboost やニューラルネットなどの学習識別器が使用可能とするなど、開発環境の改良や拡張が考えられる。

参考文献

- [1] 上東太一ほか, “Multiple Kernel Learning による 50 種類の食事画像の認識”, 電子情報通信学会論文誌 D, Vol.J93-D, No.8, pp.1397-1406, 2010.