

画像認識における、機械学習を考えてみたいと思います。基本的に Deep Learning を構成する方法についてです。これらは、今までの画像認識観点コンテンツをまとめたものでもあります。

顔認識とかの世の中に出ている画像認識技術は、基本的に特徴点の分布というものを明に使っていないものが多いかも知れませんが、画像に分布している点（画素）を選択して（画素の相互の位置関係：格は考慮しているわけです）値を見て解析していますから、その画素群を特徴点とみてよいでしょう。特徴点をいくつかのクラスターに分け、多段階の処理を経て、総合的な判断を下すというようにアルゴリズムを設計していれば、Deep Learning 技術と見なせるのではないのでしょうか。

私の好みは、特徴点を次の図形要素で捉えるものです。

- (1) 孤立点
- (2) 線分
- (3) 穴
- (4) 線の端点
- (5) 分岐点
- (6) 2つの線分の為す角

そうして、線の特徴も図形要素とします。

- (1) 図形の相対的大きさ
- (2) 曲率

さらに、認識に使うデータは、特徴点の配置関係（基本的に2項関係：格）です。

犬とか猫とか人間とかの認識は、まずモデルを機械学習します。それは、上記の特徴点と特徴点の属性とそれらの配置関係を重み付き投票に結び付けるものです。人間を認識する重み付き投票を得ていくには、無数の人間の姿の画像を特徴点とその属性とそれらの2項関係にこの弁別機から落とし込み、それが人間を表しているという重み付きにしまう。人間でないものも学習します。一つの重み付き投票ネットワーク（ニューロコンピュータ）は、入力画像の画素群を入力層に加えて、出力層は人間か人間でないかを判定するものになっている。

ということで、ひょっとして、画像の特徴点とその属性と配置関係というものは、そんなに精密なものでなくても良いかも知れません。それなりの精度であれば、十分な弁別機

ができるかも知れません。

更に、この弁別機を汎用化するには、多段階の弁別をするようにすればいいでしょう。

Deep Learning を構成するのです。例えば、

- (1) 動物と植物、人間が作ったものを弁別する段階
- (2) 哺乳類とそうでない動物とを弁別する段階
- (3) 哺乳類と犬、猫、人間を弁別する段階

おわり