

# 画像認識 1 章

tattaka

# 画像認識 (image recognition) とは

画像に写る内容を理解すること



例えば、上の写真には車と家が写っている

-> 画像を入力として [“車”, “家”] などを出力する

## 画像認識は難しい



図だと [“車”, “家”] でも [“フォルクスワーゲン”, “木”, “家”] でも  
どちらでも正しい

- ▶ 注目する場所, 認識に必要な前提知識, クラス自体の曖昧性  
によって結果が変わる

# 画像認識の歴史

- ▶ もともと**知能ロボットの視覚を実現するため**に研究がスタート
- ▶ 1970~80年代では人間のもつ知識をルールベースでプログラム->**信号を記号で表現する**アプローチ
- ▶ 1980年代以降は形や色，テクスチャ，動きなどの幾何学的・物理的な側面に焦点を当て大きな進歩
- ▶ 1999年に局所特徴**SIFT**が開発されてから多様なクラスを認識することが可能に
- ▶ 2009年には大規模データセット**ImageNet**が作成され，2012年に深層学習モデルを用いたアルゴリズムが既存手法を大きく上回った

# 画像認識の分類

- ▶ **物体認識** (object recognition)  
入力画像に写る物体を理解し、適切なラベルを付与する
  - ▶ **インスタンス認識** (instance recognition) 物体の固有な名前を出力
  - ▶ **クラス認識** (class recognition) 物体の属する概念を出力
- ▶ **シーン認識** (scene recognition)  
画像が表現する包括的な状態を理解する

これらに関連した分野として

- ▶ **画像アノテーション** (image annotation)  
狭義では複数のラベルが付与されたデータ構造を用いて入力画像に複数のラベルを付与する
- ▶ **物体検出** (object detection)  
インスタンスやクラスだけではなく、物体領域を推定する
  - ▶ **セマンティックセグメンテーション** (semantic segmentation)  
物体とその他の境界を切り分けて推定する などがある

# クラス分類の手順 1

## 1. サンプルング (sampling)

入力画像の代表的な点を抽出すること

### ▶ 疎なサンプルング (sparse sampling)

画像内のエッジやコーナなどの特徴的な点を抽出する  
特徴するためのアルゴリズムを**検出器** (detector) と呼ぶ

### ▶ 密なサンプルング (dense sampling)

一定の画素ごとに代表点を抽出する の 2 つに分けられる

## 2. 局所記述

獲得頂点の周りに小さな領域 (画像パッチやパッチと呼ぶ) を設定し、何らかの方法でパッチから特量を抽出する  
得られたパッチの特徴量と検出器から得られる情報合わせて**局所特徴** (local feature) と呼ぶ

## 3. 統計的特徴抽出

局所特徴は外乱などの影響でノイズを含むので特徴群の確率統計的構造に用いて処理すること  
主に主成分分析やフィッシャー線形判別分析などが利用される

## クラス分類の手順 2

### 4. コーディング (coding)

局所特徴を認識に有効なある次元数の特量ベクトルに変換する処理  
代表的な方法としてベクトル量子化など  
変換したい次元数と同じ数の代表点に局所特徴をクラスタリングする手法

### 5. プーリング (pooling)

ある画像領域から得られた特徴ベクトル群を 1 本のベクトルにまとめる処理  
画像一枚を代表するベクトルを **画像特徴ベクトル** (image feature vector) と呼ぶ  
代表的な手法として平均値プーリングなど

### 6. 分類

画像特徴ベクトルをクラスに割り振る処理これを行うアルゴリズムを **分類機** (classifier) と呼ぶ

シーン認識, 詳細物体認識や画像アノテーションでも上記の手順を利用することが多い

画像アノテーションの場合, 通常のクラス分類とは異なりノンパラメトリックな方法が積極的に活用される