

1. CV 任务: 图像识别; 目标检测; 目标分类; 图像分割

2. 常用特征方式: SIFT, 固定范围特征化

HOG: 梯度描述

the gap between human semantic understanding of image content and digital representation of image data

② viewpoint variation

illumination

deformation

occlusion

background clutter

intraclass variation

③ New algorithms occur; large data; advancement in computing devices

3. 数据集: CIFAR-10, CIFAR-100, ImageNet;

4. DNN 模型: CNN, RNN, GNN 等; 针对不同应用, 各网壳解决特定问题

5. Bio NN: 多神经元互联; MLP: 前向网壳感知机互联; LC: 由于输入图像线性相关来分类 $f(x) = Wx + b$

6. 损失函数: 用预测值及目标值之间 metric / d 训练, 常用回归/分类 (输出连乘值 / 预测输入类别)

$$MSE: \frac{1}{N} \sum (y - f(x))^2$$

$$MAE: \frac{1}{N} \sum |y - f(x)|$$

7. 我分及 MLP 的缺点:

① pros: easy to train, simple & infer cons: unable to solve difficult problems

② pros: handle more complex problems cons: require large number of parameters/ weight

8. CNN: 减小所需参数、受生物学激励

convolution layer: 提取特征, 前面是低级特征, 后面是高级特征, 共享 w, b

5x5, 跨度为1, 卷积2; 3x3, 跨度为1, 卷积1; 3x3, 跨度为2, 卷积2;

activation function layer: element-by-element 的非线性映射, ReLU 是最有效激活函数
与 convolution layer 结合

pooling layer: 用于减少激活层维度; 提供位置不变; 常用 max 及 average pooling; 各信道独立工作

Full-connected layer: 所有结点与下一层所有结点相连; 特征向量是输入图像的特征, 类似于线性分类器

softmax layer: 所最后一个 FC 层的输出映射成分类器中计算 loss 函数

CNN 目前最小化损失函数: 梯度下降法 (SGD, Adam 用于优化)

优点: accuracy; computation complexity, depends on application; memory footprint