

画像認識のための Few-Shot 学習

東京工業大学 情報理工学院 助教 井上 中順

実現したいこと

限られたデータからの
高精度な学習

産業界への影響

大量データの取得・共有が
困難な分野での活用
例：企業内データ，医療

主な成果と実績

- ・少量データからの学習を実現
- ・評価セット作成と国際 WS 主催

主要成果：Few-Shot Adaptation for Multimedia Semantic Indexing (ACM Multimedia 18)

概要：映像からの物体・動作・シーン検出において少量のデータから高精度なモデルを学習する

Many-Shot 学習

Training Samples → Detector

大量の教師付きデータからモデルを作成
e.g., Convolutional networks: ResNet (He 16), EfficientNet (Tan 19)

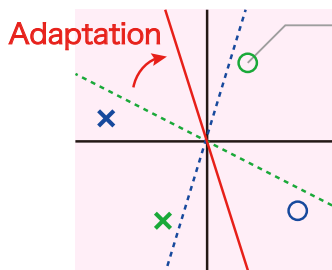
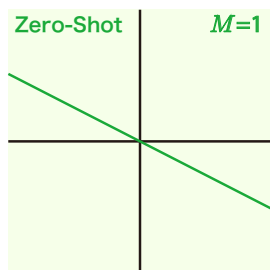
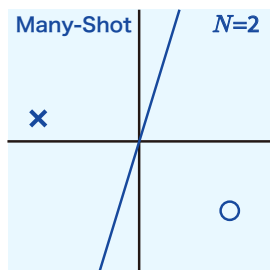
Zero-Shot 学習

Pre-Trained Detectors → Detector

既存モデルの凸結合からモデルを作成
e.g., Objects2action (Jain 15), Semantic Autoencoder (Kodirov 17)

提案手法：Few-Shot Adaptation

原理：Many-Shot 学習と Zero-Shot 学習の枠組みを統一的に定式化し、同時最適化をおこなう



Pseudo Training Samples

From a zero-shot detector for concept i ,

$$g_{ij}(x) = \sum_{j=1}^M s_{ij} w_j^T x$$

similarity between concepts i and j
extract pseudo positive/negative samples,

$$\tilde{x}_{2j} = +\frac{1}{2} s_{ij} w_j, \quad \tilde{x}_{2j-1} = -\frac{1}{2} s_{ij} w_j$$

and apply a many-shot learning optimizer.

Training Samples

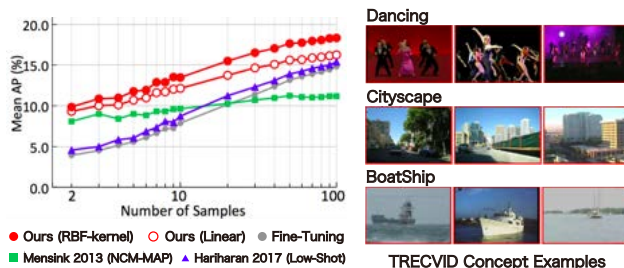
Pre-Trained Detectors

Detector

評価実験：少数サンプルでの学習時に精度が向上

Settings on the TRECVID Semantic Indexing Dataset

- Training samples: 1~100 video shots per concept
- Pre-trained detectors: ImageNet-1K + Skipgram (word2vec)
- Evaluation measure: Mean Average Precision
- Test data: 107,806 video shots, 30 concepts+background



Multi-Discipline Approach for Learning Concepts

— Zero-Shot, One-Shot, Few-Shot and Beyond —

概要：国際ワークショップの主催と Few-Shot 学習の評価用データセット作成



ICCV19 においてワークショップを主催
14 Full Papers, 12 Extended Abstracts を採録



Few-Shot Verb Dataset
動作の決定的瞬間を捉えた画像 URL 集