Hierarchical Image Classification

Prashant Kumar
Stanford University

Introduction
Can image classification via CNN benefit from Hierarchical classification?

HD-CNN

Independent layers for fine-classification \(\Rightarrow\) Depends on coarse classification accuracy.

Bilinear-CNN utilizes correlations between filters of convolutional network.

Course project idea: Use correlations between filters and coarse categorization.

CNN architecture

- **Coarse NN**: VGG13
- **Main NN**: VGG16

**Methods:**
1. Category Hierarchy (CH)
2. Coarse Visual Category (CVC)
3. End-to-End (ETE)

**Hyperparameters:**
- Trainset:Devset per label \(80:20\)
- Number of coarse classes \(10\)
- Initial Learning rate \(0.01\)
- L2 regularization \(0.001\)

VGG16 baseline

**Dataset**
- Trainset: 94.94%
- Validation set: 69.04%
- Test set: 68.74%

HD-CNN

**Overall accuracies**
- Trainset: 81.92%
- Validation set: 74.72%
- Test set: 65.41%

Coarse Visual Categories

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trainset</td>
<td>81.92%</td>
</tr>
<tr>
<td>Validation set</td>
<td>74.72%</td>
</tr>
<tr>
<td>Test set</td>
<td>65.41%</td>
</tr>
</tbody>
</table>

32 epochs.

End-to-End

**Overall accuracies**
- Trainset: 92.56%
- Validation set: 68.94%
- Test set: 68.42%

36 epochs.

Category Hierarchy

- **Trees**: chair, forest, maple tree, motorcycle, oak tree, orange, pine tree, snake, willow
- **Vehicles**: bus, cattle, bike, farm mower, mountain, pickup truck, streetcar, bus, truck, train
- **Big Carnivores**: apple, leopard, lion, tiger, wolf
- **Category 4**: cat, dog, fish, dolphin, snake, other, cat, old, duck, mouse, turtle, whale
- **Sea animals**: coral, dolphin, whale, oyster, eel, seal, sea, whale, shark, sea, turtle, shark
- **Category 7**: bed, road, table, television, wardrobe
- **Category 8**: aquarium, bike, bear, cloud, chimpanzee, elephant, hamster, sea
- **Insects**: bee, beetle, butterfly, caterpillar, cockroach, cock, cup, orchid, plant, puppy, spider, sunflower, sweet pepper, telephone, tulip
- **Category 10**: bridge, castle, crocodile, house, kangaroo, lizard, mouse, mushroom, porcupine, rabbit, shrew, skyscraper, snail, squirrel

Training Results

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Accuracy</th>
<th>Overall accuracies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trainset</td>
<td>82.00%</td>
<td>Trainset: 84.46%</td>
</tr>
<tr>
<td>Validation set</td>
<td>75.88%</td>
<td>Validation set: 69.90%</td>
</tr>
<tr>
<td>Test set</td>
<td>75.88%</td>
<td>Test set: 65.45%</td>
</tr>
</tbody>
</table>

31 epochs.

Conclusions

- Overfitting is a serious issue in Coarse NN.
- No significant change in accuracies.

References