Convolutional Neural Nets

October 24, 2013
Convolution

- Blur operation on $512 \times 512$ image of the moon’s surface
- Matrix multiplication: $512^2 \times 512^2$ matrix
  ... $256$GBytes of memory & $128$GFlops :/
- Convolution ($\text{conv}2$): $9 \times 9$ “kernel” (aka “filter”) & $20.25$ MFlops
Convolution
Orientation Maps

Miikkulainen, 2005
Edge Filters

\[ G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -2 \end{bmatrix} \ast A, \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \ast A \]

\[ G = \sqrt{G_x^2 + G_y^2} \]
Filterbanks

Histogram of Oriented Gradients (HoG)
Biologically-inspired filters: HMAX
HoG + SVM

Input image → Normalize gamma & colour → Compute gradients → Weighted vote into spatial & orientation cells → Contrast normalize over overlapping spatial blocks → Collect HOG’s over detection window → Linear SVM → Person / non-person classification

Person detection

Opponent Colors

Part-Based models

Different scenarios

Using and In

Dalal 2005
“The basic design principle is to reduce the number of free parameters in the network as much as possible without overly reducing its computational power. Application of this principle increases the probability of correct generalization because it results in a specialized network architecture that has a reduced entropy”

“LeNet” (LeCun 1989)
Krizhevsky & Hinton 2012 LSVRC Challenge (1K classes, 1M images):
convolution + ReLU + max-pooling + dropout + transforms

“Trained with stochastic gradient descent on two NVIDIA GPUs for about a week ... 650,000 neurons, 60,000,000 parameters, 630,000,000 connections”
### Object Recognition

<table>
<thead>
<tr>
<th>mite</th>
<th>container ship</th>
<th>motor scooter</th>
<th>leopard</th>
</tr>
</thead>
<tbody>
<tr>
<td>black widow</td>
<td>lifeboat</td>
<td>go-kart</td>
<td>jaguar</td>
</tr>
<tr>
<td>cockroach</td>
<td>amphibian</td>
<td>moped</td>
<td>cheetah</td>
</tr>
<tr>
<td>tick</td>
<td>fireboat</td>
<td>bumper car</td>
<td>snow leopard</td>
</tr>
<tr>
<td>starfish</td>
<td>drilling platform</td>
<td>golfcart</td>
<td>Egyptian cat</td>
</tr>
<tr>
<td>grille</td>
<td>mushroom</td>
<td>cherry</td>
<td>Madagascar cat</td>
</tr>
<tr>
<td>convertible</td>
<td>agaric</td>
<td>dalmatian</td>
<td>squirrel monkey</td>
</tr>
<tr>
<td>grille</td>
<td>mushroom</td>
<td>grape</td>
<td>spider monkey</td>
</tr>
<tr>
<td>pickup</td>
<td>jelly fungus</td>
<td>elderberry</td>
<td>titi</td>
</tr>
<tr>
<td>beach wagon</td>
<td>gill fungus</td>
<td>currant</td>
<td>indri</td>
</tr>
<tr>
<td>fire engine</td>
<td>dead-man’s-fingers</td>
<td>ffordshire bullterrier</td>
<td>howler monkey</td>
</tr>
</tbody>
</table>
+ Localization

- Chime: wine bottle, digital clock, gar, oboe, typewriter, keyboard
- Boathouse: apiary, mobile home, boathouse, fence, patio
- Scottish Deerhound: Scottish deerhound, Irish wolfhound, Leonberg, German shepherd, Tibetan mastiff
- Electric Guitar: violin, carpenter's kit, revolver, loafer, corkscrew
- Motor Scooter: motor scooter, moped, snowmobile, police van, moving van
- Sturgeon: leatherback turtle, volcano, wreck, alp, breakwater
- Violin: violin, cello, acoustic guitar, drumstick, electric guitar
- Fire Screen: fire screen, sundial, mailbag, umbrella, purse
Cancer Detection

Schmidhuber, MICCAI 2013 Grand Challenge on Mitosis Detection
Code

Matlab: https://github.com/rasmusbergpalm/DeepLearnToolbox
Python + CUDA: http://deeplearning.net/software/theano/
C++ CUDA: http://code.google.com/p/cuda-convnet/
Lua: http://www.torch.ch