Capsule Networks
A novel type of neural networks

Vladimír Kunc
FEL ČVUT

January 5, 2018
PROBLEMS OF CNNs

- it is difficult to detect spatial relationships among features
  - perspective
  - size
  - orientation
- pooling layers (if present) lose information
  - but helps with positional invariance
- rotation (etc.) robustness achieved by training filters for each possible rotation
  - blow up in the number of neurons and redundancy
  - needs many training data in different positions (e.g. the same image is randomly rotated in each epoch)
Figure: For such images the classical CNNs are more prone to failure.

Source.
SPATIAL RELATIONSHIPS AMONG FEATURES II

Figure: How a CNN evaluates features. Source.

Figure: How a capsule network evaluates features. Source.
SPATIAL RELATIONSHIPS AMONG FEATURES III

Figure: Examples of rotation. Source.
SPATIAL RELATIONSHIPS AMONG FEATURES IV

Figure: Classical CNNs dealing with rotation. Source.
SPATIAL RELATIONSHIPS AMONG FEATURES V

Figure: Capsule network dealing with rotation. Source.

Figure: Capsule network dealing with rotation. Source.
SPATIAL RELATIONSHIPS AMONG FEATURES VI

Single neuron

Single capsule with 2 neurons

(Number rotated by 20°)

Figure: Output of a (vector) capsule. Source.
INвариантность vs Еквиварианство

- Инвариантность
  - детектирование характеристик независимо от вариантов
  - потеря пространственного ориентирования
  - классические CNNs (обучение вращаемых характеристик — необходимость для огромных количеств данных)

- Еквиварианство
  - детектирование объектов, которые могут претворяться в другие
  - возможно выстраивать варианты с меньшим объемом обучающих данных
Dynamic routing

- the capsule network need to learn hierarchies among features
  - it performs hierarchical clustering
  - it selects to which capsule in the upper layer a capsule will send the output
- iterative process
  - another loop during training (harder to implement in some frameworks)
- “capsule layers may eventually behaves as a parse tree that explore the part-whole relationship” (Source).
DYNAMIC ROUTING II

Figure: Routing example Source.
Dynamic routing III

- routing-by-agreement (first proposed)
  - iterative voting
- EM routing
  - proposed in paper *Matrix capsules with EM routing*
  - matrix capsules (not just vector) — pose matrices to capture viewpoint of the object
  - EM algorithm used for routing

Figure: Matrix capsule output Source.
EXPECTED I

Figure: Hierarchical features Source.
EXAMPLE OF ARCHITECTURE (MNIST)

Figure: Architecture used for MNIST dataset Source.
MNIST

Figure: MNIST dataset Source.
**EXAMPLE OF ARCHITECTURE (SMALLNORB)**

Figure: Architecture used for MNIST dataset Source.
SMALLNORB

Figure: smallNORB dataset Source.

- 5 toy classes: airplanes, cars, trucks, humans and animals
- 18 different azimuths (0-340), 9 elevations and 6 lighting conditions
RESULTS FOR SMALLNORB I

Figure: Dynamic routing (3 iterations) Source.
**RESULTS FOR SMALLNORB**

<table>
<thead>
<tr>
<th>Test set</th>
<th>Azimuth</th>
<th>Elevation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CNN</td>
<td>Capsules</td>
</tr>
<tr>
<td>Novel viewpoints</td>
<td>20%</td>
<td>13.5%</td>
</tr>
<tr>
<td>Familiar viewpoints</td>
<td>3.7%</td>
<td>3.7%</td>
</tr>
</tbody>
</table>

Figure: Results on smallNORB compared to CNNs Source.

- better generalization with less data
The final remarks section highlights the following points:

- **Capsule networks learn hierarchical features**
  - Similarly (as we expect) to the functioning of a brain
  - Spatial relationships between features are preserved

- **Iterative process**
  - Another loop during training (harder to implement in some frameworks)

- “Capsule layers may eventually behave as a parse tree that explore the part-whole relationship” (Source).
CONCLUSION

- capsule networks are slow to learn compared to CNNs (at the moment)
  - it is easier to learn much more complex classical CNN that will perform better
  - CapsNets are just at the beginning (similar position to CNNs several years ago)
- routing algorithms are expected to develop
- implementations available in many frameworks
FURTHER RESOURCES

- https://www.youtube.com/watch?v=rTawFwUvnLE
  - Hinton’s talk "What’s wrong with CNNs"
- https://www.youtube.com/watch?v=pPN8d0E3900
  - quite good explanation of CapsNets with good visualization
  - original paper on CapsNets (Dynamic Routing Between Capsules)
- https://kndrck.co/posts/capsule_networks_explained/
- https://hackernoon.com/what-is-a-capsnet-or-capsule-network-2bfbe48769cc
  - includes code in TensorFlow framework