,,

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)

SPATIAL RELATIONSHIPS AMONG FEATURES

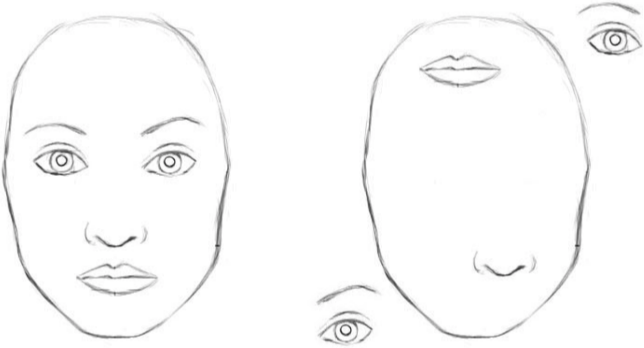

Figure: For such images the classical CNNs are more prone to failure. Source.

SPATIAL RELATIONSHIPS AMONG FEATURES II

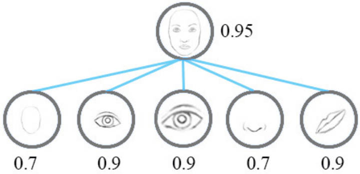

Figure: How a CNN evaluates features. Source.

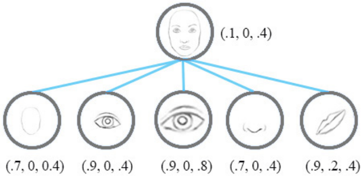

Figure: How a capsule network evaluates features. Source.

SPATIAL RELATIONSHIPS AMONG FEATURES III

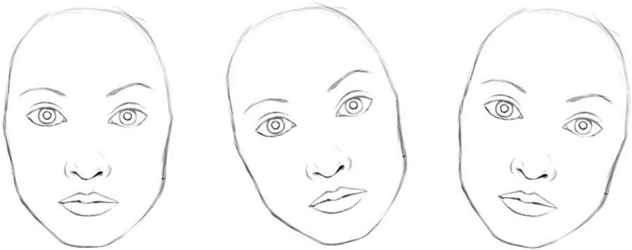

Figure: Examples of rotation. Source.

SPATIAL RELATIONSHIPS AMONG FEATURES IV

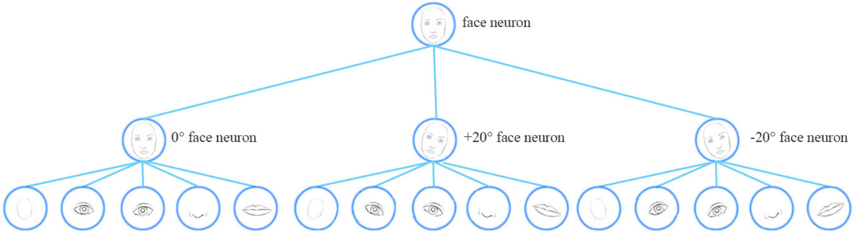

Figure: Classical CNNs dealing with rotation. Source.

SPATIAL RELATIONSHIPS AMONG FEATURES V

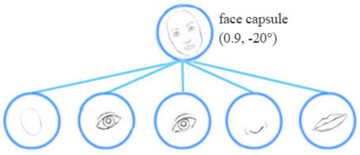

Figure: Capsule network dealing with rotation. Source.

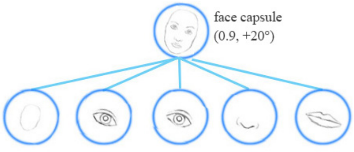

Figure: Capsule network dealing with rotation. Source.

SPATIAL RELATIONSHIPS AMONG FEATURES VI

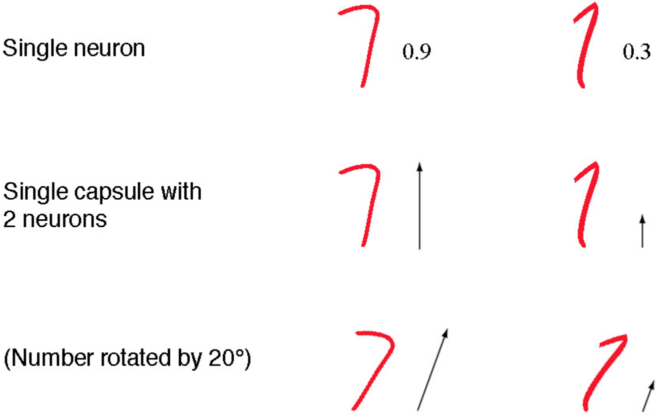

Figure: Output of a (vector) capsule. Source.

INVARIANCE VS EQUIVARIANCE

- Invariance
  - detection of features regardless of the variants
  - loss of spatial orientation
  - classical CNNs (learning rotated features — need for huge amounts of data)
- Equivariance
  - detection of objects that can transform to each other
  - it might be possible to extrapolate variants with less training data

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

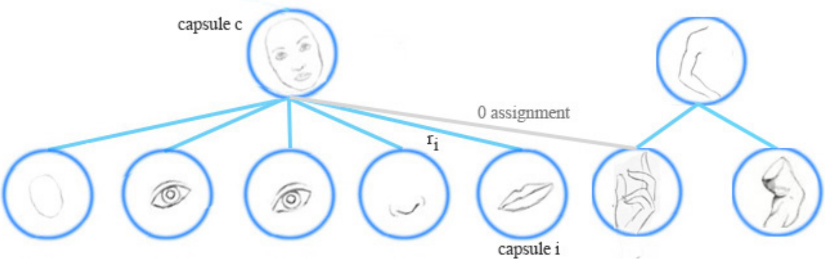

Figure: Routing example Source.

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

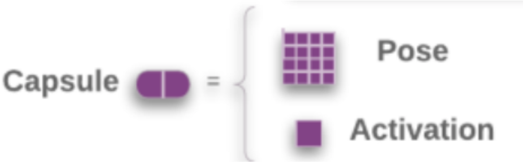

Figure: Matrix capsule output Source.

EXPECTED I

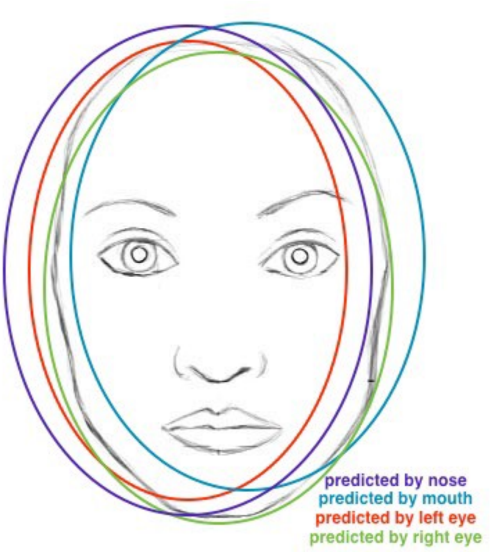

Figure: Hierarchical features Source.

Example of architecture (MNIST)

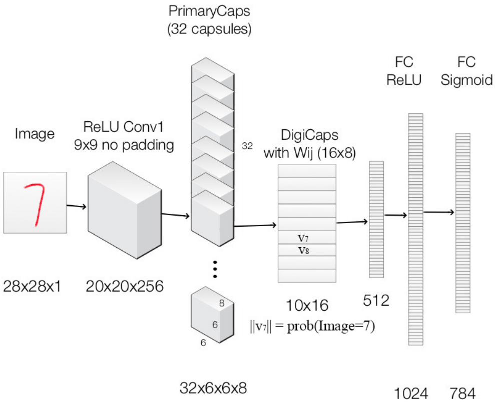

Figure: Architecture used for MNIST dataset Source.

MNIST

Figure: MNIST dataset Source.

EXAMPLE OF ARCHITECTURE (SMALLNORB)

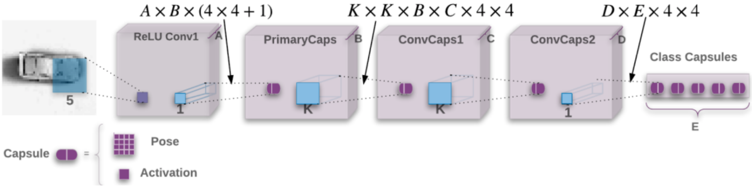

Figure: Architecture used for MNIST dataset Source.

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

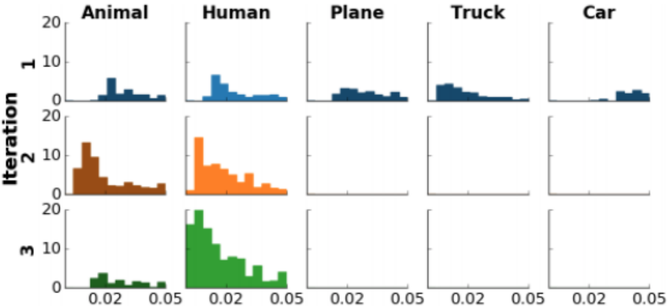

Figure: Dynamic routing (3 iterations) Source.

RESULTS FOR SMALLNORB

Test set	Azimuth		Elevation	
	CNN	Capsules	CNN	Capsules
Novel viewpoints	20%	13.5%	17.8%	12.3%
Familiar viewpoints	3.7%	3.7%	4.3%	4.3%

Figure: Results on smallNORB compared to CNNs Source.

- better generalization with less data

FINAL REMARKS

- 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 behaves 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
- https://arxiv.org/pdf/1710.09829.pdf
  - original paper on CapsNets (*Dynamic Routing Between Capsules*)
- https://jhui.github.io/2017/11/03/Dynamic-Routing-Between-Capsules/
- https://medium.com/ai%C2%B3-theory-practice-business/understanding-hintons-capsule-networks-part-i-intuition-b4b559d11
- https://kndrck.co/posts/capsule_networks_explained/
- https://hackernoon.com/what-is-a-capsnet-or-capsule-network-2bfbe48769cc
  - includes code in TensorFlow framework