

# Capsule Network with Routing Mechanism

## Part 1: Dynamic Routing

- ❑ Ou Changkun
- ❑ Email: [hi@changkun.us](mailto:hi@changkun.us)
- ❑ Homepage: <https://changkun.de>

# Agenda

## 1. Motivations of Capsule Network

- Limitations of Convolutional Networks
- “Coordinate Frame” in Human Vision

## 2. Traditional v.s. Capsule Neuron

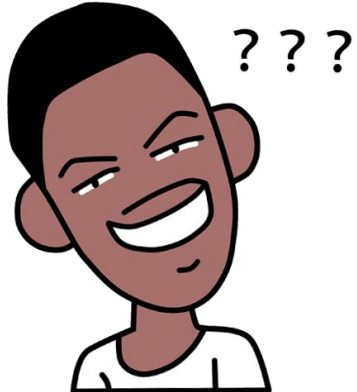
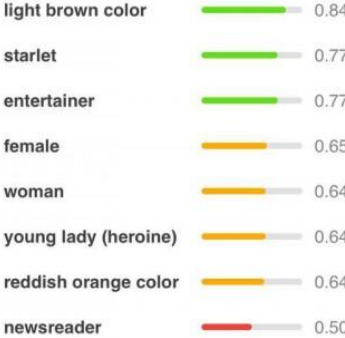
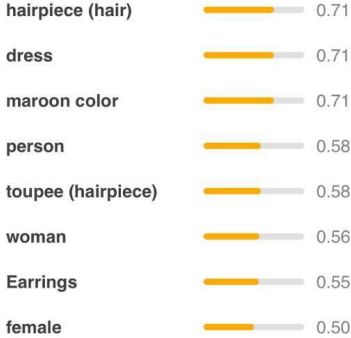
## 3. Routing Mechanism

- Dynamic Routing
- Case Study: CapsNet with Activity Vector

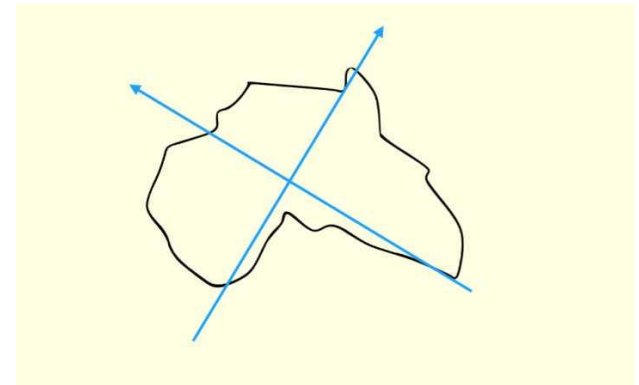
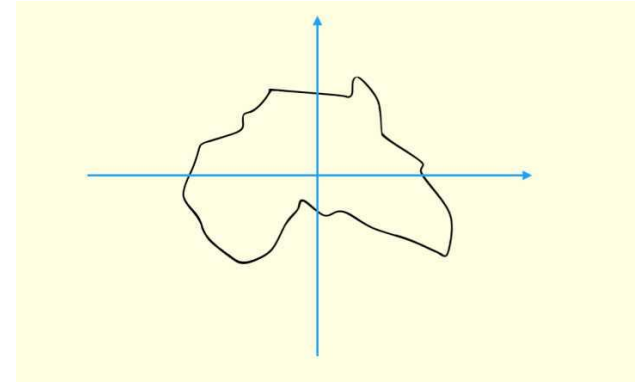
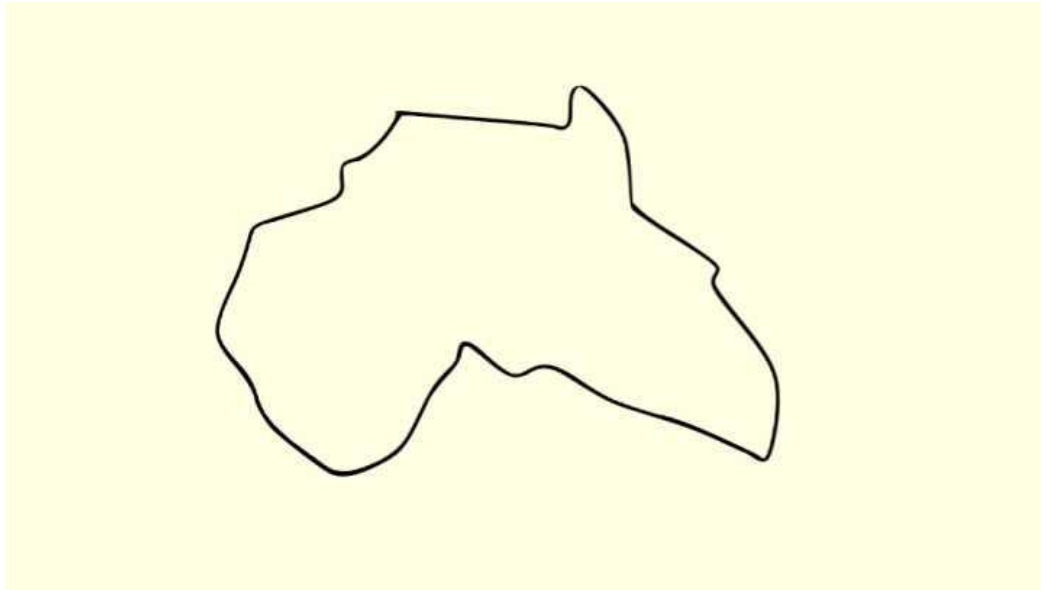
## 4. Experiments

- Interpretable Activity Vector
- CapsNet on Fashion MNIST

# Limitations of CNN

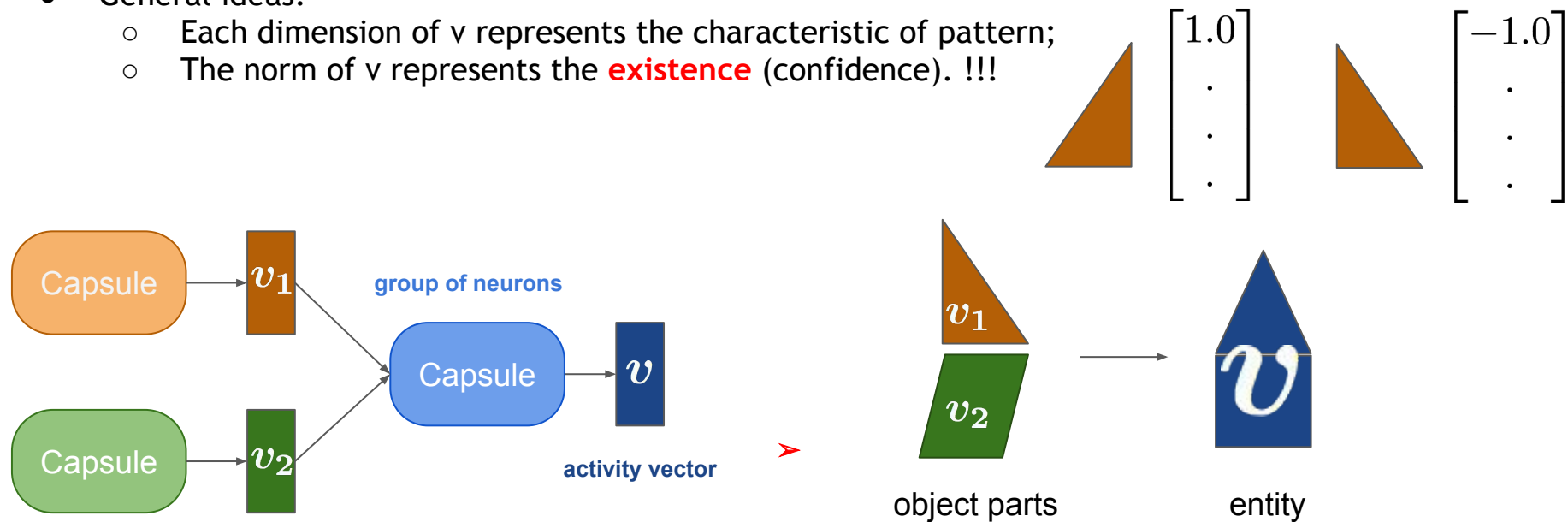


# “Coordinate Frame” in Human Vision

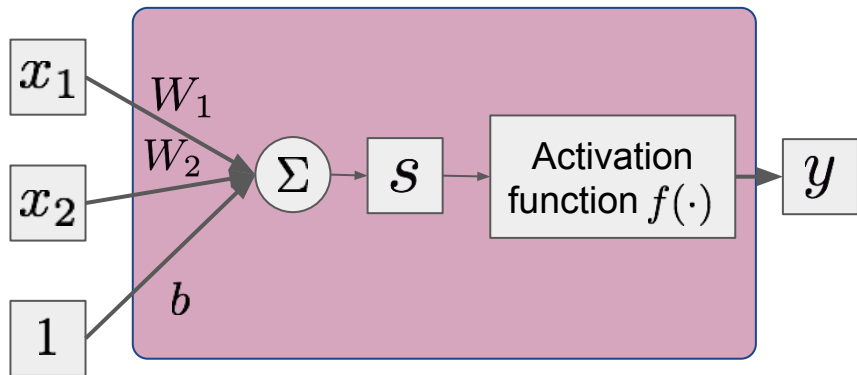


# What is a *Capsule*?

- “A *capsule* is **a group of neurons** whose **activity vector represents** the instantiation parameters of a specific type of entity such as an object or **an object part.**”
- General ideas:
  - Each dimension of  $v$  represents the characteristic of pattern;
  - The norm of  $v$  represents the **existence** (confidence). !!!



# Traditional v.s. Capsule Neuron

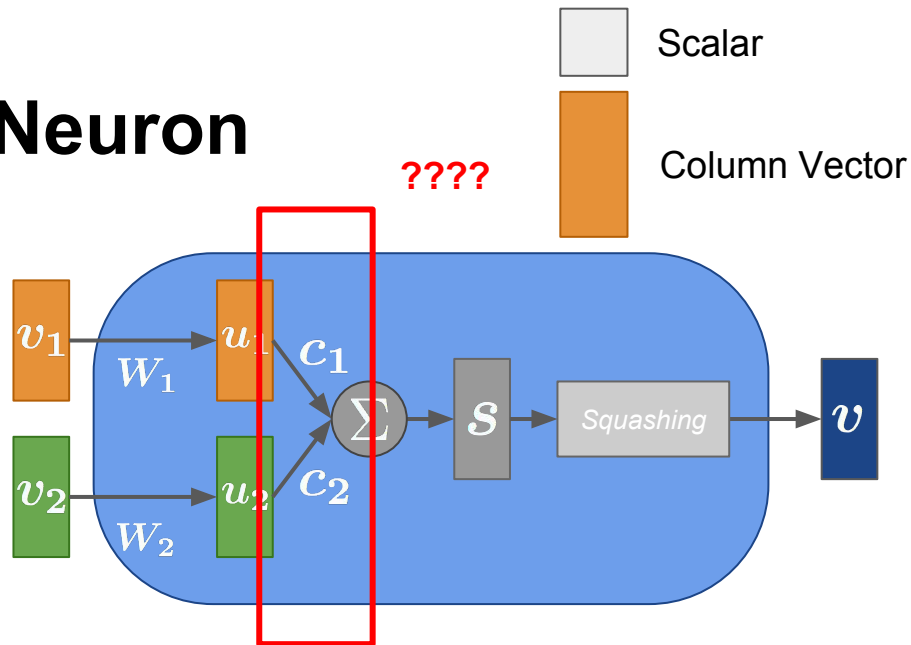


**Traditional Neuron:** Scalar  $\rightarrow$  Scalar

$$s = b + \sum W_i x_i$$

$$y = f(s)$$

$f(\cdot)$  : Sigmoid, ReLU, Maxout, etc.



**Capsule Neuron:** Vector  $\rightarrow$  Vector

$$u_i = W_i v_i \quad s = \sum c_i u_i \quad \text{????}$$

$$v = \text{Squashing}(s) = \frac{\|s\|^2}{1 + \|s\|^2} \frac{s}{\|s\|}$$

# Dynamic Routing (by Agreement)

$\|v\|$  is confidence

Initialize  $b_{11}, b_{21} = 0$

for  $r$  in range( $1 \dots T$ )

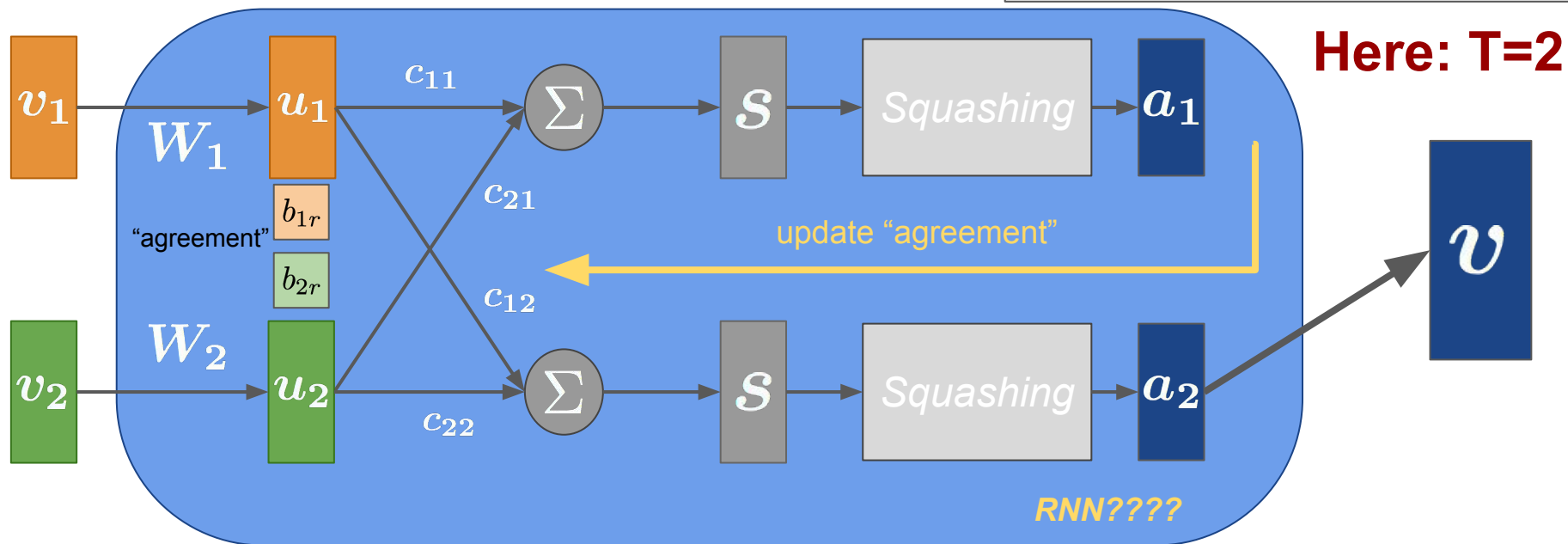
$c_{1r}, c_{2r} = \text{softmax}(b_{1r}, b_{2r})$

$a_r = \text{squashing}(c_{1r}u_1 + c_{2r}u_2)$

$b_{1(r+1)} = b_{1r} + a_r \cdot u_1$

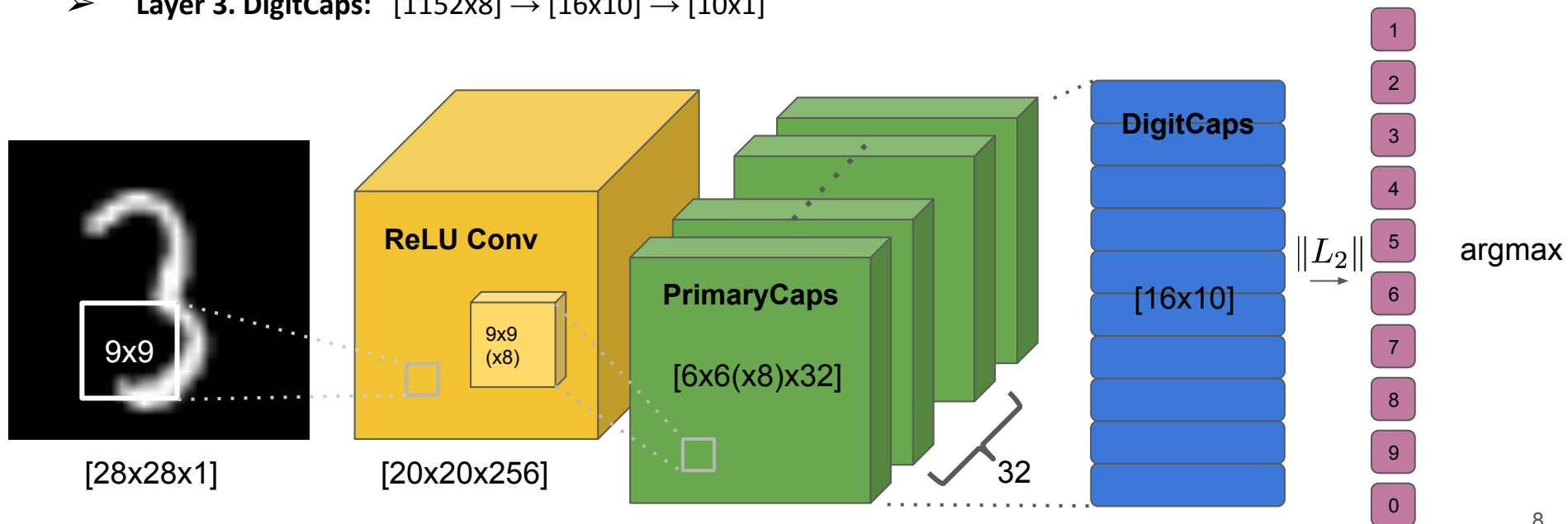
$b_{2(r+1)} = b_{2r} + a_r \cdot u_2$

**Routing Algorithm**



# A Capsule Network (CapsNet) for MNIST

- **Layer 1. ReLU Conv:**  $[28 \times 28 \times 1] \rightarrow [20 \times 20 \times 256]$
- **Layer 2. PrimaryCaps:**  $[20 \times 20 \times 256] \rightarrow [6 \times 6 \times (8) \times 32] \rightarrow [1152 \times 8]$
- **Layer 3. DigitCaps:**  $[1152 \times 8] \rightarrow [16 \times 10] \rightarrow [10 \times 1]$




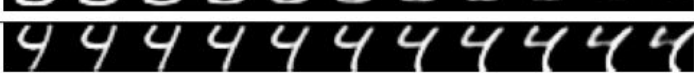

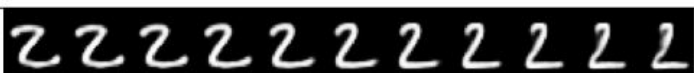


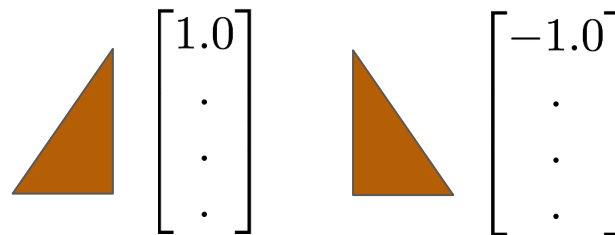


# Interpretable Activity Vector

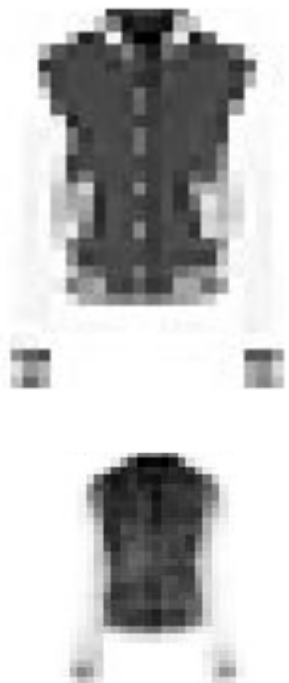
- Each dimension contains a specific information (pattern)



Scale and thickness	
Localized part	
Stroke thickness	
Localized skew	
Width and translation	
Localized part	

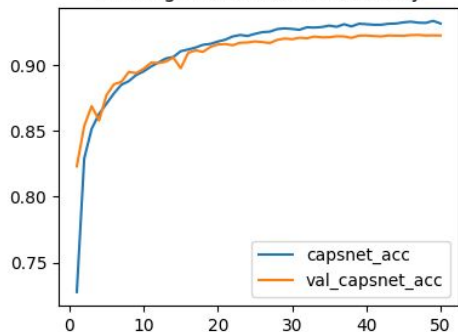
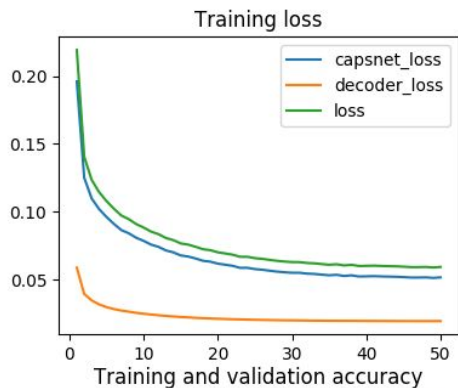


# Interpretable Activity Vector: **Fasion MNIST**

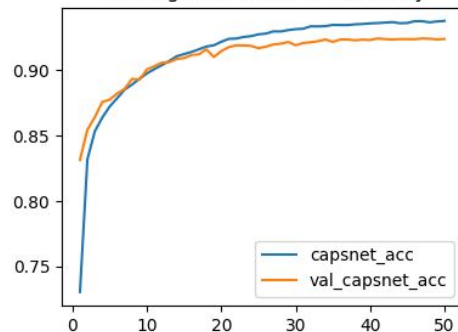
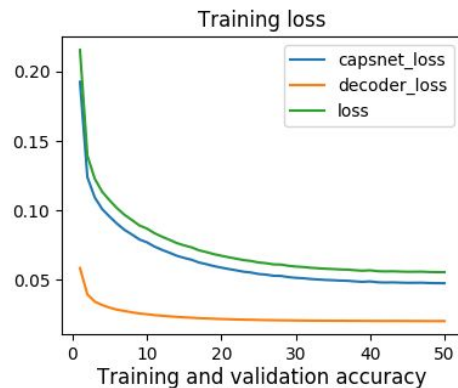


# Routing v.s. Backpropagation (on Fashion MNIST)

- Baseline(left): only using routing
- Ex1: routing+backpropagation

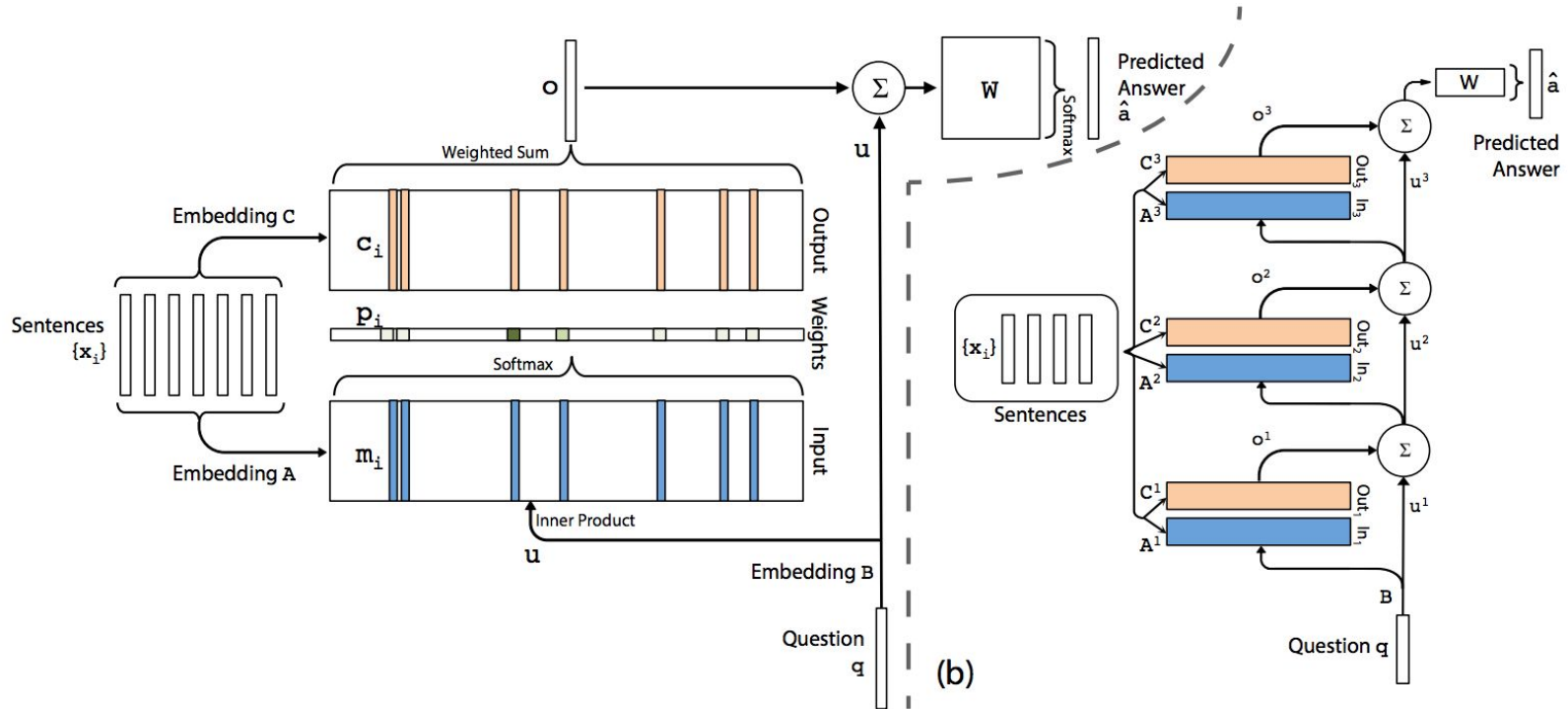


without backpropagation



with backpropagation

# Why Routing? Memory Network (Hopping Mechanism)



# Summary of CapsNet

- Key Points of Capsule:
  - **Vector** → **Vector** (**Tensor** → **Tensor**)
  - **Encapsulate** entity or its **pattern**
  - **Routing** by *agreement*
  - **Invariance** v.s. **Equivariance**
- Uncovered Topics:
  - Margin Loss
  - Reconstruction Network
  - Overlapping Handling
  - ...
- Future works:
  - Other squashing
  - Improving routing process
  - ...

# References of this Section

1. [ Hinton, G. E., Krizhevsky, A., & Wang, S. D. (2011, June). **Transforming autoencoders**. In International Conference on Artificial Neural Networks (pp. 44-51). Springer, Berlin, Heidelberg. ]
2. [ Su, J., Vargas, D. V., & Kouichi, S. (2017). One pixel attack for fooling deep neural networks. *arXiv:1710.08864*. ]
3. [ Hinton, G (2017). **What's wrong with convolutional neural nets**. <https://www.youtube.com/watch?v=Mqt8fs6ZbHk&t=562s> ]
4. [ Sabour, S., Frosst, N., & Hinton, G. E. (2017). **Dynamic Routing Between Capsules**. *arXiv:1710.09829*. ]
5. [ **Under double-blink review (ICLR 2018). Matrix Capsules with EM Routing**. ] Rating results: **4, 6, 7**
6. [ Sukhbaatar, S., Weston, J., & Fergus, R. (2015). End-to-end memory networks. In Advances in neural information processing systems (pp. 2440-2448). ]
7. [ Hung-Yi Lee (2017). Capsule. <https://www.youtube.com/watch?v=UhGWH3hb3Hk>]