

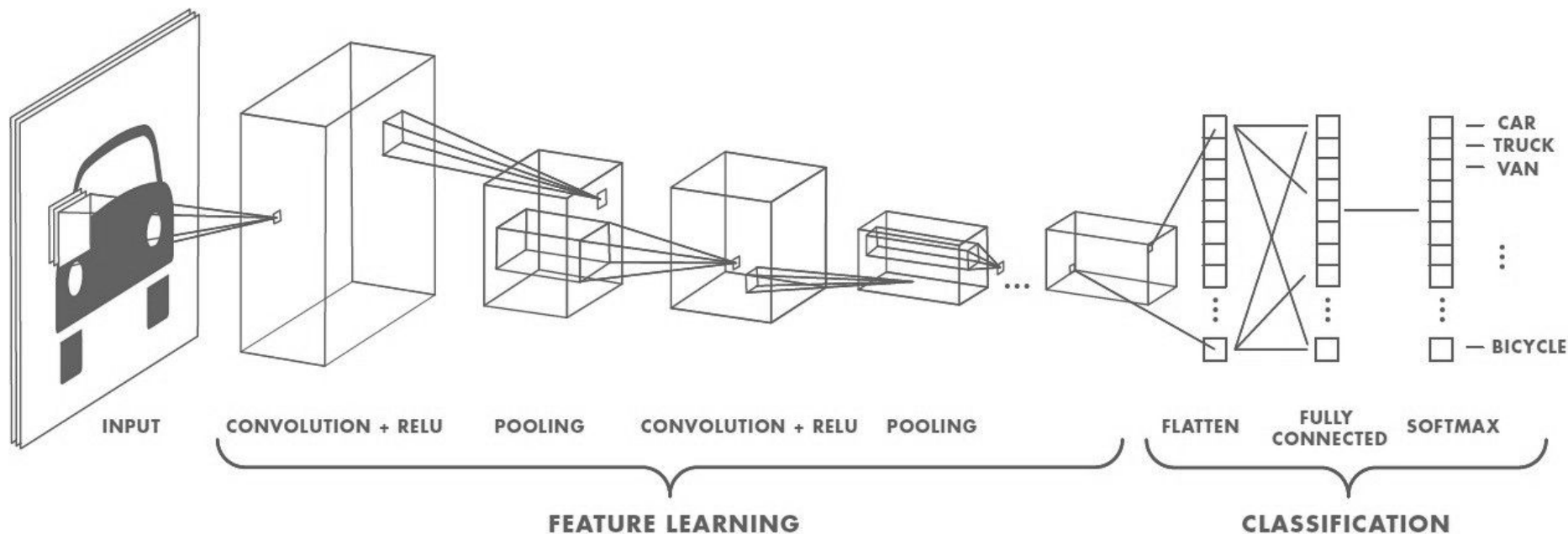
Цифровая обработка изображения

10. Ограничения сверточных нейронных сетей и
варианты решения

План занятия

- Ограничения сверточных сетей в задачах классификации
- Пространственные преобразования
- Деформируемые свертки
- Capsule Net

Классификация изображений с помощью CNN



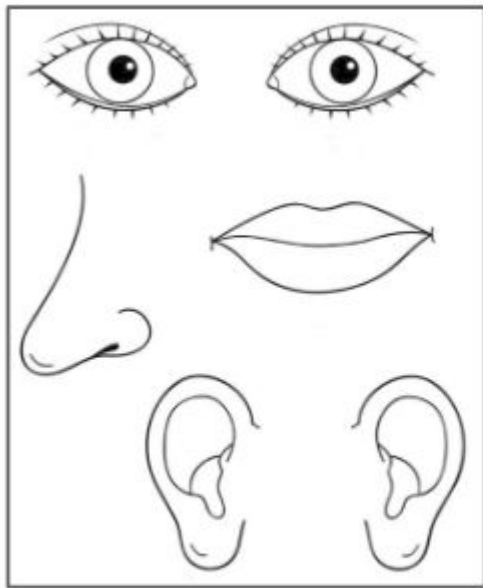
Ограничения сверточных сетей

- при классификации не учитывается пространственное расположение объектов на изображении
- знания об объекте не обобщаются на примеры с различными углами зрения

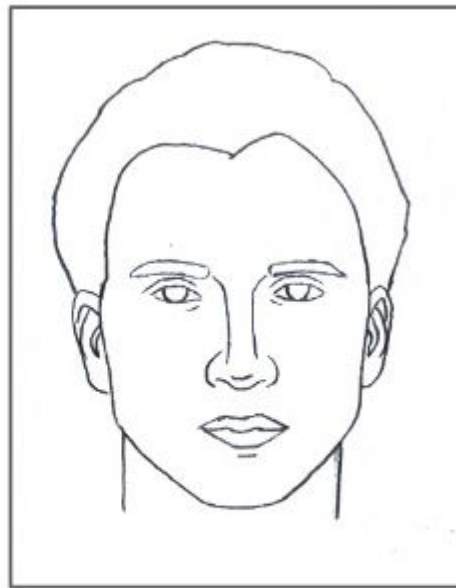
Ограничения сверточных сетей



Ограничения сверточных сетей



Face



Face

Варианты решения

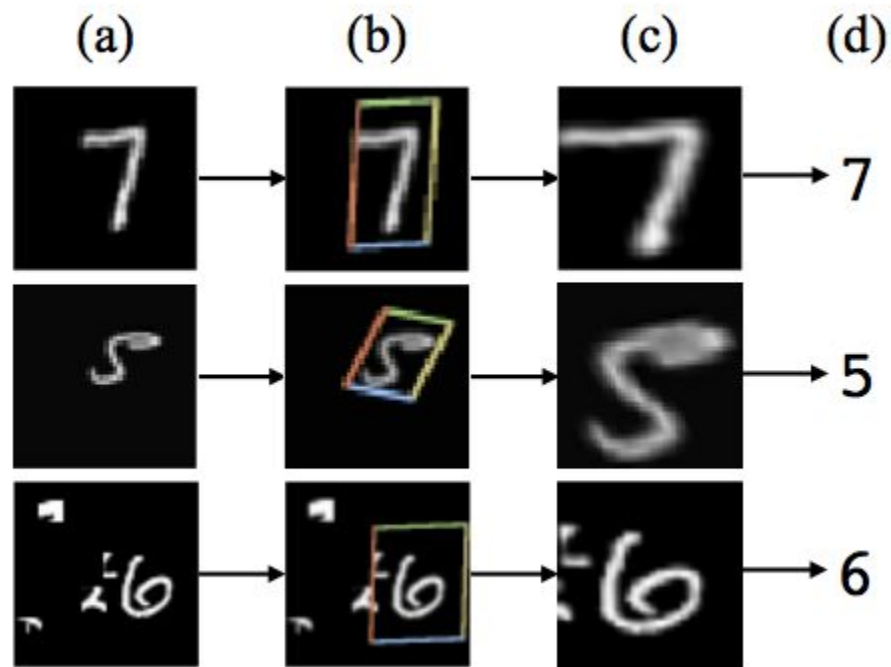
- аугментация изображений при обучении
- отдельная модель для каждого угла зрения
- применять пространственное преобразование изображения перед классификацией (Spatial Transform)
- адаптировать сверточные фильтры в зависимости от входного изображения (Deformable Convolutions)

Пространственное преобразование (Spatial Transformer)

Spatial Transformer

- дополнительный слой, который автоматически применяет пространственные преобразования к входным данным в зависимости от значений входных данных
- Localization network
- Parametrized sampling grid

Spatial Transformer



Spatial Transformer

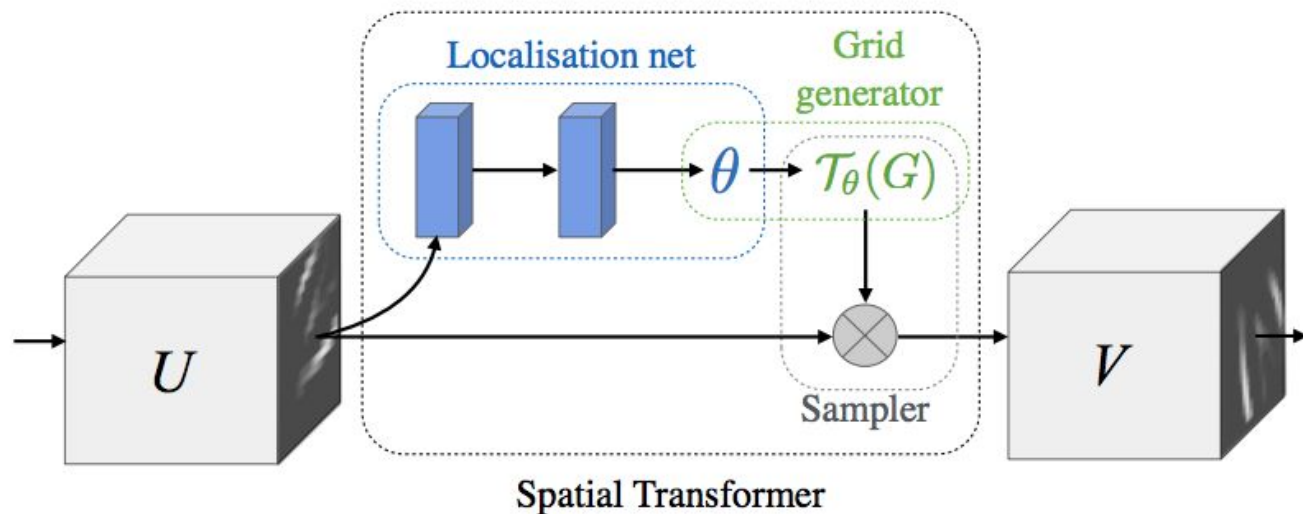


Figure 2: The architecture of a spatial transformer module. The input feature map U is passed to a localisation network which regresses the transformation parameters θ . The regular spatial grid G over V is transformed to the sampling grid $\mathcal{T}_\theta(G)$, which is applied to U as described in Sect. 3.3, producing the warped output feature map V . The combination of the localisation network and sampling mechanism defines a spatial transformer.

Spatial Transformer

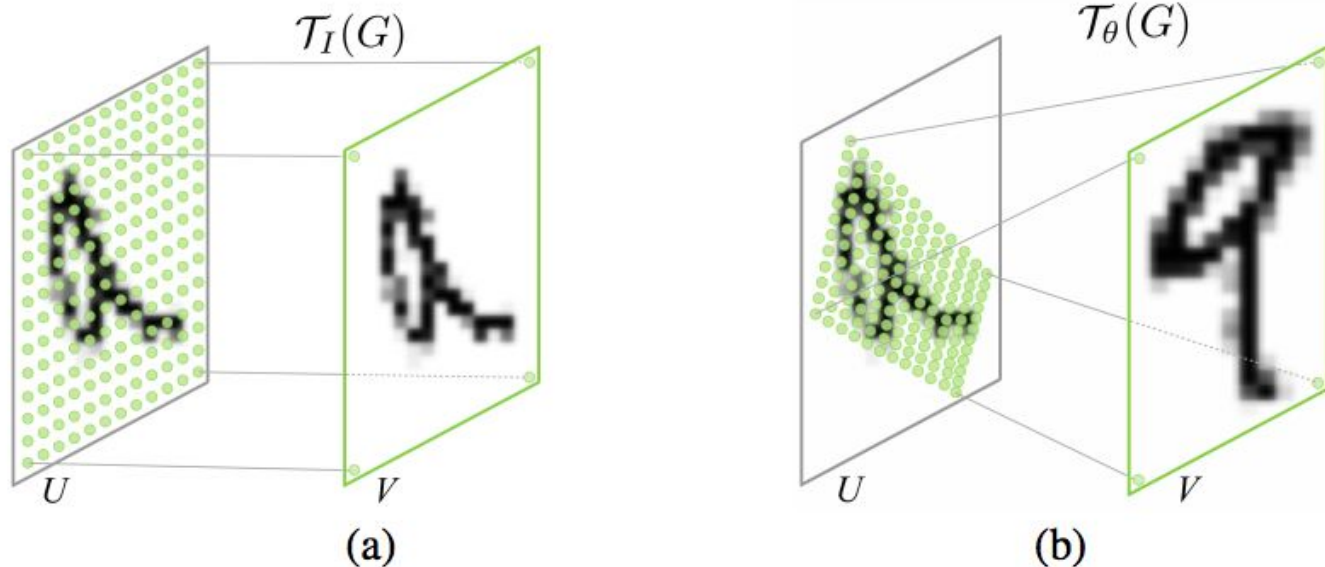


Figure 3: Two examples of applying the parameterised sampling grid to an image U producing the output V . (a) The sampling grid is the regular grid $G = \mathcal{T}_I(G)$, where I is the identity transformation parameters. (b) The sampling grid is the result of warping the regular grid with an affine transformation $\mathcal{T}_\theta(G)$.

Deformable Convolutions

[Deformable Convolutional Networks](#)

Deformable Convolutions

- пространственные изменения фильтров свертки в зависимости от исходных данных
- реализация в виде дополнительный слой в архитектуре
- слой обучается в общем пайплайне
- прирост качества в задачах сегментации и детекции
- <https://github.com/msracver/Deformable-ConvNets> (mxnet)

Deformable Convolutions

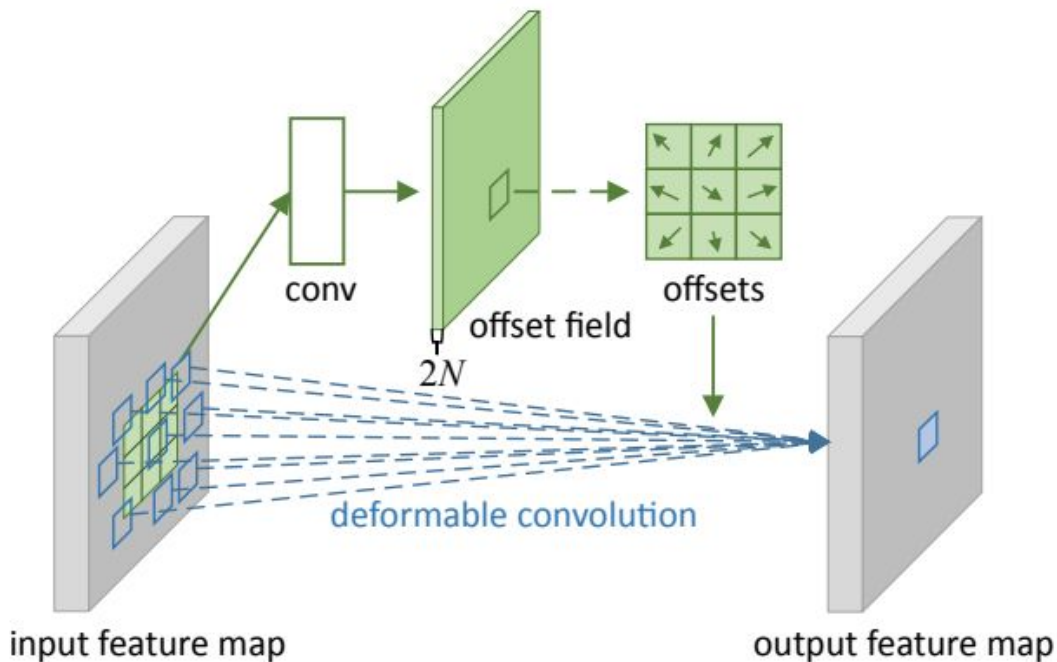
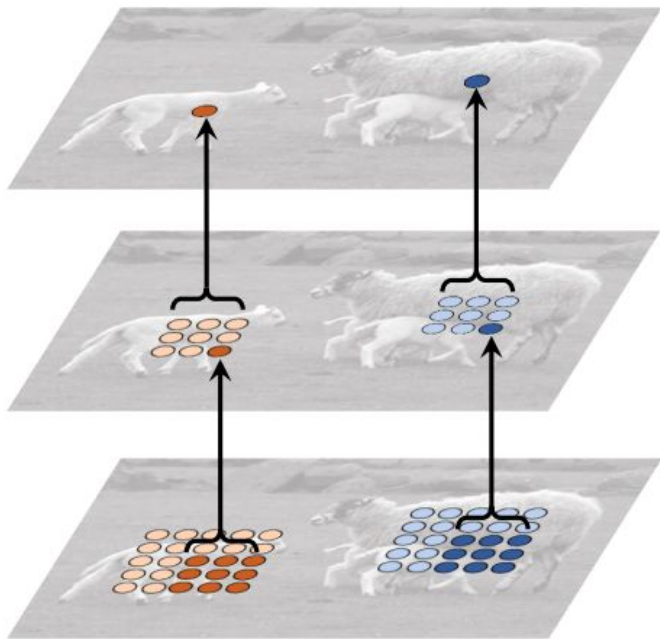
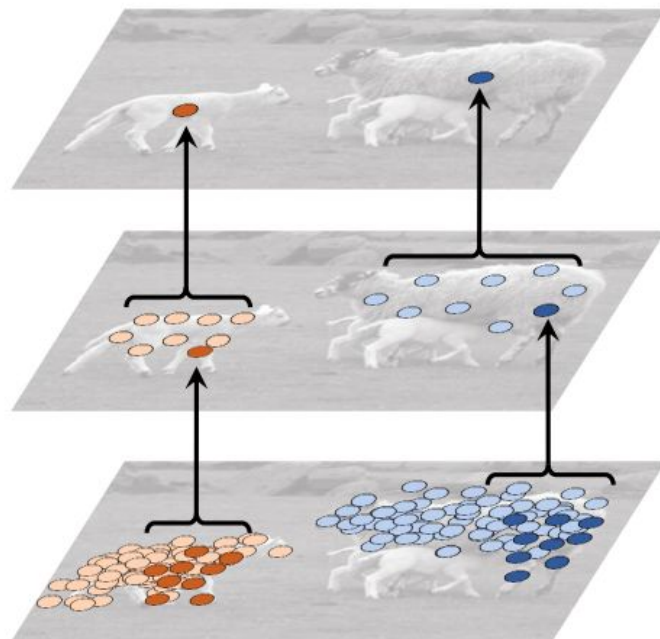


Figure 2: Illustration of 3×3 deformable convolution.

Deformable Convolutions



(a) standard convolution



(b) deformable convolution

Capsule Net

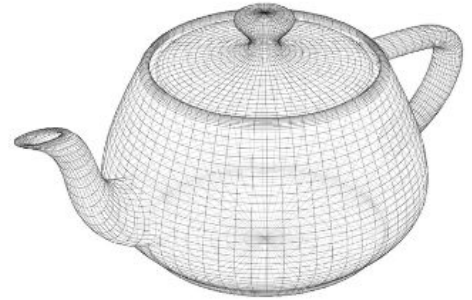
Inverse Computer Graphics

- восстановить 3D модель по компьютерному зрению (Inverse Computer Graphics)
- иерархичная модель объекта
- разбиваем объект на части, определяем положение частей относительно всего объекта
- относительное положение части объекта не зависит от угла зрения

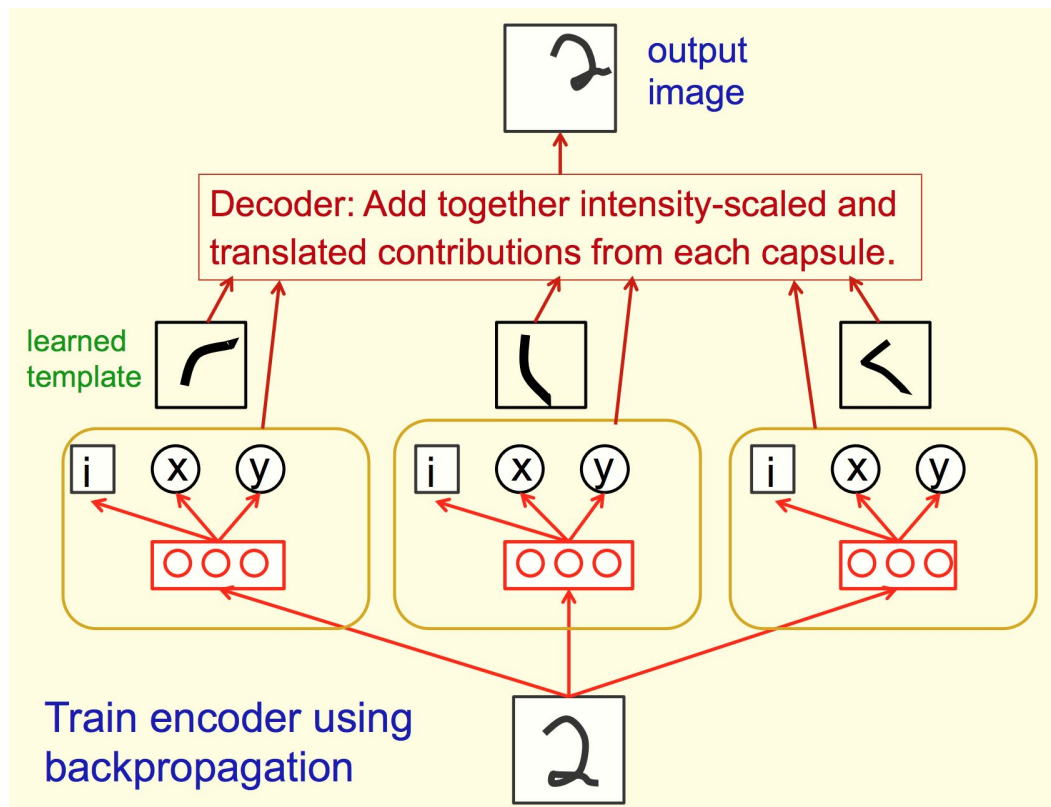
Inverse Computer Graphics



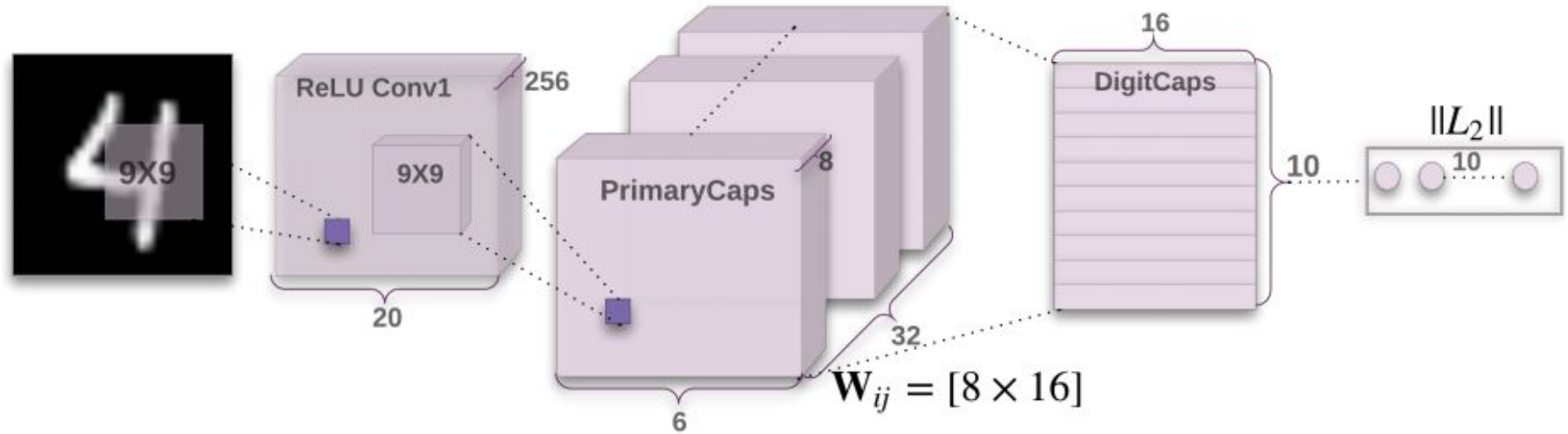
Inverse matrix
containing global
pose information



Capsule Encoder-Decoder



Capsule Net



[Dynamic Routing Between Capsules](#)

Полезные материалы

- [What is wrong with convolutional neural nets?](#)
- [Deformable Convolutional Networks](#)
- [Dynamic Routing Between Capsules](#)