

# Классические задачи компьютерного зрения

**Semantic  
Segmentation**



CAT GRASS  
TREE

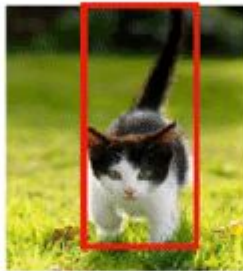
No object  
Just pixels

**Classification**



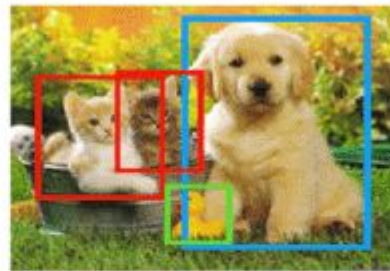
CAT

**Classification  
+ localization**



CAT

**Object detection**



CAT DOG DUCK

**Instance  
segmentation**

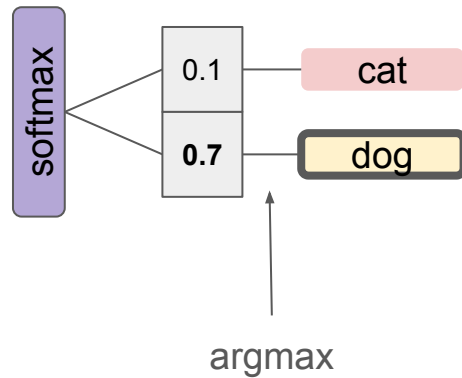
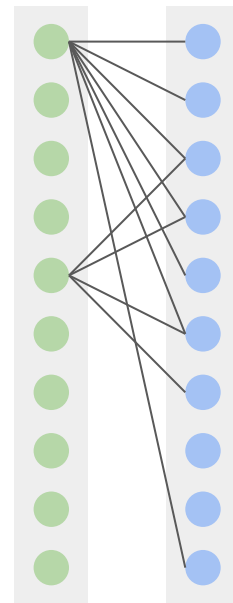
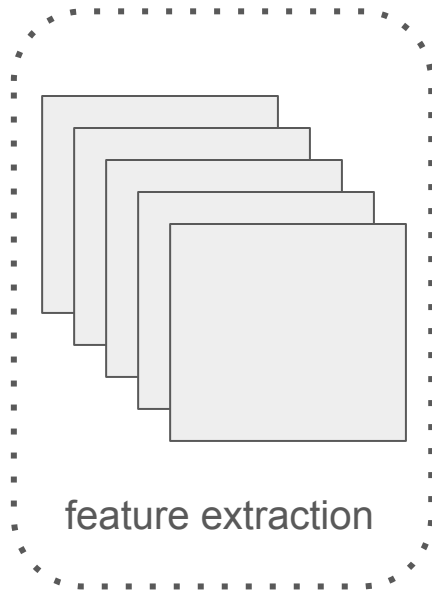
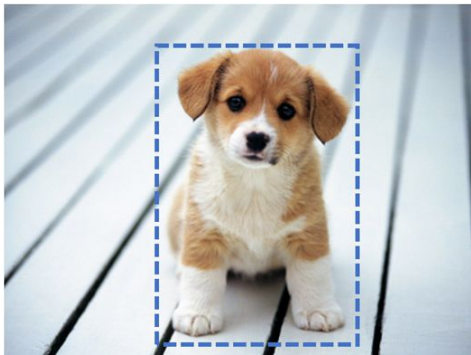


CAT CAT DOG DUC

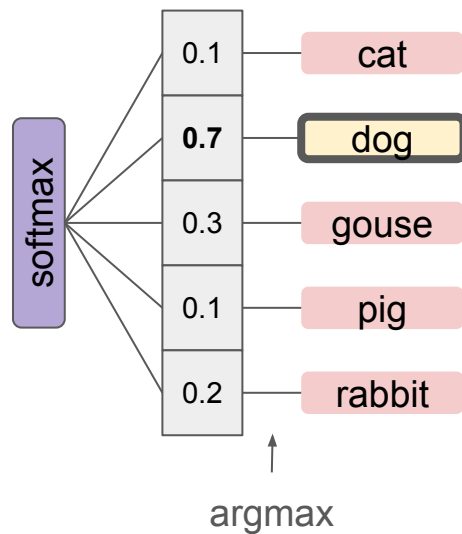
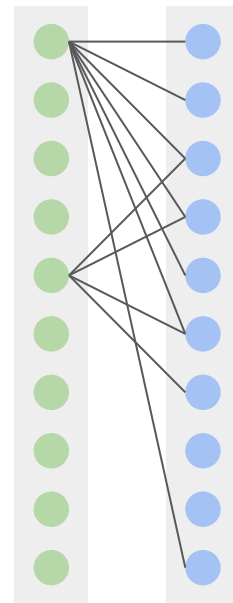
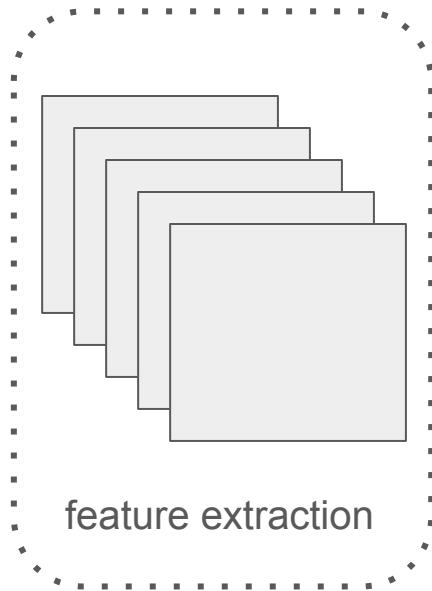
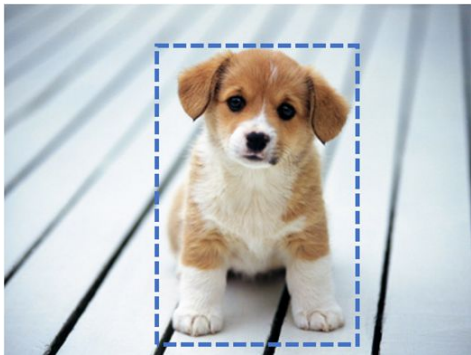
Single object

Multiple objects

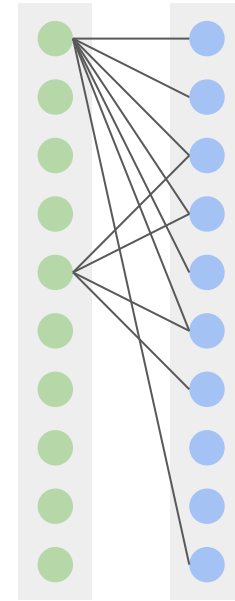
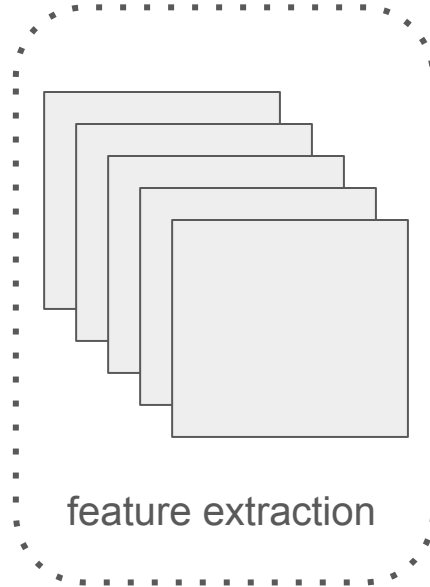
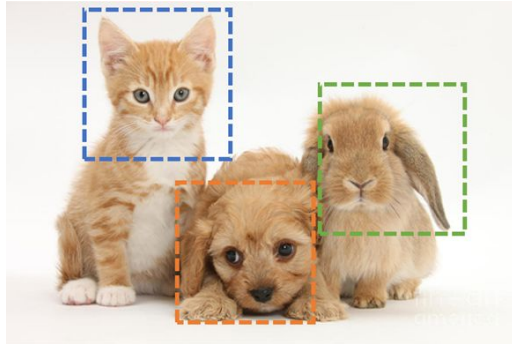
# Бинарная классификация



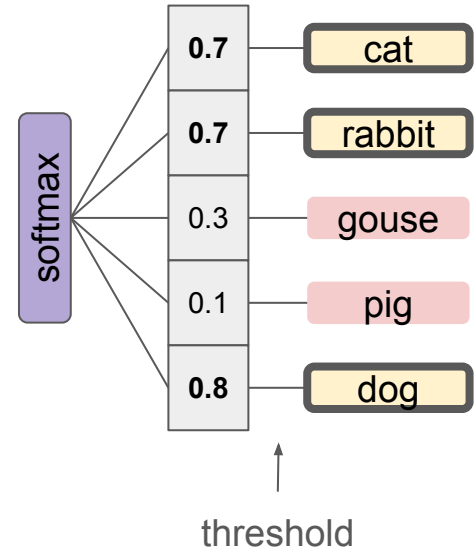
# Многоклассовая классификация



# Мягкая (soft, multilabel) классификация



fully connected



# Классификация. Функции потерь

функции потерь для типов классификации

- бинарная кросс энтропия
- MSE
- MAE
- Categorical Cross-Entropy (для многоклассовой)
- Sigmoid Cross-Entropy loss (для мягкой)
- weighted Sigmoid Cross-Entropy loss

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x^{(i)}) - y^{(i)})^2}$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |h(x^{(i)}) - y^{(i)}|$$

$$MSE = \frac{1}{m} \sum_{i=1}^m (y^{(i)} - \hat{y}^{(i)})^2$$

## Binary Cross Entropy

$$-\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))$$

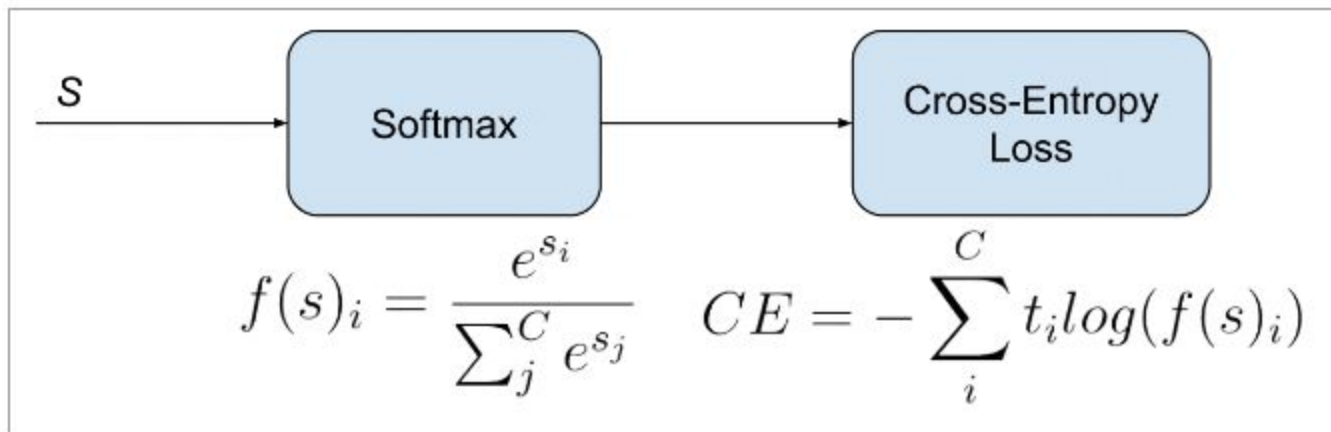
для бинарной классификации

$$\text{logloss} = -\frac{1}{N} \sum_i^N \sum_j^M y_{ij} \log(p_{ij})$$

для многоклассовой классификации

- N is the number of rows
- M is the number of classes

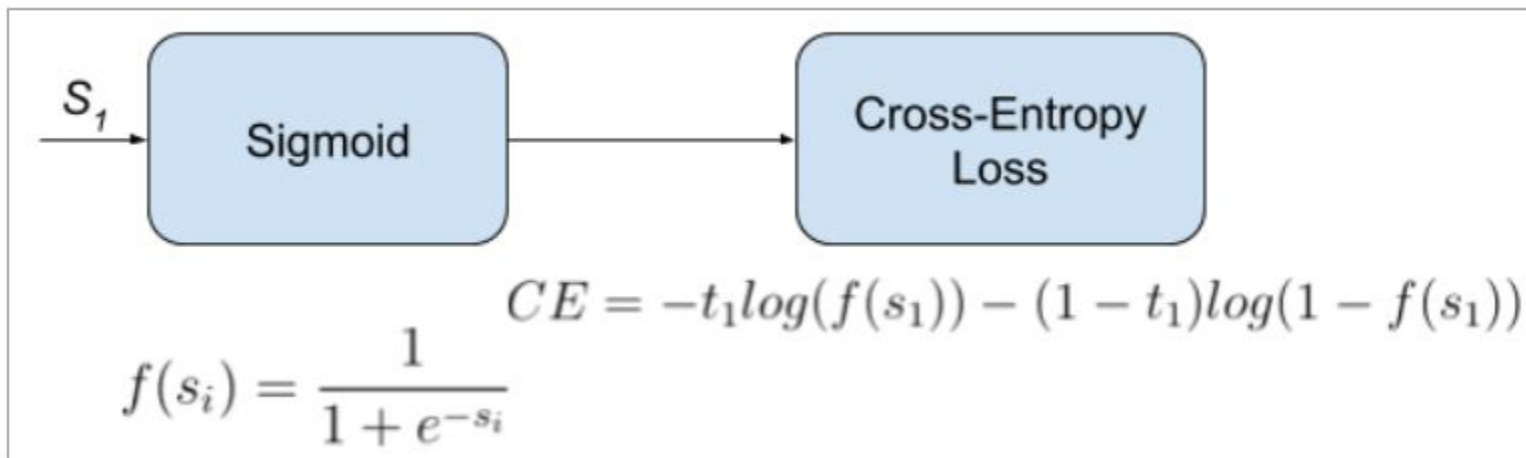
## Categorical Cross-Entropy loss



для многоклассовой классификации



## Sigmoid Cross-Entropy loss

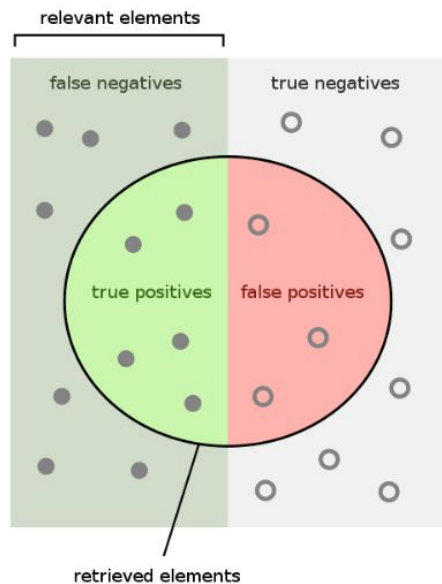


для мягкой классификации

## Еще про метрики:

- самые популярные метрики для классификации
- какой лосс лучше для передачи свойств изображений?
- как задизайнили лосс для multilabel classification
- еще про разработку лосса для multilabel classification
- про distance metrics для multilabel classification

# Классификация. Метрики



$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$F1 \text{ Score} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$$

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}$$

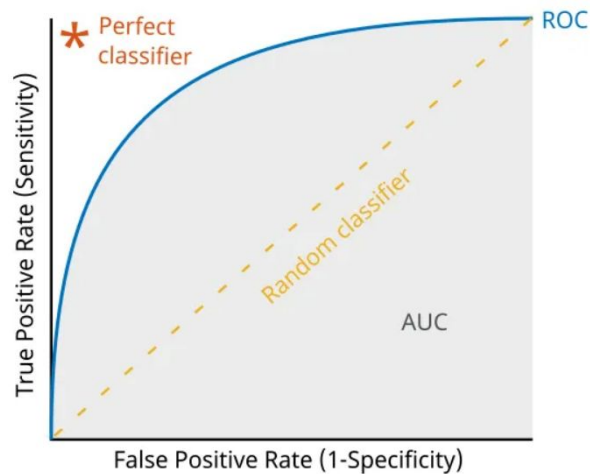
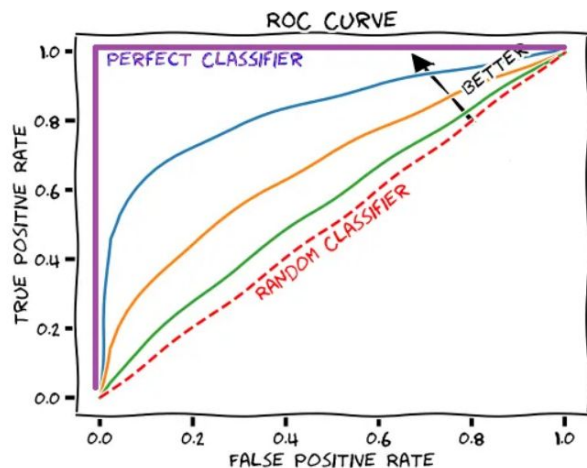
How many retrieved items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are retrieved?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

# Классификация. Метрики



cat	826	48	12	24	32	30	12	5	4	7
dog	111	801	17	18	28	13	7		2	3
horse	13	17	915	22	14	3	9	2	1	4
deer	24	13	14	898	28	14	5	2	1	1
bird	30	8	5	13	892	17	26	4	2	3
frog	27	4	1	3	16	943	5	1		
airplane	8	1	5	4	21	5	923	23	4	6
ship	4	1	1		4	2	37	931	10	10
automobile			1		2		5	5	972	15
truck	3		1		3	2	20	9	39	923

True Class

Predicted Class

# Классификация. Датасеты

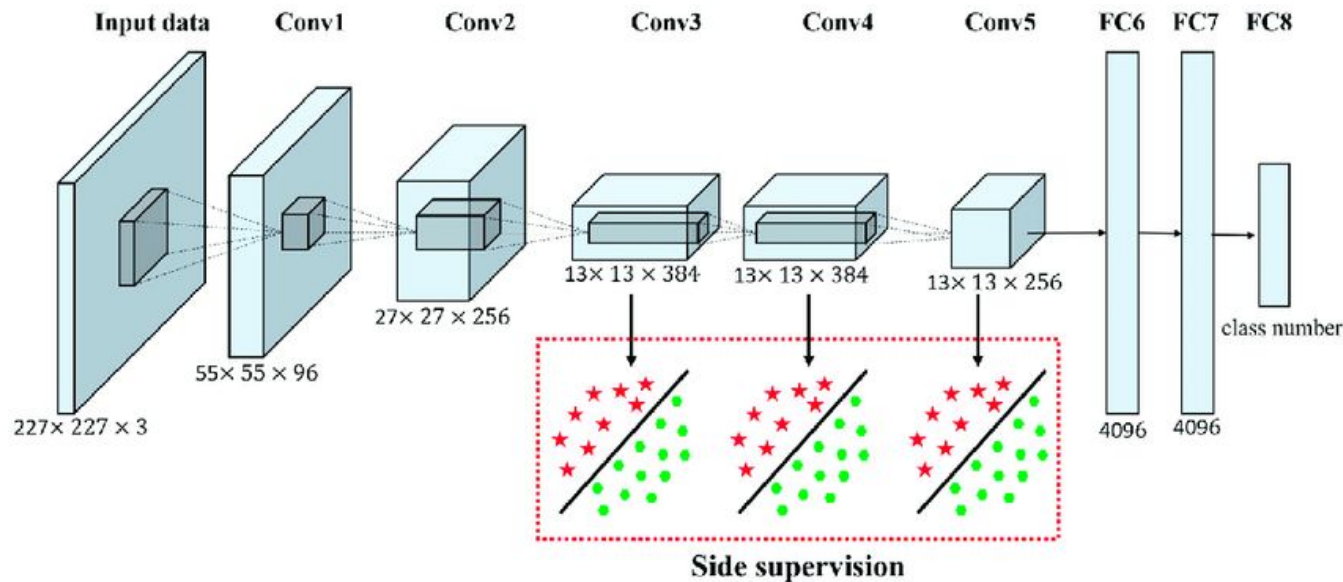
Бинарная\многоклассовая классификация	Мягкая классификация
<u><a href="#">CIFAR-10, CIFAR-100</a></u>	<u><a href="#">CelebA</a></u>
<u><a href="#">Fashion-MNIST</a></u>	
<u><a href="#">Flower Dataset</a></u>	
<u><a href="#">Food-101</a></u>	
<u><a href="#">ImageNet</a></u>	

# Где найти еще данных?

- [Kaggle](#)
- [Papers with codes](#)
- [Roboflow](#)
- наскрапить)

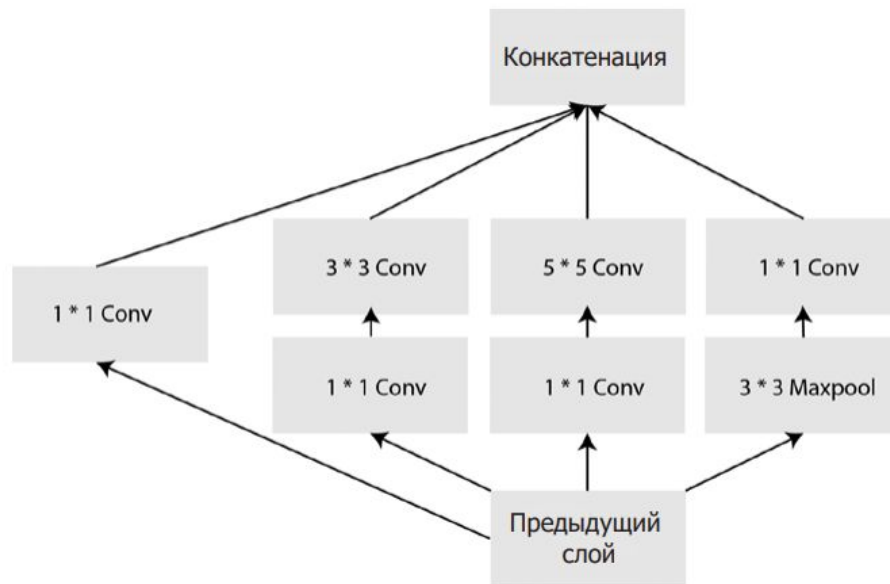
# Немного истории. AlexNet

- 2012 году уничтожила конкурентов в соревновании ImageNet
- одна из первых CNN
- использовала концепции MaxPool и DropOut



# Немного истории. Inception

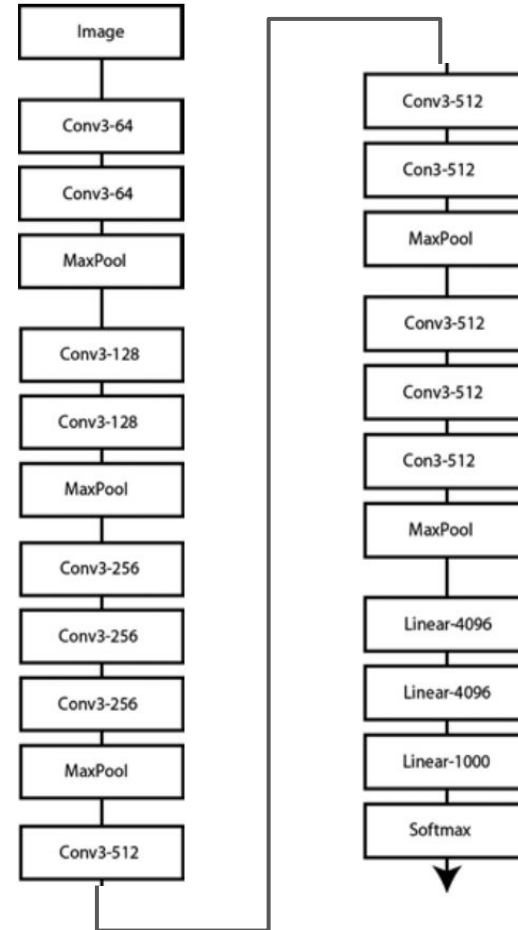
- 2014 победитель в соревновании ImageNet
- фиксированные ядра свертки
- комбинация разных ядер сверток





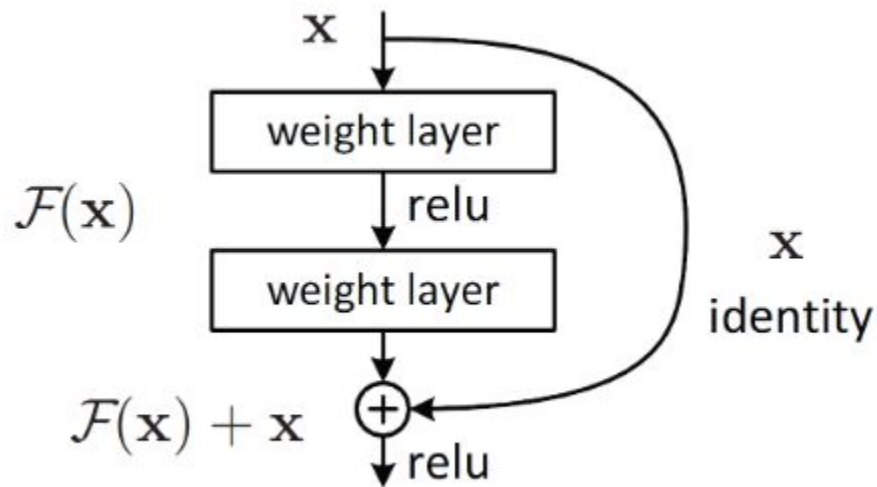
# Немного истории. VVG-16

- 2014 2ое место в соревновании ImageNe
- простые комбинации слоев
- быстро раздувается кол-во параметров



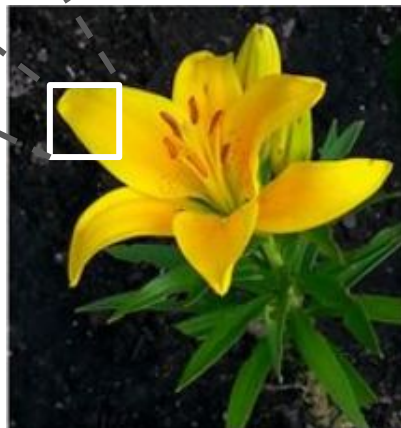
# Немного истории. ResNet

- 2015 топ-1 в соревновании ImageNet
- skip connection
- до сих пор используется во многих задачах



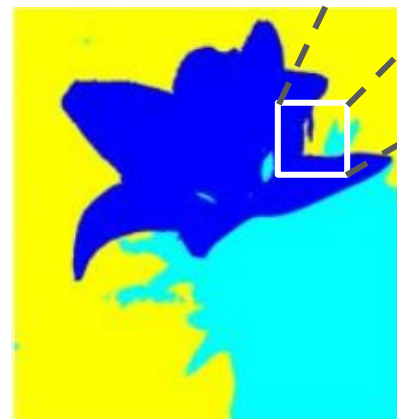
# Сегментация изображений. Постановка задачи

100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100	100	100

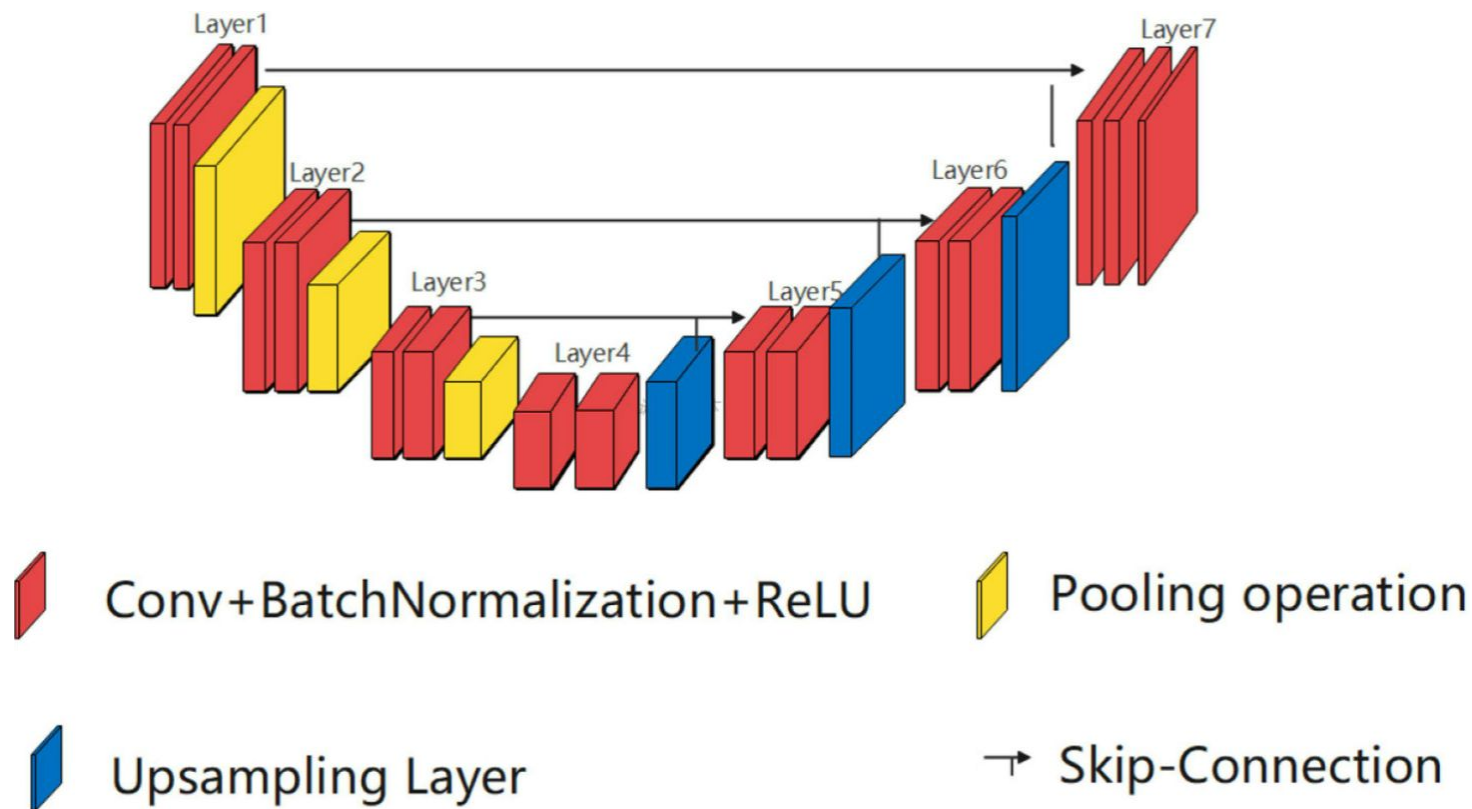


классы

0	0	1	1	1	1	0	0
0	0	1	1	1	1	0	0
0	0	1	1	1	1	0	0
1	1	1	1	1	1	0	0
2	1	1	1	1	1	0	0
2	1	2	1	1	2	2	1
2	2	2	1	1	2	2	1
2	2	2	1	2	2	2	2



# Классический U-net



# Сегментация. Функции потерь

- бинарная кросс энтропия (и ее взвешенные варианты)

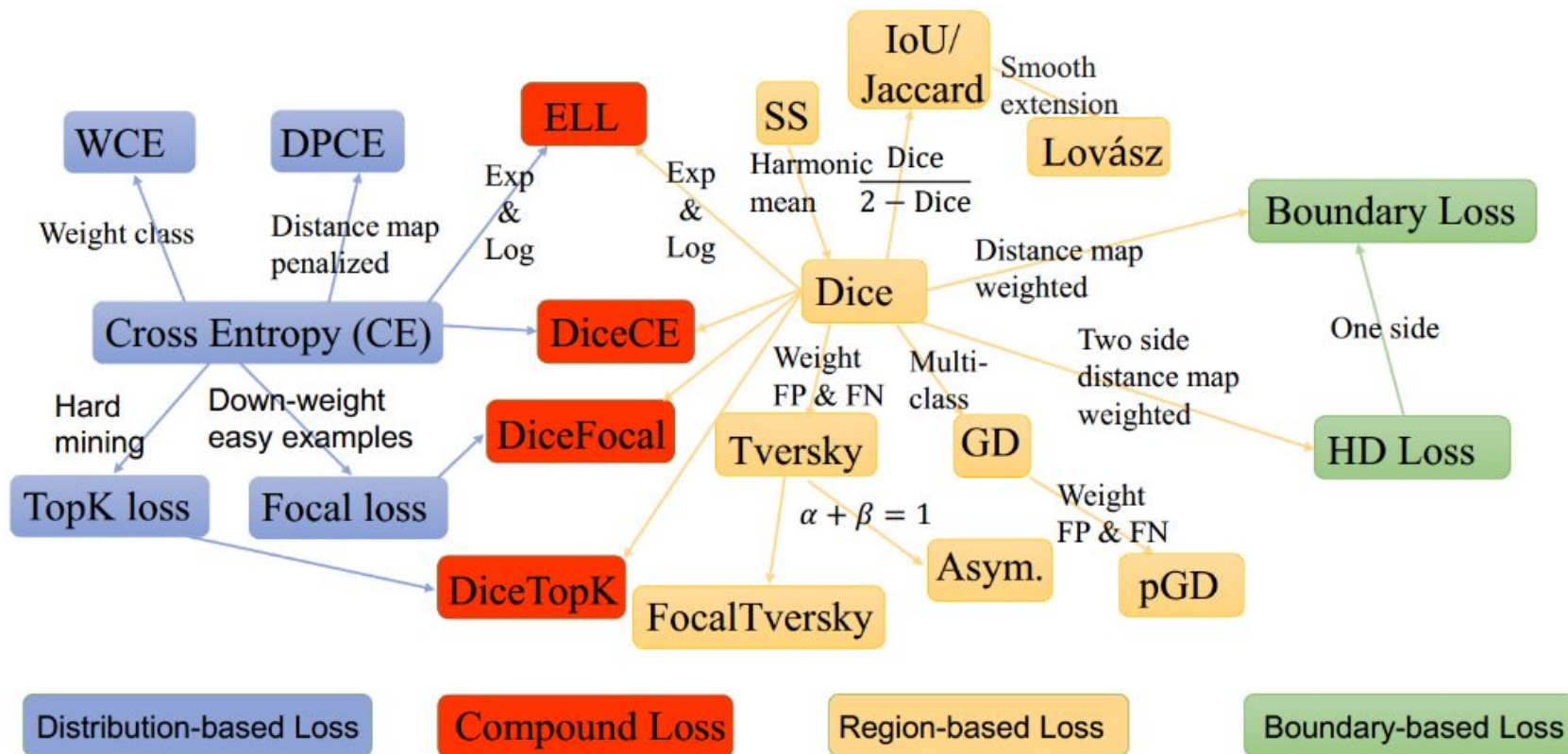
- focal loss

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t)$$

- Dice loss

$$DL(y, \hat{p}) = 1 - \frac{2y\hat{p} + 1}{y + \hat{p} + 1}$$

# Segmentations losses



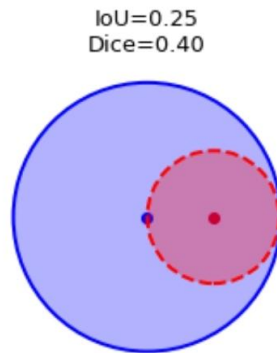
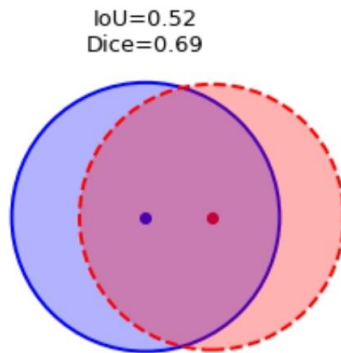
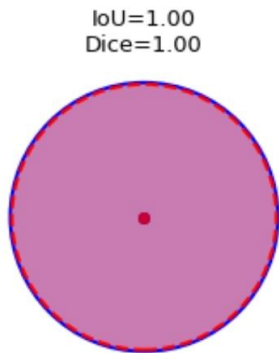
# Сегментация. Метрики

Те же, что и у классификации: accuracy, precision, recall, f1, conf. matrices.

А также:

1) Dice score  $DSC(A, B) := \frac{2|A \cap B|}{|A| + |B|}$

2) Jaccard or IoU  $J(A, B) := \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$



# Сегментация. Датасеты

- [CelebA](#)
- [COCO](#) (Microsoft Common Objects in Context)
- [Cityscapes](#)
- [PascalVOC](#)



# Сегментация. Что почитать

- [статья](#) про основные лосс функции для сегментации
- [разбор](#) лосс функций для сегментации (есть про Boundary loss)