

Свёрточные нейронные сети. Задача сегментации изображения.

Практикум на ЭВМ для 317 группы, весна 2019

Попов Артём

Кафедра ММП ВМК МГУ

План занятия

- свёрточные нейронные сети (convolutional neural networks, CNNs)
- основные идеи современных архитектур
- представления внутри свёрточных нейросетей
- задача сегментации изображений (instance segmentation)
- современные архитектуры для решения задач сегментации

ImageNet (image-net.org)

Soccer, association football

A football game in which two teams of 11 players try to kick or head a ball into the opponents' goal

1395 pictures
90.76% Popularity Percentile
Wordnet IDs

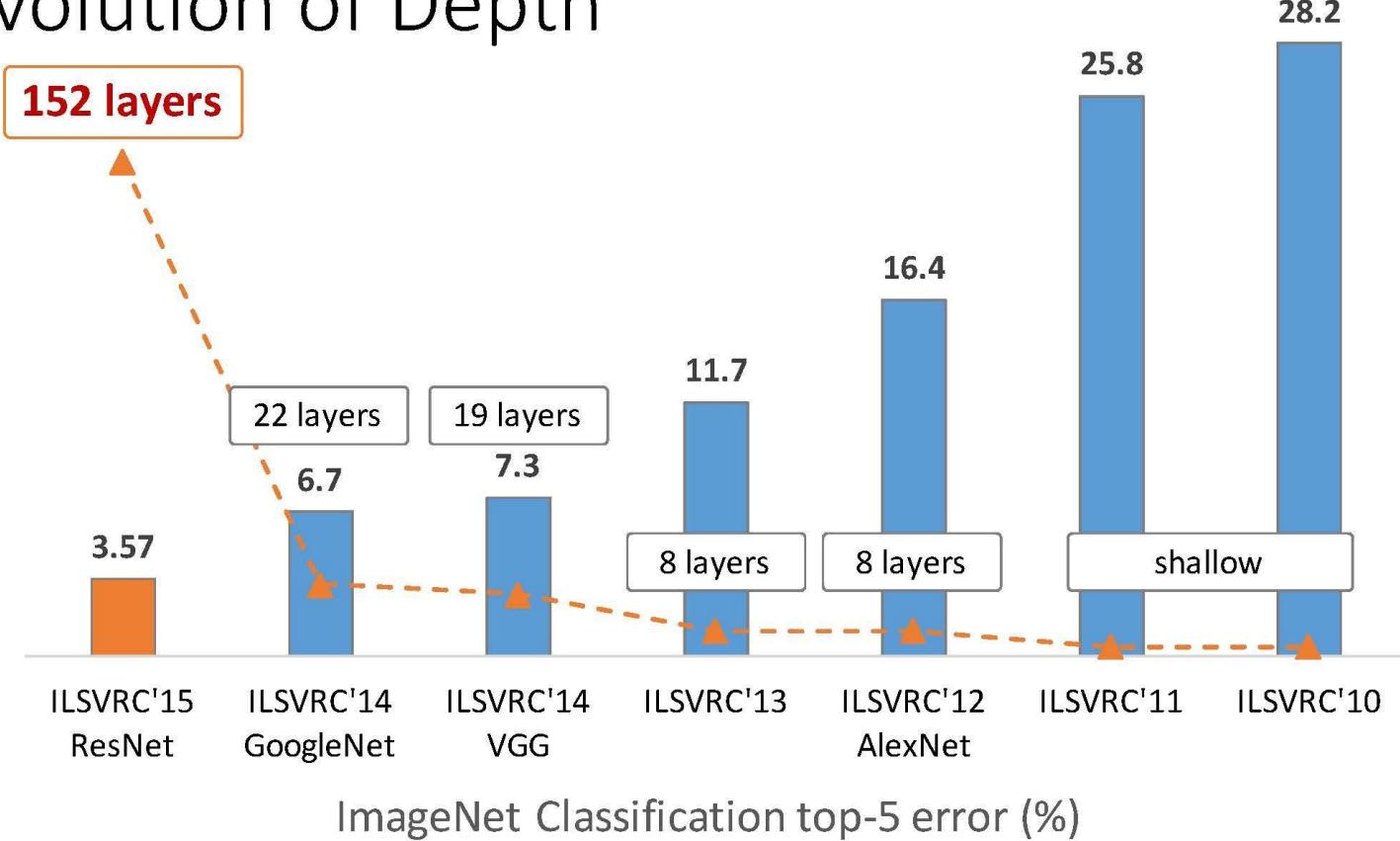
Numbers in brackets: (the number of synsets in the subtree).

- + ImageNet 2011 Fall Release (3226)
 - plant, flora, plant life (4486)
 - geological formation, formation (1)
 - natural object (1112)
 - + sport, athletics (176)
 - rowing, row (2)
 - funambulism, tightrope walking (0)
 - judo (0)
 - blood sport (10)
 - gymnastics, gymnastic exercis (0)
 - water sport, aquatics (19)
 - track and field (5)
 - outdoor sport, field sport (17)
 - + contact sport (18)
 - boxing, pugilism, fisticuffs (0)
 - wrestling, rassling, grapplin (0)
 - ice hockey, hockey, hockey (0)
 - football, football game (5)
 - rugby, rugby football, rug (0)
 - professional football (0)
 - + American football, Amer (0)
 - + soccer, association foot (0)
 - team sport (0)
 - racing (7)
 - athletic game (70)
 - riding, horseback riding, equit (0)
 - archery (0)
 - cycling (3)
 - sledding (3)
 - skating (6)

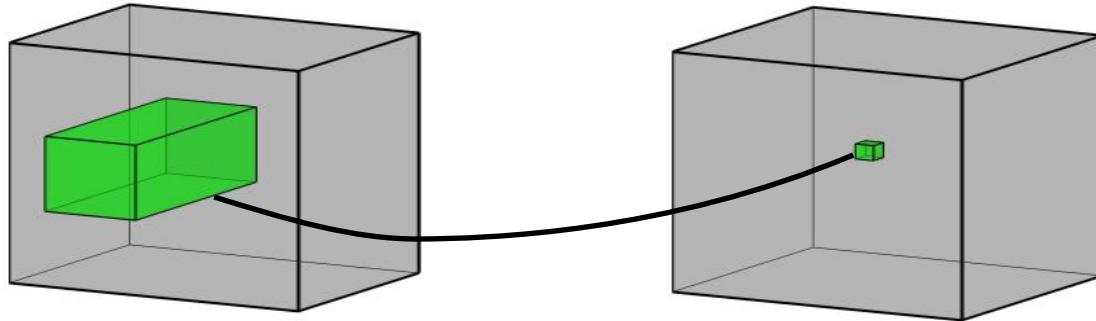


14,197,122
изображений
21841 категорий

Revolution of Depth



Операция свёртки



Операция свёртки многоканального изображения K с ядром U :

$$V(x, y, t) = \sum_{i=x-\delta}^{x+\delta} \sum_{j=y-\delta}^{y+\delta} \sum_{s=1}^S K(i - x + \delta, j - y + \delta, s, t) \cdot U(i, j, s)$$

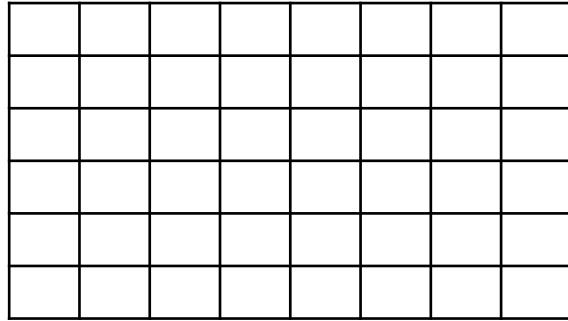
Свёртка реализуется как матричное умножение

Для случая одномерной свёртки вектора x с ядром размерности 3:

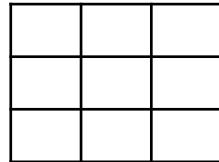
$$y = \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} = \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} x$$

Diagram illustrating the convolution operation. On the left, input vector y is shown as a vertical bar divided into six segments. In the middle, the equation $y =$ is followed by a large bracketed matrix. This matrix is a 6x6 identity matrix where each 2x2 block along the main diagonal is replaced by a 3x3 kernel matrix. The kernel matrix has yellow, grey, and orange blocks. On the right, input vector x is shown as a vertical bar divided into three segments.

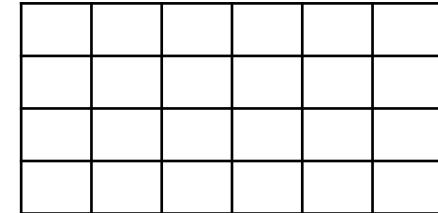
Использование паддинга (padding)



*

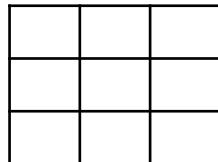


=

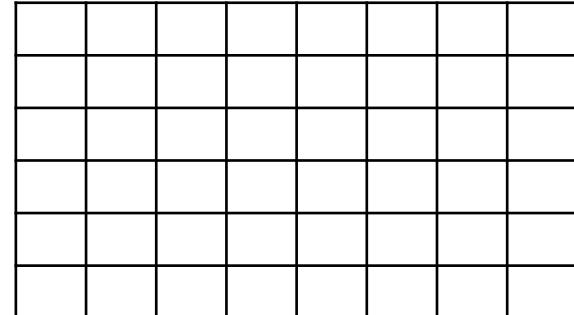


o	o	o	o	o	o	o	o	o	o
0									0
0									0
0									0
0									0
0									0
0									0
0	0	0	0	0	0	0	0	0	0

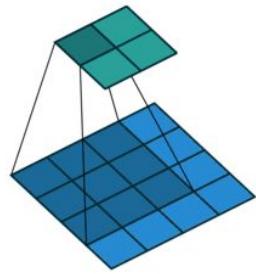
*



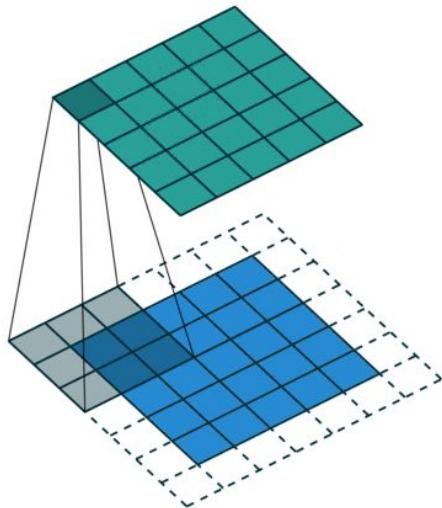
=



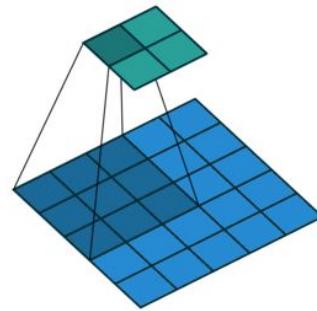
Различные параметры свёрток



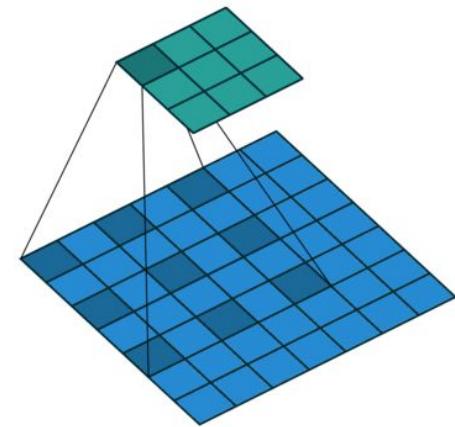
Padding
отсутствует,
без страйдов



Padding
присутствует,
без страйдов

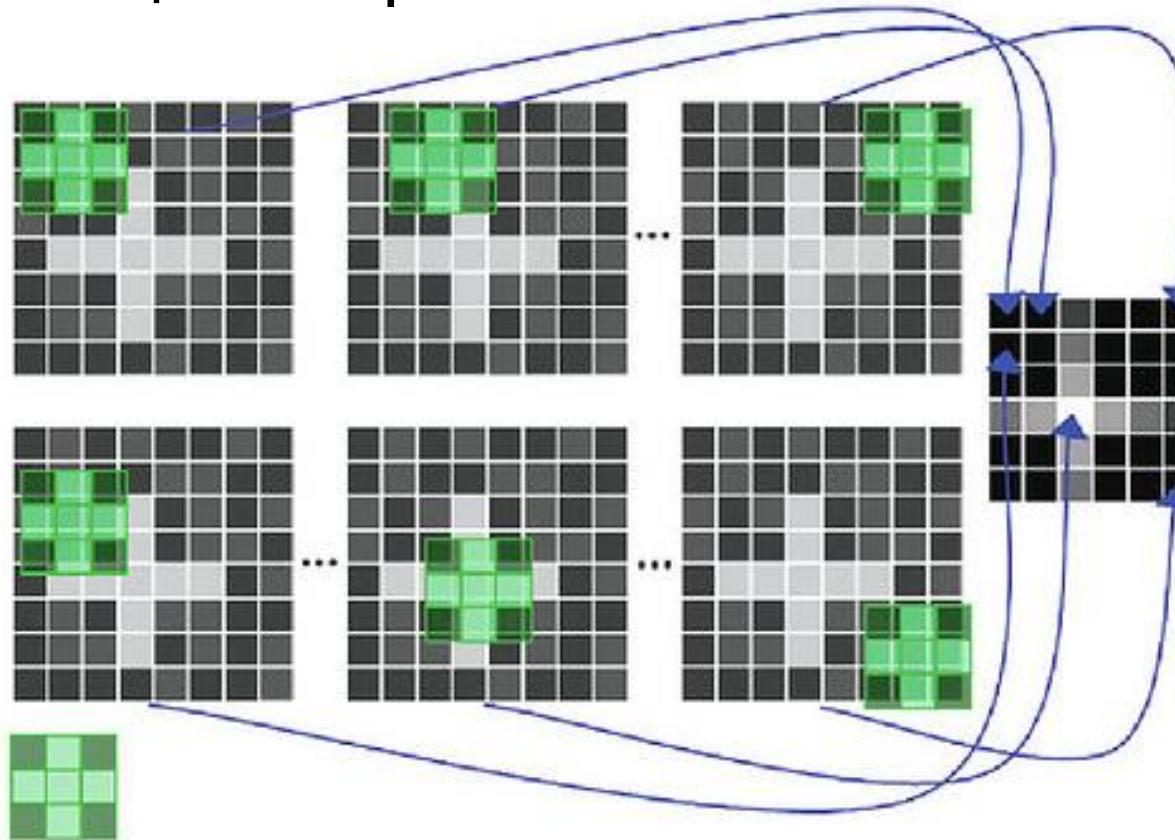


Padding
отсутствует,
Страйд = 2

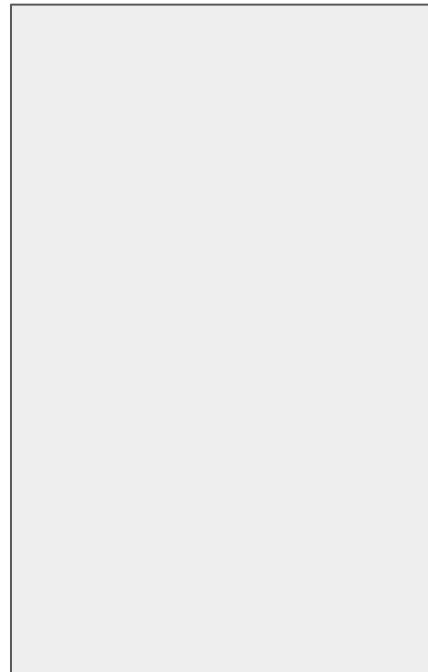


Padding
отсутствует,
без страйдов,
dilation = 2

Интерпретация свёртки: поиск шаблона



Как выглядит ядро свёртки?



Как выглядит ядро свёртки?



$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

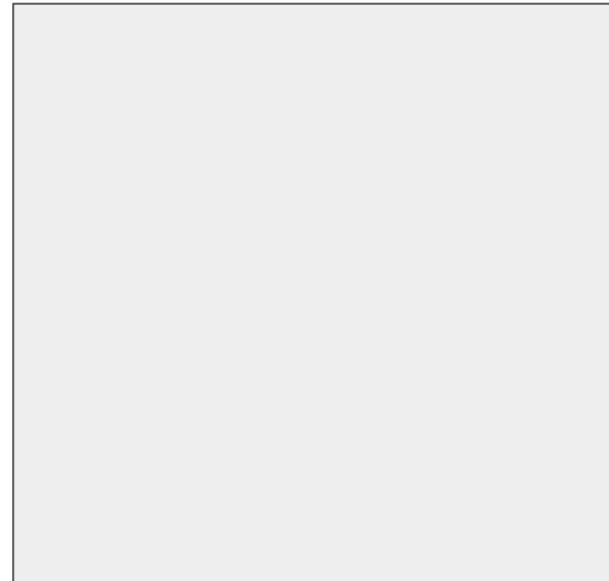
$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$



Что получится после применения свёртки?



$$\frac{1}{10} \cdot \begin{vmatrix} -1 & -2 & -1 \\ -2 & 22 & -2 \\ -1 & -2 & -1 \end{vmatrix}$$



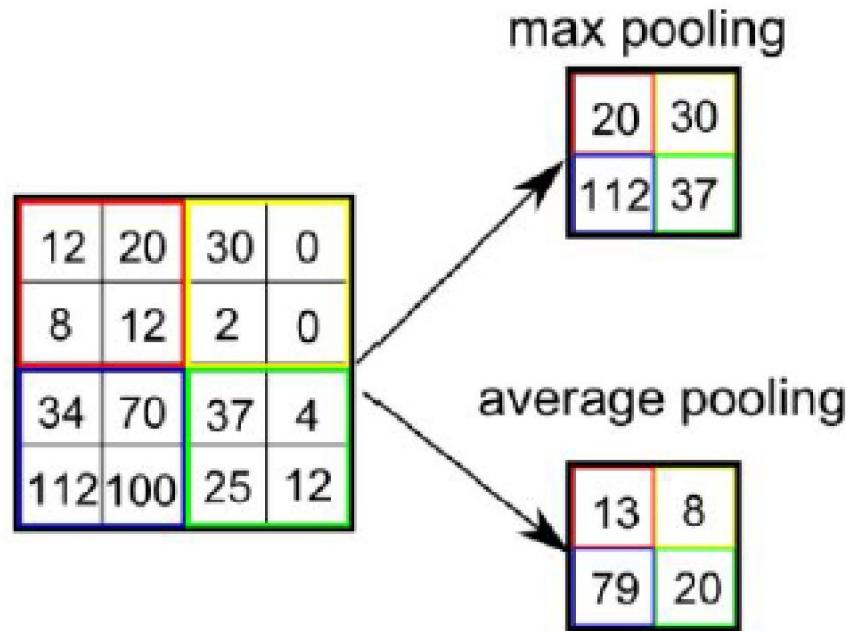
Что получится после применения свёртки?



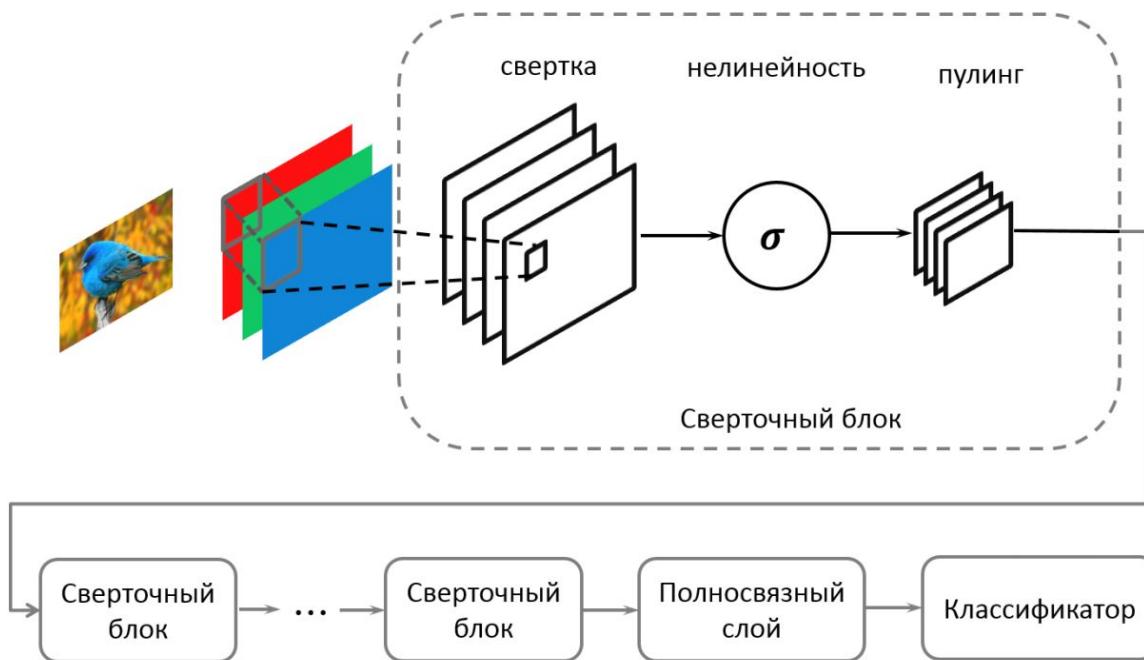
$$\frac{1}{10} \cdot \begin{vmatrix} -1 & -2 & -1 \\ -2 & 22 & -2 \\ -1 & -2 & -1 \end{vmatrix}$$



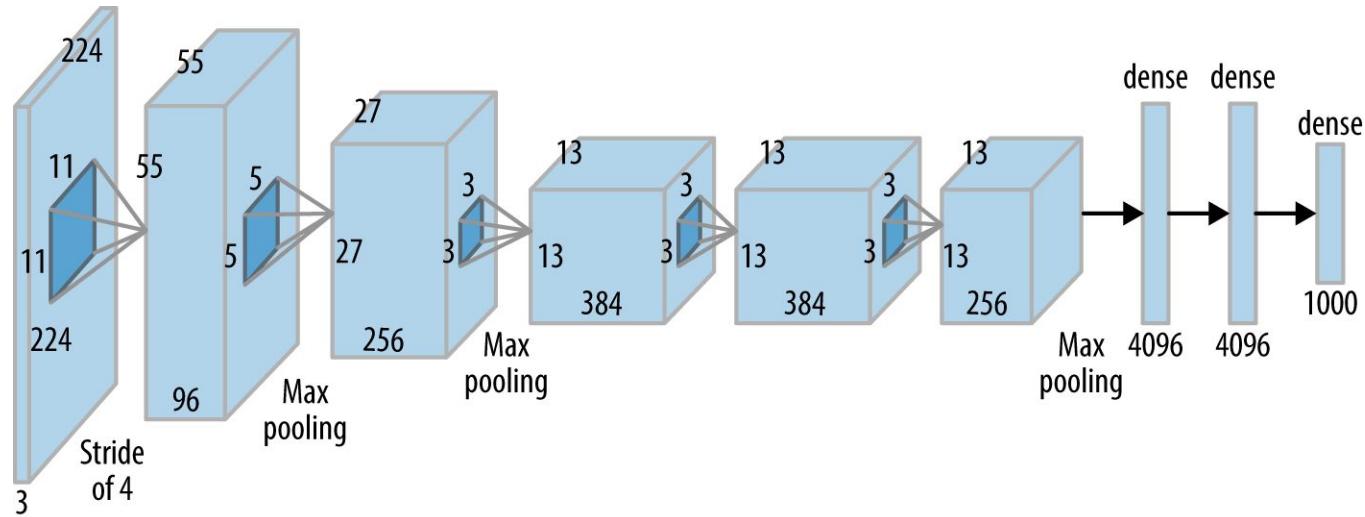
Max pooling и average pooling



Простая схема свёрточных сетей



Сеть AlexNet (2012)



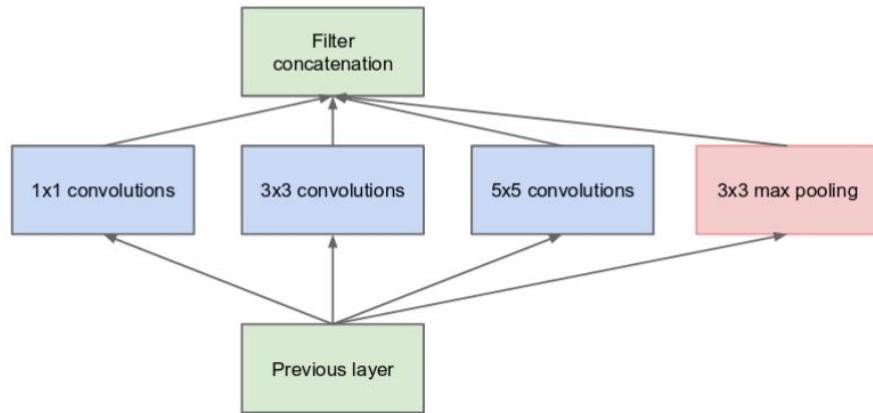
- Max pooling + ReLu + 3 полно связных слоя
- + Аугментация
- + Dropout

Известные архитектуры свёрточных нейросетей для классификации

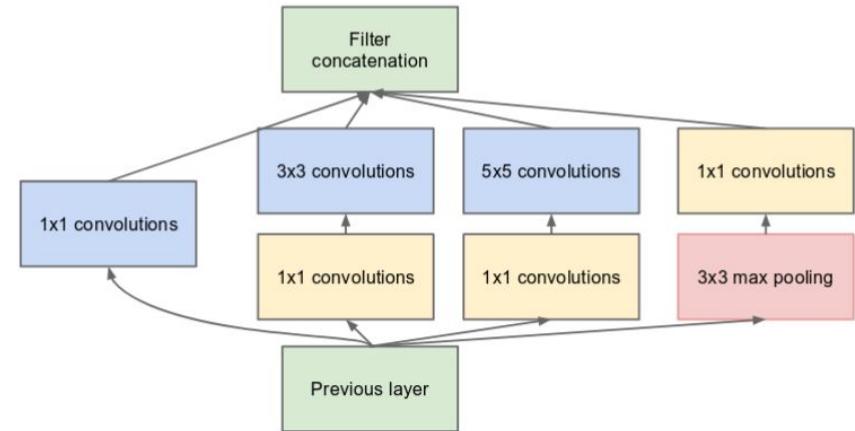
- AlexNet
- VGG
- Inception
- ResNet
- DenseNet

<https://pytorch.org/docs/master/torchvision/models.html>

Каскады свёрток

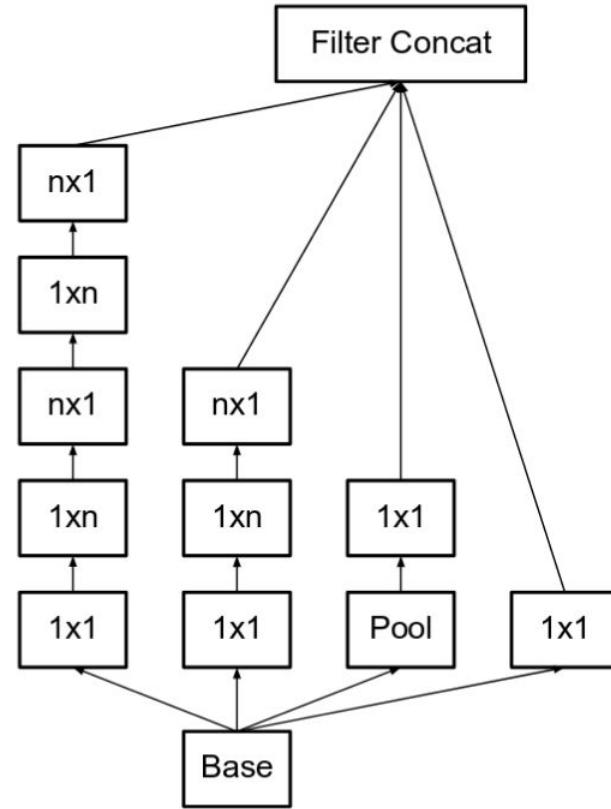
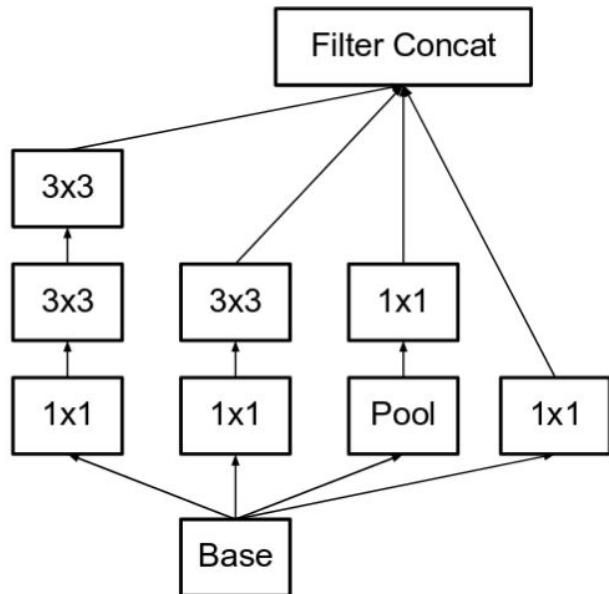


(a) Inception module, naïve version



(b) Inception module with dimension reductions

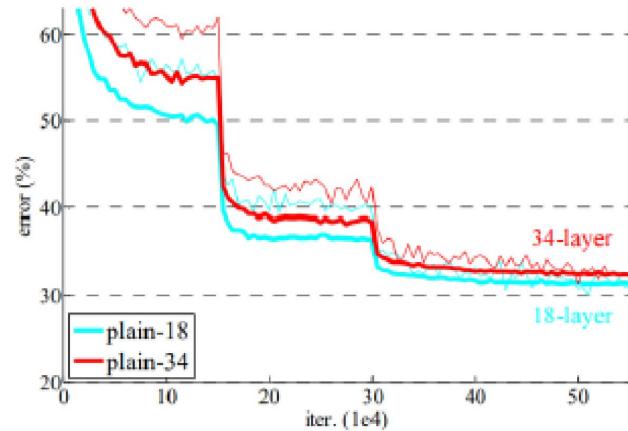
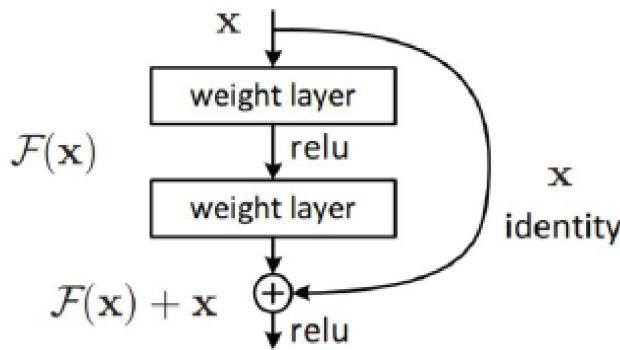
Каскады свёрток



Сеть ResNet (2015)

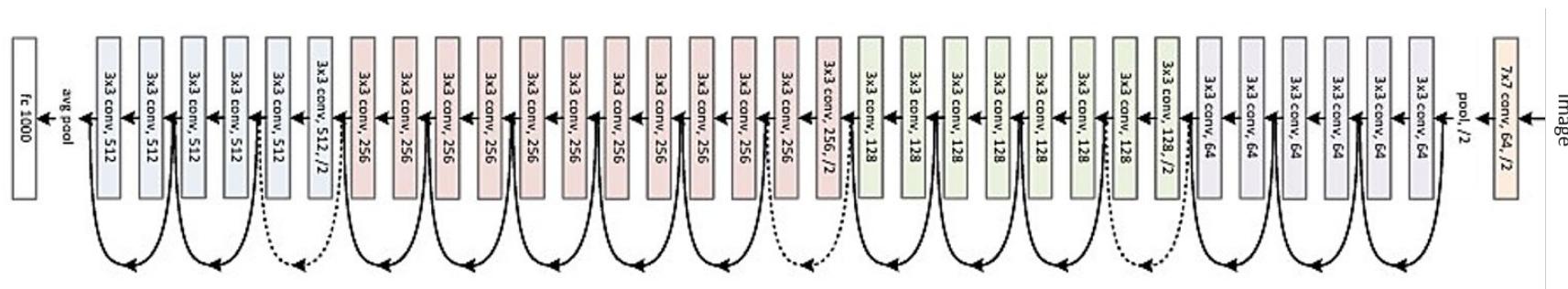
Просто добавление слоев не работает!

Основная идея: skip connections

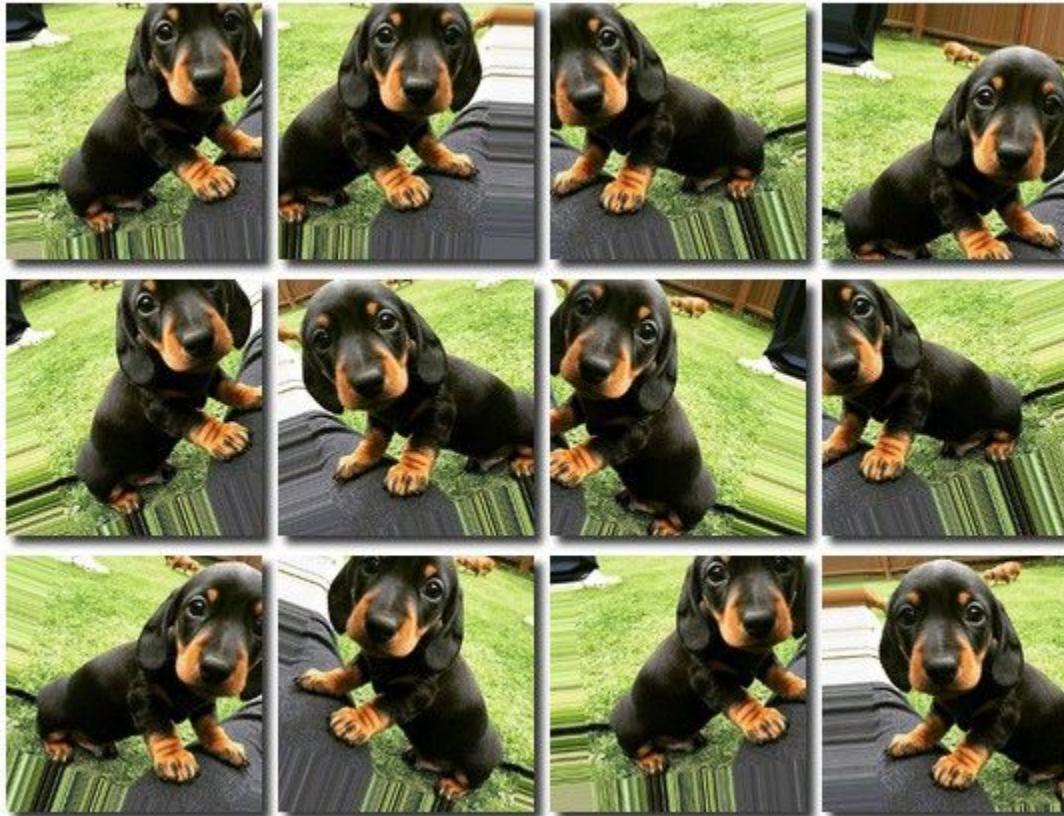


Что делать, если $\mathcal{F}(x)$ отличается по размеру от x ?

Сеть ResNet (2015)



Методы аугментации изображений

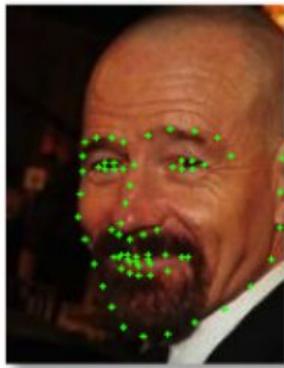
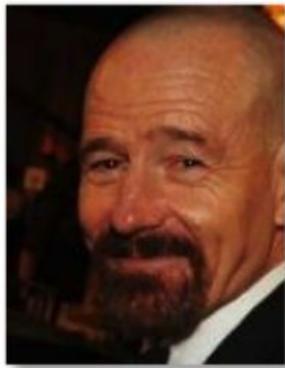


Стандартные методы

- Поворот
- Сдвиг
- Удаление части
- Обрезка
- Размытие/рэзкость
- Шум
- Осветление/затемнение



Аугментация: повороты лиц



Удаление/замена фона



Аугментация документов

Report on the Economic Well-Being of U.S. Households in 2017

Table 13. Confidence that credit card applications would be denied if they turned down a loan or investment.

Response	Percent
Yes	45
No	55

Table 14. Risk of Auto or Credit Card Being Denied.

Response	Percent
Yes	25
No	75

Figure 15. Confidence that credit card applications would be denied if they turned down a loan or investment.

The accompanying notes are a integral part of the financial statements.

STATEMENTS OF CHANGES IN NET ASSETS OF THE ENDOWMENT

In thousands of dollars

Invested	Temporarily designated	Nonvested	For the year ended
\$ 40,270	\$ 211,489	\$ 249,375	\$ 30,874
(\$ 2,047)	(1,812,453)	(2,047,453)	(2,047,453)
Realized and unrealized appreciation, net			
(\$ 7,487)	(\$ 4,959,453)	(\$ 4,959,453)	(\$ 4,959,453)
Endowment gains from operations			
(\$ 6,627)	(\$ 1,471,473)	(\$ 1,471,473)	(\$ 1,471,473)
NET CHANGE DURING THE YEAR			
\$ (10,274)	\$ (1,471,473)	\$ (1,471,473)	\$ (1,471,473)
NET ASSETS OF THE ENDOWMENT, at year end			
\$ 30,874	\$ 211,489	\$ 249,375	\$ 30,874

The accompanying notes are a integral part of the financial statements.

CASH VALUE OF THE ENDOWMENT AS OF JUNE 30, 2018

In millions of dollars

TOTAL FAIR VALUE \$74.4B

Debt

Outstanding debt increased from \$4.7 billion at June 30, 2017 to \$5.6 billion at June 30, 2018. The University's cash flow from new debt issuance over the past year was used to allow for future flexibility in the financing of major initiatives.

This concludes the summary of the key financial highlights for fiscal 2018. We encourage you to read the accompanying notes for more information regarding the University's debt portfolio and results of its financial statements.

Report on the Economic Well-Being of U.S. Households in 2017

Figure 16. Confidence that credit card applications would be denied if they turned down a loan or investment.

Financial Literacy and Comfort Investing

Among these with self-directed investment savings, including 401(k)s, IRAs, and savings outside of formal retirement accounts, conduct in managing their investment portfolios. About 70 percent of these accounts have little or no conduct managing their investments.

Some people withdraw money from their retirement accounts, such as 401(k)s and IRAs, to meet other needs despite the fact that they may incur a substantial penalty. Overall, 5 percent of non-victims have been unable to meet their basic living expenses in the past year, a portion have permanently withdrawn funds from their accounts. Funds withdrawn early are less likely to be used for basic living expenses than those withdrawn later (figure 30). These funds drive down the rate of return on these accounts by an average of 27 percent versus 19 percent (figure 30).

Report on the Economic Well-Being of U.S. Households in 2017

Table 15. Confidence that credit card applications would be denied if they turned down a loan or investment.

Response	Percent
Yes	45
No	55

Table 16. Risk of Auto or Credit Card Being Denied.

Response	Percent
Yes	25
No	75

Figure 17. Confidence that credit card applications would be denied if they turned down a loan or investment.

The accompanying notes are a integral part of the financial statements.

STATEMENTS OF CHANGES IN NET ASSETS OF THE ENDOWMENT

In thousands of dollars

Invested	Temporarily designated	Nonvested	For the year ended
\$ 40,270	\$ 211,489	\$ 249,375	\$ 30,874
(\$ 2,047)	(1,812,453)	(2,047,453)	(2,047,453)
Realized and unrealized appreciation, net			
(\$ 7,487)	(\$ 4,959,453)	(\$ 4,959,453)	(\$ 4,959,453)
Endowment gains from operations			
(\$ 6,627)	(\$ 1,471,473)	(\$ 1,471,473)	(\$ 1,471,473)
NET CHANGE DURING THE YEAR			
\$ (10,274)	\$ (1,471,473)	\$ (1,471,473)	\$ (1,471,473)
NET ASSETS OF THE ENDOWMENT, at year end			
\$ 30,874	\$ 211,489	\$ 249,375	\$ 30,874

The accompanying notes are a integral part of the financial statements.

CASH VALUE OF THE ENDOWMENT AS OF JUNE 30, 2018

In millions of dollars

TOTAL FAIR VALUE \$74.4B

Debt

Outstanding debt increased from \$4.7 billion at June 30, 2017 to \$5.6 billion at June 30, 2018. The University's cash flow from new debt issuance over the past year was used to allow for future flexibility in the financing of major initiatives.

This concludes the summary of the key financial highlights for fiscal 2018. We encourage you to read the accompanying notes for more information regarding the University's debt portfolio and results of its financial statements.

Report on the Economic Well-Being of U.S. Households in 2017

Figure 18. Lack of retirement savings and perception of preparedness.

Financial Literacy and Comfort Investing

Among these with self-directed investment savings, including 401(k)s, IRAs, and savings outside of formal retirement accounts, conduct in managing their investment portfolios. About 70 percent of these accounts have little or no conduct managing their investments.

Some people withdraw money from their retirement accounts, such as 401(k)s and IRAs, to meet other needs despite the fact that they may incur a substantial penalty. Overall, 5 percent of non-victims have been unable to meet their basic living expenses in the past year, a portion have permanently withdrawn funds from their accounts. Funds withdrawn early are less likely to be used for basic living expenses than those withdrawn later (figure 30). These funds drive down the rate of return on these accounts by an average of 27 percent versus 19 percent (figure 30).

Report on the Economic Well-Being of U.S. Households in 2017

Table 17. Confidence that credit card applications would be denied if they turned down a loan or investment.

Response	Percent
Yes	45
No	55

Table 18. Risk of Auto or Credit Card Being Denied.

Response	Percent
Yes	25
No	75

Figure 19. Confidence that credit card applications would be denied if they turned down a loan or investment.

The accompanying notes are a integral part of the financial statements.

STATEMENTS OF CHANGES IN NET ASSETS OF THE ENDOWMENT

In thousands of dollars

Invested	Temporarily designated	Nonvested	For the year ended
\$ 40,270	\$ 211,489	\$ 249,375	\$ 30,874
(\$ 2,047)	(1,812,453)	(2,047,453)	(2,047,453)
Realized and unrealized appreciation, net			
(\$ 7,487)	(\$ 4,959,453)	(\$ 4,959,453)	(\$ 4,959,453)
Endowment gains from operations			
(\$ 6,627)	(\$ 1,471,473)	(\$ 1,471,473)	(\$ 1,471,473)
NET CHANGE DURING THE YEAR			
\$ (10,274)	\$ (1,471,473)	\$ (1,471,473)	\$ (1,471,473)
NET ASSETS OF THE ENDOWMENT, at year end			
\$ 30,874	\$ 211,489	\$ 249,375	\$ 30,874

The accompanying notes are a integral part of the financial statements.

CASH VALUE OF THE ENDOWMENT AS OF JUNE 30, 2018

In millions of dollars

TOTAL FAIR VALUE \$74.4B

Debt

Outstanding debt increased from \$4.7 billion at June 30, 2017 to \$5.6 billion at June 30, 2018. The University's cash flow from new debt issuance over the past year was used to allow for future flexibility in the financing of major initiatives.

This concludes the summary of the key financial highlights for fiscal 2018. We encourage you to read the accompanying notes for more information regarding the University's debt portfolio and results of its financial statements.

Report on the Economic Well-Being of U.S. Households in 2017

Figure 20. Lack of retirement savings and self-assessed preparedness.

Financial Literacy and Comfort Investing

Among these with self-directed investment savings, including 401(k)s, IRAs, and savings outside of formal retirement accounts, conduct in managing their investment portfolios. About 70 percent of these accounts have little or no conduct managing their investments.

Some people withdraw money from their retirement accounts, such as 401(k)s and IRAs, to meet other needs despite the fact that they may incur a substantial penalty. Overall, 5 percent of non-victims have been unable to meet their basic living expenses in the past year, a portion have permanently withdrawn funds from their accounts. Funds withdrawn early are less likely to be used for basic living expenses than those withdrawn later (figure 30). These funds drive down the rate of return on these accounts by an average of 27 percent versus 19 percent (figure 30).

Report on the Economic Well-Being of U.S. Households in 2017

Table 19. Confidence that credit card applications would be denied if they turned down a loan or investment.

Response	Percent
Yes	45
No	55

Table 20. Risk of Auto or Credit Card Being Denied.

Response	Percent
Yes	25
No	75

Figure 21. Confidence that credit card applications would be denied if they turned down a loan or investment.

The accompanying notes are a integral part of the financial statements.

STATEMENTS OF CHANGES IN NET ASSETS OF THE ENDOWMENT

In thousands of dollars

Invested	Temporarily designated	Nonvested	For the year ended
\$ 40,270	\$ 211,489	\$ 249,375	\$ 30,874
(\$ 2,047)	(1,812,453)	(2,047,453)	(2,047,453)
Realized and unrealized appreciation, net			
(\$ 7,487)	(\$ 4,959,453)	(\$ 4,959,453)	(\$ 4,959,453)
Endowment gains from operations			
(\$ 6,627)	(\$ 1,471,473)	(\$ 1,471,473)	(\$ 1,471,473)
NET CHANGE DURING THE YEAR			
\$ (10,274)	\$ (1,471,473)	\$ (1,471,473)	\$ (1,471,473)
NET ASSETS OF THE ENDOWMENT, at year end			
\$ 30,874	\$ 211,489	\$ 249,375	\$ 30,874

The accompanying notes are a integral part of the financial statements.

CASH VALUE OF THE ENDOWMENT AS OF JUNE 30, 2018

In millions of dollars

TOTAL FAIR VALUE \$74.4B

Debt

Outstanding debt increased from \$4.7 billion at June 30, 2017 to \$5.6 billion at June 30, 2018. The University's cash flow from new debt issuance over the past year was used to allow for future flexibility in the financing of major initiatives.

This concludes the summary of the key financial highlights for fiscal 2018. We encourage you to read the accompanying notes for more information regarding the University's debt portfolio and results of its financial statements.

Report on the Economic Well-Being of U.S. Households in 2017

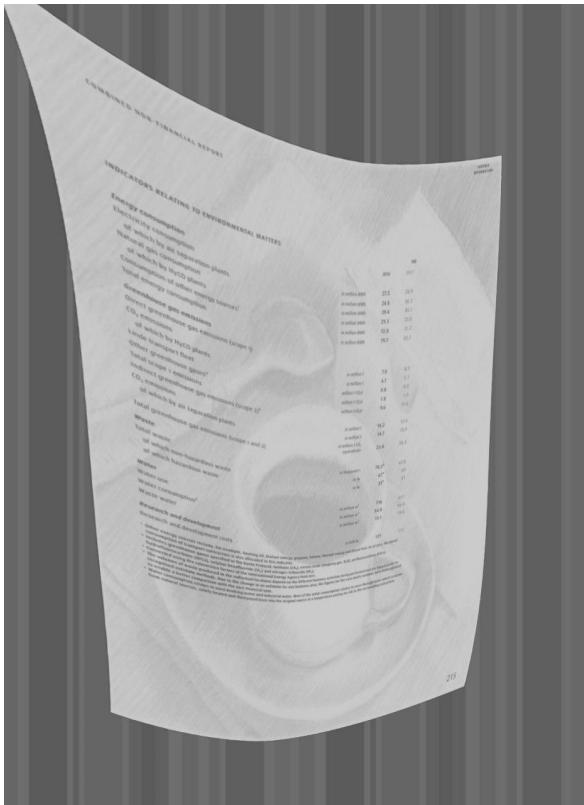
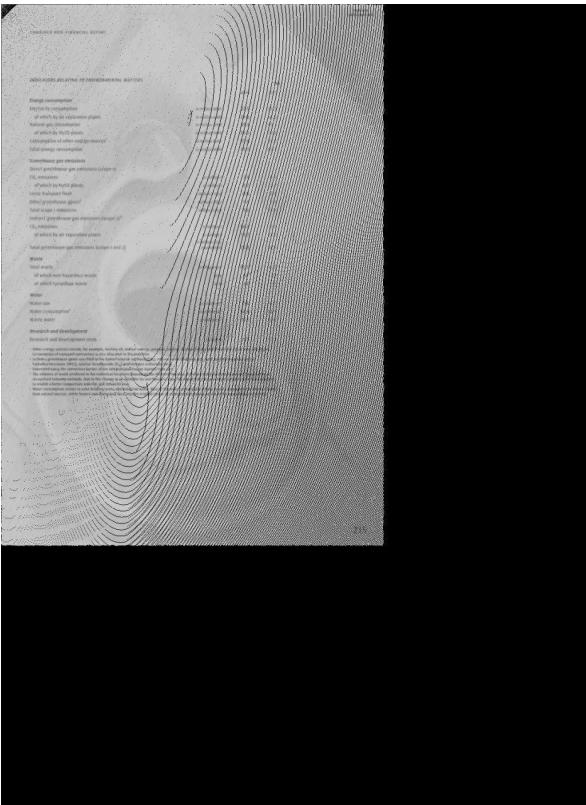
Figure 22. Lack of retirement savings and self-assessed preparedness.

Financial Literacy and Comfort Investing

Among these with self-directed investment savings, including 401(k)s, IRAs, and savings outside of formal retirement accounts, conduct in managing their investment portfolios. About 70 percent of these accounts have little or no conduct managing their investments.

Some people withdraw money from their retirement accounts, such as 401(k)s and IRAs, to meet other needs despite the fact that they may incur a substantial penalty. Overall, 5 percent of non-victims have been unable to meet their basic living expenses in the past year, a portion have permanently withdrawn funds from their accounts. Funds withdrawn early are less likely to be used for basic living expenses than those withdrawn later (figure 30). These funds drive down the rate of return on these accounts by an average of 27 percent versus 19 percent (figure 30).

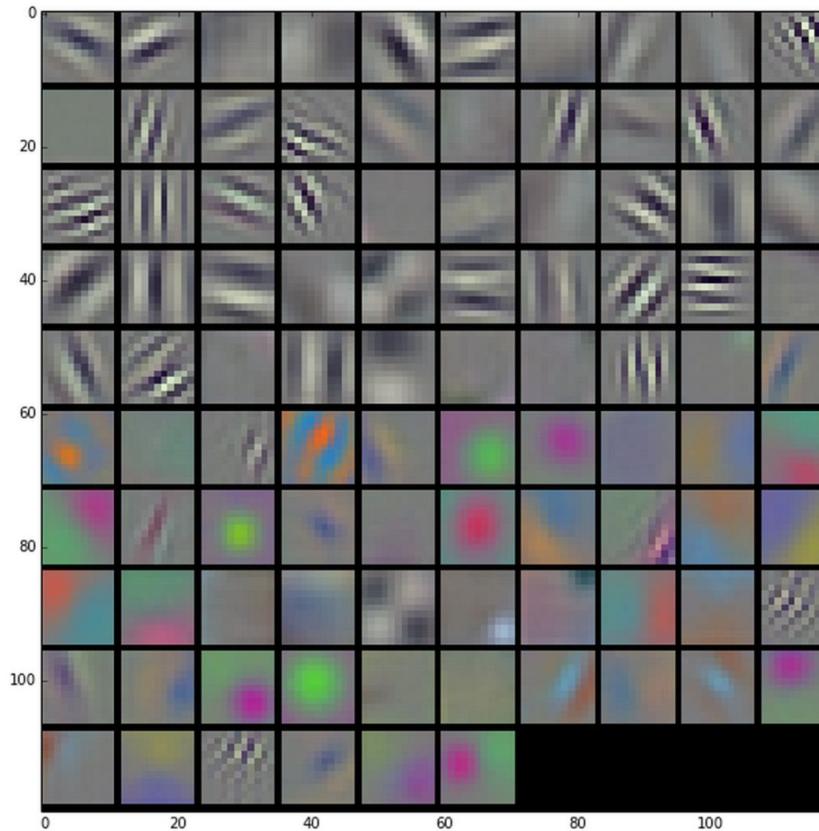
Аугментация документов



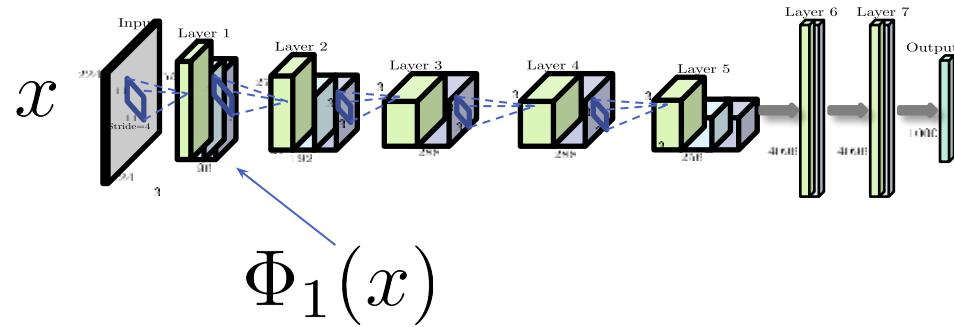
При обучении модели...



Визуализация фильтров первого слоя



Чувствительность фильтров



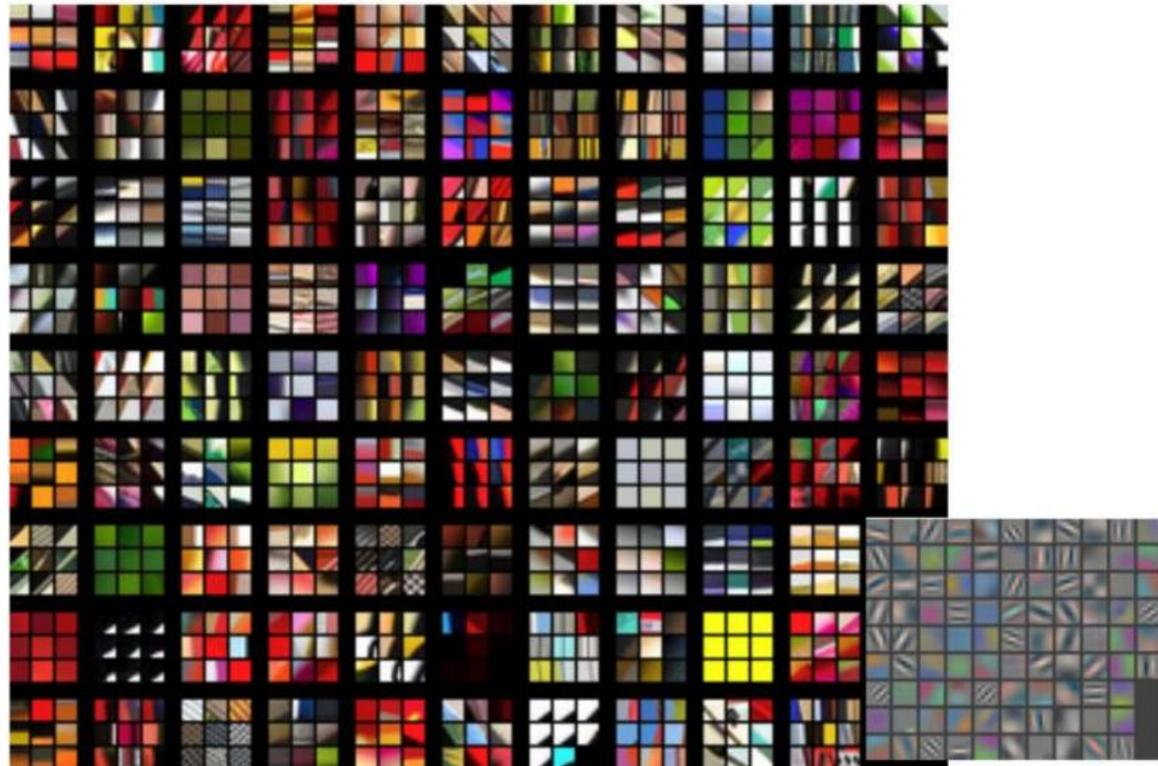
Для свёртки t выделяем картинки, на которые она реагирует сильнее всего:

$$\arg \max_{x \in S} \max_{p,q} \Phi_1(x)^t [p, q]$$

статья
[Zeiler Fergus 14]



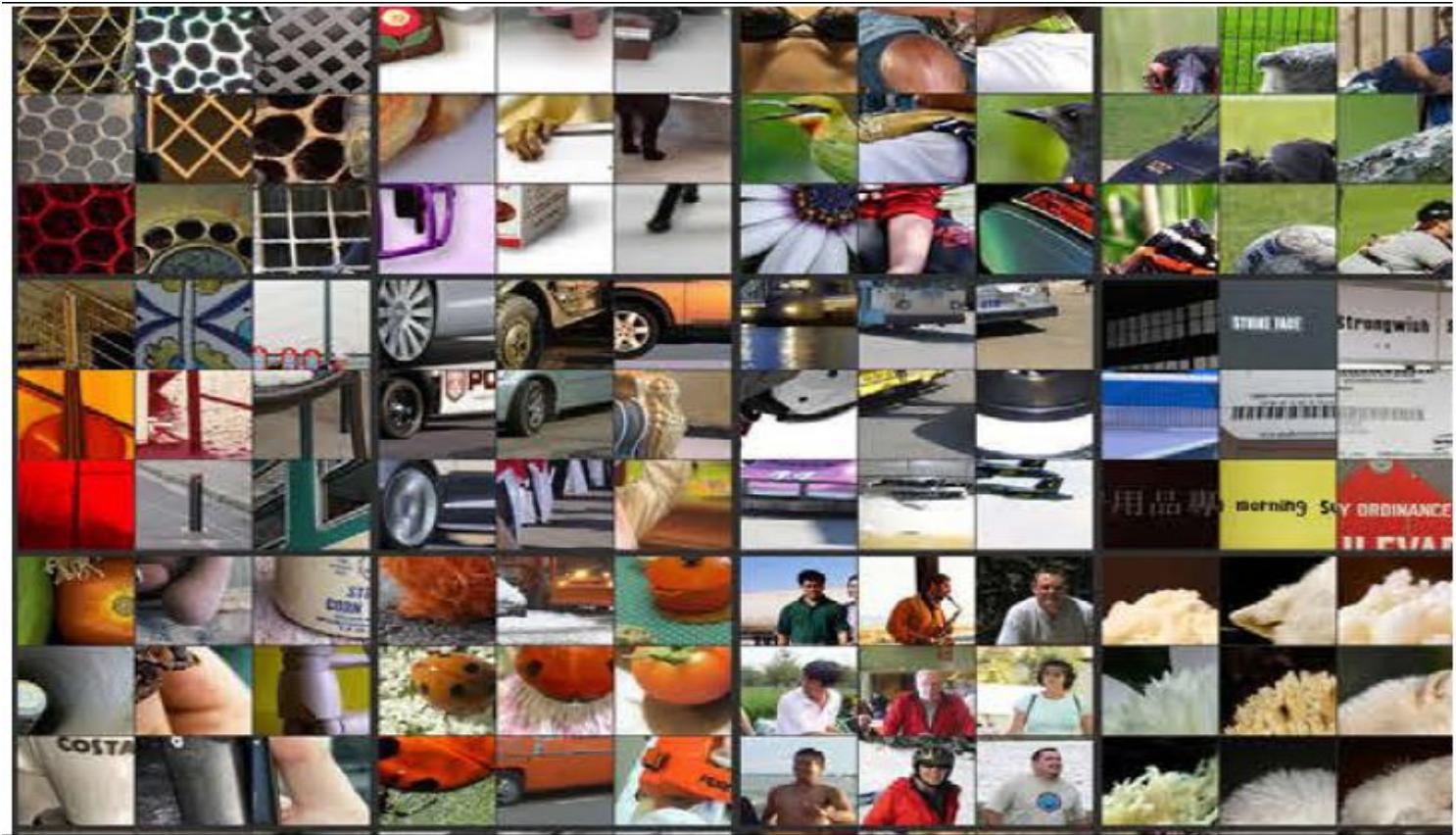
Layer 1





Layer 2





Layer 3



Layer 4



Layer 5

Генерация картинок, чувствительных к определённым нейронам

Фиксируем веса сети, находим картинку, которая будет лучше всего соответствовать классу y_0 :

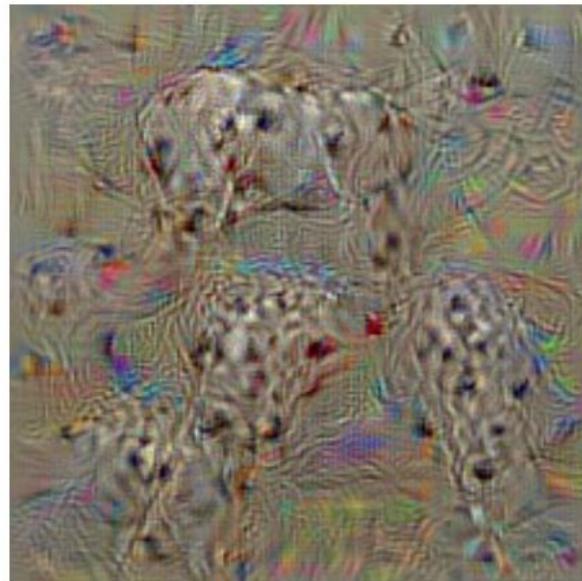
$$\hat{x} = \arg \max_x (\Phi_{\text{Last}}(x)[y_0] - \lambda R(x))$$



dumbbell



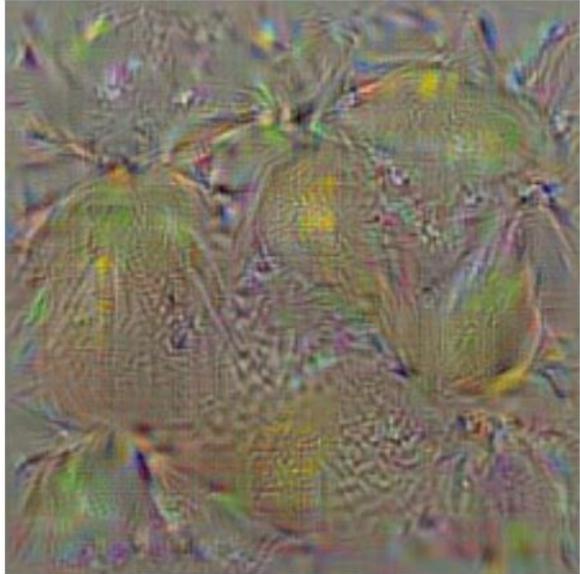
cup



dalmatian



bell pepper



lemon



husky



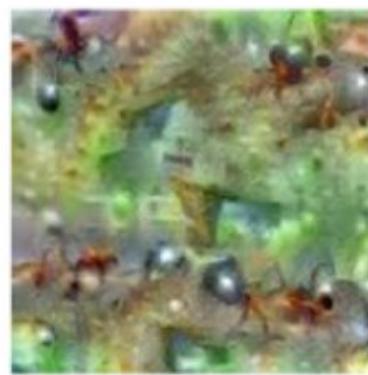
<https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>



Hartebeest



Measuring Cup



Ant



Starfish



Anemone Fish



Banana



Parachute

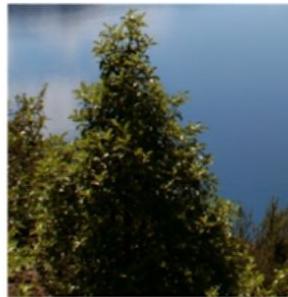


Screw

Как были получены эти картинки?



Horizon



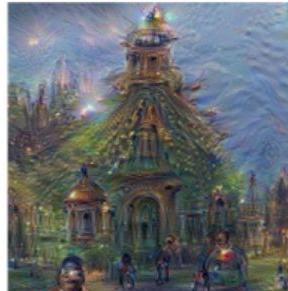
Trees



Leaves



Towers & Pagodas



Buildings



Birds & Insects

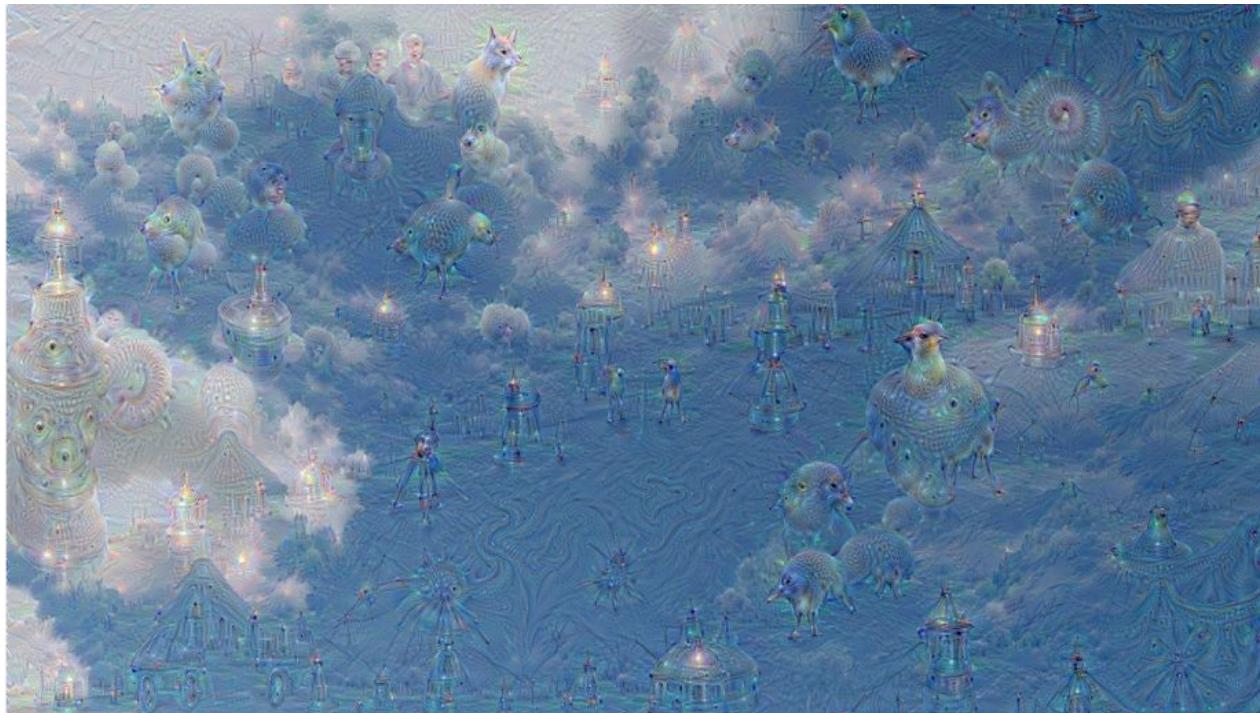
DeepDream

$$\max_x \|\Phi_k(x)\| - R(x)$$



DeepDream

$$\max_x \|\Phi_k(x)\| - R(x)$$

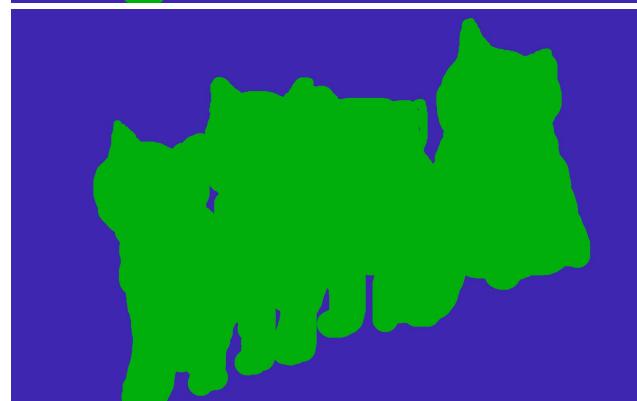
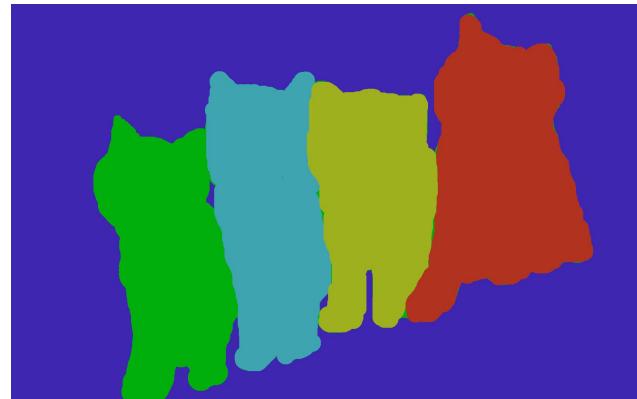


DeepDream

$$\max_x \|\Phi_k(x)\| - R(x)$$

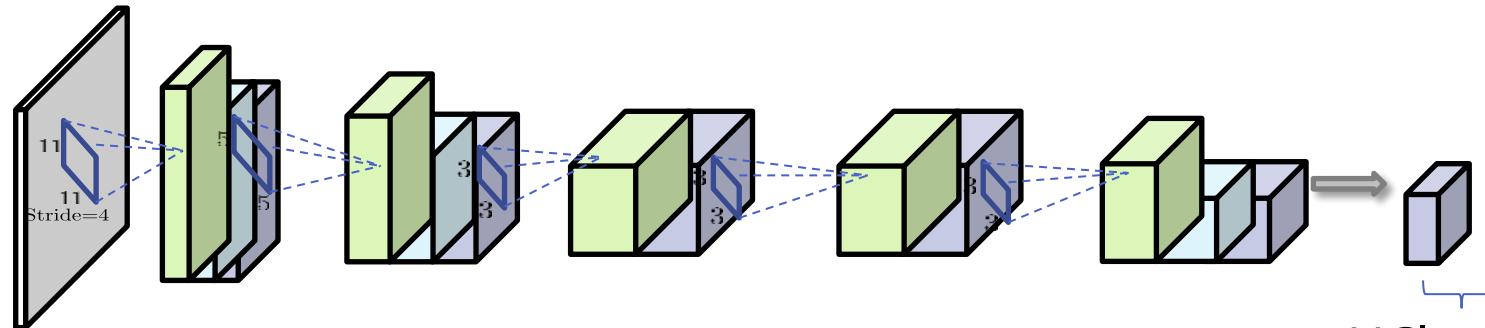
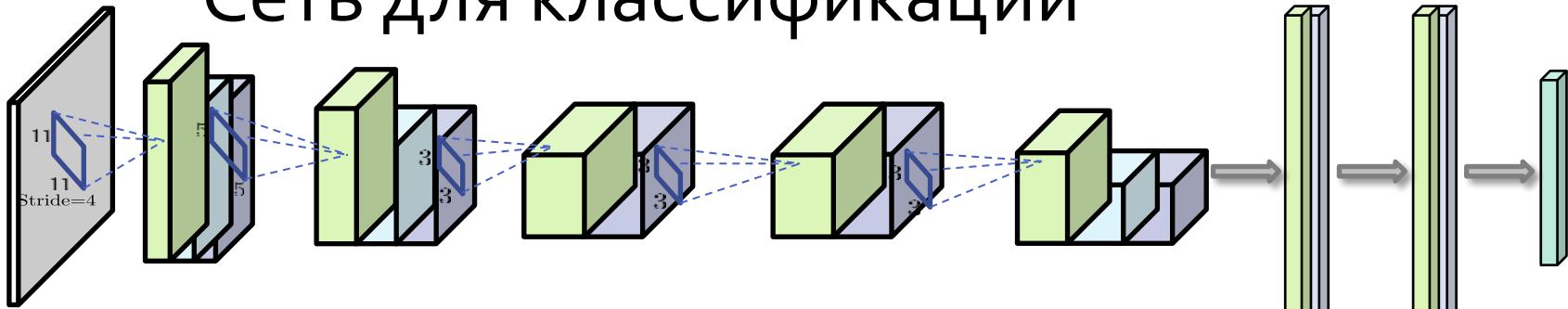


Задача сегментации изображений (instance segmentation)

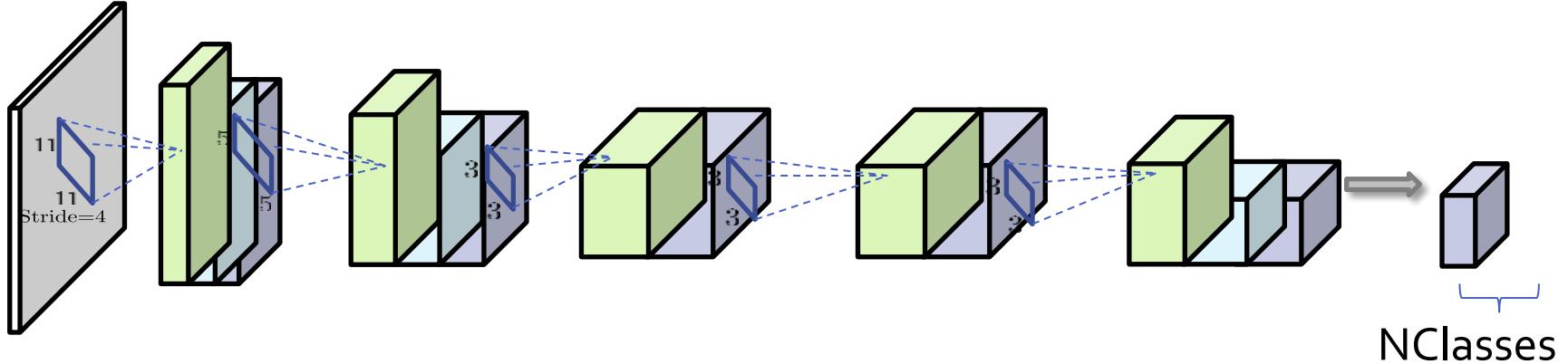


*Images from Chaosmail
Blog*

Сеть для классификации

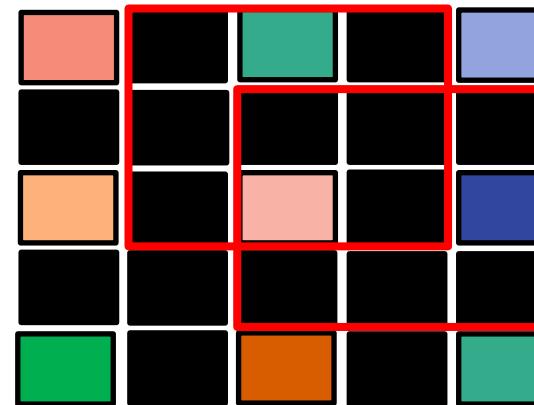
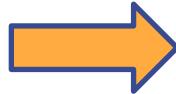
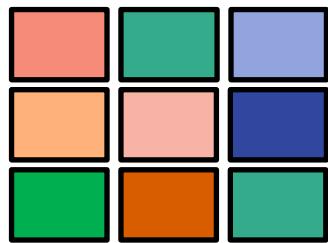


Сеть для сегментации

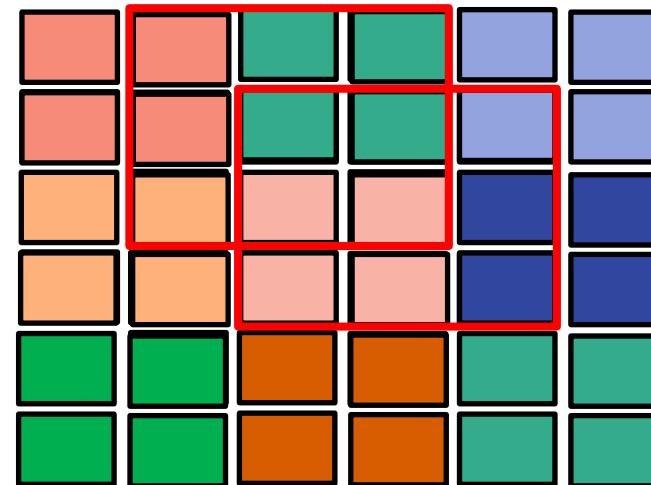
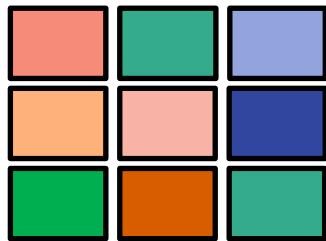


- Проблема: ответ должен быть близок по размеру ко входу
- Проблема: чем глубже тем лучше понимаем что изображено, но забываем контекст
- Receptive field на последних слоях должен быть достаточно большим

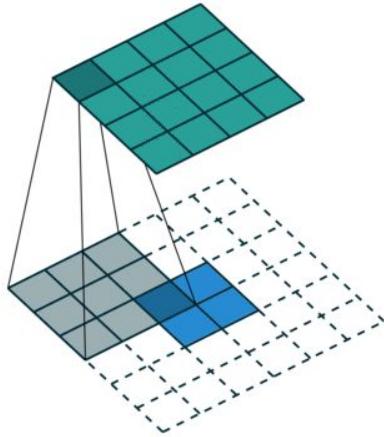
Методы “расширения” изображения



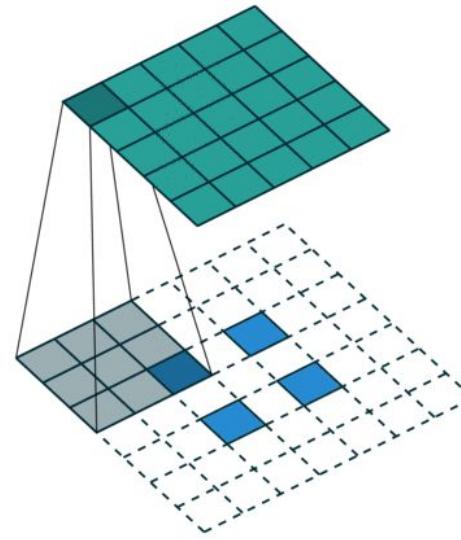
Методы “расширения” изображения



Транспонированная свёртка

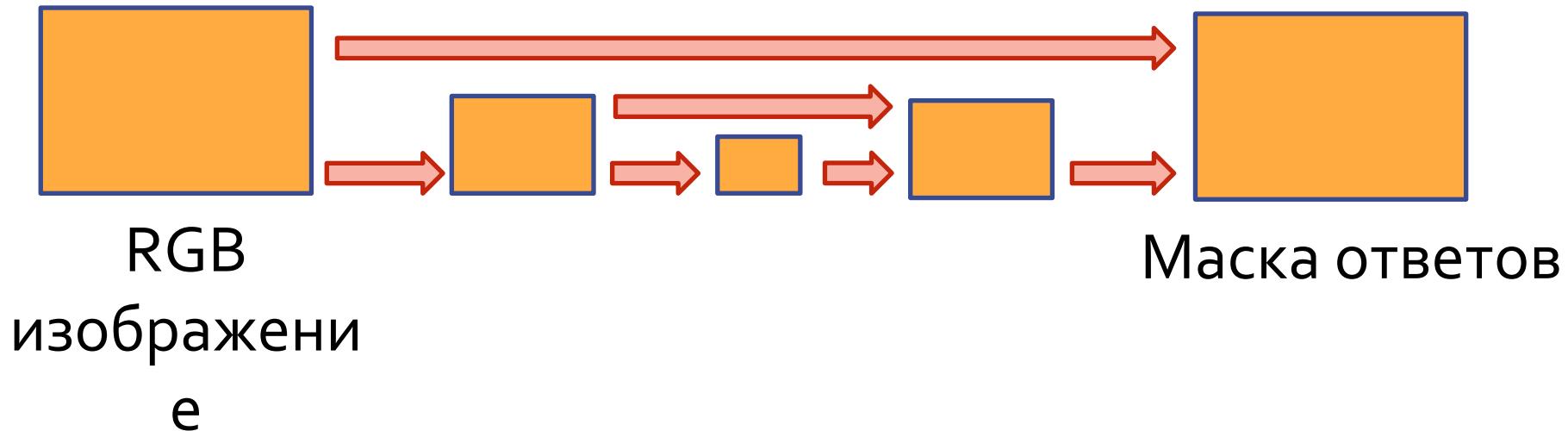


Транспонированная
конволюция
Padding отсутствует,
без страйдов

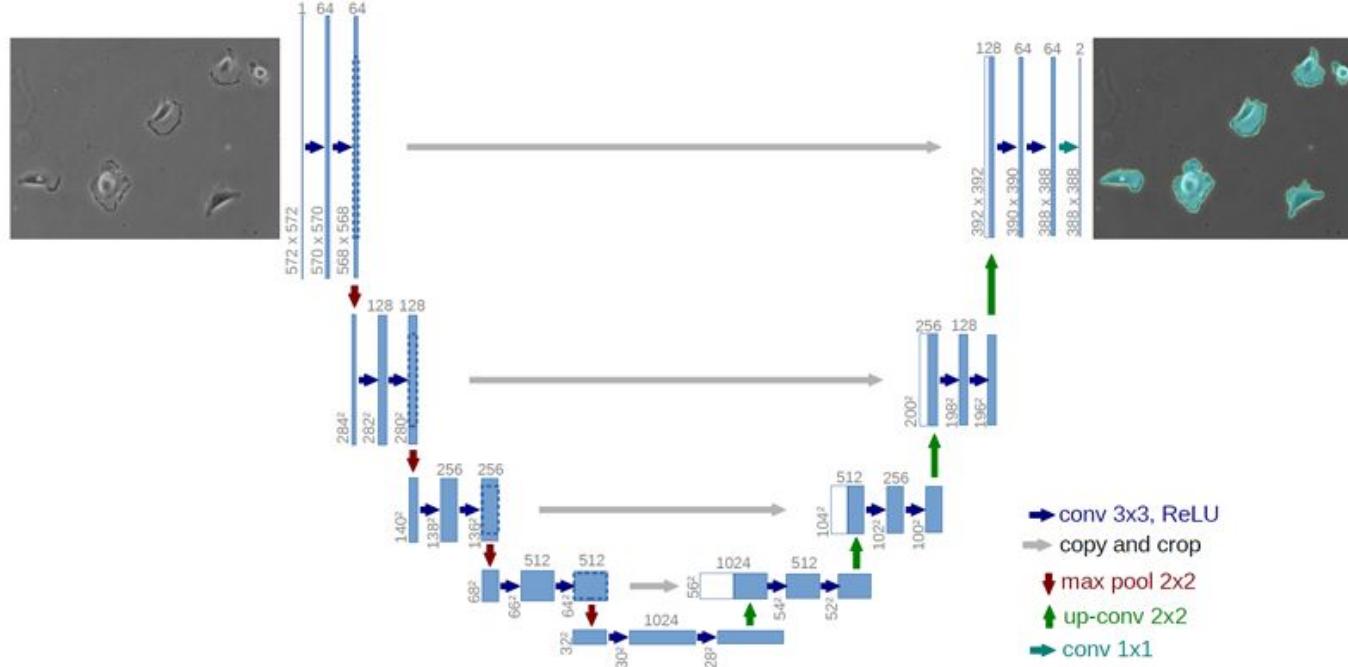


Транспонированная конволюция
Padding отсутствует,
Страйд = 2

Стандартная архитектура для сегментации

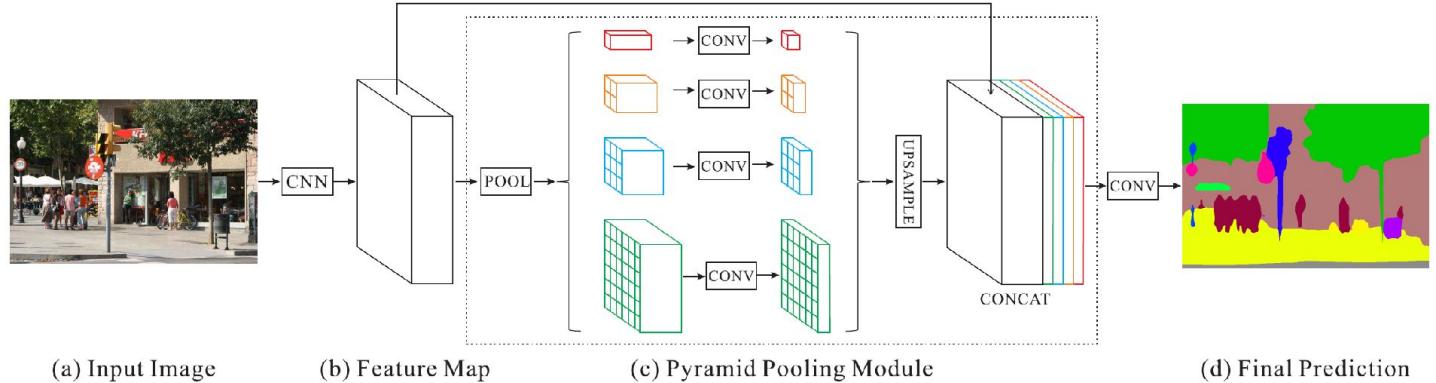


U-net



[Ronnerberger et al. MICCAI15]

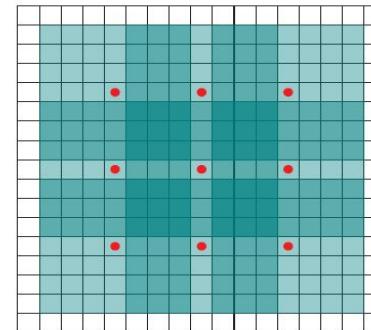
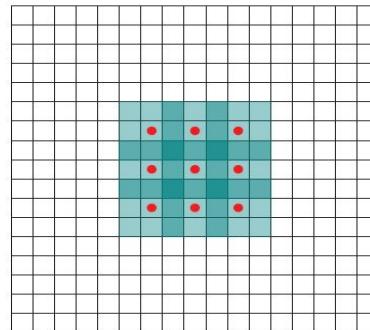
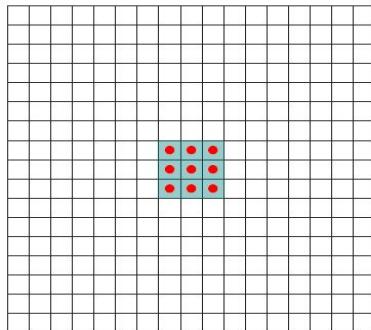
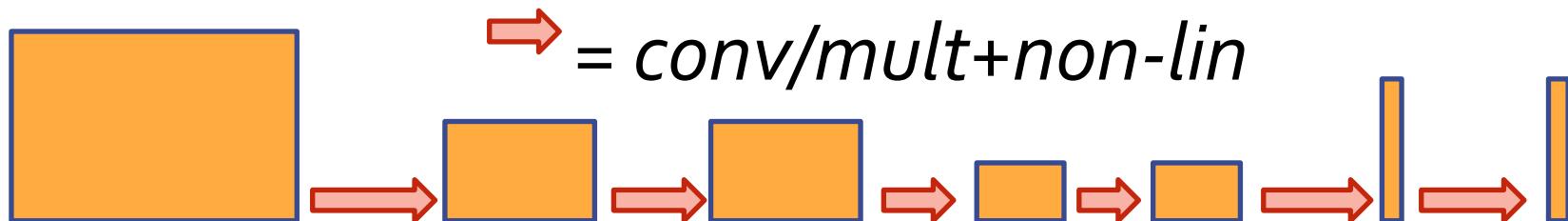
PSP-net



Method	Mean IoU(%)	Pixel Acc.(%)
ResNet50-Baseline	37.23	78.01
ResNet50+B1+MAX	39.94	79.46
ResNet50+B1+AVE	40.07	79.52
ResNet50+B1236+MAX	40.18	79.45
ResNet50+B1236+AVE	41.07	79.97
ResNet50+B1236+MAX+DR	40.87	79.61
ResNet50+B1236+AVE+DR	41.68	80.04

[Zhao et al. CVPR17]

Использование dilation в свёртках



[Yu & Koltun ICLR16]

Link-net

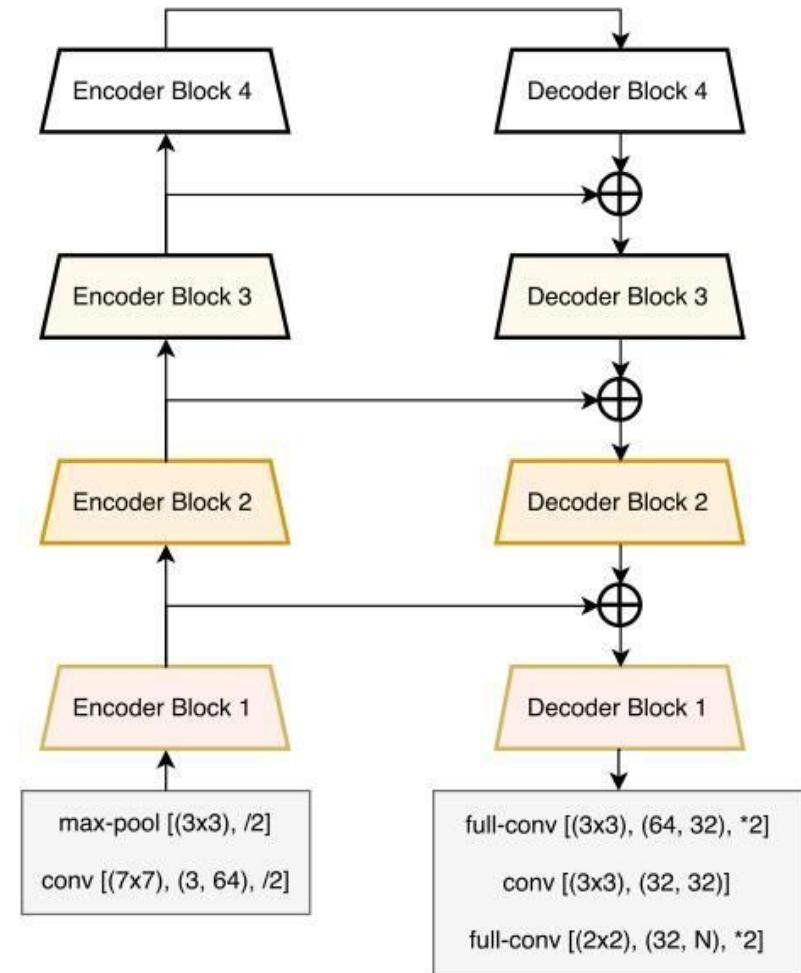
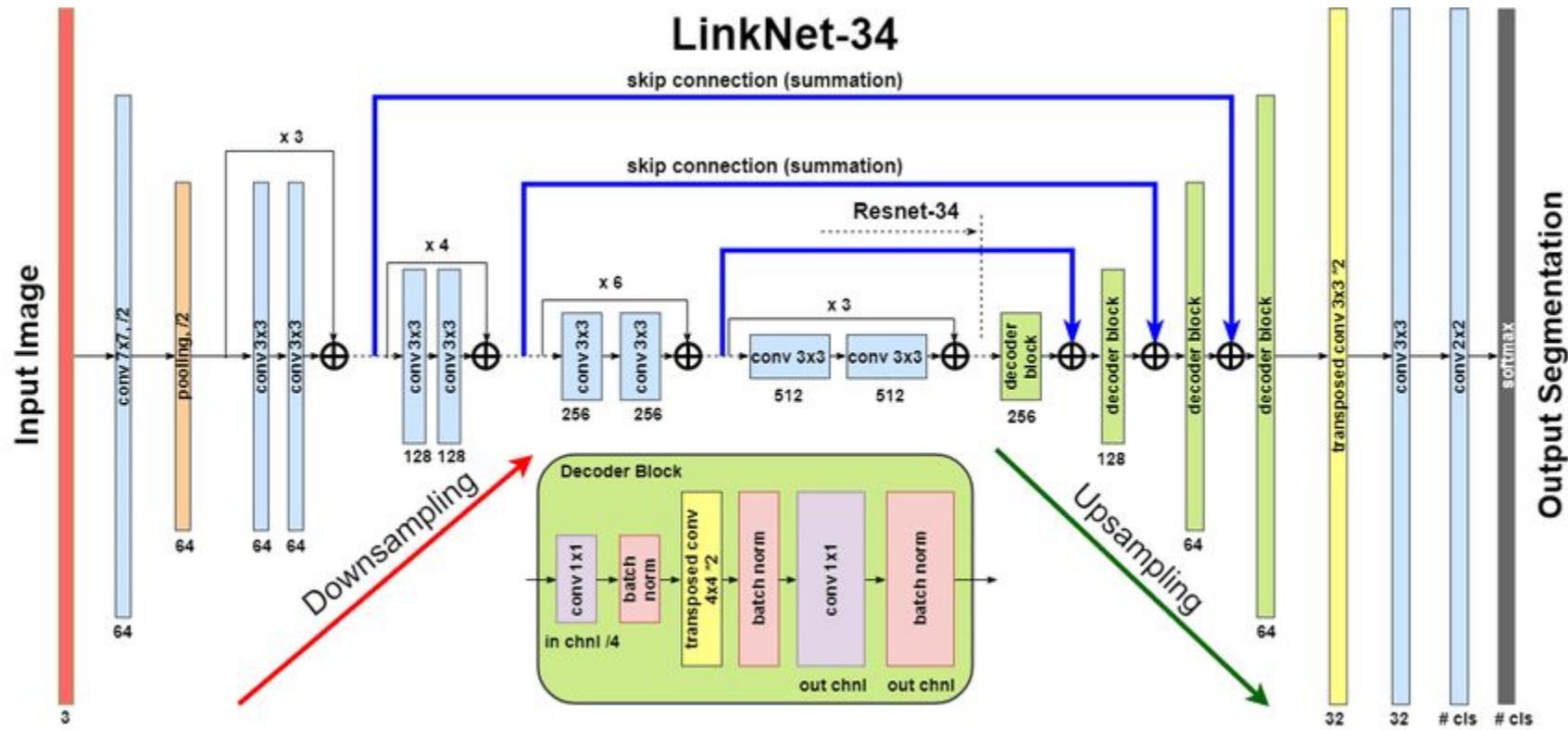


Fig. 1: LinkNet Architecture
https://blog.csdn.net/cv_family_z

Link-net



А ещё нужно рассказать про автокодировщики...