



СПбГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

Александр Калиниченко

ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ

Модуль 2. Методы искусственного интеллекта

Лекция 12. Сверточные нейронные сети

БАЗА ДАННЫХ ‘MNIST’

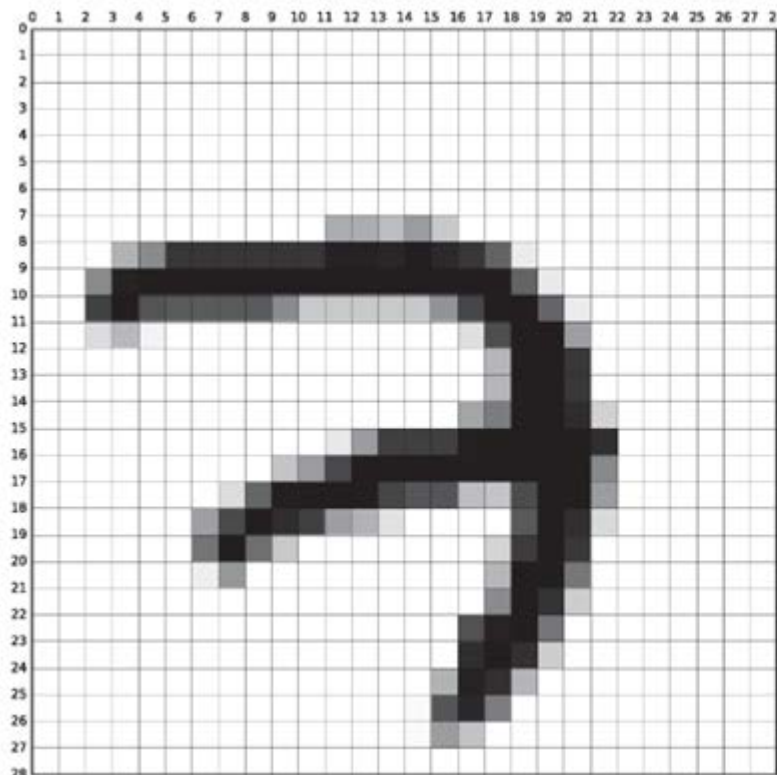
База данных **MNIST** —набор образцов рукописного написания цифр.

База данных является стандартом, предложенным Национальным институтом стандартов и технологий США («Modified National Institute of Standards and Technology») с целью сопоставления методов распознавания изображений с помощью методов машинного обучения

База данных MNIST содержит:

- **60000** изображений для обучения
- **10000** изображений для тестирования

MNIST. ПРЕДСТАВЛЕНИЕ СИМВОЛОВ



Образцы символов были нормализованы, прошли сглаживание и приведены к серому полутоновому изображению размером 28x28 пикселей

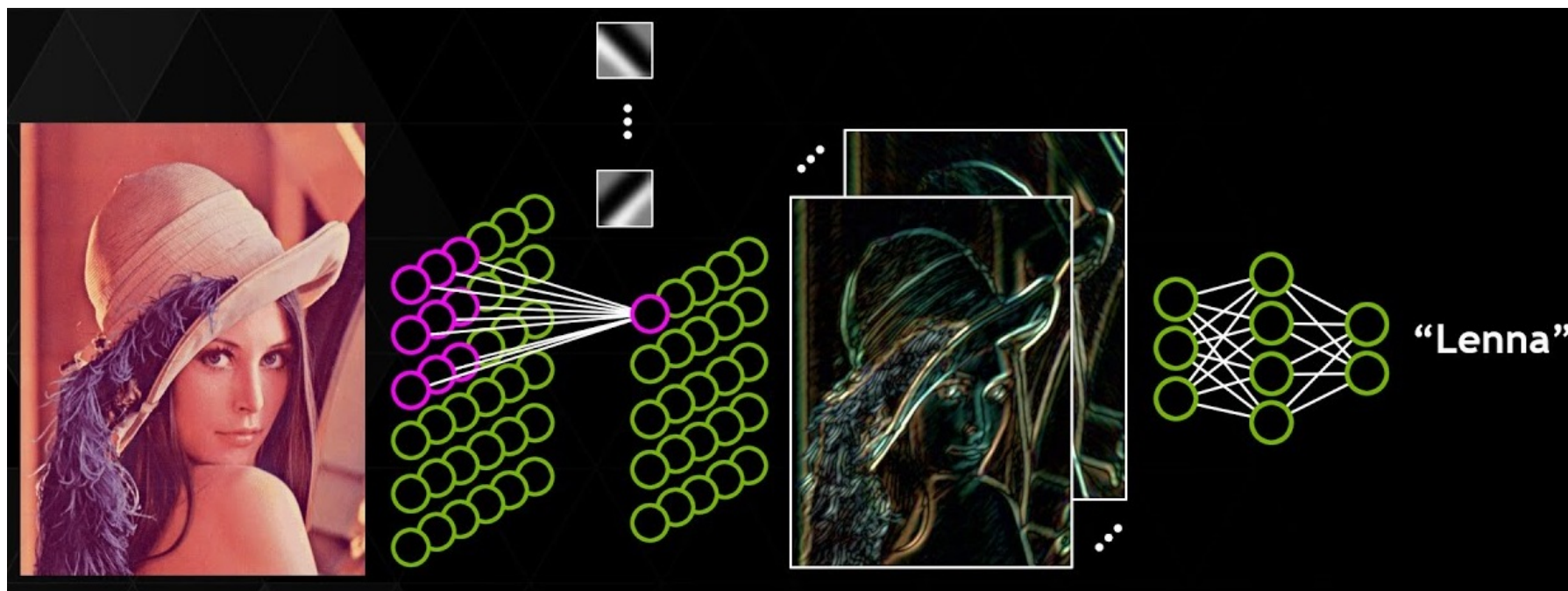
MNIST. РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Тип	Структура	Ошибка (%)
Линейный классификатор	Одноуровневый перцептрон	12
Линейный классификатор	Попарный линейный классификатор	7.6
Нелинейный классификатор	40 PCA + квадратичный классификатор	3.3
Gradient boosting	Обработка остатков на базе признаков Хаара	0.87
Нейронная сеть	2-уровневая сеть	0.7
Метод опорных векторов	Виртуальная система опорных векторов	0.56
Метод k ближайших соседей	K-NN с нелинейной деформацией	0.52
Глубокая нейронная сеть	6-уровневая сеть	0.35
Свёрточная нейронная сеть	6-уровневая сеть	0.27
Свёрточная нейронная сеть	Ансамбль из 35 CNN-сетей	0.23
Свёрточная нейронная сеть	Ансамбль из 5 CNN-сетей, 6-уровней	0.21
Случайное мультимодельное глубокое обучение (RMDL)	30 моделей случайного глубокого обучения (10 CNN, 10 RNN и 10 DNN)	0.18

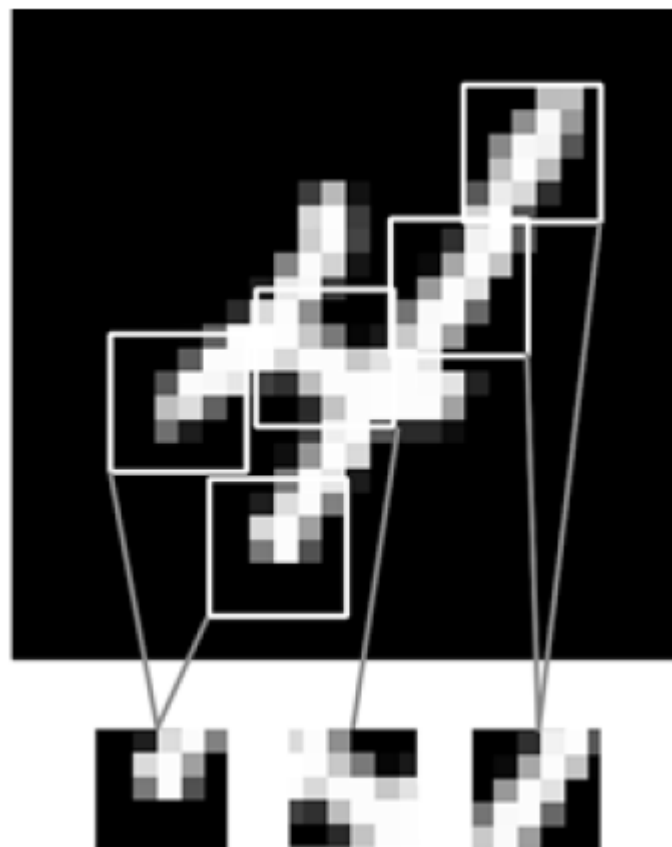
Средняя ошибка
визуального
распознавания
символов людьми
составляет около
2,5 %

СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

Свёрточная нейронная сеть — специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году и нацеленная на эффективное распознавание образов

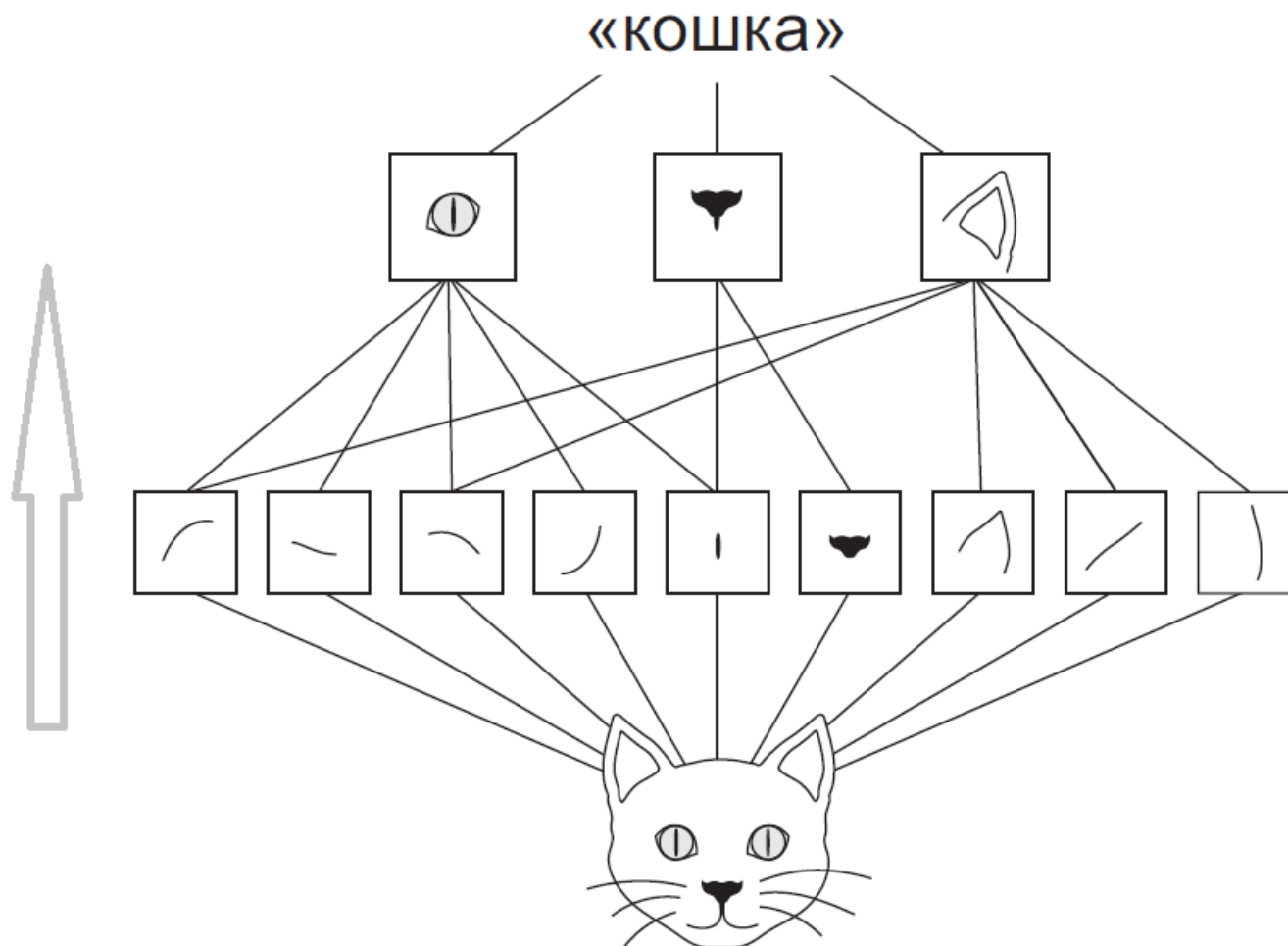


ЛОКАЛЬНЫЕ ШАБЛОНЫ



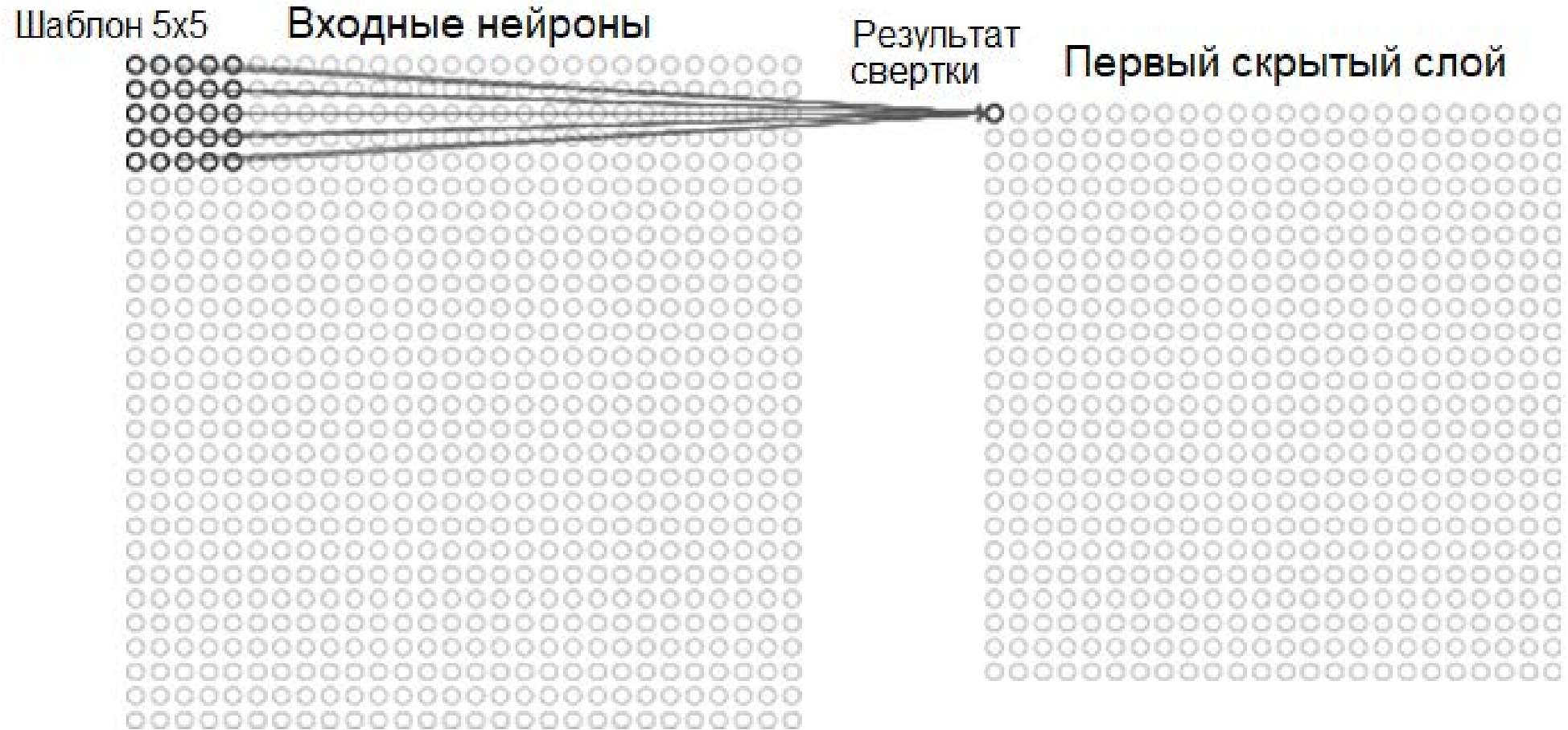
Изображение можно разбить
на локальные шаблоны,
соответствующие некоторым
типовым элементам

РАЗЛОЖЕНИЕ ИЗОБРАЖЕНИЯ



- Изображение может быть разложено на элементарные модули, из которых формируются характерные локальные объекты (глаза, уши и т.д.)
- Набор полученных локальных объектов служит основой для принятия решения о принадлежности изображения к определенному классу

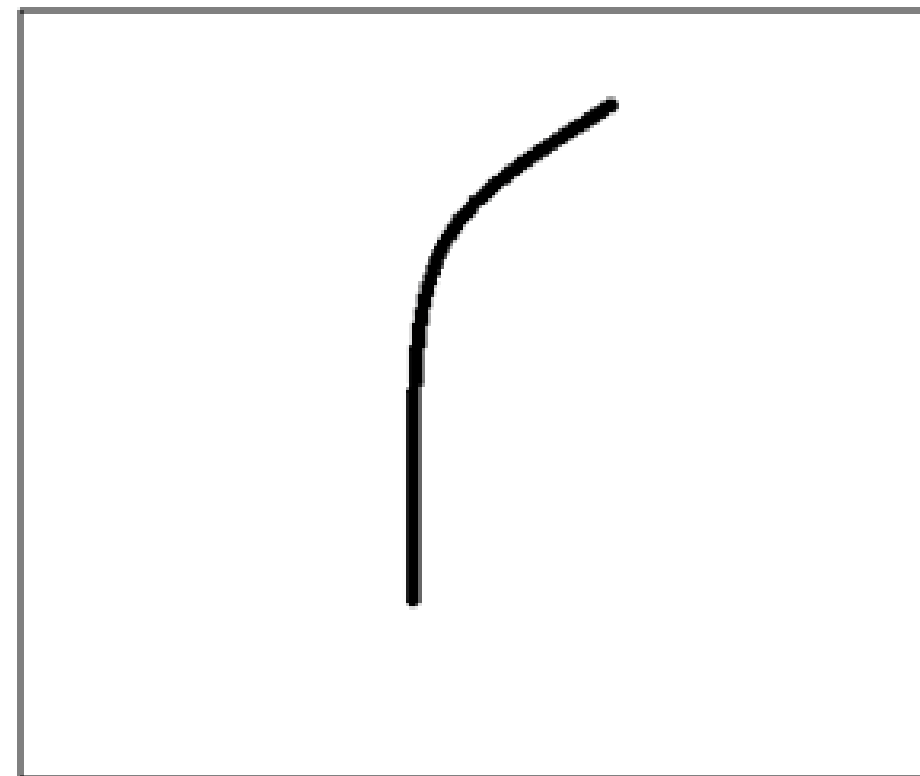
ОПЕРАЦИЯ ДВУМЕРНОЙ СВЕРТКИ В ИНС



ШАБЛОН ДЛЯ ВЫДЕЛЕНИЯ ДУГИ

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

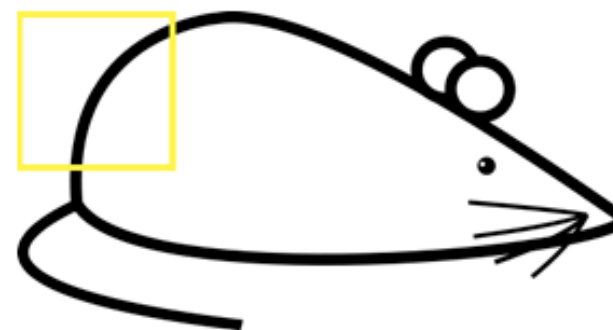


Visualization of a curve detector filter

СОВПАДЕНИЕ ФРАГМЕНТА С ШАБЛОНОМ



Original image



Visualization of the filter on the image



Visualization of the
receptive field

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

Pixel representation of the receptive
field

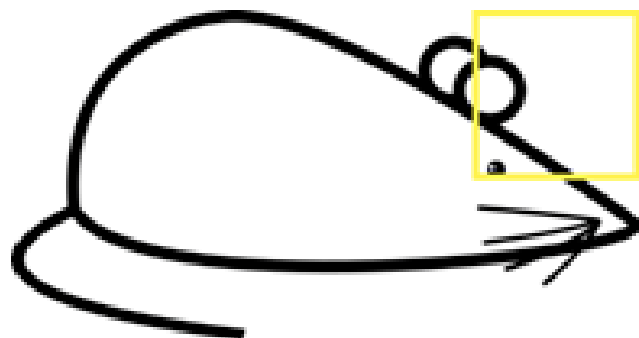
*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

Multiplication and Summation = $(50*30)+(50*30)+(50*30)+(20*30)+(50*30) = 6600$ (A large number!)

НЕСООТВЕТСТВИЕ ФРАГМЕНТА И ШАБЛОНА



Visualization of the filter on the image

0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

Multiplication and Summation = 0

ФИЛЬТР И КАРТА ОТВЕТОВ

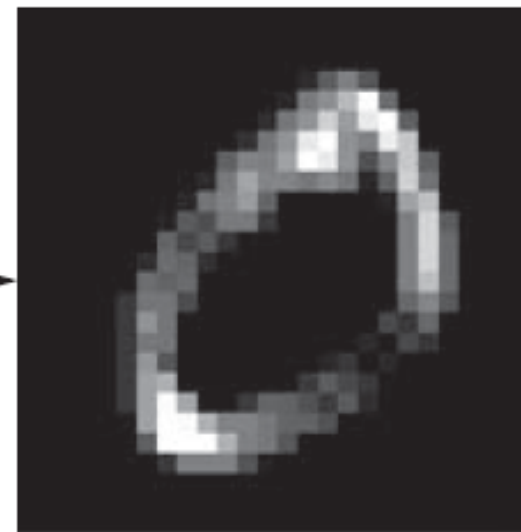
Оригинальный вход



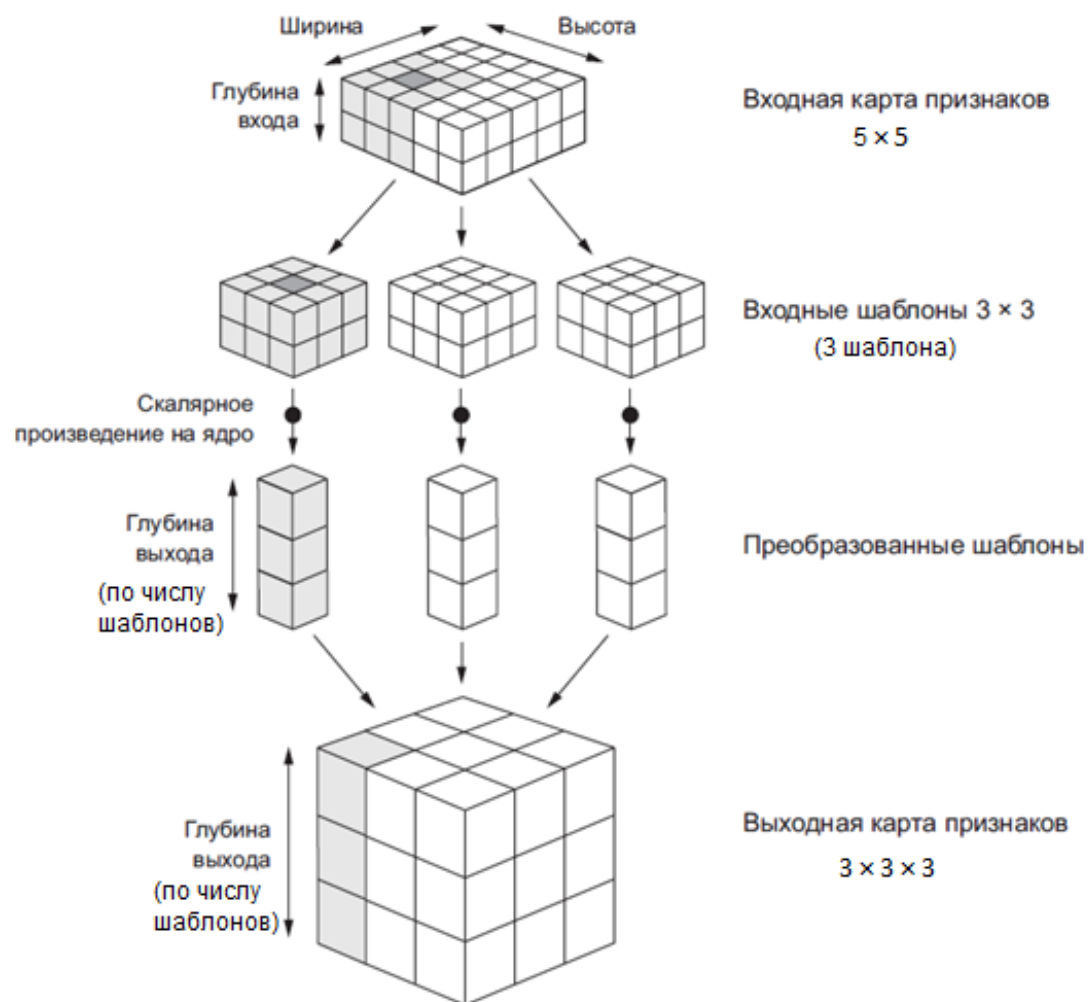
Единственный фильтр



Карта ответов,
выражает в количественной
форме присутствие шаблона
фильтра в разных участках

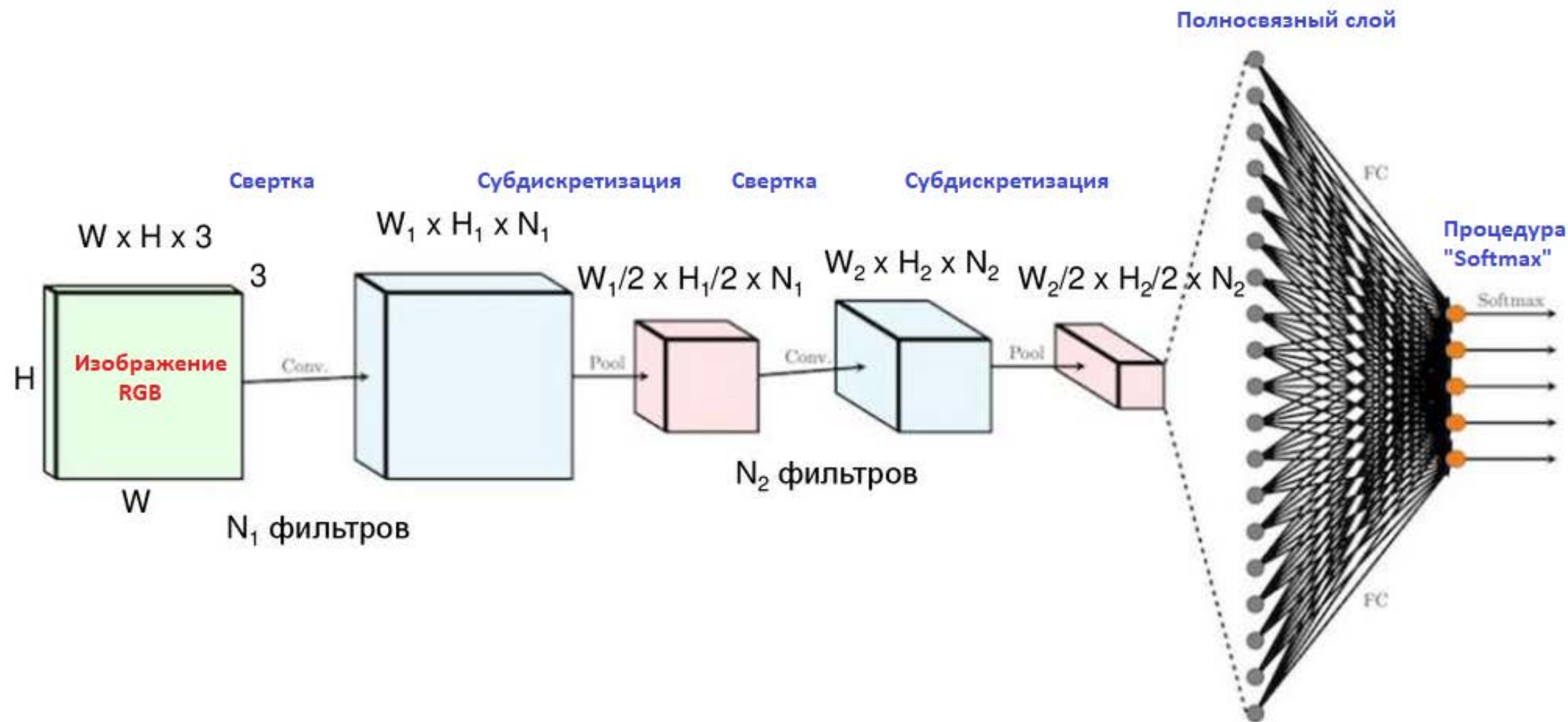


ПРИНЦИП ДЕЙСТВИЯ СВЕРТКИ

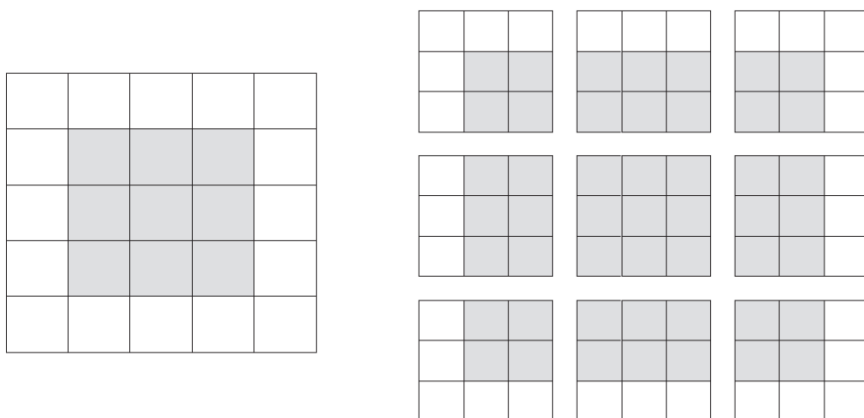


- Размер шаблонов, извлекаемых из входных данных,—обычно 3×3 или 5×5
- Глубина выходной карты признаков —количество фильтров, вычисляемых сверткой

ТРЕХКАНАЛЬНАЯ СВЕРТОЧНАЯ ИНС

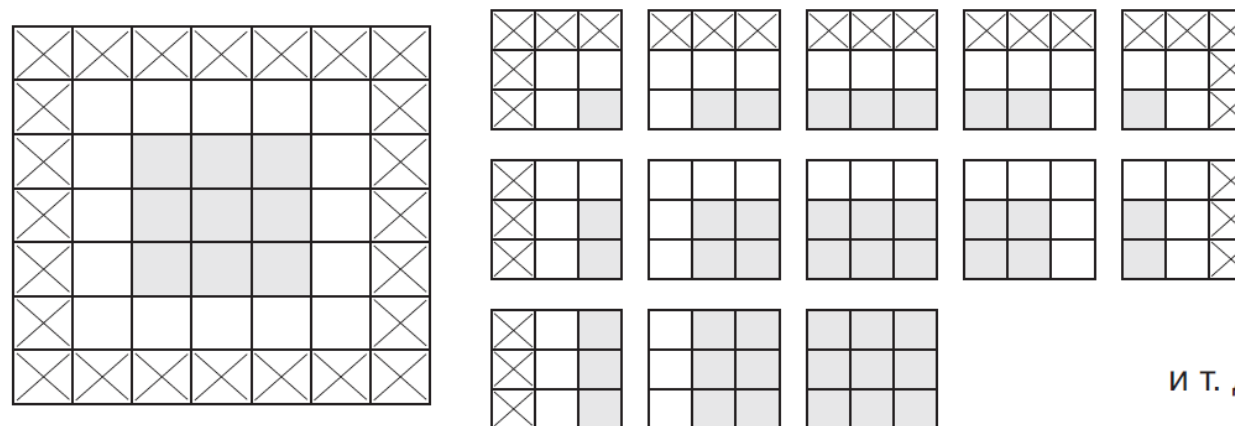


ЭФФЕКТ ГРАНИЦ И ДОПОЛНЕНИЕ



При использовании шаблона 3х3
размер выходной карты сокращается по
сравнению с входным изображением на
2 в каждом из измерений

Сохранения размера
изображения можно достичь
дополнением исходного
изображения одним рядом с
каждой стороны



и т. д.

СУБДИСКРЕТИЗАЦИЯ 2×2 (MAX POOLING)

0	1	2	1
4	1	0	1
2	0	1	1
1	2	3	1

Исходный образец

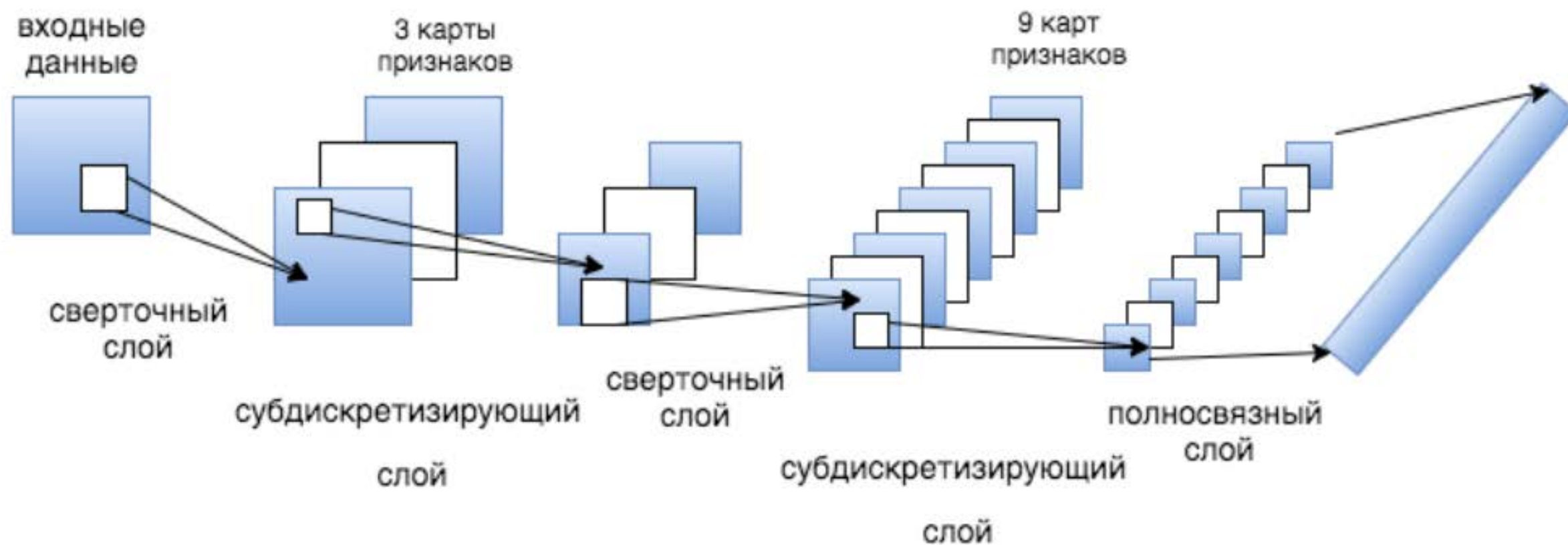
4	2	2
4	1	1
2	3	3

Субдискретизация
с шагом 1

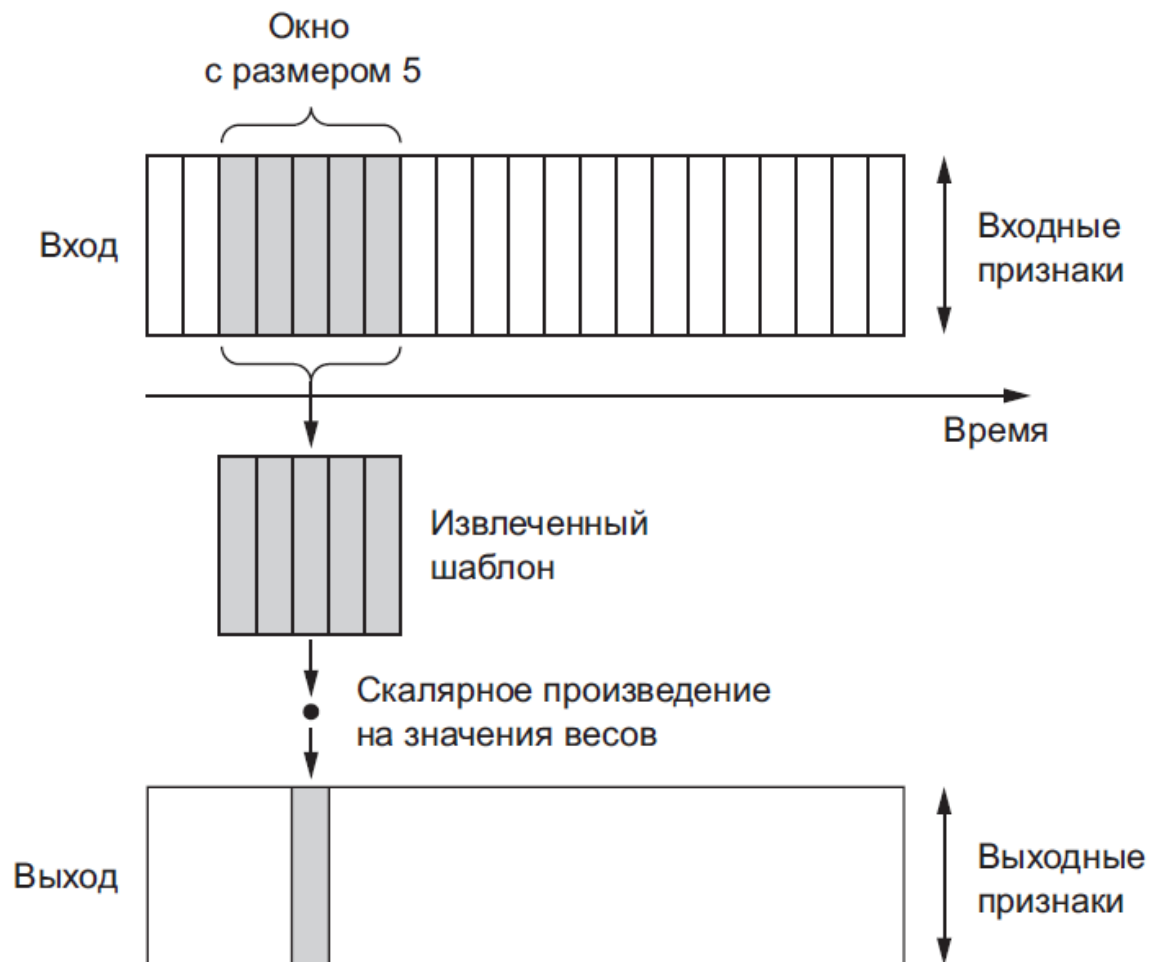
4	2
2	3

Субдискретизация
с шагом 2

БАЗОВАЯ АРХИТЕКТУРА СВЕРТОЧНОЙ СЕТИ

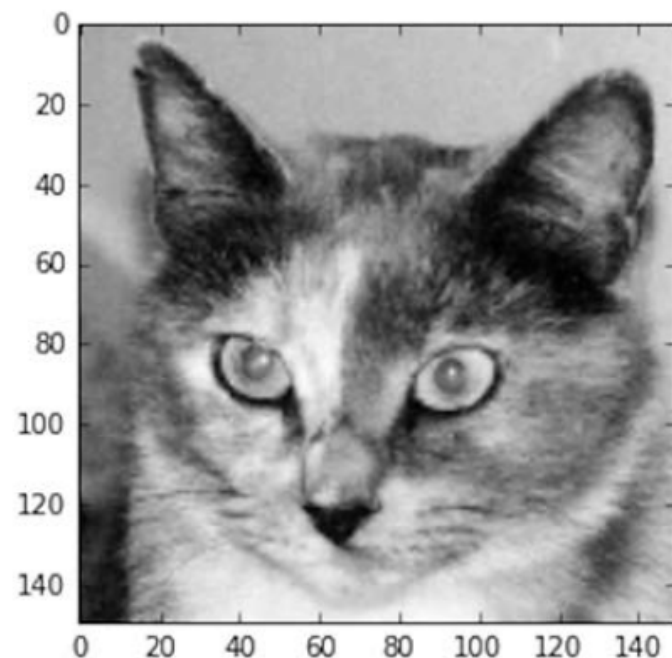


ОДНОМЕРНАЯ СВЕРТКА

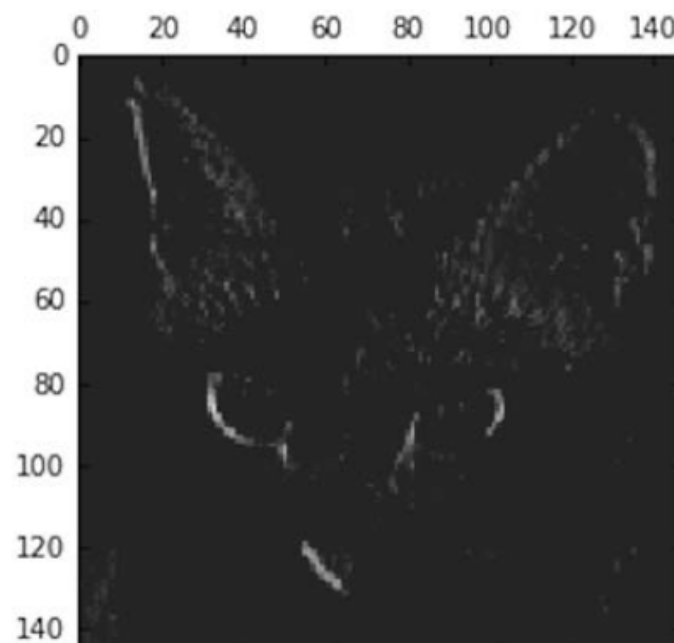


Одномерная свертка используется для анализа сигналов и последовательностей (например, текстов)

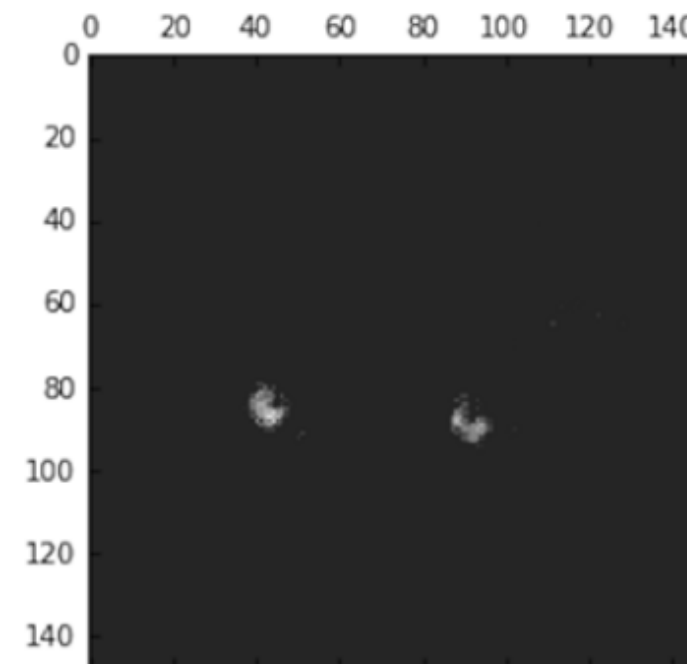
ПРЕДСТАВЛЕНИЯ ИЗОБРАЖЕНИЯ В СВЕРТОЧНОЙ СЕТИ



Исходное изображение

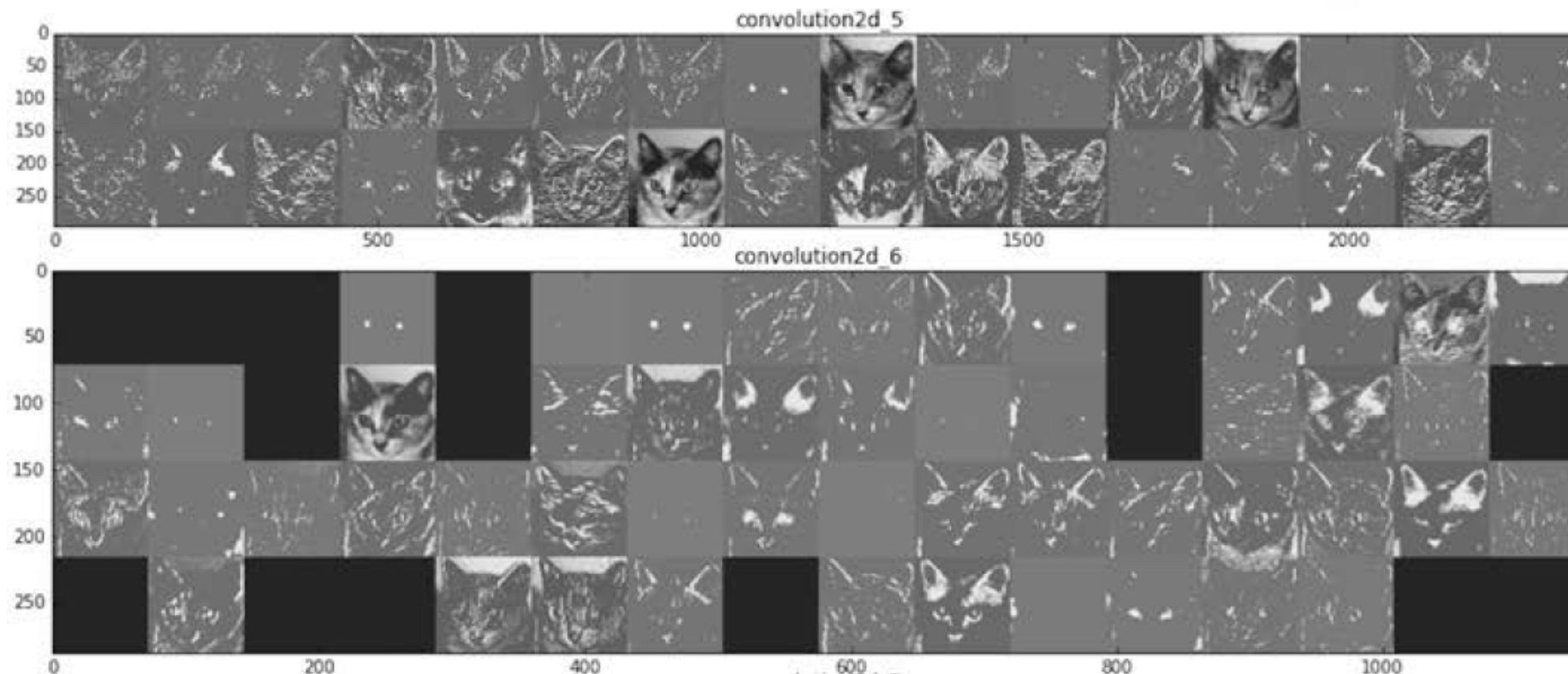


Четвертый канал
первого слоя



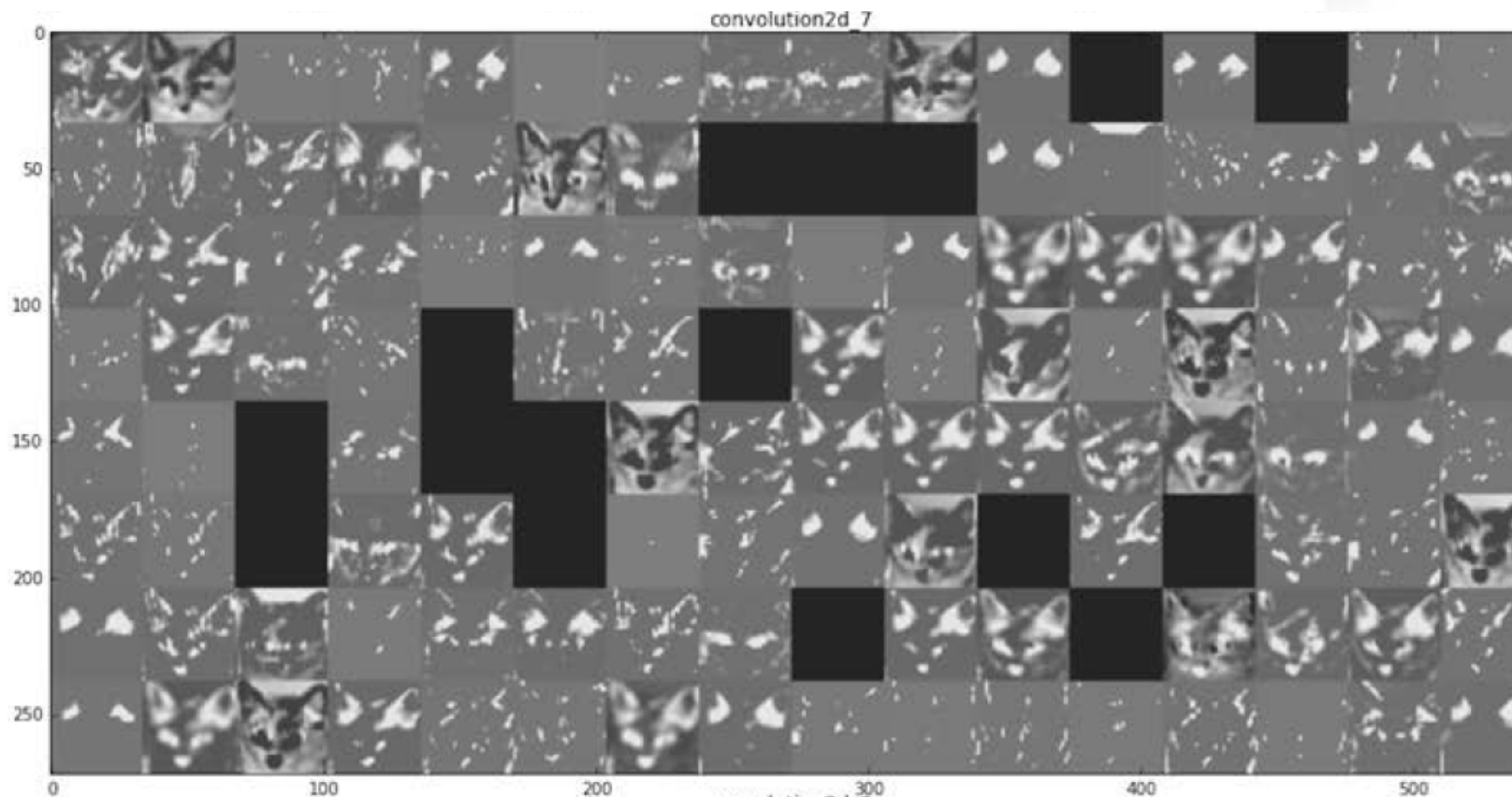
Седьмой канал
первого слоя

ПЕРВЫЕ СЛОИ СВЕРТКИ



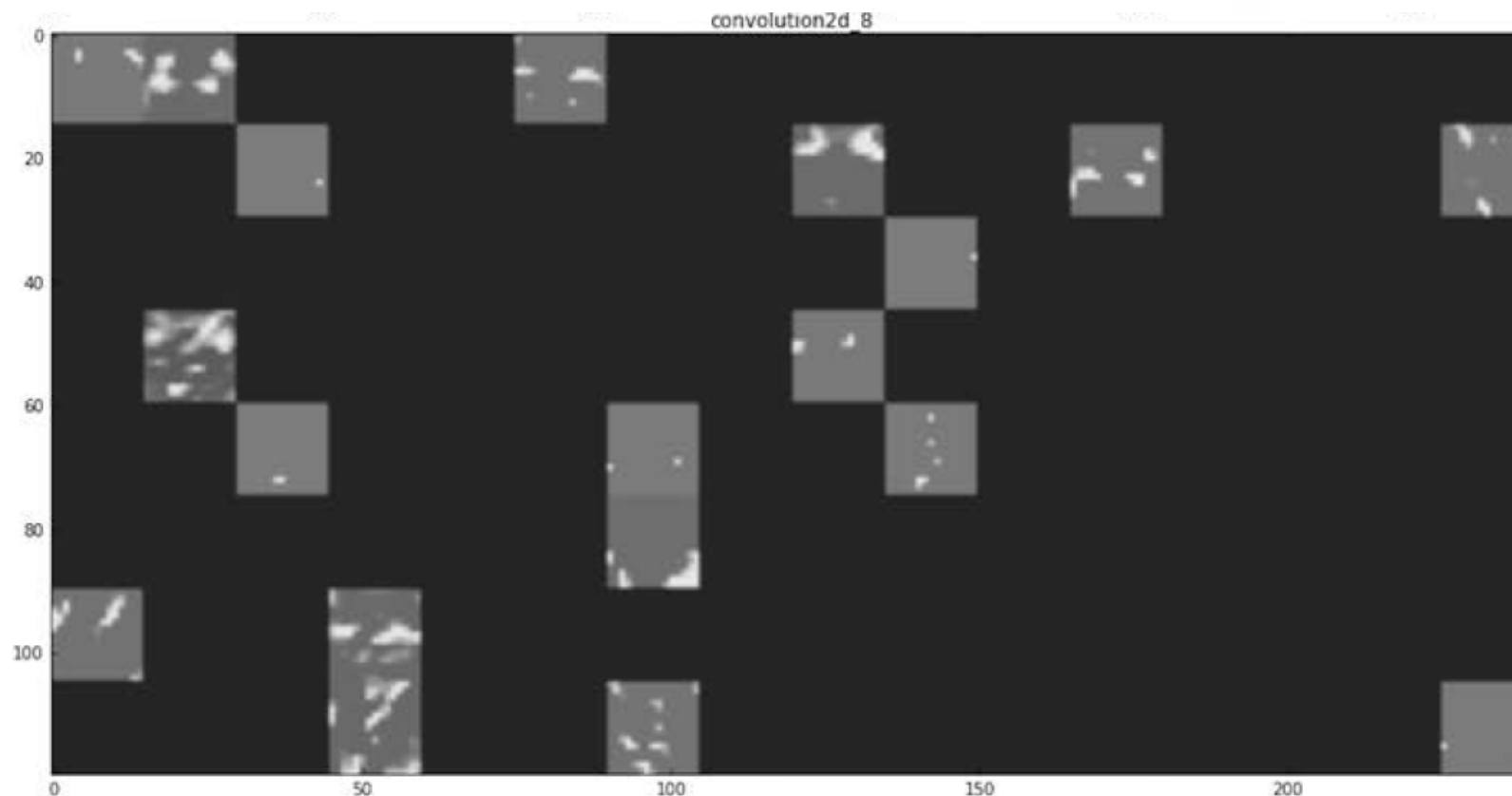
Первые слои действуют как коллекция детекторов контуров. На этом этапе сохраняется почти вся информация, имеющаяся в исходном изображении

ПРОМЕЖУТОЧНЫЕ СЛОИ СВЕРТКИ



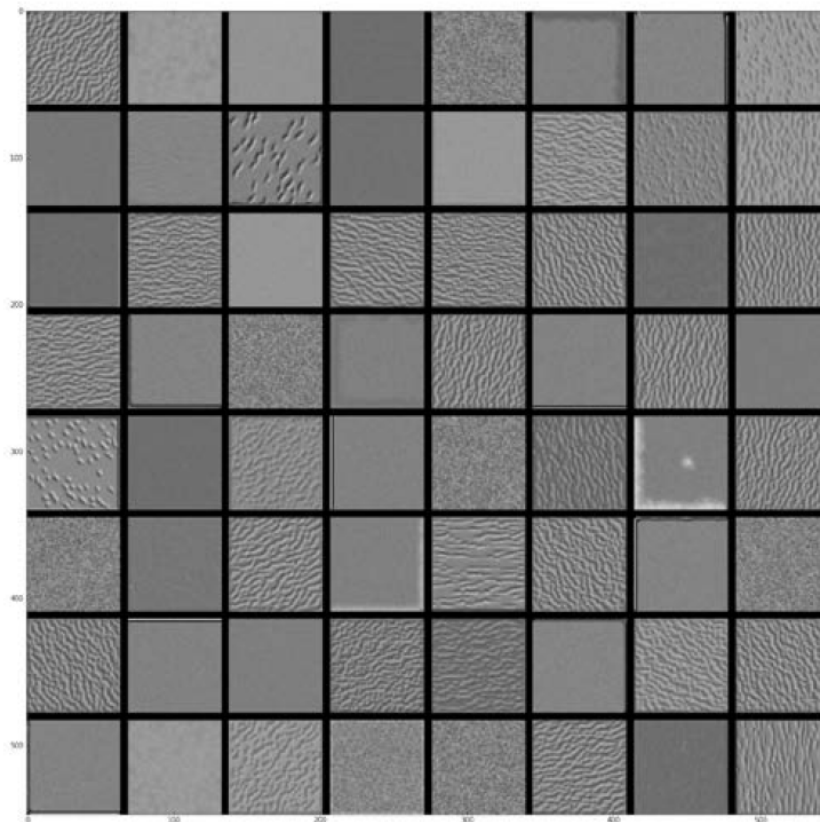
По мере продвижения по слоям, представления становятся все более абстрактными, а их визуальная интерпретация все более сложной. Высокоуровневые представления несут все меньше информации об исходном изображении и все больше — о классе изображения

КОНЕЧНЫЕ СЛОИ СВЕРТКИ

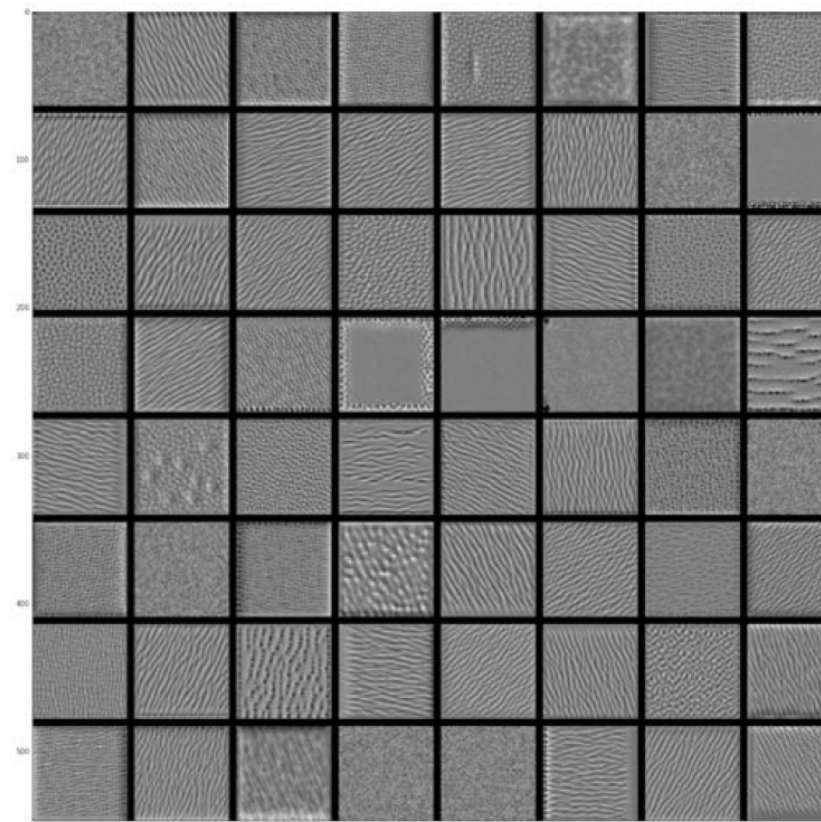


Разреженность активаций увеличивается с глубиной слоя: в первом слое все фильтры активируются исходным изображением, но в последующих слоях все больше и больше остается пустых фильтров. Это означает, что шаблон, соответствующий фильтру, не обнаруживается в исходном изображении

ШАБЛОНЫ ФИЛЬТРОВ ИЗ ПЕРВЫХ СЛОЕВ

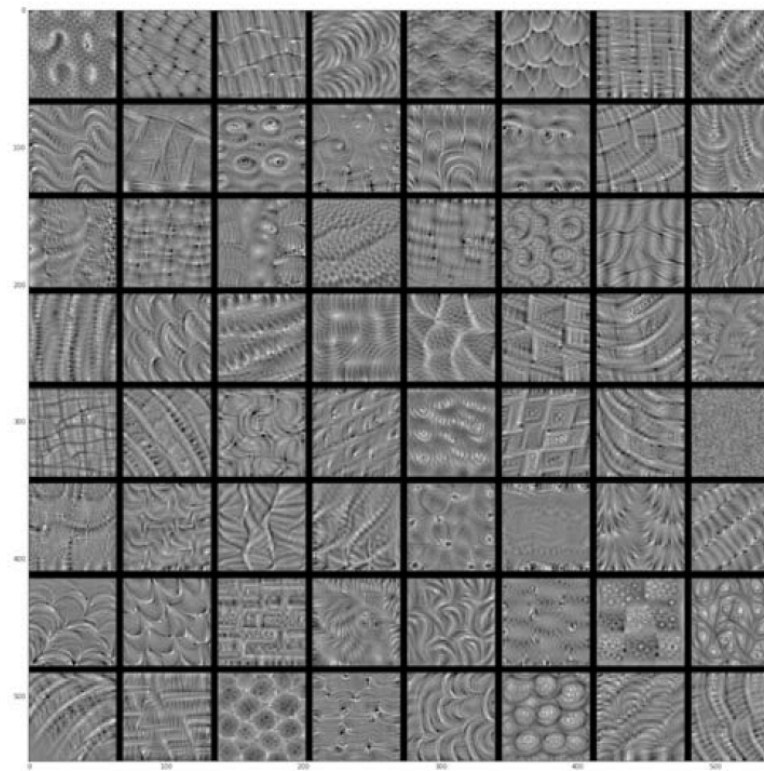
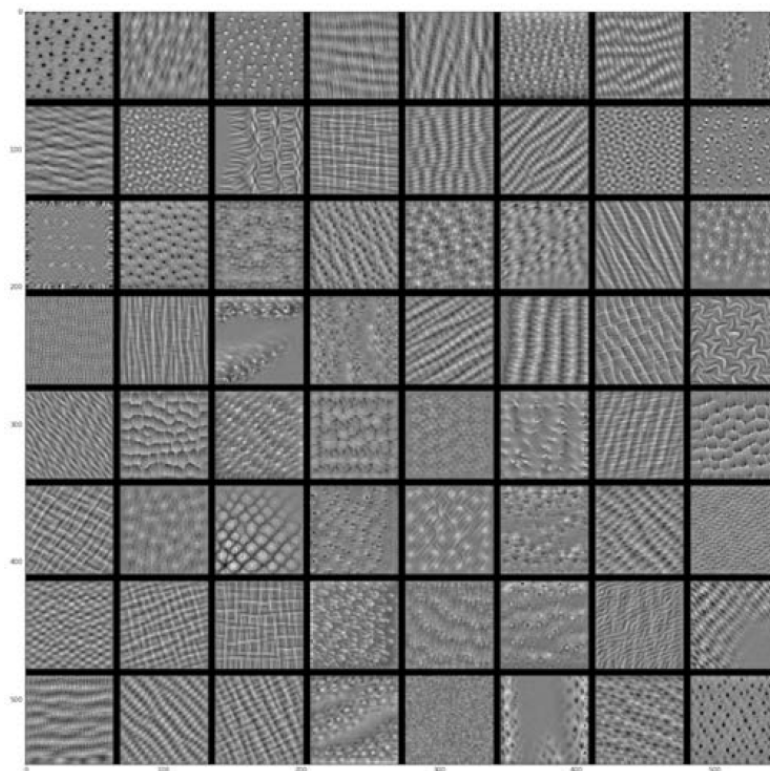


Фильтры из первых слоев в модели кодируют простые направленные контуры и цвета



Фильтры промежуточных слоев кодируют простые текстуры, состоящие из комбинаций контуров и цветов

ШАБЛОНЫ ФИЛЬТРОВ ИЗ ПОСЛЕДНИХ СЛОЕВ



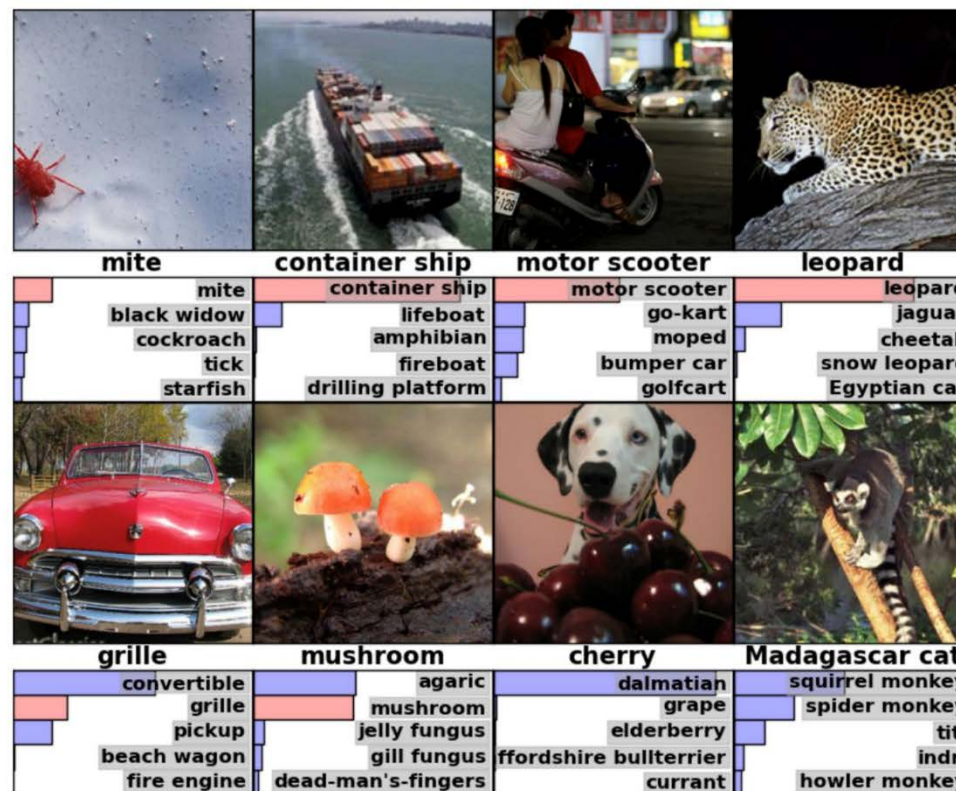
Фильтры в более поздних слоях начинают напоминать текстуры, встречающиеся в естественных изображениях, — перья, глаза, листья и т. д.

БАЗА ДАННЫХ «IMAGENET»

ImageNet Challenge



- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.



База данных **ImageNet** — проект по созданию и сопровождению массивной базы данных аннотированных изображений, предназначенная для отработки и тестирования методов распознавания образов и машинного зрения.

ПРОГРЕСС АЛГОРИТМОВ НА «IMAGENET»

