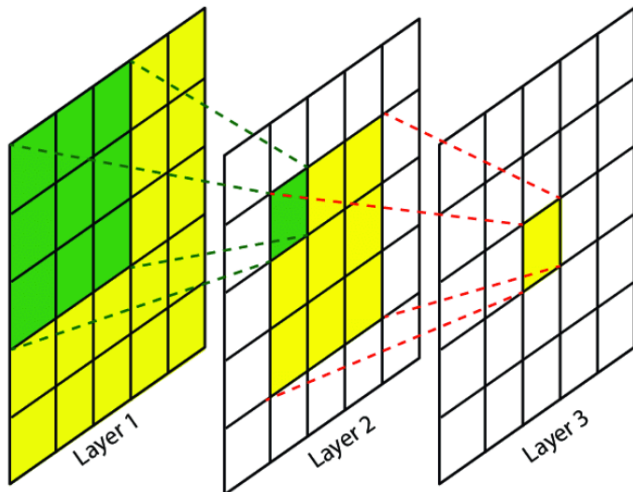


# CNN. Receptive Field

Каждый нейрон в CNN "**смотрит**" на определенный регион входного изображения. Это позволяет нейрону **выучивать** различные **паттерны**, которые есть на изображении, например, линии, какие-то детали представленных объектов.

# Receptive Field - Поле восприятия

Такой определенный регион на изображении называется **"Receptive Field"**



# Receptive Field

- Регион тензора выходного слоя, который участвовал в вычислении конкретной характеристики текущего слоя;
- Регион на изображении, куда смотрит каждая характеристика сверточной нейронной сети.

Каждой такой регион определяется свои **центром** и **размером**.

# Формулы для вычисления Receptive Field

1. Расстояния между центрами Receptive Field для характеристик текущего слоя:

$$j_{out} = j_{in} * stride_{prev}$$

На старте  $j_{in} = 1$

2. Размер RF для текущего слоя:

$$r_{out} = r_{in} + (k - 1) * j_{in}$$

$k$  -- это размер фильтра. На старте  $r_{in} = 1$

3. Центр RF:

$$start_{out} = start_{in} + \left(\frac{k - 1}{2} - p\right) * j_{in}$$

$p$  -- это размер padding.

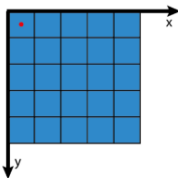
$n$  : number of features  
 $r$  : receptive field size  
 $j$  : jump (distance between two consecutive features)  
 $start$  : center coordinate of the first feature

$k$  : convolution kernel size  
 $p$  : convolution padding size  
 $s$  : convolution stride size

$$\begin{aligned}
 n_{out} &= \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1 \\
 j_{out} &= j_{in} * s \\
 r_{out} &= r_{in} + (k - 1) * j_{in} \\
 start_{out} &= start_{in} + \left( \frac{k - 1}{2} - p \right) * j_{in}
 \end{aligned}$$

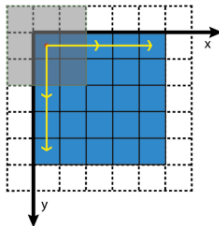
Layer 0:

$n_0 = 5$ ;  $r_0 = 1$ ;  $j_0 = 1$ ;  
 $start_0 = 0.5$



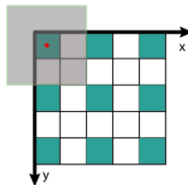
Conv1:

$k_1 = 3$ ;  $p_1 = 1$ ;  $s_1 = 2$



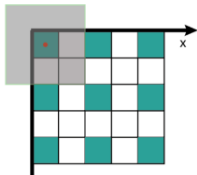
Layer 1:

$n_1 = 3$ ;  $r_1 = 3$ ;  $j_1 = 2$ ;  
 $start_1 = 0.5$



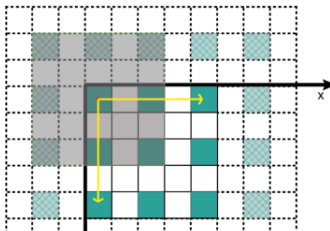
Layer 1:

$n_1 = 3$ ;  $r_1 = 3$ ;  $j_1 = 2$ ;  
 $start_1 = 0.5$



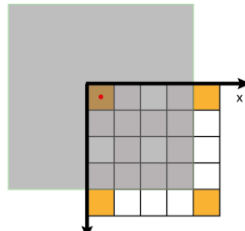
Conv2:

$k_2 = 3$ ;  $p_2 = 1$ ;  $s_2 = 2$



Layer 2:

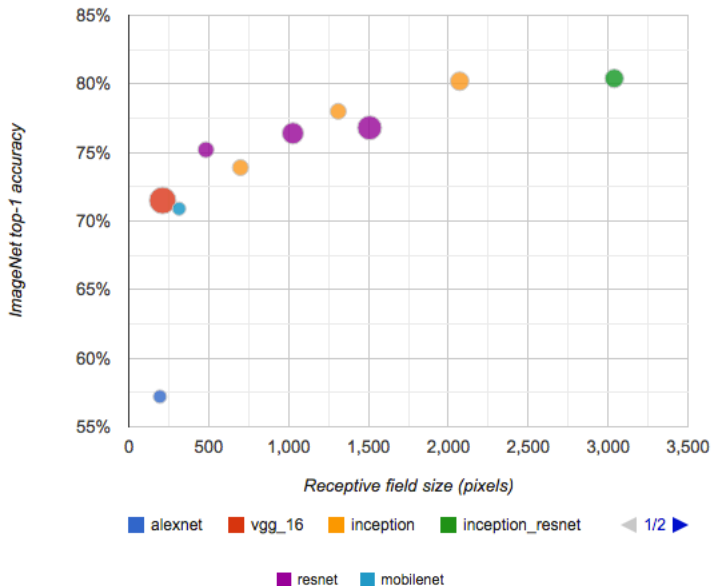
$n_2 = 2$ ;  $r_2 = 7$ ;  $j_2 = 4$ ;  
 $start_2 = 0.5$



# Что дает понять Receptive Field

- Для задач image segmentation и object detection важно понимать, какого размера Receptive Field, чтобы оценить, насколько хорошо нейрон сможет охватывать объекты на изображении.
- Object Detector плохо будет работать с большими объектами, если у него маленький Receptive Field.
- Исходя из допустимого размера Receptive Field регулируется глубина сети и, в целом, вся ее архитектура.

# Зависимость точности от размера RF



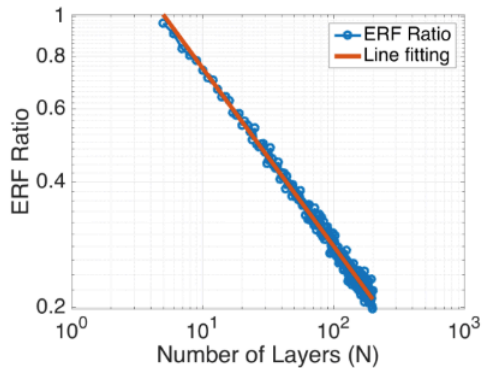
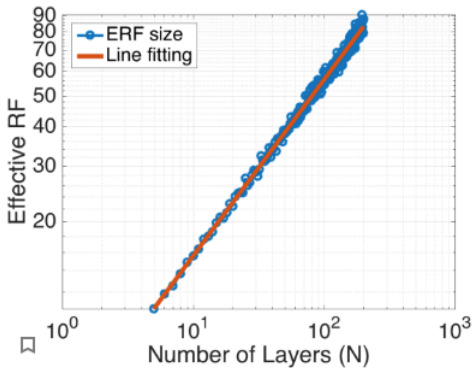


# Effective Receptive Field

- Не все пиксели, входящие в Receptive Field выходного нейрона, одинаково влияют.
- Центр RF больше всего влияет на выходной нейрон, так как могут быть еще и skip connection или другие связи, где центр RF будет чаще участвовать, чем те, что на окраине.
- Такое влияние центра RF проявляется при вычислении градиентов. Именно здесь видно, кто вносит БОЛЬШОЙ вклад.

<https://arxiv.org/pdf/1701.04128.pdf>

# Effective Receptive Field



ConvNet Model	Receptive Field (r)	Effective Stride (S)	Effective Padding (P)	Model Year
alexnet_v2	195	32	64	<u>2014</u>
vgg_16	212	32	90	<u>2014</u>
mobilenet_v1	315	32	126	<u>2017</u>
mobilenet_v1_075	315	32	126	<u>2017</u>
resnet_v1_50	483	32	239	<u>2015</u>
inception_v2	699	32	318	<u>2015</u>
resnet_v1_101	1027	32	511	<u>2015</u>
inception_v3	1311	32	618	<u>2015</u>
resnet_v1_152	1507	32	751	<u>2015</u>
resnet_v1_200	1763	32	879	<u>2015</u>
inception_v4	2071	32	998	<u>2016</u>
inception_resnet_v2	3039	32	1482	<u>2016</u>