

# Lab Meeting Presentation-4

June 21, 2017

Prerit Gupta  
Summer Intern

Ravi Kiran Sarvadevabhatla  
PhD Student

# Objective

- To plot the top 5 preferred and least 5 preferred classes for every layer of Alexnet & VGG-19 as a result from histogram plots of the maximum activated image data collected from Yosinski's toolbox.

# Characterizing Visualizations in layers of Alexnet network

(Training of Alexnet uses mean file of sketch dataset)

# Convolution Layer 1



Input Patch for Maximum Activation



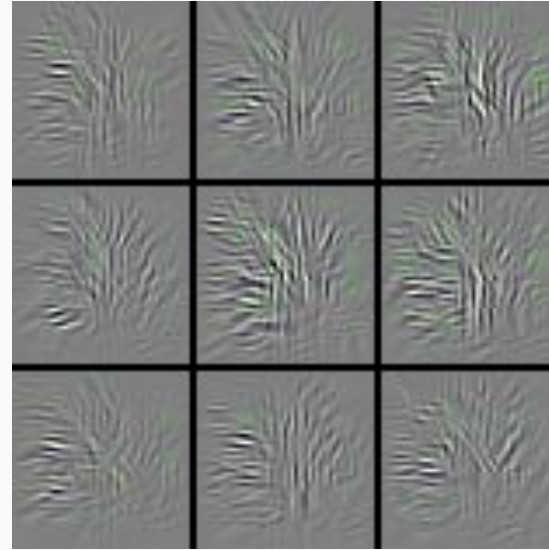
Deconvolution of the activated patch

- 11 x 11 Input Patch for 2nd filter in 1st convolutional layer.
- Maximum activation : 2652.406982
- Filter sensitive to edge transition.
- Classes with performance: zebra 92.86 mouse (animal) 71.43 fork 85.71 calculator 92.86 scissors 100.0 pig 35.71 elephant 78.57 pizza 100.0 pineapple 92.86

# Convolution Layer 2



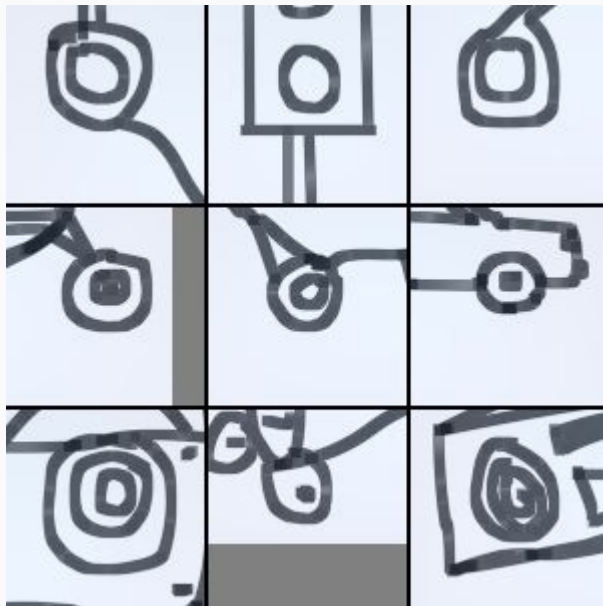
Input Patch for Maximum Activation



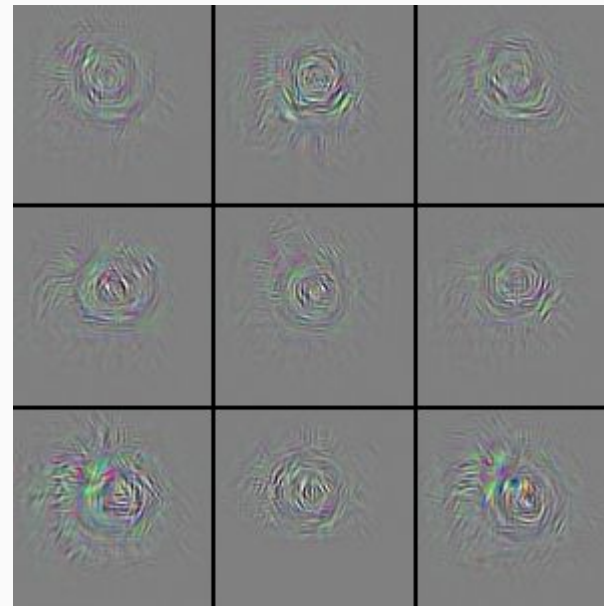
Deconvolution of the activated patch

- 51 x 51 Input Patch for 200th filter in 2nd convolutional layer.
- Maximum activation: 499.30127
- Filter sensitive to lines.
- Classes with performance: grapes 100.0 pineapple 92.86 butterfly 78.57 monkey 42.86 tree 100.0 strawberry 92.86 pineapple 92.86 potted plant 85.71 binoculars 64.29

# Convolution Layer 3



Input Patch for Maximum Activation



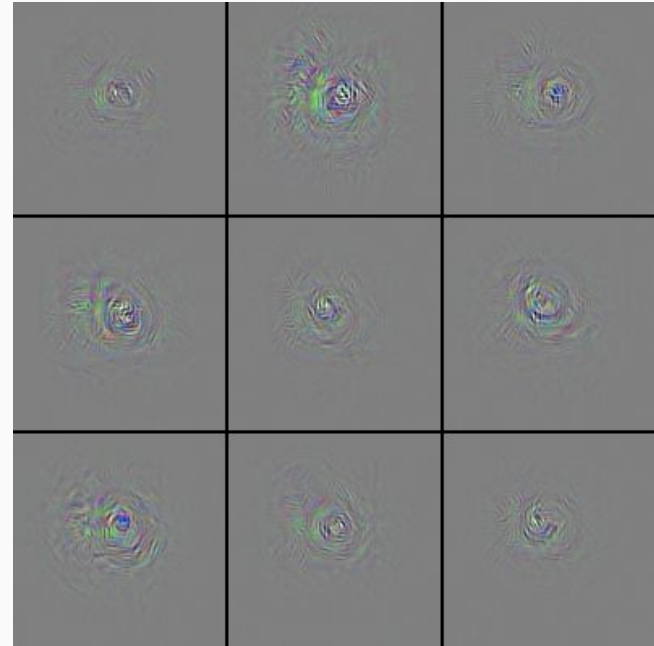
Deconvolution of the activated patch

- 99 x 99 Input Patch for 284th filter in 3rd convolutional layer.
- Maximum Activation: 428.343414
- Filter sensitive to specific parts of sketches having circles.
- Classes with performance: head-phones 100.0 traffic light 92.86 scissors 100.0 wheelbarrow 92.86 wheelbarrow 92.86 car (sedan) 92.86 camera 85.71 train 78.57 radio 57.14

# Convolution Layer 4



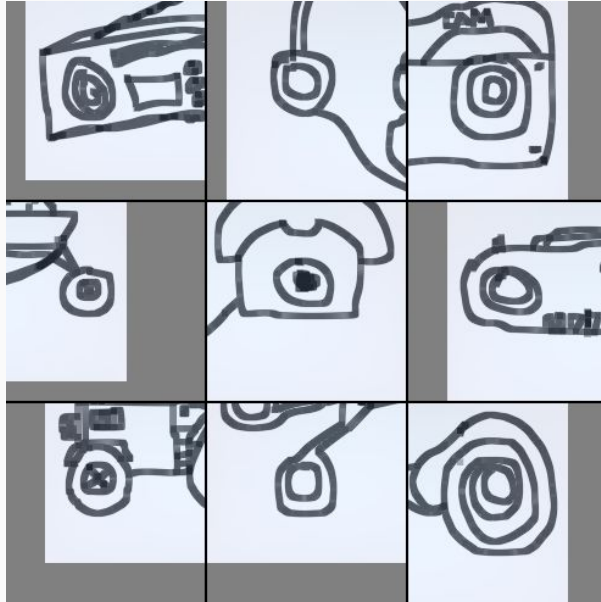
Input Patch for Maximum Activation



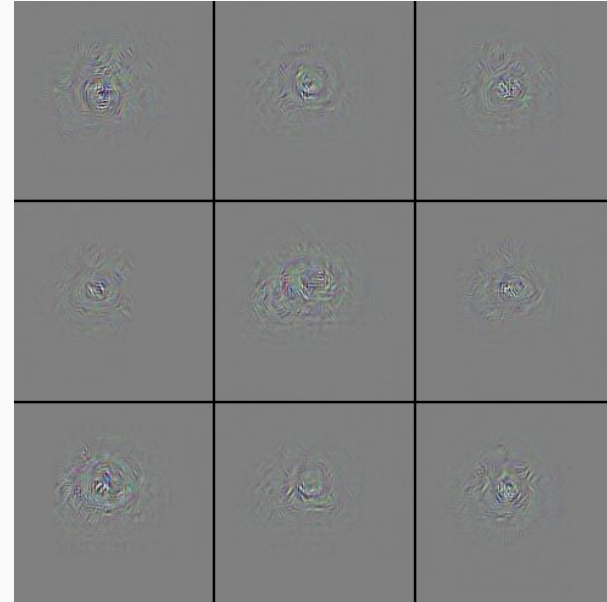
Deconvolution of the activated patch

- 131 x 131 Input Patch for 200th filter in 4th convolutional layer
- Maximum Activation: 292.267822
- Filter sensitive to specific classes like snails & more towards round & spiral objects.
- Classes with performance: wheelbarrow 92.86 camera 85.71 radio 57.14 radio 57.14 snail 100.0 head-phones 100.0 head-phones 100.0 wheelbarrow 92.86 snail 100.0

# Convolution Layer 5



Input Match for Maximum Activation

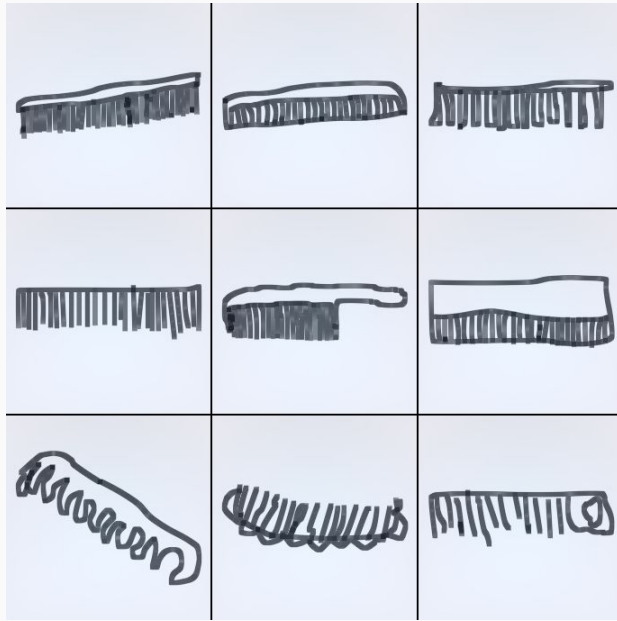


Deconvolution of the activated patch

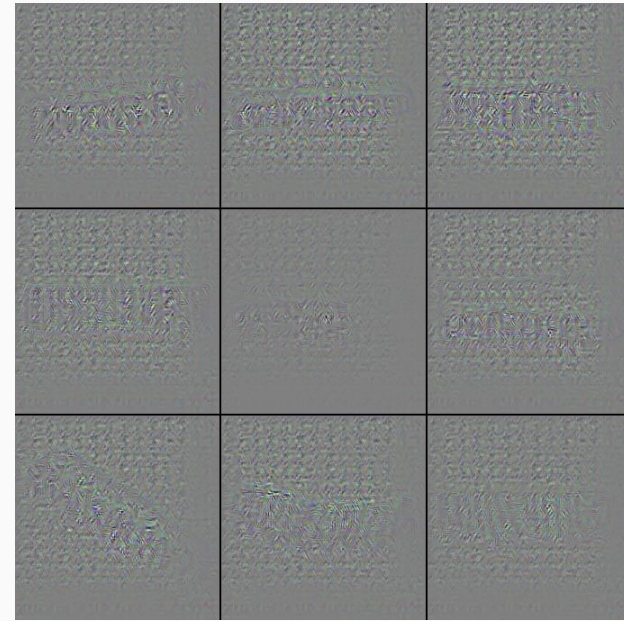
- 163 x 163 Input Patch for 173th filter in 5th convolutional layer
- Maximum Activation: 197.762451
- Filter sensitive to specific classes that have round & spiral objects.
- Classes with performance: radio 57.14 head-phones 100.0 camera 85.71 wheelbarrow 92.86 telephone 78.57 radio 57.14 tractor 85.71 scissors 100.0 snail 100.



# Fully Connected Layer 8



Input Patch for Maximum Activation



Deconvolution of the activated patch




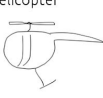




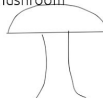


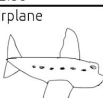
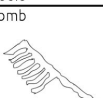
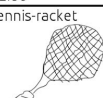


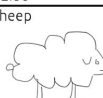



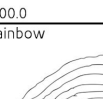
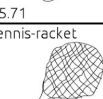
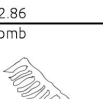


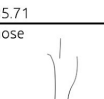
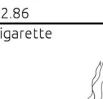
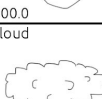
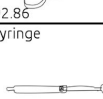
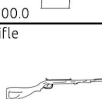
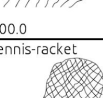



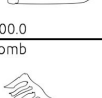
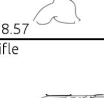
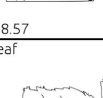
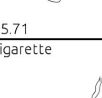
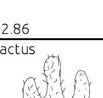
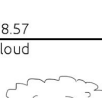
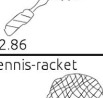
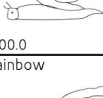
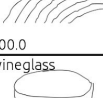

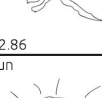
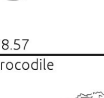
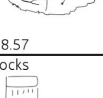
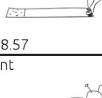
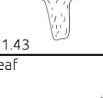

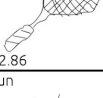
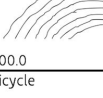
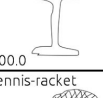
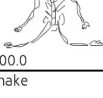

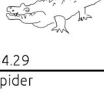
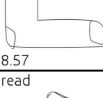
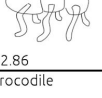
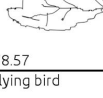

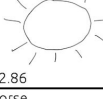

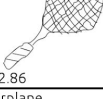


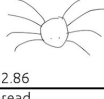
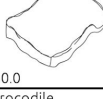


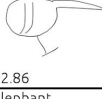
- 227 x 227 Input Patch (complete image) for 36th filter in 8th layer which is fully connected.
- Maximum Activation: 46.98156
- Filter sensitive to the comb class. (which is also 36th class in target attribute)
- Classes with performance: comb 92.86 comb 92.86 comb 92.86 comb 92.86 comb 92.86  
piano 100.0 comb 92.86 comb 92.86 comb 92.86



# Summary Plot with class performance(Alexnet with Imagenet Mean)

## Top 5 Classes

## Bottom 5 Classes

	butterfly	zebra	alarm clock	helicopter	fork	cloud	eyeglasses	mouth	mushroom	suitcase
Conv1										
	78.57	92.86	100.0	92.86	85.71	85.71	92.86	92.86	92.86	92.86
Conv2	rainbow	airplane	comb	tennis-racket	hedgehog	cloud	sheep	snowman	ear	candle
										
	100.0	85.71	92.86	92.86	85.71	85.71	92.86	100.0	92.86	100.0
Conv3	rainbow	tennis-racket	comb	mermaid	snail	nose	cigarette	cloud	syringe	rifle
										
	100.0	92.86	92.86	85.71	100.0	78.57	78.57	85.71	92.86	78.57
Conv4	tennis-racket	snail	rainbow	owl	comb	rifle	leaf	cigarette	cactus	cloud
										
	92.86	100.0	100.0	64.29	92.86	78.57	78.57	78.57	71.43	85.71
Conv5	tennis-racket	rainbow	wineglass	human-skeleton	sun	crocodile	socks	ant	leaf	pen
										
	92.86	100.0	100.0	100.0	92.86	64.29	78.57	92.86	78.57	78.57
FC8	sun	bicycle	tennis-racket	snake	pipe (for smoking)	spider	bread	crocodile	flying bird	helicopter
										
	92.86	92.86	92.86	85.71	85.71	42.86	50.0	64.29	50.0	92.86
Softmax	horse	bee	airplane	person sitting	tv	bread	crocodile	flying bird	mouse (animal)	elephant
										
	92.86	71.43	85.71	71.43	100.0	50.0	64.29	50.0	71.43	78.57

# Characterizing Visualizations in layers of VGG-19 network

# Conv1\_1 Layer



Input Patch for Maximum Activation



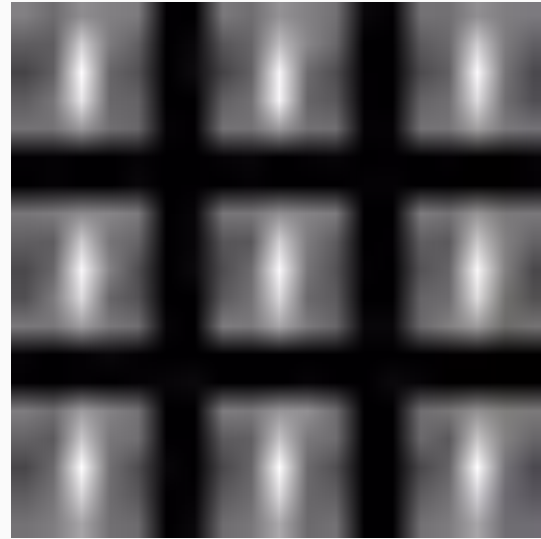
Deconvolution of the activated patch

- 3 x 3 Input Patch for 52nd filter in Conv1\_1 layer
- Maximum Activation: 852.20105
- Filter sensitive to edge transition.
- Classes with performance: chair 92.86 toilet 100.0 giraffe 100.0 horse 92.86 owl 64.29 owl 64.29 rifle 85.71 butterfly 100.0 angel 71.43

# Conv1\_2 Layer



Input Patch for Maximum Activation



Deconvolution of the activated patch

- 5 x 5 Input Patch for 1st filter in Conv1\_2 layer
- Maximum Activation: 3120.886475
- Filter sensitive to edge transition.
- Classes with performance: bathtub 78.57 snowman 100.0  
human-skeleton 85.71 angel 71.43 frog 85.71 train 85.71  
carrot 85.71 pineapple 92.86 bathtub 78.57

# Conv2\_1 Layer



Input Patch for Maximum Activation



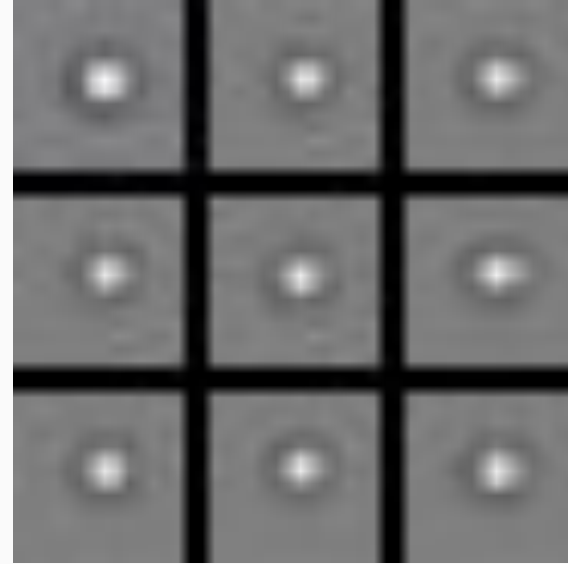
Deconvolution of the activated patch

- 10 x 10 Input Patch for 60th filter in Conv2\_1layer
- Maximum Activation: 4891.480957
- Filter sensitive to edge transition.
- Classes with performance: tennis-racket 85.71 train 85.71  
microscope 92.86 snowman 100.0 zebra 100.0 angel 71.43  
mouse (animal) 64.29 human-skeleton 85.71 snail 100.0

## Conv2\_2 Layer



Input Patch for Maximum Activation



Deconvolution of the activated patch

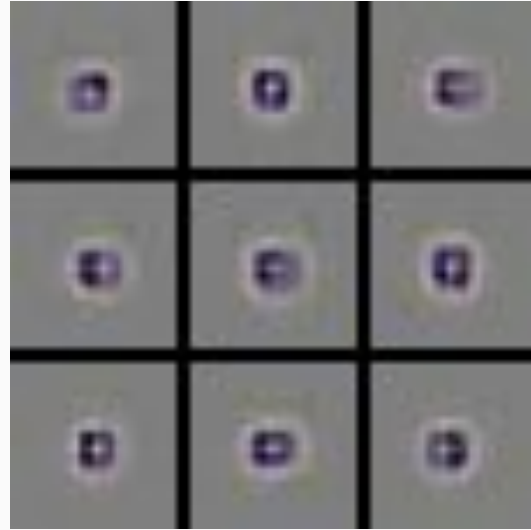
- 14 x 14 Input Patch for 80th filter in Conv2\_2 layer
- Maximum Activation: 6866.766113
- Filter sensitive to centre white patch.
- Classes with performance: spongebob 92.86 ice-cream-cone 78.57 frog 85.71 face 100.0 cannon 57.14 violin 85.71 tennis-racket 85.71 foot 71.43 trumpet 78.57



# Conv3\_1 Layer



Input Patch for Maximum Activation



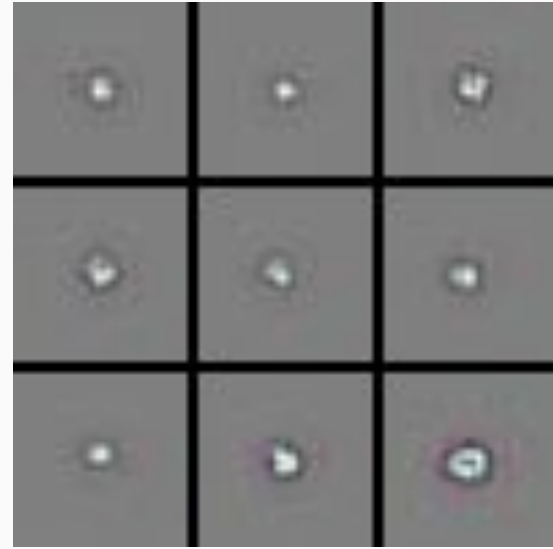
Deconvolution of the activated patch

- 24 x 24 Input Patch for 97th filter in Conv3\_1 layer
- Maximum Activation: 8785.024414
- Filter sensitive to centre black square.
- Classes with performance: camel 92.86 hedgehog 100.0 scorpion 64.29 flying bird 71.43 car (sedan) 100.0 snowman 100.0 pizza 100.0 person walking 100.0 snowman 100.0

## Conv3\_2 Layer



Input Patch for Maximum Activation



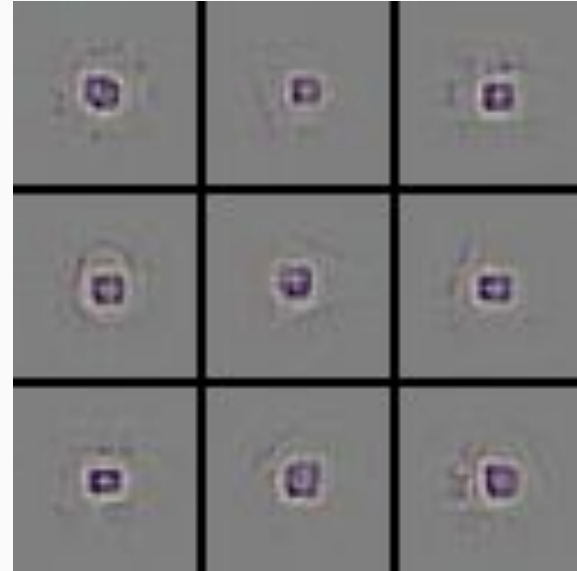
Deconvolution of the activated patch

- 32 x 32 Input Patch for 245th filter in Conv3\_2 layer
- Maximum Activation: 8998.217773
- Filter sensitive to centred circular patch.
- Classes with performance: ipod 78.57 frog 85.71 castle 92.86  
foot 71.43 butterfly 100.0 pizza 100.0 sponge bob 92.86 radio  
85.71 grapes 100.0

# Conv3\_3 Layer



Input Patch for Maximum Activation



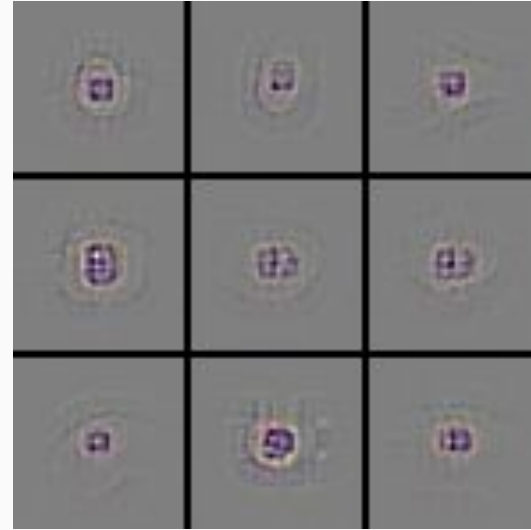
Deconvolution of the activated patch

- 40 x 40 Input Patch for 83rd filter in Conv3\_3 layer
- Maximum Activation: 11305.711914
- Filter sensitive to patches with small black squares in between.
- Classes with performance: wrist-watch 85.71 sponge bob 92.86 cow 85.71 car (sedan) 100.0 pizza 100.0 scorpion 64.29 person walking 100.0 pizza 100.0 penguin 92.86

# Conv3\_4 Layer



Input Patch for Maximum Activation



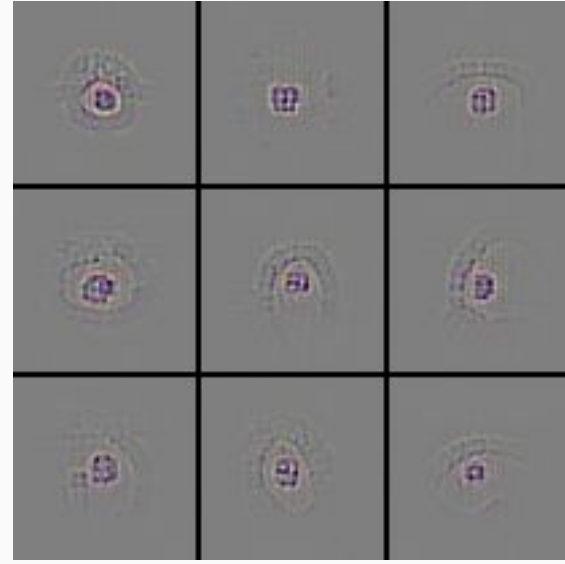
Deconvolution of the activated patch

- 48 x 48 Input Patch for 62nd filter in Conv3\_4 layer
- Maximum Activation: 20115.480469
- Filter sensitive to square enclosed within a circle.
- Classes with performance: car (sedan) 100.0 octopus 92.86  
pizza 100.0 bus 100.0 rollerblades 78.57 rabbit 85.71 person  
walking 100.0 bus 100.0 pig 85.71

# Conv4\_1 Layer



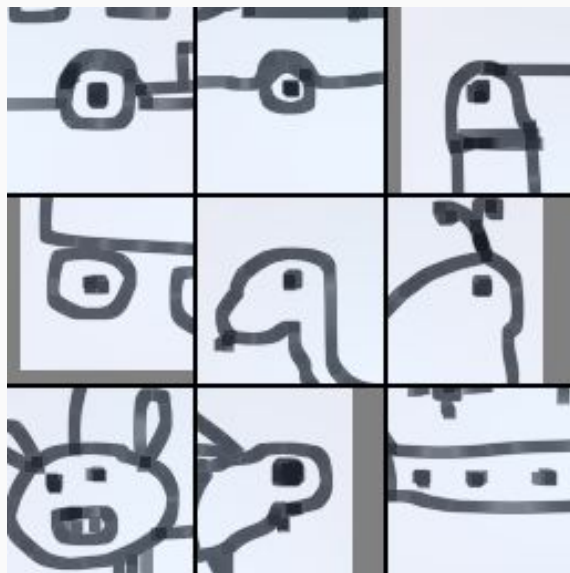
Input Patch for Maximum Activation



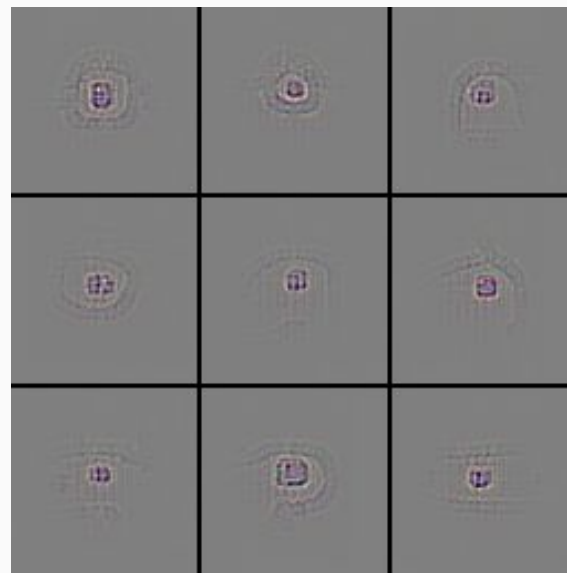
Deconvolution of the activated patch

- 68 x 68 Input Patch for 265th filter in Conv4\_1 layer
- Maximum Activation: 24987.294922
- Filter sensitive to square enclosed within a circle.
- Classes with performance: car (sedan) 100.0 pizza 100.0  
snake 85.71 rollerblades 78.57 duck 100.0 duck 100.0  
snowman 100.0 head-phones 92.86 person walking 100.0

## Conv4\_2 Layer



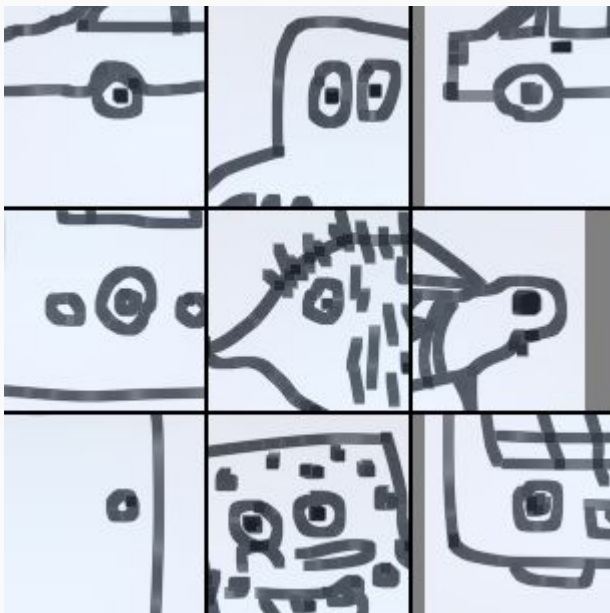
Input Patch for Maximum Activation



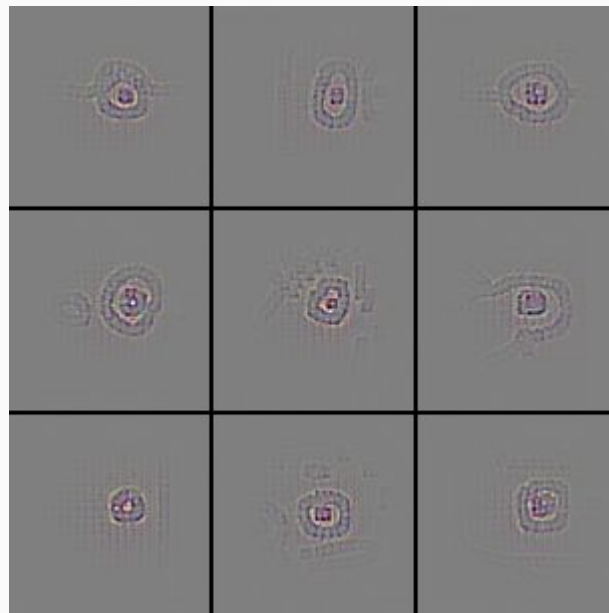
Deconvolution of the activated patch

- 84 x 84 Input Patch for 419th filter in Conv4\_2 layer
- Maximum Activation: 24043.347656
- Filter sensitive to faces, leaves etc.
- Classes with performance: bus 100.0 car (sedan) 100.0 mailbox 85.71 rollerblades 78.57 snake 85.71 snail 100.0 pig 85.71 bee 78.57 crown 71.43

## Conv4\_3 Layer



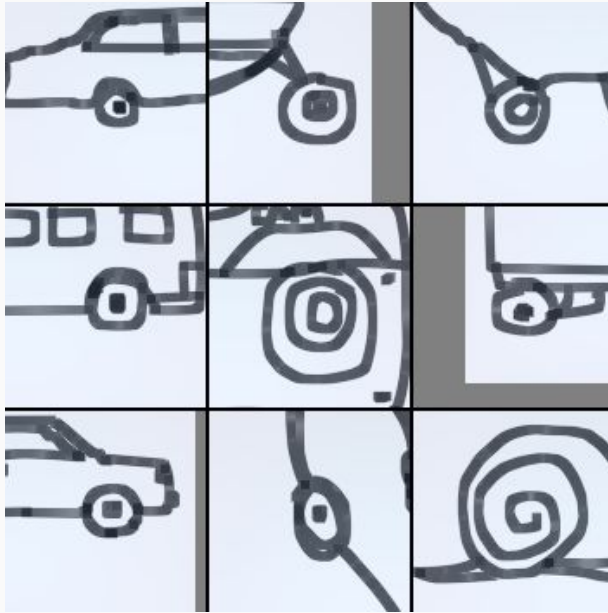
Input Patch for Maximum Activation



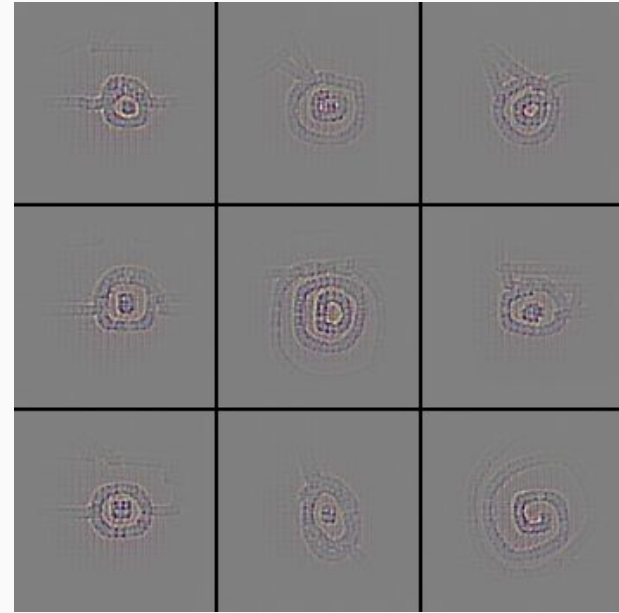
Deconvolution of the activated patch

- 100 x 100 Input Patch for 309th filter in Conv4\_3 layer
- Maximum Activation: 12583.713867
- Filter sensitive to circle enclosing dots.
- Classes with performance: car (sedan) 100.0 octopus 92.86  
car (sedan) 100.0 radio 85.71 hedgehog 100.0 bee 78.57  
door 100.0 sponge bob 92.86 radio 85.71

# Conv4\_4 Layer



Input Patch for Maximum Activation

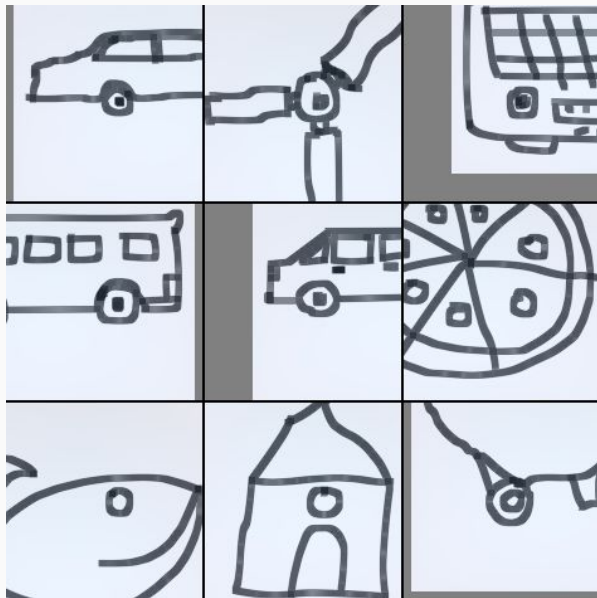


Deconvolution of the activated patch

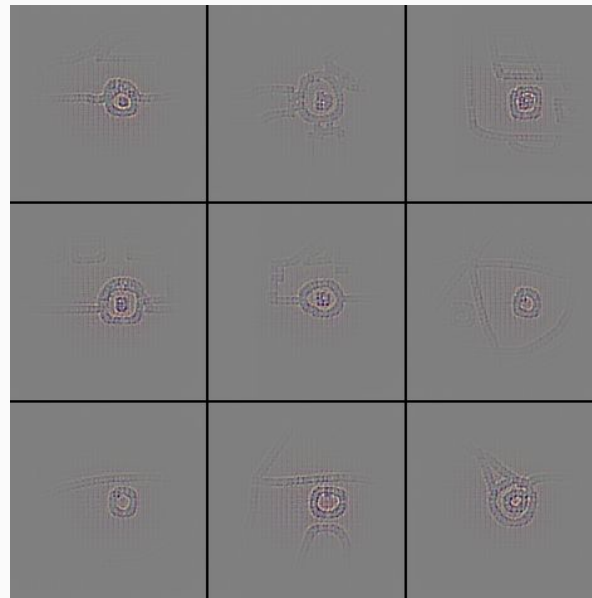
- 116 x 116 Input Patch for 319th filter in Conv4\_4 layer
- Maximum Activation: 7820.024414
- Filter sensitive to round & spiral patterns.
- Classes with performance: car (sedan) 100.0 wheelbarrow 100.0 wheelbarrow 100.0 bus 100.0 camera 85.71 rollerblades 78.57 car (sedan) 100.0 head-phones 92.86 snail 100.0



# Conv5\_1 Layer



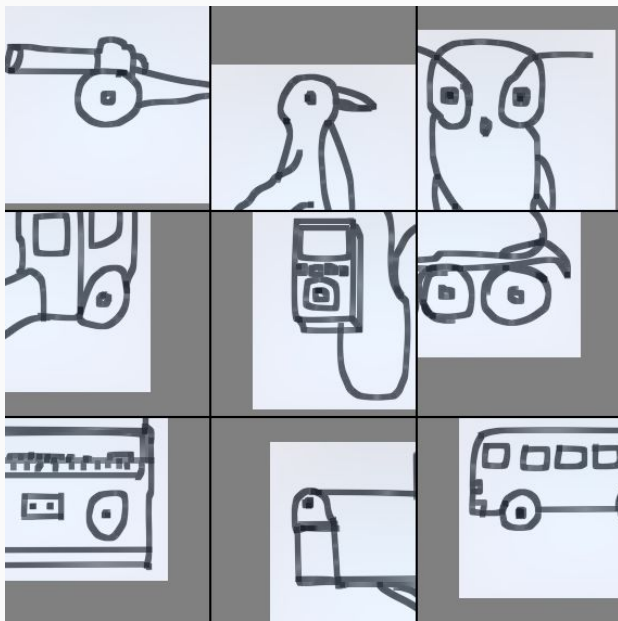
Input Patch for Maximum Activation



Deconvolution of the activated patch

- 156 x 156 Input Patch for 169th filter in Conv5\_1 layer
- Maximum Activation: 4377.005859
- Filter sensitive to random patterns.
- Classes with performance: car (sedan) 100.0 fan 85.71 radio 85.71 bus 100.0 car (sedan) 100.0 pizza 100.0 fish 85.71 church 100.0 wheelbarrow 100.0

## Conv5\_2 Layer



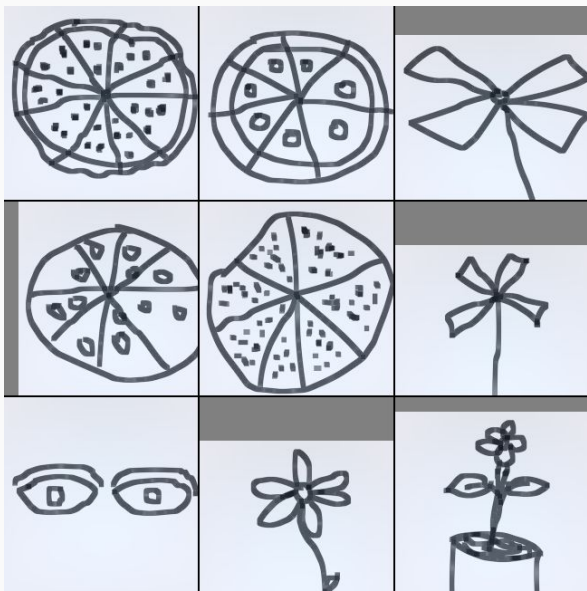
Input Patch for Maximum Activation



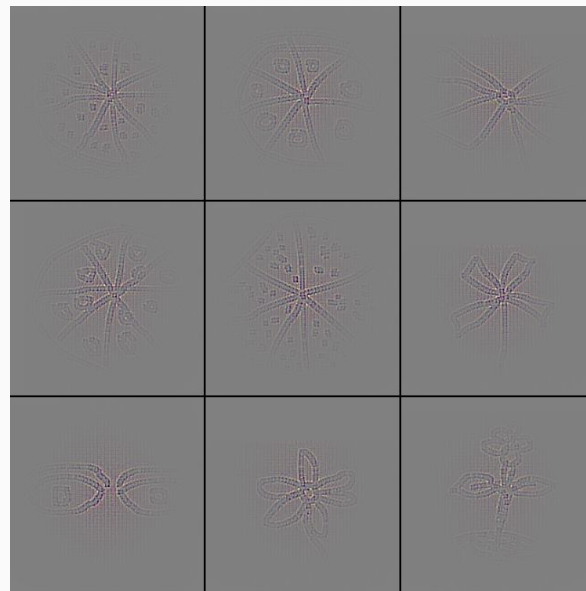
Deconvolution of the activated patch

- 188 x 188 Input Patch for 64th filter in conv5\_2 layer
- Maximum Activation: 2197.940186
- Filter sensitive to circle enclosing dot
- Classes with performance: cannon 57.14 penguin 92.86 owl 64.29 train 85.71 ipod 78.57 rollerblades 78.57 radio 85.71 mailbox 85.71 bus 100.0

# Conv5\_3 Layer



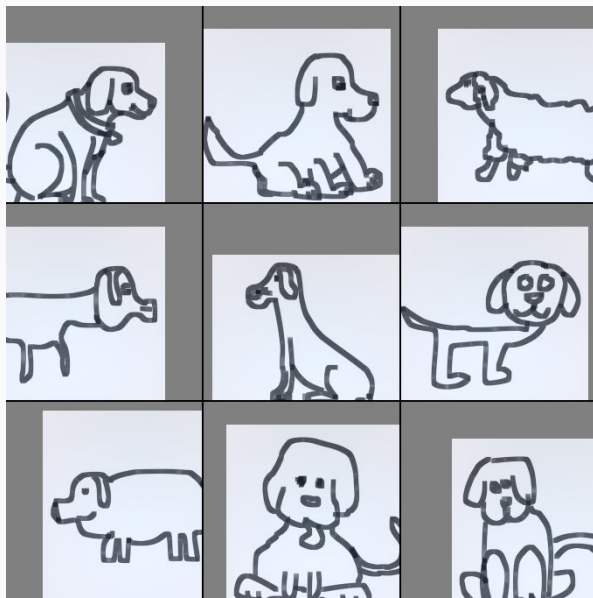
Input Patch for Maximum Activation



Deconvolution of the activated patch

- 214 x 214 Input Patch for 1st filter in Conv5\_3 layer
- Maximum Activation: 1306.398193
- Filter sensitive to specific classes like pizza, windmill, flower.
- Classes with performance: pizza 100.0 pizza 100.0 windmill 78.57 pizza 100.0 pizza 100.0 windmill 78.57 eye 100.0 flower with stem 92.86 potted plant 100.0

## Conv5\_4 Layer



Input Patch for Maximum Activation

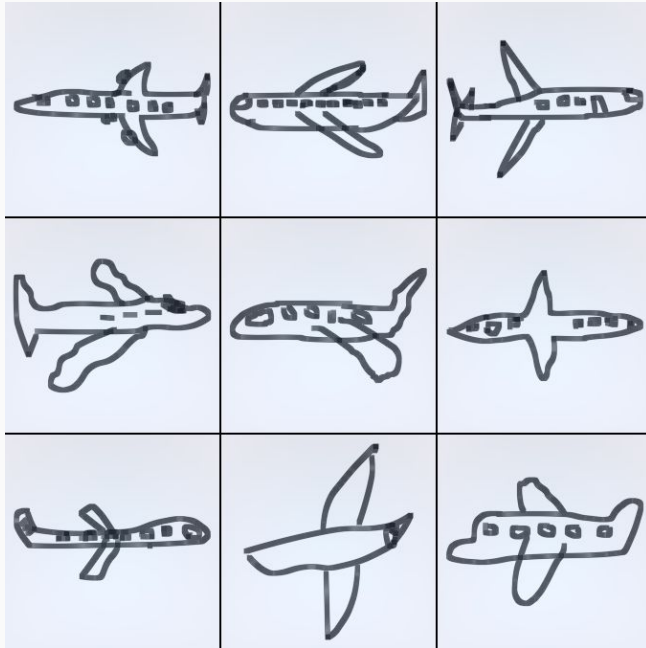


Deconvolution of the activated patch

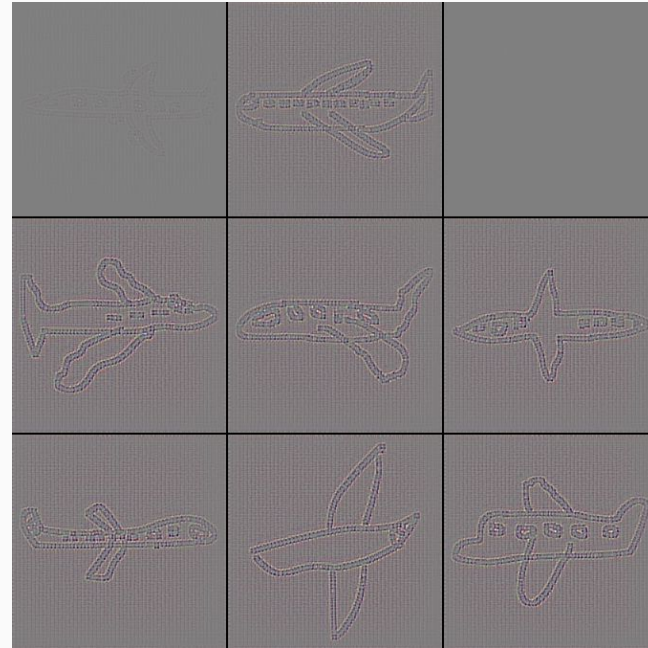
- 224 x 224 Input Patch for 352nd filter in Conv5\_4 layer
- Maximum Activation: 575.805237
- Filter sensitive to animals like dog, sheep, pig.
- Classes with performance: dog 50.0 dog 50.0 sheep 92.86  
dog 50.0 dog 50.0 dog 50.0 pig 85.71 dog 50.0 dog 50.0



# Softmax layer



### Input Patch for Maximum Activation



## Deconvolution of the activated patch











- 227 x 227 Input Patch (complete image) for 1st filter in softmax layer.
- Maximum Activation: 1
- Filter sensitive to the airplane class. (which is also 1st class in target attribute)
- Classes with performance: airplane 100.0 airplane 100.0 airplane 100.0 airplane 100.0  
airplane 100.0 airplane 100.0 airplane 100.0 flying bird 71.43 airplane 100.0

# Summary Plot (VGG-19 with Imagenet Mean)











## Top 5 Classes

## Bottom 5 Classes











Conv1\_1

airplane  100.0	angel  71.43	owl  64.29	church  100.0	crown  71.43	baseball bat  100.0	bed  71.43	bowl  92.86	calculator  85.71	comb  92.86
--	---	---	--	--	--	---	--	--	--











Conv1\_2

pineapple  92.86	zebra  100.0	bicycle  100.0	airplane  100.0	foot  71.43	ant  92.86	baseball bat  100.0	bread  57.14	cactus  85.71	calculator  85.71
---	---	---	--	---	---	--	---	--	--











Conv2\_1

grapes  100.0	zebra  100.0	hedgehog  100.0	airplane  100.0	bicycle  100.0	axe  78.57	baseball bat  100.0	bowl  92.86	chair  92.86	cigarette  85.71
--	---	--	--	--	---	--	--	---	---











Conv2\_2

grapes  100.0	airplane  100.0	zebra  100.0	dragon  64.29	hedgehog  100.0	apple  100.0	bathtub  78.57	chair  92.86	computer-mouse  64.29	ear  100.0
--	--	---	--	---	---	---	---	--	---











Conv3\_1

grapes  100.0	pizza  100.0	hedgehog  100.0	tennis-racket  85.71	zebra  100.0	hat  92.86	tooth  85.71	computer-mouse  64.29	mouth  92.86	trousers  100.0
--	---	--	---	--	---	---	--	---	--











Conv3\_2

grapes  100.0	pizza  100.0	tennis-racket  85.71	pineapple  92.86	radio  85.71	hat  92.86	pear  92.86	spoon  100.0	bowl  92.86	hand  92.86
--	---	---	---	--	---	--	---	--	--

Conv3\_3

tennis-racket  85.71	car (sedan)  100.0	pizza  100.0	grapes  100.0	sea turtle  78.57	bowl  92.86	ear  100.0	envelope  92.86	mouth  92.86	nose  100.0
---	---	---	--	---	--	---	--	---	--

Conv3\_4

tennis-racket  85.71	pineapple  92.86	hedgehog  100.0	car (sedan)  100.0	zebra  100.0	baseball bat  100.0	bowl  92.86	ear  100.0	envelope  92.86	hand  92.86
--	--	---	--	---	---	---	--	---	---



## Top 5 Classes

## Bottom 5 Classes

Conv4\_1

Conv4\_2

Conv4\_3

Conv4\_4

Conv5\_1

Conv5\_2

Conv5\_3

Conv5\_4

FC

Softmax

tennis-racket 85.71	radio 85.71	car (sedan) 100.0	pizza 100.0	hedgehog 100.0	bowl 92.86	computer-mouse 64.29	envelope 92.86	ear 100.0	hat 92.86
radio 85.71	pizza 100.0	car (sedan) 100.0	tennis-racket 85.71	hedgehog 100.0	computer-mouse 64.29	ear 100.0	tablelamp 100.0	envelope 92.86	candle 100.0
pizza 100.0	radio 85.71	car (sedan) 100.0	octopus 92.86	camera 85.71	bowl 92.86	ear 100.0	t-shirt 100.0	hammer 92.86	apple 100.0
radio 85.71	santa claus 85.71	pizza 100.0	tractor 92.86	tennis-racket 85.71	bowl 92.86	spoon 100.0	ear 100.0	cigarette 85.71	baseball bat 100.0
tractor 92.86	pizza 100.0	pineapple 92.86	radio 85.71	tennis-racket 85.71	nose 100.0	ear 100.0	tooth 85.71	bowl 92.86	cigarette 85.71
wrist-watch 85.71	pizza 100.0	pineapple 92.86	tennis-racket 85.71	radio 85.71	hat 92.86	t-shirt 100.0	mouth 85.71	nose 100.0	ear 100.0
wrist-watch 85.71	pizza 100.0	penguin 92.86	pineapple 92.86	human-skeleton 85.71	nose 100.0	stapler 100.0	eyeglasses 92.86	knife 78.57	envelope 92.86
zebra 100.0	pizza 100.0	sun 92.86	duck 100.0	church 100.0	cigarette 85.71	nose 100.0	cloud 92.86	stapler 100.0	envelope 92.86
zebra 100.0	sun 92.86	face 100.0	church 100.0	penguin 92.86	lion 42.86	dragon 64.29	squirrel 57.14	flying bird 71.43	socks 85.71
pineapple 92.86	person sitting 78.57	violin 85.71	rabbit 85.71	axe 78.57	pig 85.71	strawberry 92.86	guitar 92.86	lion 42.86	person walking 100.0



# Conclusion

- Filters in lower level layers are sensitive to edges, patterns, corners, lines and certain localized geometries present in sketches & in higher level layers are sensitive to certain classes & some specific objects in sketches.
- Deconvolution of the filters in VGG-19 shows more sharp input space activations than filters in the Alexnet.
- This comes in coherence with difference in performance where Alexnet gives 73% accuracy & VGG-19 gives 87% on test dataset.
- Convolutional layers are selective to classes with high stroke density.

# For next week ...

- Extending the usage of Yosinski's Deep Visualization Toolbox to analyze sketch CNNs fine-tuned for sketches ( GoogLeNet, ResNet)
- Visualizing & Understanding the hidden states & cell states in Alexnet + GRU Architecture to know which neurons learn the most to give the correct category label.

# Thank You

