Session 1

Due Nov 19, 2019 by 5:30am **Points** 500 **Submitting** a website url **Available** until Nov 19, 2019 at 5:30am

This assignment was locked Nov 19, 2019 at 5:30am.

2019

EIP 4

TSAI

PHASE I

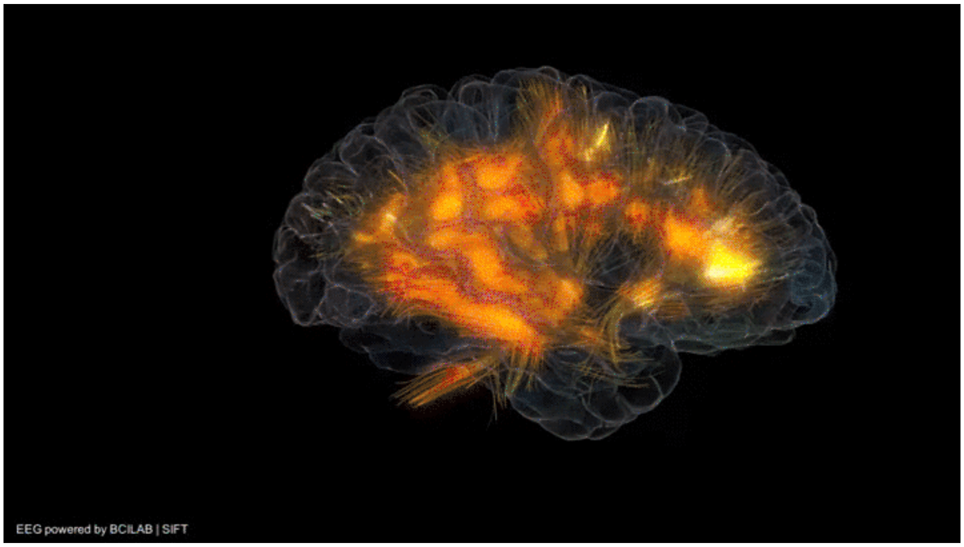
Preface:

1. We received 900+ registrations for EIP4. We'd decided to keep the participants limited to 300.
2. The professional vs student ratio is 89%.
3. 1. Phase 1:
        1. Background and basics of Convolutional Neural Networks
        2. Neural Network Architectures and Receptive Field
        3. Modern DNN Architectures, Batch Normalization and other regularization techniques
        4. Advanced Convolutions and Data Augmentation Techniques
        5. Super Convergence, or how to train CIFAR10 network in under 40 seconds to 94% accuracy
    2. Phase 2:
        1. Yolo V2 Architecture
        2. Mask RCNN Architecture
        3. Training an Object Detector from Scratch including data augmentation
        4. Human Pose Estimation and Object Segmentation
        5. Generative Adversarial Networks and Variational Autoencoders
    3. Phase 3:
        1. RNN & LSTM
        2. Reinforcement Learning Basics
        3. OpenGym & Q-Learning
        4. Deep Deterministic Policy Gradients
        5. Twin Delayed DDPG
4. All phases would end with a project-assignment, clearing which you'd be eligible for the next phase as well as certification.
5. By default, only those in the top-50 percentile would move forward.
6. For all the phases we would use Google Colab, so please make sure that you have made your personal account.
7. Each session is 2 hours and we have 15 sessions in total.
8. 2 hours per session is really too less to become a master:
    1. our aim is to get you started, covering the fundamentals.
    2. unless you spend at least 3-4 hours on your own, reading **only** the material reference, and doing assignments/quizzes, this program would fail for you
    3. quizzes are designed to test your understanding of the concepts, so please attempt the time-bound quizzes only once you feel confident
    4. assignments are designed to extend your understanding and put things into practice
9. Our relationship with you is that of between a "Driving Instructor" and a "New Driver". Our focus would be to help you drive the car and not go very deep into the concepts like what is an alternator or at what compression diesel engine works. 5 sessions are too little to do so
10. The content covered is really state-of-art and updated. **If you refer to online blogs/sources** it is given that you'll get lost. Too many sources online are covering concepts that are now obsolete. Unless we refer a source, please do not scout.
11. The content shared with you through this LMS would really take all your time as we proceed.

LET'S GET STARTED

Building the intuition!

Our Brain



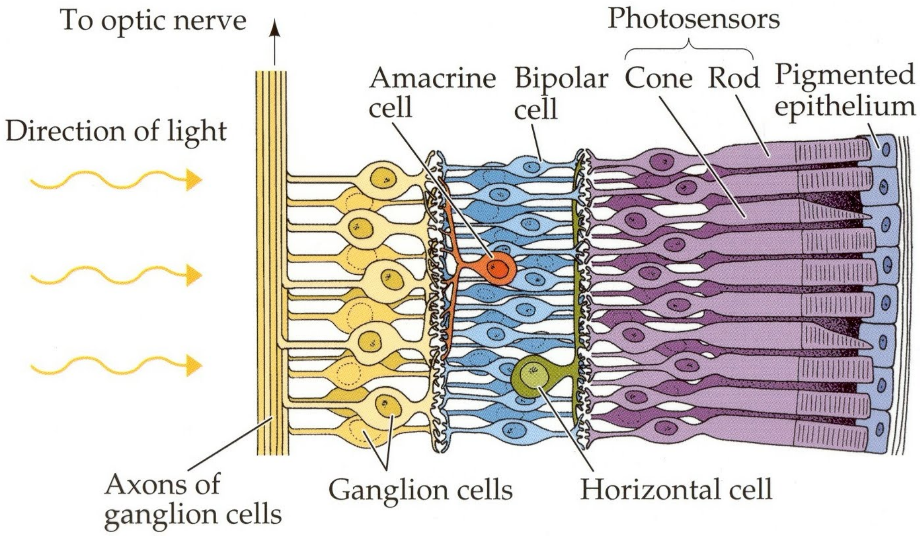
An ultra-dense connected network of flowing information!

[Today we can read brains using neural networks now!](https://www.biorxiv.org/content/10.1101/787101v2.full) [(https://www.biorxiv.org/content/10.1101/787101v2.full)](https://www.biorxiv.org/content/10.1101/787101v2.full)

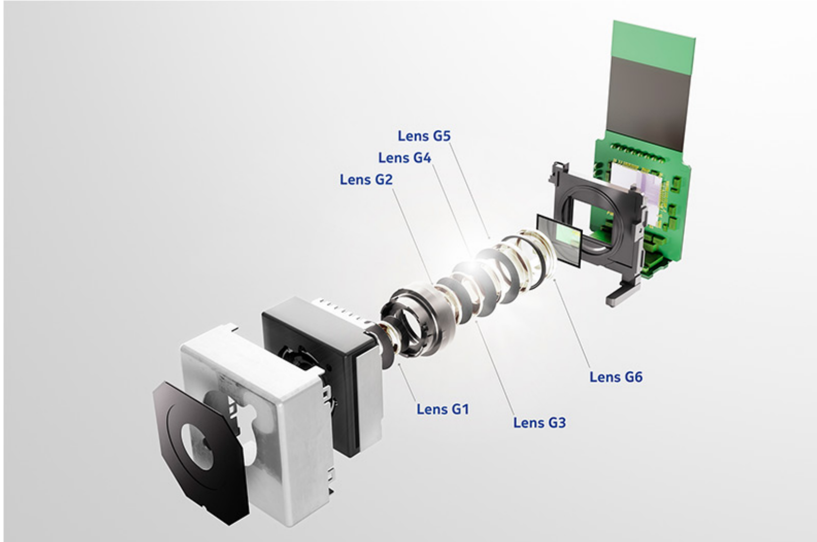
EYES

EIP is going to focus on vision, and concepts we'd learn here can be implemented on any kind of data.

Everything w.r.t vision starts with our eyes. Eye is our vision sensor and similar to our eyes, we have the camera sensor, which reacts to the light based on the amplitude of the wavelength falling over it.



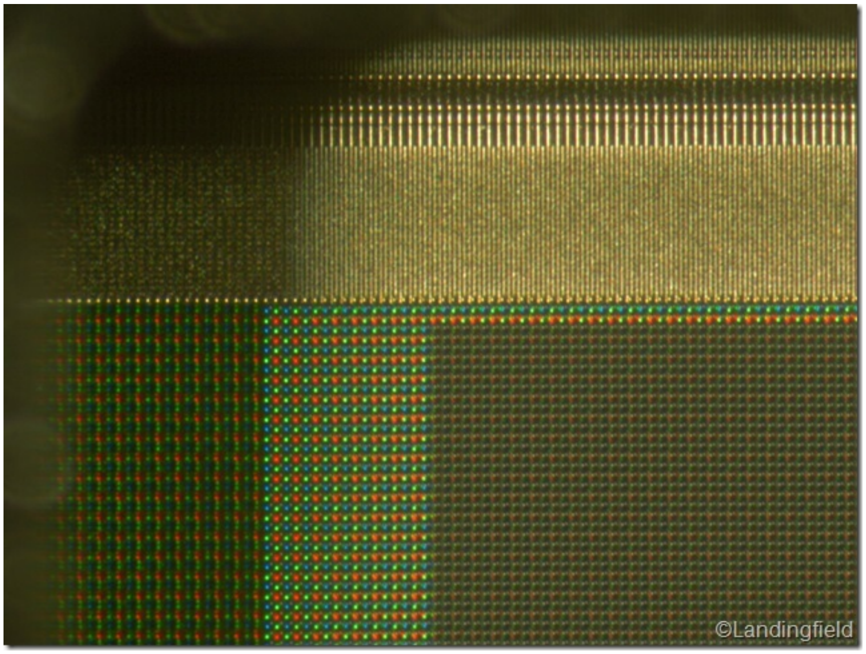
Here is how our physical camera systems look now:



This is how our retina would look under the electron microscope:

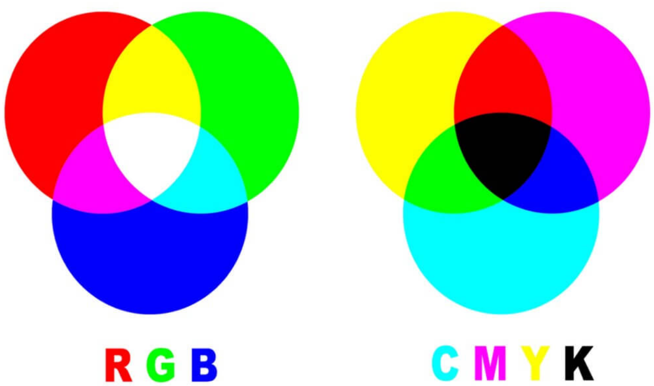
this is where what we see gets converted into rgb or bnw! As you are aware, we have 3 channels (they are not purely RGB though)

and this is how our camera sensor would look under the electron-microscope:



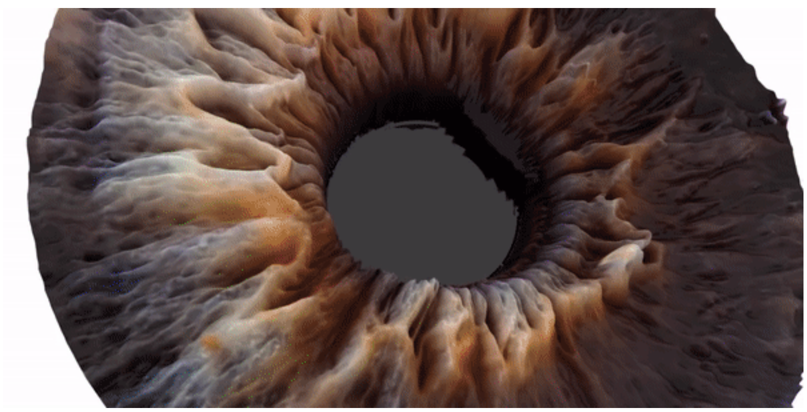
CHANNELS

We need to start to think about channels in a different manner now. You are aware of RGB, and probably CMYK as well.

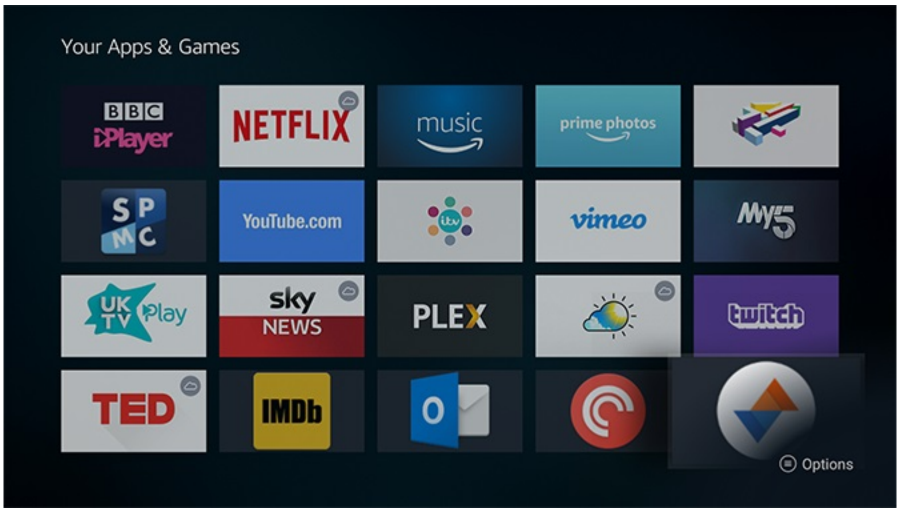


Look at this painting below, and try to see only one color at a time. That is how we need to think about the channels from now on!

Imagine that instead of breaking images into RGB, we divide it into 4 colors. (That's what we do in the newspapers already)



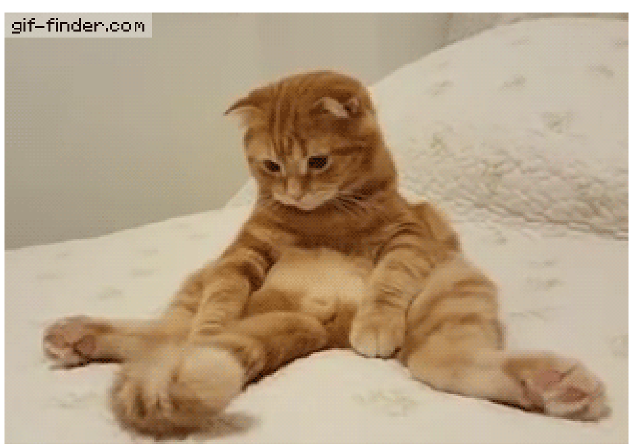
The connection to remember here is CHANNELS



THE BORED CAT EXPERIMENT

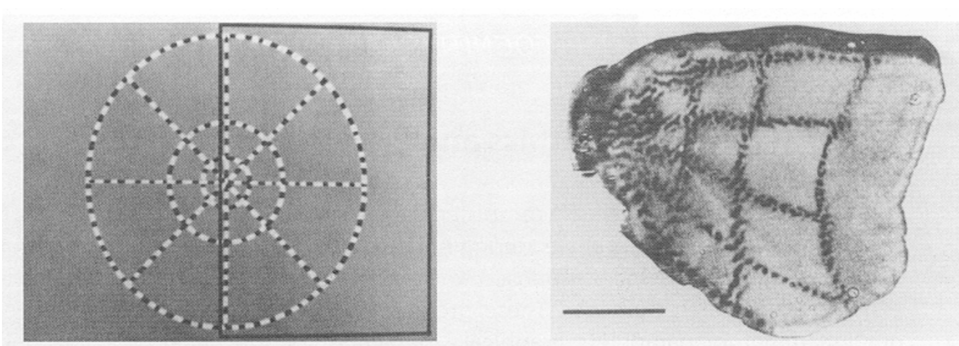
We need 2 cats for our experiments:

CAT1



CAT 2

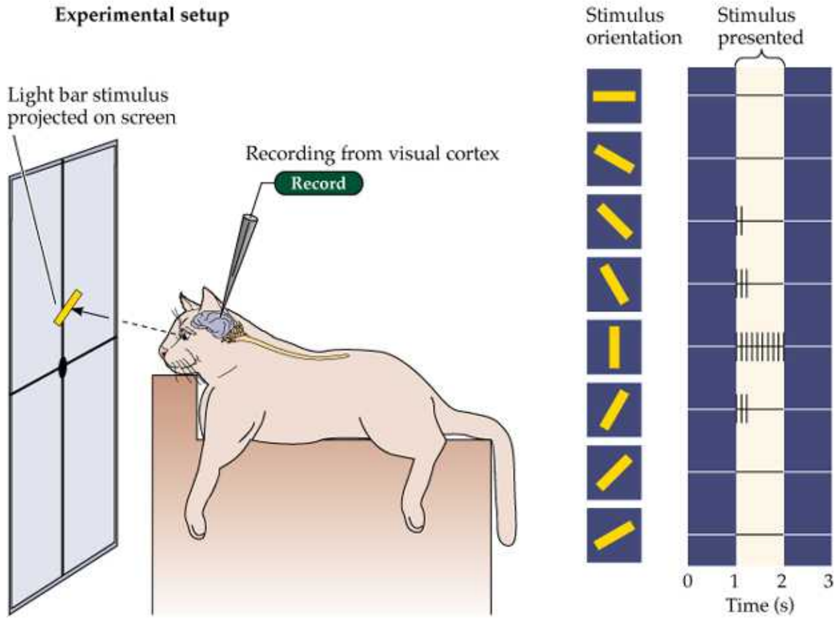
CAT1 Experiment Results:



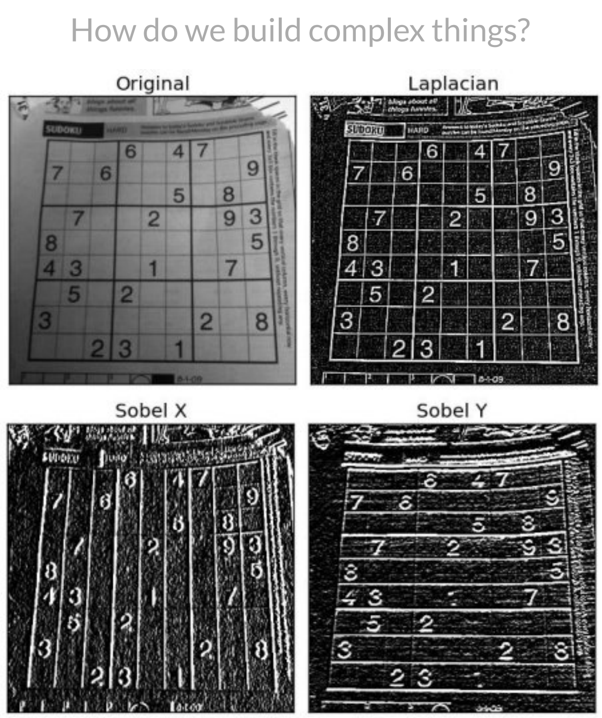
(a) (b)

When we see the image on the left, right gets "printed" on our brain!

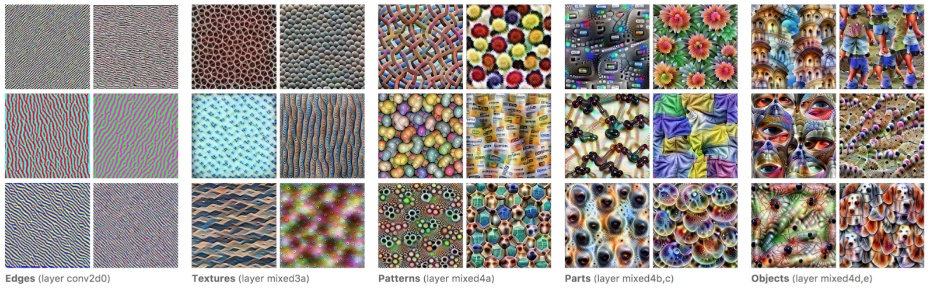
CAT#2 Experiment Results



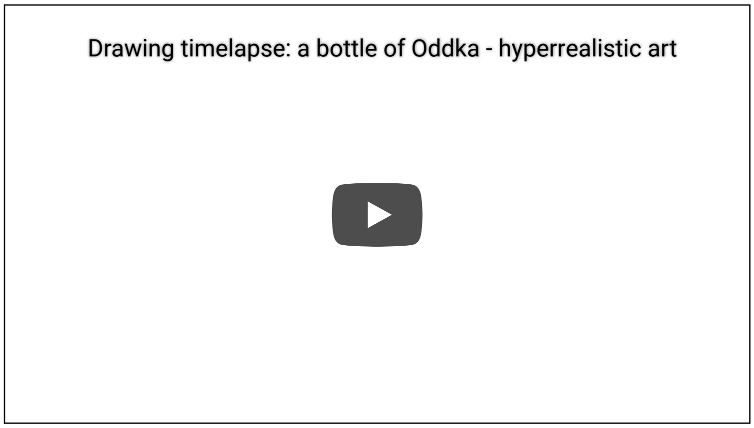
Our brain has several edge detectors (along with many other things)!



We start very simple, but then we can mix those simple building blocks to something very beautiful and detailed!



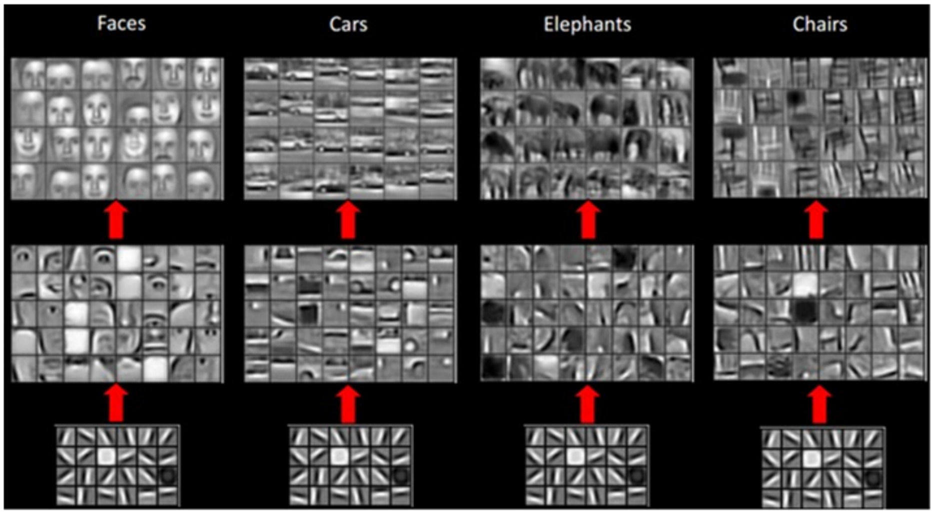
Let's look at this time lapse to appreciate how simple strokes can make something really beautiful!



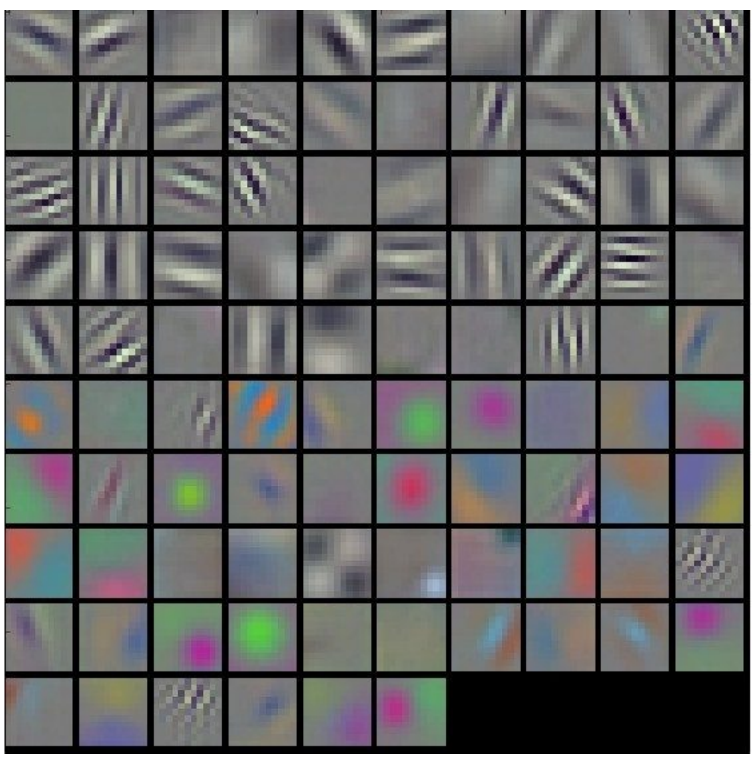
The connection to remember that very complex things can be built from simple strokes like we can make stories from only 26 alphabets.

NEURAL COMPUTING

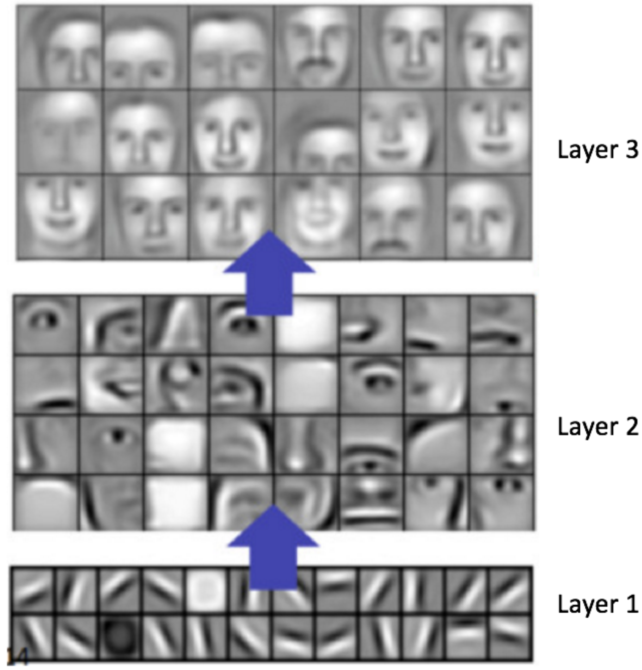
We need feature extraction methods and then combining methods



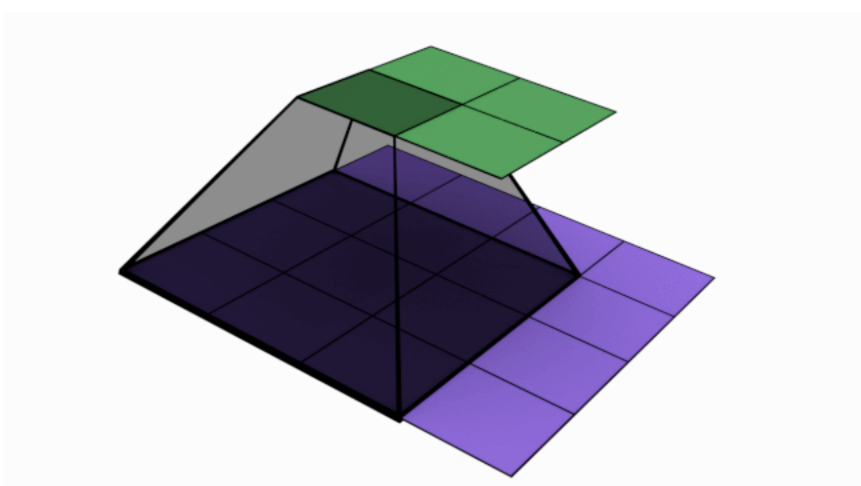
CONVOLUTIONS - FEATURES OR ALPHABETS



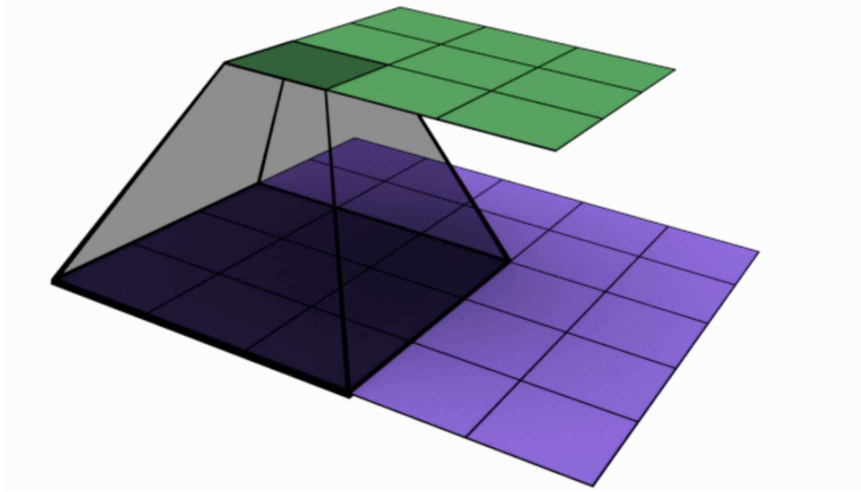
Can you image this happening now?



CORE CONCEPT - CONVOLUTION



3x3 on 4x4

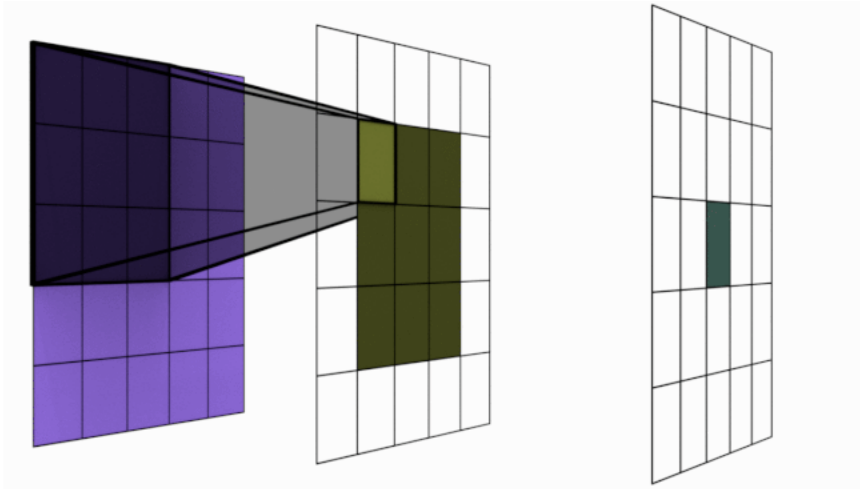


3x3 on 5x5

Let's build a network for a 400x400 image! How many layers would we need?

Why do we always use 3x3?

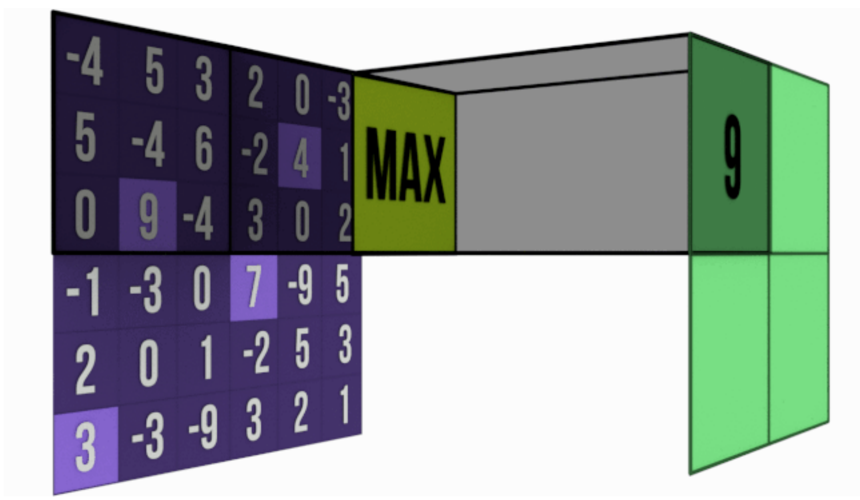
THE MOST IMPORTANT IMAGE YOU'LL EVER SEE!



Receptive Field!

Getting back to our 200 layers. What should we do to reduce the number of layers we have?

MAXPOOLING

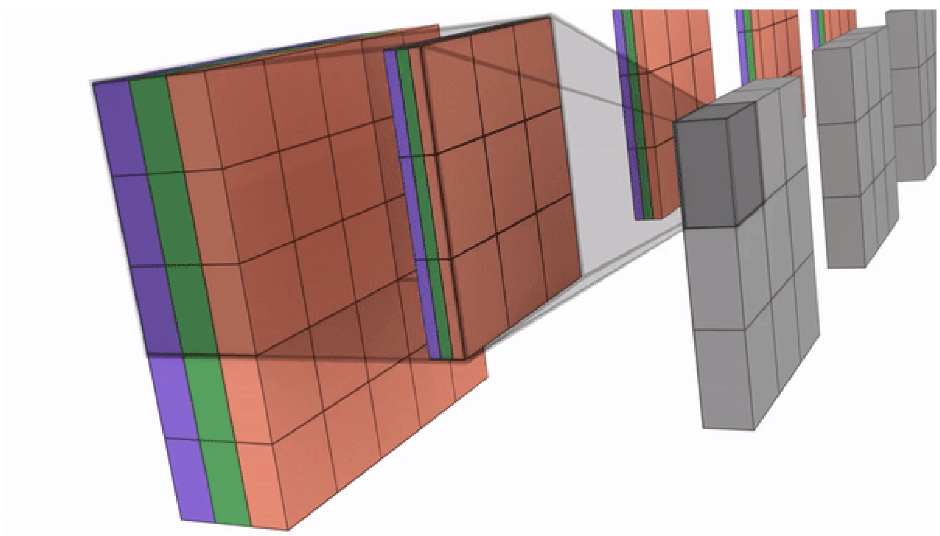


LET'S BUILD OUR NETWORK AGAIN

400 > 398 > 396 > 394 > 392 | Max Pooling
196 > 194 > 192 > 190 > 188 | Max Pooling

96 > 94 > 92 > 90 > 88 | Max Pooling
44 > 42 > 40 > 38 > 36 | Max Pooling
18 > 16 > 14 > 12 > 10...

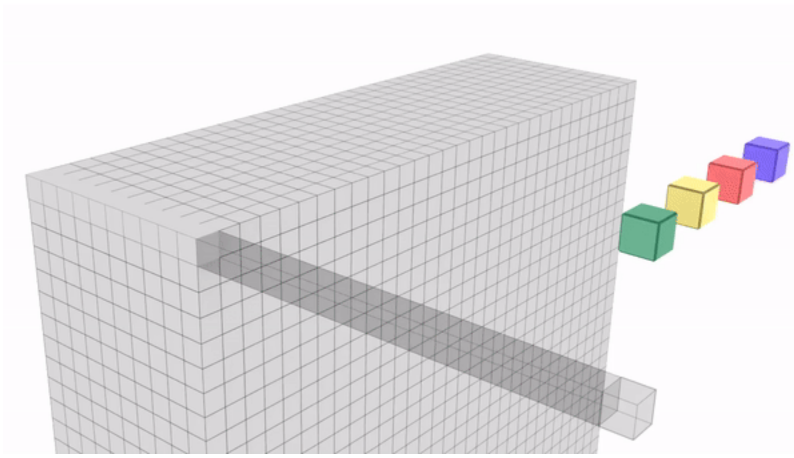
3x3 IS A LIE! HOW CONVOLUTIONS ACTUALLY WORK?



THE Z-CHANNEL DIMENSION REDUCATIONALITY

Let's build our network again!

WELCOME TO THE AMAZING CONVOLUTION OPERATOR 1x1!



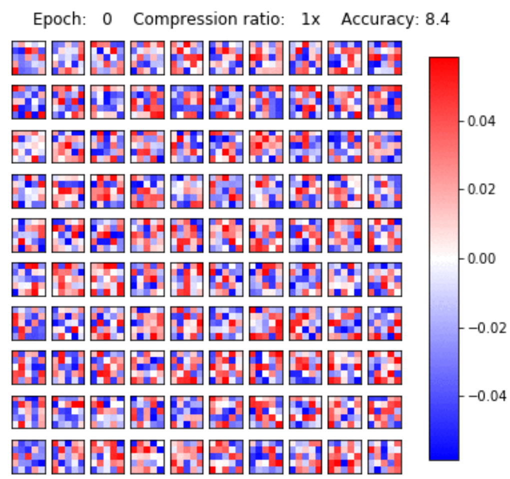
Let's build our network again!

SAMPLE NETWORK

```
input(28x28x1)
Conv2D(32, 3, 3, act=ReLU) > 26 x 26 x32
MaxPooling(2) > 13 x 13 x 32
```

```
Conv2D(10, 1, 1, act=ReLU) > 13 x 13 x 10
Conv2D(10, 13, 13, act=None) 1 x 1 x 10
SoftMax()
Compile/Fit/Optimize
```

WHAT ACTUALLY HAPPENS WHEN WE TRAIN OUR NETWORK?



CODE (https://colab.research.google.com/drive/1gZYwZdkgX

Assignment:

1. Visit this link (https://colab.research.google.com/drive/1gZYwZdkgXBJRr624SqWJ9f452BmKNkNT#scrollTo=7eRM0QWN83PV). Take sometime to understand the code.
2. Change the model in such a way that after executing the code below, your accuracy print out is more than 99.0
   1. score = model.evaluate(X_text, Y_text, verbose=0)
   2. print(score)
3. Once you are done, upload the code to your GitHub Repo. (Use this link to learn how to use GitHub).
4. Add a readme file, and write the result of your **print(score)** command in this readme.
5. Write your own definitions for:
   1. Convolution
   2. Filters/Kernels

    3. Epochs
    4. 1x1 Convolution
    5. 3x3 Convolution
    6. Feature Maps
    7. Activation Function
    8. Receptive Field
6. Upload the link to your GitHub Repo to the LMS.
7. The deadline is 1 hour before your next session starts.
8. You MUST use your own words while writing any of the above, if we find plagiarism, 0 marks would be awarded for that article.
9. The assignment is worth 500 pts.

Session Videos:

Online students should prefer this video:

Friday Video