Convolutional Neural Networks

Tyson Beach, Manas Tiwari

Questions

1. What is a filter?
2. What is the purpose of the activation function?
3. What is a big difference between CNN and FCN?

Tyson

From Knoxville

Played D2 baseball at Tusculum

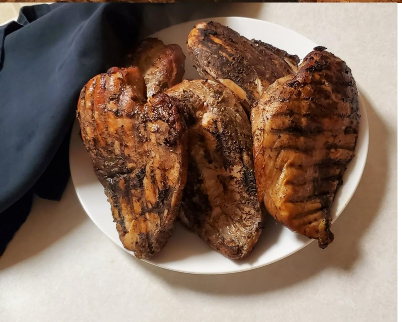
Avid extreme sporter!

Doing masters in CS

No focus/research yet :(

Really Really Good Wakeboarding Guy / VolWake

I like stuff



Manas Tiwari

Present:

- MS in Computer Science (Fall 22)
- Minor in Cybersecurity
- GTA for CS302 (Dr. Emrich)

History:

- Born - New Delhi, **INDIA** (2000)
- Bachelor’s in Information Science from Bangalore, India

Interests:

- Cricket, Swimming, Running
- Geopolitics

Festival, places and more

Ganesh Chaturthi

Holi

Panama City, FL

Smoky Mountain

Indian Loved Foods

First Love - Tea (Chai) ❤️☕
Cost: Rs 7 ~ $0.086

Pav Bhaji

Chole Bhature

Chicken Masala

Overview

CNN Intro

What is it used for?

Convolutional Layers

Pooling Layers

Flattening/Fully Connected Layers

Fully Convolutional Neural Networks

Modern Day Examples

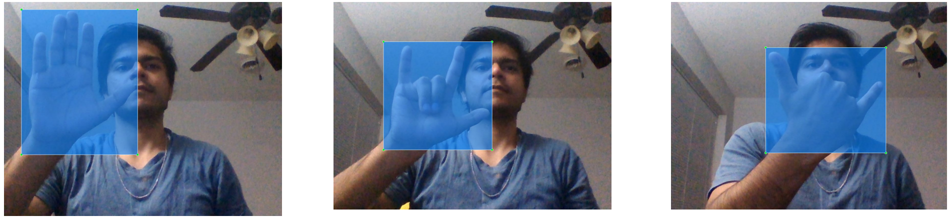
Live Demo

Intro - Convolutional Neural Network (CNN)

- Most popular neural network for analyzing images, and can also be used for data analysis and classification problems as well
- Most generally we can think of CNN as a **artificial neural network** that has a some kind of specialization to detect patterns
- This pattern detection is what makes it so useful for **image analysis**

If CNN is a form of artificial neural network (ANN) then what differentiates it from a standard multi-layer perceptron (MLP) ? - so CNN has a hidden layers called convolutional layers and this is what make it different from other neural networks. The convolutional layer receives input, transforms into someway and then outputs it to the next layer.

- Each convolutional layer has a number of filters and filters are actually what detects the patterns. For example, one filter might be used to detect edges in the image, other filter might be use detect shape in the image.
- So, this basic filters are usually at the start of the neural network and the deeper the network goes the more sophisticated these filters become. (for eg. in later layers the filter will be able to detect specific objects such as eyes, ears. etc.



Labelled Images with different hand gestures

Advantages of CNN’s

- Good at detecting patterns and features in images, videos, and audio signals.
- Robust to translation, rotation, and scaling invariance.
- End-to-end training, no need for manual feature extraction.
- Can handle large amounts of data and achieve high accuracy.

Limitations:

- Computationally expensive to train and require a lot of memory.
- Requires large amounts of labeled data.
- Interpretability is limited, it’s hard to understand what the network has learned.

Convolutional Layer

- Filters
  - Weights
    - Weights act like dails getting tuned just right to learn to classify image.
    - Weights are generally learned.
  - Scan over input to learn/extract features
  - Channels of a filter must be equal to the channels of the input.
  - It is very common for there to be many filters per layer.
- Feature Map
  - The result of convolution.
  - The values of the feature map are put through an activation.
  - Then the bias for each filter is added.
    - Bias is a threshold of how meaningfully activated a weight is.
  - Output yay!

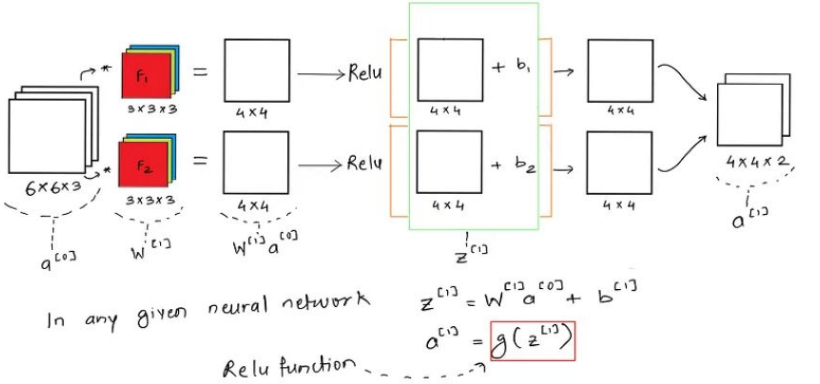
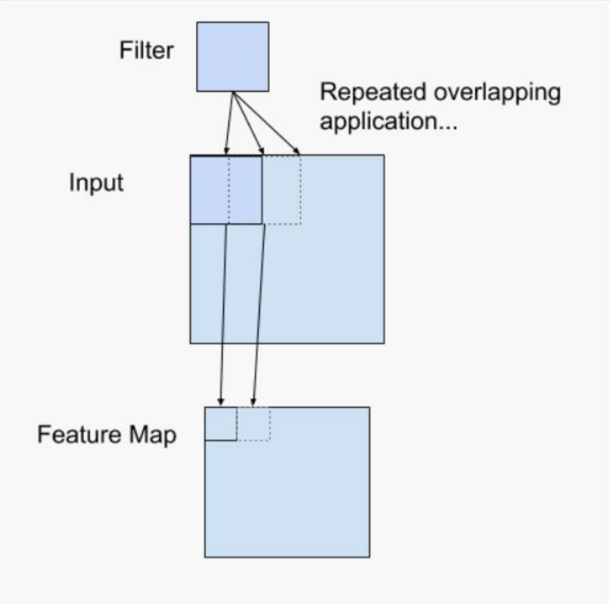


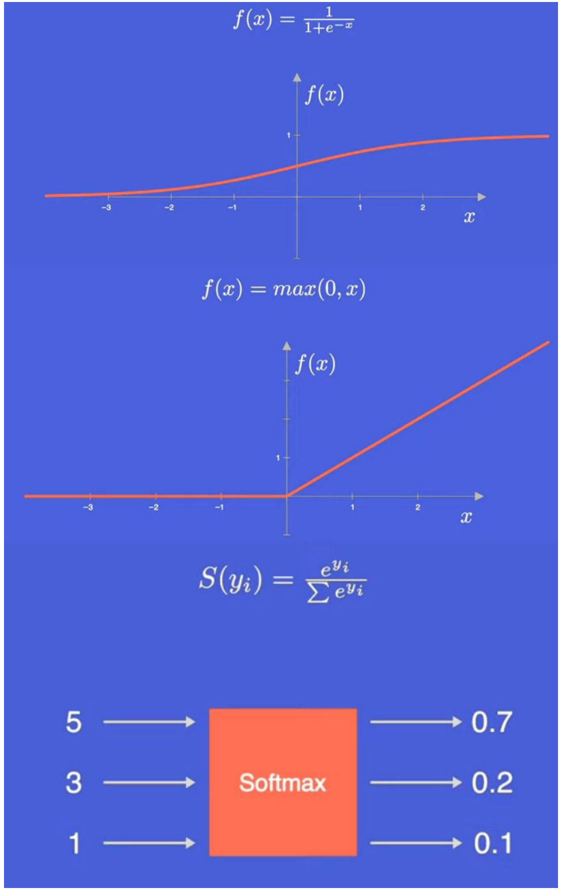
Fig 4. Single layer of convolution network only with Relu activation



Example of a Filter Applied to a Two-Dimensional Input to Create a Feature Map

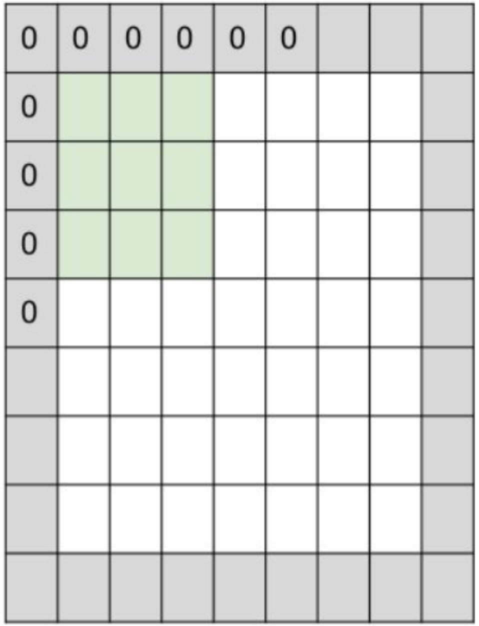
Activation

- Activation functions apply a nonlinear transformation to decide if a neuron should be activated or not (or how much).
- Without activation any neural network would just be a stacked linear regression model incapable of learning complex images.
- Popular choices are Sigmoid, ReLu, and softmax.
    - ReLu is the most popular.
    - Softmax is used most exclusively in the final layer of a CNN for classification.



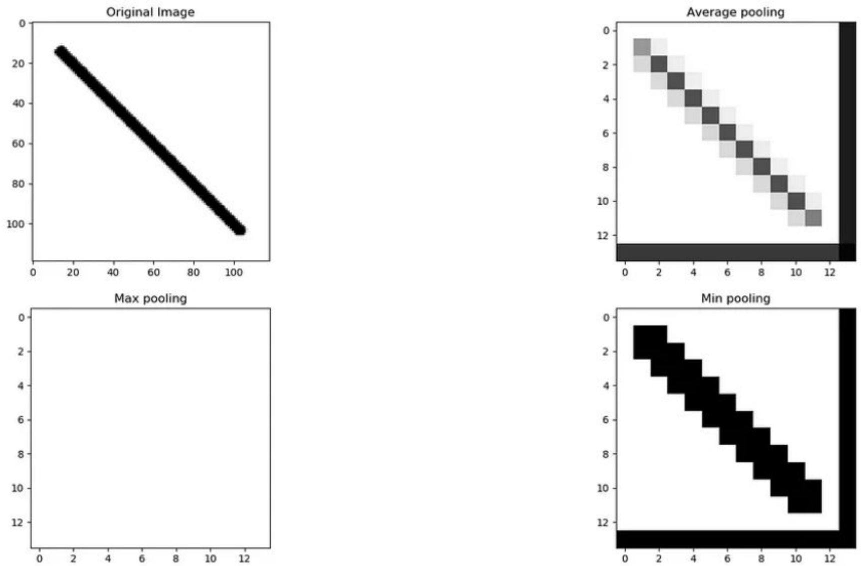
Convolutional Layer Contd.

- Padding
  - Keep info at the borders
  - Allows for deeper network design
  - Note. w/o you do have faster computation
- Stride
  - Cover more ground per iteration
  - Also means we are limiting the overlap between pixels.



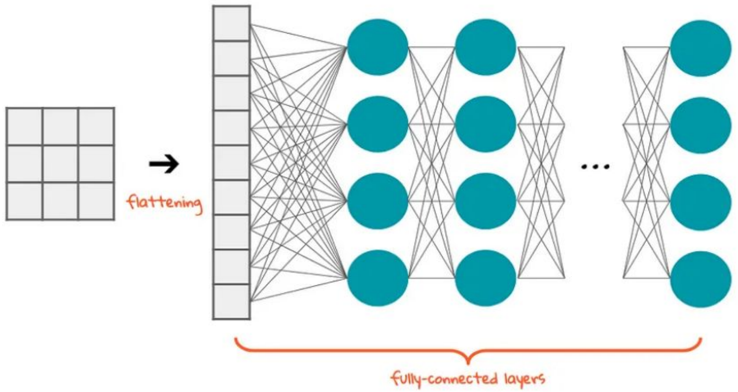
Pooling

- Min pooling
- Average pooling
- Max pooling
- Stride is applicable here as well



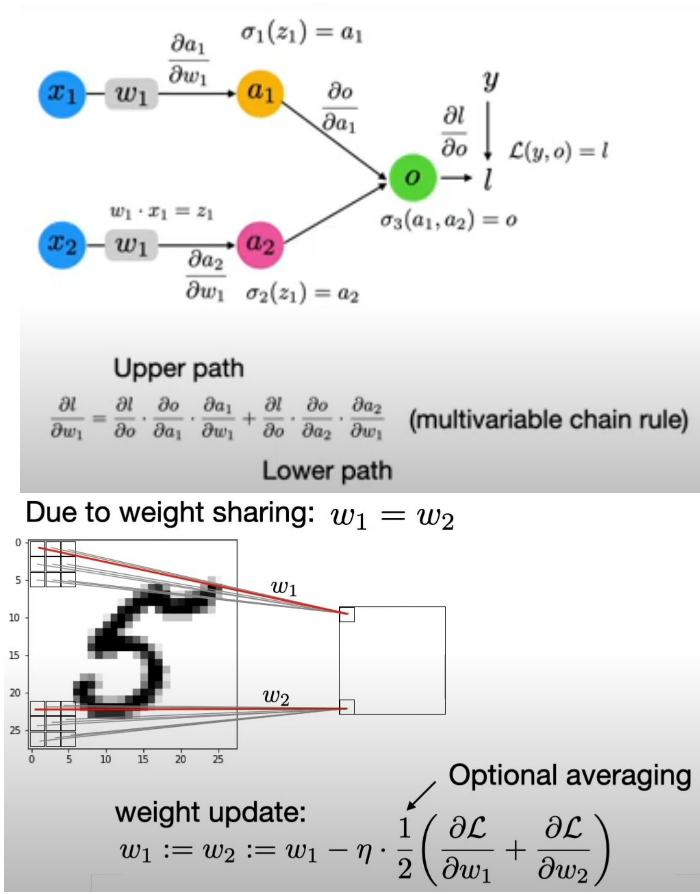
Flattening/Fully Connected Layers

- Simple, we mimic an ANN at the end to categorize/label
- Can think of the layer as a flattened filter.
- In a traditional CNN this is the classification layer.



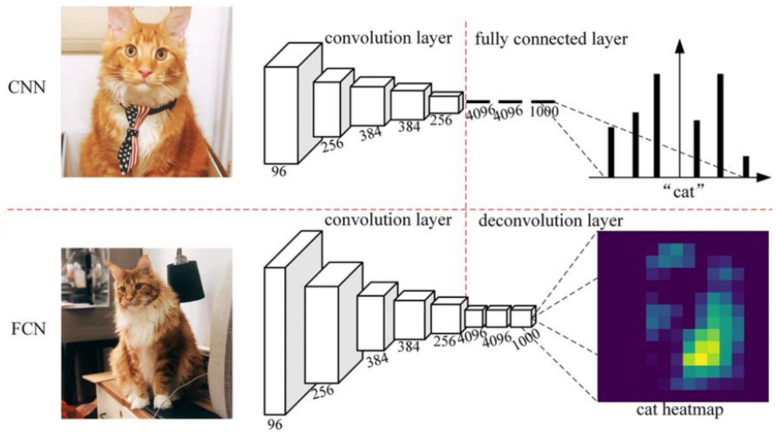
Backpropagation

- Where the ‘learning’ happens
- Get output Loss
- Multivariable Chain Rule w/ weight sharing constraint
- Partial derivatives tell us how much a function would change when we keep all but one of its input variables constant and move a slight nudge in the direction of the one variable that is not fixed.
- Use partial derivative of the loss with respect to the output
    - Partial derivative of the output with respect to the activation.
        - Partial derivative of activation output with respect to the weight
- We will use this chain for every instance the weight is being used.



FCN

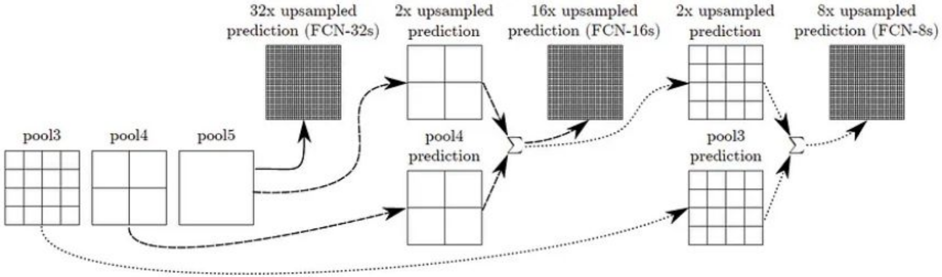
- FCN’s result in a heatmap of the image it processed.
- Upsampling/Deconvolution/Unpooling is done at the end.
- Fusing layers cleans this up! (next slide)



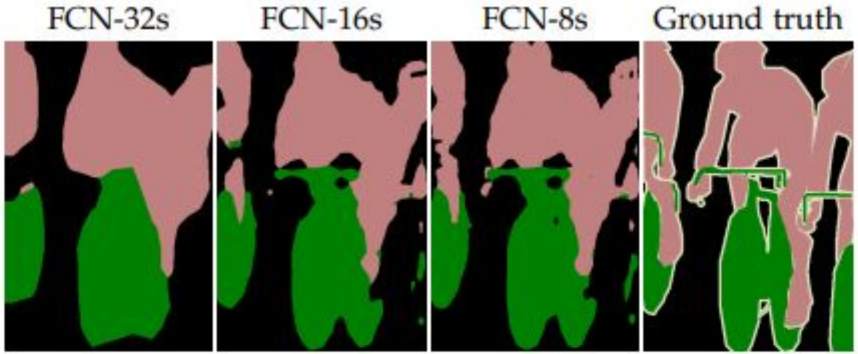
The difference between the CNN and FCN (the transforming of fully connected layers into convolutional layers by an FCN enables a classification net to output a heatmap).

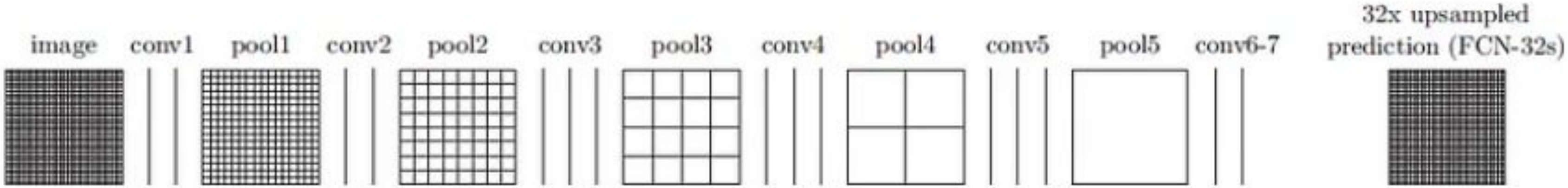
FCN

- Deep features can be obtained when going deeper but spatial information in lost.
- Fusing earlier pooling layers with slightly upsampled predictions can result in really nice images.
- Below is the architecture for a FCN take on VGGNet



Fusing for FCN-16s and FCN-8s



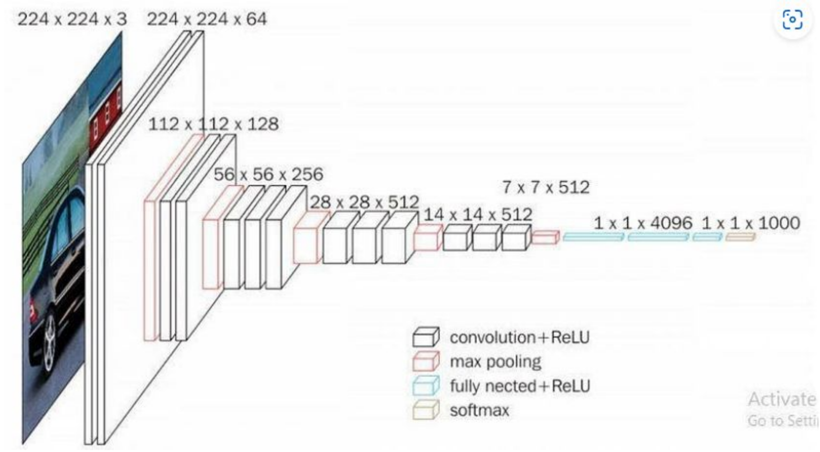


FCN-32s

Modern Day CNN's

VGGnet 16

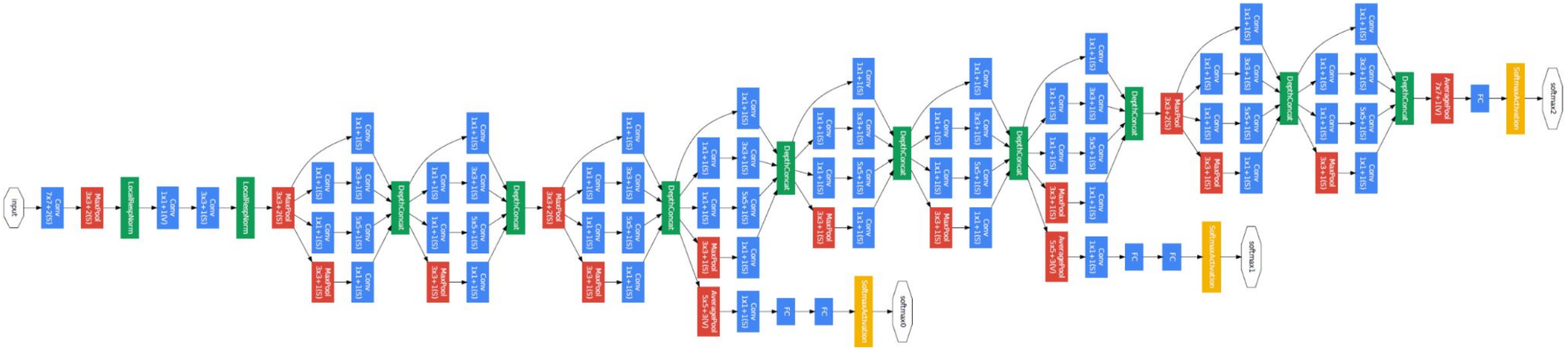
- Classified 1000 images in 1000 categories at up to 92.7% accuracy
- Won the 2014 ImageNet Challenge
- Takes 224x224 RGB image (Also uses ReLu)



Modern Day CNN's

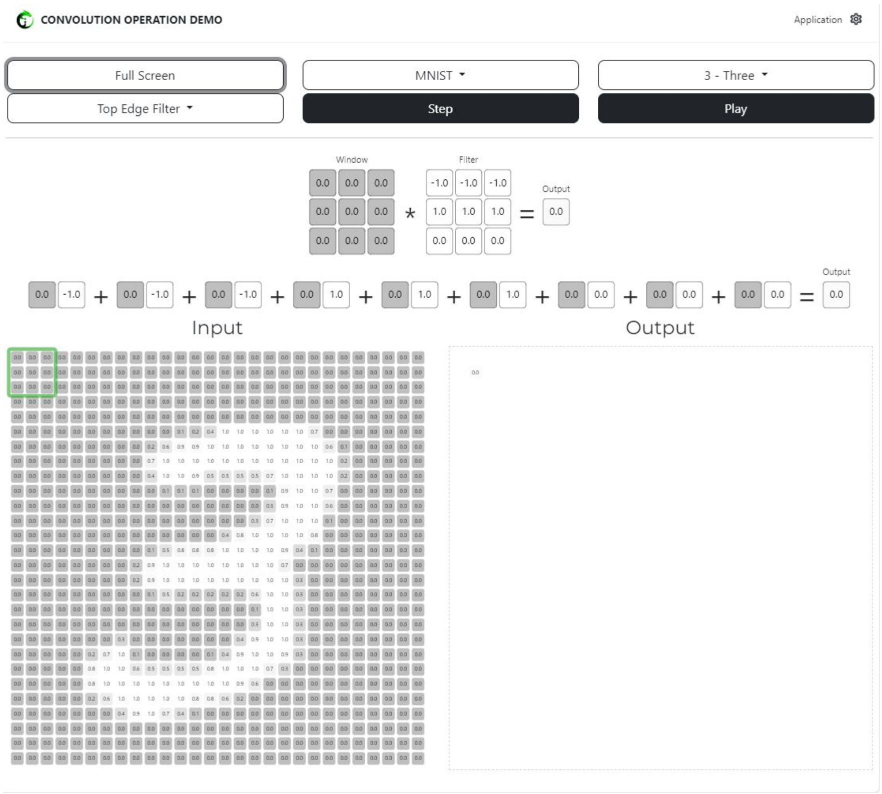
GoogLeNet

- Takes 224x224 RBG images (Uses ReLu for all convolutions)
- Won ILSRVRC 2014 taking 1st place in both classification and detection task. It has top-5 error rate of 6.67% in classification task.



Working - Convolution Operation in Neural Network

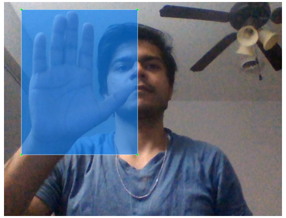
Link: https://deeplizard.com/resource/pavq7noze2



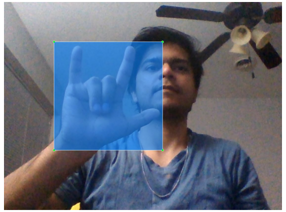
Live Demo

Hand Gesture Recognition using CNN

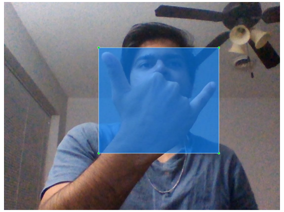
We trained our model with 5 gestures, and labelled images for input to the neural network



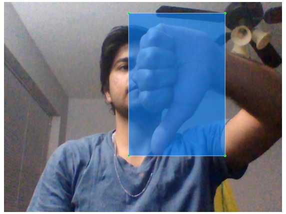
Hello



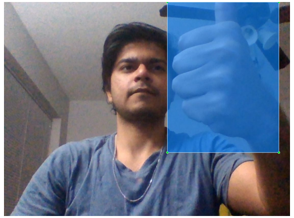
I love you



Thank you



No



Yes

References

- “Activation Functions In Neural Networks Explained | Deep Learning Tutorial” - (Youtube)
- Max Pooling vs minpooling vs average pooling | by Madhushree Basavarajaiah | Medium
- https://www.youtube.com/watch?v=pDXdlXlaCco
- https://tensorflow-object-detection-api-tutorial.readthedocs.io/en/latest/install.html
- conv neural network - Convolutional Layers: To pad or not to pad? - Cross Validated (stackexchange.com)
- https://deeplizard.com/resource/pavq7noze2
- dtybe.png (1007×569) (imgur.com)
- Review: FCN — Fully Convolutional Network (Semantic Segmentation) | by Sik-Ho Tsang | Towards Data Science
- How does Backpropagation work in a CNN? | Medium
- “L13.6 CNN’s & Backpropagation” - Sebastian Raschka (Youtube)
- I Finally Understood Backpropagation: And you can too… | by Kofi Asiedu Brempong | Towards Data Science
- Difference between Local Response Normalization and Batch Normalization | by Aqeel Anwar | Towards Data Science
- “Fully Convolutional Neural Networks for Image Segmentation” - Evan Shelhamer, Jonathan Long, Trevor Darrell

Questions?