Computational Biology Project

Day 3

Welcome to day 3!

A short overview:

- Quick recap of Neural Networks and Deep Learning from Week 1
- Introduction to Convolutional Neural Networks
- UNet and its architecture
- Significance of UNet and its layers
- Brief run-through of day 3’s notebook.

Deep Learning and Neural Networks - recap

- Deep Learning is a powerful set of techniques, approaches, and frameworks for constructing and training deep neural networks effectively.
- A deep neural network is a neural network with many hidden layers which are trained to learn patterns in data.
- Training refers to determining the best set of weights for maximizing a neural network’s accuracy.

Convolutional Neural Networks (CNN)

- A CNN is a neural network whose entries are images that allow us to encode certain properties in the architecture to recognize specific elements in the images.
- CNNs process images as tensors, which are matrices of numbers with additional dimensions.

Convolutional Neural Networks (CNN)

CNNs have two major components:

- Feature Extraction/Learning: A combination of Convolution, Pooling, and Activation layers are responsible for extracting features from an image, such as edges, shades, lines, and curves.

Convolutional Neural Networks (CNN)

CNNs have two major components:

- Feature Extraction/Learning: A combination of Convolution, Pooling, and Activation layers are responsible for extracting features from an image, such as edges, shades, lines, and curves.
- Classification: Fully Connected Layers at the end of the network, along with an appropriate activation function are responsible for generating a prediction/output.

Feature Extraction

- A feature is an individual measurable property or characteristic of something.

Feature Extraction

- A feature is an individual measurable property or characteristic of something.
- Features are the input that we feed to a classification model to output a prediction or classification.

Feature Extraction

- A feature is an individual measurable property or characteristic of something.
- Features are the input that we feed to a classification model to output a prediction or classification.
- Selecting good features that clearly distinguish the object increases the predictive power of the machine learning algorithm in use.

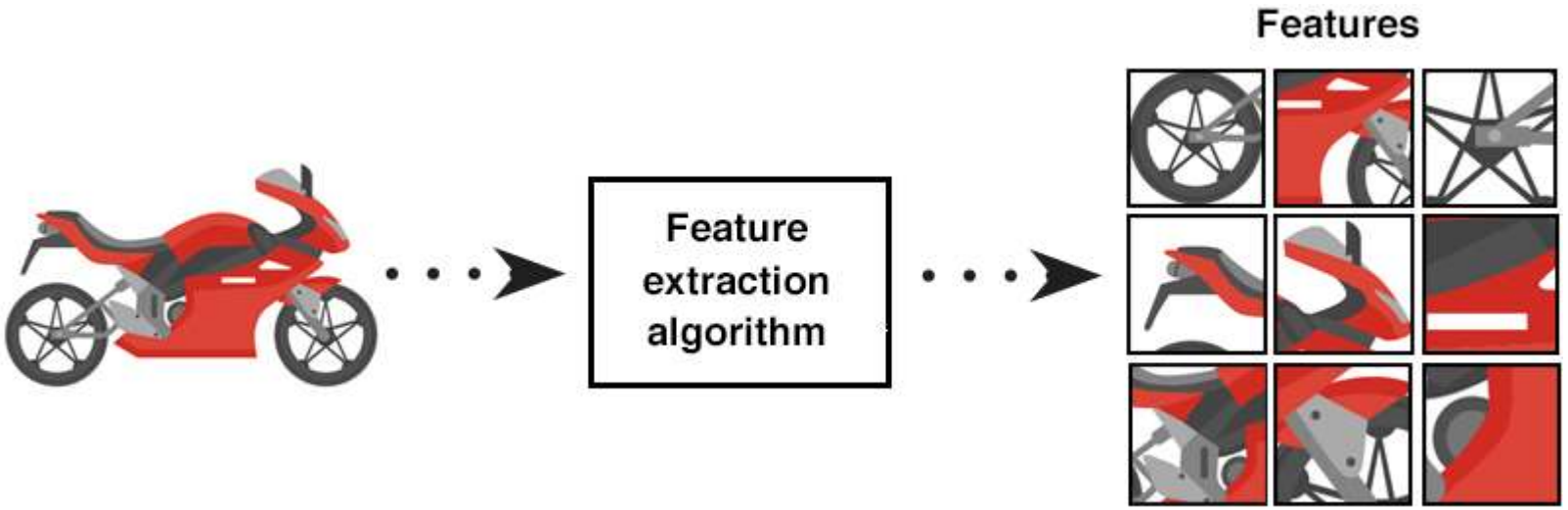
Activity

What are some of the features that can be extracted from an image using a convolutional neural network?

Feature Extraction

- In computer vision, a feature is a measurable piece of data in your image which is unique to that specific object.
- It may be a distinct color in an image or a specific shape such as a line, edge, or an image segment.
- A good feature is used to distinguish objects from one another.

Feature Extraction



Feature Extraction

It is to be noted that the feature should be efficient enough to be able to aid in detecting motorcycles in general and not just that specific one. Therefore, a feature should be repeatable.

Feature after looking at one image

Feature after looking at 1000s of images

What makes a good feature?

- A good feature is one that is invariant to image scaling, rotation and change in illumination.

What makes a good feature?

- A good feature is one that is invariant to image scaling, rotation and change in illumination.
- The extracted feature should be repeatable, precise, and distinctive to the image.

What makes a good feature?

- A good feature is one that is invariant to image scaling, rotation and change in illumination.
- The extracted feature should be repeatable, precise, and distinctive to the image.
- The detected features will be useful in classifying and recognition of images.

Convolutional Neural Networks (CNN)

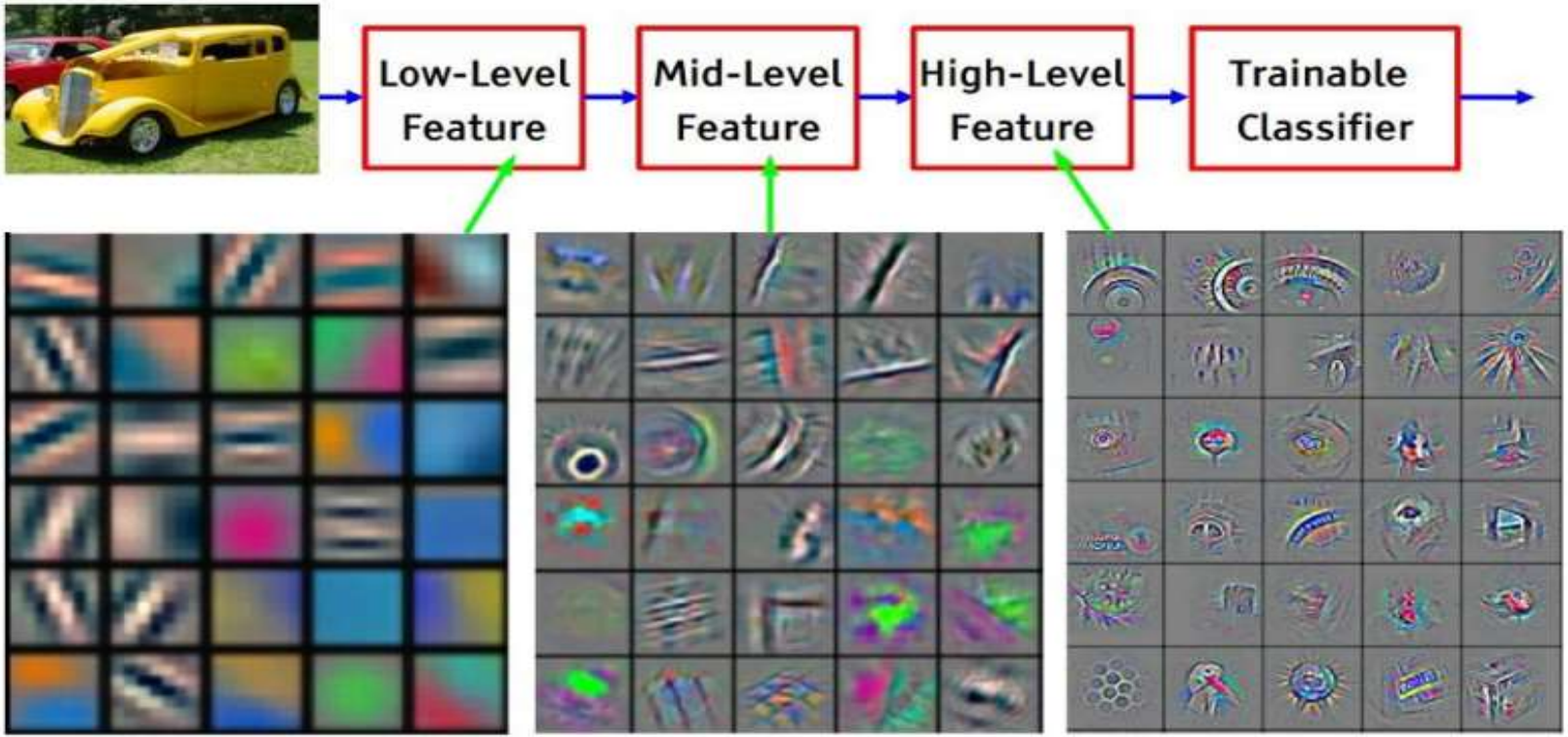
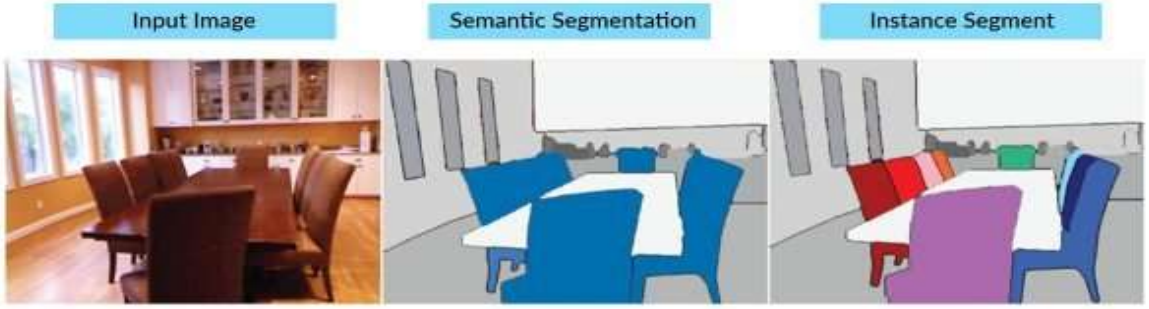
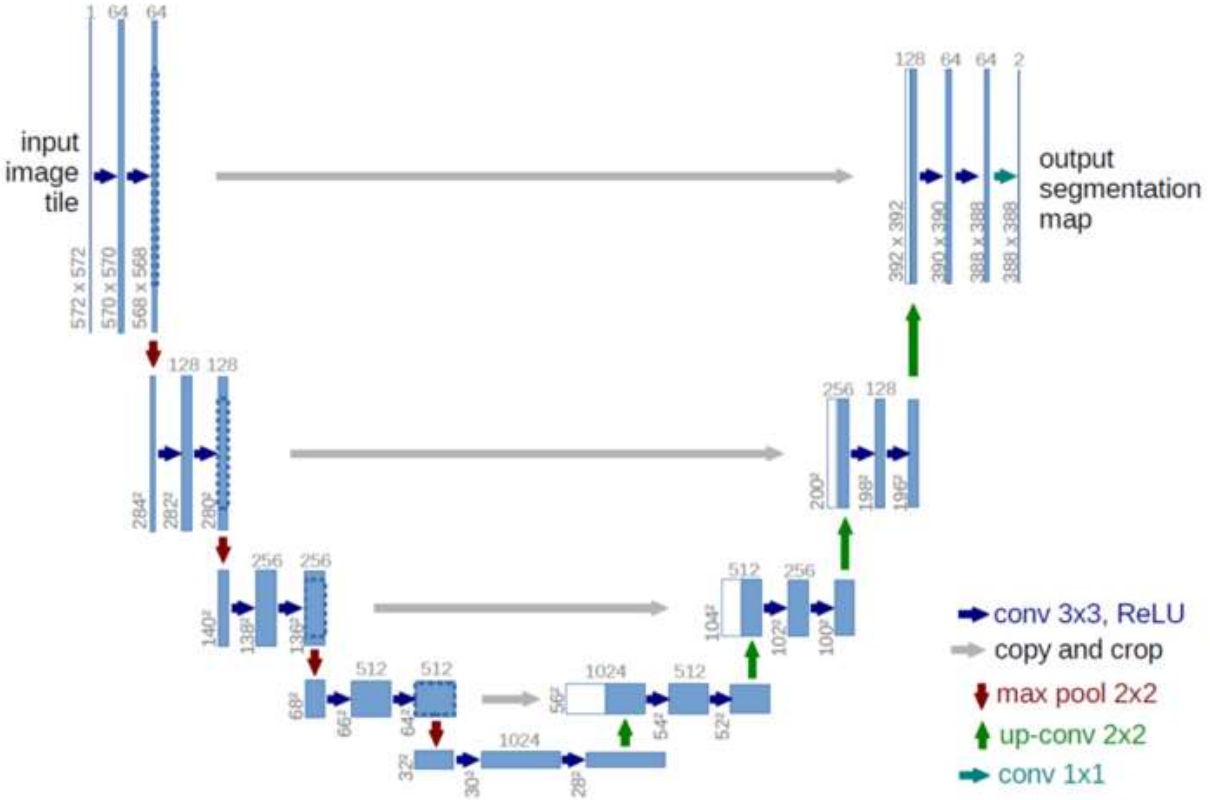


Image Segmentation - recap

- Image segmentation is the process where a network takes an image as input and outputs a pixel-wise mask.
- Algorithms are used to split and group a certain set of pixels together from the image.



UNet

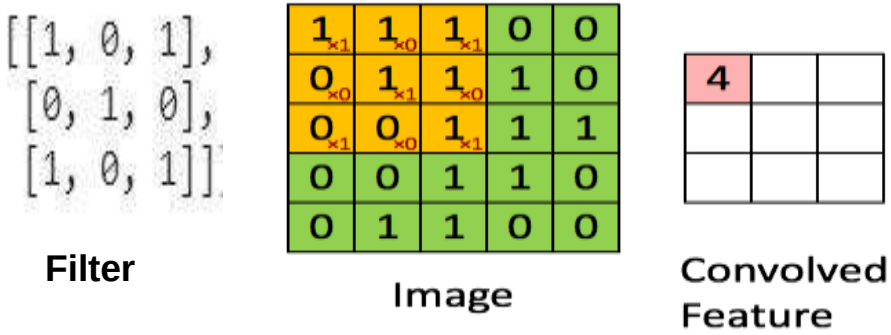

input image tile
572 x 572
570 x 570
568 x 568
output segmentation map
392 x 392
390 x 390
388 x 388
388 x 388
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1

The UNet mainly consists of the following layers:

1. Convolution layer
2. Activation layer
3. Max Pooling layer
4. Batch Normalization
5. Dropout Layer

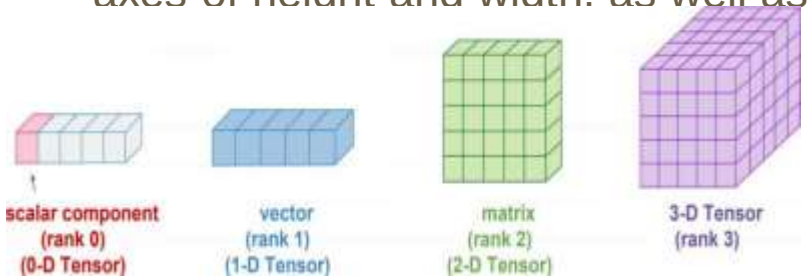
UNet - The Convolution Layer

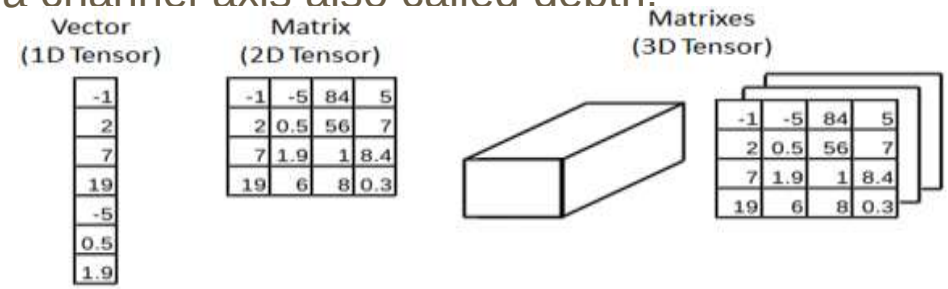
- The UNET architecture uses two 3×3 convolutional layers, each followed by a ReLU activation function at every stage of the encoder and decoder except for the bridge at the bottom.
- A filter is moved across the image to perform convolution operations. The convolution operation is an element-wise matrix multiplication between the filter values and the pixels in the image, and the resultant values are summed.



UNet - The Convolution Operation

- Once the convolution layer has learned a characteristic at a specific point in the image, it can recognize it later in any part of it.
- Convolutional layers can learn spatial hierarchies of patterns by preserving spatial relationships.
- Convolutions layers operate on 3D tensors, called feature maps, with two spatial axes of height and width, as well as a channel axis also called depth.





UNet - Activation Function

- A convolutional layer is usually coupled with an activation function.

UNet - Activation Function

- A convolutional layer is usually coupled with an activation function.
- Activation functions are mathematical equations that determine the output of a neural network by learning the complex patterns in the data.

UNet - Activation Function

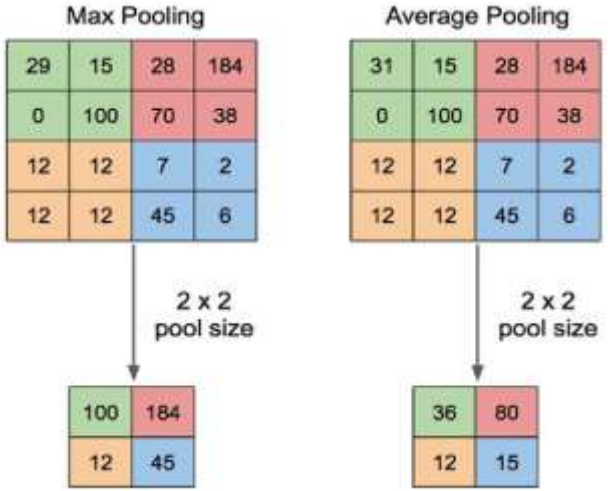
- A convolutional layer is usually coupled with an activation function.
- Activation functions are mathematical equations that determine the output of a neural network by learning the complex patterns in the data.
- There are different kinds of activation functions: Sigmoid, Tanh, ReLU, etc.

UNet - Activation Function

- A convolutional layer is usually coupled with an activation function.
- Activation functions are mathematical equations that determine the output of a neural network by learning the complex patterns in the data.
- There are different kinds of activation functions: Sigmoid, Tanh, ReLU, etc.
- ReLU provides for faster training rate, and induces non-linearity in neural networks.

UNet - The Pooling Layer

- The pooling layer scales down the amount of information the convolutional layer generated for each feature and maintains the essential information.
- Pooling also helps in avoiding overfitting.

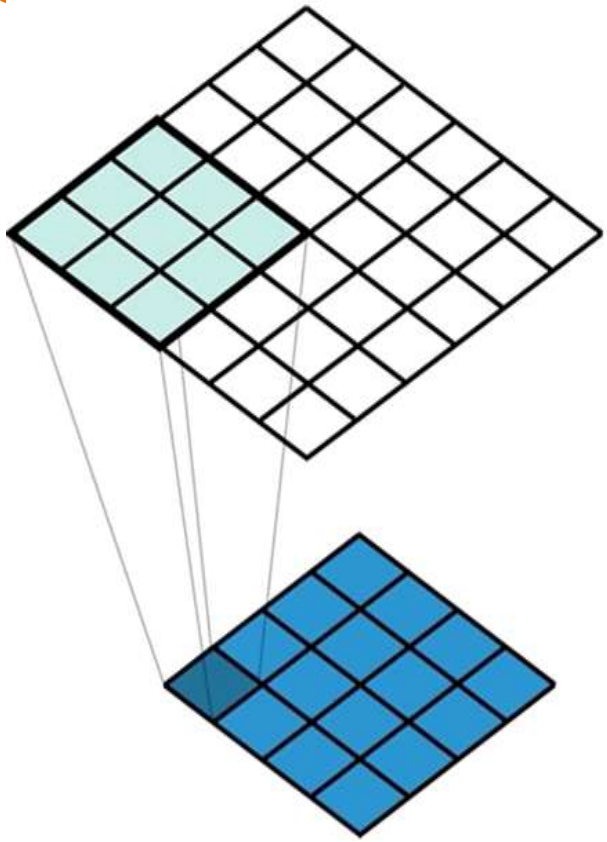


Max Pooling — Selecting the maximum value

Average Pooling — Sum all of the values and dividing it by the total number of values

UNet - Transposed Convolution

- Transposed convolution is exactly the opposite process of a normal convolution i.e., the input is a low resolution image and the output is a high resolution image.
- The input is upsampled - basically expanded while a normal convolution operation downsamples and shrinks the input image.
- The transpose of the filter matrix is taken to reverse the convolution operation.



UNet - Batch Normalization

- Batch normalization is a method used to make training neural networks faster and stabilize the network through normalization of the layers' inputs by re-scaling.

UNet - Batch Normalization

- Batch normalization is a method used to make training neural networks faster and stabilize the network through normalization of the layers' inputs by re-scaling.
- In UNet, a batch normalization layer is present in between the convolution layer and the ReLU activation function.

UNet - Batch Normalization

- Batch normalization is a method used to make training neural networks faster and stabilize the network through normalization of the layers' inputs by re-scaling.
- In UNet, a batch normalization layer is present in between the convolution layer and the ReLU activation function.
- It fixes the problem of internal covariate shift and provides regularization.

UNet - Dropout Layer

- Dropout refers to ignoring units (i.e. neurons) during the training phase of certain set of neurons which is chosen at random.

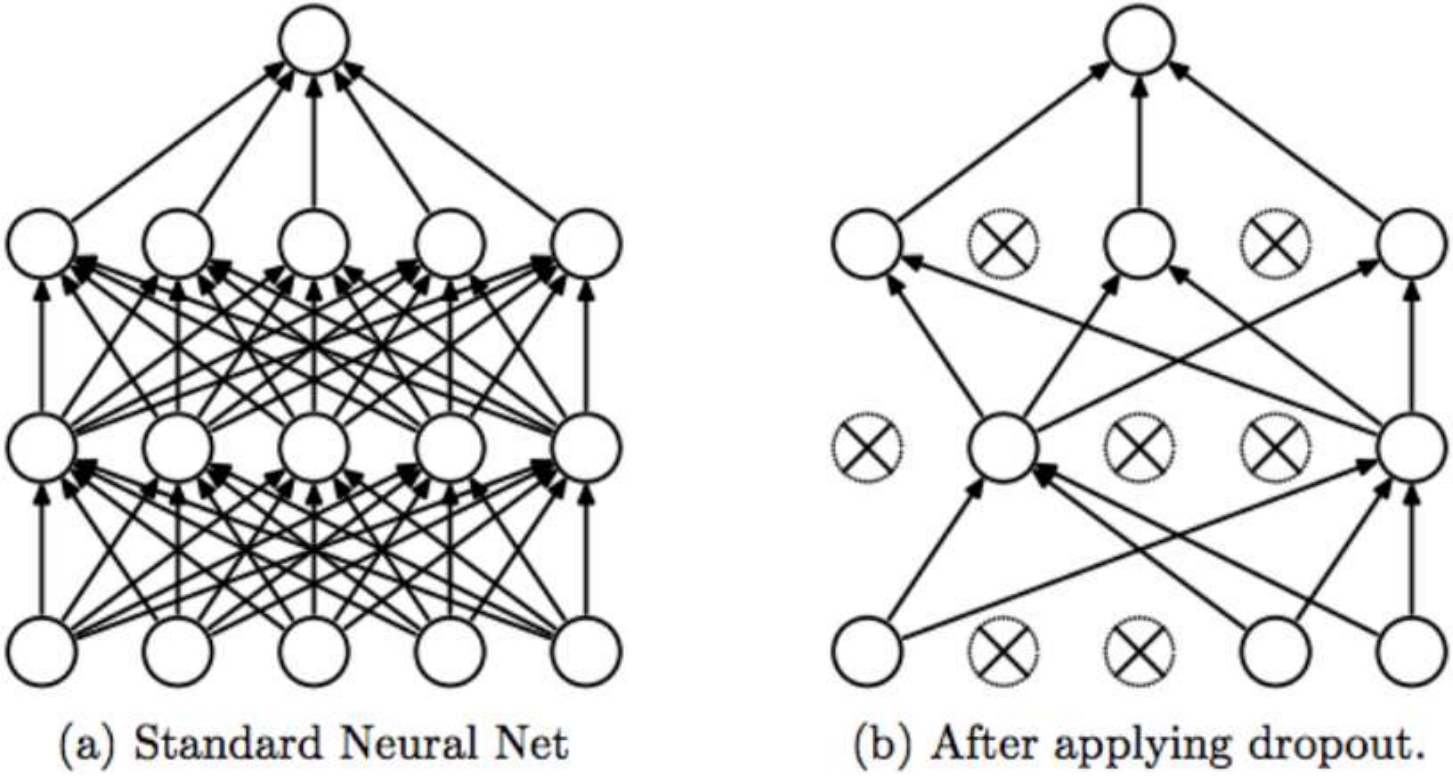
UNet - Dropout Layer

- Dropout refers to ignoring units (i.e. neurons) during the training phase of certain set of neurons which is chosen at random.
- A dropout layer is essential to prevent overfitting.

UNet - Dropout Layer

- Dropout refers to ignoring units (i.e. neurons) during the training phase of certain set of neurons which is chosen at random.
- A dropout layer is essential to prevent overfitting.
- Overfitting occurs when the model learns unnecessary information while training, and fits too closely to the training set. It becomes unable to generalize well to new data.

UNet - Dropout Layer

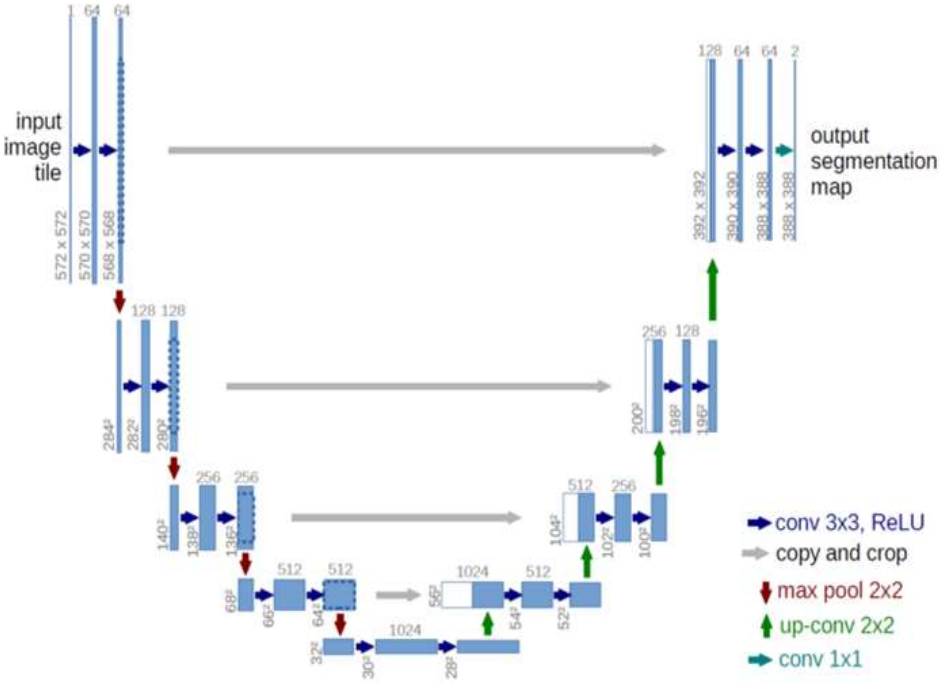


(a) Standard Neural Net

(b) After applying dropout.

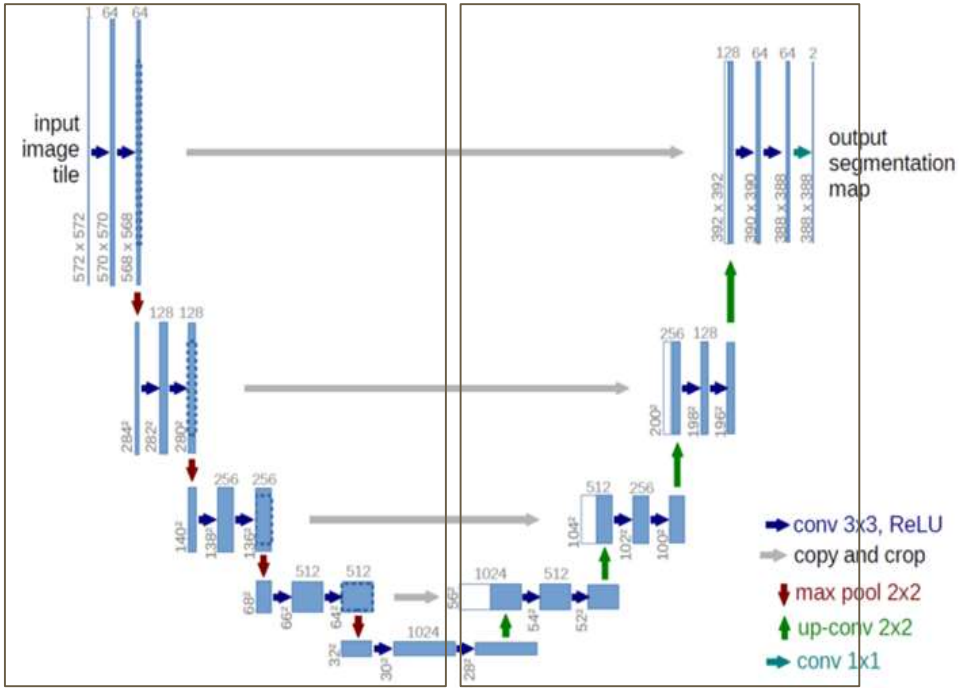
UNet

- U-Net is a convolutional neural network that was developed for biomedical image segmentation at the Computer Science Department of the University of Freiburg.
- UNET is a U-shaped encoder-decoder network architecture, which consists of four encoder blocks and four decoder blocks that are connected via a bridge.



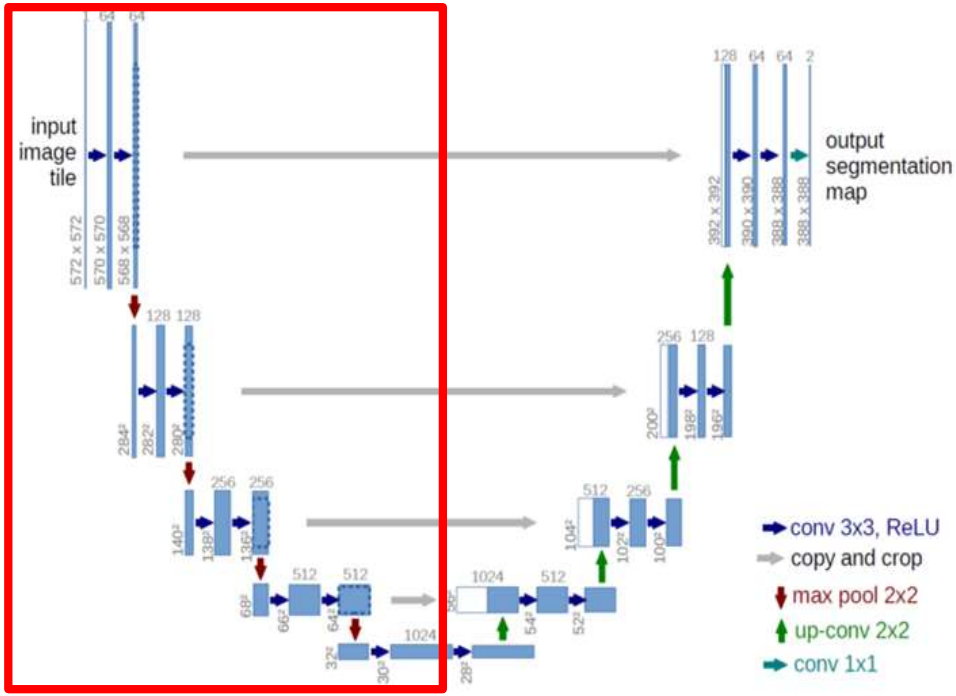
UNet

- The architecture contains two paths.
- First path is the contraction path (also called as the encoder) which is used to capture the context in the image.
- The decoder network is used to take the abstract representation and generate a semantic segmentation mask.



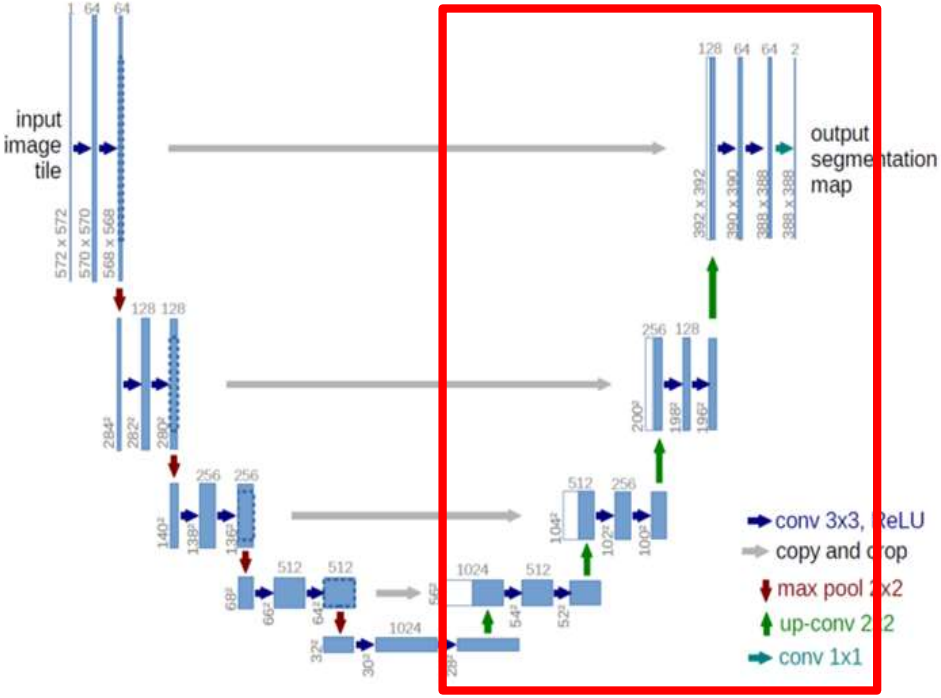
The Encoder Network (Contracting path)

- Each encoder block consists of two 3×3 convolutions, where each convolution is followed by a ReLU (Rectified Linear Unit) activation function.
- Next, follows a 2×2 max-pooling, where the spatial dimensions (height and width) of the feature maps are reduced by half.



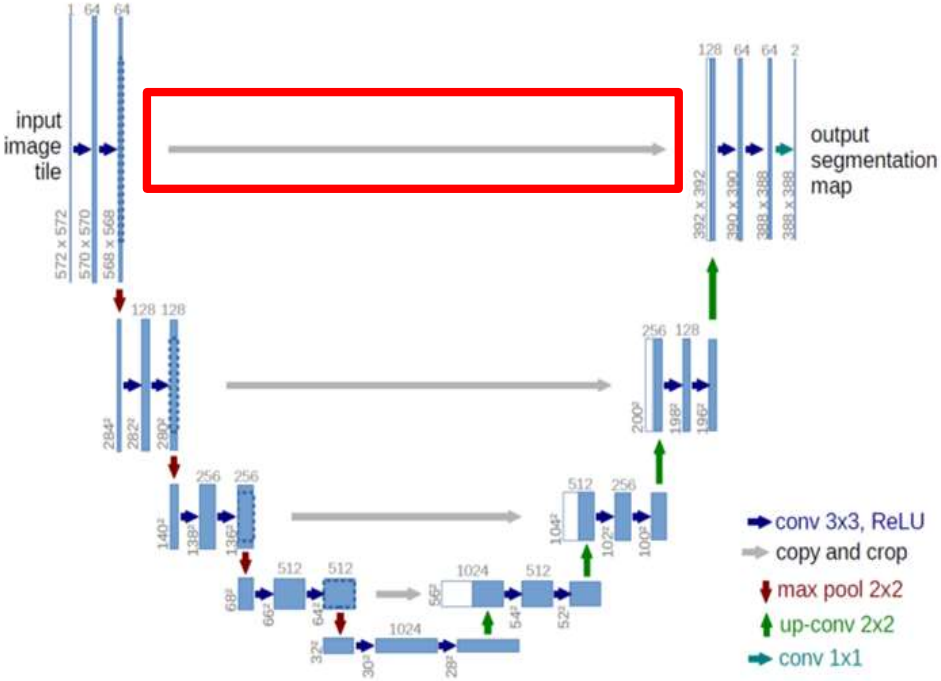
The Decoder Network (Expanding path)

- The decoder block starts with a 2×2 transpose convolution.
- Next, it is concatenated with the corresponding skip connection feature map from the encoder block.
- After that, two 3×3 convolutions are used, where each convolution is followed by a ReLU activation function.
- The output of the last decoder passes through a 1×1 convolution with sigmoid activation. The sigmoid activation function gives the segmentation mask representing the pixel-wise classification.



Skip Connections

- Skip connections provide additional information that helps the decoder to generate better semantic features.
- At every step of the decoder we use skip connections by concatenating the output of the transposed convolution layers with the feature maps from the Encoder at the same level



Day 3 - Writing UNet from scratch

- Understand the architecture of UNet and its layers
  - Activity 1: Build the model, train the first iteration
- Evaluate the models to get preliminary results
  - Activity 2: Use the model weights to predict segmentation maps
- Build variations of the UNet model
  - Activity 3: Experiment with dropout layers, re-train, and plot learning curves