

# ResNets and Model Building



# Residual Neural Networks.

- A particular way of structuring Neural Networks, Not a different type of network itself such as a CNN or RNN. It can be made of any kind of network.
- Uses skip-connections to send information past layers of the network without any augmentation. This allows for information to be delivered from the input layer to the output layer directly regardless of how many layers are in-between.



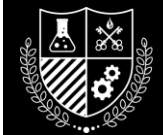
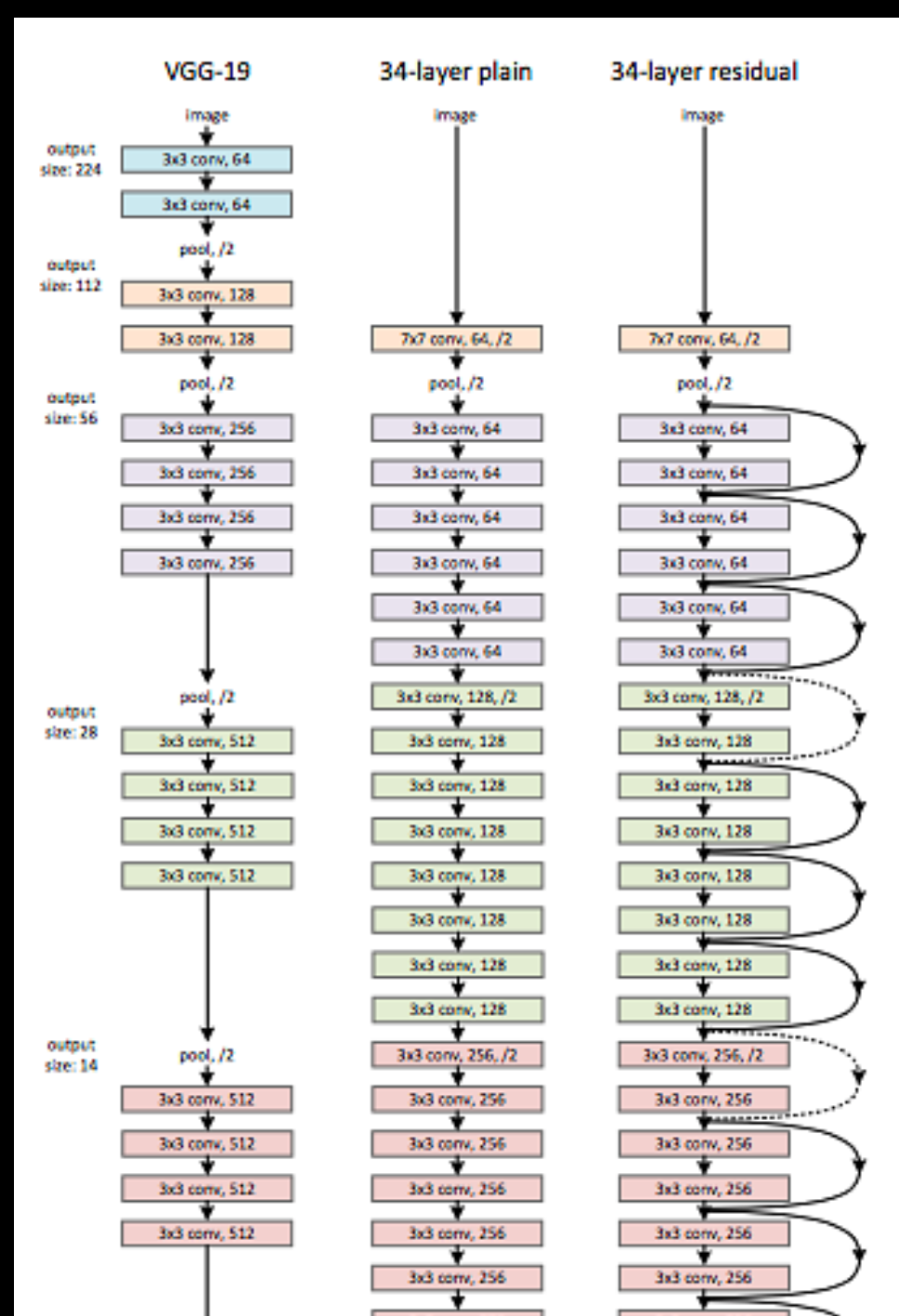
# What is so important about ResNets?

- They get rid of vanishing and exploding gradients and they speed up training time considerably while producing better results.
- It can build up layers slowly instead of all at once and it allows each layer to be specialized to some part of the network.
- It works like a random forest model.



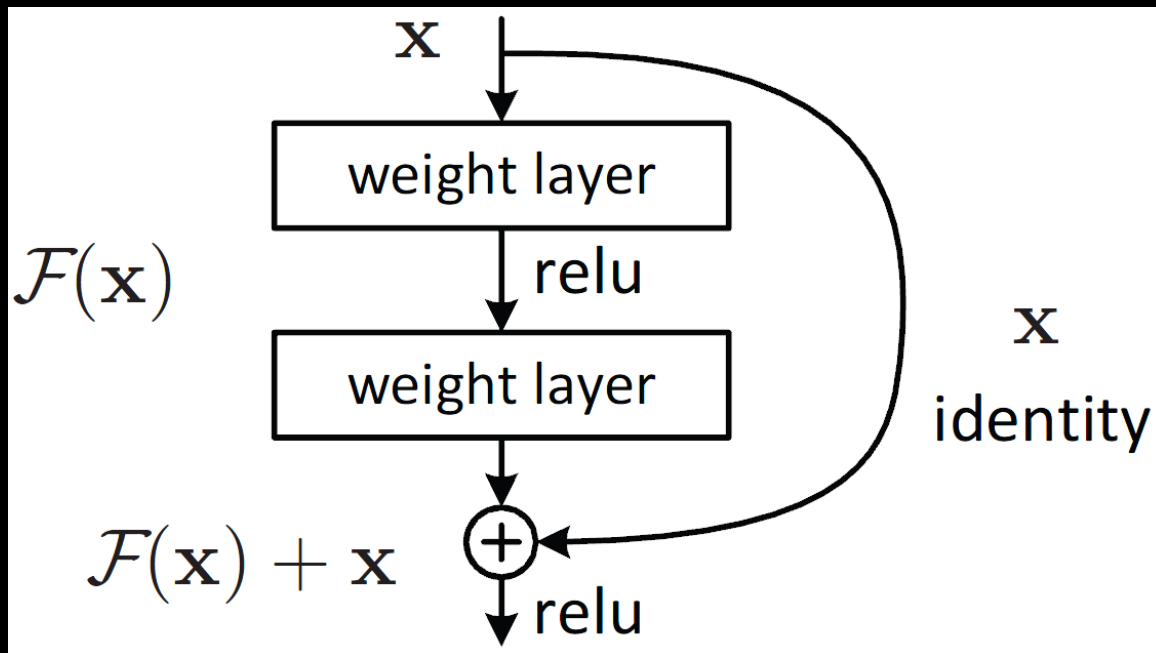
# How it works

- It starts with empty layers that are filled over time as the model gets more confident in its knowledge.
- It trains only a single residual block at a time, essentially training a 'two-layer' network.



# Math of ResNets

- The number of skip connections can be adjusted but 2-3 is typically used.



$$Y = F(x)$$

$F(x)$ , activation function

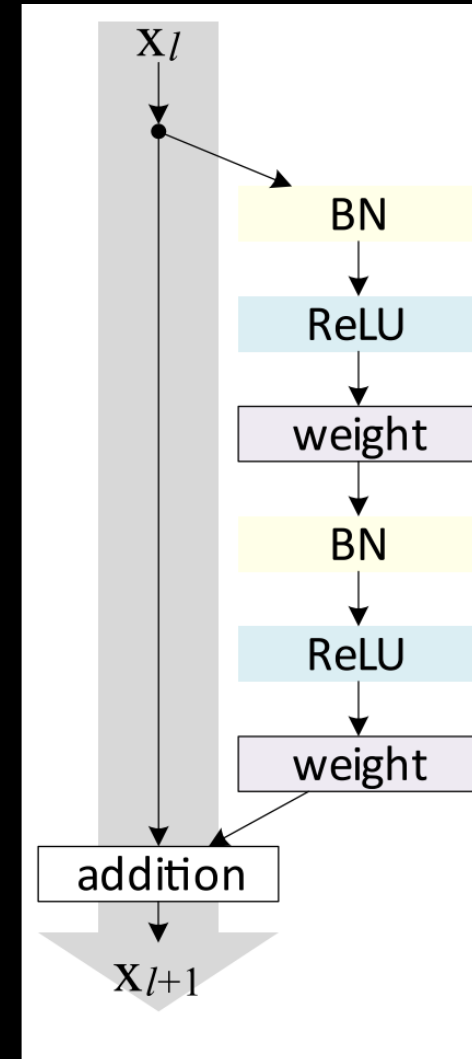
$X$ , Input

$Y$ , Output of activation function



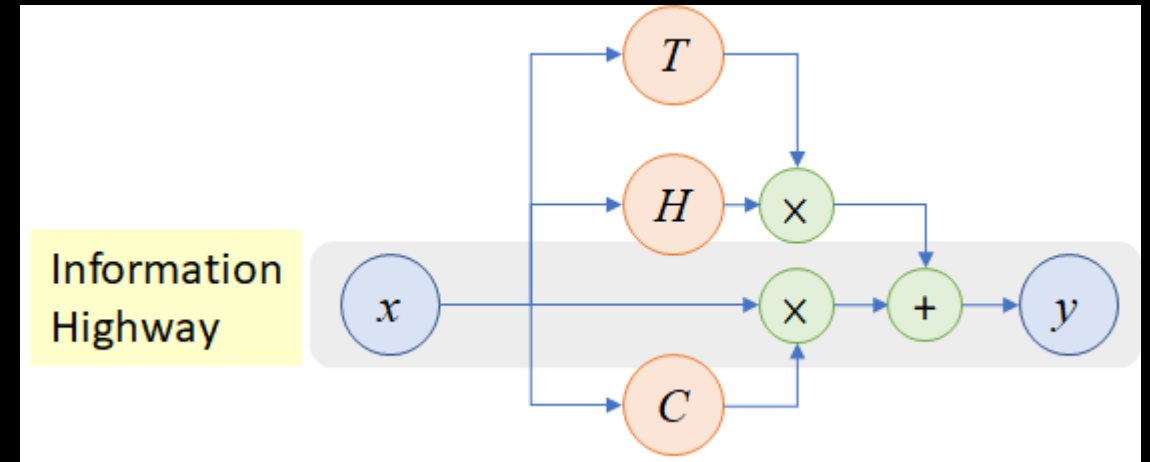
# Normal ResNet residual blocks

- They include a Batch norm and a ReLU for every skip layer in the block. This example block as two skip connections and so it has two BatchNorms and ReLU layers.

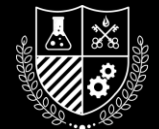


# Highway Nets

- They create an information highway that allows the input to pass directly through the network like normal ResNets.
- They allow the weights of both the Carry and Transform gates to augment the input  $X$ .

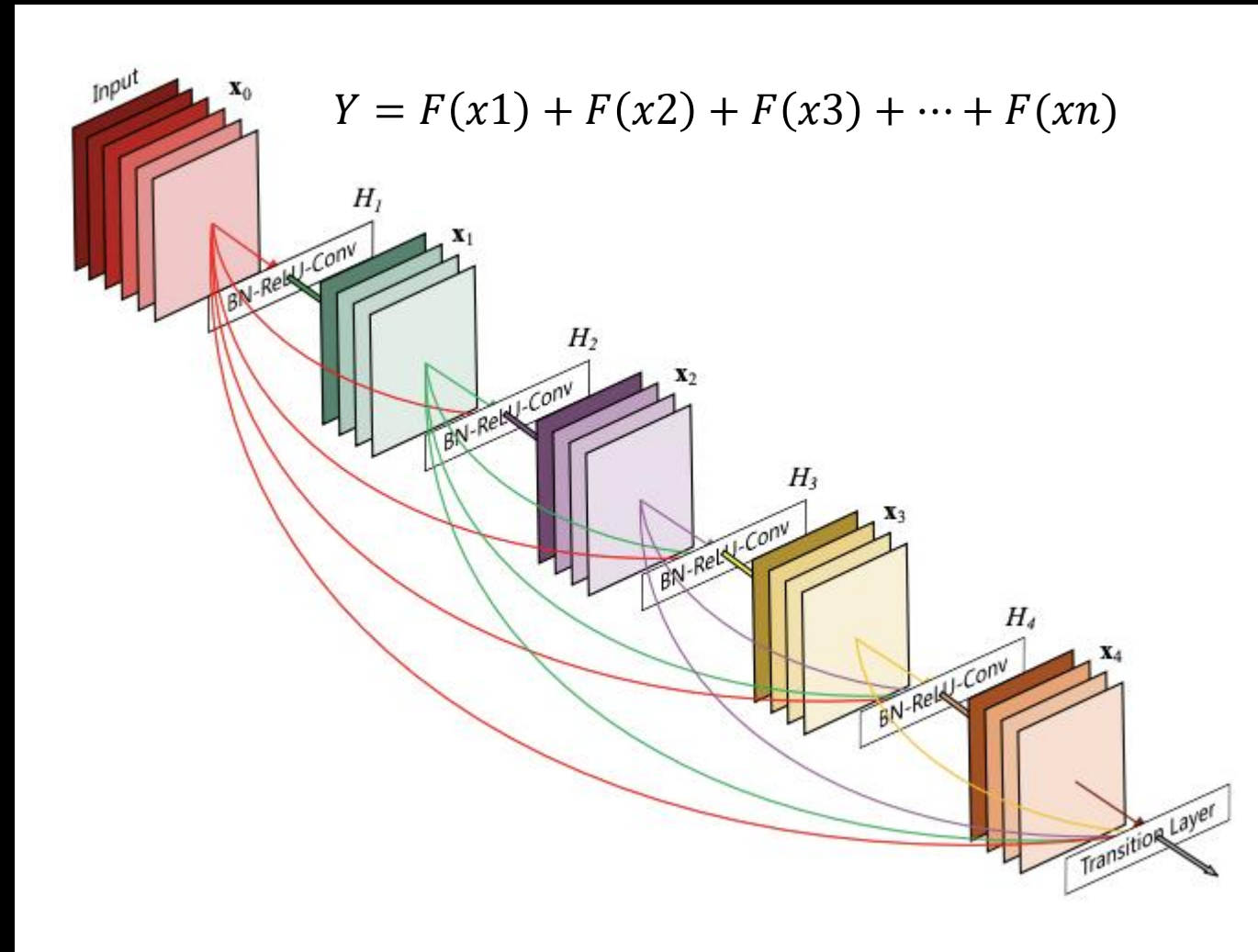


$$Y = H * T + X * C$$
$$T + C = 1$$



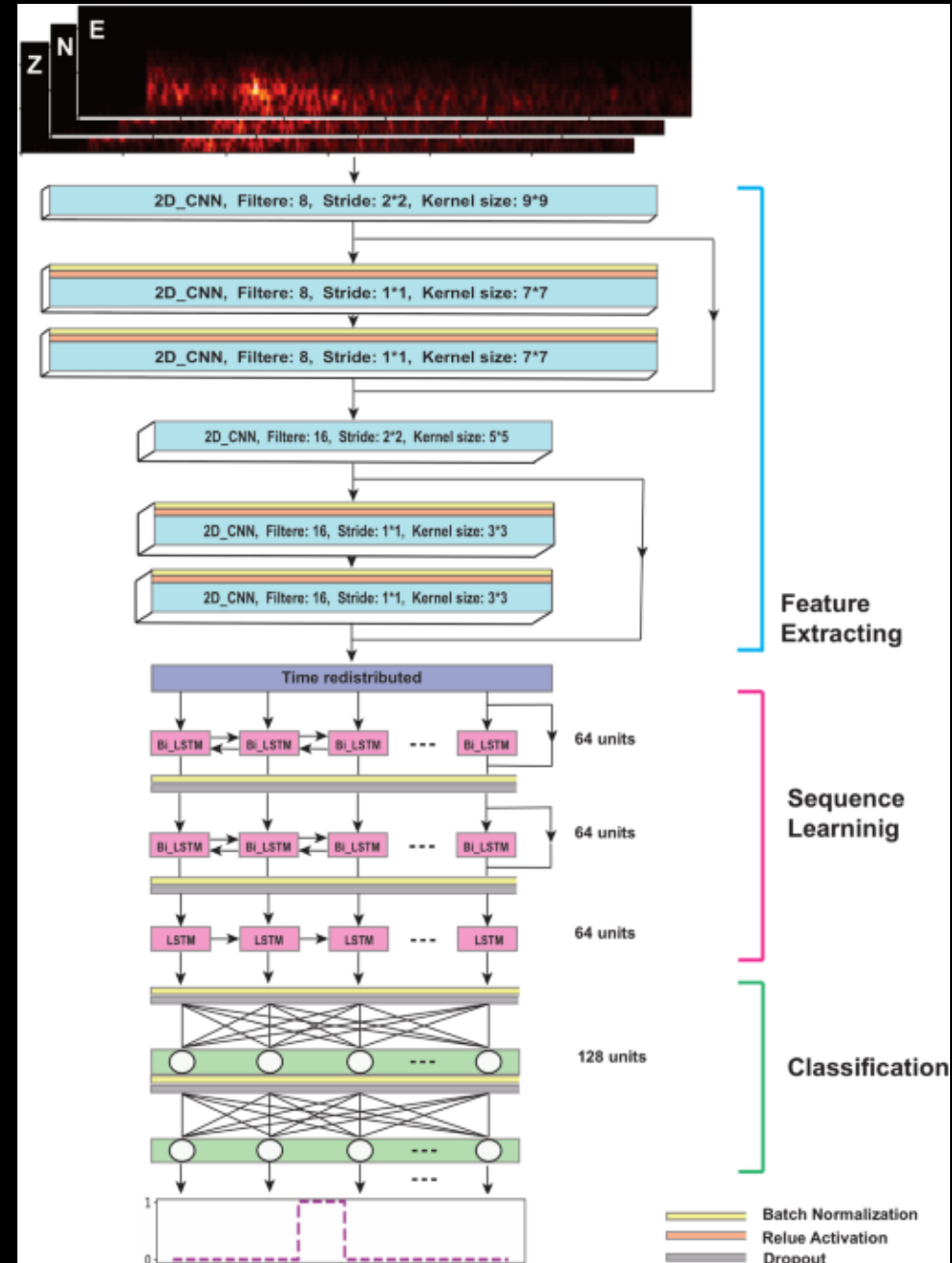
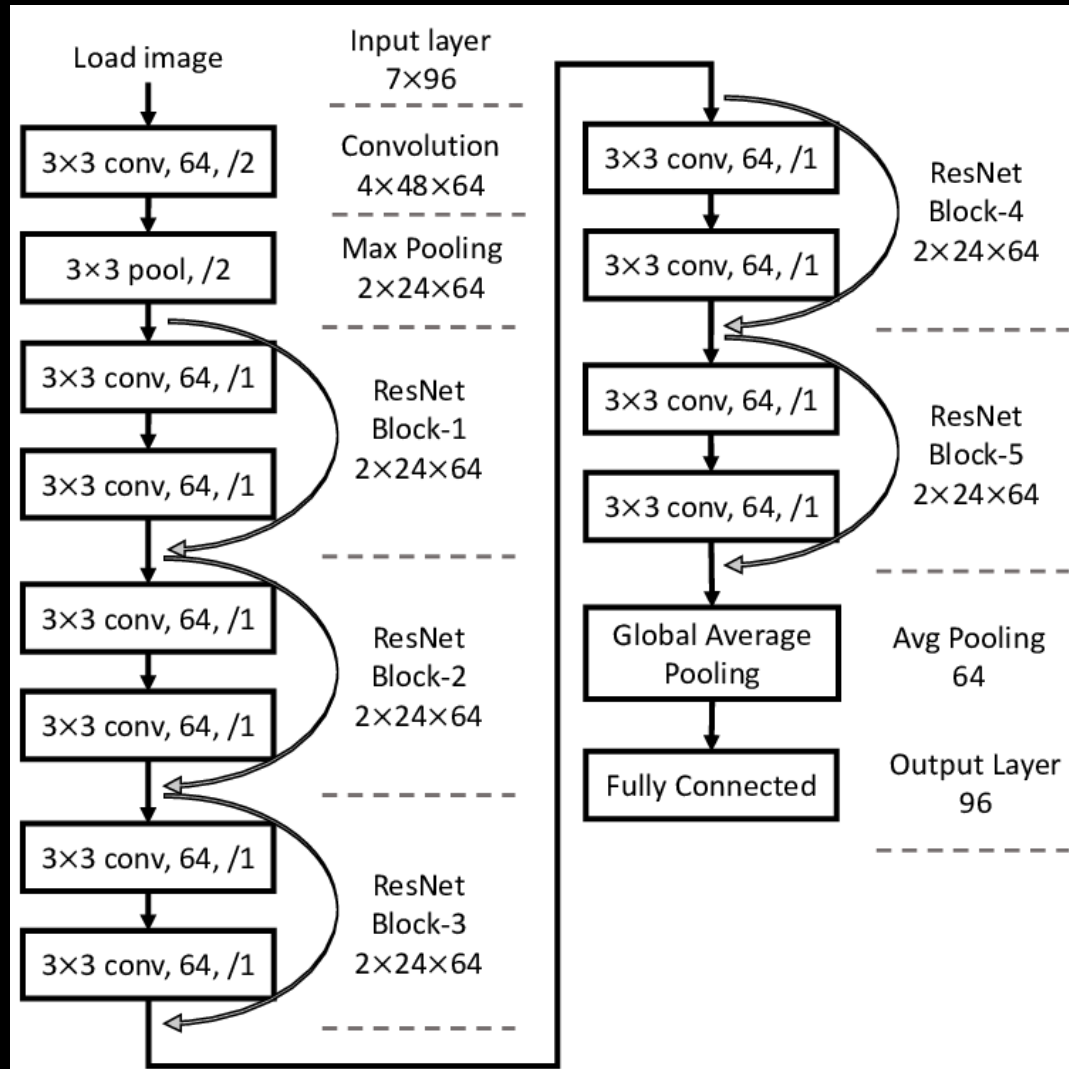
# Dense Nets

- Every layer is connected to every layer.
- Allows for direct back propagation to update individual layers. Which allows the model to have the greatest control over itself.
- Can gain deeper insight and knowledge of the data.
- State of the Art in 2018.





# Network Stacking



# Examples

- There is so much flexibility and power involved with how you setup a network. Deep learning is endlessly customizable so get used to trying new architectures!

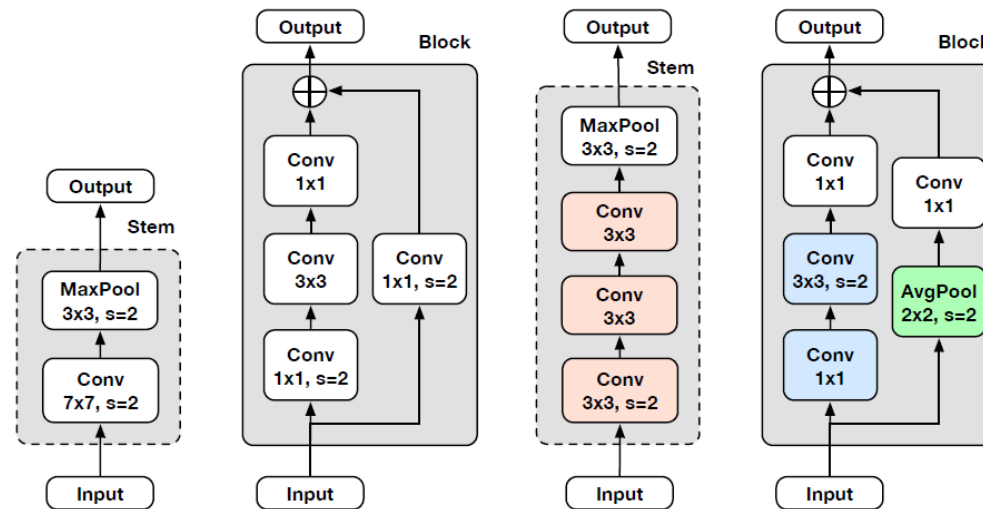
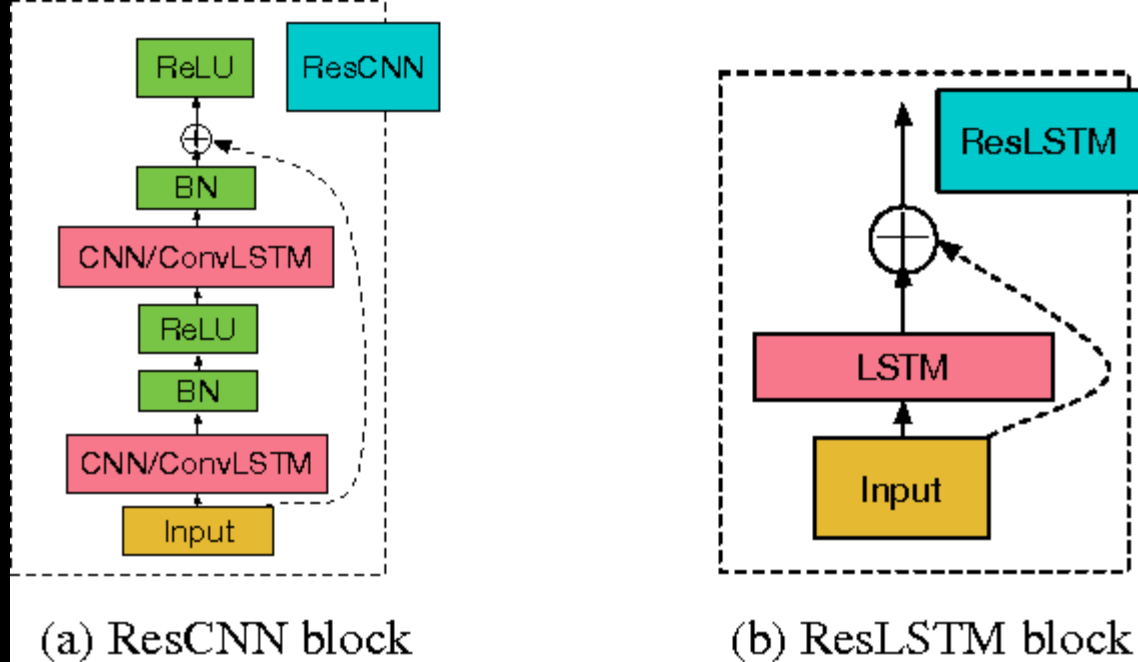


Image from the original paper



# Final Take Away

- Learn how to tune and create deep models to fit every need of your employer.
- How you structure things is just as important as what you are using!

