We all know that training a Convolution Neural Network(CNN) from scratch takes a lot of data and also compute power. So we instead use [transfer learning](https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751), where a model trained on similar data is fine-tuned as per our requirement.

<https://github.com/davidsandberg/facenet>

<https://medium.com/@BhashkarKunal/face-recognition-real-time-webcam-face-recognition-system-using-deep-learning-algorithm-and-98cf8254def7>

<https://www.hindawi.com/journals/cin/2020/7251280/>

<https://www.kaggle.com/timesler/comparison-of-face-detection-packages>

<https://ai.googleblog.com/2016/08/improving-inception-and-image.html>

<https://icml.cc/2016/tutorials/icml2016_tutorial_deep_residual_networks_kaiminghe.pdf>

<https://www.kaggle.com/byrachonok/pretrained-inceptionresnetv2-base-classifier>

<https://stackoverflow.com/questions/39582703/using-pre-trained-inception-resnet-v2-with-tensorflow>

<https://collinerickson.github.io/2018/08/04/how-to-load-a-pretrained-model-in-tensorflow/>

<https://towardsdatascience.com/how-to-use-a-saved-model-in-tensorflow-2-x-1fd76d491e69>

<https://www.tensorflow.org/lite/tutorials>

https://www.tensorflow.org/lite/convert

https://github.com/tensorflow/models/tree/master/research/slim

https://github.com/tensorflow/models/blob/master/research/slim/README.md#pre-trained-models

https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1374&context=creativecomponents

speed of the convolution will depend on the size of the inpu

The VGG-16 is one of the most popular pre-trained models for image classification



Here is a more intuitive layout of the VGG-16 Model.



The following are the layers of the model:

* Convolutional Layers = 13
* Pooling Layers = 5
* Dense Layers = 3

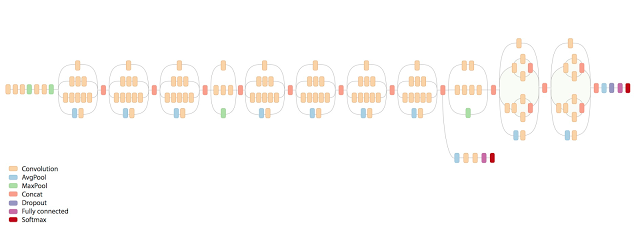
Let us explore the layers in detail:

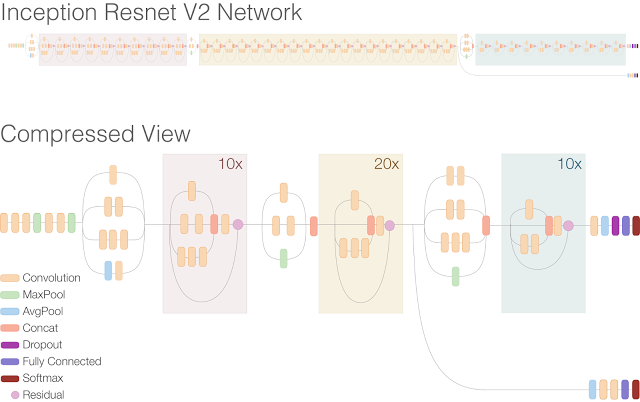
1. **Input**: Image of dimensions (224, 224, 3).
2. **Convolution Layer Conv1:**
   * Conv1-1: 64 filters
   * Conv1-2: 64 filters and Max Pooling
   * Image dimensions: (224, 224)
3. **Convolution layer Conv2:**Now, we increase the filters to 128
   * Input Image dimensions: (112,112)
   * Conv2-1: 128 filters
   * Conv2-2: 128 filters and Max Pooling
4. **Convolution Layer Conv3:**Again, double the filters to 256, and now add another convolution layer
   * Input Image dimensions: (56,56)
   * Conv3-1: 256 filters
   * Conv3-2: 256 filters
   * Conv3-3: 256 filters and Max Pooling
5. **Convolution Layer Conv4:**Similar to Conv3, but now with 512 filters
   * Input Image dimensions: (28, 28)
   * Conv4-1: 512 filters
   * Conv4-2: 512 filters
   * Conv4-3: 512 filters and Max Pooling
6. **Convolution Layer Conv5:**Same as Conv4
   * Input Image dimensions: (14, 14)
   * Conv5-1: 512 filters
   * Conv5-2: 512 filters
   * Conv5-3: 512 filters and Max Pooling
   * The output dimensions here are (7, 7). At this point, we flatten the output of this layer to generate a feature vector
7. **Fully Connected/Dense FC1**: 4096 nodes, generating a feature vector of size(1, 4096)
8. **Fully ConnectedDense FC2**: 4096 nodes generating a feature vector of size(1, 4096)
9. **Fully Connected /Dense FC3**: 4096 nodes, generating 1000 channels for 1000 classes. This is then passed on to a Softmax activation function
10. **Output layer**

Loading the Base model - We will be using only the basic models, with changes made only to the final layer. This is because this is just a binary classification problem while these models are built to handle up to 1000 classes

input\_shape = (224, 224, 3) , loss = 'binary\_crossentropy'

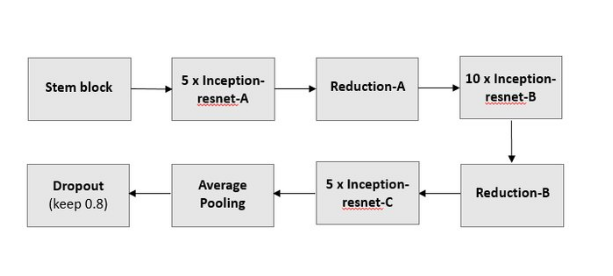
* **InceptionResnetV2**





is a convolutional neural architecture that builds on the Inception family of architectures but incorporates [residual connections](https://paperswithcode.com/method/residual-connection) (replacing the filter concatenation stage of the Inception architecture).  The network is 164 layers deep and can classify images into 1000 object categories. The network has an image input size of 299-by-299.

Inception-ResNet combines the Inception architecture, with residual connections.  In the Inception-Resnet block, multiple sized convolutional filters are combined with residual connections. The usage of residual connections not only avoids the degradation problem caused by deep structures but also reduces the training time



The basic architecture of Inception-Resnet-v2

