(1) Workloads

VGGNet： VGGNet是牛津大学计算机视觉组和Google DeepMind公司的研究员一起研发的深度卷积神经网络。VGGNet探索了卷积神经网络的深度与其性能之间的关系，通过反复堆叠3\*3的小型卷积核和2\*2的最大池化层，VGGNet成功的构筑了16-19层深的卷积神经网络。VGGNet相比之前的的网络结构，错误率大幅度下降，并取得了ILSVRC 2014比赛分类项目的第2名和定位项目的第1名。同时，VGGNet的拓展性很强，迁移到其他图片数据上的泛化性能非常好。VGGNet的结构非常简洁，整个网络都使用同样大小的卷积核尺寸和最大池化尺寸。到目前为止，VGGNet依然经常被用来提取图像特征。

VGGNet is a deep convolutional neural network developed by the Computer Vision Group at Oxford University and researchers at Google DeepMind. VGGNet explores the relationship between the depth of a convolutional neural network and its performance. By repeatedly stacking 3\*3 small convolutional kernels and 2\*2 largest pooled layers, VGGNet successfully builds a 16-19 layer deep convolution. Neural Networks. Compared with the previous network structure, VGGNet significantly reduced the error rate, and achieved the second place of the ILSVRC 2014 competition classification project and the first place of the positioning project. At the same time, VGGNet's scalability is very strong, and the generalization performance of migrating to other image data is very good. VGGNet's structure is very simple, the entire network uses the same size of the convolution kernel size and maximum pool size.

So far, VGGNet is still often used to extract image features.

Inception Net： Google Inception Net首次出现在ILSVRC 2014的比赛中就以较大的优势取得了第一名。比赛中的Inception Net通常被称为Inception V1，其最大的特点是控制了计算量和参数的同时，获得了非常好的分类性能。Inception V1中提出了Inception Module用来提高参数的利用率，虽然他的深度达到22层，但是参数量却远低于ALexNet和VGGNet。随后，在2015年和2016年又相继出现了Inception V2，Inception V3和Inception V4。Inception V2参考了VGGNet，用两个3\*3的卷积代替了5\*5的卷积用来降低参数量。同时提出了著名的Batch Normalization的方法，加快了网络的训练速度，同时收敛后的分类准确率也可以得到提高。Inception V3通过将较大的二维卷积网络拆分成较小的一维卷积网络，一方面节省了大量参数，加速运算并减轻了过拟合，同时也增加了一层非线性扩展模型表达能力，使得网络可以处理更多更丰富的空间特征，增加特征的多样性。另一方面，Inception V3优化了Inception Module的结构。随后的Inception V4在Inception V3的基础上进一步结合Resnet进一步提升准确率。

Inception Net: Google Inception Net won first place in the ILSVRC 2014 competition. The Inception Net in the game is usually called Inception V1. Its greatest feature is that it controls the calculations and parameters while achieving very good classification performance. Inception V1 proposes Inception Module to increase the utilization of parameters. Although his depth reaches 22 levels, the parameter amount is much lower than ALexNet and VGGNet. Subsequently, Inception V2, Inception V3 and Inception V4 emerged successively in 2015 and 2016. Inception V2 refers to VGGNet and uses two 3\*3 convolutions instead of a 5\*5 convolution to reduce the amount of parameters. At the same time, the famous method of Batch Normalization was proposed to speed up the training of the network, and the accuracy of the classification after convergence can also be improved. Inception V3 splits the larger two-dimensional convolutional network into a smaller one-dimensional convolutional network. On the one hand, it saves a lot of parameters, speeds up calculations, and reduces over-fitting. It also adds a layer of non-linear expansion model. Ability to express so that the network can handle more and more abundant spatial features and increase the diversity of features. On the other hand, Inception V3 optimizes the structure of the Inception Module. Subsequent Inception V4 further integrated Resnet on the basis of Inception V3 to further improve accuracy.

Resnet：Resnet由微软研究员的Kaiming He等人提出的，在ILSVRC 2015比赛中获得了冠军，取得了3.57%的top5的错误率。它使用了一种叫做“shortcut connection”的连接方式来降低网络参数，通过shortcut将block的输入和输出进行一个element-wise的加叠，这个简单的加法并不会给网络增加额外的参数和计算量，同时却可以大大增加模型的训练速度、提高训练效果，并且当模型的层数加深时，这个简单的结构能够很好的解决退化问题。在Resnet的结构中，提出了两层的残差学习单元和三层的残差学习单元。两层的残差学习单元中包含两个相同的输出通道数的3\*3卷积。三层的残差网络使用了1\*1的卷积并且是在中间3\*3的卷积前后都使用了1\*1的卷积。我们使用的50,101,152层主要是使用的三层残差结构。

Resnet: Microsoft Researcher Kaiming He and others proposed Resnet and won the championship in ILSVRC 2015. Resnet's TOP5 error rate was only 3.75%.Resnet uses a connection method called "shortcut connection" to reduce network parameters, and stack the input and output of the block. This simple addition will not add extra parameters and calculations to the network, but it can greatly increase the training speed of the model, improve the training effect, and when the number of layers of the model deepens, this simple structure can solve the problem of degradation well. There are two different learning units in Resnet's structure: two levels of residual learning units and three levels of residual learning units. The two-level residual learning unit contains two 3\*3 convolutions with the same number of output channels. The three-level residual network uses a 1\*1 convolution before and after the middle 3\*3 convolution. The 50,101,152-layer networks we use are mainly three-tier residual structures.

Mobilenet：MobileNet是2017年由Horward等人提出的，它的架构由一个作用于输入图像的标准卷积层，一个深度可分离卷积堆栈以及最后的平均池和完全连接的层组成。Mobilenet采用深度可分离的卷积来构建轻量级的深层神经网络，可以将标准卷积分解成一个深度卷积和一个点卷积（1 × 1卷积核）。深度卷积将每个卷积核应用到每一个通道，而1 × 1卷积用来组合通道卷积的输出，这种分解可以有效减少计算量，降低模型大小。MobileNet在计算量和模型尺寸方面具备很明显的优势。

MobileNet: MobileNet was proposed by Horward et al. in 2017. Its architecture consists of a standard convolution layer acting on the input image, a deep separable convolutional stack, and a final average pool and fully connected layer. Mobilenet uses a deeply separable convolution to build a lightweight, deep neural network that can integrate standard volumes into a deep convolution and a dot convolution (1 × 1 convolution kernel). Depth convolution applies each convolution kernel to each channel, and 1 × 1 convolution is used to combine the output of channel convolutions. This decomposition can effectively reduce the computation and reduce the model size. MobileNet has a clear advantage in terms of computational volume and model size.