珠海学院

**深度学习报告**

|  |  |
| --- | --- |
| 学院： | 吉林大学珠海学院计算机学院 |
| 专业： | 软件工程 |
| 设计科目： | 深度学习与计算机视觉 |
| 学生姓名： | 04191210-陈凯亮 |
| 指导教师： | 马永骏 |
| 完成时间： | 2021.12.18 |

**基于卷积神经网络实现的面部表情识别**

作者：陈凯亮

吉林大学珠海学院计算机学院 班级：软件工程2班 学号：04191210

在日常工作和生活中，人们情感的表达方式主要有：语言、声音、肢体行为（如手势）、以及面部表情等。上世纪70年代，美国著名心理学家保罗•艾克曼经过大量实验之后，将人类的基本表情定义为**悲伤、害怕、厌恶、快乐、气愤和惊讶**六种。同时，他们根据不同的面部表情类别建立了相应的表情图像数据库。随着研究的深入，**中性**表情也被研究学者加入基本面部表情中，组成了现今的人脸表情识别研究中的七种基础面部表情。

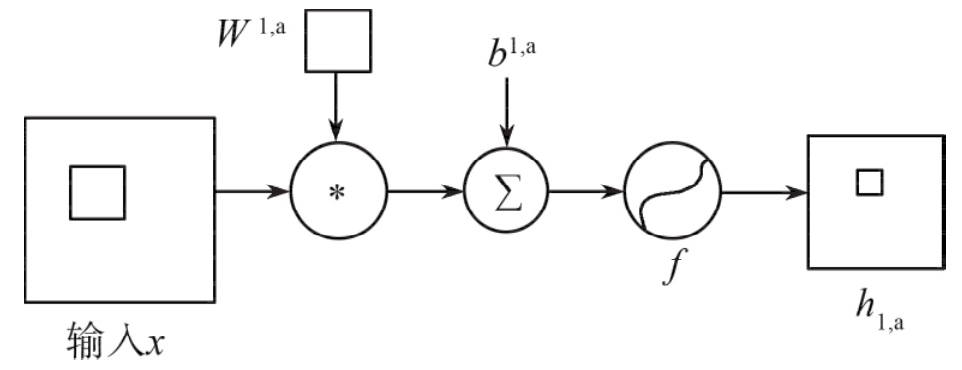


在计算机视觉中，因为表情识别作为人机交互的一种桥梁，可以更好的帮助机器了解、识别人类的内心活动，从而更好的服务人类，因此对人脸表情识别进行深入研究具有十分重要的意义。

**卷积层**

随着科学技术的发展，传统的一些基于手工提取人脸表情图像特征的方法，因为需要研究人员具有比较丰富的经验，具有比较强的局限性，从而给人脸表情识别的研究造成了比较大的困难。随着深度学习的兴起，作为深度学习的经典代表的卷积神经网络，由于其具有自动提取人脸表情图像特征的优势，使得基于深度学习的人脸表情特征提取方法逐渐兴起，并逐步替代一些传统的人脸表情特征提取的方法。深度学习方法的主要优点在于它们可通过使用非常大的数据集进行训练学习从而获得表征这些数据的最佳功能。基于深度学习的算法。该方法使用卷积神经网络，自动地提取人脸表情特征。卷积神经网络对原始图像进行简单的预处理之后，可以直接输入到网络中，使用端到端的学习方法，即不经过传统的机器学习复杂的中间建模过程，如在识别中对数据进行标注、翻转处理等，直接一次性将数据标注好，同时学习特征与进行分类，这是深度学习方法与传统方法的重要区别。相比人工的选取与设计图像特征，卷积神经网络通过自动学习的方式，获得的样本数据的深度特征信息拥有更好的抗噪声能力、投影不变性、推广与泛化能力、抽象语义表示能力。

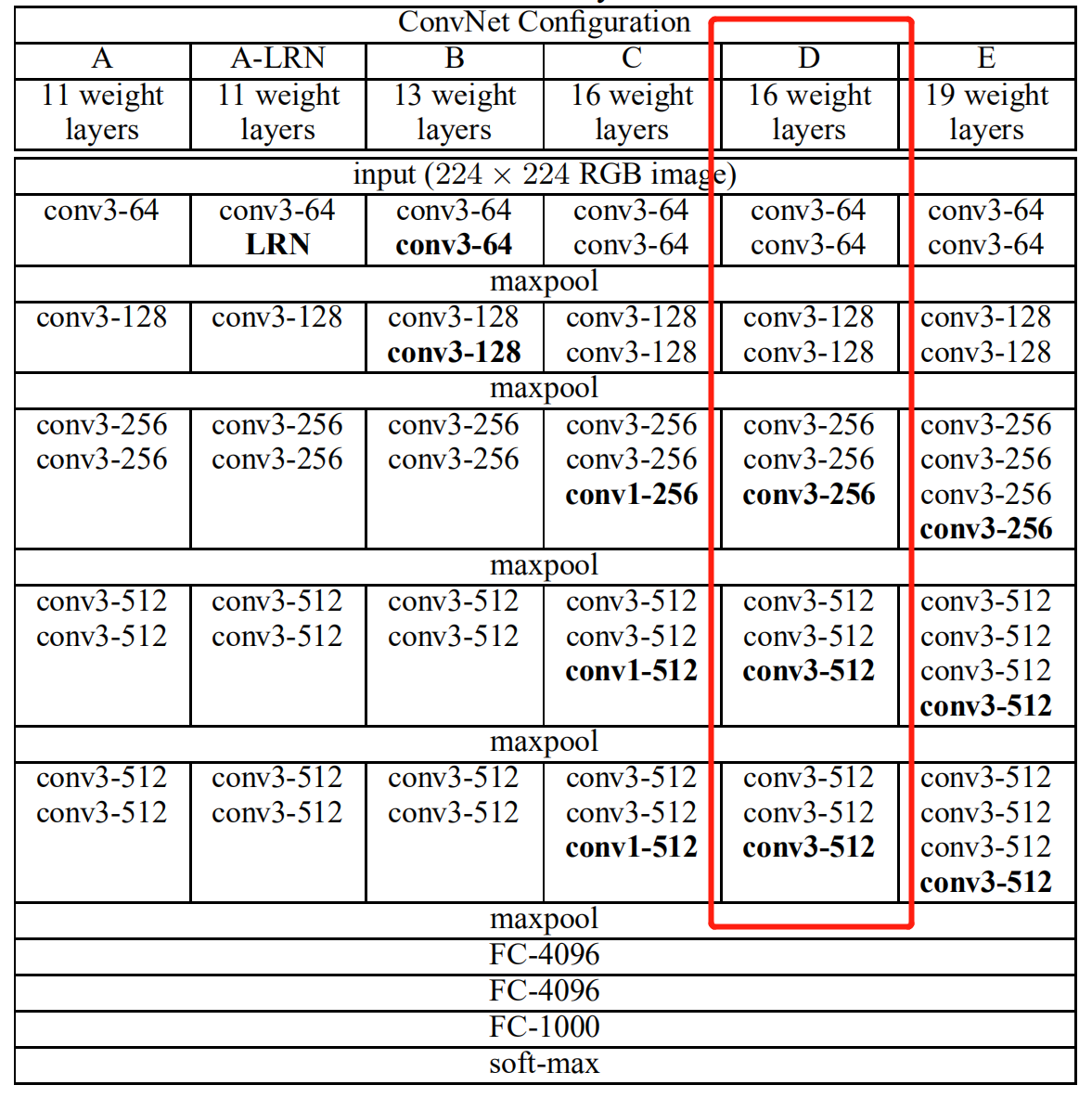
卷积层是卷积神经网络的第一层，由几个卷积单元组成。每个卷积单元的参数可以通过反向传播算法进行优化，其目的是提取输入的各种特征，但是卷积层的第一层只能提取低级特征，例如边、线和角。更多层的可以提取更高级的特征，利用卷积层对人脸面部图像进行特征提取。一般卷积层结构如下图所示，卷积层可以包含多个卷积面，并且每个卷积面都与一个卷积核相关联。



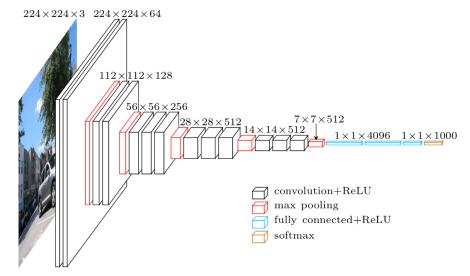
由上图可知，每次执行卷积层计算时，都会生成与之相关的多个权重参数，这些权重参数的数量与卷积层的数量相关，即与卷积层所用的函数有直接的关系。

**VGG**

随着深度学习算法的不断发展，众多卷积神经网络算法已经被应用到机器视觉领域中。VGGNet的提示更是将其推向高潮，VGGNet是首批把图像分类的错误率降低到10%以内模型，同时该网络所采用的3\times33×3卷积核的思想是后来许多模型的基础。



下面主要以的VGG16-3为例：



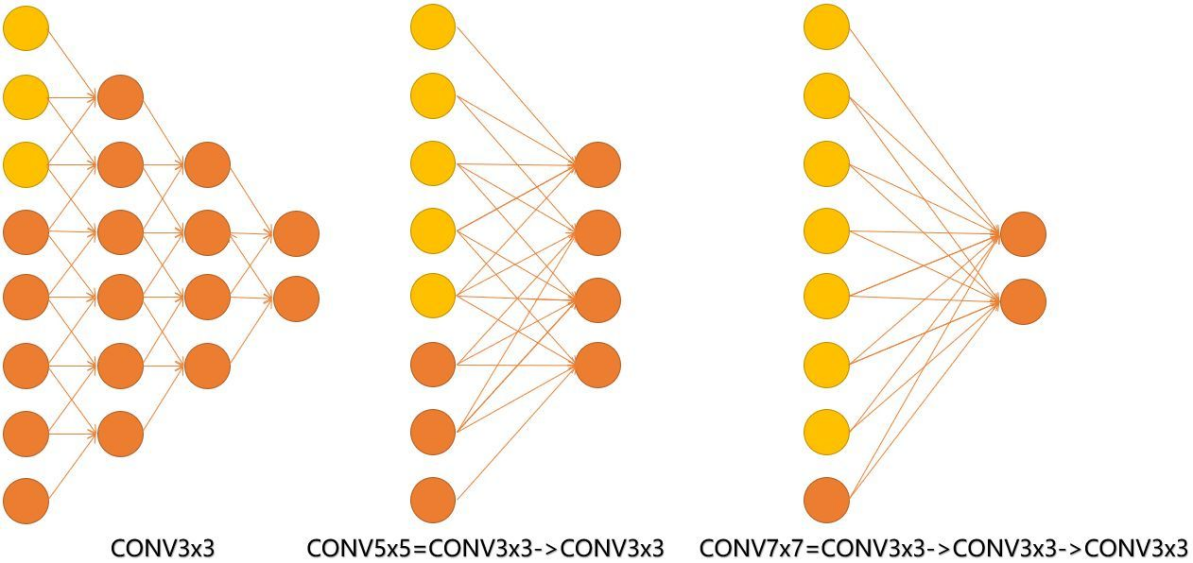
## **VGG模型的优点**

（1）**小卷积核:** 将卷积核全部替换为3x3（极少用了1x1）,作用就是减少参数，减小计算量。此外采用了更小的卷积核我们就可以使网络的层数加深，就可以加入更多的激活函数，更丰富的特征，更强的辨别能力。卷积后都伴有激活函数，更多的卷积核的使用可使决策函数更加具有辨别能力。其实最重要的还是多个小卷积堆叠在分类精度上比单个大卷积要好。

（2）**小池化核:** 相比AlexNet的3x3的池化核，VGG全部为2x2的池化核。

（3）**层数更深:** 从作者给出的6个试验中我们也可以看到，最后两个实验的的层数最深，效果也是最好。

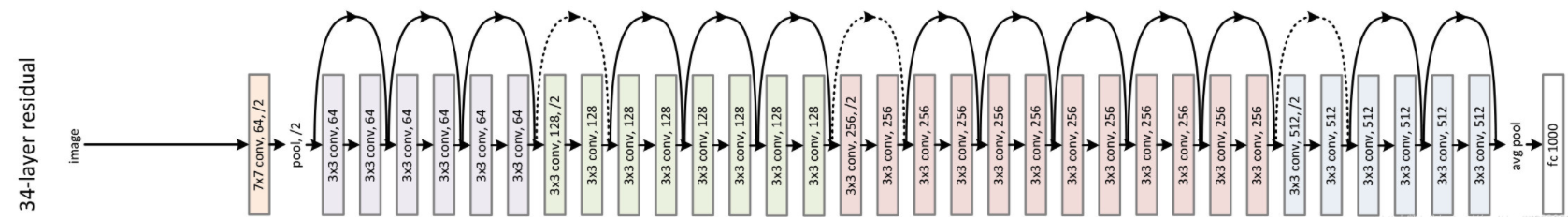
（4）**卷积核堆叠的感受野:** 作者在VGGnet的试验中只使用了两中卷积核大小：1\*1,3\*3。并且作者也提出了一种想法：两个3\*3的卷积堆叠在一起获得的感受野相当于一个5\*5卷积；3个3x3卷积的堆叠获取到的感受野相当于一个7x7的卷积。



# **基于 ResNet 的人脸面部表情识别算法**

ResNet的结构可以极快的加速神经网络的训练，模型的准确率也有比较大的提升。同时ResNet的推广性非常好，甚至可以直接用到InceptionNet网络中。

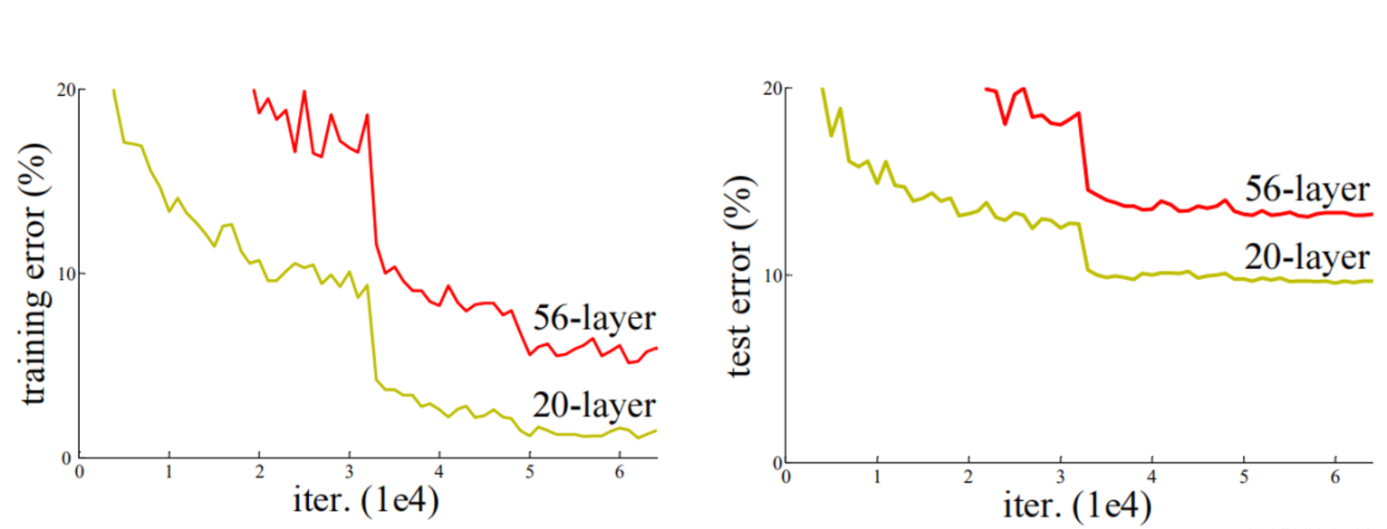
下图是ResNet34层模型的结构简图：



在ResNet网络提出之前，传统的卷积神经网络都是通过将一系列卷积层与下采样层进行堆叠得到的。但是当堆叠到一定网络深度时，就会出现两个问题。

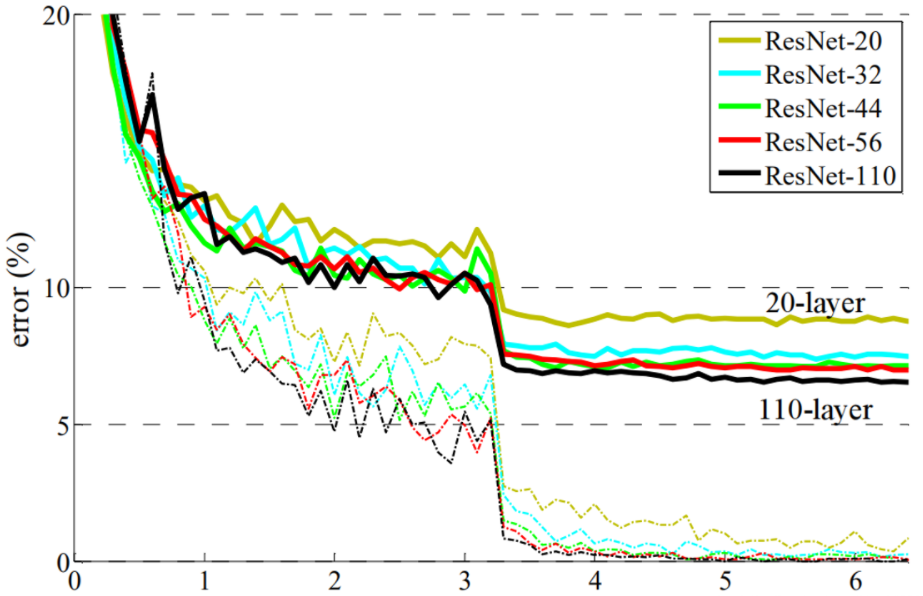
1.梯度消失或梯度爆炸。

2.退化问题(degradation problem)。



在ResNet论文中说通过数据的预处理以及在网络中使用BN（Batch Normalization）层能够解决梯度消失或者梯度爆炸问题。但是对于退化问题（随着网络层数的加深，效果还会变差，如下图所示）并没有很好的解决办法。

所以ResNet论文提出了residual结构（残差结构）来减轻退化问题。下图是使用residual结构的卷积网络，可以看到随着网络的不断加深，效果并没有变差，反而变的更好了。



# **人脸面部表情识别项目设计**

## **数据集准备**

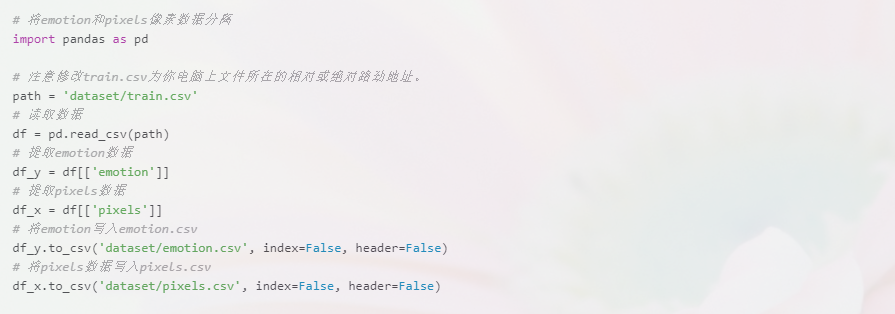
本项目是基于卷积神经网络模型开展表情识别的研究，为了尽可能的提高最终表情识别的准确性，需要大量的样本图片训练，优化，所以采用了 FER2013 数据集用来训练、测试，此数据集由 35886 张人脸表情图片组成，其中，测试图 28708 张，公共验证图和私有验证图各 3589 张，所有图片中共有7种表情。在预处理时把图像归一化为 48×48 像素，训练的网络结构是基于 CNN 网络结构的优化改进后的一个开源的网络结构，下文中会具体介绍到，通过不断地改进优化，缩小损失率，最终能达到较准确的识别出人的面部表情的结果采用了FER2013数据库。

FER2013数据集由28709张训练图，3589张公开测试图和3589张私有测试图组成。每一张图都是像素为48\*48的灰度图。FER2013数据库中一共有7中表情：愤怒，厌恶，恐惧，开心，难过，惊讶和中性。该数据库是2013年Kaggle比赛的数据，由于这个数据库大多是从网络爬虫下载的，存在一定的误差性。这个数据库的人为准确率是65% 士 5%。

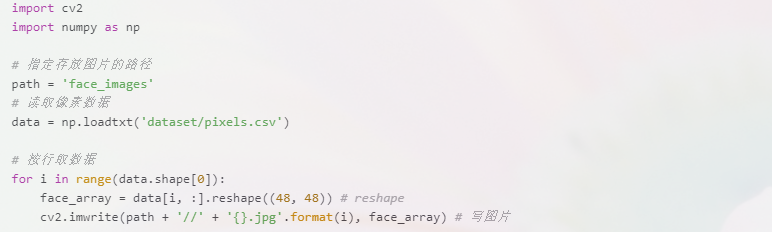
给定的数据集train.csv，我们要使用卷积神经网络CNN，根据每个样本的面部图片判断出其表情。在本项目中，表情共分7类，分别为：（0）生气，（1）厌恶，（2）恐惧，（3）高兴，（4）难过，（5）惊讶和（6）中立（即面无表情，无法归为前六类）。因此，项目实质上是一个7分类问题。



**数据集分离的代码如下：**



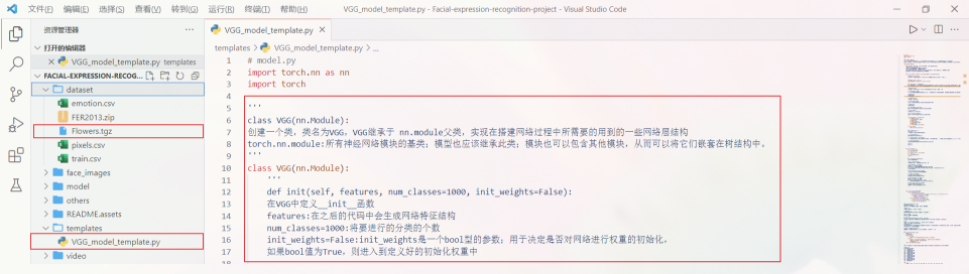
**数据可视化的代码如下：**



**训练模型代码如下：**



## **采用VGG模型优化**

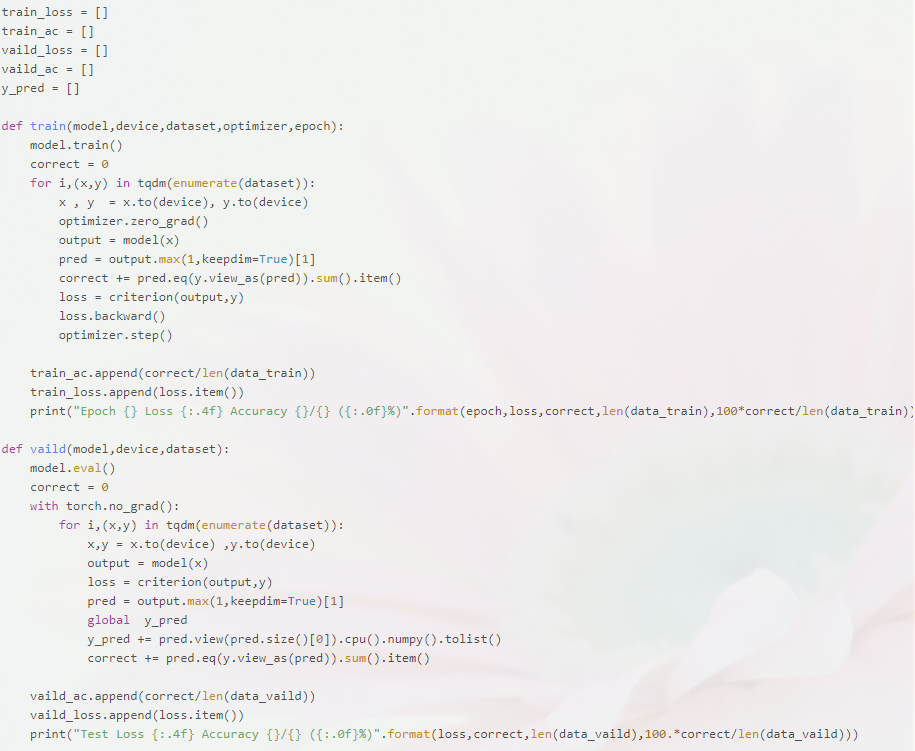


根据VGG模型的原理，我们可以通过继承nn.Module来定义我们自己的基于VGG的模型类，最后将我们自定义的VGG网络模型进行搭建。

**模型的代码如下：**



**训练模型的代码如下**：



## **采用ResNet模型优化**

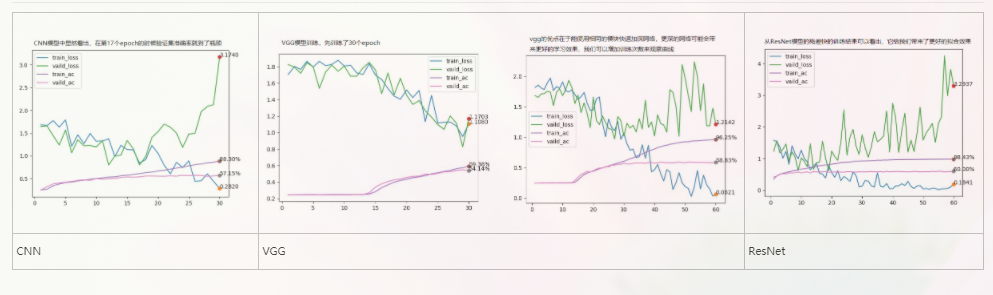
同样的，根据ResNet模型的原理，我们可以通过继承nn.Module来定义我们自己的基于ResNet的模型类，最后将我们自定义的ResNet网络模型进行搭建。

**模型的代码如下：**



训练并模型步骤和VGG的大同小异，这里不赘述。

## **模型的对比分析**



1.CNN模型中显然看出，在第17个epoch的时候验证集准确率就到了瓶颈

2.vgg的优点在于能使用相同的模块快速加深网络，更深的网络可能会带来更好的学习效果，我们可以增加训练次数来观察曲线

3.从ResNet模型的残差快的训练结果可以看出，它给我们带来了更好的拟合效果

# **参考文献**

**https://www.cnblogs.com/He-Xiang-best/p/15690252.html（小何同学的github）**

**数据集相关：**

1、FER2013数据集官网：[https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data](https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

2、Pytorch中正确设计并加载数据集方法：[https://ptorch.com/news/215.html](https://ptorch.com/news/215.html" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

3、pytorch加载自己的图像数据集实例：[http://www.cppcns.com/jiaoben/python/324744.html](http://www.cppcns.com/jiaoben/python/324744.html" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

4、python中的图像处理框架进行图像的读取和基本变换：[https://oldpan.me/archives/pytorch-transforms-opencv-scikit-image](https://oldpan.me/archives/pytorch-transforms-opencv-scikit-image" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

**CNN相关：**

1、常见CNN网络结构详解：[https://blog.csdn.net/u012897374/article/details/79199935?spm=1001.2014.3001.5506](https://blog.csdn.net/u012897374/article/details/79199935?spm=1001.2014.3001.5506" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

2、基于CNN优化模型的开源项目地址：[https://github.com/amineHorseman/facial-expression-recognition-using-cnn](https://github.com/amineHorseman/facial-expression-recognition-using-cnn" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

3、A CNN based pytorch implementation on facial expression recognition (FER2013 and CK+), achieving 73.112% (state-of-the-art) in FER2013 and 94.64% in CK+ dataset：[https://github.com/WuJie1010/Facial-Expression-Recognition.Pytorch](https://github.com/WuJie1010/Facial-Expression-Recognition.Pytorch" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

**VGG相关：**

1、VGG论文地址：[https://arxiv.org/pdf/1409.1556.pdf](https://arxiv.org/pdf/1409.1556.pdf" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

2、VGG模型详解及代码分析：[https://blog.csdn.net/weixin\_45225975/article/details/109220154#18c8c548-e4c5-24bb-e255-cd1c471af2ff](https://blog.csdn.net/weixin_45225975/article/details/109220154" \l "18c8c548-e4c5-24bb-e255-cd1c471af2ff" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

**ResNet相关：**

1、ResNet论文地址：[https://arxiv.org/pdf/1512.03385.pdf](https://arxiv.org/pdf/1512.03385.pdf" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

2、ResNet模型详解及代码分析：[https://blog.csdn.net/weixin\_44023658/article/details/105843701](https://blog.csdn.net/weixin_44023658/article/details/105843701" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

3、Batch Normalization（BN）超详细解析：[https://blog.csdn.net/weixin\_44023658/article/details/105844861](https://blog.csdn.net/weixin_44023658/article/details/105844861" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

**表情识别相关：**

1、基于卷积神经网络的面部表情识别(Pytorch实现)：[https://www.cnblogs.com/HL-space/p/10888556.html](https://www.cnblogs.com/HL-space/p/10888556.html" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

2、Fer2013 表情识别 pytorch (CNN、VGG、Resnet)：[https://www.cnblogs.com/weiba180/p/12600259.html#resnet](https://www.cnblogs.com/weiba180/p/12600259.html" \l "resnet" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)

3、OpenCV 使用 pytorch 模型 通过摄像头实时表情识别：[https://www.cnblogs.com/weiba180/p/12613764.html](https://www.cnblogs.com/weiba180/p/12613764.html" \t "https://www.cnblogs.com/He-Xiang-best/p/_blank)