**实验报告：食物分类**

**实验概述：**

分别使用：**自定义基于注意力机制的卷积网络**、预训练的**densenet121**、**mobilenet\_v2**、**resnet18**

特点：数据增强、根据验证集早停保存最优模型参数、空间注意力、通道注意力、dropout、与L2正则化等效的权重衰减、除优化自定义模型外还尝试了新旧多个预训练模型、对avif格式图片处理、保存预测结果、可视化训练过程和结果

**实验目的**

本实验旨在使用深度学习模型对五种不同类别的食物（面包、汉堡、烤串、面条、米饭）进行分类。实验数据集由全班68名同学收集，每人提供5张不同类别的食物照片，共计340张图像。

**实验数据**

数据集包含以下五个类别的食物照片：

1. 面包 (Bread)
2. 汉堡 (Hamburger)
3. 烤串 (Kebab)
4. 面条 (Noodle)
5. 米饭 (Rice)

每类食物的照片数量相等、尺寸大小和后缀格式不同。

**实验过程：**

**读取图像：**

从本地文件夹Food\_Classification\_Dataset中读取图像文件，对于train、val文件夹下的每个目录，根据图像所在的文件夹名为图像分配相应的类别标签，并将其映射为数字标签添加到labels列表，将labels列表转换为PyTorch的LongTensor类型y。

对于每个图片，使用resize((224, 224))将图像调整为224x224像素的大小。

**对于avif格式的图片，image.open并不能直接读取，还需要安装 pillow-avif库。安装了 pillow-avif 后，Pillow 才能够自动支持 AVIF 格式的读取，像处理其他图像格式一样使用 Image.open() 方法来打开 AVIF 文件。**

**数据增强、数据预处理：**

随机水平翻转图像，增加图像的水平不变性；

随机旋转图像，角度在-15到15度之间，增加图像的旋转不变性；

随机改变图像的亮度、对比度、饱和度和色调，增强模型对颜色变化的鲁棒性；

对图像的每个通道进行标准化；

验证集、测试集图像不需要增强，但仍进行必要的转换和标准化；

通过数据增强和预处理，训练数据集的图像在每个epoch中都会以不同的方式进行变换，有助于模型学习到更加鲁棒的特征，减少过拟合，提高模型在未见数据上的泛化能力。而测试集和验证集的图像则保持原样，用于评估模型的真实性能。

**模型设计：**

**模型1—自定义模型：**

使用了**卷积层**、批量归一化层、池化层、**通道注意力机制**、**空间注意力机制**及全连接层。此外，还使用了**dropout**技术来减少过拟合。

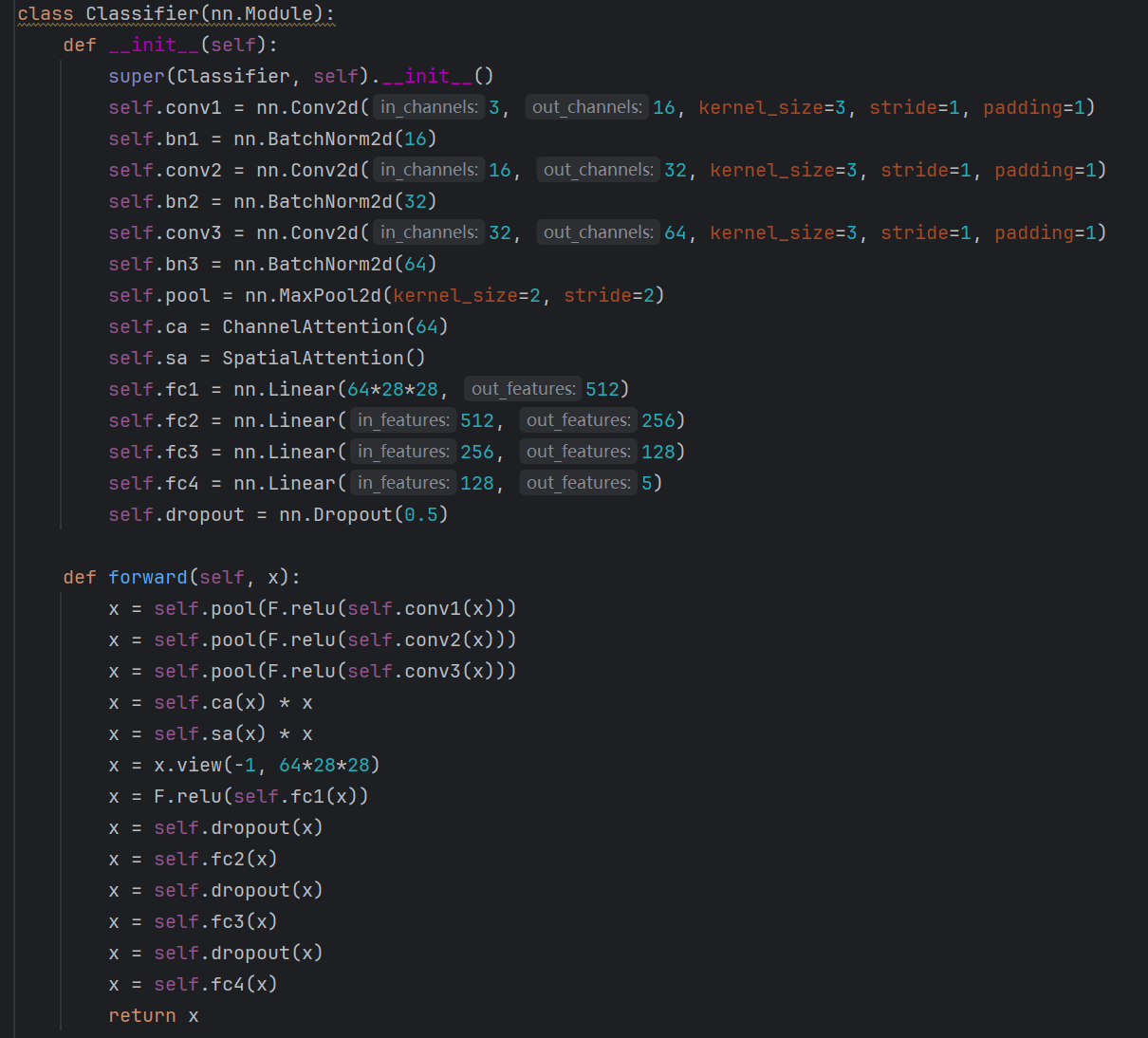
通道注意力模块的目的是在特征通道维度上应用加权，使得网络能够聚焦于更重要的特征。（实际效果一般，提升不大）

空间注意力模块是为了突出图像中的重要空间位置。（实际效果一般，提升不大）

dropout层通过在训练过程中随机丢弃一些网络连接来防止过拟合。

整个网络结构能够捕捉图像的局部特征（通过卷积层）。通道注意力和空间注意力的结合使得网络能够更加精细地调整其对特征的响应，从而提高分类的准确性。批量归一化和dropout的使用进一步增强了网络的泛化能力，使其在面对新数据时表现更加稳定。

x = self.pool(F.relu(self.conv1(x)))  
x = self.pool(F.relu(self.conv2(x)))  
x = self.pool(F.relu(self.conv3(x)))  
x = self.ca(x) \* x  
x = self.sa(x) \* x  
x = x.view(-1, 64\*28\*28)  
x = F.relu(self.fc1(x))  
x = self.dropout(x)  
x = self.fc2(x)  
x = self.dropout(x)  
x = self.fc3(x)  
x = self.dropout(x)  
x = self.fc4(x)

****

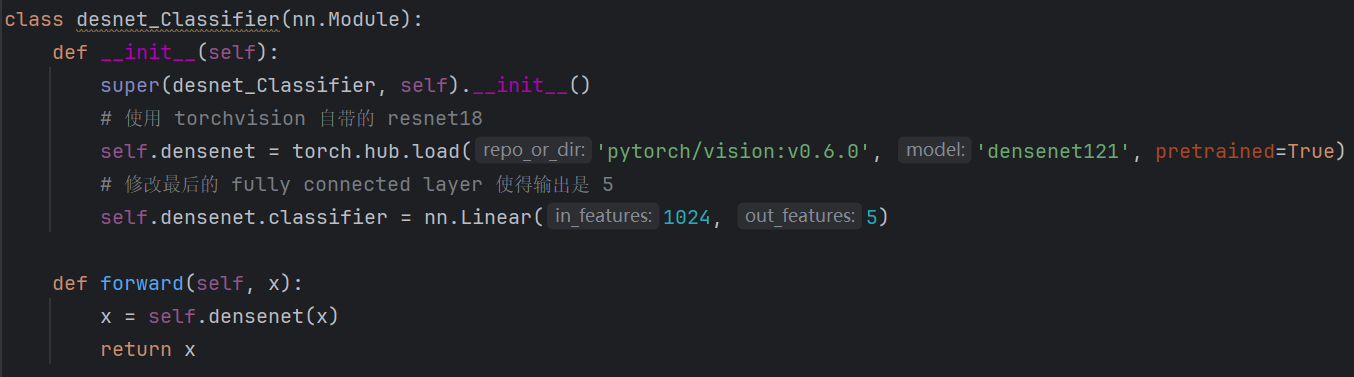
**模型2—Mobilenet：**

****

**模型3—ResNet18:**

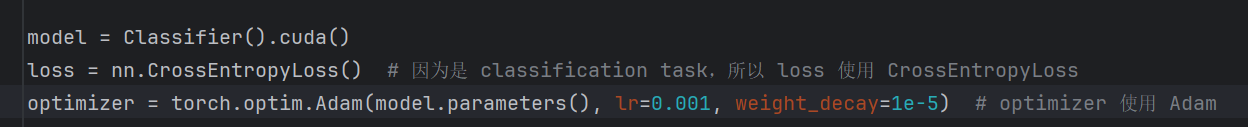
****

**模型4—Densenet121:**

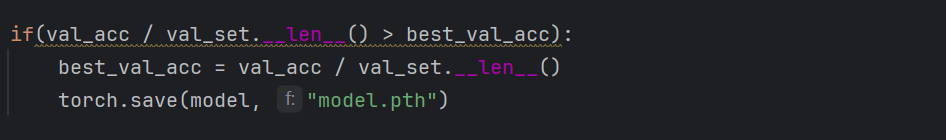
****

**训练模型：**

**权重衰减：**

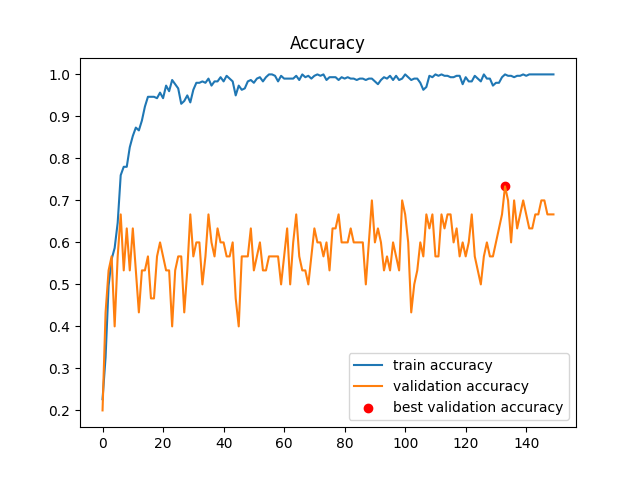
****

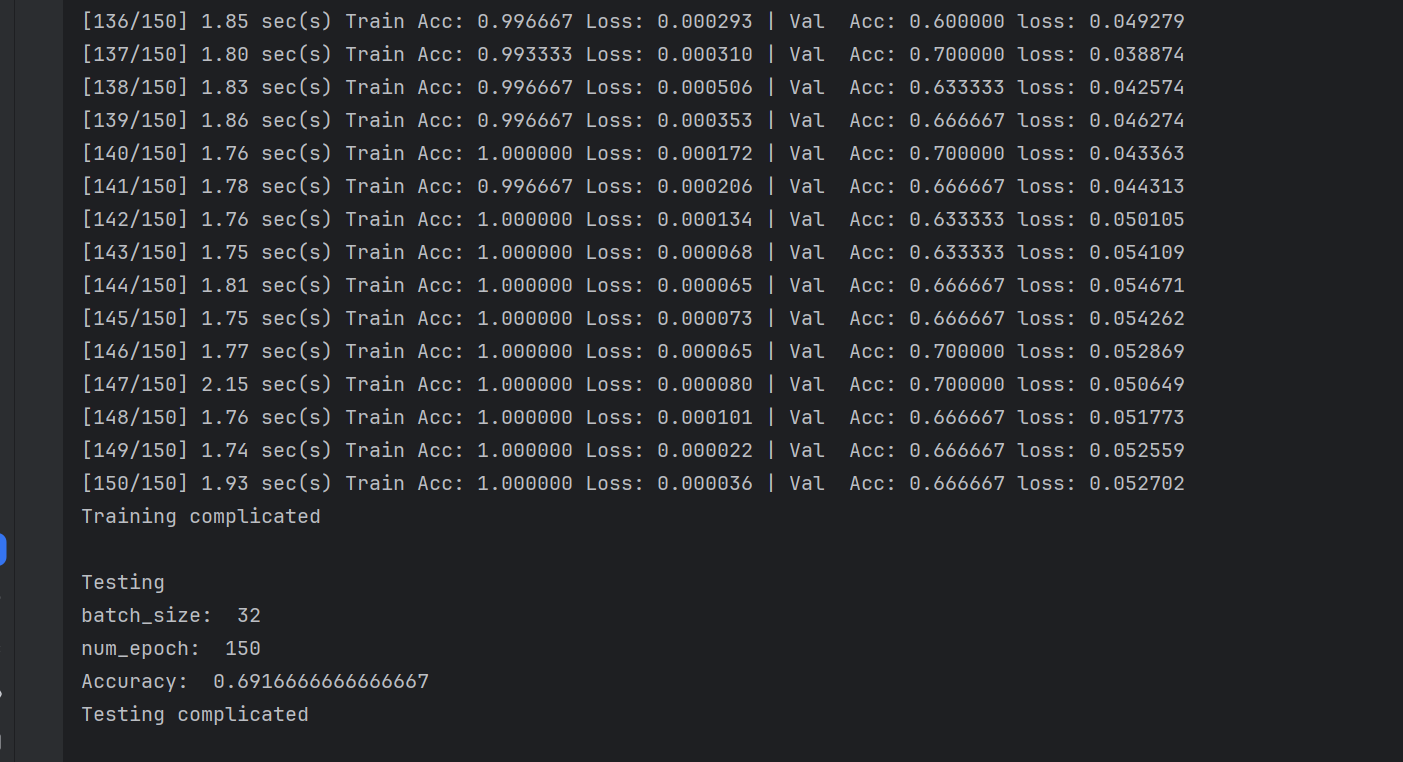
**根据验证集正确率判断最好的模型参数：**

****

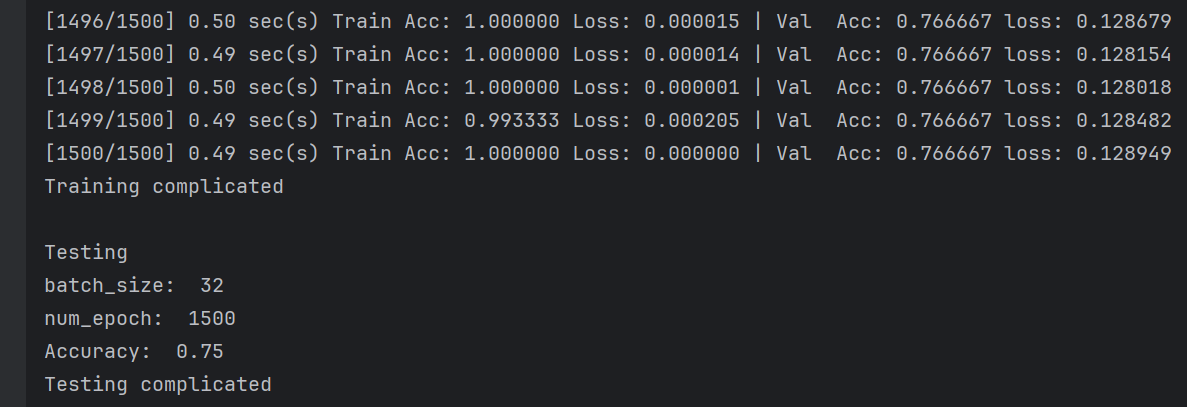
**实验结果：**

**自定义卷积网络：**

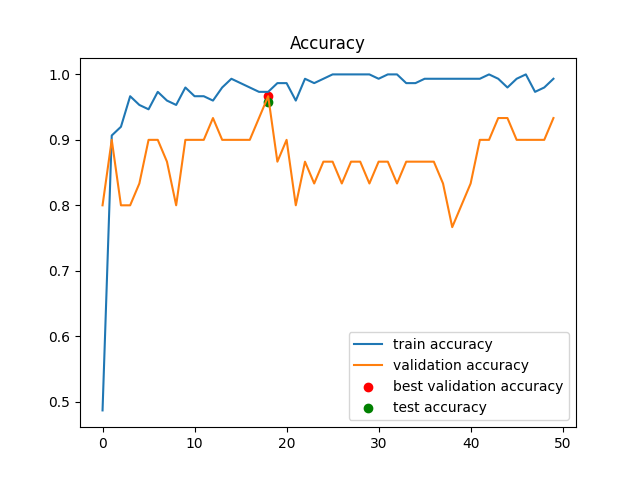
****

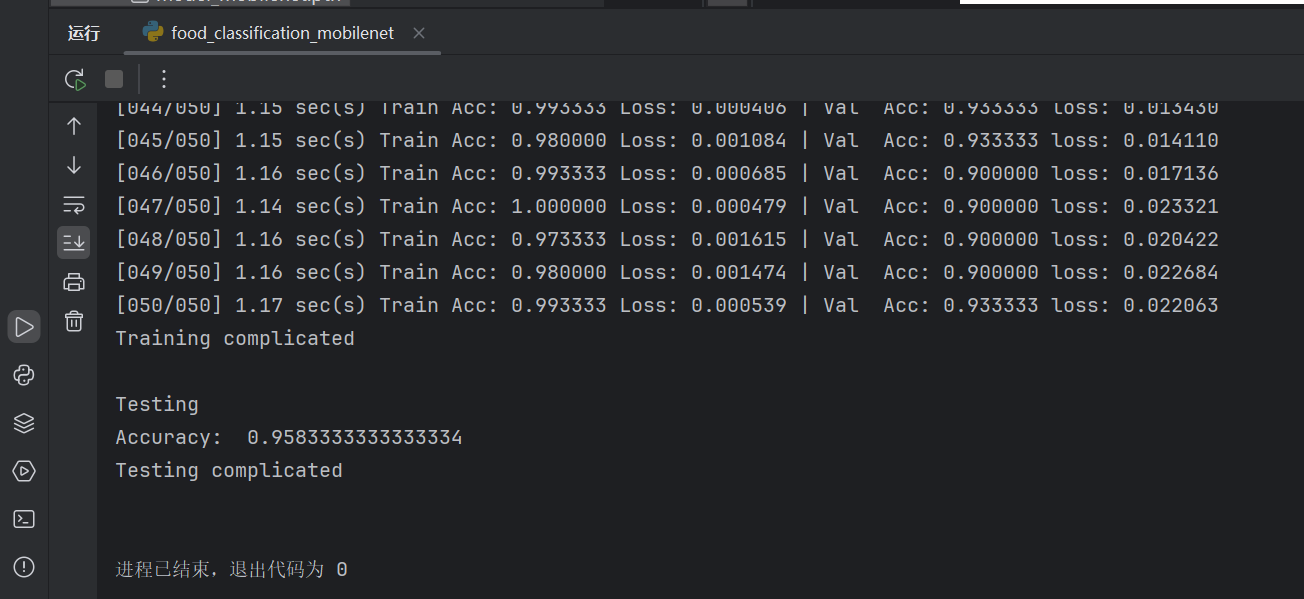
****

最优的一次：

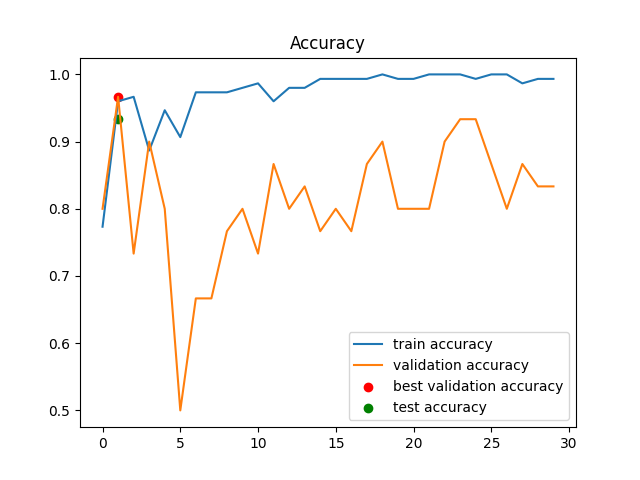


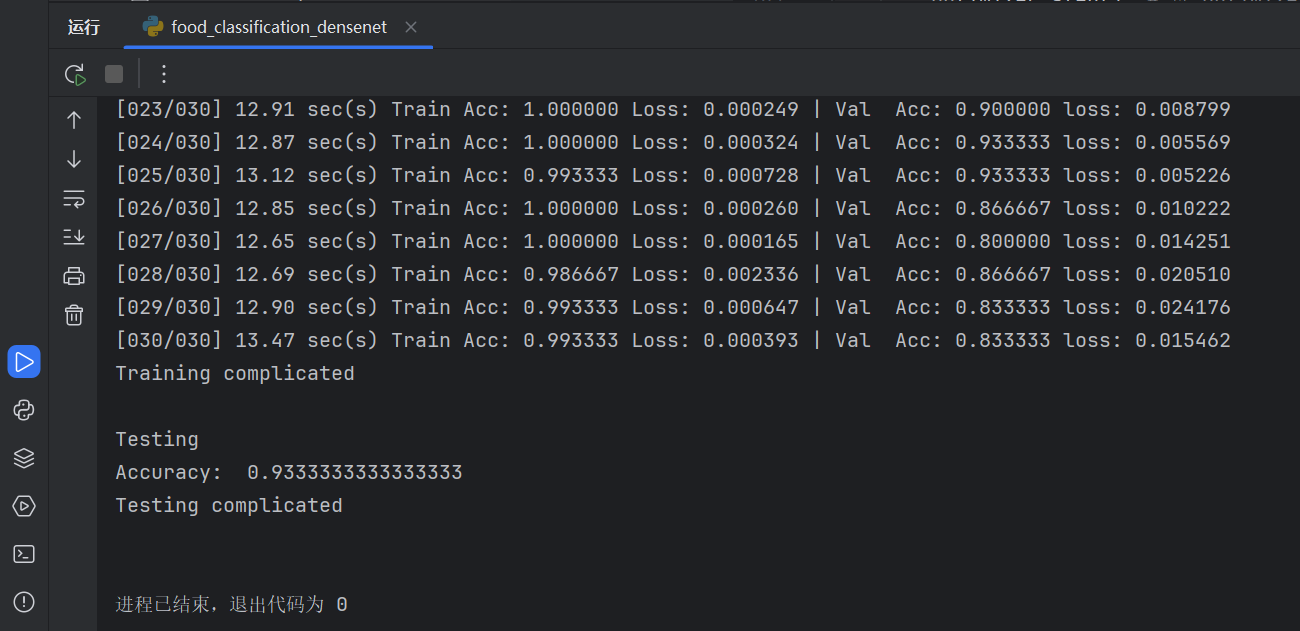
**Mobilenet\_v2:**

****

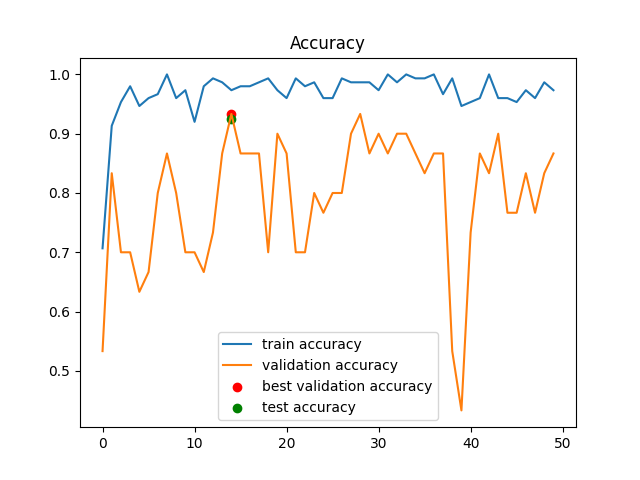
****

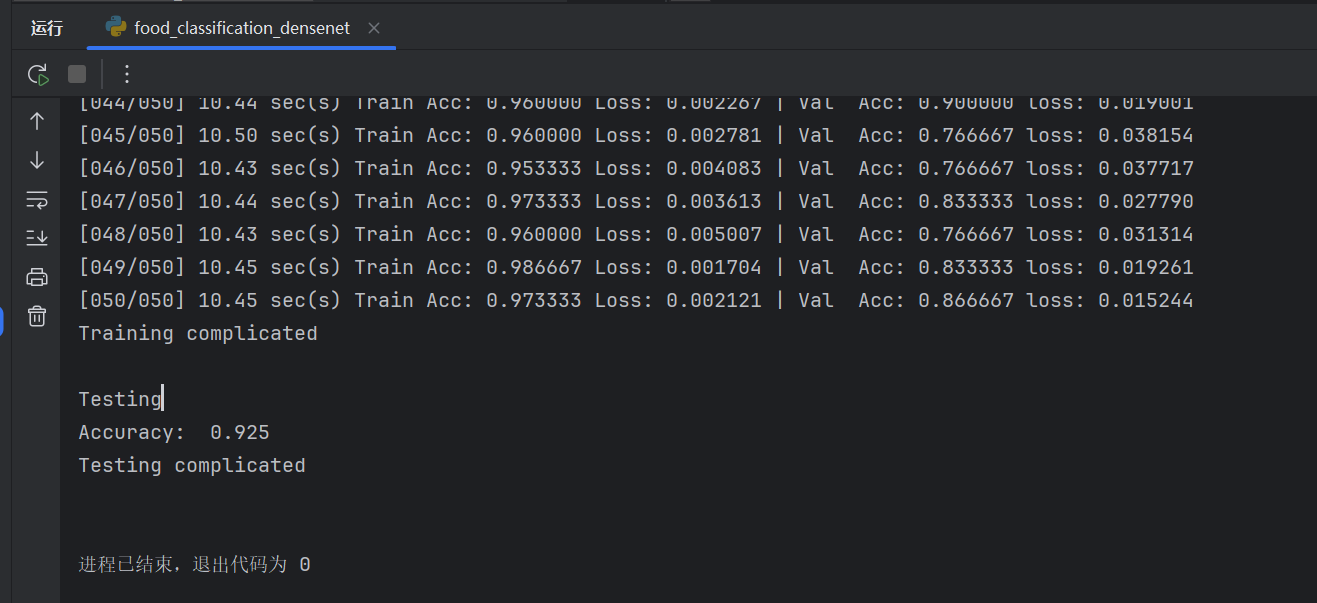
**Resnet18：**

****

****

**Densent121:**

****

****

**实验评价：**

实验结果表明，使用的深度学习模型均能够有效地对不同类别的食物进行分类。其中，Mobilenet\_v2在测试集上表现最佳，达到了95.83%的准确率，这可能得益于其轻量级架构和优秀的学习能力。Resnet18和Densenet121也展现出了较高的准确率，分别达到了93.33%和92.50%。

自定义卷积网络在测试集上的准确率相对较低，但是最高也能达到75%。考虑到自定义模型中通道注意力和空间注意力机制的实际效果一般，后续可以尝试进一步优化网络结构或调整超参数。

实验中采用了数据增强、早停法、权重衰减等技术，有效地提高了模型的泛化能力和防止过拟合。对AVIF格式图片的处理保证了数据读取的全面性。保存预测结果和可视化训练过程则为实验的复现和进一步分析提供了便利。

**后续改进：**

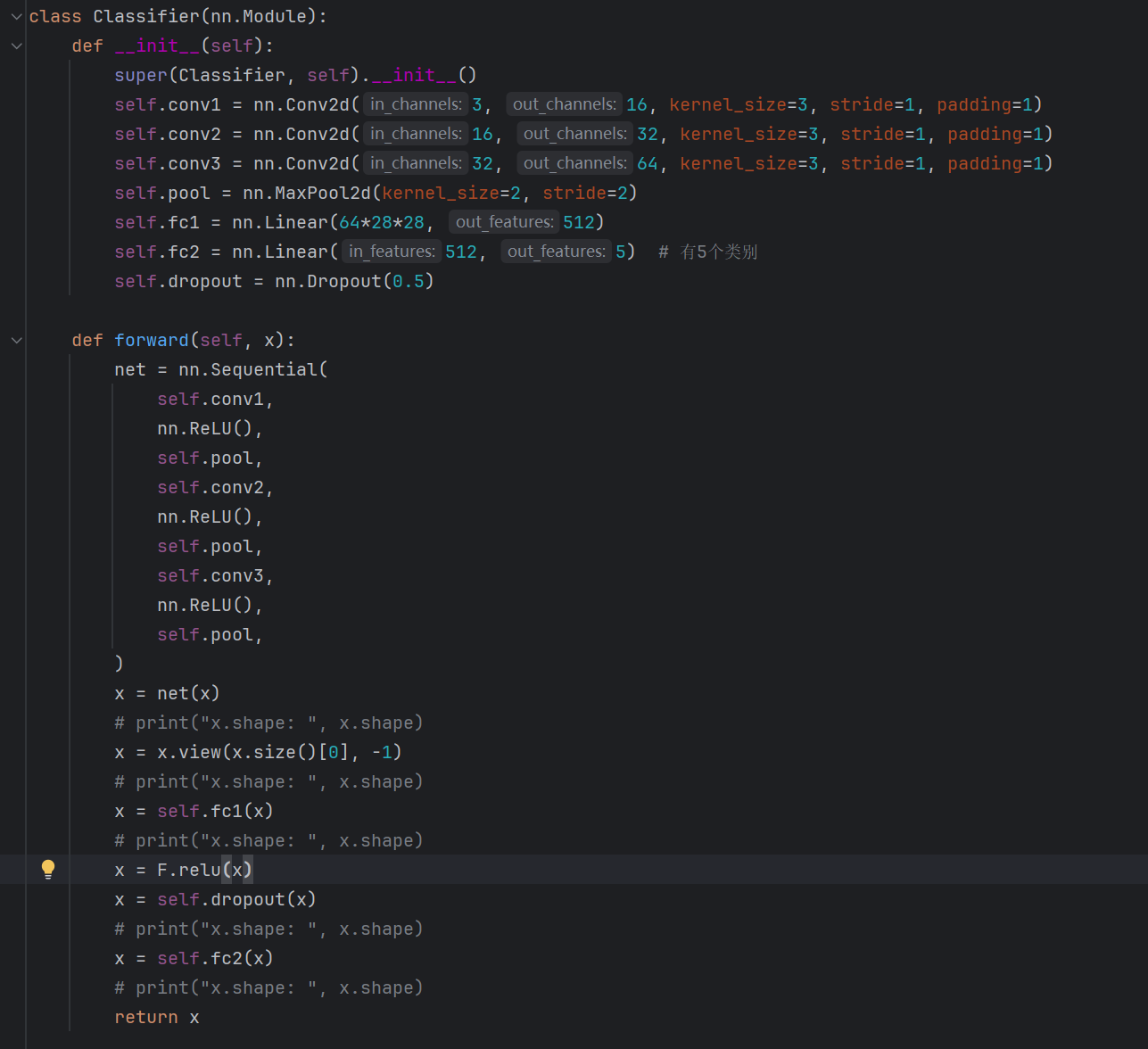
可以进一步对超参进行优化，使用**动态超参数**调整的技术等。

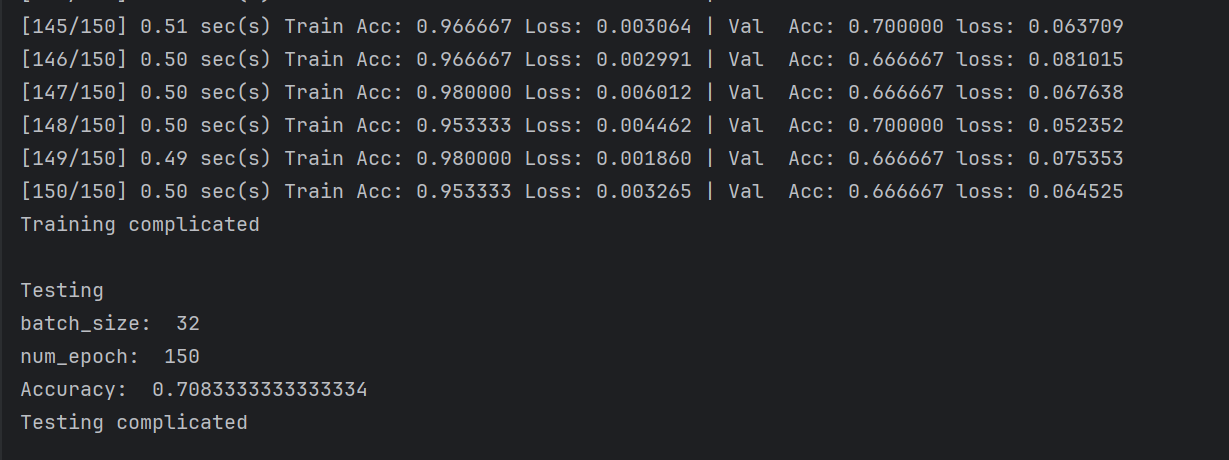
**总结：**

本实验成功展示了使用深度学习模型对食物图像进行分类的可行性，并取得了令人满意的结果。通过一系列的方法和技术，提高了模型的性能和泛化能力。未来，通过持续的改进和优化，有望进一步提升模型的分类准确率，

**自定义模型更新记录：**

Version1:





Version2、3、4:

增加了通道注意力、空间注意力、L2正则化、早停

