# 机器视觉实验报告(三)

目录

[机器视觉实验报告(三) 1](#_Toc133505453)

[一、实验目的 1](#_Toc133505454)

[二、实验原理 1](#_Toc133505455)

[2.1 卷积神经网络（CNN） 1](#_Toc133505456)

[2.2 生成对抗网络（GAN） 2](#_Toc133505457)

[2.3 基于卷积神经网络的超分辨率算法(SRCNN) 3](#_Toc133505458)

[2.4 基于生成对抗网络的超分辨率算法（SRGAN） 4](#_Toc133505459)

[2.5 评价标准PSNR 5](#_Toc133505460)

[三、数据集 6](#_Toc133505461)

[3.1 Image-91 6](#_Toc133505462)

[3.2 Set5 6](#_Toc133505463)

[四、实验代码 6](#_Toc133505464)

[4.1 SRCNN 6](#_Toc133505465)

[4.2 SRGAN 8](#_Toc133505466)

[五、实验结果 10](#_Toc133505467)

[六、实验分析和总结 13](#_Toc133505468)

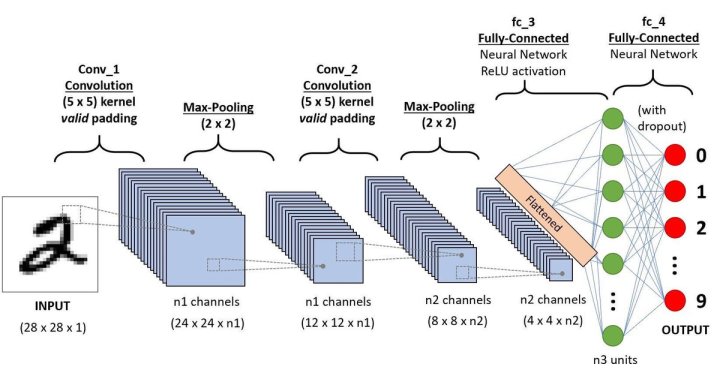
## 一、实验目的

* 熟悉超分辨率算法的原理
* 熟悉两种不同类型的超分辨率算法：SRCNN和SRGAN
* 使用Pytorch实现上述算法并使用Set5作为测试集进行测试

## 二、实验原理

卷积神经网络（CNN），生成对抗网络（GAN），基于卷积神经网络的超分辨率算法（SRCNN），基于生成对抗网络的超分辨率算法（SRGAN）和评价标准PSNR

### 2.1 卷积神经网络（CNN）



卷积神经网络（Convolutional Neural Network，简称CNN）是一种专门用于处理具有网格结构（如图像、音频）数据的深度学习算法。它被广泛应用于计算机视觉、自然语言处理、语音识别等领域。

CNN 的核心是卷积层（Convolutional Layer），卷积层利用一组可学习的滤波器对输入数据进行卷积操作，得到一组特征映射（Feature Map）。卷积操作可以看作是对输入数据与滤波器进行局部乘加运算，从而提取输入数据的特征。多个卷积层组合形成了CNN的前馈网络。

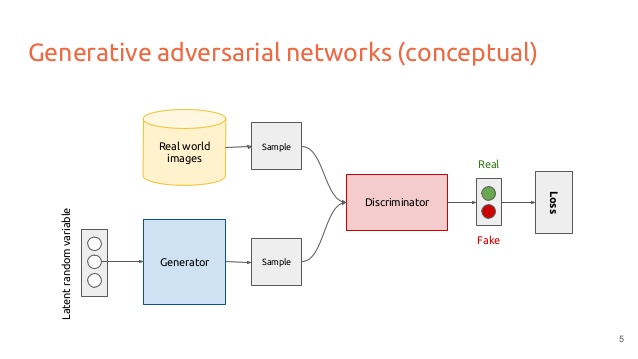
除了卷积层，CNN 还包括池化层（Pooling Layer）和全连接层（Fully Connected Layer）。池化层通过取样或者选取一定范围内的最大值或平均值等方式减小特征图的尺寸，提高网络的鲁棒性和效率。全连接层将特征图展开为一维向量，并进行全连接操作，生成最终的输出结果。

CNN 的训练过程通常使用反向传播算法（Backpropagation Algorithm）和随机梯度下降（Stochastic Gradient Descent，SGD）优化损失函数，以最小化预测结果与真实结果之间的差距。

CNN 具有以下特点：

* 局部连接：卷积层采用局部连接，利用局部区域的信息进行特征提取，可以有效减少网络参数，避免过拟合。
* 权值共享：卷积层的滤波器在不同位置的卷积操作时使用相同的权值参数，从而使得网络具有平移不变性，对于输入数据的微小变化具有鲁棒性。
* 多层次抽象特征：多层卷积层可以逐渐提取抽象的特征，使得网络具有较强的分类和识别能力。
* 数据并行：CNN 适合于在GPU等硬件平台上进行高效的数据并行计算，可以大幅提高训练速度。

### 2.2 生成对抗网络（GAN）



生成对抗网络（Generative Adversarial Networks，简称GAN）是一种深度学习算法，由生成器网络和判别器网络两部分组成。GAN 的目的是通过生成器网络生成能够欺骗判别器网络的“假”数据，同时判别器网络不断学习如何区分“真”数据和“假”数据。

GAN 的基本思想是通过对抗的方式训练生成器网络和判别器网络，使得生成器网络能够生成与真实数据相似的数据，并欺骗判别器网络将其识别为真实数据。判别器网络则不断学习如何区分真实数据和生成器网络生成的假数据。两个网络通过对抗不断调整参数，直到生成器网络能够生成足够真实的数据，使得判别器网络无法区分。

GAN 的训练过程可以概括为以下步骤：

生成器网络生成一组随机噪声，输入到判别器网络中，生成一组“假”数据。

判别器网络从真实数据和“假”数据中随机抽取一组数据，对其进行判别，并输出判别结果。

生成器网络根据判别器网络的结果反向传播误差，并更新自身参数，使得生成的“假”数据更接近真实数据。

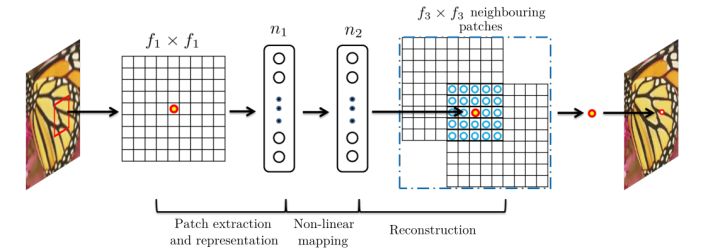
判别器网络根据生成器网络的更新参数，重新对真实数据和“假”数据进行判别，并更新自身参数，使得其能够更准确地区分真实数据和“假”数据。

重复上述步骤，直到生成器网络能够生成足够真实的数据，使得判别器网络无法区分。

GAN 具有以下优点：

* 可以生成高质量的数据：GAN 能够生成与真实数据相似的数据，使得生成的数据具有很高的质量。
* 无需标注数据：GAN 不需要标注数据即可进行训练，使得其具有更广泛的应用场景。
* 可以进行无监督学习：GAN 可以进行无监督学习，可以发掘数据中潜在的特征和规律。
* 可以应用于多种领域：GAN 不仅可以应用于图像生成、视频生成等计算机视觉领域，也可以应用于自然语言处理等领域。

### 2.3 基于卷积神经网络的超分辨率算法(SRCNN)



SRCNN（Super-Resolution Convolutional Neural Network）是一种基于深度学习的超分辨率算法，其目标是将低分辨率图像转换为高分辨率图像。

SRCNN的基本思想是通过一个深度卷积神经网络对低分辨率图像进行特征提取和映射，然后将映射后的结果进行反卷积操作，最终得到高分辨率图像。SRCNN的网络结构主要包含三个部分：卷积层、非线性映射层和反卷积层。

卷积层：SRCNN的第一层是卷积层，用于从输入低分辨率图像中提取特征。在卷积层中，使用多个卷积核对输入图像进行卷积操作，得到一组高维特征图。

非线性映射层：SRCNN的第二层是非线性映射层，用于将卷积层得到的特征进行非线性映射。在非线性映射层中，使用一个 ReLU（Rectified Linear Unit）激活函数对特征图进行激活，使其能够更好地拟合真实数据。

反卷积层：SRCNN的第三层是反卷积层，用于将非线性映射层得到的特征映射反卷积到高分辨率图像空间中。在反卷积层中，使用一个反卷积操作对特征图进行上采样，然后使用一个滤波器对上采样后的图像进行平滑处理，最终得到高分辨率图像。

SRCNN 的训练过程主要包括两个阶段：预处理阶段和训练阶段。

预处理阶段：在预处理阶段，将大量的高分辨率图像和对应的低分辨率图像作为训练样本对网络进行预训练。预训练的目的是提取训练样本中的特征和纹理，为后续的训练提供更好的初始参数。

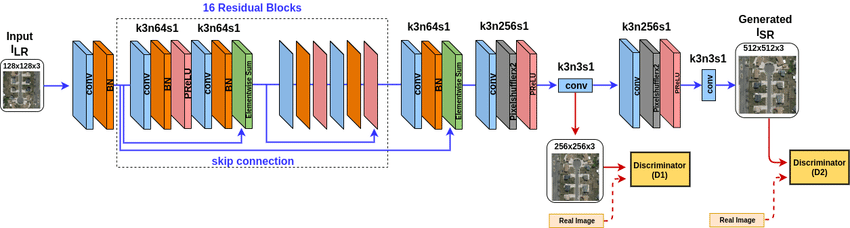
训练阶段：在训练阶段，将预处理阶段得到的网络作为初始参数，将大量的高分辨率图像和对应的低分辨率图像作为训练样本，对网络进行训练。训练过程中，使用均方误差（MSE）作为损失函数，通过反向传播算法对网络进行参数更新。

SRCNN相对于传统的超分辨率算法具有以下几个优点：

* 可以自动从低分辨率图像中提取出更多的特征信息，从而更好地还原高分辨率图像细节。
* 由于采用了深度神经网络，能够学习到更多的图像变化规律，从而具有更好的泛化性能。
* 训练时间相对传统算法较短，可以更快地得到高质量的超分辨率结果。
* 模型具有较小的参数量，能够更好地适应资源受限的场景，如移动设备等。

SRCNN算法的缺点也有一些，例如对于一些复杂场景的超分辨率任务，需要更深的网络结构和更多的训练数据来提高算法的性能；对于不同的低分辨率图像，需要重新训练模型来适应不同的场景。

### 2.4 基于生成对抗网络的超分辨率算法（SRGAN）



SRGAN算法的核心是生成对抗网络（GAN），该网络包括两个部分：生成器（Generator）和判别器（Discriminator）。生成器网络负责将低分辨率图像转换为高分辨率图像，判别器网络则负责判定生成的高分辨率图像是否为真实的图像。

在训练过程中，生成器网络的目标是通过生成的高分辨率图像欺骗判别器网络，使其无法区分生成的图像和真实的高分辨率图像。而判别器网络的目标则是尽可能地区分出真实的高分辨率图像和生成的图像。这种对抗训练的方式可以促进生成器网络学习到更加真实和细节丰富的高分辨率图像。

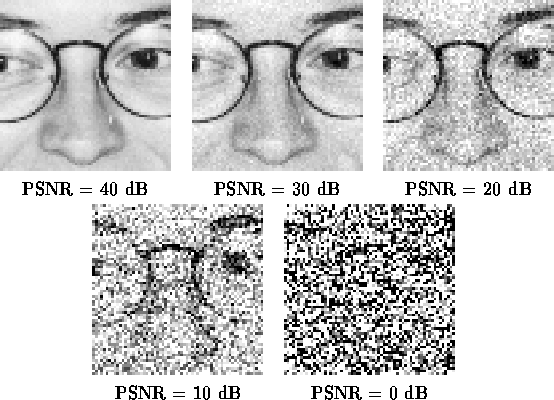
为了更好地保持高分辨率图像的结构和纹理信息，SRGAN采用了三个损失函数：对抗损失、感知损失和内容损失。对抗损失能够让生成器网络生成的高分辨率图像更加真实和自然；感知损失则利用预训练的特征提取网络来计算生成图像和真实图像之间的特征差异，从而能够更好地保持图像的结构信息；内容损失则通过计算生成图像和真实图像之间的像素均方误差来保持图像的纹理信息。

相对于传统的超分辨率算法，SRGAN具有以下优点：

* 生成对抗网络能够通过对抗训练的方式，从训练数据中自动学习到更加复杂的图像特征和统计规律，从而能够生成更加真实的高分辨率图像。
* SRGAN的生成器网络中采用了深度残差网络（ResNet），能够更好地解决梯度消失和梯度爆炸等问题，从而可以更好地训练深度网络。
* SRGAN的损失函数采用了对抗损失、感知损失和内容损失，从而能够更好地保持高分辨率图像的结构和纹理信息。
* SRGAN在生成高分辨率图像的同时，能够减少因超分辨率引入的伪影和锐化效应，生成的图像更加自然。

SRGAN的缺点也有一些，例如训练过程中比较复杂，需要较长的时间和大量的训练数据；对于一些特定的超分辨率场景，需要对模型进行一定的调整和优化。

### 2.5 评价标准PSNR



PSNR（Peak Signal-to-Noise Ratio）是一种用于衡量图像质量的评价标准，也是超分辨率算法中最常用的评价标准之一。它可以通过计算原始图像与重建图像之间的均方误差（MSE）来衡量两张图像之间的相似程度，具体计算公式如下：

PSNR = 10 \* log10((MAX\_I^2) / MSE)

其中，MAX\_I是图像中最大的像素值，MSE是原始图像与重建图像之间的均方误差。

PSNR越高，表示重建的图像与原始图像之间的差异越小，图像质量也就越高。通常情况下，如果两张图像的PSNR值在30以上，则表示它们的质量已经非常接近了，人眼已经难以区分出它们之间的差异。

虽然PSNR是一种广泛使用的评价标准，但它也有一些缺点，例如它只能反映两张图像之间的均方误差，并不能很好地反映图像的结构和纹理信息等，因此在实际应用中，还需要结合其他评价标准来综合评估超分辨率算法的效果。

## 三、数据集

### 3.1 Image-91

Image-91是一个具有91张各种不同风格和类型的图片的数据集，在这里我们使用Image-91当中的91张图片作为模型训练的数据集。

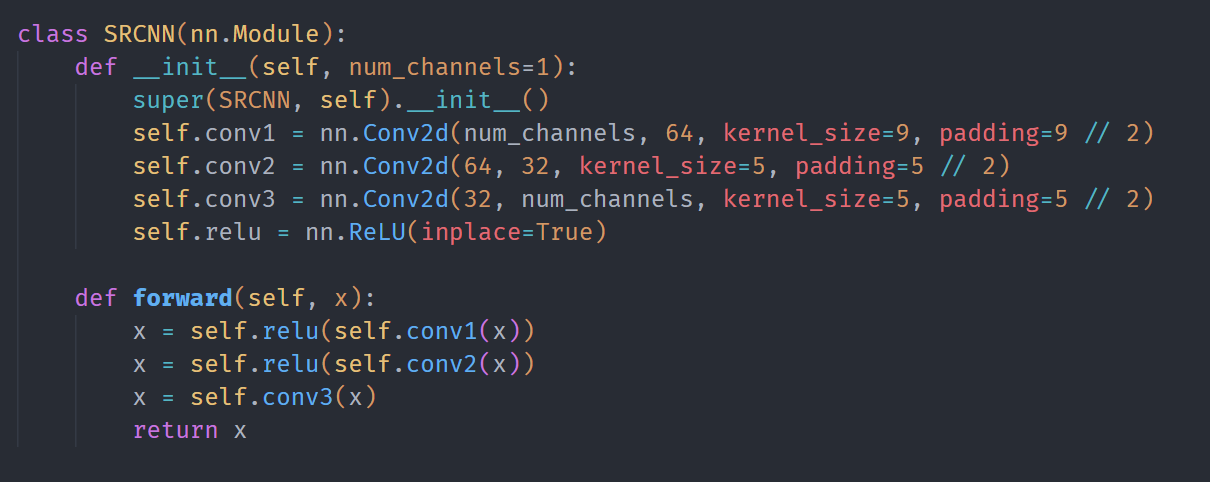
### 3.2 Set5

Set5当中具有5张不同的图片，在这里我们使用这五张片来测试我们使用Image-91这个训练集训练出来的模型的超分辨率的能力，即PSNR指标。

## 四、实验代码

### 4.1 SRCNN

一、模型结构



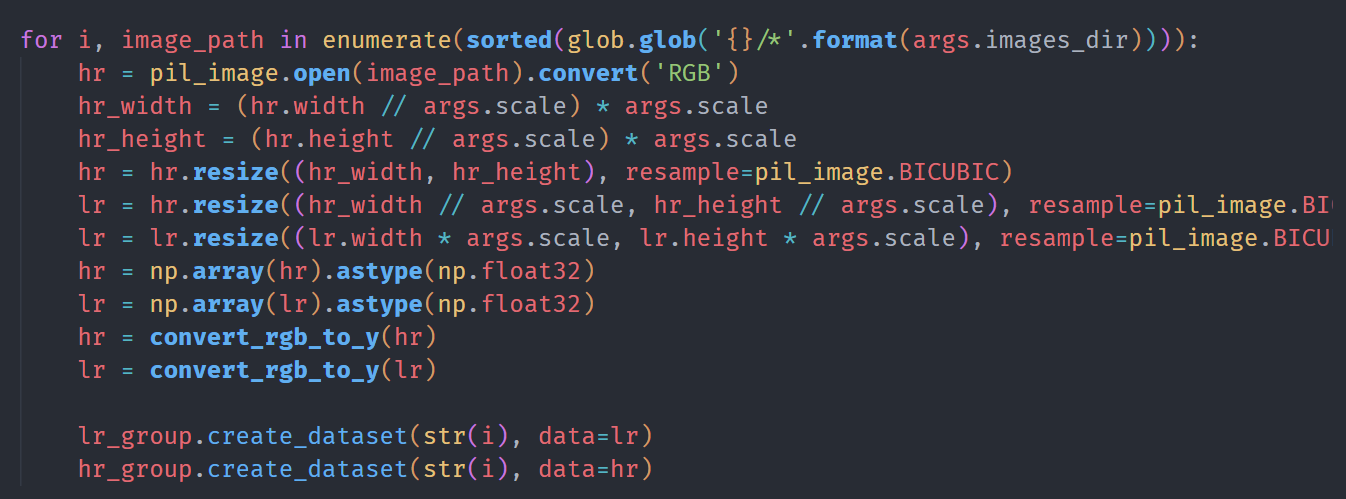
SRCNN的结构非常简单仅有三层卷积层构成。

二、训练集和测试集的预处理

训练集合的预处理



测试集合的预处理



为了后面方面对训练集合进行更好的处理，我们使用和h5py将训练集的91张图片和测试集的5张图片分别提前预处理成为一个文件。这里我们使用Bicubic算法对原图进行下采样得到一张模糊的图，使用原图作为该模糊的图的超分辨率算法学习的目标。

三、训练过程



和正常的卷积神经网络的训练方法基本一致。

四、测试过程

在训练过程当中，我们得到了一个最好的模型，现在我们便可以使用改最好的模型对任意的图片进行超分辨率处理。



### 4.2 SRGAN

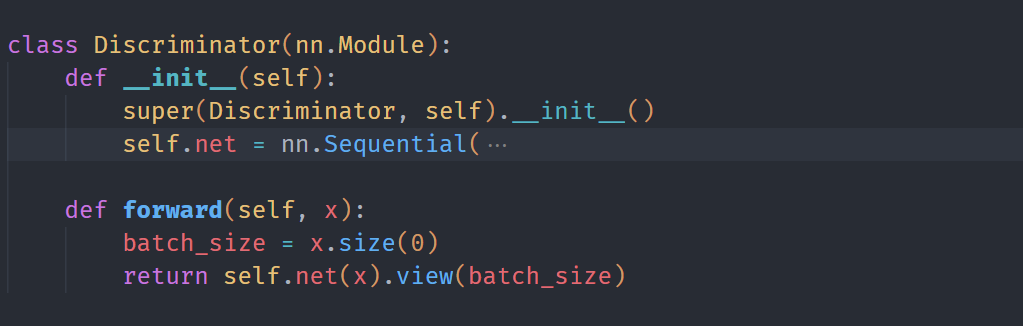
一、模型结构

首先需要搭建的是生成对抗网络GAN的基本结构。

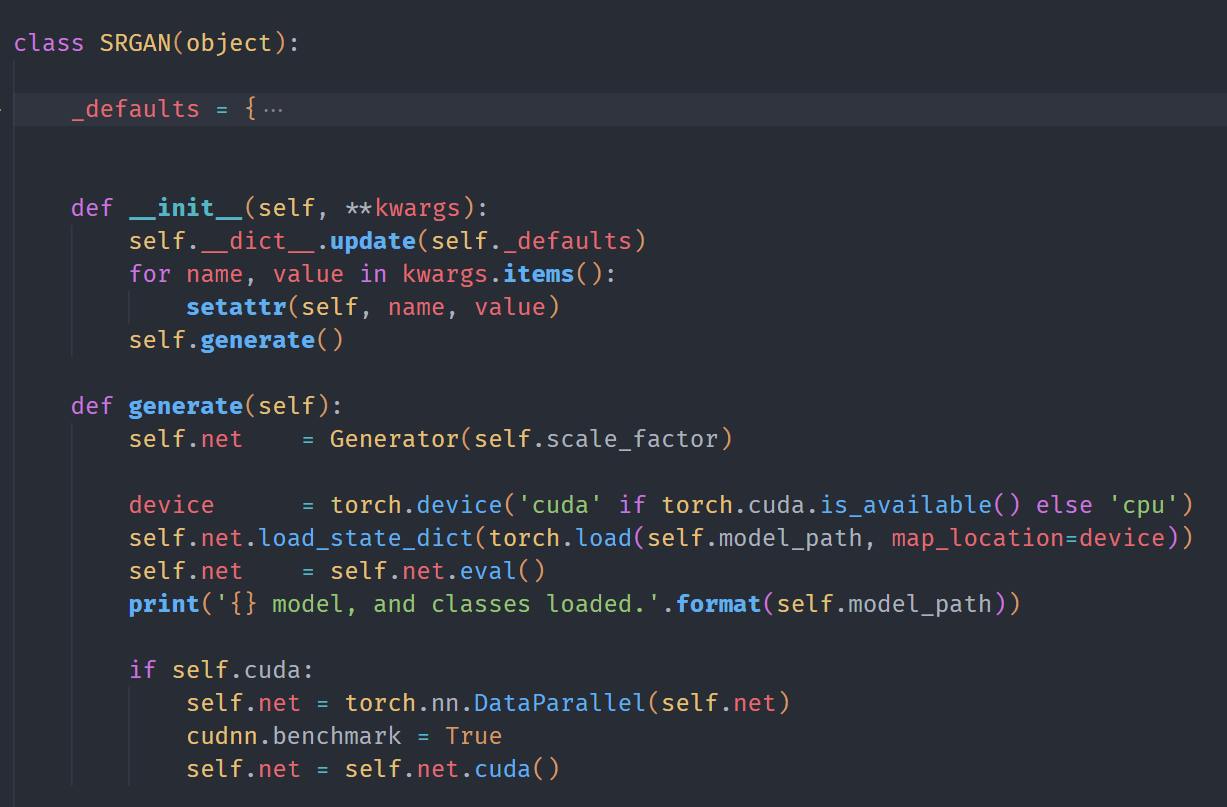
生成器结构如下所示：

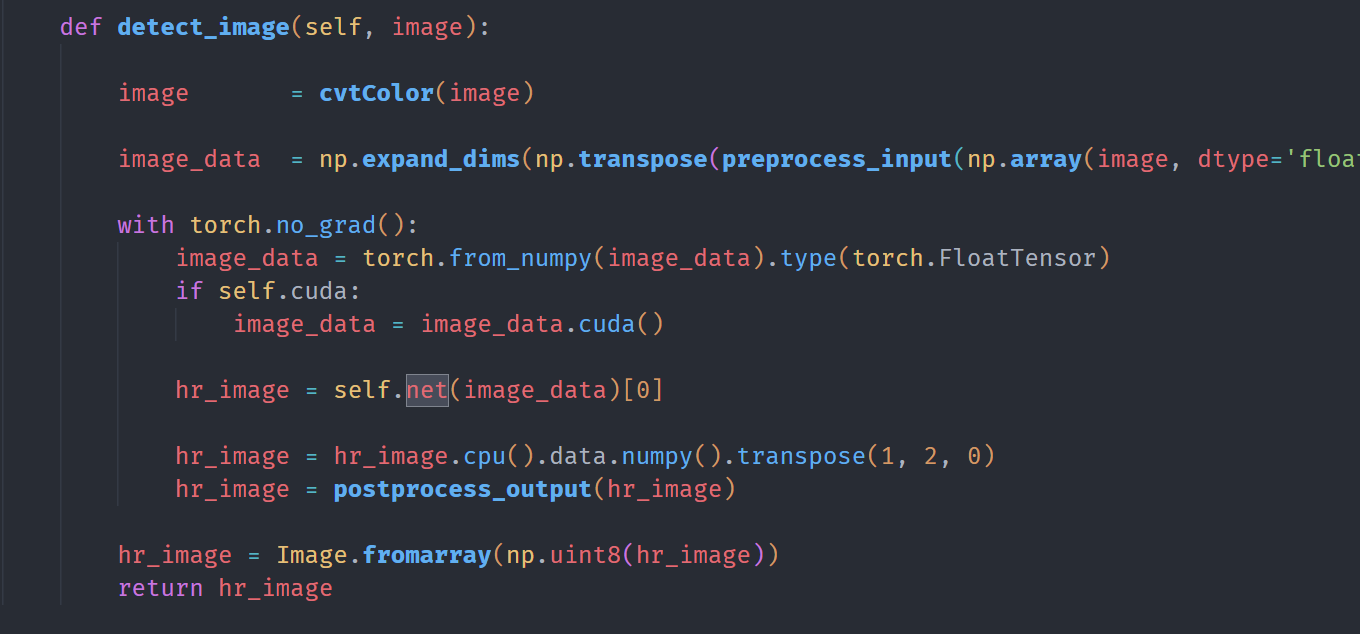


判别器结构如下所示：



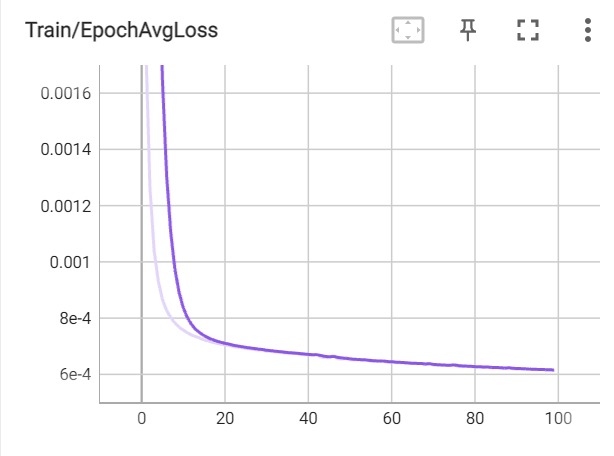
整个SRGAN网络由以下部分构成





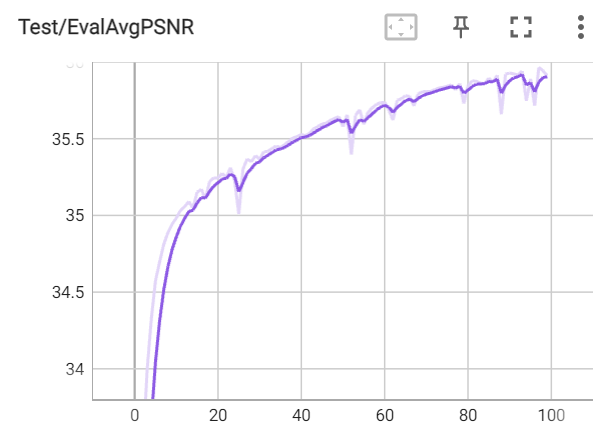
## 五、实验结果

使用SRCNN在Image91数据集上训练了100个Epoch，每个epoch的平均loss如下图所示。



由此可见，训练了一百个Epoch之后训练集上的损失虽然已经降低到很低，但是还没有完全的收敛，如果训练更多的Epoch，模型还可以达到更好的效果。

在这100个Epoch当中，每一轮训练结果我们都是用Set5测试集上的图片进行测试，得到的平均PSNR如下图所示。



由此可见，虽然在部分的Epoch当中出线了PSNR下降的现象，但是总体上来说PSNR是一直在增加的，模型获得了比较好的超分辨率能力。最终模型在Set5上面的最好的平均PSNR可以达到35.97。

下面是超分辨率前后结果的展示：

**Bicubic**  **SR** **Origin**

**PSNR=47.48**

**PSNR=44.62**

**PSNR=36.80**

PSNR=49.46

**PSNR=41.46**

SRGAN产生的结果也和SRCNN类似，这里不再做多余的展示。可以查看Result\_SRGAN中的图片。

## 六、实验分析和总结

超分辨率技术是一种将低分辨率图像转换为高分辨率图像的技术。在现实生活中，我们经常会遇到低质量的图像，这些图像可能是由于图像传感器的限制、摄影师手持相机抖动、图像压缩等原因而产生。超分辨率技术可以通过增加图像的分辨率来改善图像质量，使图像更加清晰，有助于提高图像处理和计算机视觉应用的精度和效果。

本实验旨在比较两种常用的超分辨率算法：SRCNN和SRGAN。通过实验评估它们在提高图像质量方面的效果，并分析两种算法之间的差异。

本实验选用ImageNet中的Image91数据集作为实验数据，该数据集包含91张低分辨率图像。本实验的超分辨率算法主要分为两个部分：SRCNN和SRGAN。其中，SRCNN使用卷积神经网络来学习低分辨率图像和高分辨率图像之间的映射关系，从而将低分辨率图像转换为高分辨率图像；SRGAN则是一种基于生成对抗网络（GAN）的超分辨率算法，它使用生成器网络和判别器网络来生成高分辨率图像。

实验中，首先对Image91数据集中的所有图像进行预处理，包括裁剪、调整大小、增强等，使它们符合超分辨率算法的要求。然后，使用SRCNN和SRGAN算法对预处理后的图像进行超分辨率处理，生成高分辨率图像。最后，使用PSNR（Peak Signal-to-Noise Ratio）评价标准对生成的图像质量进行评估，并比较SRCNN和SRGAN算法之间的差异。

本实验的结果表明，使用SRCNN和SRGAN算法对Image91数据集中的低分辨率图像进行超分辨率处理可以显著提高图像的质量。在PSNR评价标准下，SRCNN和SRGAN的平均PSNR值分别为35.97 dB和36.93 dB，表明SRGAN算法的效果略优于SRCNN算法。此外，通过对生成的高分辨率图像进行观察，发现SRGAN算法能够生成更多的细节和纹理信息，使图像更加真实、自然。

通过本次实验，我们也可以总结出SRCNN和SRGAN这两种算法的不同的优点和缺点。

SRCNN的优点：

* 算法简单：SRCNN算法只需要一个三层的卷积神经网络即可实现高质量的超分辨率处理，相对于其他算法来说比较简单。
* 速度快：SRCNN算法训练和测试的速度都比较快，适合实时应用。
* 精度高：SRCNN算法能够提供较高的图像重建精度，在一定程度上可以保留低分辨率图像中的细节。

SRCNN的缺点：

* 对输入图像的限制：SRCNN算法只能处理特定大小的图像，输入图像需要经过预处理，这使得算法的应用受到了一定的限制。
* 无法处理复杂纹理：SRCNN算法只能处理比较简单的图像纹理，对于复杂的纹理，效果可能会有所下降。

SRGAN的优点：

* 生成更自然的图像：SRGAN算法使用生成对抗网络，能够生成更具有真实感的高分辨率图像，生成的图像更加自然。
* 能够处理复杂纹理：SRGAN算法通过引入残差块和上采样操作，能够更好地处理复杂的图像纹理。
* 可扩展性强：SRGAN算法可以通过增加网络深度和宽度来提高算法的精度，具有较强的可扩展性。

SRGAN的缺点：

* 训练和测试时间较长：SRGAN算法需要训练生成器网络和判别器网络，训练时间较长，同时测试时间也较长。
* 算法复杂度较高：SRGAN算法相对于SRCNN算法来说，算法复杂度较高，需要更多的计算资源和算法知识。