北京郵電大學

本 科 毕 业 设 计（论文）

题目：将卷积神经网络应用于子像素的图像超分辨率实现

姓　　名　王灏

学　　院　信息与通信工程学院

专　　业　通信工程

班　　级　2014211121

学　　号　2014210592

班内序号　01

指导教师　门爱东

2018 年 6 月

北 京 邮 电 大 学

本科毕业设计（论文）诚信声明

本人声明所呈交的毕业设计（论文），题目《将卷积神经网络应用于子像素的图像超分辨率实现》是本人在指导教师的指导下，独立进行研究工作所取得的成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京邮电大学或其他教育机构的学位或证书而使用过的材料。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

本人签名：_____ 日期：_____

将卷积神经网络应用于子像素的图像超分辨率实现

摘　要

在实际应用中，受限于成像系统、传输网络等硬件的物理性能，人们通常无法获得足够高分辨率的图像。通过改良硬件设备来提升图像分辨率的方法的成本较高，因此，从软件的角度，利用图像超分辨率重建方法解决图像分辨率提升问题有更重要的现实意义和极大的应用空间。

本文对图像超分辨率重建方法的已有成果进行调研，从单图像超分辨率重建的原理出发，重点分析将卷积神经网络运用于图像超分辨率重建中的优势。本文将卷积神经网络与子像素方法相结合，搭建图像超分辨率重建网络。本文在不同的图像放缩比例设定情况下，对搭建的图像超分辨率重建网络进行测试。在重建图像质量和重建速度方面，本文的超分辨率重建网络达到了与已有研究成果可比或更优的水平。在实现单图像超分辨率重建的基础上，参考实现单图像超分辨率重建的思想，本文实现视频的超分辨率重建及其实时输出。

关键词　超分辨率　子像素　卷积神经网络

REALIZATION OF IMAGE SUPER-RESOLUTION BY IMPLEMENTING CONVOLUTIONAL NEURAL NETWORK ON SUB-PIXEL

ABSTRACT

In practical applications, it is difficult to obtain high-resolution images due to the limitation of the physical properties of hardware such as imaging systems and transmission networks. It is costly to improve image resolution by improving the hardware device. Therefore, from the software point of view, applying the image super-resolution has more important practical significance and a wider range of application.

In this paper, the existing results of image super-resolution method are investigated. Starting from the principle of single-image super-resolution, the advantages of applying convolutional neural network to image super-resolution are analyzed. In this paper, convolutional neural networks are combined with sub-pixel methods to build an image super-resolution network. The image super-resolution reconstruction network is tested under different image scaling settings. In terms of reconstructing image quality and rebuilding speed, the super-resolution network in this paper has reached a level comparable to or better than existing research results. Based on the realization of single-image super-resolution and reference to the idea of single-image super-resolution reconstruction, the super-resolution reconstruction of video and its real-time output is realized in this paper.

KEY WORDS super-resolution sub-pixel convolutional neural network

目　　录

第一章　绪论 ……1
1.1　图像超分辨率的研究背景和意义 ……1
1.2　图像超分辨率的内容和技术原理 ……2
1.3　本文的研究内容 ……3
1.4　论文的组织结构 ……3
第二章　图像超分辨率重建的研究现状 ……4
2.1　传统的重建方法 ……4
2.2　基于卷积网络的重建方法 ……4
2.3　图像超分辨率的发展趋势 ……5
第三章　将卷积神经网络应用于子像素的图像超分辨率算法 ……7
3.1　背景 ……7
3.2　网络结构 ……7
3.2.1　卷积神经网络 ……7
3.2.2　子像素卷积方法 ……8
3.2.3　子像素卷积层 ……8
3.2.4　图像缩小过程 ……9
3.2.5　主要重建结构 ……9
3.3　视频的超分辨率重建 ……9
3.4　本章小结 ……10
第四章　图像超分辨率重建实验 ……11
4.1　数据集 ……11
4.2　参数设置 ……11
4.3　图像超分辨率重建系统的性能 ……12
4.3.1　超分辨率图像的重建质量 ……12
3.2.2　超分辨率图像的重建速度 ……19
4.4　本章小结 ……21
第五章　总结与展望 ……22
5.1　论文工作总结 ……22
5.2　未来工作展望 ……22
参考文献 ……24
致　　谢 ……26

第一章　绪论

随着现代科学技术的迅猛发展，人们在日常生活中接收到的信息越来越多，对信息质量的要求越来越高。研究表明，视觉是人类最高级的感觉，所以，图像在人类的感知中扮演无可争议的重要角色[1]，是传递信息的重要载体。因此，图像信息不仅具有深刻的研究意义，同时，处理图像信息也具有非常重要的现实意义和十分巨大的应用价值。通过研究图像，挖掘图像中蕴含的特征信息，可以更好地利用图像解决现实生活中的诸多应用问题。

1.1　图像超分辨率的研究背景和意义

图像分辨率，指图像中存储的信息量，描述的是图像含有的像素点数，是衡量图像的细节呈现能力的一个重要指标，一般表示为每英寸图像内像素点个数（PPI）。随着科学技术的进步，在人们的日常生活中出现了越来越多的图像分辨率格式，从“标清”（Standard Definition）发展到“高清”（High Definition），进而发展到目前常见的“1080p”、“2K”甚至“4K”等等。更高的分辨率意味着更多的图像细节，进而意味着可能存在更大的信息量。更大的信息量蕴含更大的应用潜能。然而，在现实世界中，一方面受成像设备物理性能的限制，人们无法获取高分辨率图像；另一方面，在互联网应用中，受网络带宽、存储介质容量等的限制，用户只能存储、传输分辨率相对较低的图像。而实际上，在多数情况下，人们往往希望获得更高分辨率的图像。因此，如何高效地提高图像分辨率，进而提高图像的质量，是计算机视觉领域的一个重要的研究课题。

图像分辨率的提升有助于获取更清晰的图像，进而从图像中挖掘更多信息，实现更多应用。提高图像分辨率的出发点主要分为两个：一方面，从硬件的角度出发，提高成像设备和整体成像系统的物理性能；另一方面，从软件的角度出发，设计、优化提高图像分辨率的处理算法。

从硬件的角度，改善成像传感器等硬件设备是提高图像分辨率的最直接的方法。然而，一方面受限于高精度成像传感器高昂的价格；另一方面，对于某些特定系统（如卫星成像、医学影像等），受限于技术瓶颈和应用环境，目前已经无法从硬件的角度对成像系统的硬件设备进行高效的技术升级。如果继续改进成像系统中硬件设备的物理性能，可能带来更高的研究、生产成本，影响投入产出比。

总的来说，通过改善硬件系统来实现图像分辨率提升的方法的成本相对较高，并且很难在短期内突破技术瓶颈、克服某些特定的技术难题。因此，从软件的角度解决图像分辨率提升问题具有重要的现实意义和极大的应用空间。

图像超分辨率技术（Super-Resolution）是指从观测到的低分辨率图像重建出相应的高分辨率图像，是从软件的角度解决图像分辨率提升问题的重要技术手段。实现图像超分辨率重建的方法可以分成两类：从多张低分辨率图像重建出高分辨率图像，即多图像超分辨率技术[2,3]和从单张低分辨率图像重建出高分辨率图像，即单图像超分辨率技术

（SISR）[4]。图像的超分辨率重建为使计算机能够更好地观察、分析、处理图片提供重要的技术支持，在很多领域都具有非常重要的应用价值：

1）高清晰度电视。将普通的标清电视（SDTV）信号转换为高清晰度电视（HDTV）信号，或将视频的某一区域进行放大等，都是图像超分辨率技术重要的应用方面，显示出超分辨率技术在日常生活中重要的应用价值。

2）医学影像。随着越来越多高科技成像设备被应用于医学领域（如胃镜等），医学检查、成像的难度在逐步下降，成像质量在逐步提升。由于被检查生物体的条件所限，医学影像设备的制造需要考虑尺寸、被检查生物体感受等因素，因此，成像传感器等硬件的改良十分困难，很难同时优化硬件物理结构和成像质量。将图像超分辨率技术应用于医学影像领域，有助于弥补硬件方面的不足，进一步提升成像质量，促进医学中某些领域的发展。

3）卫星图像。受限于遥感等领域成像技术、硬件设备，以及成像时大气状况不确定等因素，原始的卫星图像可能不能满足特定的分辨率要求。特别是某些军事应用（如军事侦察等）对图像的分辨率有更严苛的要求。利用图像超分辨率技术对原始图像进行分辨率提升，有助于提升图像精度和质量，对民用级别的地图等应用和军事级别的侦察识别等应用的升级改良都具有正向推动作用。

4）面部识别。随着科学技术的发展，越来越多的设备采用面部识别技术进行安全加密。受识别时的光线、镜头质量、成像传感器质量等因素限制，拍摄到的用于识别的图像的清晰度会受到一定程度的影响。为提高安全性能，需要对被识别对象进行高精度的细节检查。图像超分辨率技术的应用可以帮助改善被识别图像的分辨率，有助于满足对高精度识别的需求。此外，更清晰的图像可以更快速地被有效识别，从而缩短识别时间，改善用户体验。

5）监控系统。受限于摄像头尺寸、存储介质容量、布控成本等因素，监控系统采集的视频图像通常具有较低的分辨率。利用图像超分辨率技术，对视频图像的关键帧进行重建得到高分辨率的视频图像，将对警方的识别、搜索等应用提供重要的证据辅助。

由于具有巨大的应用价值和发展潜力，近年来，图像超分辨率重建技术已成为计算机视觉领域最活跃的研究课题之一。

1.2 图像超分辨率的内容和技术原理

利用图像超分辨率技术（Super-Resolution）要实现的主要目标是根据低分辨率图像（I^{LR}）重建出相应的高分辨率图像（I^{SR}）。通过多图像超分辨率技术[2,3]和单图像超分辨率技术（SISR）[4]都可以实现这一目标。

多图像超分辨率技术假设多张图像是从不同角度观察同一个画面得到的低分辨率图像，通过挖掘图像的显式冗余信息，尝试实现降采样过程的反变换[5]。这种方法利用低分辨率图像之间相互补充的信息，结合高分辨率图像特征的先验信息，由一系列低分辨率图像重建一张高分辨率图像。

单图像超分辨率技术（SISR）学习单一图像的隐式冗余信息，从单一低分辨率图像

入手恢复丢失的高分辨率信息，即对单一低分辨率图像进行观测，并结合图像的一些先验信息，恢复获取图像的时候丢失的高频数据，进而重建出相应的一张高分辨率图像。由于具有计算、存储成本和运算速度上的优势，目前研究较多、应用较广的是单图像超分辨率技术（SISR）。

因为图像超分辨率问题是不适定问题，所以，不论是多图像超分辨率还是单图像超分辨率，都需要对问题的求解进行先验约束。可能的先验约束包括之前提到的低分辨率图像之间的互补信息、高分辨率图像特征的先验信息以及低分辨率图像-高分辨率图像之间的非线性映射关系等。

1.3 本文的研究内容

图像超分辨率的主要任务是提高图像的分辨率，其基本思想是根据输入的低分辨率图像（I^{LR}）恢复对应高分辨率图像（I^{HR}）的高频分量，以此重建高分辨率图像（I^{SR}）。由于图像超分辨率问题是不适定的，需要对该问题的求解进行先验约束。实际应用中，先验约束包括针对同一画面的序列低分辨率图像之间的互补信息、高分辨率图像的图像特征信息以及低分辨率图像-高分辨率图像之间的非线性映射关系等。

本文的研究内容集中在利用卷积神经网络和子像素方法实现单图像超分辨率重建的问题上，对基于高效子像素卷积神经网络的算法[5]进行深入研究。这种方法是单图像超分辨率技术（SISR）的一种，利用基于卷积神经网络的网络结构学习低分辨率图像与高分辨率图像之间的非线性映射关系，作为求解不适定的超分辨率重建问题的先验知识，并应用一个子像素卷积层对低分辨率图像进行恢复重建。在搭建图像超分辨率重建网络的基础之上，利用常用数据集对网络进行单图像超分辨率重建方面的性能测试，与已有的重建解决方案进行对比。在此基础之上，本文对视频的图像超分辨率重建进行研究。

1.4 论文的组织结构

本文的正文部分由五章组成，按如下结构组织：

第一章绪论。阐述图像超分辨率的背景、意义，分析图像超分辨率的主要内容、技术原理，梳理本文的主要工作。

第二章梳理图像超分辨率重建方法的研究现状。研究传统的方法、基于卷积神经网络的方法以及图像超分辨率的发展趋势，分析不同方法的优点和不足。

第三章研究将卷积神经网络应用于子像素的图像超分辨率重建算法。研究基于子像素卷积神经网络的重建算法的网络结构，重点分析该结构中的子像素卷积层；对网络结构中应用的技术细节进行说明，并对视频超分辨率实现方面的工作进行研究。

第四章分析利用搭建的图像超分辨率重建系统进行性能测试的结果。首先阐述性能测试实验的实验细节，并对系统在单图像超分辨率重建方面进行性能测试得到的结果进行对比分析。

第五章对本文的研究工作进行总结，并对未来的工作进行展望。

第二章　图像超分辨率重建的研究现状

在 20 世纪 60 年代，Harris[6]等人就已经提出图像超分辨率重建技术的概念。从概念提出至今，大量研究人员投入到图像超分辨率重建的研究中，并发表了大量研究成果。根据图像像素提升网络的参数是提前人为设定的还是利用计算机通过训练学习得到的，可以对图像超分辨率重建的方法进行一种分类。

2.1　传统的重建方法

传统的提高图像分辨率的常用方法之一是插值法，包括双线性插值法、双三次插值法等。这种方法利用预定义的数学公式对低分辨率图像进行重建，不需要其他训练得到的数据。插值法容易实现，但是重建出的像素点和其周围的像素点之间存在区域相似性，使得利用插值法可以生成良好的平滑图像，但不能充分重建图像的边缘区域[4]。一些图像细节可能会因此而丢失，比如变化较大、锐度较高的图像元素边缘部分可能会变模糊等等。

另一种常用的传统方法是稀疏编码方法[7]。稀疏编码方法假设任意的自然图片都能在转换区域被稀疏地表达。这个转换区域通常是一个图像元素的字典，通过发掘低分辨率图像和高分辨率图像之间的对应关系得到。然而，稀疏编码方法的缺点是将稀疏性限制引入非线性重建中时通常需要较大的计算代价[5]。

此外，传统的重建方法还有分散处理图像各部分的缺点[8]。这使得大量的图像内部特征信息、像素点之间的关系信息存在不充分发掘学习的可能。这为图像超分辨率重建技术的发展和图像超分辨率重建系统的性能提升提供可能性。

2.2　基于卷积网络的重建方法

针对传统的提高图像分辨率的方法的不足，Chao Dong 等人提出基于卷积神经网络的图像超分辨率技术（SRCNN）[8]，将卷积神经网络引入图像超分辨率重建方法中。相比于传统的双三次插值法和稀疏编码方法，基于卷积神经网络的图像超分辨率技术（SRCNN）可以消耗更少的时间成本，得到更加清晰的高分辨率图像，显示出利用深度学习方法解决图像超分辨率重建问题的优越性。

卷积神经网络是一种专门用来处理具有类似网格结构的数据（例如图像即可以被视为二维的像素网格）的神经网络[9]，已经在大量不同类型的计算机视觉处理任务（如图像分类[10]、图像监测[11]等）中取得成功。卷积运算通过三个重要思想来帮助改进机器学习系统[9]：稀疏交互、参数共享、等变表示，从而可以减少对模型存储的需求、减少计算量、提高统计效率。目前已经发展出很多基于卷积神经网络的实现单图像超分辨率重建的解决方案，比如，基于极深卷积神经网络的超分辨率技术（VDSR）[12]，基于深度递归卷积网络的超分辨率技术（DRCN）[13]等等。

虽然一些基于卷积神经网络的图像超分辨率重建技术已经取得了比较优秀的表现，

但是这些方法通常是在重建网络的第一层或第一层之前就已经完成图像分辨率的提升[8,14]，或者是在重建网络的中间层部分逐层提高图像分辨率[15]。由于卷积运算利用输入图像所有的像素点进行计算，因此这些方法会因为输入网络的图像的像素点数增加导致计算中的参数增多，从而增加计算复杂度[5]。另外，提前提高图像的分辨率不能为解决不适定重建问题提供额外信息[5]。

为进一步提高算法效率、降低计算和存储成本，Wenzhe Shi 等人提出一种高效的子像素卷积神经网络模型（ESPCN）[5]，用以更好地实现实时的单图像和视频的超分辨率。在这个模型中，Wenzhe Shi 等人创造性地引入一个高效子像素卷积层，通过一个被称为“周期性洗牌（Periodic Shuffling）”的运算操作实现子像素卷积方法。高效子像素卷积神经网络模型（ESPCN）在卷积神经网络结构的最后提高图像分辨率，相比之下具有两个优势[5]：1、可以充分降低计算和存储的复杂度，使实时高清视频的超分辨率处理成为可能；2、可以学习更精确的低分辨率图像（I^{LR}）到高分辨率图像（I^{HR}）的映射，提高重建的准确性。本文的重点研究内容集中在利用卷积神经网络和子像素方法实现单图像超分辨率重建的问题。

此外，Ying Tai 等人提出的基于深度递归残差网络的图像超分辨率（DRRN）[16]，将多路径局部残差学习、全局残差学习以及多权重递归学习相结合；Wei-Sheng Lai 等人提出的基于深度拉普拉斯金字塔网络的快速精确超分辨率（LapSRN）[17]，将传统的拉普拉斯金字塔算法和卷积神经网络相结合；Christian Ledig 等人提出的基于生成对抗网络的实感单图像超分辨率（SRGAN）[18]，将卷积神经网络和生成对抗网络相结合，利用感知损失(perceptual loss)和对抗损失(adversarial loss)增强重建图像的真实感。

2.3 图像超分辨率的发展趋势

在 2017 年的 IEEE Conference on Computer Vision and Pattern Recognition（CVPR）会议上，研究人员继续发表了有关图像超分辨率重建的研究成果，其中包括 Ying Tai 等人提出的基于深度递归残差网络的图像超分辨率（DRRN）[16]；Wei-Sheng Lai 等人提出的基于深度拉普拉斯金字塔网络的快速精确超分辨率（LapSRN）[17]；Christian Ledig 等人提出的基于生成对抗网络的实感单图像超分辨率（SRGAN）[18]等等。从传统的插值法、稀疏编码法等发展到 SRCNN，再发展到 VDSR、DRCN、ESPCN，进而发展到 DRRN、LapSRN、SRGAN 等，图像超分辨率重建技术的计算速度、计算效率越来越高，计算的成本开销越来越低。图像超分辨率重建的网络结构从传统的人为提前固定参数结构，发展到引入卷积神经网络进行参数学习的结构，进而加深网络、应用网络递归、引入其他网络结构等等。未来基于卷积神经网络的单图像超分辨率重建系统可能会继续将已有研究成果进行组合，或者引入其他结构，比如加深或改变提取图像特征的网络，更换新的由低分辨率特征映射重建高分辨率数据的网络结构等等。

随着图像超分辨率重建技术的应用变得越来越广泛，图像超分辨率重建与其他技术的结合将会更加紧密，比如将图像的超分辨率重建算法直接引入相机，减少不必要的错误对焦虚化，使摄影师能够拍摄得到更准确清晰的图像；将视频的超分辨率重建与图像

识别、图像跟踪等技术相结合，实现速度更快、精度更高的图像识别跟踪等等。

第三章　将卷积神经网络应用于子像素的图像超分辨率算法

作为算法原理部分，本章重点研究将卷积神经网络应用到子像素的图像超分辨率重建方法，分析阐述子像素卷积神经网络模型的研究背景、网络结构以及重建过程的重点环节，包括基于卷积神经网络的学习环节、利用子像素方法的图像重建环节以及输入图像的降采样生成过程，深入研究分析子像素卷积层的工作原理。此外，本章对视频的超分辨率重建算法进行研究。

3.1　背景

将卷积神经网络应用于子像素的图像超分辨率重建，在重建网络的最后一层对被重建图像进行分辨率提升，避免被重建图像在进入重建网络或在重建网络中提升分辨率。在基于卷积神经网络的重建方法中利用的卷积运算是像素对应的计算方法。输入图像的像素数越多，参与计算的数据量越多，卷积运算的复杂度就越高。在重建网络的最后一层提升图像分辨率，有利于减少重建过程中参与运算的数据量，从而降低图像重建过程的计算复杂度和存储成本，实现更高效的图像超分辨率重建和更快的视频的超分辨率重建。

3.2　网络结构

单图像超分辨率技术（SISR）的核心内容是由给定的单一低分辨率图像（I^{LR}）生成高分辨率图像（I^{SR}）。在研究和测试中，低分辨率图像（I^{LR}）通常是由相应的原始高分辨率图像（I^{HR}）进行高斯模糊并按放缩比例 r 缩小得到的。假设 I^{HR} 和 I^{LR} 有 C 个颜色通道，若 I^{HR} 的图像尺寸是 $rH\times rW\times C$，则 I^{LR} 的图像尺寸是 $H\times W\times C$，其中，H 为 I^{LR} 的高，W 为 I^{LR} 的宽。由于人眼对图像的亮度信息更敏感，本文在处理图像时，将图像由 RGB 格式转换为 YCbCr 格式，并重点对 Y 通道数据应用本文的重建方法，然后再将 Y 通道、Cb 通道、Cr 通道数据合并，经过格式转换输出 RGB 格式图像，实现图像超分辨率重建。

3.2.1　卷积神经网络

卷积神经网络是一种专门用来处理具有类似网格结构的数据（例如图像即可以被视为二维的像素网格）的神经网络[9]，已经在大量不同类型的计算机视觉处理任务（如图像分类[10]、图像监测[11]等）中取得成功。卷积运算通过三个重要思想来帮助改进机器学习系统[9]：稀疏交互、参数共享、等变表示，从而可以减少对模型存储的需求、减少计算量、提高统计效率。本文利用卷积神经网络提取图像的隐式冗余信息，结合成对训练图像学习低分辨率图像-高分辨率图像之间的非线性映射关系。由于不涉及语义信息的提取，实现图像超分辨率重建的卷积神经网络结构中不需要池化和全连接层结构。

3.2.2 子像素卷积方法

子像素卷积方法利用权重间隔$\frac{1}{r}$的卷积核，以步长$\frac{1}{r}$对低分辨率图像进行卷积。相当于先将低分辨率图像的像素点按放缩比例 r 展开，然后在原图像像素点之间插入零值像素点，再利用卷积核以步长为 1 进行标准的卷积神经网络运算。落在原图像像素点之间的权重不参与卷积运算。本文利用一种子像素卷积层实现子像素卷积方法。

3.2.3 子像素卷积层

与其他图像超分辨率重建方法不同的地方在于，将卷积神经网络应用于子像素的图像超分辨率重建方法在卷积神经网络的最后一层之后利用子像素方法对图像的分辨率进行提升，如式（3-1）。根据放缩比例 r，人为设定第 l 层卷积神经网络的特征数为 r^2，因此在第 l 层卷积神经网络之后得到 r^2 个特征映射。

$$\mathrm{I}^{SR} = f^{L}\left(\mathrm{I}^{LR}\right) = PS\left(W_L * f^{L-1}\left(\mathrm{I}^{LR}\right) + b_L\right) \qquad \text{式（3-1）}$$

$$PS(T)_{x,y,c} = T_{\lfloor x/r \rfloor, \lfloor y/r \rfloor, C\cdot r\cdot \mathrm{mod}(y,r)+C\cdot \mathrm{mod}(x,r)+c} \qquad \text{式（3-2）}$$

通过对这 r^2 个特征映射进行“周期性像素重排”操作，如式（3-2），将每个特征映射通道的子像素数据周期性重新排布得到最后重建的高像素图像，即将通过第 l 层卷积神经网络输出得到的形如$H \times W \times C \times r^2$的张量重新安排得到新的形如$rH \times rW \times C \times 1$的张量，再转换格式得到重建的高分辨率图像。

H*W*r^2C=3*3*9												T		
A0	B0	C0	0	A1	B1	C1	1	A2	B2	C2	2			
D0	E0	F0		D1	E1	F1		D2	E2	F2				
G0	H0	I0		G1	H1	I1		G2	H2	I2				
A3	B3	C3	3	A4	B4	C4	4	A5	B5	C5	5			
D3	E3	F3		D4	E4	F4		D5	E5	F5				
G3	H3	I3		G4	H4	I4		G5	H5	I5				
A6	B6	C6	6	A7	B7	C7	7	A8	B8	C8	8			
D6	E6	F6		D7	E7	F7		D8	E8	F8				
G6	H6	I6		G7	H7	I7		G8	H8	I8				

图 3-1 “周期性像素重排”之前的特征映射

为了示意“周期性像素重排”操作的基本原理，为简化问题，假设$H = W = 3$，$C = 1$，$r = 3$，则在最后一层卷积神经网络之后得到如图 3-1 所示的 9 个$3 \times 3 \times 1$的特征映射（$3 \times 3 \times 1 \times 9$）。根据算法原理，经过“周期性像素重排”操作后，得到一张$9 \times 9 \times 1$的重建图像（$9\times 9 \times 1 \times 1$），如图 3-2 示意。

rH*rW*C=9*9*1									PS(T)
A0	A1	A2	B0	B1	B2	C0	C1	C2	0
A3	A4	A5	B3	B4	B5	C3	C4	C5	1
A6	A7	A8	B6	B7	B8	C6	C7	C8	2
D0	D1	D2	E0	E1	E2	F0	F1	F2	3
D3	D4	D5	E3	E4	E5	F3	F4	F5	4
D6	D7	D8	E6	E7	E8	F6	F7	F8	5
G0	G1	G2	H0	H1	H2	I0	I1	I2	6
G3	G4	G5	H3	H4	H5	I3	I4	I5	7
G6	G7	G8	H6	H7	H8	I6	I7	I8	8
0	1	2	3	4	5	6	7	8	

转换模式		
0	1	2
3	4	5
6	7	8

图 3-2 “周期性像素重排”之后得到的重建图像及转换模式

3.2.4 图像缩小过程

实验中，为了简化实验过程，并有利于对比分析重建网络的性能，输入网络的低分辨率图像（I^{LR}）是由相应的原始高分辨率图像（I^{HR}）得到的。为了由原始高分辨率图像（I^{HR}）得到相应的低分辨率图像（I^{LR}），首先，模拟相机的点扩散函数，将原始高分辨率图像（I^{HR}）与高斯滤波器相卷积，之后按放缩比例 r 对图像进行降采样。该过程如图 3-3 所示。

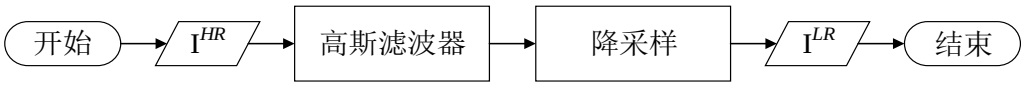

图 3-3 图像缩小过程

3.2.5 主要重建结构

为了实现单图像超分辨率重建，首先在低分辨率图像（I^{LR}）上直接应用 l-1 层卷积神经网络提取图像特征映射，之后应用子像素卷积层对低分辨率的图像映射（LR feature maps）进行整理，进而生成重建的高分辨率图像（I^{SR}）。选择 *tanh* 作为卷积神经网络中的激活函数。实现图像超分辨率重建的卷积神经网络结构中不需要池化和全连接层。本文实现图像超分辨率重建的主要网络结构如图 3-4 所示。

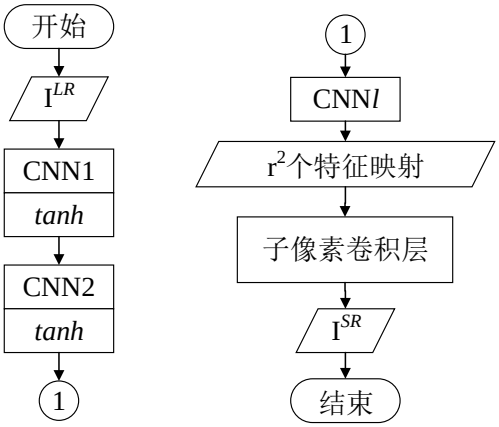

图 3-4 图像超分辨率重建网络的主要结构

3.3 视频的超分辨率重建

视频本质上是在时间轴上的一组连续图像序列的组合，其中的每一幅图像被称为一帧。当以每秒 50 帧以上的速度连续播放静止图像时，在人脑中形成连续运动的视频效果。因此，图像或视频的表示通常需要占用海量数据。由于数字视频本身独有的特点，视频图像及图像间存在空间、时间等方面的冗余。现实应用中，数字视频需要经过有效的压缩才能在存储、传输或处理方面具有实际使用价值。

在现实生活中，视频的超分辨率重建具有极大的军民实用价值，比如将标清电视信号转换为高清电视信号、提高监控视频的分辨率、提高军事侦察（卫星、无人机等）视频的分辨率。本文利用基于卷积神经网络与子像素方法结合的方法，实现了视频超分辨率的实时输出和图像保存。

针对本文的应用场景，在处理视频时，一方面可以考虑视频帧的冗余，利用视频帧

包含的相关性，整体性地处理视频图像。考虑帧内、帧间冗余的优势是：由于利用图像冗余进行压缩，视频数据量相对较小，处理速度相对较快。而这种方法的不足是：由于利用相关性，视频图像的某些数据是依赖其他图像的，因此存在错误扩散的可能。另一方面可以独立处理视频帧，即将视频的每一帧独立取出，进行独立处理之后再重新组合形成视频。这种方法的优势是：由于处理过程独立，每一帧图像的数据是原始数据，与其他帧不存在依赖关系，因此处理后的图像的质量可能相对较高。而独立处理的不足是：数据量较大，处理速度较慢。

本文实现视频超分辨率重建的基本思路是将视频的每一帧看作一张单一的低分辨率图像（I^{LR}），从视频中依次取出每一帧图像，之后依次输入重建系统网络，重建出高分辨率图像，然后再将重建出的高分辨率图像组合，重新形成视频。利用 opencv 库函数实现对图像帧的提取、格式转换、保存等操作。基本工作结构如图 3-5 所示。

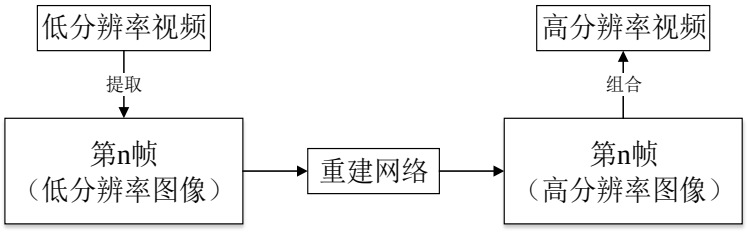

图 3-5 视频超分辨率重建的工作结构

由于本文实现视频超分辨率重建的核心思想是将视频分解为单图像再进行重建，所以，实现视频超分辨率重建的核心网络与之前章节所述的用于单图像超分辨率重建的网络是相同的。因此，视频超分辨率重建网络可以直接使用利用图像数据集训练好的网络。

3.4 本章小结

本章重点研究将卷积神经网络应用于子像素的图像超分辨率重建算法，阐述了本文实现图像超分辨率重建的算法原理、网络结构，以及重建方法的技术细节。本章还对视频超分辨率重建的实现方法进行研究。

第四章　图像超分辨率重建实验

本文的主要研究内容是基于子像素方法和卷积神经网络的单图像超分辨率重建技术，核心目标是参考第三章的算法原理以及结构模型，搭建一个将卷积神经网络和子像素方法相结合的图像超分辨率重建网络。利用本文的重建系统可以实现单图像超分辨率重建、批量单图像超分辨率重建和视频图像超分辨率重建及其实时输出。

本文参考已有研究成果，利用其实验部分使用的数据集对本文构建的重建网络进行训练，并对这一重建系统的性能进行测试，对比本文重建系统恢复图像和其它方法得到的重建图像的峰值信噪比（PSNR），比较不同重建方法的处理速度。

4.1　数据集

在网络训练环节，参考已有的研究成果[19, 20]，本文使用 T91[21]和 BSDS200[22]数据集对网络进行训练。这两个训练数据集分别提供 91 张和 200 张图片。在训练阶段，首先对训练数据集的图片进行预处理。根据放缩比例 r 对原始图像进行裁边处理，使图片的长、宽尺寸是放缩比例 r 的整倍数。处理后的高分辨率图像（I^{HR}）作真值（ground truth），对其进行高斯模糊，并根据放缩比例 r 对高分辨率图像（I^{HR}）进行降采样得到相应的用于训练的低分辨率图像输入（I^{LR}）。将高分辨率图像（I^{HR}）和低分辨率图像（I^{LR}）再分别分割成$17r \times 17r$像素和17×17像素的子图像，步长分别为$\left(17 - \sum \mathrm{mod}(f,2)\right) \times r$和$\left(17 - \sum \mathrm{mod}(f,2)\right)$。这样可以确保原始图像中的所有像素作为训练数据的真值出现且仅出现一次。

针对训练好的网络，本文中分别使用 Set5[23]、Set14[24]、BSDS300[22]、BSDS500[22]和 SuperTex136[25]数据集对重建系统的性能进行测试。同样的，在进行测试之前，需要提前根据放缩比例 r 对测试数据集的原始图像进行裁边处理，并分别生成并保存相应的高分辨率图像（I^{HR}）和低分辨率图像（I^{LR}）。其中高分辨率图像（I^{HR}）会被作为测试真值（ground truth），用于之后和重建图像（I^{SR}）相对应计算重建图像（I^{SR}）的峰值信噪比（PSNR），进行系统的性能评价。

4.2　参数设置

本文搭建的实现单图像超分辨率的重建网络的整体结构分为两部分：第一部分是主要用于学习图像内部特征（即图像内部隐式冗余信息）的卷积神经网络部分；第二部分是主要用于重建超分辨率图像的子像素卷积层部分。其中，子像素卷积层的主要操作就是“周期性像素重排”。卷积神经网络的层数 l 设定为 3 层，用于提取图像特征，其中最后一层为了之后应用子像素方法（“周期性像素重排”操作），根据放缩比例 r 人为设定特征通道数为 r^2。第 i 层的卷积核尺寸 f_i 和特征通道数 n_i 分别为：$(f_1, n_1) = (5, 64)$，$(f_2, n_2) = (3, 32)$，$(f_3, n_3) = (3, r^2)$。在卷积神经网络的每一层都选择 *tanh* 作为激活函数。使用均方误差函数（Mean Squared Error）作为网络训练的目标函数，如式（4-1）。

$$\ell\left(W_{1:L},b_{1:L}\right)=\frac{1}{r^2HW}\sum_{x=1}^{rH}\sum_{x=1}^{rW}\left(\mathrm{I}_{x,y}^{HR}-f_{x,y}^{L}\left(\mathrm{I}^{LR}\right)\right)^2 \qquad \text{式（4-1）}$$

在网络训练部分，卷积神经网络中卷积核的权重值和偏置均初始化为 0。在训练中设定学习率为 0.0001，并使用 Adam 优化器对学习率进行优化。本文使用峰值信噪比（PSNR）作为系统性能的评价指标。

4.3 图像超分辨率重建系统的性能

利用 T91 和 BSDS200 数据集对网络进行训练。首先利用 T91 数据集对网络进行训练得到第一个模型（为简便表示，表格中以 Model_91 代替），再利用 T91 和 BSDS200 数据集同时对网络进行训练得到第二个模型（为简便表示，表格中以 Model_291 代替）。使用 Set5、Set14、BSDS300、BSDS500 和 SuperTex136 数据集对两个模型分别进行测试。在本文的重建系统中，不仅实现了基于卷积神经网络和子像素方法的单图像超分辨率重建，同时，在程序中添加利用双三次插值法重建图像的部分，也得到了相应的重建图像质量数据。实验分别记录了使用子像素卷积神经网络方法和双三次插值法进行单图像超分辨率重建得到的超分辨率图像的峰值信噪比（PSNR），并记录了使用子像素卷积神经网络方法重建图像的运行时间。通过与 Wenzhe Shi 等人发表的基于高效子像素卷积神经网络模型的单图像超分辨率重建方法，及与其同时期或之前的超分辨率重建方法进行超分辨率重建图像精度、重建速度方面的对比，分析本文构建的基于卷积神经网络和子像素方法的单图像超分辨率重建系统的性能。

4.3.1 超分辨率图像的重建质量

在测试阶段，分别将放缩比例 r 设定为 2、3、4，并分别进行实验。本文使用峰值信噪比（PSNR）作为评价超分辨率图像的重建质量的参考标准。实验结果如下所示。

表 4-1 r=2 的峰值信噪比

Dataset	Scale	Bicubic (dB)	Model_91 (dB)	Model_291 (dB)
Set5	2	28.00	32.06	**32.20**
Set14	2	25.73	29.22	**29.31**
BSD300	2	25.92	28.81	**28.88**
BSD500	2	25.96	28.94	**29.01**
SuperTex136	2	23.96	27.27	**27.35**
Average	2	25.37	28.45	**28.53**

表 4-1 是放缩比例 r 为 2 时的测试结果。

表 4-2 *r*=3 的峰值信噪比

Dataset	Scale	Bicubic (dB)	Model_91 (dB)	Model_291 (dB)	ESPCN[5] (91) (dB)
Set5	3	25.01	**28.65**	28.64	30.95
Set14	3	23.30	26.15	26.15	27.56
BSD300	3	23.87	25.89	**25.93**	27.16
BSD500	3	23.81	25.89	**25.93**	27.26
SuperTex136	3	21.72	24.18	**24.20**	25.27
Average	3	23.20	25.42	**25.45**	26.69

表 4-2 是放缩比例 *r* 为 3 时的测试结果。

表 4-3 *r*=4 的峰值信噪比

Dataset	Scale	Bicubic (dB)	Model_91 (dB)	Model_291 (dB)	ESPCN[5] (91) (dB)
Set5	4	23.00	26.22	**26.44**	28.76
Set14	4	21.66	24.13	**24.30**	25.89
BSD300	4	22.48	24.18	24.18	25.68
BSD500	4	22.36	24.14	**24.29**	25.69
SuperTex136	4	20.23	22.32	**22.46**	23.63
Average	4	21.74	23.63	**23.74**	25.11

表 4-3 是放缩比例 *r* 为 4 时的测试结果。

表 4-1、表 4-2 和表 4-3 分别记录了放缩比例 *r* 为 2、3、4 时，利用双三次插值法和本文的子像素卷积法的测试结果。其中，本文的子像素卷积法的测试结果包括分别使用 T91 和使用 T91 及 BSDS200 训练网络的测试结果。表中加粗的数据是利用本文基于卷积神经网络和子像素方法搭建的单图像超分辨率重建系统测试得到的峰值信噪比较高的数据，即代表更高质量的超分辨率重建图像。

从表格中的数据可以看出，相比于双三次插值法，本文构建的将卷积神经网络应用于子像素的图像超分辨率重建系统整体上可以获得质量更高的超分辨率图像。此外，使用 BSDS200 和 T91 同时进行训练得到的网络重建得到的超分辨率图像的峰值信噪比普遍比只使用 T91 数据集训练得到的网络重建的超分辨率图像的高，说明基于相同的实验超参数设定，用更大的数据集可以训练出性能更优的重建网络。

(a) Original (b) Input (c) Bicubic (31.57dB) (d) Model_91 (35.36dB) (e) Model_291 (35.80dB)

(f) Original (g) Input (h) Bicubic (22.70dB) (i) Model_91 (27.53dB) (j) Model_291 (27.50dB)

图 4-1 *r*=2 时 Set5 的部分重建结果

图 4-1 是当放缩比例为 2 时，利用 Set5 数据集进行测试的部分重建图像，“baby”和“butterfly”。

(a) Original (b) Input (c) Bicubic (27.99dB) (d) Model_91 (34.23dB) (e) Model_291 (34.72dB)

(f) Original (g) Input (h) Bicubic (21.27dB) (i) Model_91 (22.82dB) (j) Model_291 (22.83dB)

图 4-2 *r*=2 时 Set14 的部分重建结果

图 4-2 是当放缩比例为 2 时，利用 Set14 数据集进行测试的部分重建图像，“foreman”和“baboon”。

(a) Original (b) Input (c) Bicubic (35.02dB) (d) Model_91 (38.46dB) (e) Model_291 (38.68dB)

(f) Original (g) Input (h) Bicubic (19.14dB) (i) Model_91 (21.16dB) (j) Model_291 (21.17dB)

图 4-3 r=2 时 BSDS300 的部分重建结果

图 4-3 是当放缩比例为 2 时，利用 BSDS300 数据集进行测试的部分重建图像，“3096”和“167083”。

(a) Original (b) Input (c) Bicubic (36.19dB) (d) Model_91 (39.41dB) (e) Model_291 (39.63dB)

(f) Original (g) Input (h) Bicubic (18.75dB) (i) Model_91 (20.73dB) (j) Model_291 (20.76dB)

图 4-4 r=2 时 BSDS500 的部分重建结果

图 4-4 是当放缩比例为 2 时，利用 BSDS500 数据集进行测试的部分重建图像，“70011”和“296058”。

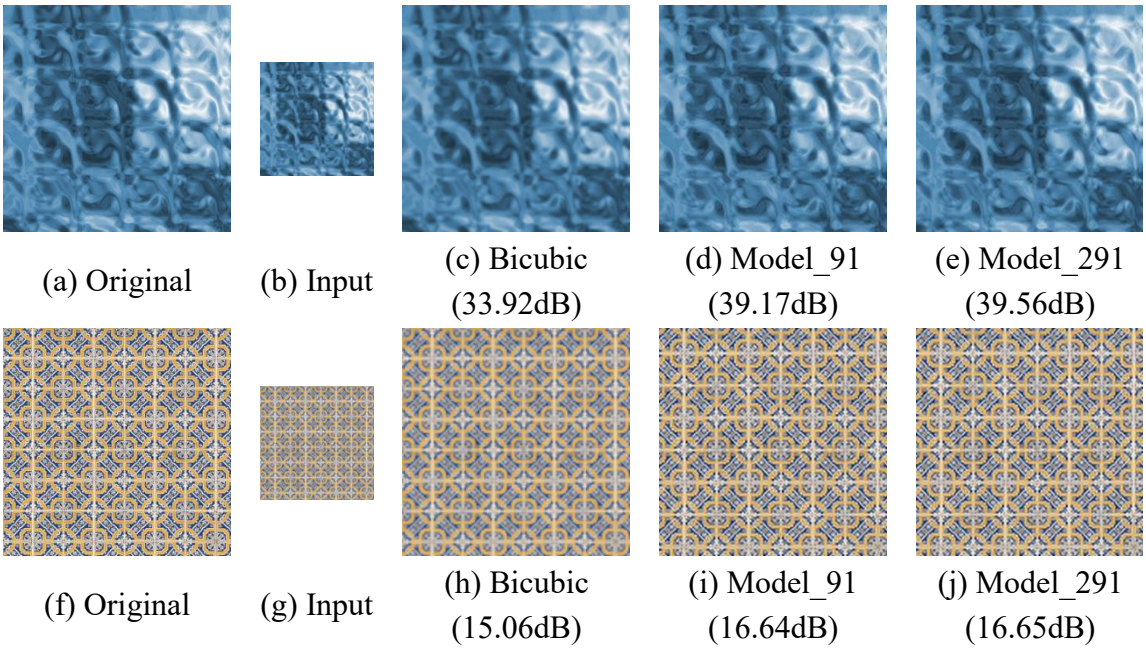

图 4-5 *r*=2 时 SuperTex136 的部分重建结果

图 4-5 是当放缩比例为 2 时，利用 SuperTex136 数据集进行测试的部分重建图像，“google_glass_687”和“google_mosaic_264”。

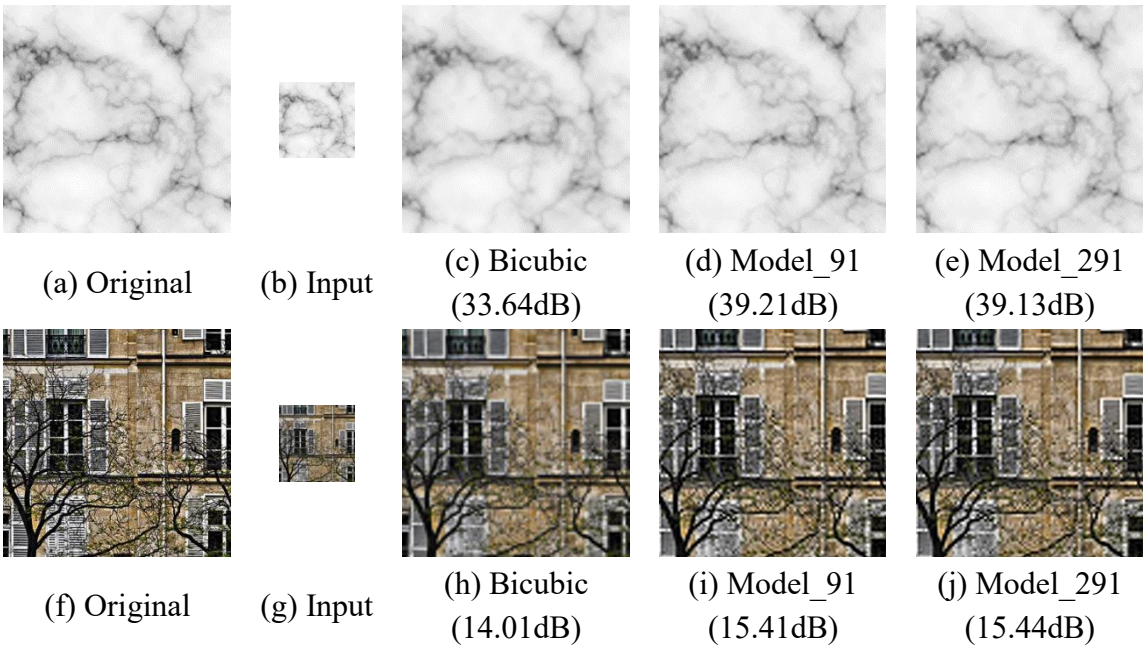

图 4-6 *r*=3 时 SuperTex136 的部分重建结果

图 4-6 是当放缩比例为 3 时，利用 SuperTex136 数据集进行测试的部分重建图像，“google_marble_337”和“google_urban_47”。

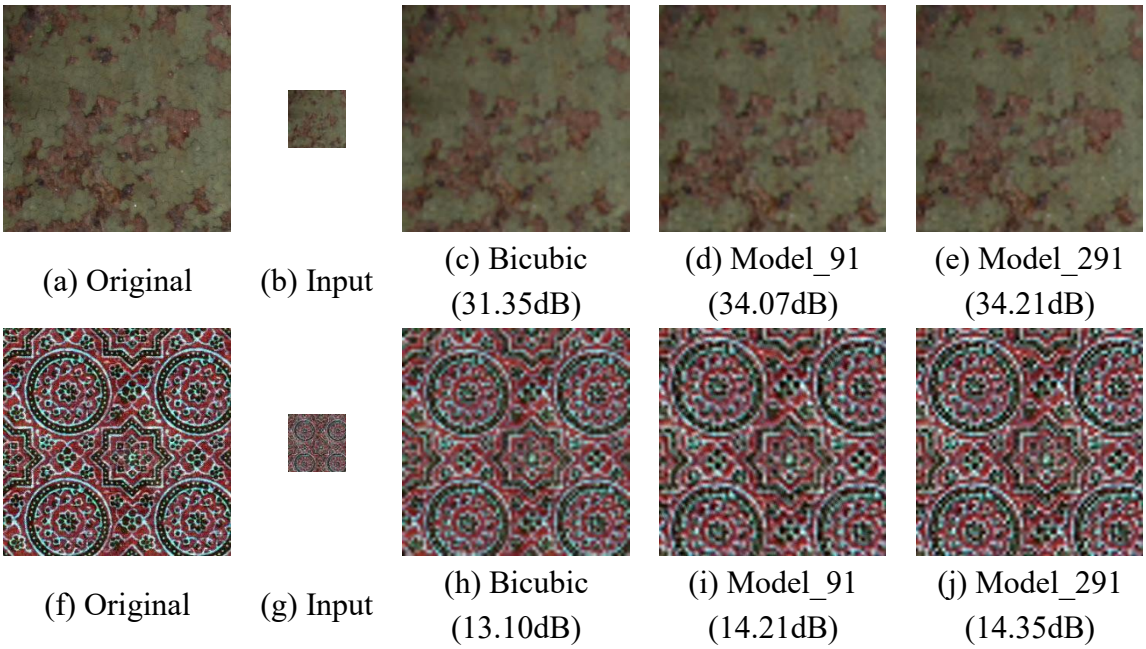

图 4-7 *r*=4 时 SuperTex136 的部分重建结果

图 4-7 是当放缩比例为 4 时，利用 SuperTex136 数据集进行测试的部分重建图像，“google_Metal_202”和“google_Fabric_634”。

图 4-1、图 4-2、图 4-3、图 4-4 和图 4-5 分别展示放缩比例为 2 时利用 Set5、Set14、BSDS300、BSDS500 和 SuperTex136 数据集对两个模型分别进行测试得到的峰值信噪比最大和最小的超分辨率重建图像。图 4-6 和图 4-7 分别展示了放缩比例为 3 和 4 时利用 SuperTex136 数据集对两个模型分别进行测试得到的部分超分辨率重建图像。重建图像下对应标注了相应的峰值信噪比结果。从图像上可以直观感受到利用基于卷积神经网络和子像素方法的重建方法得到的重建图像比利用传统的双三此插值法得到的重建图像的图像质量更高，图像更清晰。

表 4-2 和表 4-3 中最后一列的 ESPCN 数据取自 Wenzhe Shi 等人发表的基于高效子像素卷积神经网络模型（ESPCN）的单图像超分辨率重建方法的文章[5]的补充材料。数据结果显示，本文的重建系统的测试结果接近 Wenzhe Shi 等人的 ESPCN 方法的平均性能。但是，本文的实验结果与 ESPCN 的最优实验结果之间仍存在一定差距，造成这种结果的主要原因可能是：1、训练的目标函数还没有完全收敛，损失函数的结果还有继续下降的空间；2、训练过程陷入局部最优解。针对第一个可能的原因，本文通过增加训练周期观察目标函数的收敛情况。进一步实验得到的峰值信噪比的结果是：在目标函数基本收敛（损失函数值下降极其缓慢）的情况下，重建图像的峰值信噪比与增加训练周期之前测试得到的结果相比没有明显的提高。针对第二个可能的原因，在实验条件允许的情况下，经过多次实验，没有得到更优的结果。

从整体来看，如表 4-1、表 4-2 和表 4-3 的数据所示，随着放缩比例的增大，重建图像的峰值信噪比普遍下降。以不同放缩比例的平均测试结果为代表，如表 4-4 所示。

表 4-4 不同放缩比例的重建图像的平均峰值信噪比

Scale	Bicubic (dB)	Model_91 (dB)	Model_291 (dB)
2	25.37	28.45	**28.53**
3	23.20	25.42	**25.45**
4	21.74	23.63	**23.74**

从表 4-4 的数据可以看到，随着放缩比例的提高，峰值信噪比的平均结果呈现下降趋势。由于测试使用的低分辨率输入图像是由原始数据集的高分辨率图像根据放缩比例预处理得到的，重建图像的峰值信噪比的计算是以原始高分辨率图像为基准的，因此，在对重建网络进行训练的时候，根据不同的放缩比例，预处理得到不同的低分辨率-高分辨率图像数据集对，训练得到不同的模型。对于不同的放缩比例，预处理得到的低分辨率图像包含的原始信息不同，会影响重建的精度。放缩比例越大，图像缩小过程中丢失的原始信息越多，因此重建精度越低。此外，如表 4-4 所示，Model_291 的测试数据普遍大于 Model_91 的数据结果，说明相比于只使用 T91 数据集进行训练，利用 T91 数据集和 BSDS200 数据集进行训练得到的网络的重建性能更优。因此，在网络训练部分使用更大的训练数据集可以得到重建精度更高、泛化能力更强的重建网络。

表 4-5 本文与 ESPCN[5] 的峰值信噪比结果对比（r=3）

Dataset	Scale	Model_91 (dB)	ESPCN [5] (91) (dB)
Set5	3	28.65	32.55
Set14	3	26.15	29.08
BSD300	3	25.89	28.26
BSD500	3	25.89	28.34
SuperTex136	3	24.18	26.42
Average	3	25.42	27.82

相比于 Wenzhe Shi 等人提出的基于高效子像素卷积神经网络模型（ESPCN）的单图像超分辨率重建方法的实验数据[5]，本文的实验结果与其最优结果之间仍有一定的提升空间，如表 4-5、表 4-6 所示。

本文将放缩比例为 3 的测试结果与 ESPCN 文章中提供的数据进行对比，如表 4-5 所示。

表 4-6 本文与 ESPCN[5] 的峰值信噪比结果对比（平均）

Scale	Model_91 (dB)	Model_291 (dB)	ESPCN [5] (ImageNet) (dB)
3	26.15	26.17	28.11
4	24.20	24.33	26.53

此外，在 Wenzhe Shi 等人发表的基于高效子像素卷积神经网络模型（ESPCN）的单图像超分辨率重建方法的文章[5]中提供了利用 ImageNet 训练的最终模型进行放缩比例为 3 和 4 得到的重建图像的峰值信噪比的最好的平均结果，与本文的平均结果对比如表 4-6 所示。表中数据显示，本文的实验结果低于 ESPCN 的最终结果，可能的主要原因是

训练数据集不同。受限于实验室硬件条件和实验时间等因素，本文实验的训练部分使用T91 数据集和 BSDS200 数据集，与 Wenzhe Shi 等人使用 ImageNet 数据集中随机的 50000 张图像相比有巨大的数量差距，因此训练得到的网络参数会有不同，导致重建图像的质量差异。另外，Wenzhe Shi 等人在发表的文章中引用的是他们实验得到的最优的结果，在他们的文章的补充材料中有详细的测试结果数据，表明其测试结果有一定的结果范围。本文的测试结果与其平均结果接近。在表 4-2 和表 4-3 中引用的 ESPCN 的测试结果数据即为其补充材料中公布的一组实验结果。因此，本文的测试结果与最优结果之间的差距也有可能是由于实验陷入局部最优解造成的。由于实验条件所限，本文未能进行充足的实验以验证是否由于陷入局部最优解导致测试结果不能达到最优。

3.2.2 超分辨率图像的重建速度

将卷积神经网络应用于子像素的图像超分辨率重建方法的一大优势在于，相比于其他一些基于卷积神经网络的单图像超分辨率重建方法，基于卷积神经网络和子像素方法相结合的重建技术的重建速度有明显的提升。在测试实验中，分别测试并记录了当放缩比例为 2、3、4 时基于本文重建系统的图像重建时间。记录结果如表 4-7、表 4-8 和表 4-9 所示。在实验时，利用基于 GPU 的架构进行超分辨率图像重建中遇到重建图像上出现噪声的情况，而利用基于 CPU 的架构则没有噪声出现。因此以下的测试结果是在基于 CPU 的架构上实验得到的。

表 4-7 r=2 的重建时间

Dataset	Scale	Model_91 (ms)	Model_291 (ms)
Set5	2	**33.53**	35.16
Set14	2	58.32	**53.14**
BSD300	2	**55.28**	55.57
BSD500	2	**72.97**	74.90
SuperTex136	2	49.47	**49.45**
Average	2	**61.18**	61.95

表 4-7 是放缩比例 r 为 2 时的测试结果。

表 4-8 r=3 的重建时间

Dataset	Scale	Model_91 (ms)	Model_291 (ms)
Set5	3	**20.62**	21.12
Set14	3	**32.04**	32.31
BSD300	3	**39.66**	40.23
BSD500	3	**57.75**	58.17
SuperTex136	3	37.33	**37.25**
Average	3	**46.50**	46.80

表 4-8 是放缩比例 r 为 3 时的测试结果。

表 4-9 r=4 的重建时间

Dataset	Scale	Model_91 (ms)	Model_291 (ms)
Set5	4	16.75	**15.54**
Set14	4	24.19	**23.67**
BSD300	4	**33.41**	33.67
BSD500	4	50.31	**50.23**
SuperTex136	4	**33.30**	33.80
Average	4	**40.37**	40.52

表 4-9 是放缩比例 r 为 4 时的测试结果。

表 4-10 不同放缩比例的重建图像的平均重建时间

Scale	Model_91 (ms)	Model_291 (ms)
2	**61.18**	61.95
3	**46.50**	46.80
4	**40.37**	40.52

表 4-7、表 4-8 和表 4-9 分别记录了放缩比例 r 为 2、3、4 时，分别使用 T91 和使用 T91 及 BSDS200 训练网络的测试结果。表中加粗的数据是运行时间较短的数据，即代表更快的超分辨率重建速度。由于重建网络的结构相同，只有卷积核权重及偏差等网络参数上的不同，因此，在相同的放缩比例的情况下，分别使用训练得到的两个模型进行超分辨率图像重建的运行时间基本相同。

然而，如表 4-10 所示，随着放缩比例的增大，图像重建的运行时间在减小。出现这种现象的主要原因是输入图像的分辨率差异。由于重建网络中的卷积神经网络以及子像素卷积层中的运算是元素对应的，在本文研究课题中即为与图像的像素对应。当输入图像像素越多时，相应的卷积神经网络中参与运算的像素点越多，需要进行的运算次数越多，即计算成本越高。由于用于测试实验的输入图像是根据原始数据集进行高斯模糊，并依据放缩比例进行降采样缩小处理后的图像，而进行不同放缩比例测试实验的原始数据集相同，因此，经过图像预处理后的输入图像的分辨率不同，即像素点数不同。放缩比例越大，输入的低分辨率图像的像素点数越少，图像超分辨率重建的计算成本越低，处理时间越短。

表 4-11 超分辨率图像重建速度对比

Dataset	Scale	RunTime (ms)	Processor	Core	Processor Clock (MHz)
Set14	3	46.50	CPU	8	3400
		4.7	GPU	3072	745
	4	40.37	CPU	8	3400
		4	GPU	3072	745

与基于高效子像素卷积神经网络的单图像超分辨率重建方法重建超分辨率图像的

运算速度相比的结果如表 4-11 所示。其中，GPU 对应的数据为 ESPCN[5]的实验结果。若结合处理器核心数量和处理器主频将运算时间进行归一化处理后，得到的结果显示本文的重建系统在处理速度方面与 ESPCN 的基本一致。

表 4-12 超分辨率图像重建运算时间对比

Dataset	Scale	Model_91 (ms)	Model_291 (ms)	SRCNN[27] (ms)	VDSR[12] (ms)	A+[26] (ms)
Set5	2	**33.53**	35.16	180	130	580
	3	**20.62**	21.12	180	130	320
	4	16.75	**15.54**	180	120	240
Set14	2	58.32	**53.14**	-	250	860
	3	**32.04**	32.31	390	260	560
	4	24.19	**23.67**	-	250	380

参考与 ESPCN 同时期或之前的研究成果，本文构建的基于卷积神经网络和子像素方法的图像超分辨率重建系统在超分辨率图像重建速度方面有明显的优势，重建时间大幅度减少，如表 4-12 所示。重建时间的大幅降低为实时视频图像的超分辨率的实现创造了可能性。

4.4 本章小结

本章阐述了图像超分辨率重建实验的数据集选择和预处理方法，以及一些实验超参数的设定，并重点总结分析了对本文构建的图像超分辨率重建网络的测试结果。

结果显示，在重建图像的图像质量方面，相比于传统的双三次插值法，利用将卷积神经网络应用于子像素的图像超分辨率重建方法得到的重建图像的峰值信噪比大幅提高，即能得到质量更高的重建图像。随着放缩比例的增大，重建图像的平均峰值信噪比下降，即重建图像的质量呈下降趋势。主要原因在于当放缩比例增大时，由原始图像根据放缩比例经过预处理得到的输入图像的分辨率对应降低，能提供的先验信息对应减少。

在超分辨率图像的重建速度方面，相比于与基于高效子像素卷积神经网络的图像超分辨率重建方法同时期或之前的其他重建方法，利用本文的重建系统进行图像超分辨率重建可以大幅度减少重建时间，提高重建效率。

此外，本文针对视频的超分辨率重建进行了研究。利用本文构建的图像超分辨率重建系统，成功实现了对视频超分辨率的实时输出和图像保存。

第五章　总结与展望

本章对毕业设计（论文）工作的前期调研、中期实验以及后期论文工作做一个整体的总结，并对未来基于本文的单图像超分辨率重建的工作提出设想。

5.1　论文工作总结

本文首先对图像超分辨率重建的研究背景和研究意义进行了讨论，对其内容和技术原理进行了阐述，并针对单图像超分辨率重建方法的发展，包括传统的重建方法以及基于卷积神经网络的重建方法，进行了深入的研究。

本文主要参考基于高效的子像素卷积神经网络模型的图像超分辨率重建方法[5]，搭建单图像超分辨率重建网络，将卷积神经网络应用于子像素的图像超分辨率重建。首先，本文深入研究基于高效的子像素卷积神经网络模型的图像超分辨率重建方法[5]，重点研究分析了高效子像素卷积神经网络的网络结构组成及其各部分作用、图像预处理过程，以及其中创新性地引入的高效子像素卷积层的组成及其核心的“周期性洗牌”操作的工作原理。在深入研究基于高效的子像素卷积神经网络模型的图像超分辨率重建方法的基础之上，本文将卷积神经网络与子像素方法相结合，构建单图像超分辨率重建网络，实现了单图像的超分辨率重建，并利用被大量文献资料普遍使用的公共图像数据集对构建的复现网络进行重建图像质量、重建速度方面的测试，与其他已有研究成果对比分析本文网络的超分辨率重建性能。在此基础之上，本文将搭建的图像超分辨率重建网络拓展应用到视频的超分辨率重建方面。本文将视频看作一系列单图像帧的序列组合，类比单图像超分辨率重建思想，将视频的图像帧依次重建得到超分辨率图像，再组合形成超分辨率视频。本文的视频超分辨率重建部分实现了实时的超分辨率视频输出和超分辨率视频图像的保存。

5.2　未来工作展望

从传统的人为预设网络参数的图像超分辨率重建方法，发展到利用计算机训练学习得到网络参数的方法；图像超分辨率重建的网络结构从传统的参数固定结构，发展到引入卷积神经网络进行参数学习的结构，进而加深网络、应用网络递归、引入其他网络结构等等，图像超分辨率重建过程中可以挖掘出越来越多图像内部隐含的特征信息，因此对待恢复的图像有更强的适应性。未来基于卷积神经网络的单图像超分辨率重建系统可能会加深现有网络，或在此基础上引入其他结构，改变提取图像特征的网络，更换新的由低分辨率特征映射重建高分辨率数据的网络结构（上采样层）等等，继续提高重建算法的泛化能力和重建精度。另外，通过扩大训练数据集、增加训练周期等方法可以对网络的重建能力进行改善。

视频的超分辨率重建仍有极大的潜在研究空间。除了画面中某些物体元素的突然出现或消失，或者出现镜头的切换的情况，视频的一个突出特点在于其普遍的时间连续性，

因此，在某一个单独的图像帧中出现的画面，其大部分元素可能出现在这一图像帧的相邻的一段图像帧中，即部分画面共享。因此，可以挖掘视频时域上的图像特征，作为解决图像超分辨率重建问题的隐式冗余信息，并结合单图像超分辨率重建技术实现更完善、性能更优的视频超分辨率。此外，由于视频中部分片段存在部分画面共享，某些图像帧之间的画面相关性极大，因此存在图像冗余信息。当对重建系统的重建速度有更高要求时，可以考虑在不影响最终观感的前提下，从视频中提取出关键的图像帧，丢弃非必要图像帧，之后再参考单图像超分辨率重建方法进行视频的超分辨率重建。

参考文献

[1] Gonzalez R C, Woods R E. Digital image processing. Second Edition[M]. Prentice Hall. 2002:2.
[2] Borman S, Stevenson R L. Super-Resolution from Image Sequences - A Review[C]// Circuits and Systems, 1998. Proceedings. 1998 Midwest Symposium on. IEEE, 2002:374.
[3] Farsiu S, M.D. Robinson, Elad M, et al. Fast and robust multiframe super resolution[J]. IEEE Transactions on Image Processing, 2004, 13(10):1327-1344.
[4] Yang C Y, Ma C, Yang M H. Single-image super-resolution: A benchmark[C]//European Conference on Computer Vision. Springer, Cham, 2014: 372-386.
[5] Shi W, Caballero J, Huszár F, et al. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network[C]//Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on. IEEE, 2016: 1874-1883.
[6] Harris J L. Diffraction and resolving power[J]. JOSA, 1964, 54(7): 931-936.
[7] Yang J, Wright J, Huang T, et al. Image super-resolution as sparse representation of raw image patches[C]// Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008:1-8.
[8] Dong C, Chen C L, He K, et al. Image Super-Resolution Using Deep Convolutional Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(2):295-307.
[9] Goodfellow I, Bengio Y, Courville A. Deep Learning[M]. The MIT Press, 2016.
[10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
[11] Sermanet P, Eigen D, Zhang X, et al. Overfeat: Integrated recognition, localization and detection using convolutional networks[J]. arXiv preprint arXiv:1312.6229, 2013.
[12] Kim J, Kwon Lee J, Mu Lee K. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1646-1654.
[13] Kim J, Kwon Lee J, Mu Lee K. Deeply-recursive convolutional network for image super-resolution[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1637-1645.
[14] Wang Z, Liu D, Yang J, et al. Deeply Improved Sparse Coding for Image Super-Resolution[J]. 2015:370-378.
[15] Osendorfer C, Soyer H, Smagt P V D. Image Super-Resolution with Fast Approximate Convolutional Sparse Coding[C]// International Conference on Neural Information Processing. Springer, Cham, 2014:250-257.
[16] Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network[C]//The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017, 1(4).
[17] Lai W S, Huang J B, Ahuja N, et al. Deep laplacian pyramid networks for fast and accurate super-resolution[C]//Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 2017: 624-632.
[18] Ledig C, Wang Z, Shi W, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[C]// Computer Vision and Pattern Recognition. IEEE, 2017:105-114.
[19] Kim J, Kwon Lee J, Mu Lee K. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1646-1654.
[20] Schulter S, Leistner C, Bischof H. Fast and accurate image upscaling with super-resolution forests[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3791-3799.

[21] Yang J, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. IEEE transactions on image processing, 2010, 19(11): 2861-2873.
[22] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. IEEE, 2001, 2: 416-423.
[23] Bevilacqua M, Roumy A, Guillemot C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[J]. 2012.
[24] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[C]//International conference on curves and surfaces. Springer, Berlin, Heidelberg, 2010: 711-730.
[25] Dai D, Timofte R, Van Gool L. Jointly Optimized Regressors for Image Super‐resolution[C]//Computer Graphics Forum. 2015, 34(2): 95-104.
[26] Timofte R, Smet V D, Gool L V. A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution[C]// Asian Conference on Computer Vision. Springer, Cham, 2014:111-126.
[27] Dong C, Loy C C, He K, et al. Learning a deep convolutional network for image super-resolution[C]//European Conference on Computer Vision. Springer, Cham, 2014: 184-199.
[28] Dosovitskiy A, Springenberg J T, Brox T. Learning to generate chairs with convolutional neural networks[C]// Computer Vision and Pattern Recognition. IEEE, 2015:1538-1546.

致　　谢

时光荏苒，白驹过隙。拙文将毕，四年本科生活就要过去了。我很怀念它。这四年来，特别是最后做毕业设计的这段时间，承蒙家人、老师、同学的支持与帮助，我收获许多，成长许多，在此表示诚挚的谢意。

感谢姜竹青老师。在近四个月的毕业设计工作中，姜老师在每周组会上的严格指导都让我对自己的研究课题、研究方法有更深入的认识。他的“高标准，严要求”使我印象深刻，让我更加理解学术研究的严肃。

感谢门爱东老师和王海婴老师。得益于门老师和王老师的指导与帮助，我的毕业设计工作得以更加充实、完善。

感谢曾正欣学姐和卢越学长。他们给予我的技术原理、测试实验、论文写作等方面的指导，让我更有效地理解文献资料的原理方法，使我的毕业设计工作更加顺利。

感谢毕设组的小伙伴们。和他们在一起，我学习到很多，包括学习研究的方法、生活的态度等等。有他们的陪伴，我在做毕设的这段时间里感到更加快乐、充实。

感谢这四年来教过我的老师们。他们渊博的知识、严谨的治学态度和谦逊的品质让我在获得知识的同时，对“厚德博学，敬业乐群”有了更深的理解。

感谢我的舍友们。这四年来，他们给予了我很多学习上的帮助。学习之余，我的生活也因他们而更精彩。四年之缘，今当分别；“前途似海，来日方长。”

特别感谢父母。这二十余年的健康成长离不开他们的悉心呵护和陪伴。感谢他们这些年对我做出的每一个选择的支持和建议。未来充满无限的可能，父母是我坚实的依靠。

最后，向这四年来所有帮助和支持过我的朋友们表示衷心感谢！

外 文 译 文

基于高效子像素卷积神经网络的实时单图像和视频超分辨率

摘 要

当前，一些基于深度卷积神经网络的模型已经在单一图像超分辨率重建的重建精确度和计算性能上取得了巨大的成功。在这些理论方法中，利用一个单一的滤波器，在重建高分辨率图像之前，将低分辨率的输入图像提升像素重建出对应的高分辨率图像空间，使用的滤波器的核心技术通常是双三次插值法。这就意味着超分辨率重建是在高分辨率空间进行操作的。我们认为这种方法是次最优的，而且增加了计算复杂度。在本篇文章中，我们在单一的 K2 GPU 上部署一个可以完成针对 1080p 视频的实时图像超分辨率任务的第一卷积神经网络。为了实现这一目标，我们提出一个全新的卷积神经网络结构。在这一全新的结构中，图像的特征映射是在低分辨率空间中被提取出来的。此外，我们引入一个高效的子像素卷积层，用来学习一个提高像素滤波器的参数序列，应用到最后的将低分辨率特征映射提升到高分辨率图像输出的重建任务。通过这种方法，我们有效地将超分辨率重建系统中的人为设定的双三次滤波器替换为更加复杂的、针对特定特征映射训练得到的像素提升滤波器，同时，还进一步降低超分辨率重建整体的计算复杂度。我们应用公共的图像数据集和视频数据集对提出的全新的网络结构进行性能测评，得到的测试结果显示出这一网络结构的性能有明显的提高（重建图像的峰值信噪比提升 0.15dB，重建视频的图像的峰值信噪比提升 0.39dB），并且相比于已有的基于卷积神经网络的重建方法，本文提出的结构的处理速度明显更快。

一　介绍

从低分辨率图像或视频恢复重建出相对应的高分辨率图像或视频是目前数字图像处理领域的一个十分活跃的研究课题。这项任务，通常被称为超分辨率重建，在许多领域都有直接的应用，比如高清电视、医学影像、卫星成像、面部识别以及监控系统等等。一个全局超分辨率重建问题假设低分辨率数据是相对应的高分辨率数据经过低通滤波（模糊）、降采样并添加噪声得到的。这是一个极其不适定的问题，因为在这些不可逆的低通滤波以及降采样操作中丢失了很多高频信息。此外，超分辨率重建是一个从低分辨率空间到高分辨率空间的一对多的映射，可能包含多种问题的解法，因此很有必要确定正确的解决问题的方法。在许多图像超分辨率重建方法中蕴含的一个关键的假设是：部分高频数据信息是冗余的，因此可以从低频分量中被重建出来。所以，图像超分辨率重建的实现就是一个推测问题，这个问题依赖于我们对问题中图像的统计信息构建的模型。

许多重建理论假设多张图像可以被用作从不同视角得到的相同场景的低分辨率图像，即具有唯一的先前仿射变换。这些方法可以被归类为多图像超分辨率重建方法。这些方法通过应用附加信息限制不适定问题，进而挖掘图像的显式冗余信息，并尝试反转图像的降采样操作。然而，这些方法通常需要计算上十分复杂的图像登记和融合过程，这些过程的精确程度直接影响重建结果的质量。另外一套重建理论是单一图像超分辨率重建方法。这些方法寻求对图像数据中的隐式冗余信息的学习，以此从单一的低分辨率图像实例中恢复重建丢失的高分辨率信息。通常，在图像应用方面，这些高分辨率信息的表现形式是像素的局部空间相关性，在视频应用方面表现为额外的时域相关性。在这种情况下，需要引入先验信息作为重建问题的解空间的限制条件。

1.1　相关工作

单一图像超分辨率重建理论的目标是由单一的低分辨率图像恢复重建相对应的高分辨率图像。当前流行的单一图像超分辨率重建理论方法可以分类为基于边缘的方法、基于图像统计信息的方法和基于区块的方法。在文献中包含对更多一般的单一图像超分辨率重建理论方法的详细介绍。最近比较兴盛的解决单一图像超分辨率重建问题的一类方法是基于稀疏性的。稀疏编码法是一种有效的解决方案。这一方法假设任意的自然图像都能够在一个转换域中被稀疏表达。这个转换域通常是图像元的字典，通过尝试寻找低分辨率区块和高分辨率区块之间的对应关系的训练可以学习到这个字典。这个字典里可以内嵌限制不适定的超分辨率问题的必要先验知识。基于稀疏性的方法的不足在于通过非线性重建引入稀疏性限制的做法的计算成本高。

近几年，通过神经网络获得图像表达的方法也应用在单图像超分辨率重建中。这些理论应用反向传播算法，在大规模图像数据集上训练学习低分辨率图像区块到高分辨率图像区块的非线性映射。堆叠的协同式自编码器被用来对低分辨率图像进行逐层的超分辨率重建。Osendorfer 等人研究出一种基于预测卷积稀疏性编码结构的扩展形式用于单图像超分辨率重建。由稀疏性编码理论发展出的一个基于深度卷积神经网络的图像超分辨率重建理论。Chen 等人提出使用多阶段可训练的非线性反应扩散（TNRD）替代可学习权重和非线性的卷积神经网络。Wang 等人受学习迭代收缩阈值算法（LISTA）的启发，

训练了一个端到端的级联稀疏编码网络，用于挖掘图像的自然稀疏性。网络结构不限于神经网络，例如随机树在单图像超分辨率重建的实现上也取得了成功。

1.2 动机和贡献

随着卷积神经网络的发展，人们对算法的效率，尤其是其计算和存储的成本越来越重视。和已有的模型相比，深度网络模型学习非线性关系的灵活性展现出其优越的重建精度。对于将低分辨率图像超分辨率重建高分辨率图像，低分辨率图像和高分辨率图像特定对应点的像素提升是十分必要的。

在 Osendorfer 等人的研究中，图像分辨率是在网络中间层逐渐提升的。另一种流行的方法是在网络的第一层或第一层之前提升图像分辨率。然而，这种方法有一些不足。第一，在图像增强阶段之前提升图像分辨率会增加计算复杂度。这对于卷积神经网络的应用尤为重要，因为处理速度取决于输入图像的分辨率。第二，在这些方法中用到的插值法，例如双三次插值法，不能为解决不适定重建问题提供额外的信息。

在 Dong 等人的研究中简单介绍了学习提升像素滤波器的概念。然而当时并未认识到将其与卷积神经网络整合作为超分辨率重建一部分的重要性，也没有对这一方法进行深入研究。此外，如 Dong 等人所述，没有针对输出尺寸比输入的大的卷积层的高效部署，同时，像卷积网络等优化好的方案并不简单地适用这种方法。

在这篇文章中，不同于以往研究，我们提出只在网络的末端提升图像分辨率，并且从低分辨率特征映射中超分辨率高分辨率数据。这样大部分超分辨率重建操作都不需要在复杂得多的高分辨率空间进行。为了实现这一目标，我们提出一种高效子像素卷积层，用来学习提升像素的过程。

这些贡献的优势有两点：

1）在我们的网络中，像素提升是在网络最后一层进行的。这就意味着每张低分辨率图像都是直接输入网络，通过在低分辨率空间进行非线性卷积提取图像特征。由于减小了输入图像的分辨率，我们可以在一个给定的区域中有效使用一个更小尺寸的滤波器对相同信息进行整合。分辨率和滤波器尺寸的减小也降低了计算和存储的复杂度，从而能够适用于实时高清视频的超分辨率。

2）针对有 L 层的网络，我们为 n_{L-1} 个特征映射学习 n_{L-1} 个而不是 1 个像素提升滤波器。另外，不使用显式插值滤波器意味着网络隐式学习超分辨率的关键过程。因此，相比将单一固定滤波器放在第一层，网络可以学到更好更复杂的从低分辨率到高分辨率的映射。这就意味着模型可以获得额外的重建精度。

我们使用公共基准数据集中的图像和视频对提出的方法进行验证，并和已有的方法进行性能对比。我们提出的方法取得了更好的性能，并且相对具有更快的处理速度。

二　理论

单图像超分辨率技术要实现由给定的低分辨率图像 I^{LR} 生成高分辨率图像 I^{HR} 的任务，其中，低分辨率图像是由相应的原始高分辨率图像按放缩比例 r 缩小得到的。为了由高分辨率图像得到相应的低分辨率图像，首先将高分辨率图像与高斯滤波器相卷积，模拟相机的点扩散函数，之后按放缩比例 r 对图像进行降采样。假设高分辨率图像和低分辨率图像有 C 个颜色通道，若高分辨率图像的图像尺寸是 $rH\times rW\times C$，则低分辨率图像的图像尺寸是 $H\times W\times C$。

为了解决单图像超分辨率重建问题，SRCNN 的恢复是建立在低分辨率图像的提升像素及插值后的版本上而不是低分辨率图像上。为了恢复超分辨率图像 I^{SR}，使用了一个三层卷积神经网络。在这部分，我们提出一种全新的网络结构，避免在将低分辨率图像输入网络之前提升其分辨率。在我们提出的结构中，我们首先直接对低分辨率图像应用一个 l 层的卷积神经网络，然后应用一个子像素卷积层提升低分辨率特征映射重建高分辨率图像。

对于具有 L 层的网络，前 L-1 层可以如下描述：

$$f^1(\mathrm{I}^{LR};W_1,b_1)=\phi(W_1*\mathrm{I}^{LR}+b_1) \qquad 式(1)$$

$$f^l(\mathrm{I}^{LR};W_{1:l},b_{1:l})=\phi(W_l*f^{l-1}(\mathrm{I}^{LR})+b_l) \qquad 式(2)$$

其中，$W_l,b_l,l\in(1,L-1)$ 分别是可学习网络的权重和偏差。W_l 是一个二维卷积张量，大小为$n_{l-1}\times n_l\times k_l\times k_l$，其中$n_l$是第 l 层的特征数，$n_0=C$，k_l是第 l 层的卷积核尺寸。偏差 b_l 是长度为 n_l 的向量。非线性函数（或称激活函数）ϕ是对应元素运算且是固定的。最后一层 f_l 将低分辨率特征映射转换为高分辨率图像 I^{SR}。

2.1　反卷积层

在从最大池化或其他图像降采样层重建分辨率时流行使用反卷积层。这种方法已经被成功应用在可视化层激活和利用高级别网络特征生成语义分割中。在 SRCNN 中应用的双三次插值法是反卷积层的一个特殊情况。反卷积层可以被视作一个卷积核与每个输入像素以步长 r 对应元素相乘，输出结果求和也被称为反向卷积。

2.2　高效子像素卷积层

另一种提升低分辨率图像的方法是在低分辨率空间步长为 $1/r$ 的卷积。这种方法可以被简单地实现为低分辨率空间到高分辨率空间的插值、穿孔或反池化后伴随一个在高分辨率空间步长为 1 的卷积。由于卷积作用于高分辨率空间，这些措施会将计算成本提高为 r^2 倍。

一个替代方案是，在低分辨率空间用一个尺寸为 k_s，权重间隔 $1/r$ 的滤波器 W_s 做的步长 $1/r$ 的卷积可以激活 W_s 的不同部分。落在像素之间的权重不会被激活，因此不需要计算。激活模式一共 r^2 个。基于不同的子像素分布，这些模式在卷积中周期性被激活。在本文中，我们提出一种实现这种方法的操作：

$$\mathrm{I}^{SR}=f^L(\mathrm{I}^{LR})=PS(W_L*f^{L-1}(\mathrm{I}^{LR})+b_L) \qquad 式(3)$$

$$PS(T)_{x,y,c} = T_{\lfloor x/r \rfloor, \lfloor y/r \rfloor, C \cdot r \cdot \mathrm{mod}(y,r) + C \cdot \mathrm{mod}(x,r) + c} \quad \text{式(4)}$$

通过对这 r^2 个特征映射进行“周期性洗牌（periodic shuffling）”操作，将每个通道的子像素数据周期性重新排布得到最后重建的高像素图像，即将通过第 l 层卷积神经网络输出得到的形如 $H \times W \times C \times r^2$ 的张量重新安排得到新的形如 $rH \times rW \times C \times 1$ 的张量，转换格式得到重建的高分辨率图像。

注意到我们没有在最后一层在卷积的结果上应用非线性。显然，当$k_L = k_s/r$并且 $\mathrm{mod}(k_s, \mathrm{r}) = 0$时，我们的方法就是在低分辨率空间的子像素卷积。我们将新提出的层成为高效子像素卷积层（ESPCN）。这个高效子像素卷积层在网络最后直接应用一个针对特定特征映射的提升滤波器从低分辨率特征映射重建高分辨率图像。

给定一个的包含高分辨率图像的数据集，我们生成相应的低分辨率图像，然后对应像素计算均方误差。均方误差用作训练重建网络的目标函数。

注意到在训练网络时可以不做前述的周期性洗牌操作。作为备选，我们可以提前整理训练数据以匹配周期性洗牌操作之前层的输出。因此，我们提出的高效子像素卷积层在训练阶段比反卷积层快$\log_2 r^2$，比在卷积前应用像素提升方法的理论快r^2。

三　实验

详细的数据测试报告放在了补充材料中，包括原始的图像和视频数据、降采样数据、超分辨率数据、在 K2 GPU 上得到的整体和单独数据的测试结果和运行时间。

3.1　数据集

在测试中，我们使用公共基准数据集，包括被单图像分辨率研究广泛研究的 Timofte 的数据集，91 张训练图像和 Set5、Set14 两个测试数据集；分别使用 BSD300 和 BSD500 中的 100 和 200 张和提供 136 张图象的 SuperTexture 数据集进行测试。对于最后的模型，我们使用从 ImageNet 数据集中随机挑选的 50000 张图像进行训练。沿用之前的研究方法，我们在这部分只考虑 YCbCr 颜色空间中的亮度通道，因为人对亮度更敏感。我们对每一个放缩比例训练一个特定的模型。

对于视频实验，我们使用被广泛应用的 Xiph 数据集中的 1080p 高清视频。这个数据集中有 8 个长度约 10 秒的、分辨率 1920*1080 的高清视频。另外，我们还用 Ultra Video Group 数据集，包含 7 个长度 5 秒的 1920*1080 视频。

3.2　实验设置

对于 ESPCN，设置 $l = 3$, $(f_1; n_1) = (5; 64)$, $(f_2; n_2) = (3; 32)$ and $(f_3; n_3) = (3; r^2)$。参数的设定参考 SRCNN 的三层 9-5-5 模型以及 2.2 节的等式。在训练阶段，将高分辨率图像 I^{HR} 和低分辨率图像 I^{LR} 再分别分割成 17r×17r 像素和 17×17 像素的子图像，步长分别为 $\left(17 - \sum \mathrm{mod}(f,2)\right) \times \mathrm{r}$和$\left(17 - \sum \mathrm{mod}(f,2)\right)$。这样可以确保原始图像中的所有像素作为训练数据的真值出现且仅出现一次。我们使用 *tanh* 代替 *relu* 作为最终模型的激活函数。

在 100 个周期后，代价函数不在变化，停止训练。初始学习率设定为 0.01，最终学习率设定为 0.0001，学习率在代价函数变化小于阈值时变化。最终层的训练慢 10 倍。在 K2 GPU 上使用 91 张图像进行 3 倍放大的训练用了 3 小时，最终模型用 ImageNet 的需要 7 天。我们使用峰值信噪比作为模型评价的指标。使用 Matlab 实现计算 SRCNN 的峰值信噪比和 Chen 的模型。

四　结论

在本文中，我们解释了在网络第一层使用不变的提升滤波器会得到比可变滤波器效果差的单图像超分辨率结果，并且需要更高的计算复杂度。为了解决这一问题，我们提出在低分辨率空间而不是高分辨率空间提取图像特征。为了实现这一目标，我们提出一个全新的子像素卷积层，能实现由低分辨率数据到高分辨率空间的超分辨率，而且和反卷积层相比，在训练阶段需要极少的额外计算成本。在基准数据集上的测试 4 倍放大得到的结果表明，相比于其他已有的使用更多参数的基于卷积神经网络的模型，我们的网络的运行速度更快（大于 10 倍速度），重建效果更好（图像+0.15dB，视频+0.39dB）。这使得我们的模型成为第一个可以只用一枚 GPU 实现实时高清视频超分辨率的模型。

五　未来工作

在处理视频信息时，可以假设相邻的视频帧中都包含一个画面的大多数信息，除了画面切换或物体偶然出现消失。这为视频的超分辨率提供了额外的隐式冗余信息。时空网络十分流行，因为它们在人的动作识别中充分利用视频中的时域信息。在未来，我们将研究将 ESPCN 应用于时空网络中，利用 3D 卷积从相邻的几张图像帧恢复特定图像帧。