

超分辨率重构遥感图像分类研究

谢海平¹ 李高原¹ 杨海涛² 赵洪利²

¹ 航天工程大学研究生院 北京 101416

² 航天工程大学航天信息学院 北京 101416

(xiehp18@163.com)

摘要 通常情况下,超分辨率技术能够使图像获得更好的视觉质量,从而有利于图像的目视解译。然而,超分辨率技术能否提升图像在应用于分类、识别等其他更高级计算机视觉任务时的效果呢?结合计算机视觉技术的发展现状,首先训练一个具有良好性能的分类模型,用于对经过不同超分辨率方法处理后的图像进行分类。利用分类准确率来衡量超分辨率重构处理对图像分类任务的影响。在遥感分类数据集上的实验显示,相比插值方法,经过超分辨率处理的图像在分类模型中取得了更高的分类准确率。这一结果说明,经过良好设计的超分辨率模型不仅能使低质图像获得更好的视觉质量,在分类模型中也能够被更准确地归入所属类别。证实了经过良好设计的超分辨率模型对于图像分类任务具有促进作用,并且还能有效提升低质图像在计算机视觉中的感知效果。

关键词: 超分辨率重构;图像分类;深度学习;图像处理;计算机视觉

中图分类号 TP751

Classification Research of Remote Sensing Image Based on Super Resolution Reconstruction

XIE Hai-ping¹, LI Gao-yuan¹, YANG Hai-tao² and ZHAO Hong-li²

¹ Graduate School, Space Engineering University, Beijing 101416, China

² Institute of Space Information, Space Engineering University, Beijing 101416, China

Abstract Usually, super-resolution technology makes the image get better visual quality, which is conducive to the visual interpretation of the image. However, will super-resolution technology improve the effect of image in the application of classification, recognition and other more advanced computer vision tasks? Combined with the development of computer vision technology, a new image classification model is trained to classify the images processed by different super-resolution methods. By measuring the classification accuracy of different types of images, the promotion effect of different super-resolution methods on image classification task is measured. Experimental results on remote sensing classification data set show that the well-designed super-resolution model has obvious effect on enhancing low-quality images. Compared with the interpolation method, the image processed by super-resolution achieves higher classification accuracy in the classification model, which proves that the well-designed super-resolution model can promote the image classification task. It is confirmed that super-resolution can effectively improve the perception effect of image in computer vision.

Keywords Super resolution reconstruction, Image classification, Deep learning, Image processing, Computer vision

1 引言

图像是信息的载体,图像质量对于图像信息提取和图像判读解译具有十分重要的影响。观察图像最重要的目标是从图像中获取所需的信息。

在图像退化与超分辨率过程中,高分辨率图像损失了大量高频信息,从而退化为低分辨率图像,而超分辨率的目标在于尽可能地重构出退化过程中损失的高频信息。由于存在信息损失和重构,高分辨率图像、低分辨率图像和重构图像所包含的信息各不相同,可用于判读解译的信息也不同。

对于超分辨率来说,重构图像所含信息量越接近高分辨率图像,则该模型的重构效果越好。通过与高分辨率图像进行比较,判断超分辨率重构图像中是否包含在图像退化过程

中损失的原图像关键信息,就可以从是否重构出原图像损失的关键信息的角度来对超分辨率模型的重构能力进行评价。因此,对于超分辨率模型,其重构图像能否为判读解译提供更多的有效信息,能否提升图像的判读解译效果,也可以成为评价超分辨率模型的一条标准。

在计算机图像处理中,超分辨率属于初级视觉处理任务,常被用来提升低分辨率图像的视觉质量,以便于后续的判读解译。因而超分辨率技术使图像获得了更好的视觉感受,能够获得更好的目视解译效果。然而,超分辨率技术能否提升图像在其他计算机视觉任务中的效果呢?如图像分类、目标识别、语义分割等^[1]。本节主要研究了超分辨率技术对遥感图像分类任务是否有促进作用这一问题,并通过设计一系列实验,来探究超分辨率对于遥感图像分类问题的影响。

2 相关工作

2.1 遥感图像分类

图像是信息的载体,图像质量对于图像信息提取和图像判读解译具有十分重要的影响。观察图像的最重要的目标在于从图像中获取所需的信息。图像分类是指将图像信息所反映出的特征差异按照一定的规则对其类别属性进行划分的过程。对于人类视觉而言,指出图像的类别是一个十分简单的问题。然而对计算机而言,数字图像是由一个个像素数值组成的数值矩阵,对其进行分类首先需要从数值矩阵中提取出图像特征,但依据图像特征描述对图像进行分类是指十分复杂的问题。

遥感图像分类是指对遥感图像进行分析和信息提取,其目的是获得遥感图像中所存各部分地物的信息。长期以来,由于遥感图像判读的复杂性,由判读人员进行目视解译一直是遥感图像解译的主要实施方法,即通过直接观察遥感图像中色调分布、大小形状、空间关系这3种地物特征并进行分析,进而获得图像中地物目标的信息。然而,由于遥感图像数据量大,所含信息复杂,使用目视解译效率往往很低,难以满足高效、大批量处理信息的要求。

与人工判读解译不同,计算机图像判读解译将计算机作为图像判读解译的主体,通过读取输入图像的像素矩阵,从而获取所需的数据和信息。得益于计算机视觉技术的发展,近年来,深度学习在很多传统的图像处理任务中得到应用,其不仅能够应对大批量的数据处理任务,很多时候还表现出不亚于人工判读解译的准确率。

传统的图像分类方法依赖于人工设计的规则和模板,在复杂语义理解中鲁棒性往往较差,其应用也受到诸多限制。传统方法需要使用人工设计特征,不仅高度依赖于设计人员的经验,同时存在操作繁琐、效果不佳的问题。而卷积操作能够自动提取图像特征,取代了传统方法中手工设计特征的操作,因此基于卷积结构的计算机视觉模型逐渐受到人们的关注。

20世纪60至70年代,生物神经学领域的研究学者发现,猫的视觉皮层中既有简单细胞,也有复杂细胞,简单细胞对图像的边缘、角点等细节区域更加敏感,复杂细胞对图像的空间关系更加敏感,能够对不同角度、不同位置的图像进行处理,并据此提出了“感受野”的概念。日本学者 Fukushima 在感受野概念的基础上,受视觉神经系统对外部环境的感知机制的启发,提出了神经认知模型。该模型被认为是首个可实现的卷积神经网络。1989年,LeCun等对神经认知模型进行了重大改进,在网络中使用卷积层和下采样层相结合的方式,于1998年提出并设计了能够用于手写数字识别的网络(LeNet),初步形成了现代卷积神经网络的雏形。

理论上,卷积神经网络可以被视作一类特殊的多层感知机结构或者前馈神经网络。卷积操作能够提取图像特征,在基于深度学习的图像处理中,这一特性十分重要。从本质上来说,卷积神经网络是表达输入与输出之间映射关系的一种方式。通过大量数据的训练,卷积神经网络能够自动学习输入与输出之间的端到端表示,而无需借助人为了拟定的数学公式。

2012年,Alex等提出的卷积神经网络模型“AlexNet”在

图像分类赛事中取得了优异成绩,效果远超传统方法。此后,随着网络的加深以及结构的改进,模型效果越来越好,出现了VGG-net,Resnet等深度学习结构。在2015年的ImageNet视觉识别比赛(ILSVRC)中,基于残差网络的网络模型获得了图像分类和物体识别两个赛道的冠军,分类准确率已经超越了人眼的水平。换言之,计算机视觉在某些领域中已经具备和人眼视觉系统几乎一样的信息提取能力。

2.2 超分辨率图像分类

图像分类即将图像按照语义信息划分到所属的类别,在遥感图像判读解译中具有十分重要的研究意义与价值。在计算机视觉研究中,图像分类是最为核心的问题之一,也是很多其他高级视觉任务的基础,如目标检测、图像分割、行为识别等。因此,开展图像分类的相关研究对于计算机视觉来说十分重要。

在遥感图像中,每个像素所代表的地面区域的尺寸大小与遥感图像的空间分辨率是对应的,该尺寸是限制地表地物能否在遥感图像中被分辨出来的重要因素^[2]。在遥感图像中,由于图像退化导致图像的分辨率降低,反映在遥感地物中,地物的可分辨能力下降了,从而极大地影响了遥感图像的判读解译。因此,超分辨率重构的目标之一就是提高退化图像的可解译能力。

从人眼观察的角度来说,图像退化后图像中的目标会愈发难以分辨,清晰度大幅下降。从计算机视觉的角度来说,由于关键像素的损失,图像退化同样会对计算机视觉的信息提取产生影响^[3]。

对低质图像进行超分辨率重构的目标在于尽可能地恢复出图像在退化过程中损失的信息部分,以获得更好的视觉感受以及可解译能力。从可解译性的角度考虑,超分辨率模型的重构能力越强,重构出的图像就能够提供更丰富的有效信息,从而获得更好的判读解译效果。

超分辨率任务属于低级视觉处理任务,其结果通常被用于更高级的视觉信息处理与分析(如分类、目标识别检测、图像分割等图像解译任务)。目前超分辨率的研究主要集中在两个方面:一是获得更好的视觉效果,即更好的人眼感知质量;二是获得更接近原图像的结果,即重建图像与降质前的图像之间的像素偏差尽可能小,并未考虑超分辨率结果是否有利于更高级别的视觉处理与分析^[4]。

以图像二分类为例,输入有猫和狗两类图像,需要将图片划分到正确的类别。对于人眼视觉来说,这样的任务并不难,稍有认知能力的人都能判断。对于一个经过良好训练的计算机分类模型来说,分类准确率也能达到很高的水平。

图像退化以后,由于高频部分信息的损失,图像用于分类的关键信息很可能丢失,导致无法正确分类。经过超分辨率重构以后,超分辨率是否重构出这部分关键信息,直接关系到图像能否被归入正确的类别^[5]。图像退化与超分辨率过程示意如图1所示。



图1 图像退化与超分辨率重构

Fig. 1 Image degradation and super resolution reconstruction

对低质图像进行超分辨率重构的目标在于尽可能地恢复出图像在退化过程中损失的信息部分,以获得更好的视觉感受以及可解译能力^[6]。从可解译性的角度考虑,超分辨率模型的重构能力越强,重构出的图像就能够提供更丰富的有效信息,从而获得更好的判读解译效果。

由以上分析可知,在图像退化与超分辨率过程中,高分辨率图像损失了大量高频信息,从而退化为低分辨率图像,而超分辨率的目标在于尽可能地重构出退化过程中损失的高频信息。由于存在信息损失和重构,高分辨率图像、低分辨率图像和重构图像所包含的信息各不相同,可用于判读解译的信息也不同。

2.3 图像分类精度评价

总体分类精度为全体被正确分类的样本数目除以整体样本数。图像分类中通常使用混淆矩阵来对结果进行评价。混淆矩阵也被称为误差矩阵,是一个用于表示每个图像真实类别与预测类别的比较阵列。通常情况下,阵列中的每一列代表样本经由分类器得到的预测类别,列中数值代表该预测类别中样本对应的实际类别;每一行代表样本对应的实测数据的类别(即真实类别),行中数值代表实际类别中样本对应的预测类别。在混淆矩阵中,所有被正确分类的样本有规律地分布在矩阵的对角线上,其数值代表每类样本被正确归类的样本数目。整体样本数等于所有参与分类的样本总数。该项指标表示参与分类的随机样本被正确归入所属类别的概率。

一般来说,分类精度能够综合反映出分类器对各类样本的判断情况,是一个不可忽视的指标。但是在使用准确率对分类模型进行评价时,需要注意正样本与负样本的数量差距。例如,假如待检样本有 50 例,其中正样本数目为 45,负样本数目为 5,即便分类器将所有的样本判定为正样本,准确率仍然能够达到 90%,但是很难体现出分类器的分类效果。为了避免出现这种情况,参与分类的各类样本数量应保持一致。

3 基于超分辨率的遥感图像分类研究

3.1 研究方案

本节主要研究超分辨率对图像分类任务的影响。首先需要训练一个遥感图像分类模型,用于对图像进行评估。通过对图像进行不同的预处理操作后,输入到训练好的分类模型中进行图像分类,评估超分辨率对图像分类的影响。主要的研究步骤如图 2 所示。

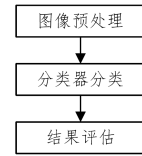


图 2 研究步骤

Fig. 2 Research steps

3.2 图像预处理

我们将选用的图像分类数据集划分为训练数据和测试数据两个部分。训练数据中的图像主要用于训练分类模型和超分辨率模型,测试数据中的图像则用于评估超分辨率重构图像的分类效果。

对于测试数据,通过进行不同的预处理,获得不同分辨率条件下的数据。测试超分辨率重构图像应用于分类模型的效果。首先对测试数据集进行下采样,以获得低分辨率图像;随后对低分辨率图像进行超分辨率,以获得超分辨率图像。同时,使用双三次插值作为超分辨率组的对比结果。下采样方式采用常用的双三次下采样。

经过数据处理后,得到了 HR 图像、LR 图像、插值图像以及超分辨率图像的各组成数据,将各组成数据输入到训练好的超分辨率模型中进行分类,并比较分类结果。

3.3 分类器分类

图像分类的研究一直是计算机视觉的重要内容。2012 年 Krizhevsky 等^[7]设计的卷积神经网络结构 Alex-Net 在图像分类比赛中取得了远超传统方法的效果,由此深度学习备受人们关注。与传统的机器学习算法相比,深度学习算法最特别之处在于端到端的学习方式,即利用卷积神经网络,实现从数据中提取图像特征,在学习之前无需经过特征提取。经过大量数据的训练,卷积神经网络能够学习到图像对应的类别表示。

在 Alex-Net 之后,出现了一系列的卷积神经网络模型,这些模型在图像分类比赛中不断取得更好的成绩。随着网络结构的变革,网络越来越深,结构越来越巧妙,图像分类的错误率也越来越低。2015 年,基于残差网络的图像分类算法的 top-5 错误率低至 3.5% 左右,而同时期的人眼识别错误率为 5.1% 左右,其识别能力已经超越了人眼^[8]。

本文选择残差网络 resnet-152 作为图像分类模型,网络结构如图 3 所示^[9]。

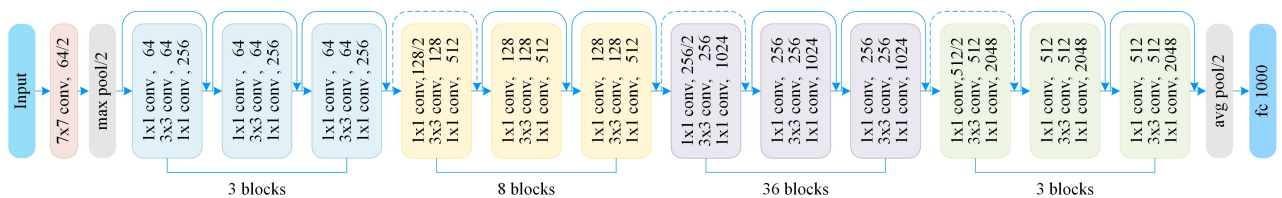


图 3 Resnet-152^[9]结构示意图

Fig. 3 Structure of Resnet-152^[9]

在用于图像分类的众多模型中, resnet-152 网络^[9]具有以下几个特点:首先,该网络结构成熟,分类效果得到了人们的公认,在多个数据集上的分类准确率超越了人眼;其次,该网络深度足够深,能够达到十分优异的分类性能,同时残差结构收敛快,训练耗时大大减少;此外,该模型已经集成到多个深度学习框架中,无需自行构建网络。

4 实验与分析

4.1 实验数据

本文主要用到两个数据集。数据集一是由 UC Merced 计算机科学实验室整理的遥感图像分类数据集 UC Merced Land-Use Dataset^[1],数据集二为超分辨率领域常用的

2K 数据集 DIV2K。

UC Merced Land-Use Dataset 原始数据来自美国国家调查局所拍摄的航空遥感影像,经过 UC Merced 计算机科学实验室的整理,主要用于研究土地利用。该数据集经过了预处理,属于较为纯净的遥感图像数据,共有 21 个类别,每个类别各 100 张的场景图像,每张图像像素大小为 256×256 ,分辨率为 0.3m。

我们从中选择 8 类感兴趣的目标图像用于实验,分别是飞机、棒球场、建筑物、高速路、港口、跑道、储油罐、网球场,每类图像 100 张。

将所选择的图像分类数据集中的每类图像按照 6:4 的比例划分为两个部分。一部分作为训练集,用于分类模型和超分辨率模型训练,包含 8 类图像,每个类别各 60 张;另一部分作为测试数据集,用于评估超分辨率对图像分类效果的影响,包含 8 类图像,每个类别各 40 张。

4.2 图像预处理

由于遥感图像在成像过程中存在大量的干扰,成像质量往往较差。考虑到图像之间具有相似性,本文首先使用高质

量的包含丰富图像结构信息的 DIV2K 数据集对超分辨率模型进行预训练,再使用光学遥感图像数据集对该模型进行训练,以提升超分辨率的效果,获得更高质量的超分辨率重构图像。图像的下采样均使用双三次下采样完成。

超分辨率实验中,我们选择 Sc^[10],VDSR^[11]和 EDSR 算法^[12]用于对比。其中,Sc 算法是一种基于传统机器学习的稀疏表示方法^[10],VDSR 算法是首个将残差思想应用于超分辨率的深度学习算法^[11],EDSR 算法是目前超分辨率领域中重构能力最好的算法之一^[12]。

重构图像示例如图 4 所示,可以看到,EDSR 算法重构效果最好,重构结果最接近原始图像。进一步计算各种算法在测试集上进行不同倍数的超分辨率后的客观评价指标 PSNR、SSIM 的表现(平均值)^[13-14],如表 1 所列。可以看到,各种算法重构结果的客观评价指标值相比 Bicubic 插值法均具有明显提升,说明了超分辨率算法的有效性能够有效提升图像质量。对比 Sc 算法和 VDSR 算法,EDSR 算法在各个放大倍数(2x, 3x, 4x)下的 PSNR 和 SSIM 值均具有显著的提升,展现出了该深度算法的优越性。

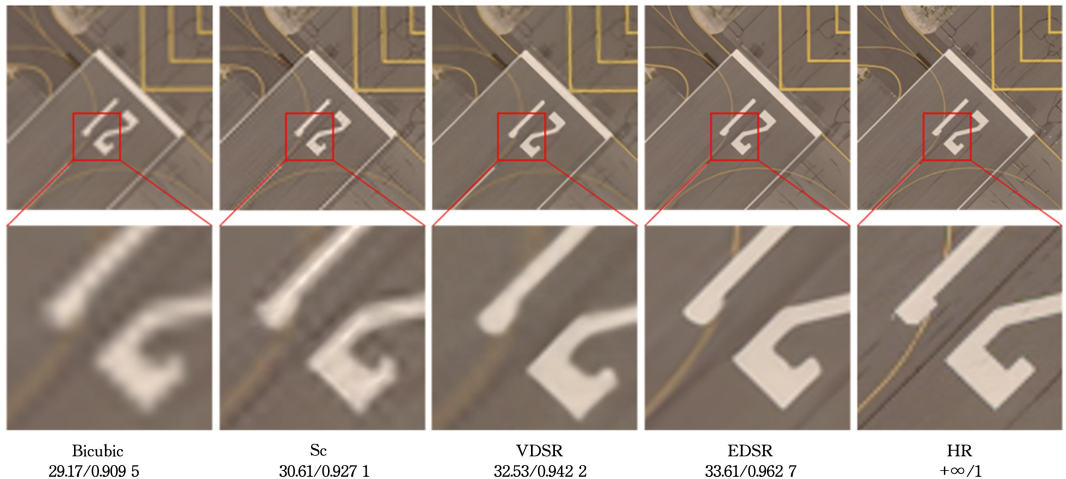


图 4 图像超分辨率结果(4x)

Fig. 4 Image super resolution result(4x)

表 1 超分辨率重构图像客观评价指标对比

Table 1 Comparison of objective evaluation indexes of super resolution reconstructed images

倍数	Bicubic		Sc ^[10]		VDSR ^[11]		EDSR ^[12]	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
2	30.77	0.9089	31.84	0.9234	32.36	0.9028	33.95	0.9419
4	25.53	0.7306	26.52	0.7801	27.06	0.7583	27.69	0.7932

4.3 结果与分析

首先将测试图像输入到分类模型中进行测试,得到原图像的分类准确率。对未经处理的初始图像的分类结果显示,图像平均分类精度达到 95%以上,显示出本文训练的分类模型具有良好的分类能力。

随后将经过不同处理的测试图像输入到分类模型中分别进行分类,得到各组数据的分类结果。统计分析各组图像的分类结果,得到表中的分类精度。表 2 与表 3 分别列出了图像在 2 倍和 4 倍放大倍数下经过不同分辨率处理的图像的分类精度。

表 2 图像分类精度(2x)

Table 2 Image classification accuracy (2x)

类别	倍数为 2x 时的图像分类精度/%					
	原图	下采样	Bicubic	Sc ^[10]	VDSR ^[11]	EDSR ^[12]
飞机	97.5	95	95	95	95	95
棒球场	90	82.5	82.5	97.5	95	87.5
建筑物	100	92.5	95	97.5	100	100
高速路	95	85	87.5	85	87.5	92.5
港口	92.5	87.5	87.5	90	90	92.5
跑道	97.5	85	85	87.5	87.5	92.5
储油罐	97.5	95	97.5	95	97.5	97.5
网球场	95	90	92.5	92.5	92.5	95
平均	95.6250	89.0625	90.3125	92.5000	93.125	94.0625

表3 图像分类精度(4x)

Table 3 Image classification accuracy (4x)

类别	倍数为4x时的图像分类精度/%					
	原图	下采样	Bicubic	Sc ^[10]	VDSR ^[11]	EDSR ^[12]
飞机	97.5	85	90	92.5	92.5	92.5
棒球场	90	70	72.5	75	75	80
建筑物	100	80	80	87.5	87.5	90
高速路	95	75	72.5	75	77.5	82.5
港口	92.5	77.5	80	90	85	85
跑道	97.5	72.5	75	75	77.5	80
储油罐	97.5	85	85	85	85	90
网球场	95	75	80	82.5	82.5	85
平均	95.625	77.5	79.375	82.8125	82.8125	85.625

由不同分辨率处理的图像分类精度可知,分辨率变化对图像分类结果具有较大的影响。

经过2倍下采样以后,图像分类准确率下降了约6.8%。经过4倍下采样以后,图像分类准确率下降了约19%。经过双三次插值或超分辨率处理后,图像分类准确率又有所提高。

3种超分辨率方法的分类结果都优于双三次插值。原因是,对于低分辨率图像来说,双三次插值并没有引入任何额外的有效信息,因此对图像分类效果的提升也十分有限。而超分辨率在重构图像并提升图像分辨率的同时,引入了外部数据信息进行重构,使图像具有更好的细节信息。

此外,在视觉效果和客观评价指标上都表现最优异的EDSR算法,其重构图像分类准确率在各算法中也是最高的。

总的来说,可以得出以下几点结论:

(1)随着图像的分辨率降低,图像分类的准确率也随之下降,这是由于图像分辨率降低导致图像损失了大量的信息,进而影响了分类模型的信息提取;

(2)对低分辨率图像进行双三次插值能够略微提高图像的分类准确率,但是提升程度很有限,这是因为双三次插值并没有引入额外的有效信息;

(3)相比双三次插值,超分辨率方法能够显著提高图像的分类准确率。这是因为经过良好训练的超分辨率模型能够重构出更多的图像信息,从而为图像分类提供更多的信息支持。

结束语 本文主要研究了超分辨率技术对图像分类是否有促进作用这一问题,首先训练了一个神经网络分类模型,随后利用经过下采样、插值、超分辨率等方法获得不同的分辨率处理的图像进行图像分类实验。在遥感图像分类数据集上的实验显示,超分辨率重构技术能够有效提升退化图像在计算机视觉分类模型中的分类效果。随着图像退化倍数变大,分类效果的提升越明显。本文研究对于超分辨率技术在计算机视觉中的应用具有重要意义。下一步的研究重点将是如何设计超分辨率模型以获得更好的图像分类效果。

参考文献

- [1] CHENG G, HAN J, LU X. Remote Sensing Image Scene Classification: Benchmark and State of the Art[J]. Proceedings of the IEEE, 2017, 105(10): 1865-1883.
- [2] MA L, LIU Y, ZHANG X, et al. Deep learning in remote sensing applications: A meta-analysis and review[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 152: 166-177.
- [3] VAN OUWERKERK J D. Image super-resolution survey[J]. Image and Vision Computing, 2006, 24(10): 1039-1052.

- [4] SMORENBURG A, LEDEBT A, FELTHAM M G, et al. The positive effect of mirror visual feedback on arm control in children with Spastic Hemiparetic Cerebral Palsy is dependent on which arm is viewed[J]. Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale, 2011, 213(4): 393-402.
- [5] NASROLLAHI K, MOESLUND T B. Super-resolution: a comprehensive survey[J]. Machine Vision and Applications, 2014, 25(6): 1423-1468.
- [6] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C] // European Conference on Computer Vision. Springer, Cham, 2016: 391-407.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2): 1106-1114.
- [8] JING L, YANG X, TIAN Y. Video You Only Look Once: Overall Temporal Convolutions for Action Recognition[J]. Journal of Visual Communication and Image Representation, 2018, 52 (APR.): 58-65.
- [9] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 770-778.
- [10] YANG J, WRIGHT J, HUANG T, et al. Image Super-Resolution Via Sparse Representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [11] KIM J, LEE K J, LEE M 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.
- [12] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR). 2017: 136-144.
- [13] MAO X J, SHEN C, YANG Y B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections[C] // Proceedings of the 30th International Conference on Neural Information Processing Systems. 2016: 2810-2818.
- [14] HORE A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C] // 2010 20th International Conference on Pattern Recognition. IEEE, 2010: 2366-2369.



XIE Hai-ping, born in 1996, postgraduate. His main research interests include deep learning and image processing.



YANG Hai-tao, born in 1978, Ph.D., associate professor. His main research interests include image processing and decision making.