

交互式深度学习计算机视觉理论教学研究

邓丽珍 张冉冉

南京邮电大学通信与信息工程学院, 中国·江苏 南京 210000

【摘要】为了让学生更好的理解并掌握基于深度学习的计算机视觉,本课程采用了理论和实践一体化并结合交互式的教学方法。本文以基于深度学习的图像分割算法为例,讲解理论知识和对应的实例,采用交互式方法让学生参与到实验过程中。在本例中当学生选择不同的交互点时会产生不同的图像分割结果。这种教学方法将可以让学生对所学知识更加融会贯通。

【关键词】计算机视觉;深度学习;交互式;图像分割

Interactive Deep Learning Computer Vision Theory Teaching Research

Deng Lizhen, Zhang Ranran

School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210000, Jiangsu, China

[Abstract] In order to allow students to better understand and master computer vision based on deep learning, this course adopts the integration of theory and practice and combines interactive teaching methods. This paper takes the image segmentation algorithm based on deep learning as an example to explain the theoretical knowledge and corresponding examples, and uses an interactive method to allow students to participate in the experimental process. In this example, different image segmentation results are produced when students select different interaction points. This teaching method will allow students to become more integrated with what they have learned.

[Key words] computer vision; deep learning; interactive; image segmentation

深度学习是计算机视觉^[1]最重要的研究方向之一,目前基于深度学习的计算机视觉研究方向非常热门,其模型和算法发展迅速,工业应用范围广,如在图像分割^[4],人脸识别^[5],自动驾驶^[6]等方面的应用。而且在学术研究和项目研究上都取得了累累硕果,吸引了各个学科的研究人员加入这一领域做进一步研究。

“计算机视觉”课程是电子类专业的一门重要的专业课程,为了让学生接触到最前沿的计算机视觉研究方向,培养学生求真务实的科学态度、独立思考的科学精神、团队合作的科学素养,掌握计算机视觉领域基本的原理、概念、方法及关键技术,初步具备在该领域分析和解决问题的能力,为今后从事相关领域的研究和开发打下扎实的基础,所以开设了本课程“基于深度学习的计算机视觉”。在学习本课程之前,需要用到的理论和实践知识涉及:《高等数学》、《矩阵论》、《概率论》、《最优化》以及编程语言Python,对初学者来说将会面临一些挑战,所以对于本课程的讲授需要做到理论和实践一体化^[2],并采用交互式学习方法让学生真正的动手动脑参与到实验中,在实验过程中再次回顾课堂教师讲解的原理和算法,对这些知识点做到真正理解。教师也可以适当给学生介绍计算机视觉研究方向的当前领先技术,在提高学生理论学习、实践能力的同时,培养学生们对基于深度学习的计算机视觉研究方法的兴趣,挖掘学生们在这一领域的潜力。

1 交互式深度学习计算机视觉课程教学研究

本课程涉及到的其他学科的相关知识较多,原理和概念较综合,模型和算法较复杂,为了让学生更好的掌握本课程,本文采用理论和实践一体化并结合交互式学习模式。首先,教师要对课程的理论基础进行剖析讲解;其次教师在教学中充分利用实践活动,讲解知识点对应的实例的逻辑思想和代码;然后在交互式学习方法下,不仅仅学生自己动手,同时让学生与教师,学生与学生之间配合,在不断的循环合作中完成实验作业。交互式的学习方法与之前学生生硬的运行教师提供的或者网络现有的代码的学习方法相比,可以使学生更好的掌握、理解理论知识点和算法。

本文对于基于深度学习的计算机视觉课程的教学方法的研究

主要创新点是:交互式学习方法。从具体实施教学的三个方面进行研究:课程理论讲解、理论基础对应实例、交互式学习。与传统的教学方法相比,本课程采用的理论与实践一体化并结合交互式的教学方法可以让学生更好的掌握基于深度学习的计算机视觉的原理和算法,激励学生的创造性思维。

1.1 课程的理论基础讲解

任何课程的讲授都要以理论知识为基础,而基于深度学习的计算机视觉课程包含很多基本原理与基础算法。为了防止学生觉得本课程难以学习,让学生快速入门本课程,在讲授本课程的理论时,教师可以采用从浅入深,结合具体例子,使用流程图等方法,讲解基础算法并扩充前沿发展,使学生理解和掌握知识点并从中获益。

本文以深度学习对图像做分割为例,教学的主要任务是讲解图像分割知识点,引导学生学好理论知识。简要给出图像分割^[3]的类型:语义分割,实例分割,全景分割。其中本文讨论是基于卷积神经网络(CNN)的图像语义分割,该算法的思想是先采用下采样加上上采样,再提取出多种特征值,对每个特征值赋予权重进行加权,最后采用卷积神经网络分析每个像素点的特征,完成图像分割。

1.2 理论基础对应实例

实例是采用深度学习算法对二维医学图像做图像分割^[7],目标图像例如:图1(a)、(b)磁共振图像。

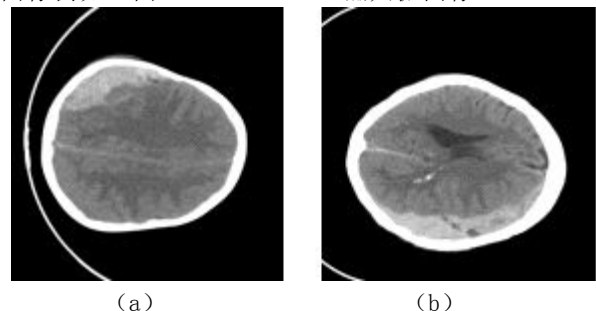


图1 磁共振图像

1.3 交互式学习

对于本课程采用的交互式学习方法, 学生不是直接运行教师汇编的代码或者网络上已有的代码进行自动图像分割, 而是在此基础上需要学生进行交互。

该例题的流程如图2所示:

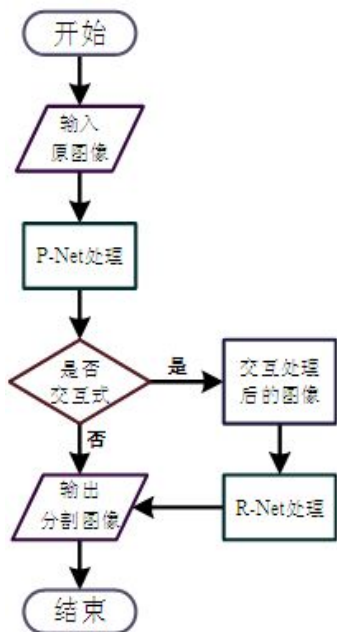


图2 算法流程图

图像分割算法的伪代码和注释如图3所示:

```
import cv2
def P_Net(image): #CNN卷积神经网络
Input_image = cv2.imread(image_path)#读取原输入图像
Initial_Segmentation = P_Net(Input_image)#得到一个自动分割图像
User_answer = input("Agreed by the user")
#判断用户对结果是否满意, 是否需要进行第二次图像分割
if User_answer:
Finally_segmentation = Initial_Segmentation
#交互点或者交互区域
User_Interactions_image = input(User_Interactions)
#再次得到一个自动分割图像
Refined_segmentation = P_Net(User_Interactions_image)
#在交互作用下分割更准确
Finally_segmentation = Refined_segmentation
```

图3 图像分割算法的伪代码和注释

根据深度学习原理, 输入目标图像在实验前需要做预处理, 用神经网络的训练集的均值和标准差对目标图像做归一化。之后通过垂直、水平翻转或者随机旋转图像角度和随机缩放因子增加

数据, 其中旋转角度的范围是 $\left[-\frac{\pi}{8}, \frac{\pi}{8}\right]$, 随机缩放因子的范围

是 $[0.8, 1.25]$ 。使用交叉熵损失函数^[8]和随机梯度下降法 (SGD)^[9]进行优化, 其中批量大小设置为 1, 动量设置为 0.99, 权重衰减设置为 0.0005。每迭代 5000 次, 学习速率都减一半。由于适当初始化 P-Net^[10]和 CRF-Net (f)^[11]有助于提高联合训练的收敛速度, 因此我们用该方法分三步对 P-Net 进行训练:

(1) P-Net 的初始学习速率是 0.001, 最大迭代次数 100000。

(2) CRF-Net (f) 中的配对网络要进行预先训练。

(3) P-Net 和 CRF-Net (f) 联合训练, 初始学习率为 0.000001 和最大迭代次数 50000。

在使用 CRF-Net (f) 训练 P-Net 后, 学生可以进行交互标记出第一次自动分割结果未分割的部分和过度分割的部分。根据算法流程图 (图 2) 可以清晰的得出图像分割的步骤: 首先使用卷积神经网络 P-Net 对每个训练图像进行自动分割; 其次通过交互在每个错误分割的区域随机标记 n 个像素; 然后使用交互标记的图像对训练数据做初始分割, 通过随机梯度下降法实现 R-Net^[12] 和 CRF-Net (fu) 的联合训练, 得出采用交互式方法的图像分割结果。

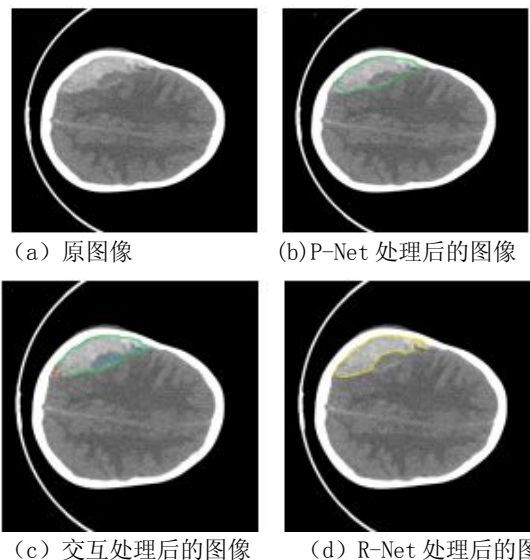


图4 磁共振图像 (a) 分割过程中的图像

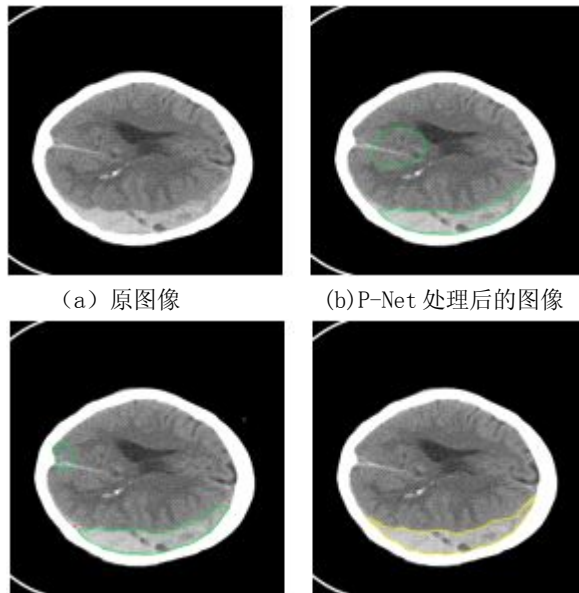


图5 磁共振图像 (b) 分割过程中的图像

在该例题中, 第一次通过卷积神经网络 (P-Net) 对图像自动分割的结果为图中绿色包围区域 (如图 4 (b)、图 5 (b)); 学生在此图片基础上进行交互, 标记的红色点代表未充分分割的部分, 蓝色点代表过度分割的部分, 图 4 (c) 和图 5 (c) 是两种标记图像; 把标记图像作为第二次卷积神经网络 (R-Net) 的输入图像, 得到图像分割结果为图中黄色包围区域 (如图 4 (d)、图 5 (d))。通过二张磁共振图像分

割的过程可以得出交互式图像分割的结果比自动图像分割的结果要好很多。

通过实验可以得出：对图像采用卷积神经网络（P-Net）做自动分割时，输入图像得到的图像分割的结果是相对固定的，不能根据数据、图像和不同的情况进行具体分析，运行的结果很难达到理想目标；而采用交互式方法对自动分割的图像标记交互点，对二维医学图像产生的分割结果非常好。

在本实验中，学生通过交互参与图像分割可以体会到自动分割与图像分割的差别，学生在实验的基础上可以进一步理解和掌握基于深度学习的图像分割的原理和算法。交互式学习方法还有可能激发学生的思考和探索能力，学生可以与教师讨论或者自己查找问题的解决方法，还可能挖掘部分学生对计算机视觉研究方向感兴趣，为他们日后投入到此研究方向奠定基础。

2 新教学模式对教师的要求

与之前的一般的教学方法相比，理论和实践一体化与交互式学习相结合的教学模式需要讲授课程的教师有更高的专业素养。首先，教师要有很强的专业知识，因为对于刚刚学习本课程的学生而言，可能会有些吃力，所以讲授课程时教师需要把理论知识与实际问题结合在一起，从浅入深，从旧到新引导学生慢慢理解新知识，减轻学生的学习压力。其次，授课教师要有清晰的逻辑，可以使用思维导图或流程图给学生讲授算法，让学生学习的知识点理解更加清楚。然后教师要结合实例，用例子让学生理解怎样应用学习的原理和算法，做到理论与现实相结合，而且讲授本课程采用交互式学习方法，代码和算法将会变得更加复杂，这需要教师有很强的编程能力。最后，交互式学习方法与以往的教学方法不同的，不再是让学生简单的搬运代码，而是为了让学生通过交互式方法学好本门课程。这种学习方法会提高了学生对知识点的理解，学生还可能提出一些新颖的问题，需要教师有较强的科研能力以及了解该领域的前沿发展知识。

3 结束语

本文根据“基于深度学习的计算机视觉”课程的特点，并结合之前的教学经验，提出理论和实践一体化与交互式学习相结合的教学方法。从课程理论基础讲解、理论基础对应实例、交互式学习三个部分研究该方法的具体实施过程。理论是一切的基础，课程理论基础的讲授是最重要的，教师需要有较强的专业知识和逻辑能力。理论基础对应的例题讲解是让学生对课堂上知识点的再次巩固和理解，教师需要认真讲解代码让学生清晰其中的编程逻辑。最重要的交互式学习方法学生不只是运行案例代码还要对参数和输入做交互式选择，运行过程会因这些操作而不同，最终产生不同的结果传达给学生，通过此学习方法学生会对于理论知识的理解产生质的变化。

交互式学习方法与学生直接运行教师提供的实现相应方法的源代码的学习方法相比，学生编程能力、逻辑思维和创造力更强。所以本课程的讲授采用交互式教学方法，在学生参与的过程中，让同学对所学的理论知识更加理解，提高学生的编程和逻辑思维能力，更有可能让学生对该领域感兴趣，积极参加基于深度学习的计算机视觉研究方向的大学生创新创业项目或者选做该研究方向的毕业设计，最终获益良多。

理论的学习需要实践。虽然在学习和科研的过程中，理论是一切的基础，但只有实践才可以进一步理解掌握理论知识，在思考中和反复的实践中学生才可以更好的掌握知识，挖掘自己的兴趣。

参考文献:

- [1] 高常鑫, 桑农, 王岳环. 大学本科开设计算机视觉课程教学的探讨[J]. 湖北第二师范学院学报, 2015, 32 (08): 96-99.
- [2] 朱虎, 卢官明, 谢世朋. “数字电视原理”理论实践一体化教学研究[J]. 电子世界, 2014 (13): 54.
- [3] 卢旭, 刘钊. 基于深度学习的图像语义分割技术综述[J]. 软件导刊, 2021, 20 (01): 242-244.
- [4] Minaee S, Boykov YY, Porikli F, et al. Image segmentation using deep learning: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
- [5] Harikrishnan J, Sudarsan A, Sadashiv A, et al. Vision-face recognition attendance monitoring system for surveillance using deep learning technology and computer vision[C]//2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN). IEEE, 2019: 1-5.
- [6] Dickmanns E D, Zapp A. Autonomous high speed road vehicle guidance by computer vision[J]. IFAC Proceedings Volumes, 1987, 20 (5): 221-226.
- [7] Wang G, Zuluaga M A, Li W, et al. DeepIGeoS: a deep interactive geodesic framework for medical image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 41 (7): 1559-1572.
- [8] Qu Z, Mei J, Liu L, et al. Crack detection of concrete pavement with cross-entropy loss function and improved VGG16 network model[J]. IEEE Access, 2020, 8: 54564-54573.
- [9] Klein S, Pluim J P W, Staring M, et al. Adaptive stochastic gradient descent optimisation for image registration[J]. International journal of computer vision, 2009, 81 (3): 227.
- [10] Tovar E, Vasques F, Burns A. Supporting real-time distributed computer-controlled systems with multi-hop P-NET networks[J]. Control Engineering Practice, 1999, 7 (8): 1015-1025.
- [11] Nobis F, Geisslinger M, Weber M, et al. A deep learning-based radar and camera sensor fusion architecture for object detection[C]//2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF). IEEE, 2019: 1-7.
- [12] Viprey V F, Corrias M V, Kagedal B, et al. Standardisation of operating procedures for the detection of minimal disease by QRT-PCR in children with neuroblastoma: quality assurance on behalf of SIOPEN-R-NET[J]. European Journal of Cancer, 2007, 43 (2): 341-350.