

随机过程在图像分割中的应用

摘要

图像分割是图像处理领域中的一大关键问题。本文从实际生活中的场景出发，通过自动驾驶等近几年图像分割取得的最新进展，引入图像分割的概念，并简要、科普地讲解它的定义和一些实际应用。之后分析二者与随机过程的关系，即其背后的原理是随机过程中一类重要的概念——马尔科夫随机场和马尔科夫链，列出数学定义，并通俗地进行解释。然后再通过图像分割算法的发展历史，找寻图像分割理论及算法发展过程中的关键节点及解决关键性问题的关键人物，探讨推动发展的内在动力。最后结合自身学习生活，对比差距，总结出这一探究过程给自己带来的学业等方面的启示。

关键词 图像分割 随机过程 马尔科夫链 发展历史

引言

图像分割是计算机视觉研究中的一个经典难题,已经成为图像理解领域关注的一个热点,图像分割是图像分析的第一步,是计算机视觉的基础,是图像理解的重要组成部分,同时也是图像处理中最困难的问题之一。所谓图像分割是指根据灰度、彩色、空间纹理、几何形状等特征把图像划分成若干个互不相交的区域,使得这些特征在同一区域内表现出一致性或相似性,而在不同区域间表现出明显的不同。简单的说就是在一副图像中,把目标从背景中分离出来。对于灰度图像来说,区域内部的像素一般具有灰度相似性,而在区域的边界上一般具有灰度不连续性。

由于问题本身的重要性和困难性,从 20 世纪 70 年代起,图像分割问题就吸引了很多人员研究,并为之付出了巨大的努力。虽然到目前为止,还不存在一个通用的完美的图像分割的方法,但是对于图像分割的一般性规律则基本上已经达成的共识,已经产生了相当多的研究成果和方法。

本文首先对图像分割算法及其在日常生活中的应用进行科普性质的介绍,并就其与随机过程中的两个重要概念——马尔科夫随机场和马尔科夫链之间的关系进行探讨。本文重点在于通过总结图像分割理论及算法发展历史上的几个关键节点及关键人物,来揭示推动发展的内在动力,并总结出其对于当代大学生日常生活的重要启示,同时也对目前正在使用的各种图像分割方法进行了一定的归纳总结。

1 图像分割技术

1.1 图像分割技术简介

在处理随机信号的过程中,我们往往首先将信号采样得到有效数据,再进一步通过信号的规律,如随机过程的统计规律和可以化为不同三角级数叠加的频率特性等,进行预处理,而后才能将有限的采样数据在综合域分析,利用多种手段得到信号中的有效信息。将信号处理技术类比到图像处理技术,图像就是对一个客观事物的二维反映,其每一个像素点对应信号处理的采样点。每一个像素点包含的信息确实是有限的,但将各像素点在整个图像的分布中综合考虑,可以发现各个像素点的分布规律具有潜在的关联。对应于随机信号的频域分析等手段,图

像处理技术也应具备相应的预处理手段，而目前最直观的，也是最基础的手段，就是将获得的像素点集，依据各像素点的分布规律，定性的分为聚类或划分边界，这就引出了图像分割技术。

从概念的角度来总结，图像分割就是指将图像分成若干具有相似性质的区域的过程，从数学角度来看，则是将图像划分成互不相交的区域。图像分割将图像中有意义的特征部分提取出来，其有意义的特征有图像中的边缘、区域等，这是进一步进行图像识别、分析和理解的基础。虽然目前已研究出不少边缘提取、区域分割的方法，但还没有一种普遍适用于各种图像的有效方式。因此，对图像分割的研究还在不断深入之中，是目前图像处理中研究的热点之一。

1.2 图像分割技术的应用

从图像分割技术的应用来理解。在我们观察物体时，总是先看到物体的主体或关键部分，摄影中常常对背景进行虚化处理，都是提取图像中的关键特征、对图像进行识别的方法。图像分割在目前飞速发展的数字图像处理技术中，对图像的有效识别具有重要意义，因为后续利用计算机对图像进行一系列的处理都建立在正确的图像识别的基础上。本文中所要介绍的图像分割即是图像识别和计算机视觉之中至关重要的预处理，没有正确的分割就不可能有正确的识别。



图1 图像分割示例

近些年来，随着深度学习技术的逐步深入，图像分割技术有了突飞猛进的发展，该技术相关的场景物体分割、人体前背景分割、人脸人体 Parsing、三维重建等技术已经在无人驾驶、增强现实、安防监控等行业都得到广泛的应用。在通信领域中，图像分割技术对可视电话等活动图像的传输也很重要，设备需要把图像中活动部分与静止的背景分开，还需把活动部分中位移量不同的区域分开。利用

图像分割技术，对不同运动量的区域用不同的编码传输，可有效降低传输所需的码率。

2 图像分割与随机过程之间联系

2.1 马尔科夫随机场和马尔科夫链

如上文所述，我们需要寻找一种方法来反映图像的基本规律，在此引入了一个重要的数学模型，即概率无向图模型，也称马尔可夫随机场模型。这同样是图像分割技术的理论基础，为了能够深入理解图像分割技术以及不同算法的关联性，首先介绍随机过程中的一类重要概念——马尔科夫链和马尔可夫随机场。

人们在实际中常遇到具有下述特性的随机过程：在已知它目前的状态(现在的条件下，它未来的演变(将来)不依赖于它以往的演变(过去)。这种已知“现在”的条件下，“将来”与“过去”独立的特性称为马尔可夫性，具有这种性质的随机过程叫做马尔可夫过程。液体中微粒所作的布朗运动，传染病受感染的人数，原子核中自由电子在电子层中的跳跃，人口增长过程等等都可视为马尔可夫过程。还有些过程(例如某些遗传过程)在一定条件下可以用马尔可夫过程来近似。

马尔可夫链由俄国数学家 A.A.马尔可夫于 1907 年提出，是一组具有马尔可夫性质的离散随机变量的集合。具体地，对概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 内以一维可数集为指数集的随机变量集合 $\mathbf{X} = \{X_n : n \geq 0\}$ ，若随机变量的取值都在可数集内 $X = s_i, s_i \in \mathcal{S}$ ，且随机变量的条件概率满足

$$p(X_{t+1} | X_t, \dots, X_1) = p(X_{t+1} | X_t)$$
，则 \mathbf{X} 被称为马尔可夫链，可数集 $\mathcal{S} \subseteq \mathbb{Z}$ 被称为状态空间，马尔可夫链在状态空间内的取值则称为状态。

马尔可夫链可被应用于蒙特卡罗方法中，形成马尔可夫链蒙特卡罗，也被用于动力系统、化学反应、排队论、市场行为和信息检索的数学建模。此外作为结构最简单的马尔科夫模型，一些机器学习算法，例如隐马尔科夫模型、马尔科夫随机场和马尔科夫决策模型均以马尔可夫链为理论基础。

当给每一个位置中按照某种分布随机赋予相空间的一个值之后，其全体就叫做随机场，而马尔科夫随机场则是一类具有马尔科夫性的随机场。马尔科夫随机场又被称为概率无向图模型，或者马尔科夫网络，它是一类无向图模型，也即两

个点之间并没有明确的前后以及方向关系。虽然两个点之前存在相互作用，但是这个作用仅仅在附近的点与点之间，与更远处的点或者最前面的点没有任何关系。马尔科夫随机场可以很好地描述空间连续性，被广泛地应用于图像处理、纹理迁移等机器学习图像处理方法中。

2.2 图像分割问题的处理思路

数字图像的本质是用工业相机、摄像机、扫描仪等设备经过拍摄得到的一个大的二维矩阵，该矩阵的元素称为像素，其值称为灰度值。因此对于数字图像的一系列处理在数学上就可抽象为对矩阵中的元素进行运算与变换。图像处理技术一般包括图像压缩，图像增强和复原，图像匹配、描述和识别三大部分，而图像分割则正是属于图像匹配、描述和识别的范畴。

从数学角度来看，图像是一个典型的马尔科夫随机场。在图像中每个点都有可能和周围的点有所牵连，但是和远处的点或者初始点没有关系，离这个点越近对这个点的影响也就越大。基于这一点，我们即可建立起图像分割与随机过程之间的联系。图像处理问题即为一种基于随机过程中的马尔科夫随机场和马尔科夫链的数学问题，其中图像分割问题本质上就是一种基于局部区域的分割方法。根据统计决策和估计理论中的最优化准则即可确定分割问题的目标函数，求解满足这些约束条件下的最大可能分布，就可以将分割问题转化为优化问题。以基于神经网络的分割方法为例，它的基本思想就是通过训练多层感知机来得到线性决策函数，然后用决策函数对像素进行分类来达到分割的目的。

迄今为止进行分割仅有的依据是图像中像素的亮度及颜色，这就导致了由计算机自动处理分割时将会遇到各种困难。例如，光照不均匀、噪声、图像中存在不清晰的部分以及阴影等因素常常会引发分割错误。因此图像分割及其所应用的数学原理需要得到进一步的研究。人们希望引入一些人为的知识导向和人工智能的方法，以纠正某些分割中的错误，得到更加理想的分割效果。

3 科学家和科学家精神

3.1 中国科学家

3.1.1 何恺明

图像分割是计算机视觉领域图像处理工作中极为重要的一环，而图像的去雾

算法就是图像分割衍生出来的一个重要应用。在去雾算法的研究中，中国的青年科学家何恺明曾在 2009 年提出了一个非常有效的解决方案，他基于无雾图像中每一个局部地区都可能会有阴影的这一特征，总结出无雾图像中每一个局部地区都应当至少具有一个颜色通道较低的值这一规律，他将这一规律命名为 Dark Channel Prior，基于这一规律，最终完成论文《Single Image Haze Removal Using Dark Channel Prior》并获得了 CVPR 2009 最佳论文奖。

在论文中，他所设计的去雾算法效果如下



图2 何恺明去雾算法效果

其中，左上为原始图片，右上为暗通道提取结果，左下为考虑雾气因素的暗通道导向图，右下则是最终的去雾效果，可以看出去雾的最终效果还是极为明显的。

何恺明的研究也在这一领域内得到了高度认可。而在他分享自己心路历程的文章中，我们还注意到这一非常杰出的研究成果最初竟来源于他打游戏时受到的启发。这一故事也告诉我们，生活中的很多不起眼的细节往往是很多重大问题的关键所在，在生活中，我们也应当时刻带着思考来观察身边的一切。而另外，我

们可以说何恺明的这一出色研究成果有一定的运气成分，但俗话说的好，“机会总是留给有准备的人”，何恺明本身的知识积累也非常深厚，作为清华的保送生，他在 2003 年的广东省高考中还获得了满分的成绩，而在大学期间，何恺明对于各种各样的知识更是广泛涉猎，因此才能够由游戏联想到科研，同时基于他沉稳且有耐心的性格与善于自我反思、自我完善的特质，使得他对于在游戏中的发现没有浅尝辄止，也使得他所总结的规律没有成为浮于表面的简单发现，而是在诸多实验结果与理论支持下的完备方案。这也告诉我们，观察生活、思考生活固然可以为我们提供更多的机会，而只有扎实的理论基础与知识体系才能让我们抓住这些机会，因此作为大学生的我们不仅要善于观察生活，更要脚踏实地地摄入知识，以自己的知识体系来审视生活，才能真正做到在生活中收获重要启示。

3.1.2 高文

中国工程院院士高文的经历同样给予人启示，通过高文院士的经历，我们学到了在科研过程中如何开拓思路、打开思维，从而找寻新的切入点。作为在计算机视觉领域同样具有突出贡献与诸多杰出成果的科研工作者，高文院士通过他正在参与建设的“数字视网膜”计划为我们阐释了在学习乃至研究的过程之中，交叉思维的重要性。

“数字视网膜”计划是高文院士基于对生物学知识的了解而产生出来的想法。现有的监控视频在应用过程中普遍与机器分析相结合，而如果前端不支持，就要完全靠云计算，需要大量的算力。我们的眼睛效率非常高，于是乎“有没有办法让计算机识别学一学生物系统的眼睛？”就成为了一个全新的问题。

眼睛通过视网膜获取光学信号，视网膜里有两种比较关键的细胞，感光细胞和锥状细胞。我们平时看东西都是通过这些细胞把信号收集来，传到大脑去。眼睛视网膜的和最后传到大脑里的细胞数存在差值，眼睛视网膜细胞数是后面脑区接收的数量的 126 倍。也就是说，视网膜上有 126 个细胞，最后汇总到连接到脑的一个神经元上。人的视觉系统，在传达信息时经过若干层，每传一层信息就进行缩减，一直传到脑神经的连接地方。这就给高文院士一个很好的启发：脊椎动物视网膜设计得非常巧妙，这种结构对视觉认知非常有帮助。

而现有的城市大脑反而比较像病态的人的视觉系统。从神经系统解释自闭症和癫痫症的根源，自闭症就是人一生下来，末端神经和脑端的神经几乎全连接，

神经系统所有的连接都相似且都连着。婴儿接受外界刺激、学习的过程，实际把全连接的网络进行了增强和剪裁，有些连接变得越来越粗，有些连接慢慢就萎缩掉。一个正常的孩子，该衰减的衰减，该增强的增强，最后是正常的。如果小孩在发育的时候没有经常进行剪裁，这个孩子长大以后就会犯自闭症。

现在我们的城市大脑系统，没有进行任何信息的提取，每个摄像头在系统里同等重要，这是一个“自闭症系统”。如果有的摄像头特别重要，比如有的摄像头能识别出车牌号，这些连接就太粗壮，系统很容易变成“癫痫症系统”。现有的城市大脑系统最容易犯的两种病，就是“自闭症”或者“癫痫症”。因此，高文院士认为，需要有一个系统，像人的视觉系统一样，把信息汇总以后，缩减以后再往上送。

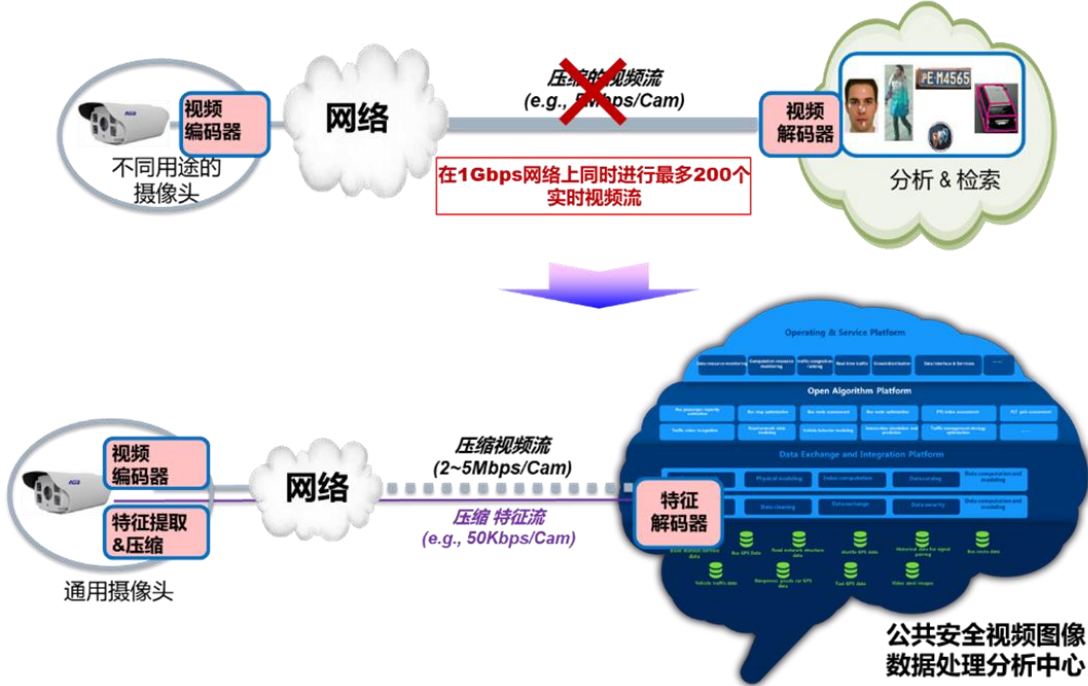


图3 现有多用途摄像头与未来通用摄像头应用对比

而如今，“数字视网膜”系统已经广泛应用于全国各地，上述诸多问题的解决也使得我国“天网工程”建成了全世界覆盖范围最广、实时分析效率与总体识别精确度最高的监控系统，服务于我国公安治理与广大人民生命与财产安全的保障。

高文院士的这一经历给予我们的启发是：在学习和探索研究方向的过程中，我们应当要扩宽自己的视野，不应闭门造车，要去尝试跨学科、跨领域进行知识的摄取。有的时候，不把所有时间花在一个科目或者一个领域上，转而投身涉猎其他领域知识的行为，看似在做无用功，却很可能为自己的研究思路打开了一扇

新的大门。而这样的交叉思维，恰恰也是现在的科研领域中所大力推崇的，放在大学生的身上，也是在告诉我们，不应读死书，而是需要做到将知识活学活用，融会贯通，才能够在诸多现有知识的基础之上，找到新的方向、开拓新的领域。

3.2 外国科学家

3.2.1 Olaf Ronneberger

菲兹保大学的 Olaf Ronneberger 教授于 2015 年提出了一种新的应用于生物医学领域的图像分割方法，即 U-Net 方法。这一方法相较于之前其他应用深度学习和神经网络理论的图像分割方法，多尺度信息更加丰富，能够满足生物医学学科的特殊需求，掀起了生物医学与计算机视觉相结合的潮流。

Olaf Ronneberger 教授提出 U-Net 这一方法并不是异想天开，而是敏锐地观察社会现状，基于了工程实际应用中的需求。今日糖尿病仍是困扰着人类的一大病症。据不完全统计，美国约有 3000 万名糖尿病患者，而全球范围内这一数字则超过了 4 亿。而糖尿病会引起诸多并发症，视网膜病变则是其中较为严重的一种，给糖尿病患者带来了身体上的痛苦和生活上的极大不便。但是，在美国每年只有大约一半的糖尿病患者会按照建议检查自己的眼睛，在世界其他地区这一比例则更低。其中一个重要原因就是眼科医生的短缺。这个问题在低收入和中等收入国家尤为严重。为解决这一问题，人们想到是否可以利用人工智能，通过图像分割的方法检测相对容易发现的糖尿病视网膜病变以及其他常见的眼病，提高大规模筛查的速度和准确性，以此改善眼科服务欠缺地区的现状。

这项工作的难点在于生物医学对于视觉任务的要求往往很高，目标输出不仅要包括目标类别的位置，还要使得图像中的每个像素都拥有类标签。在生物医学图像分割中，可获得的训练数据很少，训练集大小存在限制，同时常见对于相接触的同一类的多个对象进行分割的问题。对于这类问题，传统的使用卷积神经网络的图像分割方法很难实现较好的效果。

Olaf Ronneberger 教授通过分析图像分割筛查糖尿病视网膜病变这一实际应用的需求，提出了几种全新的方法来解决这一需求中的两大核心问题。解决训练集大小存在限制这一问题的方法是进行数据扩大。他及他的研究团队使用在粗糙的 3×3 点阵上的随机取代向量来生成平缓的变形。而解决对相接触的同一类的多个对象进行分割这一问题的方法则是利用加权损失，使得基于相邻细胞的分界的

背景标签在损耗函数中有很高的权值。这一方法即是今日在图像分割领域广为人知的 U-Net 方法，实现了在训练集很小的情况下也可以达到很好的分割效果。这一方法不仅被应用于检测早期病症的视觉特征等生物医学手段，更是被广泛地应用于卫星导航地图、电子显微镜等诸多不同学科领域，推动了图像分割理论及算法的发展。



图4 应用 u-net 的卫星地图分割

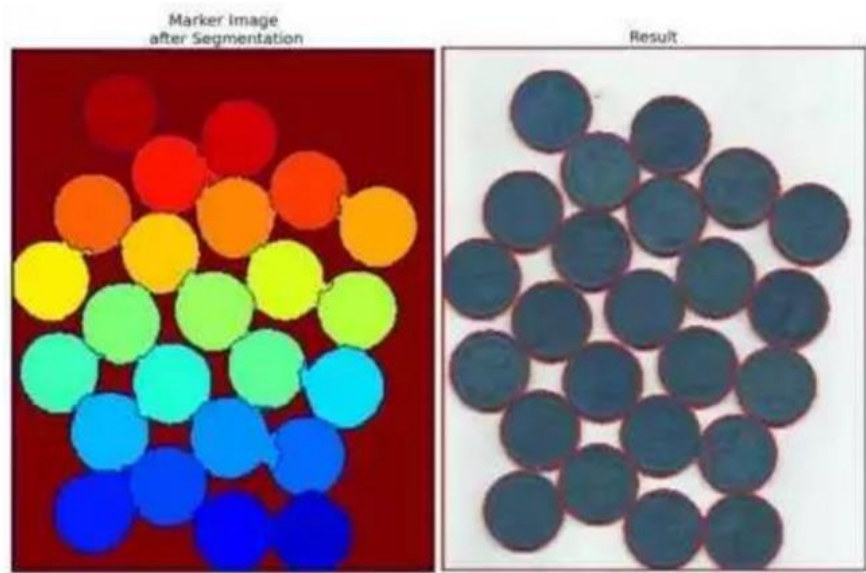


图5 应用 u-net 分割电子显微镜下的细胞

由 Olaf Ronneberger 教授提出 U-Net 方法这一样例，我们体会到在进行新的科学技术方法的探索时，不应闭门造车，而应关注工程实际，关心社会需求。Olaf Ronneberger 教授从筛查糖尿病视网膜病变这一具体的工程实际应用出发，通过

解决其中的关键问题得以获得了成功，推动了整个图像分割领域的发展。试想，倘若 Olaf Ronneberger 教授没有注意到这一实际应用，或许人们也不会去关心如何解决图像分割方法在小样本、多对象情境下的适用性这样的问题，这也就很难催生出多尺度信息更加丰富的新的图像分割算法。

遑论 U-Net 这一图像分割领域的新方法，历史上很多我们耳熟能详的伟大发现也都是由工程实际的需要催生的。19 世纪末，由于冶金以及照明设备制造等的需要，人们急需找到黑体辐射强度和辐射频率的关系。从经典物理学出发推导出的维恩定律在低频区域与实验数据不相符，而在高频区域，从经典物理学的能量均分定理推导出瑞利-金斯定律又与实验数据不相符，在辐射频率趋向无穷大时，能量也会变得无穷大，这种结果被称作为“紫外灾变”。普朗克使用插值法将两个公式化成了一条公式也即普朗克公式，并为了解释这个半经验公式的准确性提出了能量量子化假设，开启了量子力学的新时代。历史和当代的事例均启示我们要多关注社会现状与产业发展，从工业实际的需要中得到关于新的研究方向的灵感。同时我们也应思考怎样才能使得自己的研究成果能够更好地造福大众，提高人们生活的幸福感，推动社会生产领域的发展与变革，而这也是科技进步的意义之所在。

4 感想总结

4.1 遇到的问题及解决

在本学期研究之初，我们对马尔科夫随机场的原理和图像分割及纹理迁移的实现进行了探究，并尝试结合了重要的思想方法。但经过老师的点评和分析，我们发现这一阶段的研究没有牢牢把握“科普”这一点要求，忽视了对此过程中体现出的重要思想方法的分析。

因此，在研究的第一阶段结束后，我们确定了接下来的研究重点是科普方面，通过二者的发展历史关键节点来提炼出重要的思想方法及对我们的学习生活的启示。

我们在小组内进行多次讨论，经过不断完善和优化，形成了第二阶段的研究成果，汇报展示也获得了老师的肯定。在第二次汇报中，我们首先从实际生活中的场景出发，通过自动驾驶等近几年图像分割取得的最新进展，引入图像分割的概念，并简要、科普地讲解它的定义和一些实际应用，例如在摄影中的应用；之

后我们讲述了二者与随机过程的关系，即它背后的原理是随机过程中一类重要的概念——马尔科夫随机场和马尔科夫链，列出数学定义，并通俗地进行解释；然后再通过图像分割算法的发展历史，找寻它们发展过程中的关键节点，是谁解决了这一关键性问题，以及他为什么能解决这一问题；最后每位小组成员结合自己的学习生活，对比差距，总结出这一探究过程给自己带来的学业等方面的启示。

对于第二阶段的研究，张有光老师在点评中指出，演讲过程中演讲者声音稍小，影响了台下老师与同学的听讲效果，同时，在展示内容方面，在启示与展望部分，我们的准备内容与演讲词有些空泛，稍显套路性，没有能很好地体现出本组成员对于本组题目的独立思考。

然而，小组成员却在此时遭遇了研究的瓶颈，我们之前的工作执着于在理论知识上死磕，局限于公式和 Matlab 模拟，结果不仅没有做到对知识的深入理解，在如此长的探究时间中也没能聚焦优秀前人的经历，更没能从中获得心路历程上的指导。这导致我们小组的工作效率极度低下，在比较长的探究时间中不能对零散的理论进行探究、收集与融合。

幸运的是，我们迎来了第二次和张老师的交流。老师一语道破我们的核心问题——研究不深，并给予了我们在思考与展示上的建议，即要将故事和理论融入生活，这样的探究与分享才是有意义的。如图像分割背后的理论，对于没有研究图像分割的同学来说，图像分割的纯理论是枯燥乏味且难懂的，然而如果将这些理论与实际生活相结合，不仅能使得听讲者产生共鸣，也能够诱导他们产生兴趣，进而有机会也可以进一步对理论的学习展开协同讨论。

根据张老师的建议，我们在第三阶段的研究中，通过大牛们的博客等渠道获得其第一手的经验与体会，而对于一些历史较为悠远的杰出科学家，我们也会通过公众号等渠道着重对其成长过程与发现问题、解决问题的过程进行关注与总结。同时，在总体上，我们规避了过度沉迷图像分割这一领域，而是将重点放在了随机过程上，毕竟随机过程知识的学习才是本课的主题。我们的学习探究应当为课内基础知识服务，从而让我们学习随机过程的效率最优化、收获最大化，这是我们应当一以贯之的核心。

在三个阶段的研究中，我们小组遇到了几次小困难和一次重大瓶颈，但都在张老师的指导和小组成员的共同努力下一一克服，完成了全部研究过程，加深了

对马尔科夫随机过程的理解，并对随机过程这门学问产生了更加强烈的探究愿望。

4.2 收获与启发

通过研究，我们知道图像分割算法的实现是一个循序渐进的过程，我们应当要看到前人方法的优点和不足之处，并积极思考如何才能对其进行改进。

同时，我们应当掌握前沿的数学知识并将其应用于自己所在的领域，并应当时刻提醒自己数学的重要性。著名数学家牛顿一生取得了无数成果，但大都建立在其微积分等数学突破的基础之上，数学的重要可见一斑。

在 U-Net 的例子中，我们体会到关注工程实际的重要性。U-Net 就是从医疗图像分割这一具体的工程实际应用出发，通过解决关键问题得以成功的。我们要多关注社会现状与产业发展，从工业实际的需要中得到新的研究方向的灵感。著名物理学家普朗克正是从工业生产实际中出发，发现并解决了黑体辐射问题。

随着深度学习在计算机视觉领域的巨大成功,以卷积神经网络为基础的图像语义分割方法取得了突破性的进展，将图像分割带到了一个新的高度。如何继续提高分割算法的精度，以及降低分割算法的复杂度是值得继续研究的问题。在日常的学习生活中我们应多将所学知识联系实际，不能仅仅是应付考试，而应积极思考它们在实际应用中的重要价值。