

医学影像人工智能处理技术

医学影像与人工智能的结合已经在医疗领域引起了广泛的关注和研究。目前人工智能技术已被用来改善医学影像的获取、分析、应用,并提高患者的诊断、检测、治疗等环节的效率和准确性。本章将主要介绍常用医学影像类型、医学影像人工智能的典型方法、任务和应用。



5.1 医学影像类型

本章首先介绍常用医学影像数据,包括 X 射线影像、超声影像、CT 影像、MRI 影像等。此外,医学影像还有很多其他类型,如病理学影像、内窥镜影像、骨扫描图像等,在这里不一一介绍,有兴趣的读者可以自行查阅文献学习。

5.1.1 X 射线影像

X 射线摄影依赖人体组织之间的密度和厚度的差别成像。X 射线透过人体不同组织结构时,被吸收的程度不同,所以到达胶片上的 X 射线量具有差异,形成明暗或黑白对比的影像。普通 X 射线成像因速度快、成本低廉,在临床中应用范围很广。常用的 X 射线图片包括胸腹部平片、骨关节摄影、乳腺 X 射线摄影检查等,也称之为 X 光影像或 X 线片。图 5-1 为脊柱 X 射线影像和手骨 X 射线影像的例子,图 5-2 为迈瑞医疗的 DigiEye 330/350 系列数字 X 射线成像系统医疗器械。

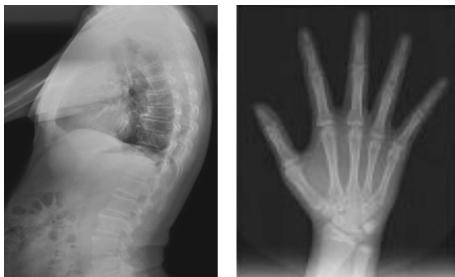


图 5-1 脊柱 X 射线影像和手骨 X 射线影像的例子

在医学影像领域,X 射线设备已经成为最常用的医学影像设备之一,具有如下优点和局限性。

1. X 射线检查的优点

(1) 快速成像: X 射线设备成像速度快,通常只需要花费较短的时间,有助于



图 5-2 迈瑞医疗的 DigiEye 330/350 系列数字 X 射线成像系统医疗器械

医生快速做出初步诊断,为患者提供及时的治疗。

(2) 适用范围广: X 射线可以用于不同部位和器官的成像,包括胸部、骨骼、牙齿、胃肠道等,这使得 X 射线可以应用在多个医学领域。

(3) 成本较低: 与 MRI 或 PET 扫描等相比,X 射线设备相对便宜,维护成本也相对低。目前,即使是在偏远地区,在医院、诊所和医疗保健机构中,也可以经常看到 X 射线设备。

2. X 射线检查的局限性

(1) 辐射暴露: X 射线具有电离辐射,暴露于辐射源可能会对人体组织造成损伤,长期或过量的辐射暴露可能增加癌症的风险。

(2) 不适用于特定患者群体: 由于辐射风险,X 射线可能不适用于孕妇等人群。

(3) 有限的软组织分辨率: X 射线在显示骨骼结构方面效果较好,但在显示软组织(如肌肉、脂肪和器官)方面效果较差。对于某些疾病,X 射线影像可能不足以提供足够的信息。

(4) 造成图像伪影: X 射线影像可能受到骨骼叠加效应的影响,导致图像上出现伪影,这可能需要进一步的评估。

5.1.2 超声影像

超声影像,也称为超声检查或超声波检查,是一种利用超声波传播和回波原理获取人体内部结构的医学影像技术。现代超声仪器的探头通过线性、扇形或其他扫描方式,向人体发射超声波。当超声波遇到具有不同声阻抗的两种组织交界面时,会产生回波。这些回波被探头接收后,经过信号放大和信息处理,最后显示在屏幕上,形成断层图像,称为超声影像。图 5-3 为超声影像的例子,图 5-4 为飞利浦彩色超声诊断系统 Affiniti70 设备。

超声影像可用于医学诊疗的方方面面,例如乳腺病变的筛查、肝脏肿瘤的检查、引导活检操作等。通常会使用灰度超声、彩色多普勒超声等。灰度超声利用超声波的反射和传播来获取人体内部组织的图像。与传统的二维超声图像相比,灰度超声图像能够提供更加细致的人体组织结构信息。彩色多普勒超声通过多普勒效应、多普勒频移测量和彩色编码等技术,提供人体内部组织和血管的实时和高清晰度图像,可用于探测血流速度和方向,帮助医生进行准确的诊断。由于具有多普勒功能的超声扫描仪能够检测血液流动情况,因此在人工智能医学影像的研究中,也同时基于灰度超声和多普勒超声数据来完成分析任务。其优点和局限性如下。



图 5-3 超声影像的例子



图 5-4 飞利浦彩色超声诊断系统 Affiniti70 设备

1. 超声检查的优点

(1) 无辐射性：超声检查不会产生电离辐射，对人体的伤害较小。因此，超声检查在孕妇和其他需要避免 X 射线等辐射的人群中更为适用。

(2) 非侵入性：超声检查是一种非侵入性的检查方法，不会对患者造成介入的影响。

(3) 实时性：在超声检查过程中可以实时观察到人体内部器官的动态变化。医生可以根据这些实时图像进行准确的诊断和评估，这在心脏超声、胎儿监测、活检引导等应用中尤为重要。

(4) 多角度成像：超声检查可以通过多个角度进行成像，有助于更全面地观察身体状况。这一特点使得超声检查可以有效地应用在多角度器官观察、复杂病变诊断的任务中。

(5) 便捷快速：相比于其他成像设备，超声检查可以在短时间内完成，是一种高效的成像工具。

(6) 成本低：相对于其他成像技术(如 CT 影像或 MRI 影像)，超声检查成本相对较低。

2. 超声检查的局限性

(1) 图像质量受体质影响：超声影像的质量受到患者的体质(如体重、身体脂肪含量)和气体干扰(如胃肠气体)等因素影响。

(2) 组织穿透力有限：超声波穿透骨骼或空气的能力相对较弱，有时候难以检测到某些深层的病变组织，此时，需要结合其他检查方法进行综合判断。

(3) 对操作人员的依赖性强：超声检查的操作和解读都依赖医生的经验和技能。对于一些复杂的病变,需要具备丰富经验的医生才能做出准确的诊断。

(4) 受视窗限制：对于某些器官和组织,超声检查可能受到视窗的限制,难以获得全面的影像。因此,在实际的检查中,针对同一病灶,超声技师通常采集多张影像数据。

5.1.3 CT 影像

计算机断层扫描,通常称为 CT 扫描,是一种常见的医学影像技术。它使用 X 射线环绕患者身体扫描,通过计算机技术将人体内部的结构以横截面的方式呈现出来。这些图像被称为断层图像,相比于传统的 X 射线图像,可以提供更详细的病变信息。这些断层图像堆叠在一起,形成三维体积数据,可以直观地展示身体基本解剖结构、病变等信息。CT 扫描能够以非侵入式的方式获得高分辨率、多方位的横截面图像,常用于诊断、筛查和手术指导。图 5-5 为胸部 CT 扫描图像案例,图 5-6 为飞利浦的 CT 扫描设备 Spectral CT 7500。

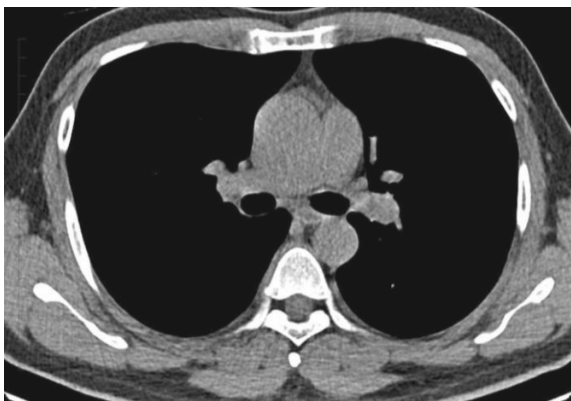


图 5-5 胸部 CT 扫描图像案例

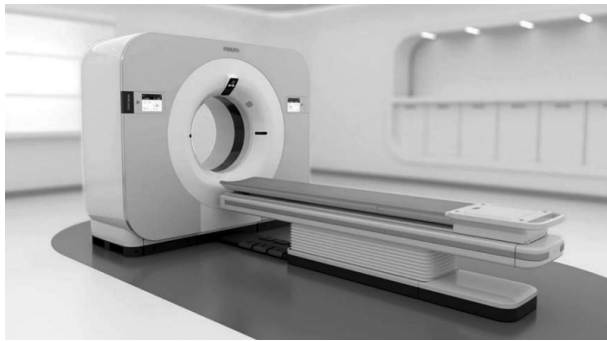


图 5-6 飞利浦的 CT 扫描设备 Spectral CT 7500

CT 检查的优点和局限性如下。

1. CT 检查的优点

(1) 图像分辨率高：CT 扫描可以提供高分辨率、清晰的内部结构图像,能够清晰地显示骨骼、器官和组织结构。

(2) 组织密度分辨能力强：可将密度差别小的软组织及其病变分辨出来,可清晰显示解剖结构复杂部位的病变情况,提供疾病诊断的证据。

(3) 多平面重建: 通过重新构建 CT 图像, 能提供不同方向(如横断面、冠状面、矢状面)的视图, 有助于显示微小结构和病变。

(4) 适用范围广: CT 检查可用于身体的各个部位, 包括头部、胸部、腹部和盆腔等。

2. CT 检查的局限性

(1) 辐射暴露: CT 检查使用 X 射线, 产生的电离辐射对人体有一定的损害。需要特别注意儿童和孕妇等特定人群。

(2) 成本较高: 相对于超声检查设备, CT 设备的购置和维护成本相对较高, 这可能会反映在患者的医疗费用上。

(3) 对某些软组织的分辨率低: 对于某些软组织的成像, 如肌肉和神经, CT 检查的分辨率低。

(4) 无法评估功能性信息: CT 检查提供的是静态的结构图像, 无法提供关于器官功能的信息, 如心脏的收缩功能或肺部的通气功能。

5.1.4 MRI 影像

磁共振成像(MRI), 也称为核磁共振成像, 是一种利用强磁场和无线电波来生成人体内部高质量图像的医学影像学技术。目前 MRI 已应用于全身各系统的成像诊断, 包括颅脑、心脏、关节、骨骼、软组织及盆腔等。图 5-7 为 MRI 影像的例子, 图 5-8 为飞利浦的 3T 核磁共振系统 MR 7700。

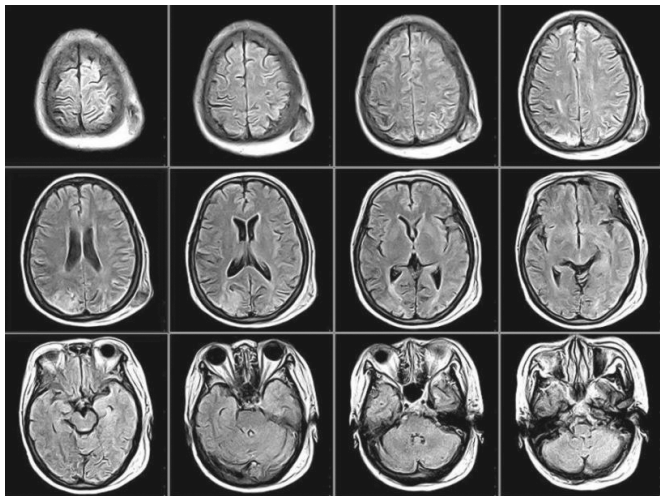


图 5-7 MRI 影像的例子

MRI 是一种综合的成像技术, 可用于评估退行性疾病、炎症性疾病以及癌症等各种病理组织, 提供多参数评估。近年来, 随着设备与技术的不断发展, MRI 序列不仅能提供解剖信息, 还能提供功能和代谢信息。MRI 检查的优点和局限性如下。

1. MRI 检查的优点

(1) 无辐射: MRI 不使用离子辐射, 对患者不产生辐射暴露的风险。这使得 MRI 成为适用于孕妇和儿童等敏感人群的影像检查方法。

(2) 对比度高: MRI 能提供较高的组织对比度, 能够清晰地区分不同类型的组织, 包括



图 5-8 飞利浦的 3T 核磁共振系统 MR 7700

软组织、肌肉、骨骼和血管等。

(3) 多参数成像: MRI 可以生成多种参数的图像,如 T1 加权图像、T2 加权图像、扩散加权图像和灌注加权图像等。这些图像提供了丰富的信息,有助于医生做出准确的诊断,制订适当的治疗计划。

(4) 多平面成像: MRI 允许在多个平面(矢状、冠状、轴位)上获取图像,提供了更全面的解剖信息。医生可以选择他们最需要的平面来查看图像。

(5) 功能性成像: 功能 MRI(fMRI)允许观察大脑活动,对神经科学研究和临床神经学研究具有帮助。

2. MRI 检查的局限性

(1) 成本较高: MRI 设备的购置和维护成本相对较高,这使得 MRI 检查的费用通常比其他成像技术(如 X 射线或超声检查)更高。

(2) 耗时: MRI 检查通常需要较长的时间,需要患者保持相对静止,这通常会导致患者的排队等待时间增加。

(3) 不适用于某些患者: 由于 MRI 使用强磁场,因此对于某些患者可能不适用,如植入心脏起搏器、内耳植入物、植入大型金属物体的患者,以及一些焦虑症或幽闭恐惧症患者。

(4) 对运动敏感: MRI 对患者的运动比较敏感,对于年幼的儿童、病情严重的患者或无法保持静止的患者,检查难度较大。

(5) 噪声大: MRI 扫描期间会产生噪声,这可能会令一些患者感到不适,通常需要提供耳塞或耳罩来减轻噪声。

(6) 对某些组织的成像限制: MRI 在某些情况下对骨骼和气体的成像效果不如 X 射线或 CT 扫描。骨骼成像通常采用 X 射线或者 CT 成像技术。

MRI 检查适合于需要高对比度和多参数成像的情况。医生通常会根据患者的具体情况和临床需求来决定是否进行 MRI 检查,以及如何有效地使用这些图像来做出诊断和治疗决策。



5.2 医学影像的分类方法

5.2.1 引言

医学影像的分类任务是指将医学影像数据分为不同类别的任务,标签是分类别的。在

实际研究中,根据数据是否有标签,以及标注的数据量,选择有监督学习、半监督学习、无监督学习等技术路径来完成分类任务(详见第3章)。以有监督学习为例,这意味着将从医学影像设备中获取的图像或者数据归类到特定的病理状态、器官结构、疾病类型等类别中。在有监督学习中,医学影像分类模型需要从标记了类别信息的训练数据中学习,然后对新的未标记的医学影像数据进行分类,例如将X射线影像分类为正常或异常,将乳腺超声图像分类为良性病变和恶性病变等。

图像分类模型需要处理特征抽取、分类器训练等关键步骤。传统的图像分类主要依据手动特征提取和机器学习方法,其中常用的特征提取方法包括尺度不变特征转换(Scale Invariant Feature Transform, SIFT)算法、方向梯度直方图算法(Histogram of Oriented Gradient, HOG)等,常用的机器学习方法包括贝叶斯、支持向量机(Support Vector Machine, SVM)、决策树、 k 最近邻、随机森林等(见第2章)。传统的医学影像分类方法通常需要手动设计特征,在处理复杂的海量医学影像数据时会出现性能瓶颈。随着深度学习技术的发展,越来越多的医学影像分类任务采用了基于深度学习的方法,这些方法可以自动学习和提取特征,目前已在许多任务中取得了极大的进展。

基于深度学习的医学影像分类方法是利用深度学习网络进行自动特征学习和医学影像分类的一种方法,目前已经应用在疾病的诊断、筛查、预后等大量的医学影像任务中。本节主要介绍具有代表性的基于深度学习的分类方法,并在介绍过程中展开说明网络模型中所涉及的关键技术。希望读者可以在学习后了解经典网络模型,学习该模型中的关键技术,并尝试使用深度学习技术完成医学图像分类。

5.2.2 AlexNet 方法

AlexNet是具有重大历史意义的网络结构。2012年,Hinton和他的学生Alex Krizhevsky设计了AlexNet网络,在当年的ImageNet分类挑战赛中获得了冠军,Top-1的分类准确率达到57.1%,Top-5的分类准确率达到80.2%,这远远超过传统的机器学习分类算法的性能。AlexNet的成功标志着人工神经网络重返历史舞台。这一成功的背后有若干关键因素,这些因素也构成了当前计算机视觉领域的核心技术。AlexNet架构示意图如图5-9所示。

1. 激活函数

Sigmoid是人工神经网络最常用的非线性激活函数之一,它能够把输入的连续实值“压缩”到0~1之间。但是它也有一些缺点。第一,具有梯度消失问题。也就是说,Sigmoid函数的值域为(0,1),当网络的层数较多时,经过多次链式求导后,梯度会不断地减小,甚至趋于0,这就导致了梯度消失问题。第二,Sigmoid的输出不以0为中心。当输入呈现较大值或者较小值时,Sigmoid的输出接近1或者0,这使得激活函数的输出不以0为中心,不利于梯度下降优化算法的训练。

为了解决Sigmoid激活函数的问题,AlexNet网络选择采用ReLU激活函数。使用ReLU函数更有助于模型收敛。ReLU函数图像见图5-10。可以看到,当输入信号小于0时,ReLU的输出为0;当输入信号大于0时,输出等于输入。

2. Dropout

集成学习(Ensemble Learning)可以将多个预先训练好的不同模型结合起来进行预测,

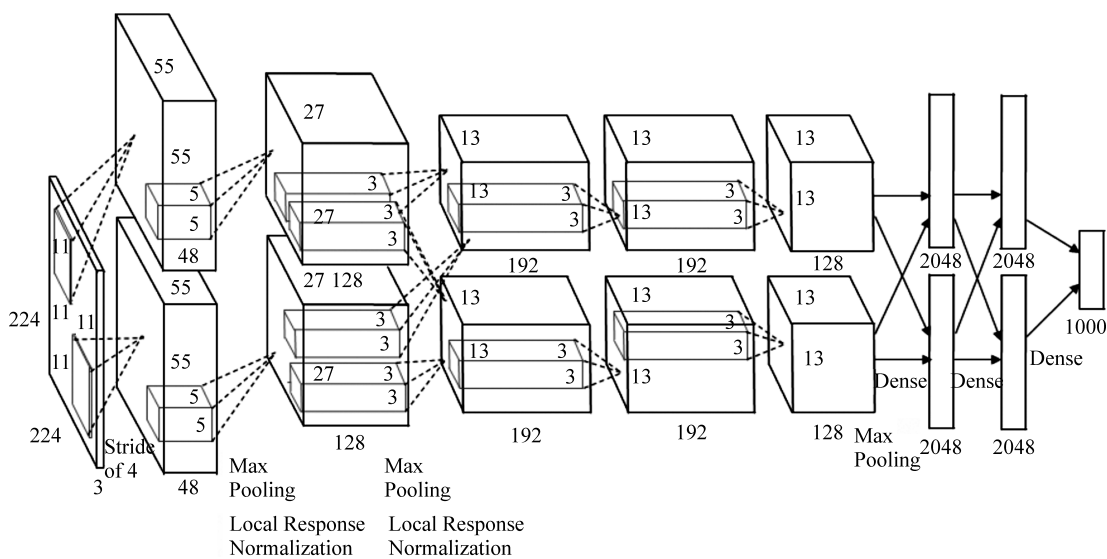


图 5-9 AlexNet 架构示意图

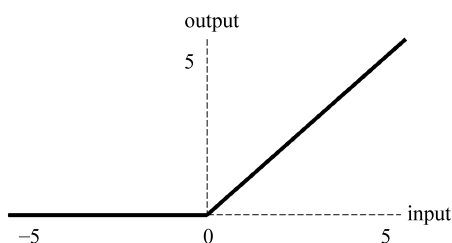


图 5-10 ReLU 函数图像

能够有效降低测试误差。但是由于训练耗时长，因此应用受到一定限制。为解决这一问题，AlexNet 提出了 Dropout 方法，其核心思想在于以一定的概率随机将每个隐藏层神经元的输出置零。这些被排除的神经元不参与信息的前向传播和梯度的反向传播。因为神经元不依赖特定的其他神经元而存在，所以 Dropout 技术降低了神经元之间复杂的相互依赖关系，这迫使神经网络学习更具有鲁棒性的特征。Dropout 是 AlexNet 网络的一项创新技术，目前已应用在各种神经网络的训练中。

3. 多 GPU 实现

鉴于 2012 年的硬件条件约束，单个 GTX 580 GPU 只有 3GB 内存，这限制了其上训练网络的最大规模。因此 AlexNet 的作者将网络部署在两个 GPU 上。

AlexNet 所采用的并行方案是在每个 GPU 中放置一半神经元，且约束 GPU 间的通信只在某些神经网络层中进行。例如，第 3 层的神经元的核需要从第 2 层中所有神经元的核来映射输入，而第 4 层的神经元的核只需要从第 3 层中位于同一 GPU 的那些核映射输入。目前，跨 GPU 的并行化技术功能强大，不需要通过主机内存就能够直接从另一个 GPU 的内存中读出和写入。

4. 局部响应归一化

在 AlexNet 架构中也引入了局部响应归一化 (Local Response Normalization, LRN)。与常见的 Tanh 和 Sigmoid 不同，这一架构采用了 ReLU 激活函数。引入 LRN 的原因在于鼓励侧向抑制，这是神经生物学中的一个概念，指的是神经元抑制其邻居神经元的活动。在深度神经网络中，侧向抑制的目的是进行局部对比度增强，以确保在下一层的激活中突出显示局部区域内的最大像素值。

LRN 本身是一个不可训练的层,它用于对特征图中局部邻域内的像素值进行平方归一化。根据定义的邻域,有两种 LRN 类型,分别是通道间的 LRN(Inter-Channel LRN)和通道内的 LRN(Intra-Channel LRN)。通道间的 LRN 侧重于跨通道的归一化,而通道内的 LRN 则关注同一通道内的归一化。这有助于调整特征图中不同通道之间或同一通道内的活动水平,以增强模型的鲁棒性和其他性能。

5.2.3 VGG 方法

2014 年,牛津大学计算机视觉组研发了 VGG 网络。VGG 网络尝试将所有的卷积核大小设定为 3×3 ,滑动步长设定为 1。虽然卷积核变小会减小神经网络的感受野,但该方法通过串联多个卷积核以达到和较大的卷积核具有相同感受野区域的目的。这种简单易实现的网络模型对于后续 CNN 的相关模型设计具有很大的启发。

VGG 的主要创新在于用多个较小卷积核叠加感受野代替一个大的卷积核,例如,2 个 3×3 的卷积核叠加,可代替一个 5×5 的卷积核;3 个 3×3 的卷积核叠加,可代替一个 7×7 的卷积核,如图 5-11 所示。VGG 采用了较小的固定的 3×3 卷积核,其层数增加带来了更强的非线性,使模型的判别能力更强。

在网络架构方面,VGG 设计了更多的全连接层和卷积层,并且将一部分卷积层组合成块。

经典卷积神经网络大多呈现下面的序列表达:

- 带填充以保持分辨率的卷积层;
- 非线性激活函数,如 ReLU;
- 池化层,包括最大池化和均匀池化等。

一个 VGG 块由一系列卷积层组成,再加上用于空间下采样的最大池化层。在最初的 VGG 论文中,作者采用带有卷积核、填充(padding)为 1(保持高度和宽度)的卷积层和步幅(stride)为 2(经过每个块后特征图减半)的最大池化层。除了全连接层以外,VGG 网络将 AlexNet 中的层替换成了多个 VGG 块。通过具有不同次数的重复块得到不同的架构,如 VGG-16、VGG-19 等。具体架构如图 5-12 所示。

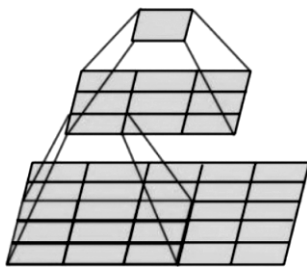


图 5-11 卷积核叠加

5.2.4 GoogLeNet 方法

谷歌在 2014 年的 ILSVRC 竞赛中使用了 GoogLeNet 网络,并获得了冠军。为了提升网络性能并减少网络参数量,GoogLeNet 引入了 Inception 结构。为了方便增加和修改,GoogLeNet 使用了模块化的架构,同时增加了 Softmax0 和 Softmax1 两个辅助分类节点来解决梯度消失问题。GoogLeNet 网络有多个版本,包括 Inception v1、Inception v2 等。

GoogLeNet 的创新可以归纳为以下四点:

- 引入 Inception 结构,融合不同尺度的特征信息;
- 使用 1×1 的卷积核进行降维以及映射处理;
- 添加两个辅助分类器帮助训练;
- 丢弃全连接层,使用平均池化层,大大减少了模型参数。

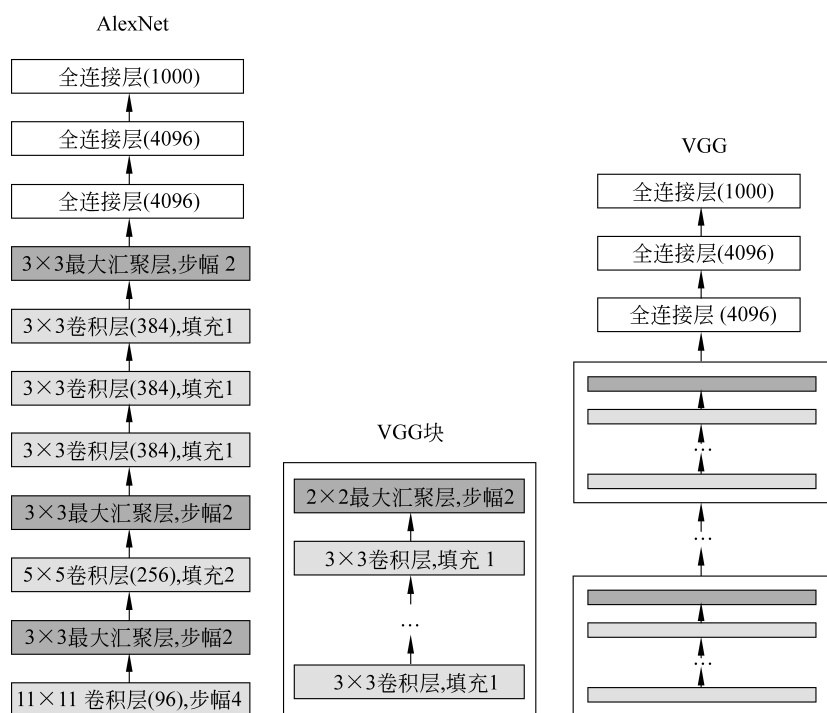


图 5-12 AlexNet 和 VGG 架构示意图

1. Inception 结构

GoogLeNet 提出了 Inception 结构,如图 5-13 所示,它可以并行处理多个分支的特征矩阵,然后将它们的输出按深度进行拼接,从而在不影响网络性能的情况下增加了网络的深度和宽度,并控制模型的参数数量。

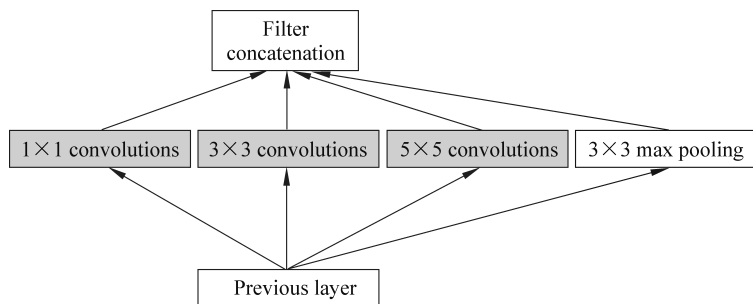


图 5-13 Inception 结构

在 Inception 结构的基础上,通过在每个分支中添加 1×1 大小的卷积层进一步引入降维结构。这一降维过程有助于减小特征图的深度,减少模型的训练参数和计算量。通过这种方式,网络能够更加高效地进行特征提取,同时保持较低的复杂度。

2. 辅助分类器

AlexNet 和 VGG 都只有 1 个输出层。GoogLeNet 有 3 个输出层,其中两个是辅助分类层,其结构如图 5-14 所示。在模型中,辅助分类器(auxiliary classifier)与主干网络并行。在训练模型时,将两个辅助分类器的损失函数乘以权重值(在论文中,取值为 0.3)加到网络