

# 人工智能医学影像诊断的原理和应用

## 临菲电科

医学影像主要有 x 线、计算机断层扫描 (CT) 和磁共振成像 (MRI) 等等, 传统上是医生人工操作对影像进行分析, 为疾病诊断给出依据。随着图像处理的发展应用, 出现了计算机辅助诊断。当前, 由于人工智能技术的发展, 特别是深度学习的应用, 人们正试图将其应用于医学影像分析, 以实现智能诊断, 从而提高诊断速度和诊断准确性, 使病人迅速获得正确的治疗, 此外还能弥补医生的不足。

人工智能在医学影像中的应用, 其作用大体上可分为两个层面: 一是增强成像效果, 包括摄影和图像处理, 提供更加能够诊断疾病的影像; 二是分析诊断, 利用人工智能技术对影像进行分析, 从而给出诊断结论。本文主要关注后者。

## 1 医学图像处理与影像分析

### 检测/定位

检测的主要目的是识别图像中的感兴趣的特定区域, 并给出其周围的一个边界, 例如核磁共振扫描中的脑肿瘤。定位也是用于检测任务的另一个术语。在医学图像分析中, 检测通常被称为计算机辅助检测。计算机辅助检测系统的目的是检测病人早期的异常迹象。肺癌和乳腺癌的检测可以看作是计算机辅助检测的常见应用。

### 分割

在医学图像分析中, 深度学习正被广泛地应用于不同形式的图像分割, 包括计算机断层扫描 (CT)、X 射线、正电子发射断层扫描 (PET)、超声、磁共振成像 (MRI) 和光学相干层析成像 (OCT) 等。分割是通过自动或半自动地勾画图像边界, 将图像分割成不同的有意义的片段 (具有相似的特征) 的过程。在医学成像中, 这些片段通常对应于不同的组织类别、病理、器官或某些其它生物结构。

### 配准

医学图像的配准是指将多幅图像在共同的解剖空间对齐, 是医学图像分析中的一项常见任

务。图像配准是通过变换使源图像与目标图像对齐，在应用深度学习之前就受到了广泛的关注。深度学习的出现使得神经网络渗透到了医学图像配准中。

## 分类

图像分类是医学影像分析和计算机视觉等相关领域的重要问题。在医学影像学的背景下，图像分类实际上就是诊断——计算机辅助诊断。当前深度学习在这方面的应用研究正在取得新的进展，本文将重点讲述其基本原理，并给出几个研究例子。

## 2 基于深度学习的医学影像诊断

医学影像诊断是“图像分类”非常重要的一个应用领域。在医学图像检查的分类中，通常有一个或多个图像作为输入，而最简单的输出是只有一个诊断变量（即是否存在疾病）。本节以此简单的情形作为背景，多个诊断变量的情形与此类似。

当前研究最多的影像诊断还是基于卷积神经网络（CNN）。CNN 用于医学影像诊断的过程类似于一般的图像识别过程，见图 1 所示。其基本原理是：CNN 获取原始像素的输入图像，并通过卷积层、整流线性单元（RELU）层和池化层对其进行变换，完成特征提取，然后输入到完全连接层中，该层计算各分类的分数或概率，最高得分（或最高概率）者即为最后的分类结果。

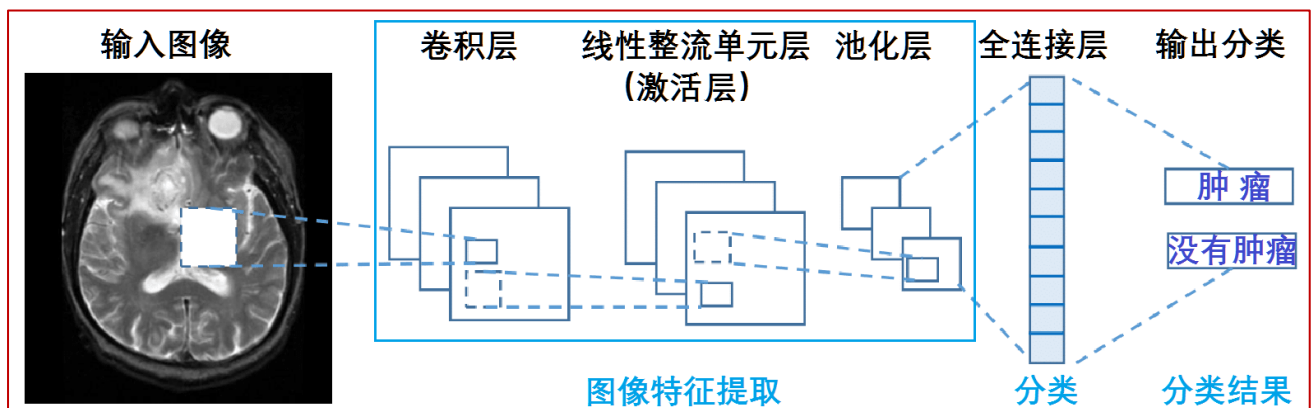


图 1 卷积神经网络影像诊断示意图

### 卷积层

卷积是两个数组（可由矩阵转换为数组）的运算，一个数组由图像中某个位置的输入值（如

像素值) 组成, 另一个是滤波器 (或核)。计算输入与滤波器的点积得到一个输出。按一定的步长将滤波器移到图像中的下一个位置, 重复上面计算过程, 直到覆盖整个图像, 生成特征 (或激活) 映射。

### 线性整流单元层

线性整流单元 (RELU, Rectified Linear Unit) 层是一个将负输入值设置为零的激活函数, 即当输入值  $x$  小于 0 时, 输出  $f(x)$  为 0; 当输入值  $x$  大于等于 0 时,  $f(x) = x$ 。RELU 简化、加速了计算与训练, 并且有助于避免消失梯度问题。其它一些激活函数还有 sigmoid、tanh、leaky RELU, 等等。

### 池化层

池化层的作用是减少参数数量以及图像的大小 (宽度和高度, 但不是深度)。最大池化是常用的方法, “最大” 是指获取最大的输入值而丢弃其它值。其它池化还有平均池化等。

### 全连接层

“全连接” 是将前一层中的每个神经元都连接到全连接层中的每个神经元。可以有一个或多个完全连接的层。这一层的任务是计算出分类中各种可能类别的概率, 最终实现分类。

基于深度学习的影像分析示意图如图 2 所示, 其基本结构是将多个卷积层、激活层和池化层堆叠起来。在空间维进行压缩, 并根据学习到的特征映射数量进行扩展之后, 所有特征被映射到全连接层上, 由最后一个全连接层的激活函数给出分类概率, 最后输出分类结果。

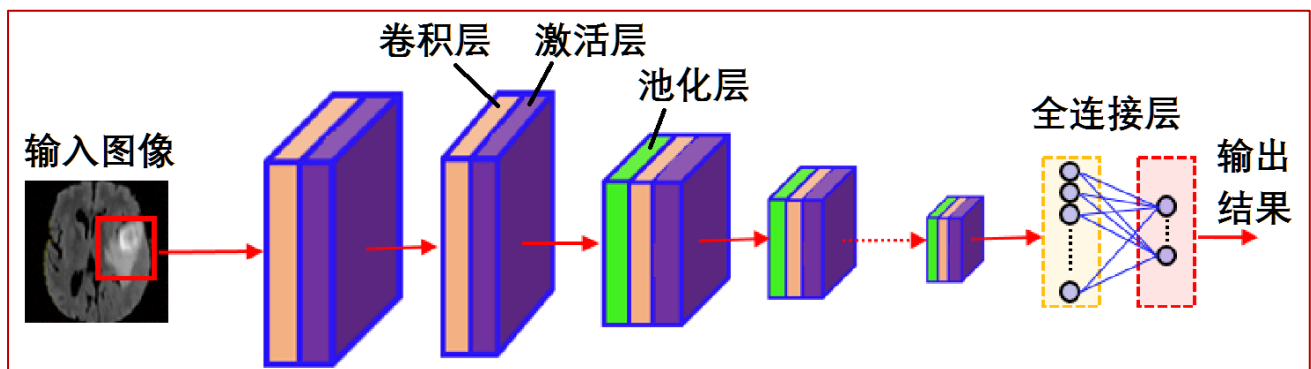


图 2 基于深度学习的医学图像分析示意图

### 3 训练与学习

要让机器能根据医学影像/图像自动给出疾病的判断结果，首先必须要对机器进行训练与学习。

#### 训练

训练是建立一个输入—输出关系的过程。用已知的数据（样本）及其结论（标签）作为输入，让机器在知道输入数据及其应该获得的正确结论的前提下，调整内部参数，从而通过这些参数“记住”输入数据与正确结论之间的关系。可见，“训练”其实就是“教导”机器的过程。经过训练后，机器便建立起了反映输入与输出关系的一种“模型”。这种“模型”类似于函数关系，以后输入新的数据，通过模型的计算，就可以获得结论（输出结果）。

#### 学习

学习是遵循某种规则（学习算法）调整神经网络内部参数的过程。机器学习可分为三类：监督学习、无监督学习和强化学习。在当前的研究中，大部分医学影像诊断是采用监督学习。它的基本原理是：计算输出结果，并与应有的正确结论（标签）进行比较，计算出误差，依据此误差去调整神经网络中各神经元之间的连接系数（即权值  $W_{ij}$ ），然后又根据新的输出计算误差，再调整  $W_{ij}$ 。重复上述过程，直到完成全部训练数据。如图 3 所示。

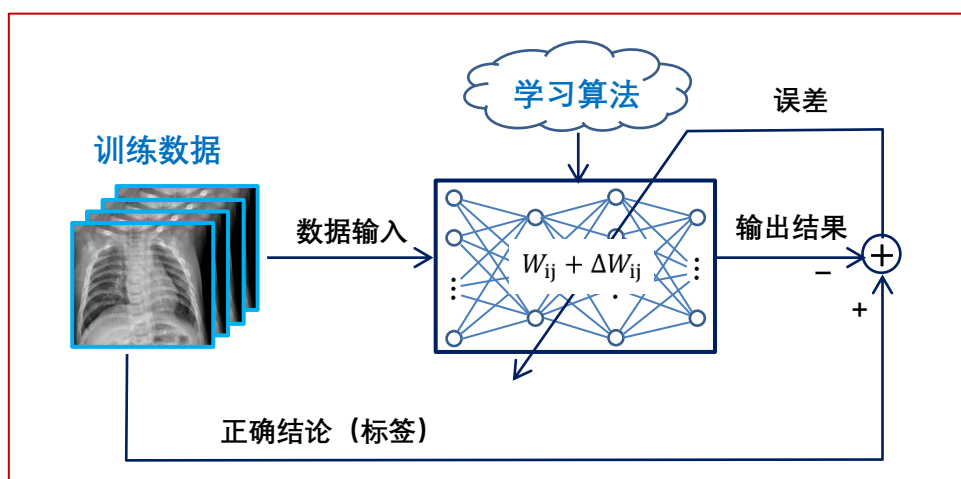


图 3 图监督学习示意图

## 4 典型例子

近几年，深度学习在医学影像分析中的研究获得了显著进展，限于篇幅，下面给出三个研究例子。

### 皮肤癌诊断

2017 年，斯坦福大学 A. Esteva 等人在《Nature》发表了利用深度学习实现皮肤癌诊断的论文（“Dermatologist-Level Classification of Skin Cancer With Deep Neural Networks,” Nature, vol. 542, no. 7639, p. 115, 2017）。皮肤癌的诊断一般是先临床筛查，然后有可能再做皮肤镜分析、活检和组织病理学检查。该文阐述了使用深度卷积神经网络进行的皮肤病变分类。该系统使用的是预先在 ImageNet 数据集上训练的 Google Inception v3 CNN 架构，并在自己的数据集（包含 2,032 种不同疾病的 129,450 张临床病变图像）上使用像素和标签输入进行训练微调，如图 4 所示。757 个训练分类，其由皮肤病新分类法和将疾病映射到训练分类的划分算法来定义。最后的推理分类要笼统一些，由一个或多个训练分类组成（例如，恶性黑色素细胞病变——黑色素瘤类）。推理分类的概率是根据分类结构对训练分类的概率进行求和。系统性能的测试使用经活检证实的临床图像，其测试性能与 21 位经过认证的皮肤科医生进行对比。使用两个关键的二元分类：角质形成细胞癌与良性脂溢性角化病、恶性黑色素瘤与普通的痣。第一个二元分类代表对最常见癌症的识别，第二个则代表对最致命皮肤癌的识别。结果表明，CNN 在这两项任务中都取得了与所有专家同等的性能，其能力水平堪比皮肤科医生，从而展示了人工智能在皮肤癌诊断中的前景。

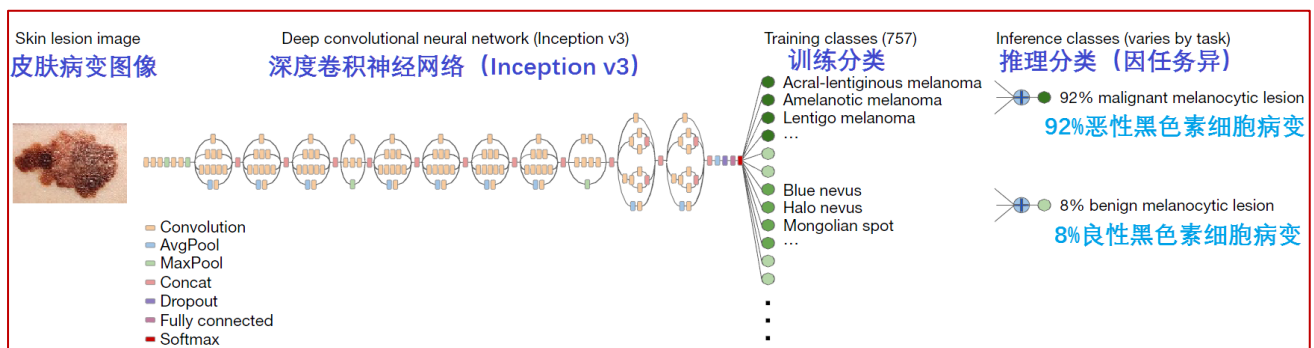


图 4 深度卷积神经网络架构（采用 Google Inception v3）

## 间质性肺病诊断

瑞士伯尔尼大学 Marios Anthimopoulos 等人于 2016 年发表了利用深度 CNN 实现间质性肺疾病诊断的论文（Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network, IEEE Transactions on Medical Imaging, Vol. 35, No. 5, May 2016）。间质性肺病（ILD）是以弥漫性肺实质、肺泡炎症和间质纤维化为病理基本病变。伯尔尼大学的论文提出并评价了一种用于 ILD 模式分类的卷积神经网络（CNN），如图 5 所示。该网络由 5 个卷积层，核为  $2 \times 2$ ，激活函数是 LeakyReLU，采用平均池化，全连接层最后一层的激活函数是 softmax，给出 7 个分类的概率分布。7 个输出对应的分类是：健康、毛玻璃不透明度（GGO）、微结节、实变、网状、蜂窝和 GGO/网状的组合。训练和评估的数据集包含 14696 个图像。这是一个针对特定问题设计的深度 CNN。在一个具有挑战性的数据集中，对比分析证明了有效性。该系统旨在为 ILDs 提供鉴别诊断，作为放射科医生的辅助工具。

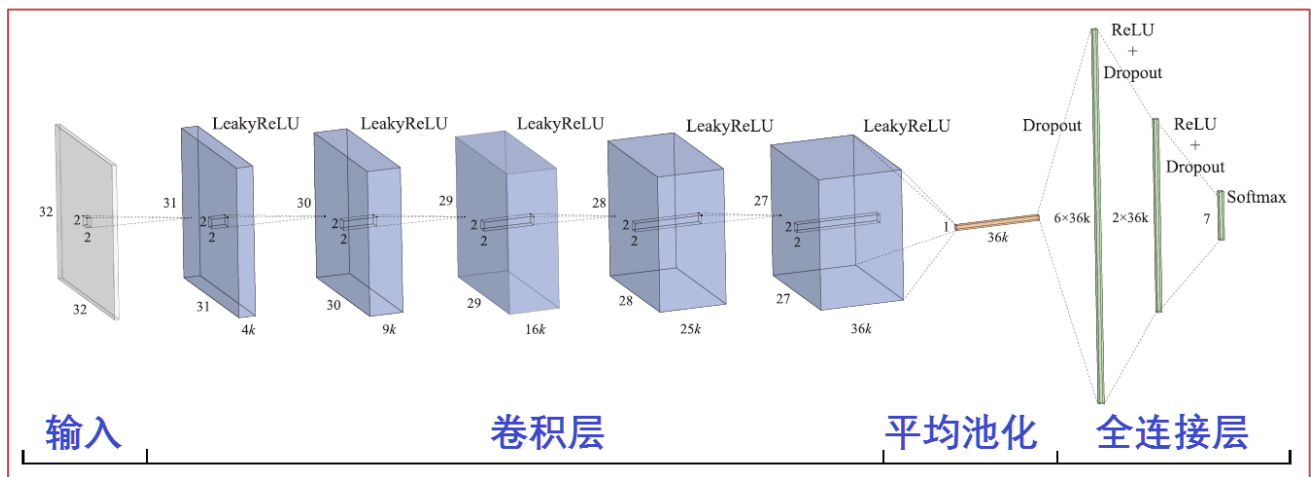


图 5 肺模式分类的 CNN 结构

## 乳腺癌诊断

美国德克萨斯大学埃尔帕索分校 Wenqing Sun 等人关于深度 CNN 用于乳腺外诊断的论文（Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data, Comput Med Imaging Graph, 2016），给出了一个基于图的半监督学习（SSL）方案，使用深度 CNN 诊断乳腺癌。CNN 通常需要大量的标记数据进行训练和参数微调，而该方案只需要训练集中的一小部分标签数据，其余大量数据为无标签数据。诊断系统包括四个模块：数据加权、特征选择、分割协同训练数据标注和 CNN，如图 6 所示。该研究使用 3158 个感兴趣区

域 (ROIs), 每个感兴趣区域包含从 1874 对乳房 X 线照片中提取的肿块, 其中的 100 个 ROIs 作为标签数据, 其余的作为无标签数据, 实验结果表明该方案的准确度为 0.8243。对于混合数据与采用同样数量的有标签数据相比, 后者方案的精度较前者高 3.75%。不过, 混合数据方案 (SSL) 的优点是可以利用无标签数据提高准确度, 这在缺乏大量标签数据时是很有意义的。

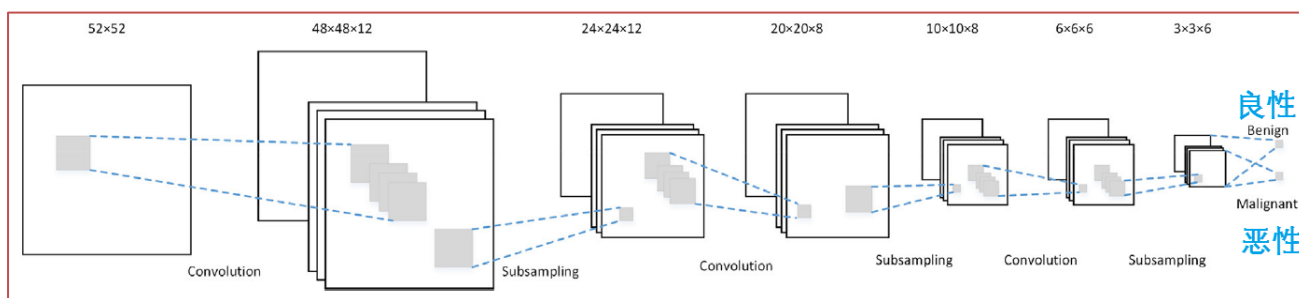


图 6 乳腺癌诊断用的 CNN 结构

## 5 结束语

用人工智能实现医学影像智能分析, 关键还是神经网络的训练和学习。训练的效果与数据密切相关。一是需要用大量的数据“教导”机器, 对于医学影像诊断而言, 即需要大量的病例影像。尤其是监督学习, 需要很大的有标签数据。二是数据的正确性, 即原始数据与结论 (标签) 的正确关系, 显然, 如果使用错误的的数据, 训练后得出的模型就可能导致错误。

许多用于医学影像分析的公共数据集已经开始出现, 而且还会有更多的公共数据集出现, 这将为人工智能机器学习提供极大的数据支持。然而, 数据是过去的, 对于一些突发的新型疾病, 比如新病毒传染病, 由于初期缺乏足够的有标签数据, 对于机器学习是一个很大的挑战, 而学习算法的进步则是应对这一挑战的关键。



临菲信息技术港



临菲信息技术港公众号



临菲学堂