



计算机工程与应用
Computer Engineering and Applications
ISSN 1002-8331,CN 11-2127/TP

《计算机工程与应用》网络首发论文

题目: 多模态融合的膝关节损伤预测
作者: 陆莉霞, 邹俊忠, 郭玉成, 张见, 王蓓
网络首发日期: 2020-10-21
引用格式: 陆莉霞, 邹俊忠, 郭玉成, 张见, 王蓓. 多模态融合的膝关节损伤预测. 计算机工程与应用.
<https://kns.cnki.net/kcms/detail/11.2127.TP.20201021.1444.004.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

多模态融合的膝关节损伤预测

陆莉霞¹, 邹俊忠¹, 郭玉成², 张 见¹, 王 蓓¹

1.华东理工大学 信息科学与工程学院, 上海 200237

2.清影医疗科技(深圳)有限公司, 广东 深圳 518083

摘 要: 膝关节磁共振成像(MRI)是诊断膝关节损伤的首选方法。然而, MRI 影像的人工诊断是费时的, 而且容易出现诊断错误。为了更准确地预测膝关节损伤, 辅助临床医生做出诊断, 本文提出一种多模态特征融合的深度学习方法, 用于检测一般异常、前交叉韧带撕裂和半月板撕裂。首先, 提取梯度方向直方图(Histogram of Oriented Gradients, HOG)特征和局部二值模式(Local Binary Pattern, LBP)特征, 经 contact 融合后利用 PCA 选取特征贡献度超过 95%的特征作为传统特征; 其次, 在 VGG16 模型的基础上加入金字塔融合的思想, 将多个 feature map 的信息融合作为深度特征; 最后将传统特征和深度特征经多层神经网络的能量模型进行相关性融合, 作为多模态的特征, 并得到预测概率。实验结果表明, 上述模型在一般异常、前交叉韧带撕裂和半月板撕裂下 ROC 曲线下的面积(AUC)值分别为 0.9410、0.9708 和 0.8479, 与传统特征和深度特征的效果相比, 具有明显的优势, 可以实现更有效的预测。

关键词: 膝关节损伤; MRI 影像; 多模态; 特征融合

文献标志码:A 中图分类号: doi: 10.3778/j.issn.1002-8331.2002-0095

陆莉霞, 邹俊忠, 郭玉成, 等. 多模态融合的膝关节损伤预测. 计算机工程与应用

LU Lixia, ZOU Junzhong, GUO Yucheng, et al. Prediction of knee injury based on multimodal fusion. Computer Engineering and Applications

Prediction of knee injury based on multimodal fusion

LU Lixia¹, ZOU Junzhong¹, GUO Yucheng², ZHANG Jian¹, WANG Bei¹

基金项目: 国家重点研发计划科技冬奥专项(No.2018YFF0301102, No.2018YFF0301105); 国家自然科学基金(No.61773164); 上海市自然科学基金(No.16ZR1407500)。

作者简介: 陆莉霞(1996—),女,硕士研究生,研究领域为计算机视觉、深度学习;邹俊忠(1960—),通信作者,男,教授/博士生导师,研究领域为模式识别与智能系统应用;郭玉成(1989—),男,博士,研究领域为生物信息学、深度学习;张见(1976—),男,博士,工程师;王蓓(1976—),女,博士,副研究员。

1. East China University of Science and Technology, School of Information Science and Engineering, Shanghai 200237, China

2. Tsimage Medical Technology, Shenzhen, Guangdong 518083, China

Abstract: Magnetic resonance imaging (MRI) of the knee is the preferred method for diagnosing knee injuries. However, the interpretation of knee MRI is time-intensive and prone to errors. In order to assist clinicians to diagnose more accurately, a deep learning model based on multimodal fusion is proposed to detect general anomalies, anterior cruciate ligament (ACL) tears and meniscus tears. First, HOG features and LBP features are contacted as low-level semantic features after selecting the features with over 95% contribution by PCA. Secondly, feature pyramid network on the basis of VGG16 model merges multiple feature maps as high-level semantic features. Finally, the low-level semantic features and the high-level semantic features are fused by correlation to get the prediction probability through an energy model which consists of multi-layer neural network. The experiments show that the AUC of the general anomaly, ACL tear and meniscus tear are 0.9410, 0.9708 and 0.8479, which have certain advantages over the results of the high-level semantic features or the low-level semantic features.

Key words: knee injury; MRI image; Multimodality; feature fusion

1 引言

膝关节主要包括韧带、关节面、肌肉、肌腱等多个组织结构,是人体结构中最复杂、最易发生损伤的关节之一。作为临床上一种常见的骨损伤类型,膝关节损伤多发生在半月板、膝关节韧带等位置,患者多表现为不同程度肿胀、疼痛等,且随着现代交通业、建筑业的不断发展,该疾病发生率呈现出逐年增高的趋势,不仅给患者生活带来了极大的不便,还影响到患者的生活质量^[1]。

随着 X 线、CT、MRI 等影像学检查的普及,膝关节损伤诊断的准确率显著上升,为临床早期治疗提供客观依据。MRI 检查具有较高的软组织分辨率,能够采集到更为丰富的图像信息。通过矢状位、横轴位等方位的 MRI 影像能够对膝关节软组织及周围结构予以全面的探查,获得详细的、准确的图像,因而 MRI 检查对前十字韧带撕裂、半月板撕裂的临床诊断表现出较高的准确性^[2,3]。近年来,膝关

节磁共振成像(MRI)已成为诊断膝关节损伤的首选方法,一个用于分析膝关节 MRI 影像的自动化系统,可以筛选出高危患者,并帮助临床医生做出更准确的诊断。

传统医学影像分析主要采用边缘检测、纹理特征、形态学滤波以及构建形状模型和模板匹配等方法。Haralick 等人^[4]使用灰度共生矩阵来提取乳腺组织的纹理特征;Swain 等人^[5]使用颜色直方图来提取眼底图像的颜色特征。

MRI 影像的多维度、多平面的特性限制了传统医学影像分析方法对膝关节 MRI 的应用^[6],而深度学习能够自动学习多层特征,非常适合于辅助诊断医学影像^[7]。近年来,深度学习已经超越了传统的医学影像分析方法,并在膝关节 MRI 影像领域取得了重大进展。Prasoon 等人^[8]提出了一种新的体素分割系统,该系统集成了三个二维 CNN 网络,分别用于 MRI 三维影像的 xy 、 yz 和 zx 三个

平面的膝关节胫骨软骨的分割。Liu Fang 等人^[9]开发了一套基于深度学习的全自动膝关节 MRI 软骨损伤检测系统,该系统有两个 CNN 网络组成,第一个 CNN 网络进行软骨和骨骼的快速分割;第二个 CNN 网络评估了分叶软骨组织的结构异常。吴恩达团队^[10]首次提出了用于预测膝关节损伤的深度学习模型——MRNet,该模型通过 MRI 影像对膝关节常见的损伤进行预测,提高 MRI 诊断的质量。

本文通过将膝关节矢状面、冠状面和轴向面的 MRI 影像的多种特征融合,对前交叉韧带撕裂、半月板撕裂和一般异常三种常见的膝关节损伤进行分类预测。通过各个对比试验,验证本文方法的有效性和可行性。

2 方法介绍

对于膝关节 MRI 影像的特征提取方法,主要分为传统特征提取方法和深度学习特征提取方法。一般来说,传统方法提取的特征底层语义特征,例如边缘特征和纹理特征。而深度学习如卷积神经网络 (Convolution Neural Network, CNN) 等,提取的图像特征包含更多的细节信息,能表达高层语义信息。底层特征语义信息与高层语义信息对于膝关节损伤的诊断均有重要意义,因此本文提出用于膝关节诊断的一种多模态特征融合网络,如图 1 所示。

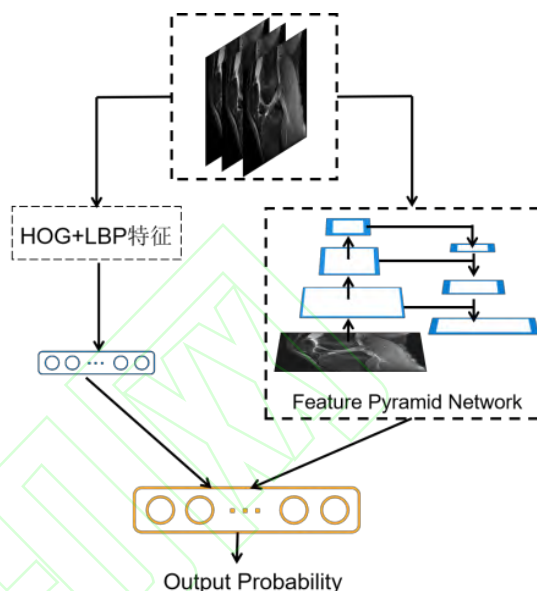


图 1 多模态特征融合网络流程图

首先,提取边缘特征 HOG 特征和纹理特征 LBP 特征,经 contact 融合后利用 PCA 选取特征贡献度超过 95%的特征作为传统特征;其次,在 VGG16 网络模型的基础上加入金字塔融合的思想,将多个 feature map 的信息融合作为深度特征;最后,通过一个多层神经网络将传统特征和深度特征经进行深度融合,并得到预测概率。

2.1 数据预处理

膝关节 MRI 影像的一个样本的数据大小为 $n \times 256 \times 256$,其中 n 为一组 MRI 影像序列所包含的静态图片数目。

由于膝关节主要位于 MRI 的中央位置,因此先将所有静态图片进行裁剪操作,并截取图像中心 224×224 区域。接着进行归一化操作,像素值被归一化到 $[0,255]$ 之间,归一化是为了消除其他变换函数对图像变换的影响。在训练阶段,数据进行数

据增强，包括水平翻转和旋转角度为 25° ，水平和垂直方向移动 0.1 的仿射变换。

2.2 传统特征提取

2.2.1 HOG 特征

HOG 特征提取算法是 Dalal 等^[11]提出的一种特征提取算法，核心思想是通过计算图像局部区域梯度，并将每个局部区域中各像素点梯度的方向直方图级联。图 2 为 HOG 特征提取算法的基本流程图。

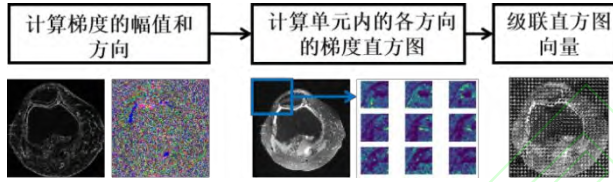


图 2 HOG 特征提取算法的基本流程图

HOG 特征提取的具体步骤如下：

(1) 为消除光照和噪声的影像，首先对图像进行图像灰度化和伽马标准化处理。

(2) 用一维中心对称算子 $k = [-1 \ 0 \ 1]$ 及其转置计算图像横坐标和纵坐标方向的梯度，表达式为：

$$g_x(x, y) = H(x+1, y) - H(x-1, y) \quad (1)$$

$$g_y(x, y) = H(x, y+1) - H(x, y-1) \quad (2)$$

其中， $H(x, y)$ 表示像素在 (x, y) 处的灰度值。得到像素点的梯度幅值和梯度方向为：

$$g(x, y) = \sqrt{(g_x(x, y))^2 + (g_y(x, y))^2} \quad (3)$$

$$\theta(x, y) = \arctan\left(\frac{g_y(x, y)}{g_x(x, y)}\right) \quad (4)$$

(3) 将图像分割为若干个小块 (block)，每个小块由 4 个单元 (cell) 组成，每个单元由 8×8 像素组成，块的滑动步长为 1 个单元。将 $\theta(x, y)$ 在 $[0, \pi]$ 区间上分为 9 个区间，即每 20° 为一个方向 (bin)。单元 (cell) 中的每一个像素点都为直方图通道进行加权投票，权重为 $g(x, y)$ ，这样就得到每个单元内的 9 个方向的梯度直方图。

最后按顺序级联所有 block 的直方图向量，得到图像的 HOG 特征 μ_{HOG} ，维度为 7056 维。

2.2.2 LBP 特征

原始的 LBP 模式算子为 3×3 的局部邻域，以邻域中心点像素的灰度值 $I(c)$ 为基准，对周围其他 8 个像素点做二值化处理，比 $I(c)$ 大的邻域像素值置为 1，否则为 0。将这 8 个二进制数按顺序排列得到一个二进制序列即为中心点像素的 LBP 值。

$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} 2^p \times s(I(p) - I(c)) \quad (5)$$

$$\text{其中, } s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

为对原始的 LBP 模式进行降维，并在数据量减少的情况下能最好的表示图像的信息，本文采用均匀局部二值模式 (Uniform Local Binary Pattern, ULBP)^[12]。当某个 LBP 所对应的循环二进制数从 0 到 1 或从 1 到 0 最多有两次跳变时，该 LBP 所对应

的二进制就称为一个等价模式，其余情况为非均匀模式。通过这样的改进，模式数量由原来的 2^p 种减少为 $p(p-1)+2$ ，二进制模式的种类大大减少，而不会丢失任何信息。

本文用局部二进制编码直方图表示 LBP 图像的特征，将 LBP 图像分为 8×8 块区域，获取每个区域的 LBP 编码直方图，继而得到整幅图像的 LBP 编码直方图，一张 LBP 图像共有 3776 维特征，记为 μ_{LBP} 。

2.2.3 HOG 特征与 LBP 特征融合

经上述方法提取到的 HOG 特征和 LBP 特征均存在特征维度过高，维度之间存在耦合的问题，本文通过主成分分析(PCA)^[13]来达到降维的效果。PCA 不仅减少了特征集的维数，同时避免了重要信息的丢失，保持特征集中对方差贡献最大的特征。

设向量 $X = (x_1, x_2, \dots, x_n)^T$ ，PCA 的具体步骤如下：

(1) 零均值化，得到 X' ： $x_i \leftarrow x_i - \frac{1}{n} \sum_{i=1}^n x_i$ ；

(2) 计算协方差矩阵 $X'(X')^T$ ；

(3) 对协方差矩阵 $X'(X')^T$ 做特征值分解；

(4) 取最大的 m 个特征值所对应的特征向量 w_1, w_2, \dots, w_m ，可以证明第 i 个主成分的方差等于

协方差矩阵的第 i 个特征值 λ_i ，因此主成分 w_i 的贡

献率为 $\frac{\lambda_i}{\sum_{i=1}^n \lambda_i}$ 。

将 HOG 特征和 LBP 特征归一化后，经 contact 融合后的特征为 $[\mu_{HOG}, \mu_{LBP}]$ 。经 PCA 降维后，将特征向量按对应特征值大小从上到下按行排列，前 1325 项的累计贡献率超过 95%，为 95.43%。因此将原特征集降维到 1325 维得到传统特征集

$\mu = (\mu_1, \mu_2, \dots, \mu_{1325})$ 。

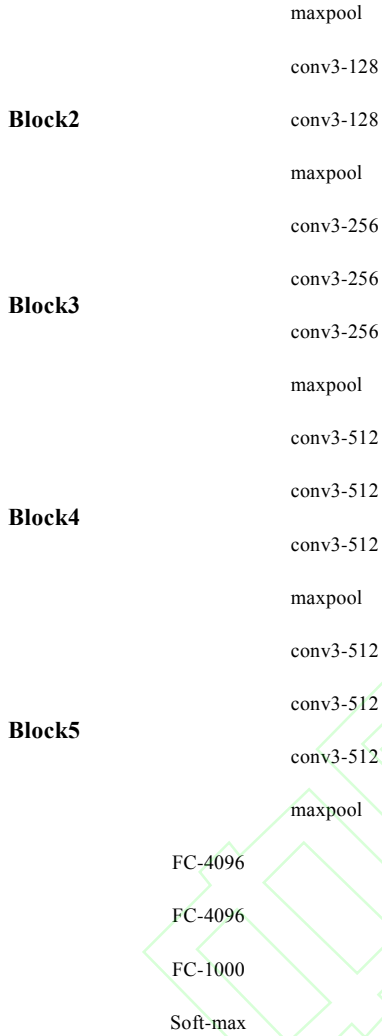
2.3 深度特征提取

近年来，卷积神经网络在医疗图像诊断领域的应用得到了广泛发展迅速，并已取得不错的成果。卷积神经网络通过对图像的卷积操作，可自动完成特征的提取，且这些特征拥有更高级的语义信息，鲁棒性更强^[14]。

VGG16 网络^[15]是 2014 年由牛津大学的视觉几何组 (Visual Geometry Group) 提出，该网络共有五个卷积块，每个卷积块后都是最大池化层，最后三层为全连接层，VGG16 的结构配置见表 1。

表 1 VGG16 结构配置

VGG16NetConfiguration	
Block1	conv3-64 conv3-64



VGG16 网络利用最后一层特征来进行分类,这种方法的优点是速度快、需要内存少。缺点是仅仅关注深层网络中最后一层的特征,却忽略了其它层的特征,但多种信息的融合可以在一定程度上提升诊断的准确率。

由于卷积神经网络中浅层可以学习到比较通用的低级特征,如基本的灰度和边缘信息等,深层可以学到特征的更多细节。因此本文在 VGG16 网络

的基础上加入 Top-down 结构,将提取的特征层逐步搭建为特征金字塔,并进行像素叠加,网络结构如图 3 所示。

通常深度学习被认为在大量有标注的图像上进行训练才可以达到良好的效果,在医学图像中很难有大量带有标注的病理图像用于分割训练,因此本文采用迁移学习的方法,直接使用预训练好的 VGG16 的第二至第五个卷积块 (Block)。将每个卷积块的最后一个卷积层得到的特征映射提取出来,分别为 $F(n)$, 其中 $n = 2, 3, 4, 5$ 。 $F(n)$ 经上采样 (upsampling) 与 $F(n-1)$ 经 1×1 卷积处理后的结果进行像素点融合,再采用 3×3 的卷积核对融合结果进行修正,消除了上采样的混叠效应,得到新的特征映射 $F'(n-1)$ 。金字塔融合的公式为:

$$C(n-1) = f_{1 \times 1}(F(n-1)) \quad (6)$$

$$F'(n-1) = f_{3 \times 3}(F(n) + C(n-1)) \quad (7)$$

获得最后一层融合层的 $F'(2)$, 依次通过 BN 层、ReLU 层、自适应最大池化层和全连接层。BN 层不仅能够加快模型的收敛速度,还能提升分类效果^[16]经过 BN (Batch Normalization) 层。设 $x^{(k)}$ 为 $F'(2)$ 第 k 维特征, BN 层为 $x^{(k)}$ 引入了两个可学习

的参数 $\gamma^{(k)}$ 和 $\beta^{(k)}$, 并利用无差估计输出第 k 维特征 :

$$y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \beta^{(k)} \quad (8)$$

其中 $\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}$, $\mu_B = \frac{1}{m} \sum_{i=1}^m x_i$ 为每个

Batch size B 的平均值 , $\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2$ 为每个

Batch size B 的方差。

ReLU 层加入非线性因素的 , 增加模型的表达能力不够。ReLU 激活函数的表达式为 :

$$f(x) = \max\{0, x\} \quad (9)$$

自适应最大池化层与标准的 MaxPooling 区别

在于 , 自适应最大池化层 AdaptiveMax Pooling 会根据 $\text{input_size}(In)$ 来控制输出 $\text{output_size}(Out)$ 。

stride 和 kernelsize 分别为 :

$$\text{stride} = \text{floor}(In / Out) \quad (10)$$

$$\text{kernelsize} = In - (Out - 1) \times \text{stride} \quad (11)$$

$$\text{padding} = 0 \quad (12)$$

全连接层可以看做 $h \times w$ 的全尺寸卷积 , h 和

w 为前一层输出的大小 , 最终得到 1024 维通过卷

积神经网络提取的深度特征 $g = (g_1, g_2, \dots, g_{1024})$ 。

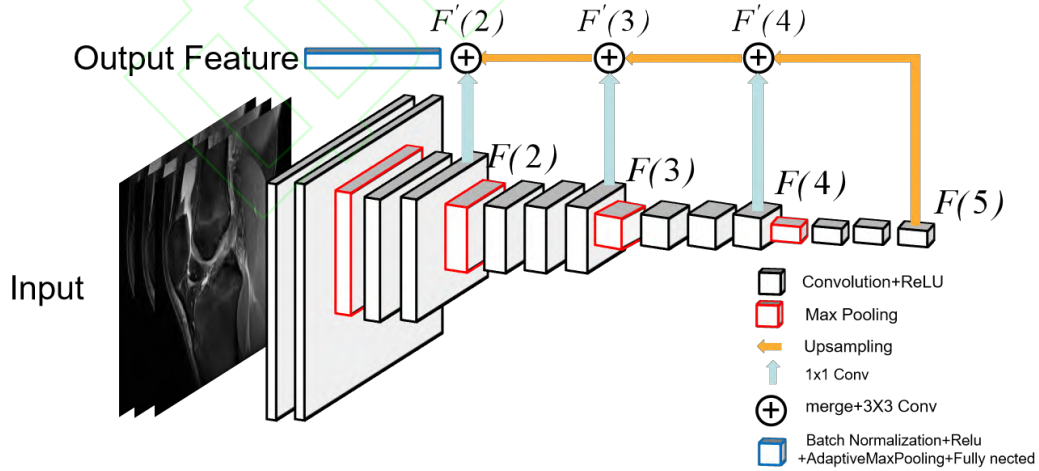


图3 加入金字塔融合的 VGG16 网络结构图

2.4 多模态特征自适应融合

由于提取到的传统特征和深度特征维度均较高 , 若直接级联两个特征向量作为最后的特征会大

大增加时间成本和计算成本 , 因而在多模态特征融合时要减少融合后特征的维度。同时考虑到不同模态的特征在融合结果中所起到的作用 , 以及保留多

模态的相关性，本节构建了一个多模态特征融合模块，该模块包含一个神经元个数小于特征维数的隐藏层和一个 Sigmoid 层，通过最大化特征层的能量比重来训练整个网络，如图 4 所示。隐藏层对融合后特征进行非线性降维，Sigmoid 层将融合后特征映射到 (0,1) 区间，即预测概率。令特征向量 $x = (\mu, g)$ ，前向传播公式为：

$$h_j = (x_i + \beta_i)\gamma_{ij} + \alpha_j \quad (13)$$

$$S = \sigma\left(\sum_{i=1}^m (x_i + \beta_i)\gamma_{ij} + \alpha_j\right) \quad (14)$$

其中 $\sigma(z) = \frac{1}{1 + e^{-z}}$ ， β_i 为可见层的偏差， α_j

为隐藏层的偏差， h_j 为隐藏层向量。为了最大可能的拟合多模态的特征，本节利用能量模型来调整参数^[17]。定义能量函数：

$$E(x, h|\theta) = -\sum_{i=1}^m \beta_i x_i - \sum_{j=1}^n \alpha_j h_j - \sum_{i=1}^m \sum_{j=1}^n x_i \gamma_{ij} h_j \quad (15)$$

其中 $\theta = \{\gamma_{ij}, x_i, h_j\}$ ， $E(x, h|\theta)$ 表示模块的总能量。

定义边缘概率分布：

$$P(x|\theta) = \frac{1}{Z(\theta)} \sum_h e^{-E(x, h|\theta)} \quad (16)$$

$$P(h|\theta) = \frac{1}{Z(\theta)} \sum_x e^{-E(x, h|\theta)} \quad (17)$$

其中 $Z(\theta) = \sum_{x, h} e^{-E(x, h|\theta)}$ ，定义优化函数：

$$\theta_{opt} = \arg \max_{\theta} \sum_{t=1}^T \lg P(x_t|\theta) \quad (18)$$

其中 T 为样本数量。当函数 $\psi(\theta)$ 取得最大值时，特征层的能量比重大，隐层的能量小。当在网络内传输数据时，数据流的方向也是能量衰减的方向，多次迭代之后，网络能量呈衰减趋势，网络趋于有序或者概率分布趋于集中。

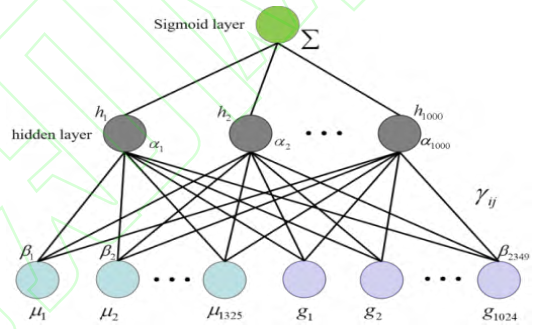


图 4 多模态特征融合网络图

本文将偏差 β_i 、 α_j 初始值设为服从平均值为 0，标准差为 0.1 的正态分布的随机数，权重 γ_{ij} 初始值设为服从平均值为 0，标准差为 0.01 的正态分布的很小的随机数，隐藏层神经元设为 1000 个。

3 实验结果与分析

实验时，所使用的 CPU 型号为 3.4GHz、16 核 AMD1950x，内存容量为 64GB，GPU 为 12G 显存 NVIDIA TITAN-V。软件方面，使用 Python 作为设计语言，配合 opencv 等视觉库进行代码编写。深度学习模型框架使用 Pytorch 以完成对模型的构建，训练与预测。

本文网络模型使用 Adam 优化器，增加模型收敛稳定性和收敛速度，同时也能得到优良的分类结

果。学习率经过调试设置为 $1e-5$,权重衰减为 0.01 ,模型运行 50 轮。

3.1 评价指标

实验过程中使用了公开的膝关节 MRI 竞赛数据集^[10]对本研究中的网络模型进行训练和验证。本研究的核心任务为检阳,因而本文采用了以下评价指标:准确率 Accuracy、Recall 和 AUC(Area Under Curve) 值定量的评估不同模型的性能。

准确率为模型预测正确的样本占有所有参与预测的样本的比例。定义为:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

召回率 $recall = \frac{TP}{TP + FN}$ 度量有多个正例被分为正例。

其中, TP 为真阳性样本数量, TN 为真阴性样本数量, FP 为假阳性样本数量, FN 为假阴性样本数量。

AUC 值被定义为 ROC 曲线下与坐标轴围成的面积。由于 ROC 曲线一般都处于 $y = x$ 这条直线的上方,所以 AUC 的取值范围在 0.5 和 1 之间。AUC 越接近 1.0, 检测方法真实性越高。

3.2 数据集

本次研究中的膝关节 MRI 影像数据集来自于斯坦福大学医学中心整理的从 2001 年 1 月 1 日至 2012 年 12 月 31 日期间内的 1370 个膝关节核磁共振检查数据集——MRNet 数据集^[18]。该数据集包含 1104 例(80.6%)异常 (Abnormal) 检查, 其中前

交叉韧带撕裂 (ACLtear) 319 例(23.3%), 半月板撕裂(Meniscustear)508 例(37.1%)。194 例(38.2%)患者同时发生前交叉韧带撕裂和半月板撕裂。每个样本包括矢状面加权序列 (sagittalplane)、冠状面加权序列 (coronalplane) 和轴向面加权序列 (axialplane)。这些序列的图像数量从 17 幅到 61 幅不等(平均 31.48 幅, SD 7.97 幅)。

数据集被分为训练集(1130 个样本,1088 名患者)和验证集(120 个样本,111 名患者),训练集与测试集的数据分布见表 2。

表 2 训练集与测试集的数据分布见表

Statistic	Training	Validation
Numberofexams	1130	120
Age,mean(SD)	38.3(16.9)	36.3(16.9)
Numberwithabnormality(%)	913(80.8)	95(79.2)
NumberwithACLtear(%)	208(18.4)	54(45.0)
Numberwithmeniscustear(%)	397(35.1)	52(43.3)
NumberwithACLandmeniscustear(%)	125(11.1)	31(25.8)

若将矢状面加权序列、冠状面加权序列和轴向面加权序列同时放入上述多模态特征融合网络中,不仅对硬件要求较高,运行时间较长,而且不同序列的特征表达差异较大,预测准确度不高。

因而本研究针对每个任务(异常、前交叉韧带撕裂、半月板撕裂)和序列类型(矢状面、冠状面、轴向面)训练了不同的 9 个模型。

在验证阶段,采用回归模型得到异常、前交叉韧带撕裂、半月板撕裂的预测概率,回归模型结构图如图 5 所示。

值也更高，因而在偏态的样本中更稳健。

本文模型在卷积神经网络中融入低级语义信息，丰富了模型的特征表达。由表 3 中 ours 的结果可知，在融合了传统特征与深度学习特征后，各性能指标均得到一定程度的提升，尤其是准确率和 AUC 值。在异常、前交叉韧带撕裂、半月板撕裂这三项任务中，最高准确率为 92.50%、92.50% 和 81.83%，最高 AUC 值为 0.9613、0.9742 和 0.8574。同时，本文模型对一般异常的阳性样本的检出率均高于 94%，对前交叉韧带撕裂和半月板撕裂阳性样本的检出率也较高。这对于优先发现高危患者并协助临床医生进行诊断有着重要意义。

由表 3 中 ours 的结果可知，在一般异常预测任务中，轴向面序列表现最佳，准确率、召回率和 AUC 值分别为 92.50%、97.95% 和 0.9613；在前交叉韧带撕裂预测任务中，矢状面序列更有优势，准确率、召回率和 AUC 值分别为 92.50%、94.44% 和 0.9728；在半月板撕裂预测任务中，轴向面序列表现较好，准确率、召回率和 AUC 值分别为 81.83%、82.92% 和 0.8465。此结果与半月板位于两侧胫骨与股骨外侧髁间^[19]而前交叉韧带位于膝关节滑膜外侧^[20]的膝关节结构相符合，因此，本文实验结果具有一定的科学性。

表 3 各模型效果对比

Model	Task	Series	Accuracy	Recall	AUC
Modela	Abnormal	axial	0.8467	0.9053	0.8248
		coronal	0.8250	0.8642	0.8476
		sagittal	0.8417	0.8869	0.8682
	ACL tear	axial	0.7583	0.8029	0.8583
		coronal	0.8083	0.8333	0.8440
		sagittal	0.8583	0.8029	0.9010
	Meniscustear	axial	0.6583	0.6649	0.6968
		coronal	0.6667	0.6346	0.6971
		sagittal	0.6917	0.6766	0.7130
	Abnormal	axial	0.9160	0.9789	0.9356
		coronal	0.8958	0.9474	0.8535
		sagittal	0.9000	0.9579	0.9179
Modelb	ACL tear	axial	0.9000	0.9068	0.9341
		coronal	0.8767	0.8670	0.9113
		sagittal	0.9183	0.9074	0.9666

OURS	Meniscustear	axial	0.8058	0.7969	0.8016
		coronal	0.7967	0.7790	0.7907
		sagittal	0.7684	0.7791	0.8112
	Abnormal	axial	0.9250	0.9795	0.9613
		coronal	0.9083	0.9484	0.9305
		sagittal	0.9167	0.9589	0.9402
	ACL tear	axial	0.9083	0.9074	0.9742
		coronal	0.8936	0.8704	0.9654
		sagittal	0.9250	0.9444	0.9728
	Meniscustear	axial	0.8183	0.8292	0.8465
		coronal	0.8033	0.7930	0.8385
		sagittal	0.7866	0.8092	0.8574

鉴于每个任务均只有阳性和阴性两种可能，为了更好的展示回归模型的效果，将回归模型预测概率大于或等于 0.5 的认为是阳性样本，小于 0.5 的

认为是阴性样本，则在一般异常、前交叉韧带撕裂和半月板撕裂下的实验结果如表 4。

表 4 回归模型各性能指标

Task	Accuracy	Recall	AUC
Abnormal	0.9176	0.9522	0.9410
ACL tear	0.9098	0.9074	0.9708
Meniscus tear	0.8047	0.8124	0.8479

由表 4 可知，本文模型在一般异常、前交叉韧带撕裂和半月板撕裂预测任务中均有较好的性能指标。尤其是一般异常和前交叉韧带撕裂预测任务，准确率和召回率均超过 90%，AUC 值均高于 0.94，说明本次研究所应用的方法能够在这两项膝关节损伤预测中为临床医生提供可靠的基础诊断。由于半月板主要由纤维软骨构成，而 MRI 影像对软

骨组成成分具有局限性^[21]。在这种情况下，本文模型在半月板撕裂预测任务中的 AUC 值仍达到了 0.8479。综合来看，本文方法在膝关节损伤诊断方面具有一定的价值。

表 5 对比了本文算法与 MRNet 竞赛^[18]前三名结果和现有文献中最好结果的平均 AUC 值。

表 5 不同膝关节损伤检测模型性能比较

Model	Average AUC
ours	0.920
mrnet-baseline ^[18]	0.917
Bien et al ^[10]	0.916
dc_baseline ^[18]	0.911
Triple-MRNet ^[18]	0.904

可以发现,在对膝关节损伤进行预测时,表 5 中 6 种深度学习网络 AUC 值均高于 90%,都能够取得较好的预测结果,验证了深度学习强大的特征学习能力。Bien 等^[10]在论文中提出的 MRnet 以及 mrnet-baseline 都是以 alexnet 模型为基础的,AUC 值虽已达到 0.917,但受 alexnet 模型本身深度和广度的限制^[22],卷积层能学习到的特征的数量和种类有限。本文模型以更深层的 vgg16 模型为基础,并在该卷积神经网络的基础上加入了多模态的融合特征,包括多层卷积层的融合特征以及 MRI 影像的纹理和边缘特征,让模型有了更强大的表达能力,以此来获得更好的预测效果。因此本文模型的平均 AUC 值略高与其他深度学习网络,为 0.920。

4 结束语

本文提出了用于膝关节诊断的一种多模态特征融合网络。首先,通过对膝关节 MRI 影像的预处理,从传统与深度学习两方面提取膝关节损伤的多模态特征;然后,通过多层神经网络对特征进行快速的相关性融合;最后,用 logistic 回归整合各个模型的

概率。通过不同的实验,对上述所提到的三种模型进行了比较,同时对回归模型进行了分析,结果显示本文方法在膝关节损伤预测方面具有较大的优势和较高的实用价值。在未来的研究工作中,将尝试更多模态的特征融合,例如 MRI 影像在时间维度上的特征。同时将此应用在膝关节 MRI 的算法应用在其他医学图像类型和不同人体器官的诊断上。

参考文献:

- [1] 李显,赵力,王淑丽,李雪梅.骨性关节炎患者膝关节软骨损伤的关节镜与 MRI 诊断分级研究[J].中国矫形外科杂志,2017,25(1):29-32.
- [2] Oei E H G, Nikken J J, Verstijnen A C M, et al. MR Imaging of the Menisci and Cruciate Ligaments: A Systematic Review[J]. Radiology, 2003, 226(3):837-848.
- [3] Rangger C, Klestil T, Kathrein A, et al. Influence of Magnetic Resonance Imaging on Indications for Arthroscopy of the Knee[J]. Clinical Orthopaedics and Related Research, 1996, 330(330):133-142.
- [4] Wan S, Lee H C, Huang X, et al. Integrated local binary pattern texture features for classification of breast tissue imaged by optical coherence microscopy[J]. Medical Image Analysis, 2017, 38:104-116.
- [5] Manivannan S, Cobb C, Burgess S, et al. Sub-Category Classifiers for Multiple-Instance Learning and its Application to Retinal Nerve Fiber Layer Visibility Classification[J]. IEEE Transactions on Medical Imaging, 2017, PP(99):1-1.
- [6] Doi K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential[J]. Computerized Medical Imaging and Graphics, 2007,

- 31(4-5): 198-211.
- [7] Zeiler M D, Fergus R. Visualizing and Understanding Convolutional Networks[C]. European Conference on Computer Vision. Springer, Berlin, 2014.
- [8] Prasoon A, Petersen K, Igel C, et al. Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network[C]. Medical image computing and computer-assisted intervention: MICCAI. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Berlin, Heidelberg, 2013.
- [9] Fang L, Zhaoye Z, Alexey S, et al. Deep Learning Approach for Evaluating Knee MR Images: Achieving High Diagnostic Performance for Cartilage Lesion Detection[J]. Radiology, 2018:172986-1990/5000.
- [10] Bien N, Rajpurkar P, Ball R L, et al. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet[J]. PLoS Medicine, 2018, 15(11).
- [11] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, 2005.
- [12] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7):971-987.
- [13] Raiko T, Ilin A, Karhunen J. Principal Component Analysis for Large Scale Problems with Lots of Missing Values[C]//Machine Learning: ECML 2007, 18th European Conference on Machine Learning, Warsaw, Poland, September 17-21, 2007, Proceedings. Springer-Verlag, 2007.
- [14] Wang G G, Deb S, Gandomi A H, et al. Opposition-based krill herd algorithm with Cauchy mutation and position clamping[J]. Neurocomputing, 2016, 177(C):147-157.
- [15] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.
- [16] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//International Conference on International Conference on Machine Learning. JMLR.org, 2015.
- [17] Rumelhart D, McClelland J. Learning and Relearning in Boltzmann Machines[M]. Parallel distributed processing: explorations in the microstructure of cognition, vol. 1. MIT Press, 1986.
- [18] Stanford M L Group. A Knee MRI Dataset And Competition[EB/OL].[2018-11-27].<https://stanfordmlgroup.github.io/competitions/mrnet/>.
- [19] 齐滋华,李传福,张晓明. 膝关节半月板撕裂的 MRI 诊断[J]. 实用放射学杂志, 2008(1):78-80+94.
- [20] 严和强. CT 与 MRI 诊断膝关节前交叉韧带损伤的准确率及其影像学特征[J]. 医疗装备, 2019(16).
- [21] Lian H W, Ping H S, Yao H C, et al. Low-field MRI of the Knee Meniscus Diagnosis of the Clinical Value[J]. Progress in Biomedical Engineering, 2009.
- [22] Bottou L, Chapelle O, Decoste D, et al. Scaling Learning Algorithms toward AI[M]//Large-Scale Kernel Machines. MIT Press, 2007.