**肝病医疗综述调研研究报告**

**1. 引言**

**1.1 先写肝病现在面临的情况**

脂肪肝病（FLD）和肝硬化是临床上常见的肝病，发病率高。 脂肪肝通常被描述为脂肪在肝细胞中的积聚。非酒精性脂肪性肝病（NAFLD）的发病率在世界不同地区的人口的9至36.9％之间；即使在军队，非酒精性脂肪性肝病的发病率约为17.1％的无人机的船员和潜艇。肝硬化是晚期肝病的结果。其特征在于通过纤维化和再生结节替换肝组织。肝硬化最常见的是酒精中毒，乙型肝炎和丙型肝炎以及脂肪性肝病。肝硬化是世界上死亡的主要原因。在欧洲，2002年有95,609名男性和53,123名女性死于肝硬化，欧洲不同地区的年龄校正死亡率差异很大。并发症如腹水，食管静脉曲张出血。肝性脑病和肝肾综合征是这类疾病的主要死亡原因。

肝炎是肝脏炎症的统称。通常是指由多种致病因素--如病毒、细菌、寄生虫、化学毒物、药物、酒精、自身免疫因素等使肝脏细胞受到破坏，肝脏的功能受到损害，引起身体一系列不适症状，以及肝功能指标的异常。该疾病可分为急性或慢性。急性肝炎是疾病的快速，急性和疼痛的发作。急性症状对患者更痛苦，但其疗程有限，很少持续超过1或2个月。慢性肝炎是持续超过6个月的肝脏炎症。有五种不同类型的肝炎病毒：A，B，C，D和E。甲型和戊型肝炎是急性型，而乙型，丙型和丁型肝炎是慢性型。慢性肝炎导致肝硬化，其引起肝实质细胞的破坏。

**1.2 其次写如今的研究方向，包括对肝癌疾病的检测，分类，预测等。**

1.2.1脂肪肝和肝硬化患者的脉搏诊断信号分析

中医在脂肪肝和肝硬化两种疾病的治疗中起着重要的作用，脉诊可以帮助诊所医生在诊断和治疗过程中，包括处方和评价。经验丰富的医生可以通过脉搏感觉感受到病人和健康人之间的差异;有人甚至可以从脂肪肝患者中分离肝硬化患者。

分析脉冲收集仪收集的脉搏信号，收集3组人群：健康志愿者，脂肪肝患者和肝硬化患者。本研究使用无监督学习的主成分分析（PCA）和监督学习的最小二乘回归(LS)和最小绝对收缩与选择算子(LASSO)。

本研究中建立了基于无监督学习的PCA和监督学习的LS和LASSO的组合的机器学习方法，为计算机辅助中医诊断成为可能。此外，这项研究可能为中医临床诊断的脉搏诊断科学提供一些重要证据。

1.2.2使用血清与微阵列反应后的标记进行神经网络学习进行肝癌诊断

从过去到现在已经研究了更准确和更容易的疾病诊断方法。通常，通过对有限数量的标志物反应来进行疾病诊断。然而，具有有限数量的标记，不能保证准确的疾病诊断为了解决该问题，近年来已经积极研究使用微阵列的疾病诊断。

微阵列由许多未指定的标记物组成，血清与它们反应能用于疾病诊断。通常，微阵列具有许多数百至数十万个感测点的阵列。需要这么多标记的原因是研究者不知道某种疾病的确切标记。但是这么多标记中有许多是无用标记，Sangman Kim 和Seungpyo Jung等人在2014年提出一种特征选取的方法，大大的减少标记的数量，而且结果显示精确度也没有降低，同时节省时间，降低成本（Kim S, Jung S, Park Y, et al. Effective liver cancer diagnosis method based on machine learning algorithm[C]// International Conference on Biomedical Engineering and Informatics. IEEE, 2014:714-718.）。

1.2.3基于内容的图像检索（Content-Based Image Retrieval，CBIR）的肝病诊断技术

随着医学成像技术的迅猛发展和PACS 系统（Picture Archiving and Communication Systems，影像归档和通信系统）的普及，传统的基于文本方式的数据库管理方式已经渐渐无法满足大规模医学图像数据库的检索需要，为此基于内容的图像检索技术（CBIR）成为近年来该领域中的研究热点。CBIR 技术将成为从海量医学图像数据库中检索图像的一种重要手段。医学图像CBIR系统可以帮助医生在海量数据库中快速寻找具有类似病理特征并已确诊的医学图像，提高疾病诊断的准确率。

大多数CBIR 系统使用医学图像的灰度、纹理、形状等低级特征，这些特征能较好地表征一幅图像。临床经验表明医学图像的视觉信息对诊断结果有很大影响。目前，人们已经提出了各种各样的图像纹理、形状描述子，但仍然不理想。Smith 和Chang 将提取的小波子带系数的统计特性（均值和方差）作为纹理。

南方医科大学生物医学工程学院医学信息技术研究所的于梅等人，在2010年和2011年提出了，本文提出了一种检索方法，用基于非张量积小波滤波器组的纹理全局特征提取方法，提取腹部图多个方向的全局特征，然后结合肝部病灶的局部特征，检索相关图像。这个检索系统针对不同的肝部疾病种类，采用不同的算法，提取该类疾病专有的图像特征向量。用户输入待诊断的图片后，利用已列的算法提取其特征向量，并在该疾病种类的图像数据库中搜寻其匹配图像。实验结果表明算法能提高肝部疾病诊断的准确率，可以作为医生的辅助诊断手段。

1.2.4使用人工神经网络诊断由丙型肝炎病毒引起的肝病

对于CT图像的肝肿瘤分割是肝癌诊断和治疗的重要任务。然而，由于外观的可变性，模糊边界，异质密度，形状和损伤的大小，这是困难的。Li等人（2015）提出了一种基于卷积神经网络（CNN）的自动方法，用于从CT图像分割病变。

**2. 简介**

**肝病医疗方面的描述？目前不知道有没有这个内容**

**对肝病这种疾病的评价指标，比如说【准确度，流行率，灵敏度，错误率和特异性等】**

**3. 基于肝病研究的模型（对应于鄂博说的第二部分吧）**

**3.1 国内外研究的机构**

**3.1.1**

2015年中国中医科学院实验研究中心、清华大学计算机科学与技术系、北京市鼓楼医院中医、朝阳区太阳宫卫生保健中心和天津大学第一教学医院石学敏院士办公室共同研究使用机器学习对脂肪肝和肝硬化患者的脉搏诊断信号分析。

**3.1.2**

日本立命馆大学信息理工学部教授尖端ICT医疗健康研究中心负责人陈延伟的研发小组开发出了利用人工智能（机器学习）辅助肝部肿瘤图像诊断的系统。开发的“相似多时相CT图像检索系统”。系统要使用被称作“单一相”、“动脉相”、“门脉相”、“延迟相”的多张肝脏的CT图像，这些图像是在造影剂注入后经过不同时间拍摄的。系统会提取各相关系（共现关系）的特征，并计算与过去的病例数据库的相似度，最后输出相似病例。

　　检索主要由两步构成。首先是从CT图像中分离出内脏器官（肝脏）和肿瘤，分别制作内脏器官和肿瘤的三维模型。然后提取各相关系的特征（多时相共现特征）。现在，采用机器学习分离内脏器官和肿瘤，3分钟左右即可制成内脏器官和肿瘤的三维模型。今后，还将在提取多时相共起特征时引进机器学习。

　　这项研究是与浙江大学附属医院共同进行的，并在该院进行了先导研究。以实习生和第1、2年的研修医生为对象，在参考检索系统结果和不参考的情况下，比较了5种肿瘤（共计14个病例）的诊断精度和自信度。结果表明，不管是实习生还是研修医生，其诊断精和自信度都得到提高。

　　陈延伟等人还在研究根据医用图像制作每个患者的计算解剖模型。今后，将开发计算解剖模型与机器学习相结合的肝脏疾病诊断辅助系统。

**3.2 肝病研究的模型**

3.2.1Snake模型

Snake模型 (即著名的活动轮廓模型)实质上是一种边缘检测算子，它是一个以能量最小化的样条表示对象的轮廓或表面，其形变受许多不同的能量项约束。主要用于图片分割，在肝病的研究中可以应用到CT图像中的肝的分割、肝肿瘤的边界识别等等。其基本思想是使用连续曲线来表达目标边缘，并定义一个能量泛函使得其自变量包括边缘曲线，因此分割过程就转变为求解能量泛函的最小值的过程，一般可通过求解函数对应的欧拉(Euler．Lagrange)方程来实现，能量达到最小时的曲线位置就是目标的轮廓所在。

  它是一个自顶向下定位图像特征的机制，用户或其他自动处理过程通过事先在感兴趣目标附近放置一个初始轮廓线，在内部能量（内力）和外部能量（外力）的作用下变形外部能量吸引活动轮廓朝物体边缘运动，而内部能量保持活动轮廓的光滑性和拓扑性，当能量达到最小时，活动轮廓收敛到所要检测的物体边缘。

3.2.2医学图像分割技术

图像分割是一个提取感兴趣区域的过程，其分割结果可以为随后的疾病诊断、治疗方案规划以及治疗效果评估等提供参考。CT 由于具有较高分辨力，能更清晰的彰显解剖结构和病变组织等特点，使其广泛地应用到许多系统的疾病诊断。因此，研究图像分割方法在CT 图像中的应用具有非常重要的意义。与基于区域的分割法相对应的是基于边界的分割法，该类方法通常利用区域之间灰度不连续性划分出各个区域的边界。该类方法包含串行和并行边界检测两种，其中串行边界检测主要通过查找梯度值高的像素，然后将这些像素连接起来构成图像的边缘，而并行边界检测通常使用边界检测算子，如Sobel 算子、Prewitt 算子以及拉普拉斯算子等。由于梯度以及其它一些算子依赖于图像的导数，因此使得基于边界的分割方法抗噪性较差。

3.2.3检测肝炎的神经网络模型

（1）前馈反向传播神经网络（FFNN）

FFNN是一种监督网络，属于MLP类别。该架构包括在输入层的19个神经元（因为有19个属性用于诊断肝炎），两个隐藏层每个包含10个神经元，输出层只有一个神经元。 使用训练和验证集训练网络。验证集通过监视错误停止训练过程。在输出层处计算误差，该误差是期望输出和由神经元产生的输出之间的差。

（2）广义回归神经网络（GRNN）

GRNN由四层组成：输入层，模式层，求和层和输出层。在实现GRNN时，输入层有19个神经元，输出层只有一个神经元。 在第一层中使用径向基础传递函数（radbas），而在第二层中使用线性传递函数（purelin）。 使用欧几里得距离加权函数（dist）来计算加权输入。 GRNN的输出属于两个类之一。

（3）自组织地图（SOM）

SOM是一种无监督的神经网络，使用竞争性学习方法。 输出神经元在竞争的过程中被激活，因此一次只有一个输出神经元开启。 基于欧几里德距离选择胜者神经元。