{Objective:} Pneumonia detection is one of the most crucial steps in pneumonia diagnosing system. Clinical information of patients plays an important role in detection of pneumonia. In is paper, a Multimodal Data Diagnosing Network(MDDNet) is described for clinical pneumonia detection.

{Method:} MDDNet is based on deep learning neural network and analyzes multimodal data. We use Recurrent CNN, which can keep 3-D spatial information and reduce the need of calculation resource, to capture visual features from CT image data. Each slice of CT is transformed into one 3-channel(Lung Window, High Attenuation, Low Attenuation) image which can provide more information of lung density. Meanwhile, patient clinical information like complaint, age and gender is adopted and provides more abundant information to improve the accuracy of pneumonia detection. A Long Short Term Memory(LSTM) network is used to analyze semantic features of patient complaints and provides information which image data cannot provide, like how many days the patient has been ill. Information about age and gender can provide priori information since age and gender is associated with certain kinds of pneumonia. CT visual features, complaint semantic features, patient age and gender will be fused together and calculate joint distribution to predict whether these cases are pneumonic.

{Results:} We analyze 1002 clinical cases from The First Affiliated Hospital of Army Medical University. Our model achieves 0.945 in accuracy, and has a very balanced performance in sensitivity and specificity. As far as we know, we are the first to detect pneumonic cases based on large scale clinical raw data using multimodal clinical data.

{Conclusion:} Our method proves that multimodal data provides more abundant information than image data only and improves the accuracy of pneumonia detection.

{Significance:} Our model can be extended and include more kinds of clinical data to give out more reliable and explainable detection results.

1.利用临床多模态信息对肺炎进行检测

2.利用RCNN处理CT序列。由于医生需要根据CT计算各器官的大小，因此不能随意拉伸、裁剪图片。完整的CT保留了大量的冗余信息，比如骨骼组织、正常的心血管组织等。如果直接用3-D CNN，这些冗余信息会给计算机带来沉重的负担。RCNN利用一个CNN统一从CT的2-D图像序列提取视觉信息，利用LSTM获取序列的前后关系，保留3D信息同时减少卷积计算量。

3.利用3 channel(Lung Window, High Attenuation, Low Attenuation)处理CT。Lung window提供肺部正常组织的信息以及肺部整体状况；High Attenuation提供肺部密度异常增高的信息，Low Attenuation提供肺部密度异常降低的信息。三个通道互相补充，既保持了对正常肺部组织的信息提取能力，又增加了对肺部异常组织的信息提取能力。

4.利用LSTM分析主诉文本的语义信息，与CT影像信息相关联。比如“右侧 季肋部 疼痛 二月”，“右侧 季肋部”提供方位信息，“疼痛”提供症状信息，“二月”提供生病时长信息，方位信息、症状信息、时长信息与CT影像相互应证，增强了CT提取的视觉信息。

|  |  |  |
| --- | --- | --- |
| **方位信息** | **症状信息** | **时间信息** |
| 胸背 | 咳嗽 | 月余 |
| 左上胸 | 隐痛 | 一周 |
| 双上胸 | 恶心 | 年余 |

5.利用年龄、性别为模型的决策提供先验信息。由于男性有更高的吸烟比例，因此男性肺部CT出现阴影的可能性更大，患肺炎的概率更大；同时由于老人、儿童抵抗力较青壮年低，因此老人儿童患肺炎的概率也更大。另外，年龄和性别与胸腔、肺部的形态具有直接联系，一般来说，成年男性的胸腔比成年女性胸腔大，成年人胸腔比未成年人胸腔大。

# 概括总结：

1.利用临床多模态信息对肺炎进行检测。

2. RCNN利用CNN从2-D CT图像序列提取视觉信息，LSTM提取序列的前后关联关系，既保留3D信息同时减少计算量。利用RCNN处理CT序列,会导致沉重的计算负担完整CT序列,保留了大量冗余信息。

3.利用3 channel处理CT。Lung window提供肺部正常组织的信息以及肺部整体状况；High Attenuation提供肺部密度异常增高的信息；Low Attenuation提供肺部密度异常降低的信息。三个通道互相补充，既保持了对正常肺部组织的信息提取能力，又增加了对肺部异常组织的信息提取能力。

4.利用LSTM分析主诉文本的语义信息。文本中包含的方位信息、症状信息、时长信息与CT影像相互应证，对CT提取的视觉信息起到增强作用。

5.利用年龄、性别为模型的决策提供先验信息。不同年龄、不同性别在胸腔、肺部的形态上具有差异性。

三通道CT影像序列分别表示了肺部不同且具有互补性的特征信息，主述文本描述了被检查者对其影像特征的主观感受，不同年龄和性别被检查者的影像形态特征具有一定的差异性。

* 三通道CT影像序列分别表示了肺部不同且具有互补性的特征信息，互补性对诊断具有定性判断作用。
* 主述文本描述了被检查者对其影像特征的主观感受，相似影像特征的主观感受具有相似性，相似性对诊断具有辅助增强判断作用。
* 不同年龄和性别的被检查者的影像形态特征具有一定的差异性，其可为医生的决策提供先验信息。
* 网络具有医生的综合判断能力，且计算复杂度低。

数据特征（客观性）

设计相应的网络结构