{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.

需要进一步思考如下问题：

特点

1.利用多临床多模态信息对肺炎进行检测

2.利用RCNN处理CT序列（序列之间存在冗余性），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)为模型提供更丰富的（不要用不确定的语言描述，科学语言是严谨、准确。三通道分别表征了肺部的什么特征，且三通道信息是否具有互补性？如有，互补性是什么？由于互补性只有融合才能准确刻画肺炎特征）肺部密度信息

4.利用LSTM分析主诉文本的语义信息（哪些信息，是否有格式化或表格？这些文本信息与图像信息有什么关联，是否具有一致性和关联性？如有，具体表现是什么？），获得图像无法获取的信息，比如生病的时长等

5.由于年龄、性别与某些特定肺炎（这个词语不准确，）有着很强的联系(年龄与图像有什么关联，性别与图像有什么关联，男性和女性图像分别有什么特征？肺病在不同年龄阶段的表现有什么特征？这些不同表现为差异性)，利用年龄、性别为模型的决策提供先验信息

四大特性：CT序列图像冗余性、三通道互补性、文本与CT图像关联性、年龄性别与CT图像的差异性