大綱

introdiction

根據研究發現covid-19致死的病例，多數人含有心血管疾病，

而心血管疾病會有突發性的特徵，

大量醫學表明，當心臟驟停時，最佳搶救時間在4分鐘內。

因此如果我們能提前檢測到細微跡象並且採取有效措施，

就可以避免心血管疾病患者死亡。

目前最直接有效的臨床檢查各種心血管疾病的方法是心電圖（ECG）

雖然醫生和醫院可以通過檢查來了解 COVID-19 患者的健康狀況，

但這種傳統的診斷和治療缺乏即時性和連續性，

所以捕捉到 COVID-19 患者的特徵難度較大。

因此，

及時有效的監控分析和診斷 COVID-19 患者的心電圖信號就會是非常重要。

目前穿戴式醫療設備可以測量心跳、心律、血壓、血糖計等，

可以用來幫助我們跟踪 COVID-19 患者的健康狀況

並且提供個性化的醫療保健解決方案。

可是現有的無線通信技術存在高延遲和低速度。

而且，醫療量能的不足所以依靠人工識別心電圖診斷方法已經是不可行的

所以本篇論文提出了一個新的系統，是由5G及時加上深度學習的方式來監測患者的心血管。

介紹整體結構圖，它由四個模塊組成。

首先區塊負責數據的採集和傳輸。

第二區塊負責傳輸和存儲，

第三區塊負責用於深度學習的訓練。

先介紹第一個區塊

主要透過各種感測器、控制器、處理器和電源等穿戴設備的耗材。

來完成採集心電圖數據。

再藉由手持裝置通過5G網絡上傳到雲服務器。

與傳統的網路相比5G網路有較低的延遲、大頻寬和大連結。

可以滿足ECG高流量密度、高連接數密度和超高流動性。

也改善了網絡的頻譜效率，所以它非常適合本文的即時系統的環境。

第二階段是傳輸和存儲，

Kafka優點是他可以有效的處理高吞吐量的資料，且有較高的性能。每一個節點支持超過上千的客戶端同時連線，最重要的優勢是不會有數據丟失的問題。

Flink框架本质上来说是流引擎，但是在实际工作当中，实现了流处理和批处理的统一。

在Flink上資料可以被作為 無界 或者 有界 流來處理。

1.無界資料流從有定義流的開始，但沒有定義流的結束。它們會無休止地產生資料。無界流的資料必須持續處理，即資料被讀取後需要立刻處理。我們不能等到所有資料都到達再處理，因為輸入是無限的，在任何時候輸入都不會完成。處理無界資料通常要求以特定順序讀取事件，例如事件發生的順序，以便能夠推斷結果的完整性。

2.有界資料流 有定義流的開始，也有定義流的結束。有界流可以在讀取所有資料後再進行計算。有界流所有資料可以被排序，所以並不需要有序讀取。有界流的處理通常被稱為批處理。

所以Fink的優點是同時支持高吞吐、低延遲、高性能。

因此整體的架構為患者使用手持設備傳送資料首先會通過 TCP 讀取框架，先經過緩衝區，在到我們KAFKA系統，接下來再經由Flink架構下訓練和預測，最後片別需不需要通知管理人員來處理。

第三階段是深度學習的訓練

數據再進行訓練前，我們需要先預處理數據

首先心電圖訊號過於微弱，所以非常容易受電子設備或其他外在因素所干擾，因此本文選用小波轉換8層將心電信號分解，由於心電信號的波長、頻率和振幅雜訊頻率差異較大，所以選用軟閾值處理方式，即不同閾值用於量化處理 不同的變換尺度。最後，心電圖信號根據低頻係數重構小波分解得到的第8層的 以及每一層的高頻係數，來得到去噪的心電圖信號

其中評價心電圖訊號的公式有，

SNR信噪比 越大雜訊越小，MSE 平均方誤差 越小代表失真程度越小

s(i) 表示 包含噪聲的原始心電信號，x(i) 表示 去噪心電信號，

N 表示收集到的心電圖信號長度。

心電信號去噪後， 由於一個長時間的心電圖訊號，會需要較長的時間進行運算，因此波形需要定位和分割

我們知道由於 PQ 波約 0.3 秒，STU 波之約 0.4 秒，因此一次完整的心跳約為 0.6-0.8 秒。再加上 R 波斜率較大的特性和一次完整心跳的時間，來進行心跳分割。

數據增量

基本上正常的心電圖訊號會比異常的訊號多很多，因此需要進行數據增量，本篇論文用的方法是用SMOTE的方式合成新的異常心電圖樣本數

合成的方式簡單來說，找出異常心電圖樣本的中心點，接著在異常心電圖樣本中挑一點與中心點之間隨機形成新的異常心電圖樣本。

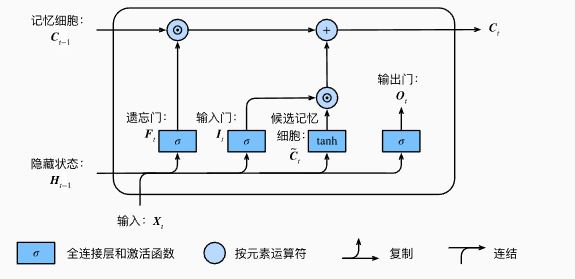
代表第i個心電圖異常的樣本和 代表 。

​CNN

**卷積神經網路**是一種[前饋神經網路](https://zh.wikipedia.org/wiki/%E5%89%8D%E9%A6%88%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C)，它的人工神經元可以回應一部分覆蓋範圍內的周圍單元，對於圖像處理有出色表現

LSTM

1. Memory Cell: 將運算出的數值記憶起來，以利下個cell運用
2. Output Gate: 控制是否將這次計算出來的值output，若無此次輸出則為0
3. Forget Gate: 控制是否將Memory清掉(format)



LSTM的優勢在於擅長處理時間序列中有規律性的事務，例如股票或是銷售分析等。

CNN 和 LSTM 都可以對圖像輸入進行分類，所以將兩種網絡模型結合起來 以提高分類效率。如果輸入增加時間 CNN本身無法處理的特徵，組合 兩者中更有利。由於心電圖 信號是根據收集的生理信號 隨著時間的推移，它包含了豐富的時域特徵。對於LSTM很適合此種可以保存 上下文信息，並且依賴於遠程時間。

講圖的架構

實驗

本篇論文使用 MIT-BIH 的數據庫，包括了 47 名受試者的4000 多筆 24 小時的心電圖，其中有 48 筆時常為 30 分鐘的紀錄文件。

然後分成5類 根據AAMI標準，即N（正常心跳）， S（心房過早），V（心室過早），F（融合）指的是當心臟的自然節律和[起搏器](https://en.wikipedia.org/wiki/Pacemaker)的脈衝同時激活心室的同一部分時，就會發生心室融合搏動，Q（沒有標記的心跳）

對於 S、V、F 和 Q，首先選擇 20% 作為 測試集，並使用剩餘的 80% 生成新的 數據。 為了保持實驗的可信度， 本文采用十折交叉驗證方法。

在介紹實驗解結果之前需要了解如何去評估

第一種是

由準確度、特異度和靈敏度來作為判斷指標，其中準確度用來表示模型的準度，特異度來表示真陰性的機率，靈敏度是來表示真陽性的機率。

第二個指標是

ROC 曲線：又稱為分類模型，是一種用座標圖式的分析工具，X=0代表沒有偽陽性，Y=1 代表著沒有偽陰性，簡單的判斷方式可以用 ROC 曲線所覆蓋面積來判斷該模型的性能。

從這張圖可以看出，通過預測 測試集上的不同模型，發現分類 傳統機器學習 SVM 效果較差 與深度學習方法相比，我們提出方法準確度比其他深度學習方法要高，他們的準確度、特異度和靈敏度分別達到了 99.29%、99.53% 和 97.77%。

接下來是ROC曲線分析

可以看到這四個模型的ROC曲線都有很大的 偏離 45 度對角線。其中， SVM模型下面積小，訓練好的SVM 模型對冠心病患者的心電信號具有一般的分類效果。這表示傳統機器學習數據量大，對特徵較多的數據特徵提取效果較差， 使模型的泛化能力變差；下面積 CNN和LSTM差別不大。也說明深度學習模型可以有效在一定程度上提取心電數據的單邊特徵，使模型泛化更好；可以看出我們的面積提出的方法是最大的，表明 CNN+LSTM 模型不僅能提出特徵 心電信號本身，也準確提出其中的時域特徵。將它們組合起來進行分類 可以有效提高分類。

總結，我們採用5G發送和接收來自可穿戴醫療設備的數據。

再來，應用Flink流式數據處理框架訪問心電圖數據。

最後，我們使用CNN和LSTM得到自動預測 COVID-19 患者的心血管健康狀況

實驗結果表明此方案可以很好地解決上述問題，

將心血管疾病的預測準確率提高到 99.29%。

最後論文中有提到，未來希望可以使用對抗網路來解決信號資料量不均的問題

。