



深度神經網路之呼吸音訊降噪系統

A Deep Neural Network-Based Breathing Audio Noise Reduction System

指導教授: 朱元三

專題生: 林子恩、鍾岳儒

摘要

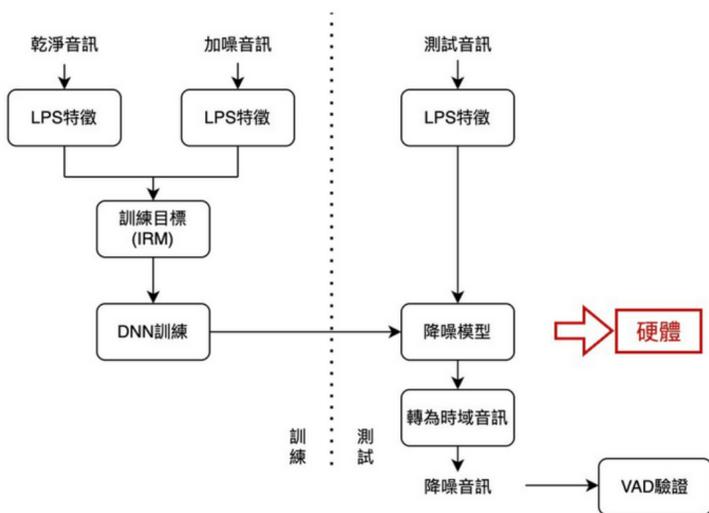
呼吸音訊的頻率及特徵檢測可以及早發現許多身體異狀。但錄製呼吸音訊時，會受到環境許多噪音的干擾，因此如何將噪音去除留下乾淨之呼吸音訊為一大課題。許多傳統上降噪方法，透過線性轉換實現降噪，但無法對非線性的噪聲做處理，而大部分的噪聲都是以非線性呈現，而本研究針對戶外常見之噪音進行降噪處理，在特徵擷取部分採用呼吸音訊之對數功率譜(Log Power Spectrum, LPS)經過深度神經網路(Deep Neural Network, DNN)來學習呼吸音訊及噪音音訊之間的特徵關係，並以理想比值遮罩形式(IRM)做為訓練目標，讓錄製之帶噪呼吸音訊經過此降噪系統後能降低噪音的影響盡可能得到乾淨之呼吸音訊，以利於後期呼吸音訊相關之應用。

實驗結果

將乾淨呼吸音訊與噪音（蟲鳴聲、鳥叫聲以及風切聲）以3種SNR值(-2.5dB、0dB、5dB)進行加噪，並計算降噪前後之準確率。以下為不同神經元數進行訓練並降噪的結果。

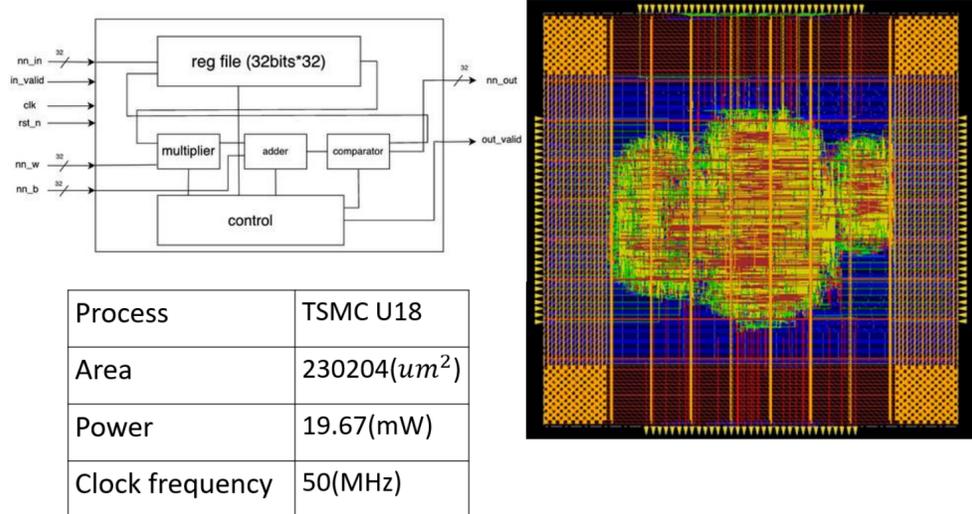
隱藏層神經元數	5db 準確率提升		0db 準確率提升		-2.5db 準確率提升	
256	0.58%	-10.62%	6.03%	-7.96%	7.75%	-5.41%
128	2.59%	-9.73%	6.89%	-7.08%	7.75%	-1.77%
64	6.04%	4.43%	11.20%	5.31%	12.93%	9.73%
32	6.04%	3.54%	10.34%	7.08%	12.93%	11.50%
16	6.04%	5.31%	13.79%	7.96%	12.93%	12.39%

系統流程



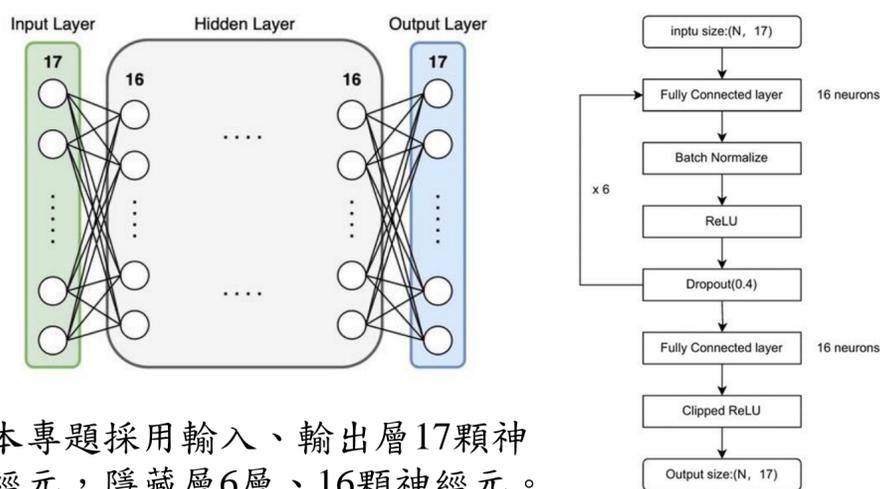
本實驗需先收集呼吸音訊，接著將乾淨呼吸音訊與噪音音訊以不同之SNR值進行加噪，並透過加噪音訊及乾淨音訊算出IRM作為訓練之Label。深度神經網路(DNN)訓練階段，先將乾淨呼吸音訊以及加噪呼吸音訊進行LPS特徵擷取，再來會以IRM為目標，進行DNN之訓練。訓練完之降噪模型會將測試音訊作為模型輸入得到DNN輸出，並將輸出與音訊做處理得到降噪音訊，再以VAD和ZCR判斷呼吸器及吸氣評估降噪系統效果。

硬體實現



硬體加速器架構主要藉由 register file 存取每層輸出結果，並且作為下一層的輸入，當下一層的所有神經元都算出來後，這些值會作為新一層計算的輸入，透過有限狀態機的方式控制乘法器及加法器的運作，並且透過比較器實現 ReLU 的行為，藉由不斷的循環完成具有六層隱藏層之神經網路。

DNN架構



本專題採用輸入、輸出層17顆神經元，隱藏層6層、16顆神經元。

結論

在本專題中，透過深度神經網路（DNN）技術實現了針對戶外噪音的呼吸音訊降噪系統，成功去除了多種類型的背景噪音，使VAD、ZCR準確率提高5-13%。在硬體實現方面，透過有限狀態機設計並完成 Verilog 硬體實作，使得處理時間從 20 ms大幅縮短至75 us，達到約 250 倍的加速效果。未來計劃使用pipeline或systolic array，將速度提升，以及浮點數運算轉為定點數運算，以節省所需面積。