



應用於 CNN 降噪還原系統與其硬體加速器

組別: 晶片系統組 學生: 賴柏亦、蔡政宏 指導教授: 朱元三

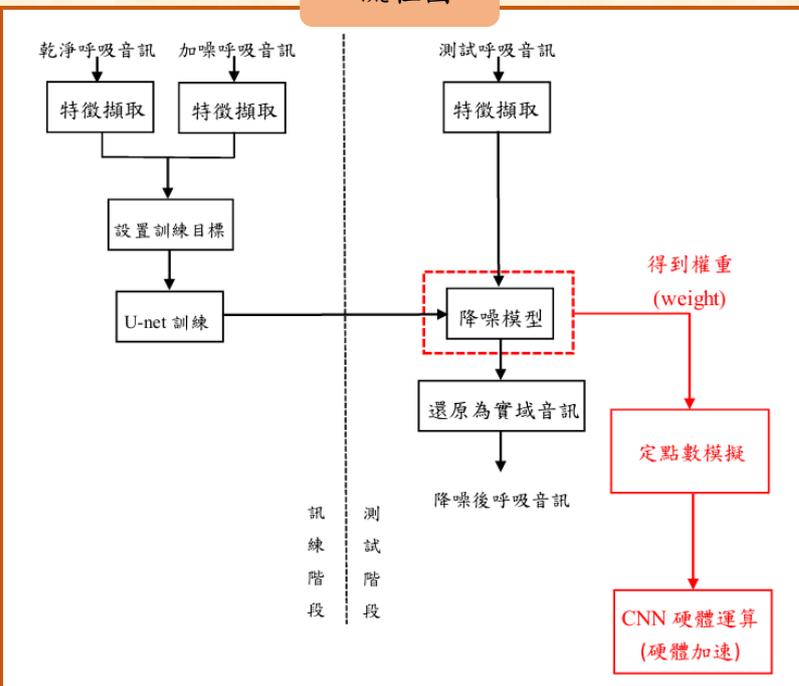
摘要

卷積神經網絡 (CNN) 在圖像降噪和還原上表現突出。透過多層卷積濾波器自動學習圖像特徵，CNN 能準確去除噪音，比傳統手動特徵提取更高效、適應性更強。特別是在 GPU 加速下，CNN 可以快速處理大數據集，常見架構如 ResNet、自編碼器和 U-Net 適合不同降噪需求。這些技術廣泛應用於醫學圖像、攝影和衛星圖像等領域，提升診斷精度和圖像質量。透過自動學習噪聲特徵，CNN 在多個領域顯著提高圖像質量。

實作方法

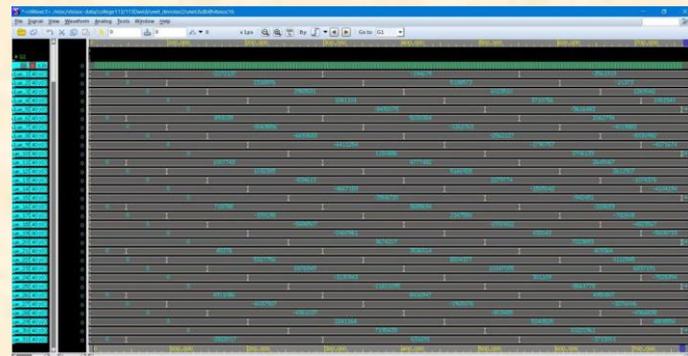
在訓練階段，本研究將乾淨與加噪的呼吸音訊進行特徵擷取，並以遮罩 (Mask) 為目標進行 Label 分類，以訓練卷積神經網絡 (CNN) 模型來進行降噪測試。音訊經過 46.4ms 音框化，兩音框間重疊 34.8ms，並使用漢明窗防止邊緣資訊遺失，隨後透過短時傅立葉轉換及對數功率運算取得 LPS 特徵。本論文採用 U-Net 模型，其 Decoder 和 Encoder 各包含卷積、批量歸一化、激活、Dropout 和最大池化層等層結構，以遮罩學習噪音特徵並抑制，保留呼吸音訊特徵。

流程圖

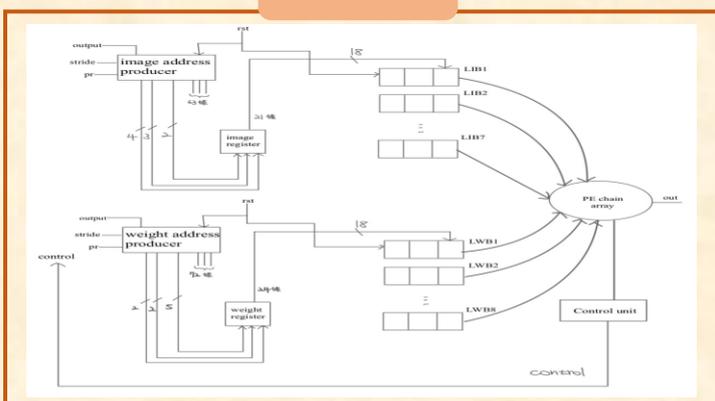


實驗結果

SNR	-2.5dB		-1dB		0dB		2.5dB	
	in	ex	in	ex	in	ex	in	ex
呼吸種類								
總共	125	160	125	160	125	160	125	160
降噪前正確數目	99	111	102	112	102	113	104	119
降噪前準確率	79.2%	69.3%	81.6%	70%	81.6%	70.6%	83.2%	74.3%
降噪後正確數目	114	119	116	126	116	122	114	122
降噪後準確率	91.2%	74.3%	92.8%	78.7%	92.8%	76%	91.2%	76.3%
提升	12%	5%	11.2%	8.7%	11.2%	5.4%	8%	2%



硬體架構



結論

本專題利用乾淨和加噪的呼吸音訊，經短時傅立葉轉換提取 LPS 特徵，並將其輸入深度神經網絡，以 IRM 遮罩 (MASK) 作為訓練目標進行模型訓練。透過平均 PESQ 和呼吸聲分類準確率的測試結果顯示模型性能提升。模擬結果表明，隨著 SNR 提高，呼吸聲的準確分類數量顯著增加。在硬體方面，成功使用 PE-chain 架構完成卷積層運算，相較於純軟體運行，速度更快。