

國立臺北大學資訊工程學系111年專題

基於遠程光電容積脈搏波特徵及機器學習的深度偽造影片偵測方法

Video Deepfake Detection using Remote Photoplethysmography Feature and Machine Learning

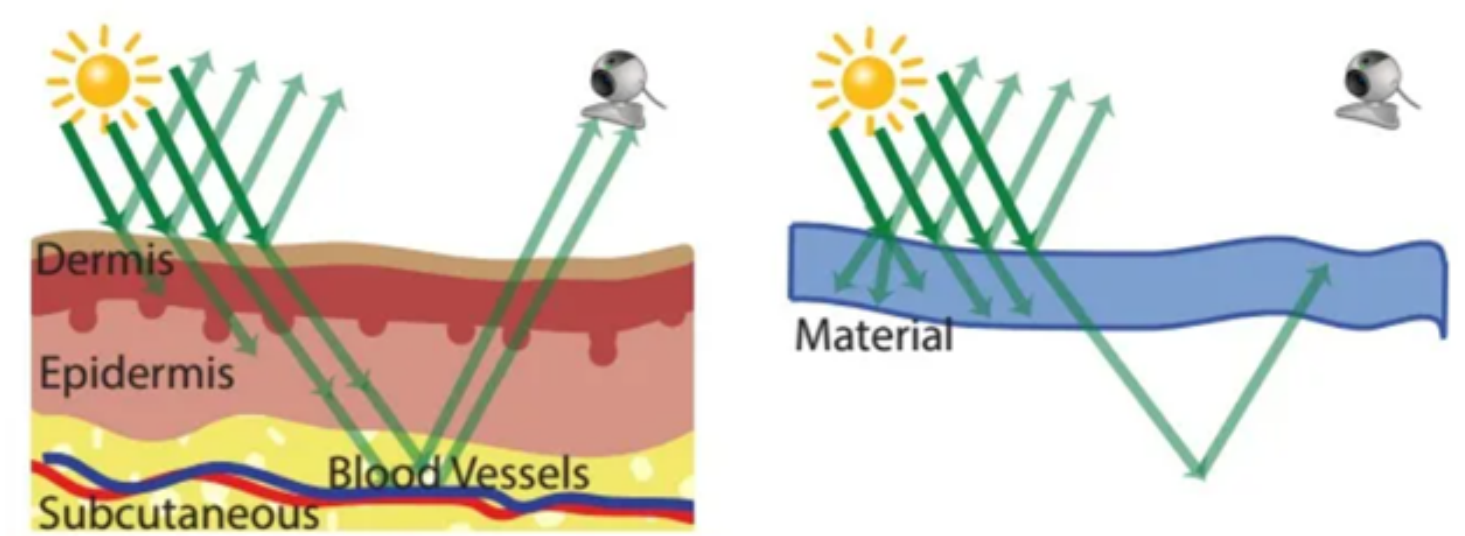
呂霈芳、謝采軒、林翊婷、施巧軒

摘要

隨著人工智慧和機器學習技術的進步，製作高度逼真的虛假內容變得相對容易。這種技術在開放的網絡平台上迅速流行，引起了虛假資訊和侵犯隱私權的嚴重問題，也由此技術帶動了deepfake偵測工具快速的崛起。

此研究將採用人體特徵作為偵測影片真偽的主要依據，並使用數學特徵和圖形特徵作為輸入來完成一套基於遠程光電容積脈搏波的深度偽造偵測方法。

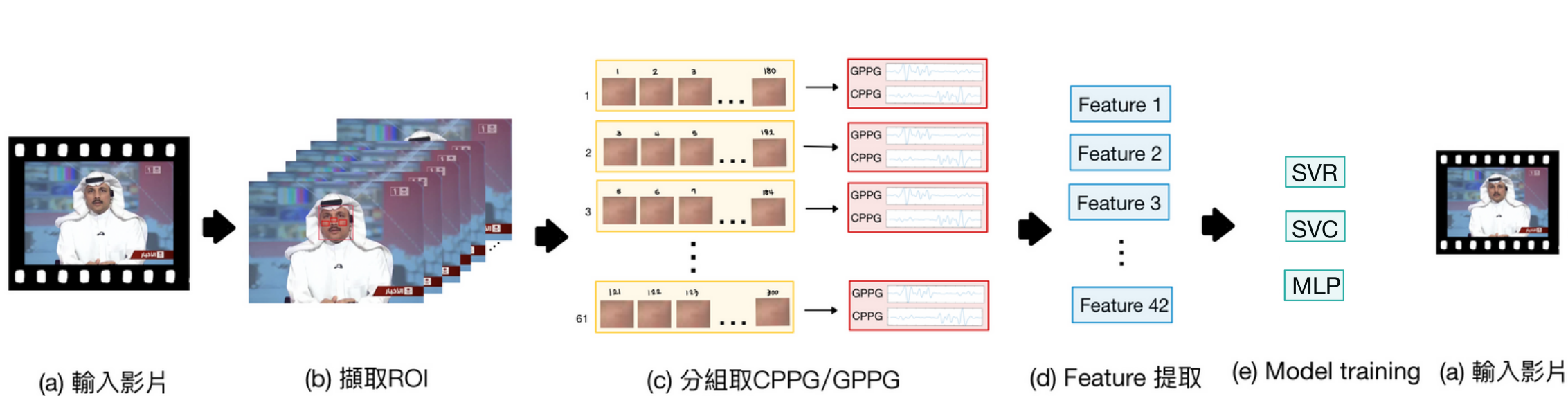
原理簡介



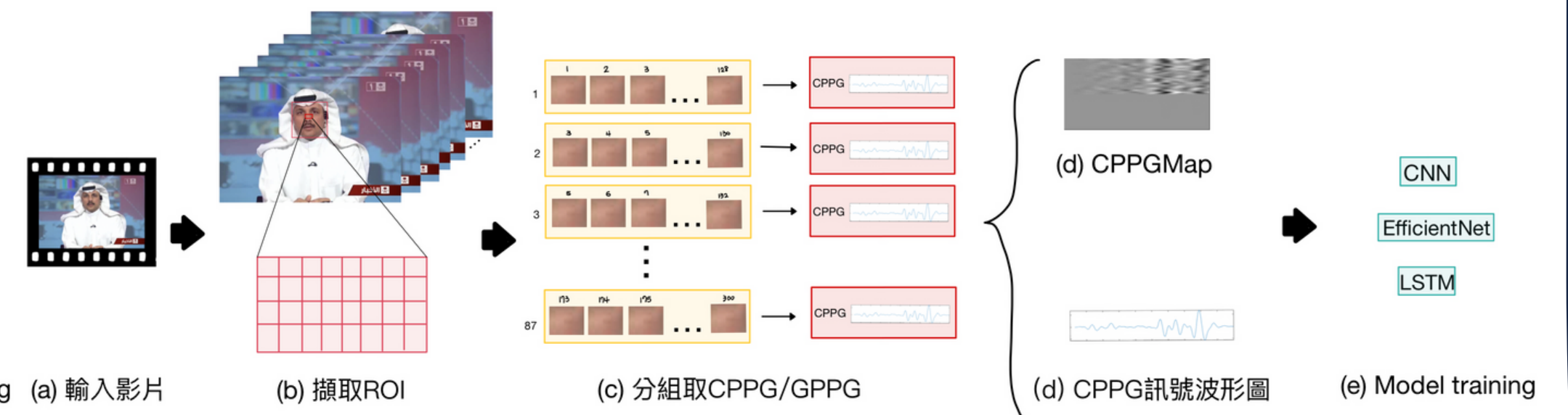
- **RPPG(Remote Photoplethysmography)**
用反射光線測量皮膚亮度變化
- **CPPG**
HVC中彩度channel的RPPG
- **GPPG**
RGB中綠色channel的RPPG

系統架構

數學特徵：



圖形特徵：



成果展示

Model	資料集	特徵	Accuracy	Precision	Recall	F1
MLPRegression	DFDC	1,7,8	0.62	0.6	0.67	1.75
MLPRegression	DFDC	1,7,8	0.59	0.58	0.58	1.65
SVC	Celeb	15,20,19,28,24,4,3	0.75	0.81	0.73	0.74
SVC	FF	15,20,19,28,24,4,3	0.72	0.71	0.75	0.73
SVC	FF	15,20,18,19,28,24,4,3	0.73	0.72	0.75	0.73
SVC	Celeb	15,20,19,28,24,4,3,30	0.76	0.82	0.72	0.75
SVC	Celeb	15,20,19,28,24,4,3,30	0.76	0.82	0.73	0.75
SVC	Celeb	15,20,19,28,24,4,3,30,26	0.76	0.8	0.76	0.76

數學特徵的部分成果



Model	Dropout	資料集	Input	Accuracy
EfficientNetV2B0	0.2	Celeb	Cppgmap	50%
EfficientNetV2S	0.2	Celeb	Cppgmap	59%
EfficientNetV2S	0.5	Celeb	Cppgmap	50%
EfficientNetV2M	0.3	Celeb	Cppg波形圖	73%
EfficientNetV2L	0.3	Celeb	Cppg波形圖	62%
EfficientNetV2L	0.3	FF	Cppg波形圖	49%
CNN		Celeb	Cppg波形圖	70%
LSTM		Celeb	Cppg波形圖	65%

圖形特徵的部分成果

結語與展望

這次的專題研究中，我們嘗試了許多方法，包含數學特徵部分的三種模型以及多種feature組合和圖像特徵部分的三種模型架構以及兩種輸入。最終兩邊最好的結果分別是SVM的76%以及EfficientNet的73%。

未來我們期盼能嘗試看看Vision Transformer (ViT)，以及研究更多feature的挑選方法。讓我們能夠開發出準確率更高deepfake偵測軟體給民眾使用，使deepfake的鑑定能擴展到民間，更程度的避免錯誤資訊的傳播。