目錄

壹、簡歷

基本資料

生日：2000/05/06

手機：0972387353

地址：臺中市大安區中山南路 79 巷 5 號

Email：as26872076@gmail.com

畢業學校：

- 中山醫學大學醫學資訊學系
- 臺中市立臺中女子高級中等學校
- 臺中市立順天國中
- 文昌國民小學

謝雅竹

Ya-Chu Hsieh

自傳

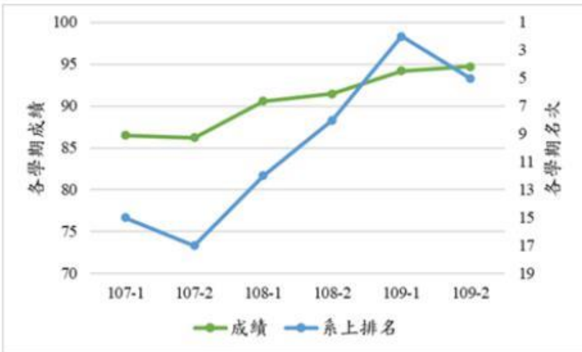
- 在校成績為系上第 8 名，於大二上學期加入「健康照護實驗室」，在指導教授的帶領下參加許多競賽、論文發表、學術研討會及課外專業培訓課程
- 除了修習系上課程，也與校內外單位合作進行系統設計、開發，提升程式撰寫能力
- 擔任「程式語言」高等教育與深耕計畫之課程助教，學習溝通和解決程式問題的能力
- 利用課餘時間開發「**基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類輔助診斷系統**」，榮獲 110 年科技部**大專學生研究計畫**補助，計畫編號：110-2813-C-040-096-E

學習成績與排名

中山醫學大學醫學資訊學系 107 年度至 109 年度下學期

學業成績平均：**90.39**，累積排名：**第 8 名**

學期		成績	系上排名
大一	上學期	86.5	15
	下學期	86.25	17
大二	上學期	90.56	12
	下學期	91.48	8
大三	上學期	94.21	2
	下學期	94.68	5



校外實習

實習時間	實習單位	實習專案	技術
2021/03~2021/08	工業技術研究院 服科中心	結合穿戴式感測裝置 進行腿部健康預測	Web Bluetooth API、Web Storage API、 HTML5、JavaScript、BLE 5.0

校內外競賽紀錄

主辦單位	競賽名稱及組別	名次(比率)
中華民國模糊學會	「2020 International Conference on Fuzzy Theory and Its Applications」 **Best Conference Paper Award**	Second Place
教育部、經濟部	「2020 第 25 屆大專校院資訊應用服務創新競賽」 **資訊應用組八(IP8-04)**	第一名 (7.69%)
	「2020 第 25 屆大專校院資訊應用服務創新競賽」 **Open Data創意應用開發組(Open Data-13)**	第一名 (7.69%)
	「2020 第 25 屆大專校院資訊應用服務創新競賽」 **資訊應用組六(IP6-09)**	第三名 (23.07%)
	「2020 第 25 屆大專校院資訊應用服務創新競賽」 **教育開放資料組(EDUOD-06)**	佳作 (30.76%)
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 **資訊應用組七(IP7-04)**	第三名 (23.07%)
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 **資訊應用組六(IP6-02)**	第三名 (23.07%)
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 **大桃園資料競技場創新應用組(TY Data Arena-07)**	佳作 (30.76%)
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 **資訊應用組十(IP10-12)**	人氣獎 (7.6%)
教育部	「第二屆智慧專題聯網專題實作競賽」 **智慧醫療或智慧農業應用**	金牌 (6.67%)
	「第二屆智慧專題聯網專題實作競賽」 **智慧醫療或智慧農業應用**	銀牌 (13.33%)
	「2020全國大專校院智慧創新暨跨域整合創作競賽」 **物聯網組**	跨域整合 特別獎(4.54%)
高雄科技大學	「2021第十七屆全國電子設計創意競賽暨學術研討會」 **學術研討組-通訊領域**	優秀論文獎
	「2021第十七屆全國電子設計創意競賽暨學術研討會」 **學術研討組-醫電領域**	最佳論文獎
	「2020 第十六屆全國電子設計創意競賽」 **大專資通類**	第三名 (6.06%)
	「2020 第十六屆全國電子設計創意競賽」 **大專電子類**	佳作 (5.79%)
中山醫學大學	「2019中山醫學大學全國大專校院創新、創意及創業競賽」 **創新科技組**	第三名 (28.57%)
	「2019中山醫學大學全國大專校院創新、創意及創業競賽」 **創新科技組**	佳作 (57.14%)

校外課程研習

時間	課程名稱
2020/08/01~2020/08/29 每週六	**臺北市生物產業協會—** 2020「台灣發展智慧醫療之新創趨勢與實務應用」系列論壇
2021/03/26	**臺大醫院智慧醫療中心—** 如何準備智慧醫療研究資料?從石油到汽油
2021/05/21	**科技部人工智慧生技醫療創新研究中心—** AI醫療資訊攻防戰-資安理論與實務探討

學術發表

時間	發表文章
2021	I Miao Chen, Pin Yu Yeh, Ting Chi Chang, **Ya Chu Hsieh**, Chiun Li Chin*, "Based on LCNet Model and Multi-Feature to Evaluate Gait Data," 2021. (Submitted to ***Measurement***, state: **Under Review**)
2021	Hao-Hung Tsai, Chia-Shin Wei, **Ya-Chu Hsieh**, I-Miao Chen, Pin-Yu Yeh, Darren Shih, Chiun-Li Chin*, "Calcification clusters and lesions analysis in mammogram using multi-architecture deep learning algorithms," *Biomedical Engineering: Applications, Basis and Communications* (***BME***), 2021. (**EI Compendex**)
2021	Jung-Mao Lu, Chiun-Li Chin, I-Miao Chen, **Ya-Chu Hsieh**, Ting-Chi Chang, Pin-Yu Yeh*, "GAN-based SSD Segmentation Algorithm to Assist the Character Recognition of Seven-Segment Display Digits," *Chun Shan Medical Journal* (***CSMJ***), 2021.
2021	秦群立、陳怡妙、葉品郁、**謝雅竹**、張婷淇、賴彥名、丁敬訓，結合空拍機與TensorRT Pose之訓練報告能力系統，*第十七屆全國電子設計創意競賽暨學術研討會*，2021。
2021	秦群立、陳怡妙、葉品郁、**謝雅竹**、張婷淇、陳婷、李杰祐，基於DL-RMS多重架構驗證群聚鈣化點之特徵抽取、篩選及分類方法，*第十七屆全國電子設計創意競賽暨學術研討會*，2021。
2020	**Ya-Chu Hsieh**, Chiun-Li Chin, Chia-Shin Wei, I-Miao Chen, Pin-Yu Yeh, Ru-Jiun Tseng, "Combining VGG16, Mask R-CNN and Inception V3 to identify the benign and malignant of breast microcalcification clusters," *2020 International Conference on Fuzzy Theory and Its Applications* (***iFUZZY 2020***), 2020.
2020	Chiun-Li Chin, Ming-Chen Hsu, I-Miao Chen, Pin-Yu Yeh, Ting-Ya Chang, **Ya-Chu Hsieh**, "Based on Mask R-CNN Tooth Position Labeling And Periodontal Disease Identification," *2020 International Conference on Fuzzy Theory and Its Applications* (***iFUZZY 2020***), 2020.
2020	Chiun-Li Chin, Ting-Ya Chang, Pin-Yu Yeh, **Ya-Chu Hsieh**, I-Miao Chen, "The analysis of high-risk group of angina recurrence using smart health-box and deep learning algorithm," *2020 International Automatic Control Conference* (***CACS 2020***), 2020.

實務經驗

程式語言能力	實作案例
程式語言程式設計： Python 深度學習框架： Tensorflow、Keras、Pytorch	• 基於深度學習辨識血糖及血壓機之七段顯示器系統 • 結合TensorRT Pose之訓練報告能力系統 • 結合BeautyGAN之妝容模擬系統
嵌入式硬體與行動軟體設計實作： Arduino-C、Android、 JAVA、SQLite 嵌入式硬體： Arduino、Raspberry Pi、 Arduino Jetson Nano、 Arduino Jetson TX2、Arduino Jetson Xavier NX	• 「得心應手」心絞痛預測及偵測系統 • 「聲之形」聽障輔助行車安全系統 • 「一鏡到底」穿搭建議系統 • 「明鏡照形」妝容模擬及化妝輔助系統
網際網路程式設計： PHP、MySQL、 HTML、JavaScript、CSS	• 「孩好有您」年長者照護系統 • 「We can help」聽障輔助學習及溝通系統 • 中華民國模糊學會--2020 International Conference on Fuzzy Theory and Its Applications暨中華民國第二十八屆模糊理論及其應用研討會網站

論文閱讀經驗

讀過/報告過/聽過的SCI 論文	期刊名稱	Impact Factor(2019)	Rank(2019)
Deep Learning for Sensor-based Activity Recognition: A Survey	*Pattern Recognition Letters*	3.756	Q2
Visual-audio emotion recognition based on multi-task and ensemble learning with multiple features	*Neurocomputing*	5.719	Q1
Automatic segmentation of tumors and affected organs in the abdomen using a 3D hybrid model for computed tomography imaging	*Computers in Biology and Medicine*	4.589	Q2
Deep Learning-Based Gait Recognition Using Smartphones in the Wild	*IEEE Transactions on Information Forensics and Security*	6.013	Q1

研究規劃

貳、自傳

一、家庭背景

出生於南投市，隨後因父母工作因素舉家遷徙到臺中市大甲區，家中成員共四位，除了我以外還包含父親、母親以及弟弟，由於父母皆從事服務業，工作十分忙碌，無暇顧及我和弟弟，因此在遇到困難時會憑自身經驗或另尋方法解決，從小就培養了獨立自主的能力。在學習方面，父母從不要求成績，並對於我感興趣的才藝也都給予支持，因此培養許多才藝，包含彈吉他、烏克麗麗等，在壓力來臨時，也能透過這些興趣來陶冶性情以及抒發壓力。

二、人格特質

待人處事謙遜有禮，不吝於向人請教。大學期間曾擔任藍韻口琴社幹部中的總務一職，任職期間做事有條理且認真負責，對於社團的每筆債務瞭如指掌，避免了多餘的開銷導致入不敷出，在開會時也勇於發表自己的看法，積極與幹部溝通與討論，尋求完成事情的最佳解。

三、求學歷程

1. 國小

就讀國小的期間，利用課餘的時間參加管樂社，獲得了台中市全國學生音樂比賽團體組的優等，且在參加社團的同時也不忘兼顧課業成績，在國小畢業時獲得楷模獎的殊榮。

2. 國中

國中時期，除了持續參加管樂社並定期代表學校比賽以外，因得老師及同學的信任而任職副班長一職，擔任老師與同學之間的溝通橋樑。在學習上，認真勤勉，曾獲得全學年第一名，並利用課餘時間參加英文補習班以增進外文能力並通過全民英檢初級。

3. 高中

通過國中的努力進到了臺中女生的第一學府後，體驗到開放且自由的學習風氣，在學期間擔任國文科的小老師，輔助老師處理成績以及作業等事務，並將消息公告給同學知悉。在高中的資訊課開始接觸到一些基礎的程式語言，從此對於資訊相當感興趣，當時資訊課的成績還拿了班上第一。

4. 大學

大學期間擁有豐富社團經歷，在兼顧學業的同時積極學習以及研究，透過參與論文發表、學術研討會及資訊競賽等，訓練口頭發表及臨場反應，並讓所學以及研究成果接受外界檢驗，吸收更多元的看法，最後皆取得相當不錯的成果。此外，在學期間與校內外的各單位合作開發系統的經歷，使我累積豐富的實務經驗。

1. 精進專業能力，同時兼顧學業

- ✓ 大一下學期加入「**健康照護實驗室**」，平時同實驗室成員互相交流學習，在教授的推薦下同實驗室成員與校內外多個單位合作開發系統，實際學得系統分析與設計的經驗。

- ✓ 參與多項資訊競賽、學術研討會及發表論文，將實際所學運用於實戰，從中獲得反饋並自省以精進自身能力。
- ✓ 系上累計排名為第八名，大一到大三學業平均為 90.39 分，班上人數共 41 位。
- ✓ 跨領域與中山醫學大學附設醫院醫學影像部醫師合作並經常同指導教授討論，完成基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類輔助診斷系統之研究，此研究也獲得了 110 年科技部大專生研究計畫的殊榮。

2. 多方參與系上及社團活動代表系上與友校進行學術交流

- ✓ 大一時參與系上舉辦至日本筑波大學參訪的**學術交流活動**。
- ✓ 參加管樂社及口琴社並代表學校參加**全國音樂比賽**，均獲得優等的佳績。
- ✓ 大二時擔任**宿營**的**領隊**，時常與幹部開會討論活動流程、熟稔活動時間及事項，也讓宿營活動順利且流暢地舉辦完成。

3. 擔任助教輔導學生，同時提升自身能力

- ✓ 在大二及大三上學期分別擔任不同老師的程式語言助教，除了能在輔導學弟妹學習程式語言的同時提升自身的程式能力、應對進退能力以及表達能力外，還能作為老師及學生的溝通媒介，幫助老師理解學生的學習需求。

四、未來展望

1. 短期目標

短期我將專注於完成大學學業，除了會繼續論文的投稿外，也會持續維護與校內外單位合作開發的系統，學習與非資訊人員溝通的技巧。在進入研究所後，也會積極投入教授的計畫中，並利用課餘時間精進英文能力。

2. 中期目標

除了持續加強英文能力外，在就讀研究所的期間認真修課之餘，積極參與指導教授計畫，從中獲得實戰經驗及溝通技巧，並透過實習的方式，提前適應職場生活。

3. 長期目標

期望能進入高科技產業工作，在工作穩定後回報家人，並帶他們出國遊玩。另外，也希望自己在出社會後，也能維持積極學習的態度，持續養精蓄銳，保有學習的熱誠以學習更多新知。

參、就學計畫

一、就讀動機

大學時期為了增加學習及實作的經驗，多次參與資訊相關競賽，作品受到不少評審的青睞，獲得了不錯的獎項，如教育部及經濟部主辦的「2020 第 25 屆大專校院資訊應用服務創新競賽」(**第一名，7.69%**)、教育部主辦的「第二屆智慧專題聯網專題實作競賽」(**金牌，6.67%**)、教育部主辦的「2020 全國大專校院智慧創新暨跨域整合創作競賽」(**跨領域整合特別獎**)及高雄科技大學主辦的「2020 第十六屆全國電子設計創意競賽」(第二名，6.06%)等，也參加過國際研討會，除了論文被接受外，也受邀參加口頭發表，獲得了 Best Conference Paper Award 的 **Second Place** 的榮耀，還在大學就學期間發表多篇論文，其中包含一篇 **EI journal paper**，一篇與醫學相關的**期刊論文**和一篇 **SCI paper** (Under Review)。

另外，為了增加實務應用的經驗，在工研院的服務系統科技中心實習半年，因此我認為將來在研究所，我有足夠的能力完成教授交代的工作，此外，大學時我還修習完資訊重要的基礎科目，包含**線性代數(70 分)**、**微積分(92 分)**、**計算機組織與結構(81 分)**、**機率與統計(79 分)**、**資料結構(88 分)**，**作業系統(84 分)**，同時我也規劃在研究所時會修 XXX 等課程，以此提升我的專業知識與能力。

二、入學前計畫

若有幸錄取，定會妥善規劃時間，持續增強自身能力，不虛度甄試錄取至研究所開學前的九個月時間。表一為研究所入學前的詳細計畫內容。

表一、入學前計畫表

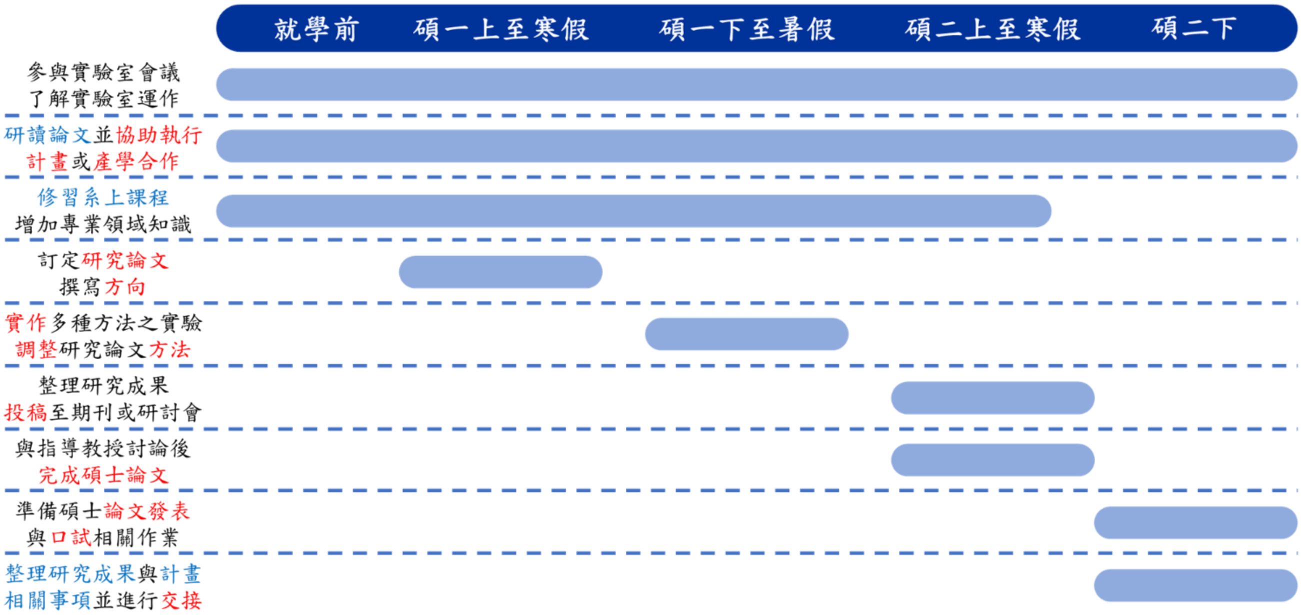
項目	內容及目標
程式能力	精進各類程式語言能力，如C、Python、Java等，並持續維護合作計畫系統，以報考大學程式能力檢定 CPE 檢視能力提升的程度。
英語能力	無論是在何種階段，英文皆是重要且應具備的能力，因此在就學前會積極加強英語的聽說讀寫能力，並以參與多益TOEIC來檢驗英文能力，了解不足之處。
研究能力	透過閱讀SCI paper了解近期研究發展，並多方參考及蒐集與研究所領域相關之論文，以為未來做準備。
影像處理	閱讀並實作影像處理書籍範例、參加各校影像處理相關課程。
人工智慧	透過閱讀書籍、瀏覽網路影片及參加校外課程等管道學習深度學習領域知識。

三、研究所兩年規劃

研究規劃

我將就讀研究所的兩年分為五個時間段，分別為就學前、碩一上至寒假、碩一下至暑假、碩二上至寒假、碩二下，並在各個階段訂下目標，如表二所示，藉由此方式鼓勵自己按部就班的朝目標前進，提升自我價值。

表二、研究所兩年學習計畫甘特圖



修課規劃

現今資訊日新月異，擁有一顆靈活的頭腦可以應付資訊快速變動的情況，具備專業知識更成為自身在業界立足的基石之一，因此在了解研究所課程後，我擬定了欲修習的科目，如表三所示，期望透過修習這些課程，提升自身的競爭力及專業知識技能，以面對未來的挑戰。

表三、欲修習科目

年級	欲修科目
研一上學期	專題研究、專題討論、類神經控制系統、訊號控制及對策、數位影像處理
研一下學期	專題研究、專題討論、深度學習於電腦視覺、生醫電子電路設計
研二上學期	專題研究、專題討論、機器學習中的數學原理
研二下學期	專題研究、專題討論、碩士論文

四、自我充實

就讀研究所的期間，除了致力於修課及研究以外，在平時也會安排一些課餘活動充實自我，提高競爭力，自我充實之計畫表如表四所示。

表四、個人能力充實表

項目	內容
與領域相關之實習	找尋與研究所研究領域相關之實習，提早熟悉職場氛圍，同時了解當時業界需求，以此精進自身能力，相信日後在求職時能夠更如魚得水。
英語能力精進	維持英語學習的習慣，除了常閱讀英文論文及雜誌外，定期安排英文檢定考試檢視一段時間的學習成果。
運動健身	提升自身價值及能力必須建立在身心靈狀況良好的基礎上才能達成，因此養成良好的運動習慣來維持健康及體力是不可或缺的。
學術交流	就讀研究所期間積極參與國內外學術研討會，與他人交流以增廣見聞，並虛心接受其他學者的建議，提升研究能力。

五、未來規劃

1. 從事資訊領域方面的工作，在有所成時能至各校演講，分享個人經驗
2. 若有機會期望能出國工作，體驗國外資訊領域職場氛圍，並學習其優點
3. 事業有成時能有更多時間陪伴家人，安排多些娛樂活動或是待在家一同看電影

肆、研究計畫

1. 研究計畫主題

Early breast cancer prediction using Multi-View feature and SCR-CNN

2. 動機與目的

乳房 X 光攝影檢查是現今篩檢早期乳癌的主要方式，並且 WHO 證實此檢測方法能有效偵測出無症狀的早期乳癌，不僅可以實現預測早期乳癌之目標，還能降低乳癌死亡率。目前在臨床方面大多是透過計算機輔助(CAD)系統輔助放射科醫師快速進行乳房攝影影像的判讀以此減少放射科醫師的工作負擔[1]，然而 CAD 系統大多屬於專家系統，多數僅能處理特定或是具有標準答案的問題，並且需仰賴領域專家制定系統的判斷標準，造成難以正確檢測乳房攝影影像中各種不同形狀、大小的鈣化點特徵，導致 CAD 系統在進行鈣化點良惡分類的準確率上仍有很大的提升空間。

另外，雖然乳房 X 光攝影檢查是現今最普及的乳癌檢測方法，但對於乳房緻密度高的婦女而言，乳房攝影影像中的乳腺組織與乳房鈣化點群聚處的特徵相近，在乳房攝影影像上皆呈現白色，導致醫師不易區分其差異，如圖 1 所示。而且在**乳房 X 光攝影檢查**中，每位受檢者**共會拍攝 4 張影像**，其中**各包含 2 張 CC view** 與 **MLO view**，如圖 2 所示，醫師需分別透過肉眼判讀不同方向拍攝的乳房攝影影像，方能確定受檢者是否罹患早期乳癌。若是以人工方式同時判讀多張乳房攝影影像時，將容易受到醫師的經驗或疲勞程度而影響診斷早期乳癌的標準，使得醫師難以準確分辨出惡性群聚鈣化點，導致錯失治療的黃金時機。

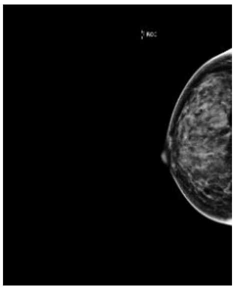


圖 1、緻密型乳房

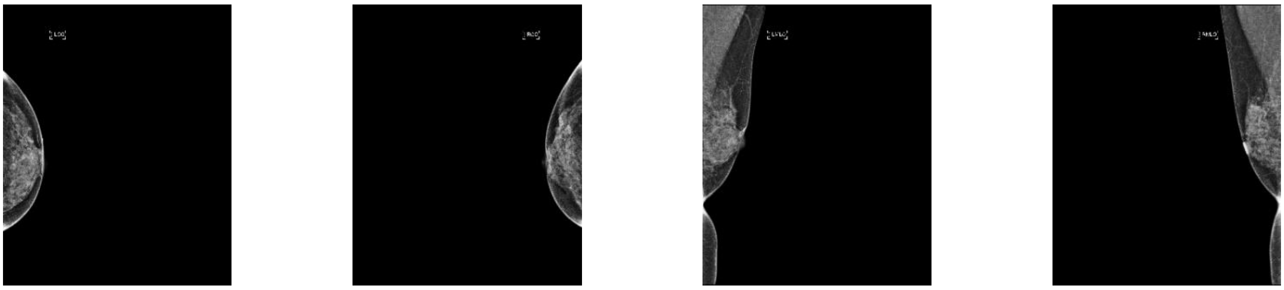


圖 2、每位受測者的 CC view 與 MLO view 影像，
由左至右為 LCC、RCC、LMLO、RMLO

近年來深度學習快速發展，已有研究將深度學習應用於醫療領域[2]，像是利用 CNN 模型獲取皮膚影像的特徵，以分類出良性與惡性的皮膚病變[3]，證明深度學習能有效輔助臨床醫療診斷。因此本研究為解決上述問題，採用 **SCR-CNN** (Siamese Cascade R-CNN)，**結合 Siamese Neural Network** [4]與 **Cascade R-CNN** [5]深度學習演算法，並將 CC view 與 MLO view 同時輸入至 SCR-CNN，自動偵測並切割出乳房攝影影像內鈣化點群聚處的微鈣化點，降低乳腺組織及複雜背景等雜訊對於判斷乳房鈣化點群聚處的影響，最後再將 **CC view** 與 **MLO view** 中相似的**特徵結合**，以**實現精準切割及辨識良性與惡性群聚鈣化點之目的**。本研究期望能縮短找尋鈣化點群聚處再切割單顆鈣化點的時間，解決像素點重複計算的問題，如圖 3 所示，還能直接透過鈣化點的形狀等特徵，預測受檢者是否可能罹患早期乳癌，加快診斷流程及節省等候報告結果的時間。

圖 3、像素點重複計算之比較圖，
可觀察到左圖像素點重複計算的狀況較右圖嚴重

3. 預期方法

為了減少乳房攝影影像中的乳腺組織或複雜背景干擾，同時還能學習 CC view 與 MLO view 之間的相似特徵，達成準確找出乳房攝影影像中的微鈣化點並分辨其良惡之目標，本研究所提出的 **SCR-CNN** 深度學習演算法會**透過兩個 Cascade R-CNN** 深度學習模型分別**分離出 CC view 和 MLO view** 兩種不同視圖中重要的**乳房攝影影像特徵**，**再透過 Siamese Neural Network 組合兩種不同視圖的特徵**，最後找出鈣化點群聚處中的微鈣化點，並分析其良惡，SCR-CNN 架構圖如圖 4 所示。

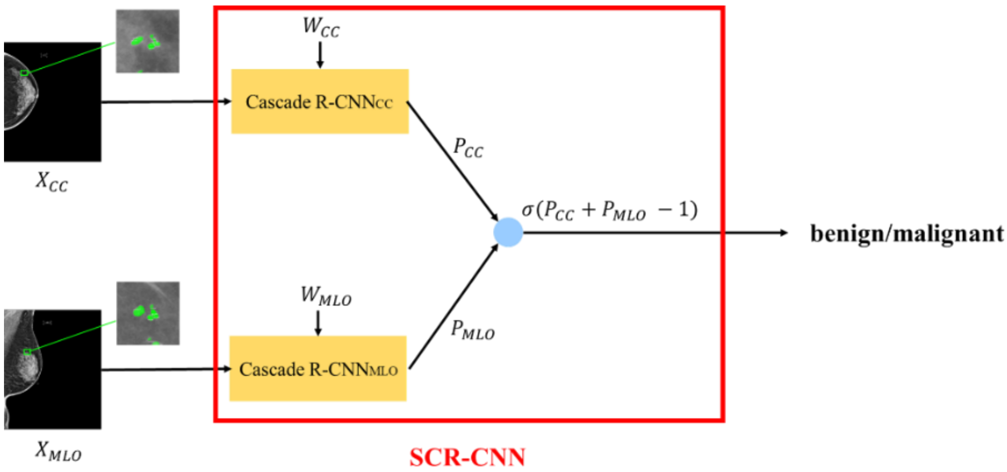


圖 4、SCR-CNN 架構圖

接下來將介紹資料蒐集、SCR-CNN 深度學習演算法中所使用到的視圖組合方法以及 Cascade R-CNN 深度學習模型。

A. 資料搜集

由於乳癌為國人婦女癌症發生的第一位，因此國民健康署對與此議題一直非常關注並且補助 65 歲以上、未滿 70 歲的婦女每 2 年接受 1 次乳房 X 光攝影檢查，每年均有不少女性接受此檢測並拍攝乳房攝影影像，另外，本研究實際詢問放射科醫師每年所拍攝的乳房攝影影像張數，得知某家中部醫學中心每年約有 40 人拍攝乳房攝影影像，且每人會拍攝共 4 張影像，包含左、右各 2 張的 CC View 及 MLO View 影像，因此本研究預期和醫院放射科醫師合作，除了獲取近 10 年，約 1600 張乳房攝影影像外，還能與放射科醫師討論並向其學習專業領域知識，使我們的研究更具可信度。

B. 視圖組合方法

CC view 和 MLO view 的視圖組合方法會透過 Siamese Neural Network 來完成，其前向傳播公式如下：

$$\mathrm{S}(W_{CC}, W_{MLO}) = \sum_{t=1}^{n} log\sigma((2Z_t - 1)(P_{t,CC} + P_{t,MLO} - 1)) \tag{1}$$

其中W_{CC}和W_{MLO}分別為 Cascade R-CNN$_{\mathrm{CC}}$ 深度學習網路及 Cascade R-CNN$_{\mathrm{MLO}}$ 深度學習網路的參數，σ代表 sigmoid 激發函數，Z_t表示每個 pixel 良惡的 label，數值為二進制，不是 1 就是 0，$P_{t,CC}$代表 Cascade R-CNN$_{\mathrm{CC}}$ 深度學習網路的輸出，$P_{t,MLO}$則代表 Cascade R-CNN$_{\mathrm{MLO}}$ 深度學習網路的輸出。而兩種視圖的分類誤差關係如下：

$$\frac{\partial S}{\partial W_{CC}} = \sum_{t=1}^{n}\left(Z_t - \hat{\mathrm{Z}}_t\right)\frac{\partial P_{t,CC}}{\partial W_{CC}}，\frac{\partial S}{\partial W_{MLO}} = \sum_{t=1}^{n}\left(Z_t - \hat{\mathrm{Z}}_t\right)\frac{\partial P_{t,MLO}}{\partial W_{MLO}} \tag{2}$$

其中$\hat{\mathrm{Z}}_t = \mathrm{P}(Z_t | X_{t,CC}, X_{t,MLO}; W_{CC}, W_{MLO})$。

C. Cascade R-CNN 深度學習模型

在訓練 Cascade R-CNN 深度學習模型時，會將輸入的乳房攝影影像通過卷積層進行特徵抽取，接著將結果輸入至池化層篩選特徵，獲得特徵圖，再通過 H 進行檢測與分類，得到目標區域的 Bounding Box 及類別，Cascade R-CNN 的架構圖如圖 5 所示，而 Cascade R-CNN 所採用的級聯(cascade)訓練方式，除了能讓每次網路輸出的 Bounding Box 在輸入池化層後得到**最佳化的 IoU 閥值**，還能保證在樣本數不減少的情況下訓練出高品質的偵測模型。

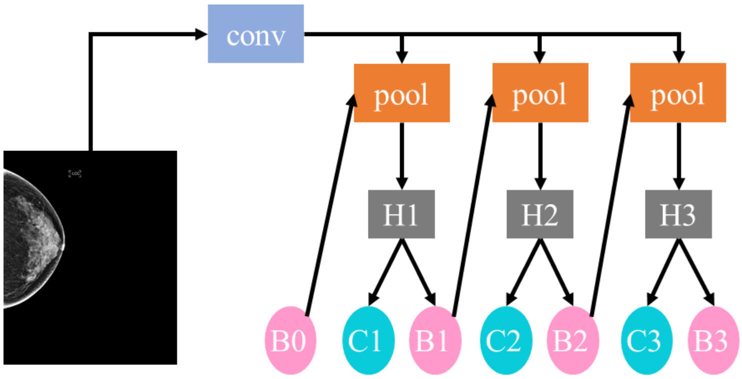


圖 5、Cascade R-CNN 之架構圖

4. 參考文獻

[1] H. Harvey, E. Karpati, G. Khara et al., "The Role of Deep Learning in Breast Screening," Current Breast Cancer Reports, vol. 11, pp. 17-12, 2019.

[2] J. Wang, X. Yang, H. Cai et al., "Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning," Scientific Reports, vol. 6, 2016.

[3] A. Esteva, K. Chou, S. Yeung et al., "Deep learning-enabled medical computer vision," npj Digital Medicine, vol. 4, 2021.

[4] A. J. Bekker, H. Greenspan, J. Goldberger, "A multi-view deep learning architecture for classification of breast microcalcifications," 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 726-730, 2016.

[5] Z. Cai, N. Vasconcelos, "Cascade R-CNN: Delving into High Quality Object Detection," The IEEE Conference on Computer Vision and Pattern Recognition, pp. 6154-6162, 2018.

1. 研究計畫主題

Using multi-architecture deep learning algorithm for depth perception of medical images and videos

2. 動機與目的

醫學影像是醫師進行病症診斷的重要依據，在手術前醫師一般會透過判讀 X 光攝影、電腦斷層或磁振造影等醫學影像進行術前評估，以制定出手術方法與治療方針，同時透過影像得知實際手術範圍與位置，進行手術模擬與規劃。另外，在進行手術時，醫師會使用手術顯微鏡來放大手術位置，讓人體血管與組織結構變得更清晰，以方便進行手術治療。但是不同手術顯微鏡的產生的影像解析度不同，且會受到手術顯微鏡放置角度的影響，所以當手術顯微鏡所拍攝的 2D 影像**無法**讓醫生確切**瞭解影像深度**時，可能會造成**誤判目標位置**[1]，進而導致手術事故的發生。由此可見，降低醫師**誤判空間深度**的機率，是目前臨床手術上遇到的重要挑戰之一。

深度感知(depth perception)是現今醫療領域十分重視的研究方向[2]，並且已有許多研究透過深度學習方法進行深度估計[3]。本研究將 2D 影像透過 **ESRGAN** 模型自動**提高影像**之**解析度**，讓手術顯微鏡影像中的血管與組織結構能清晰易見，再利用 **DepthNet** 模型將 **2D 影像轉為**具有深度的**深度圖**，以獲得立體視覺，讓醫師在手術臺上能進行精密的血管與神經縫合，大幅提升手術精確度，同時也提升手術醫療品質。

3. 預期方法

為了讓醫生能夠在**手術過程**中，可以**同時查看醫學影像**之**深度圖**，並且為了提高醫學影像生成深度圖的準確率，本研究利用**增強式超解析度生成對抗網路(ESRGAN)**[4]**提升**醫學影像的**解析度**，接著使用 **DepthNet**[5]深度學習模型，用以**生成**醫學影像之**深度圖**，以輔助醫生精準判斷空間深度，使手術過程更加順利且流暢，本研究之方法流程圖如圖 1 所示。

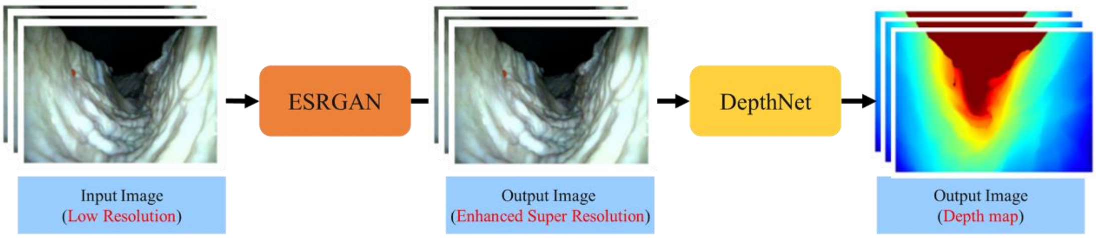


圖 1、方法流程圖

A. 資料蒐集

本研究會與某中部醫院的合作，透過醫院提供的心血管影像及手術顯微鏡拍攝的序列影像作為本研究的訓練及測試資料集，分別占整體的 80%及 20%。另外，現今多數關於深度圖的研究會將 KITTI 資料集作為模型訓練的資料集，而 KITTI 資料集中由 389 對立體影像和光流圖所組成，且以 10Hz 取樣頻率擷取出的影像皆屬於街景影像，能清楚評估模型準確率，因此本研究會先透過 KITTI 資料集進行 DepthNet 模型的預訓練，以確保模型具有一定的辨識能力。

B. Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN)

本研究藉由訓練 ESRGAN 中的 Generator Network 及 Discriminator Network 進行對抗，以生成增強式的高解析度醫學影像，相較於一般 GAN 學習去判斷影像是或假，ESRGAN 學習判斷 Generator Network 生成影像 和 Ground Truth 誰更為接近真實的影像，此架構有助於 Generator Network 恢復更真實的紋理細節，ESRGAN 架構圖如圖 2 所示。

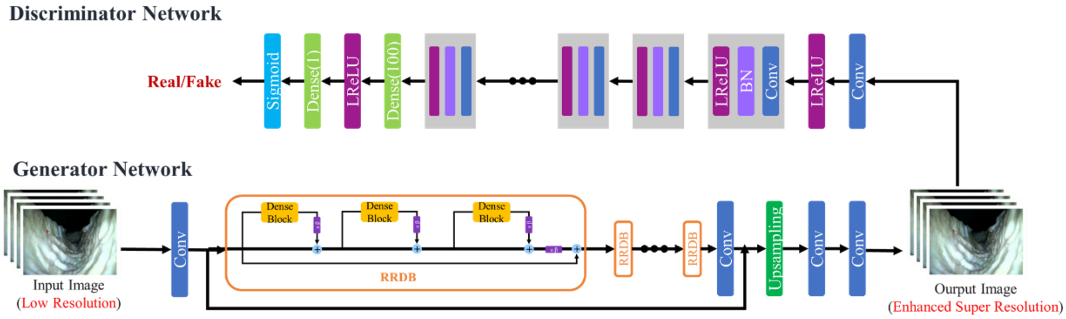


圖 2、ESRGAN 架構圖[4]

(1) Generator Network

Generator Network 用於生成影像，並在訓練的過程中根據 Discriminator Network 的回饋進行調整並再次生成影像，藉由此方式讓生成影像接近於目標影像。在 Generator Network 中，利用 Convolution layer 抽取影像特徵，還加入了密集殘差網路(RRDB)作為網路的基本建構單元，使網路具有大容量，更易於訓練。

(2) Discriminator Network

Discriminator Network 用於評估 Generator Network 所生成的影像特徵，在 Discriminator Network 的最後一層使用 Sigmiod 作為激發函數，以輸出增強式高解析度生成影像的機率值，作為 Generator Network 下一次生成影像前的回饋值。

(3) Loss Functions

這裡所採用的 Discriminator Network 是 Relativistic average Discriminator (RaD)，用以嘗試估計真實影像相對來說比假影像更逼真的機率，其表示為：

$$D_{R_a}(x_r, x_f) = \sigma(C(x_r)) - E_{x_f}[C(x_f)] \quad (1)$$

其中σ是 Sigmoid 激發函數，$C(x_r)$是 Discriminator Network 的輸出，$E_{x_f}[\cdot]$表示取 batchsize 中所有假資料的平均值。

i. Discriminator Network

在 ESRGAN 中 Discriminator Network 的 Loss Function 定義如公式(2)所示。

$$L_D = L_D^{R_a} = -E_{x_r}\left[\log\left(D_{R_a}(x_r, x_f)\right)\right] - E_{x_f}[\log\left(1 - D_{R_a}(x_f, x_r)\right)] \quad (2)$$

ii. Generator Network

ESRGAN 中 Generator Network 的 Loss Function 定義為：

$$L_G^{Ra} = -E_{x_r}\left[\log\left(1 - D_{R_a}\left(x_r, x_f\right)\right)\right] - E_{x_f}\left[\log\left(D_{R_a}\left(x_f, x_r\right)\right)\right] \tag{3}$$

其中x_f代表輸入的低解析度影像，x_r代表 Generator Network 生成的增強式高解析度生成影像。

iii. Total Loss

最後 Generator Network 的 Total Loss 為：

$$L_G = L_{percep} + \lambda L_G^{Ra} + \eta L_1 \tag{4}$$

其中L_{percep}為感知損失函數，在此使用 L1 Loss，而L_G^{Ra}為生成器的損失，$\lambda = 5\times 10^{-3}$，$\eta = 0.01$，而$L_1$是 Content loss，其公式為$L_1 = E_{x_i}\|G(x_i) - y\|_1$。

D. DepthNet

DepthNet 深度學習模型包含 8 層的 ConvLSTM layer、5 層的 Deconvolution layer 及 5 層的 Convolution layer，類似於 U-Net[6]是一個 Encoder-Decoder 的架構，Encoding 階段的目的主要學習 N 個序列影像間的時空關係以預測深度圖，會將經過 ESRGAN 所獲得的增強式高解析度序列醫學影像輸入至 **ConvLSTM layer 獲取序列醫學影像**之間的**時空關係**與特徵，其中每層 ConvLSTM layer 包含 N 個狀態；而 Decoding 階段的目的是學習重建 N 個單獨的深度圖，由 Deconvolution layer 及 Convolution layer 交替組成，並將 Encoding 階段所產生的特徵圖與 Decoding 階段進行拼接，最後使用一層 1×1 的 Convolution layer 將張量轉換為一個深度圖，DepthNet 架構圖如圖 3 所示。

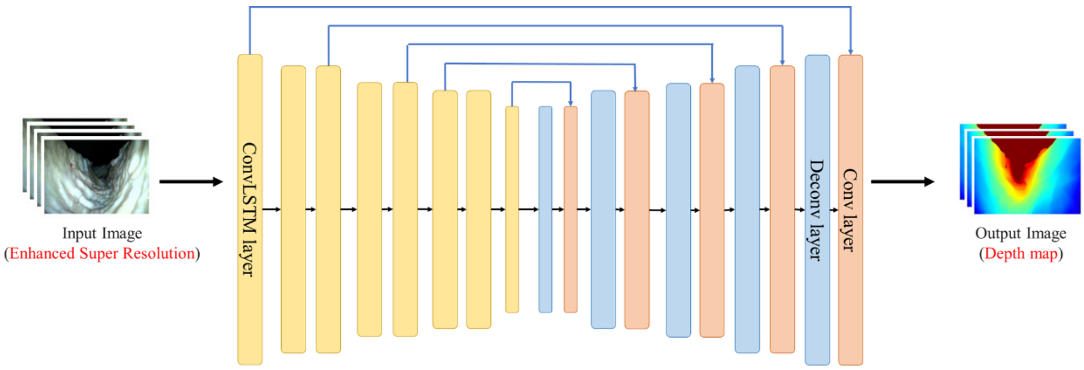


圖 3、DepthNet 架構圖[5]

5. 參考文獻

[1] R. Bogdanova, P. Boulanger, B. Zheng, “Depth Perception of Surgeons in Minimally Invasive Surgery,” *Surgical Innovation*, vol. 23, pp. 515-24, 2016.

[2] K. Zhou, X. Meng, B. Cheng, “Review of Stereo Matching Algorithms Based on Deep Learning,” *Computational Intelligence and Neuroscience*, vol. 2020, 2020.

[3] Y. Ming, X. Meng, C. Fan, H. Yu, “Deep learning for monocular depth estimation: A review,” *Neurocomputing*, vol. 438, pp.14-33, 2021.

[4] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong et. al., “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks,” *Proceedings of the European conference on computer vision (ECCV) workshops*, 2018.

[5] A. C. Kumar, S. M. Bhandarkar, M. Prasad, “Depthnet: A recurrent neural network architecture for monocular depth prediction,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018.

[6] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234-241, 2015.

伍、在校成績證明

中山醫學大學
Chung Shang Medical University

中山醫學大學歷年成績單

2021/9/11

學系：健康管理學院醫學資訊學系　學號：0758025 姓名：謝雅竹　修業總學分：127　畢業總成績：

107學年度(107 年 9 月到 108 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
物件導向程式設計(一)	必			3	99
計算機組織與結構	必			3	81
健康資訊輔具概論	必			2	96
微積分	必	2	92	2	73
管理學	必	2	85		
線性代數	必			2	70
醫務管理概論	必			2	88
醫學資訊概論	必	3	91		
大學之道(一)(大學之道)	必	1	90		
大學之道(二)(大學之道)	必			1	95
. NET程式設計	選			3	98
C語言	選	3	79		
MATLAB程式設計	選			3	71
資訊倫理	選	2	90		
管理數學	選	2	84		
體育-足球	選			1	82
體育-網球	選	1	83		
智慧財產權法概要(社會領域)	選	2	89		

107學年度(107 年 9 月到 108 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
日本日常生活禮儀(院特色通識領域)	選			2	96
日本旅遊文化(院特色通識領域)	選	2	84		
法醫與生活(醫學人文核心通識)	選	2	86		
學業總分			1903		2070
學業平均成績			86.5		86.25
修習學分數		22		24	
實得學分數		22		24	
累計學分數		22		46	
全系人數		45		45	
名次		15		17	
名次占全系人數百分比		33.33		37.78	
軍訓成績					
體育成績					
服務教育成績					
操行成績			84		88

108學年度(108 年 9 月到 109 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
作業系統	必			3	84
研究方法與生醫資料分析	必			3	83
健康雲與大數據	必	2	87		
資料庫系統	必			3	86
資料結構	必	3	88		
機率與統計	必	3	79		
臨床醫學網路暨資訊安全	必	3	99		
醫用程式語言	必			3	95
醫資專題導引	必	1	89		
生醫訊號處理概論	選	3	100		
專題研究(一)	選			1	100
醫用電子學	選			2	99
醫學影像處理	選			3	99
醫學影像傳輸系統	選			2	91
體育-有氧舞蹈	選	1	89		
體育-流行舞蹈	選			1	94
變態行為與現代生活(社會領域)	選	2	90		
進階程式語言(科普領域)	選			2	96

108學年度(108 年 9 月到 109 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
從生活中談微生物及免疫(醫學人文核心通識)	選			2	90
學業總分			1630		2287
學業平均成績			90.56		91.48
修習學分數		18		25	
實得學分數		18		25	
累計學分數		64		89	
全系人數		44		42	
名次		12		8	
名次占全系人數百分比		27.27		19.05	
軍訓成績					
體育成績					
服務教育成績					
操行成績			84		84

備註:
1. #抵免 ##校際選課 @暑修 &免修 ◇跨校雙主修/輔系 *不及格 ◎成績未送出 P檢定通過 F檢定不通過
2.抵免及校際選課學分，成績不列入當學期學分及平均成績，惟列入畢業修業總學分，不列入畢業總成績。
3.暑修學分及成績不列入當學期學分及平均成績，惟列入畢業修業總學分及畢業總成績。
4.跨校雙主修/輔系修讀科目學分及成績列入當學期與畢業學分及成績。
深化醫能力百分百: P(147)

最低畢業學分數： 128 (必修學分數75 學分)

全系人數41人,該生1至3年級下學期學業成績總平均90.39,
名次為第8名,名次占全系19.51%

中山醫學大學
Chung Shang Medical University
教務處
Academic Affairs Office

中山醫學大學
Chung Shang Medical University

中山醫學大學歷年成績單

2021/9/11

學系：健康管理學院醫學資訊學系　學號：0758025 姓名：謝雅竹　修業總學分：127　畢業總成績：

109學年度(109 年 9 月到 110 年 6 月)

科目	必選修	第一學期 學分	第一學期 成績	第二學期 學分	第二學期 成績
大數據管理系統	必	3	88		
電子病歷實務	必	2	95		
實習	必			3	96
管理資訊系統	必	2	79		
優質網路健康照護	必			2	97
醫用行動軟體設計	必			3	99
醫用資訊系統分析與設計	必	3	97		
醫院資訊系統實務	必			2	93
醫學資訊專題(一)	必			1	100
醫療資訊管理	必			2	95
專題研究(二)	選	1	100		
深度學習	選	2	98		
圖型識別概論	選	3	99		
離散數學	選	3	98		
宮崎駿動畫賞析(文藝領域)	選			2	85
生命之歌(科普領域)	選			2	91
輻射與人生(科普領域)	選			2	96
學業總分			1790		1799
學業平均成績			94.21		94.68
修習學分數		19		19	
實得學分數		19		19	
累計學分數		108		127	
全系人數		41		41	
名次		2		5	
名次占全系人數百分比		4.88		12.2	
軍訓成績					
體育成績					
服務教育成績					
操行成績			86		85

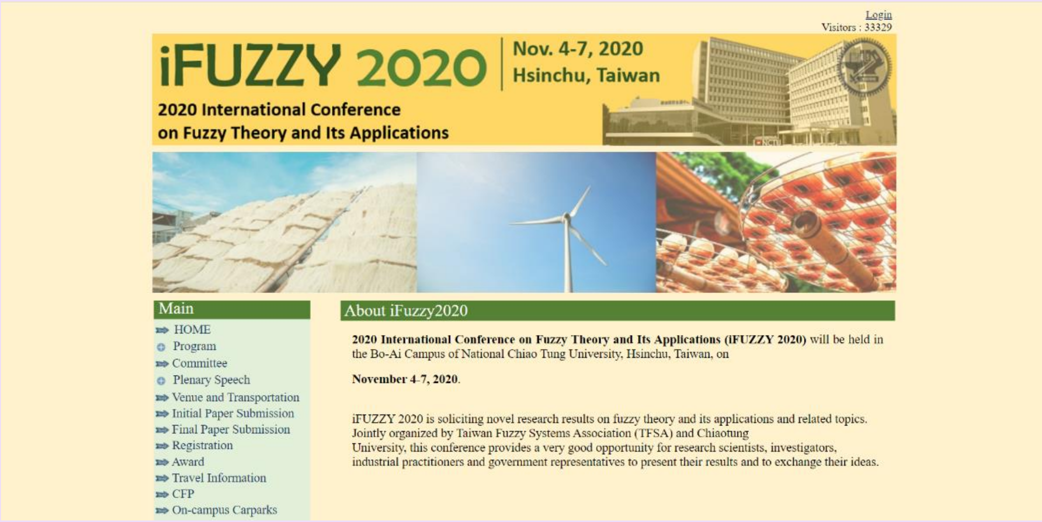
備註:
1. #抵免 ##校際選課 @暑修 &免修 ◇跨校雙主修/輔系 *不及格 ◎成績未送出 P檢定通過 F檢定不通過
2.抵免及校際選課學分，成績不列入當學期學分及平均成績，惟列入畢業修業總學分，不列入畢業總成績。
3.暑修學分及成績不列入當學期學分及平均成績，惟列入畢業修業總學分及畢業總成績。
4.跨校雙主修/輔系修讀科目學分及成績列入當學期與畢業學分及成績。
深化醫能力百分百: P(147)

最低畢業學分數： 128 (必修學分數75 學分)

全系人數41人,該生1至3年級下學期學業成績總平均90.39,
名次為第8名,名次占全系19.51%

中山醫學大學
Chung Shang Medical University
教務處
Academic Affairs Office

陸、校內外合作計畫

2020 International Conference on Fuzzy Theory and Its Application 暨中華民國第二十八屆模糊理論及其應用研討會網站	
合作單位	中華民國模糊學會
開發團隊	**謝雅竹**、張婷淇、葉品郁、陳怡妙
指導教授	秦群立 老師
系統說明	此系統為中華民國模糊學會舉辦「2020 International Conference on Fuzzy Theory and Its Application 暨中華民國第二十八屆模糊理論及其應用研討會網站」，提供各式研討會活動時程、論文提交資訊、場地交通等重要訊息，供主辦方將研討會資訊公開呈現給來自各方的投稿者。  圖 1、「2020 International Conference on Fuzzy Theory and Its Application 暨中華民國第二十八屆模糊理論及其應用研討會網站」主介面圖
開發程式	✓ 程式語言：PHP、JavaScript ✓ 標籤語言：HTML ✓ 樣 式 表：CSS ✓ 資 料 庫：MySQL
系統狀態	已上線，在研討會舉辦期間皆運作良好
系統連結	https://neuroengineering.nctu.edu.tw/ifuzzy2020/

iFUZZY2020 感謝狀

感 謝 狀

(109)TFSA-C-1091101011

秦群立 教授所指導之

中山醫學大學 醫學資訊學系　謝雅竹 君

協助 2020 International Conference on Fuzzy Theory and Its Applications 暨 中華民國第二十八屆模糊理論及其應用研討會，網站製作及維護管理。

特頒發此狀，以資感謝。

中華民國模糊學會

理 事 長　鄭錦聰

中　華　民　國　一百零九　年　十一　月　七　日

<table>
<tr><th colspan="3">We can help 筆記系統</th></tr>
<tr><td>合作單位</td><td colspan="2">醫學資訊學系及醫療產業科技管理學系</td></tr>
<tr><td>開發團隊</td><td colspan="2">**謝雅竹**、張婷淇、葉品郁、陳怡妙</td></tr>
<tr><td>指導教授</td><td colspan="2">秦群立 老師</td></tr>
<tr><td>系統說明</td><td colspan="2">此系統為聾朋友聽覺輔助學習系統，透過資訊集結大眾的力量建立筆記的資料庫，輔助聾朋友了解老師的上課內容，甚至是同學們整理後的資訊，解決聽覺嚴重受損的聾朋友於學習上所遇到的不便。

圖 2、「We can help 筆記系統」主介面圖</td></tr>
<tr><td>開發程式</td><td>✓ 程式語言：PHP、JavaScript
✓ 標籤語言：HTML</td><td>✓ 樣式表：CSS
✓ 前端框架：Bootstrap
✓ 資料庫：MySQL</td></tr>
<tr><td>系統狀態</td><td colspan="2">已上線</td></tr>
<tr><td>系統連結</td><td colspan="2">http://140.128.137.11/new_note/notice.php</td></tr>
</table>

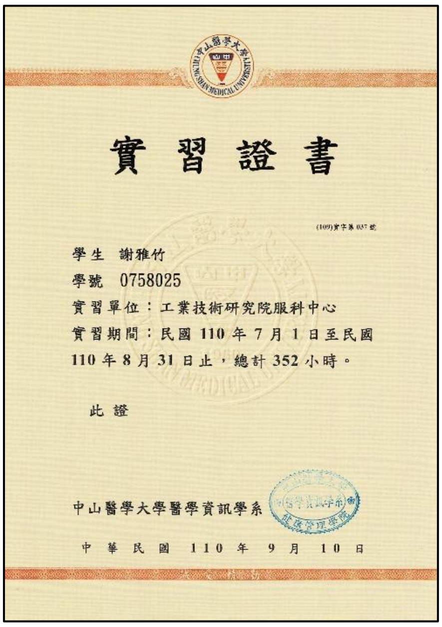
<table>
<tr><th colspan="3">學生活動應用系統</th></tr>
<tr><td>合作單位</td><td colspan="2">中山醫學大學-學生事務處</td></tr>
<tr><td>開發團隊</td><td colspan="2">**謝雅竹**、張婷淇、葉品郁、陳怡妙</td></tr>
<tr><td>指導教授</td><td colspan="2">秦群立 老師</td></tr>
<tr><td>系統說明</td><td colspan="2">此系統為各社團進行活動申請、經費申請與借閱器材之線上資訊平台。透過系統輔助讓老師更快速掌握審核各申請單所申請之資料，達到減少紙本資料且方便管理的目的。系統主介面如圖 3(a)所示，目前正在進行系統介面的改版，如圖 3(b)所示，預計在 2021 年底完成。


(a)


(b)
圖 3、「學生活動應用系統」主介面圖</td></tr>
<tr><td>開發程式</td><td>✓ 程式語言 ：PHP、JavaScript
✓ 標籤語言 ：HTML</td><td>✓ 樣 式 表 ：CSS
✓ 前端框架 ：Bootstrap
✓ 資 料 庫 ：MySQL</td></tr>
<tr><td>系統狀態</td><td colspan="2">已上線，介面改版中，預計 2021 年底完成</td></tr>
<tr><td>系統連結</td><td colspan="2">http://140.128.142.94/osa/index.php</td></tr>
</table>

視光系實習志願選填與成績計算系統	
合作單位	中山醫學大學視光學系
開發團隊	**謝雅竹**、張婷淇、葉品郁、陳怡妙
指導教授	秦群立 老師
系統說明	此系統共分為兩個部分，分別為成績計算系統及志願選填系統，其介面如圖 4(a)及圖 4(b)所示，用以提供視光系系秘書製作實習成績計算及實習志願選填文件，透過輸入特定格式的基本資料 Excel 檔案，即可使用系統內各項功能進行批次處理，以精準一次到位的方式達到行政作業所需的格式，完成文件後可匯成 PDF 檔案，減輕人力核對大量資料的負擔，有效縮短作業時程。  (a) (b) 圖 4、「視光系實習志願選填與成績計算系統」主介面圖
開發程式	✓ 程式語言：Visual Basic for Application (VBA)
系統狀態	已上線
系統連結	雲端連結

柒、校外實習

實習單位	工業技術研究院 服科中心	
實習期間	**2021/03/01~2021/08/31**	
指導員	蔡明杰、張簡文昇	
實習成績	**96 分**	
組員	**謝雅竹**、張婷淇、葉品郁、陳怡妙	
系統名稱	結合穿戴式感測裝置進行腿部健康預測系統	
實習證明	 工業技術研究院 Industrial Technology Research Institute 實習證明書 中山醫學大學 醫學資訊學系 工程人才 謝雅竹 自110年3月1日至110年8月31日，於工業技術研究院 「產學研工程人才實務能力發展基地計畫-健康物聯網應用主題班」 完成實習，共計480小時 特致此狀，以茲證明  實 習 證 書 學生 謝雅竹 學號 0758025 實習單位：工業技術研究院服科中心 實習期間：民國 110 年 7 月 1 日至民國 110 年 8 月 31 日止，總計 352 小時。 此 證 中山醫學大學醫學資訊學系 中 華 民 國 110 年 9 月 10 日	

A.實習內容

◆動機與目的

現今已有許多疾病可透過穿戴式感測裝置所獲得的人體步態數據進行預測，例如帕金森氏症與肌少症等。其中肌少症是近年來受到高度重視的疾病，主要成因為肌肉量減少及肌肉功能衰退。目前在臨床上，早期肌少症的診斷大多會要求病患重複多次進行相似的測驗像是直線行走、起身行走等，再依據醫師自身的經驗進行診斷，但早期肌少症一般並不會產生明顯的不適感，須等到病患發現肌力下降、平地行走困難且須倚靠支撐物方能起身等症狀，影響日常生活的行為時，才會前往醫院進行檢查。而肌少症若未及早發現並給予適當治療，嚴重時則可能會導致病患失能。

因此我們進行了「**結合穿戴式感測裝置進行腿部健康預測**」的專案研究與系統實作，透過自行開發出的穿戴式感測裝置來蒐集受測者步行過程中的加速度訊號與表面肌電訊號，再利用自建的測驗平台、步態數據指標演算法以及深度學習模型，以預測受測者是否有罹患肌少症的風險及其嚴重性。本系統不僅讓受測者能提早發現並更加重視骨骼肌流失的問題，還能提供步態指標給醫師參考，輔助醫師評估受測者的健康狀況。

◆系統流程

首先本系統會要求受測者於人體特定位置配戴穿戴式感測裝置，以獲取站立、坐姿以及行走時的人體活動訊號，接著將蒐集到的訊號上傳並儲存至伺服器的資料庫，系統會將蒐集到的訊號透過折線圖的方式呈現於系統介面上，以提供使用者參考。另外，系統同時還會對蒐集到的訊號透過自建的步態數據指標演算法進行處理與分析，以計算步態指標，並運用深度學習模型評估肌少症風險程度，最後將指標與預測結果呈現於系統介面上，方便受測者及醫師參考其預測結果。「**結合穿戴式感測裝置進行腿部健康預測**」系統流程圖如圖 1 所示。

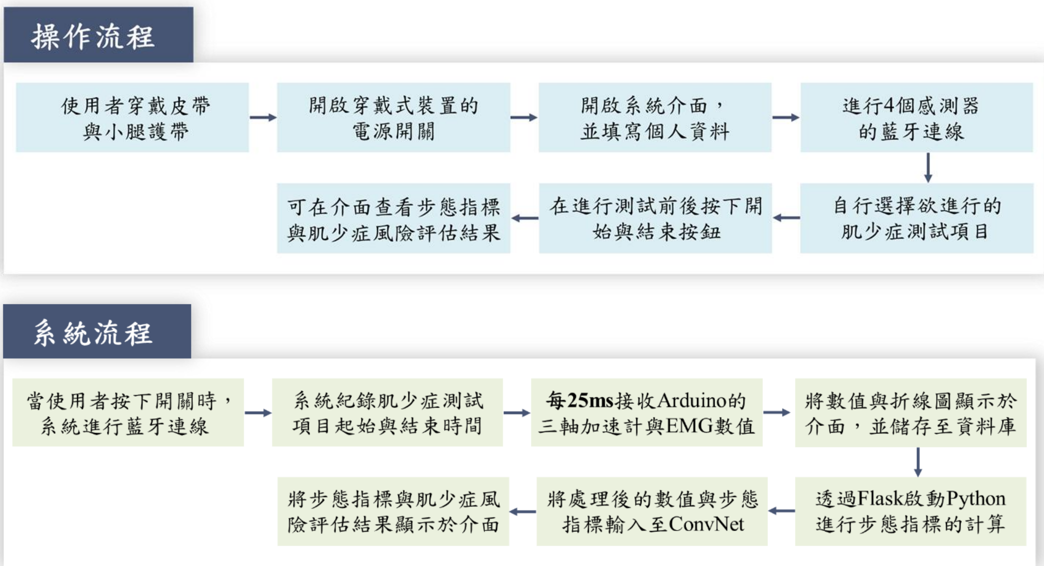


圖 1、系統流程圖

◆ 測驗項目流程

在 2020 年歐洲肌少症工作小組提出肌少症定義與判斷，說明可以透過受測者的肌肉力量、肌肉質量與體能表現來預測是否有罹患肌少症的風險。此外，歐洲肌少症工作小組還提供自行檢測肌肉力量、肌肉質量與體能表現三部分的測驗項目及判斷標準，方便醫師進行肌少症的診斷。我們參考歐洲肌少症工作小組發布的肌少症判斷標準，透過「4 公尺行走測試」與「起身行走測試」瞭解受測者的肌力與體能表現，並預測是否有罹患肌少症的風險。

1. 「4 公尺行走測試」

受測者一開始為站立姿勢，接著直線行走 4 公尺後停下，若受測者的行走速度每秒少於 0.8 公尺，則表示罹患肌少症的風險高。圖 2 為「4 公尺行走測試」的示意圖。

圖 2、「4 公尺行走測試」的示意圖

2. 「起身行走測試」

受測者一開始為坐姿，接著從椅子起立，直線行走 3 公尺，轉身後再直線走回椅子並坐下，若總測驗時間大於 20 秒，則表示患有肌少症的風險高。圖 3 為「起身行走測試」的示意圖。

圖 3、「起身行走測試」的示意圖

◆ 系統開發工具

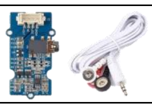
本系統可分為軟體與硬體兩部分，在硬體方面，系統以 Arduino Nano 33 BLE Sense 為控制核心，結合 EMG Detector 等多項硬體感測模組以作為數據蒐集裝置。而在軟體部分可細分為三部分，分別為步態數據蒐集的測驗平台、步態數據指標演算法及深度學習預測模型，以下會詳細介紹。系統示意圖如圖 4 所示。

圖 4、系統示意圖

1. 硬體

本系統所使用到的硬體元件，如 Arduino Nano 33 BLE Sense、EMG Detector 等，其用途及功能如表 1 所示。我們是使用滑動開關控制一個 3.7V 的聚合物鋰電池是否供電，並搭配一個鋰電池充電板，方便為穿戴式感測裝置進行充電。若是電源開啟時，電池會以並聯的方式連接到工作電壓為 3.3V 的 Arduino Nano 33 BLE Sense 以及 EMG Detector。最後以每 25ms 的頻率擷取步態數據，並透過 Arduino Nano 33 BLE Sense 內建的的藍牙 BLE5.0 將接收到的步態數據傳輸至系統測驗平台。

表 1、硬體設備表

	**Arduino Nano 33 BLE Sense：** 作為硬體感測器的主要核心，內含 9 軸慣性感測器與 Bluetooth 5.0
	**Grove-EMG Detector：** 用於收集肌肉收縮的電訊號
	**3.7V 聚合物鋰電池：** 用於提供硬體感測器運作之電源
	**鋰電池充電保護板：** 作為聚合物鋰電池的充電板
	**滑動開關：** 作為穿戴式裝置的開關

我們一共完成 4 組可充電式的穿戴式感測裝置，並且將硬體大小控制在 5x6.5 公分，以方便使用者攜帶。硬體電路設計圖如圖 5 所示，硬體電路實作圖如圖 6(a)與(b)所示。

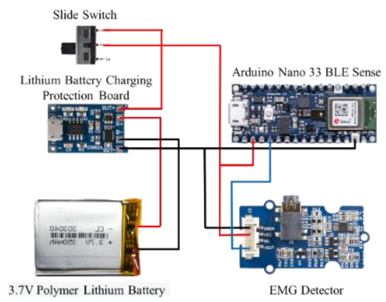


圖 5、硬體電路設計圖

(a) (b)

圖 6、硬體電路實作圖，(a)為硬體電路的正面，(b)為硬體電路的反面

我們參考多篇與穿戴式感測裝置配戴位置相關的文獻之後，決定將穿戴式感測裝置安裝於人體的左、右腳小腿處以及左、右腰位置，用於測量人體步態數據。接著有文獻提到將 EMG Detector 放置於在特定肌肉位置的皮膚表面，可用於瞭解人體健康狀況。因此我們將 EMG Detector 的電極貼片貼於大腿的股四頭肌位置，以及左、右腰的豎脊肌位置，用於蒐集人體肌電訊號。圖 7 為受測者實際配戴穿戴式感測裝置之示意圖。

圖 7、受測者實際配戴穿戴式感測裝置之示意圖

2. 軟體

本系統在軟體的部分主要分為三大區塊，分別為人體姿勢與步態訊號收集、網站系統及大數據分析/深度學習，圖 8 為軟體各區塊所使用到的開發環境、開發框架、函式庫及資料庫的示意圖。

圖 8、軟體開發所使用開發環境、開發框架、函式庫及資料庫示意圖

1) 步態數據蒐集的測驗平台

我們建置出一個測驗平台系統，以獲得穿戴式感測裝置蒐集到的人體活動訊號，測驗平台的介面及架構如圖 9 所示。在測驗平台前端的部分，使用 HTML5、JavaScript 及 CSS 完成 Single-page 的 Web App 介面，以方便使用者操作及理解為設計理念，可相容於 iOS 及 Android 平台。而為了整合軟硬體之間的溝通，我們採用 Web Bluetooth API 接收透過 BLE 5.0 傳送的多個三軸加速度訊號與 EMG 訊號，並利用 Web Storage API 將數據暫存於瀏覽器中，每接收 50 筆資料即封裝一次，再利用 Ajax 技術搭配 PHP 後端程式語言將資料傳至 MySQL 資料庫中並儲存，以避免同時傳送過多的資料導致數據的遺失，最後，在測驗完畢後會透過 Flask 框架啟動 Python 程式，進行步態指標的計算，以及評估罹患肌少症的風險程度，並將結果顯示於系統網頁供使用者參考。

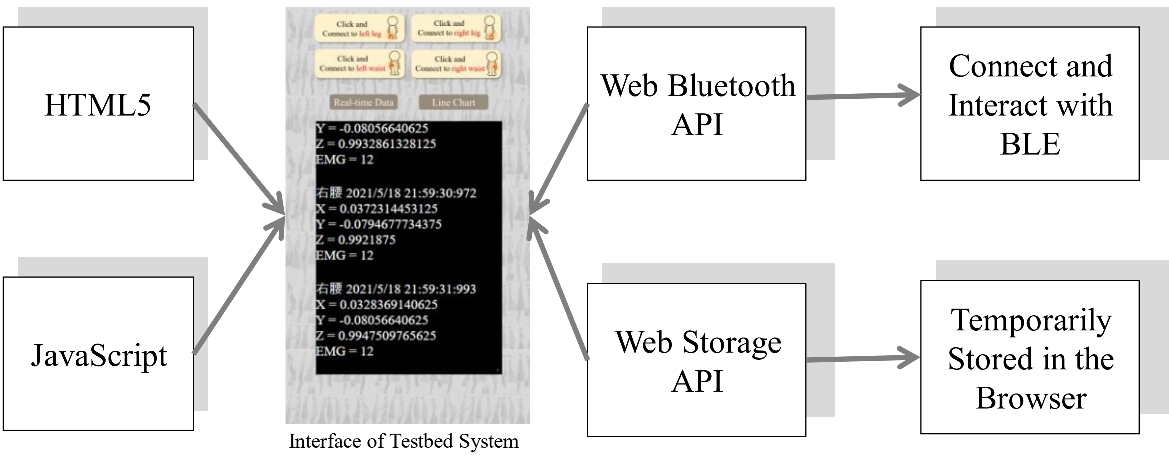


圖 9、測驗平台系統介面及其架構

2) 步態數據指標演算法

為了量化穿戴式感測裝置蒐集到的三軸加速度訊號及肌電訊號，我們參考多篇文獻，透過 9 個公式換算三軸加速度訊號及肌電訊號，最後計算出步長、步幅、步數、步速、4 個三軸加速度值的平方根、4 個 EMG 值的 RMS、2 個三軸加速度值的 RMS、4 個 EMG 平均值及 4 個人體左右側對稱性等 22 項指標，並將這些指標進行標準化，除了能將這些指標用於訓練深度學習預測肌少症風險程度外，還能將指標提供給醫師參考，作為診斷的依據，用於訓練深度學習模型指標如圖 10 所示。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	id	species	step_len	stride	num	speed	acc_r	acc_l	acc_rw	acc_lw	emg_r	emg_l	emg_rw	emg_lw	RMS_r	RMS_l	RMS_rw	RMS_lw	RMS_acc	RMS_acc_lw
2	1	中高風險	-0.53899	-0.53899	0.098268	0.264829	-0.2895	-0.17635	-0.36342	-0.08339	-0.09558	0.1165	1.150383	0.284446	-0.22278	0.06598	0.444082	1.158906	-0.16065	-0.15721
3	2	低風險	1.007378	1.007378	-0.83621	0.051065	-0.27866	-0.54972	-0.27383	-0.19412	1.163184	-0.01684	-1.22116	0.367715	1.191823	-0.08381	0.297539	-1.13527	-0.09846	-0.0966
4	3	低風險	1.208329	1.208329	-0.83621	0.047003	-0.26222	-0.57184	-0.3881	0.048335	0.350223	-0.44335	-0.45638	-0.54848	0.361724	-0.476	-0.24095	-0.37367	2.987886	2.619325
5	4	低風險	0.644942	0.644942	-0.70271	0.053316	-0.41334	-0.66622	-0.54533	-0.65933	-0.47858	-0.21178	-0.58577	1.00175	-0.53104	-0.23231	1.134656	-0.6167	-0.13149	-0.12749
6	5	低風險	-0.43522	-0.43522	-0.03523	0.442495	-0.27067	-0.01681	0.267148	0.035365	-0.43589	-0.20003	-0.96691	0.522372	-0.60034	-0.2897	0.40507	-1.00657	-0.16067	-0.15722
7	6	中高風險	-1.35303	-1.35303	1.700228	0.046353	-0.39816	0.008296	-0.19409	-0.15873	-0.24753	-0.29006	0.652267	0.451656	-0.39133	-0.39086	0.319974	0.650215	-0.16061	-0.15717
8	7	中高風險	-0.54372	-0.54372	0.098268	0.270961	-0.2895	-0.17635	-0.36342	-0.08339	-0.09558	0.1165	1.150383	0.284446	-0.22278	0.06598	0.444082	1.158906	-0.16065	-0.15721
9	8	低風險	-0.59106	-0.59106	0.231765	0.523113	-0.30653	-0.4864	-0.2851	-0.50724	3.448569	0.677253	-0.06472	0.471851	3.761215	0.699024	0.344271	-0.0839	-0.16067	-0.15723
10	9	低風險	-0.27735	-0.27735	-0.16872	0.361191	-0.25528	-0.42898	-0.13702	0.256511	1.663746	-0.21321	-0.54667	0.976376	1.729153	-0.30451	0.951397	-0.57683	-0.16066	-0.15722
11	10	低風險	-0.11523	-0.11523	-0.30222	0.493155	-0.45428	-0.60093	0.050917	0.415383	-0.20538	-0.0658	-0.57909	-0.80919	-0.34458	-0.13886	-0.44341	-0.60997	-0.16067	-0.15722
12	11	低風險	-0.89412	-0.89412	0.765751	0.468686	-0.28117	-0.01868	-0.02356	-0.05085	-0.43589	-0.19934	-0.95875	0.525786	-0.60035	-0.28891	0.409178	-0.99827	-0.16067	-0.15722
13	12	中高風險	-1.79714	-1.79714	1.833724	0.078217	-0.4292	-0.35556	-0.32232	-0.52783	-0.24978	-0.28563	-0.94516	0.438349	-0.39385	-0.38586	0.303978	-0.98418	-0.15905	-0.15572

圖 10、用於訓練深度學習模型的各項指標

3) 深度學習預測模型 – LCNet

我們提出 LCNet 深度學習演算法，此模型由一層卷積層和兩層全連接層組成，將上述提到的步態數據指標及類別作為輸入資料進行訓練，其中肌少症的類別分為分為兩類，分別為高肌少症罹患風險及低肌少症罹患風險，最後模型的訓練成功率為 100%，測試成功率為 94%。

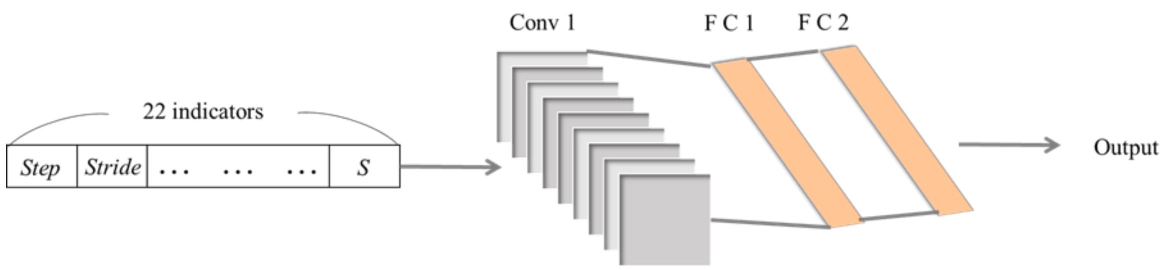


圖 11、LCNet 深度學習模型架構圖

◆ 結論

本研究的主要貢獻在於實作出完整且具有連貫性的步態分析與肌少症預測的系統，讓**受測者在家便能監測自己的肌肉力量與體能表現狀況**，無須等到身體出現警訊才前往醫院看診，導致延誤治療時間。另外，本系統開發出的穿戴式感測裝置具有功耗低、體積小且藍牙連線範圍更加廣泛的優勢，並且能蒐集慣性訊號與生醫訊號，更準確地擷取人體活動訊號，期望未來能應用於更多不同疾病的預測與評估。

◆ 附錄

1) 實習專案成果影片

2) 於 2021/09/19 投稿至 *Measurement*，目前狀態為 **Under Review**

B.實習心得

從這幾個月在工研院的實習中，我學到了如何從無到有做出**專題**的**過程**，從題目發想、文獻探討、方法實踐以及結果發表，都是由我們、指導老師以及部門主管一步步開會所討論出來的。由於工研院的研究主題為**健康物聯網**，因此我們最後討論出來的主題是透過配戴感測器來預測腿部健康，目前專注於研究肌少症的議題，當然未來也可應用於其他疾病的預測。在討論方法的過程當中，**部門主管**總能提出一些對我們有幫助的見解**指點**我們，例如我們對於受測者的**檢測項目**有些許疑慮時，他建議我們可以透過計時起走測試(TUG)來蒐集受測者行走時的步態參數，再開完會討論後我們再去查詢與 **TUG 測試**相關的資料後，發現這是蒐集步態數據時常用的測試項目之一，且也有多篇文獻是採用這個測試項目，有種豁然開朗的感覺；此外，因為要同時蒐集腰部及腿部的步態數據，因此我們需要智慧型可與多個蒐集裝置配對，我們根據網路上的資料了解到 **BLE5.0** 的技術可以完成我們的需求，並將我們的想法與指導老師討論，覺得可行便開始進行實作，但由於這項技術是第一次接觸，且又要與多個裝置進行配對，因而增加了難度，但同組的組員們會幫忙查資料，而負責嘗試的組員也會詳細記錄嘗試的內容與結果，在這樣的良性循環中我們才得以快速地完成**多裝置連線**，而我身為負責嘗試的組員，在參與這次研究中深深感受到了**詳細記錄**的重要性，即便記錄麻煩且耗時，但真的幫助很大，避免重複作業，也能**清楚了解項目**的執行**進度**。另外在蒐集數據的過程中，我們發現電路有不穩的狀況，與指導老師討論後決定在硬體裝置的外殼上坐加強以保護內部的電路，但由於市面上所販售的保護殼皆不符合我們設計的硬體裝置大小，因此我們在網路上查了許多資料後發現壓克力板是可以切割也可以黏合的，因而才完成我們硬體外殼的加強，且蒐集的硬體裝置在此次的加強後，電路不穩的情況也改善了。

在執行工研院專案的過程中，除了能夠學習到新技術、專案開發等外，我認為收穫最大的還是學會如何**定期**與部門主管**匯報階段性**的**成果**，就算目前只進行到一半，還是要能夠為目前的成果做一個總結，且**報告時長**的**拿捏**也是很重要的，像是此次學習總共分為 9 個不同的研究主題，每個月都會定期開會並由每組負責報告目前的進度，但主管不一定每次都有足夠的時間聽完所有的報告，有時是半小時的報告，有時是 10 分鐘，甚至只有 2 分鐘可以報告，這是我在實習中感到最震撼但也認為是最重要的部分。除了研究外，在工研院的實習中也提供多堂線上課程的實習，包含新技術的趨勢如 AI 與 5G 未來的發展趨勢以及專案的規劃與分析等，此外工研院也提供許多企業的介紹與參訪，讓我們能夠更加了解目前業界的發展走向及所需人才，工研院也提供許多就業的**媒合機會**，為前來的實習生搭建一個就業的渠道，幫助非常大。這次實習的經驗對我來說是難得的經驗，不僅體驗了職場的運作模式還學習到很多知識以及獲取寶貴的經驗。

圖為專案的研究過程、開會紀錄及最後與實習結束親自前往工研院與主管合影的照片

捌、專業領域優秀表現

一、專業證書與專業課程

2020「台灣發展智慧醫療之新創趨勢與實務應用」系列論壇	
主辦單位	**臺北市生物產業協會**
開課講師	衛生福利部醫事司石崇良司長、奇美醫院林宏榮首席副院長、台北榮民總醫院郭萬祐主任等 20 位講師
修課日期	2020/08/01~2020/08/29，每週六 09:00~14:40


台北市生物產業協會 一〇九年度
大健康產業人才培訓課程
TBA109-M-150
結業證書
謝雅竹 君 於中華民國一〇九年八月一日至
一〇九年八月二十九日期間 參加本協會主辦之
「二〇二〇台灣發展智慧醫療之新創趨勢
與實務應用」系列課程 時數共計三十小時
修業期滿 成績合格
特頒此證
台北市生物產業協會
理事長 張立言
中華民國一〇九年八月二十九日

如何準備智慧醫療研究資料？從石油到汽油	
主辦單位	**臺大醫院智慧醫療中心**
開課講師	臺大醫學研究部李宜家副主任、臺大智慧醫療中心李建璋副主任、臺大人工智慧中心計畫博士後研究人員李佳達博士、臺大影像醫學部李文正副主任
修課日期	2021/03/26

AI 醫療資訊攻防戰-資安理論與實務探討

主辦單位	**科技部人工智慧生技醫療創新研究中心**
開課講師	財團法人電信技術中心王慶豐副主任、吳勁像工程師、林家樑副主任
修課日期	2021/05/21

教育部人工智慧技術及應用人才培育計畫
110年實習前訓練課程結業證書
CERTIFICATE OF PRE-INTERNSHIP TRAINING COURSES

茲證明:
PRESENTED TO: 謝 雅 竹

於 110 年 6 月 2 日及 110 年 6 月 3 日參加「教育部補助人工智慧技術及應用人才培育計畫分項之人工智慧人才產業鏈結 110 年實習前訓練課程」，共計 12 小時，修課期滿，特頒此結業證書，以茲證明。

Has successfully completed the Ministry of Education - Artificial Intelligent Talent Cultivation Program from June 2 to June 3 ,2021 . A total of 12 hours .

教育部人工智慧技術及應用人才培育計畫
Artificial Intelligence Talent Cultivation Program

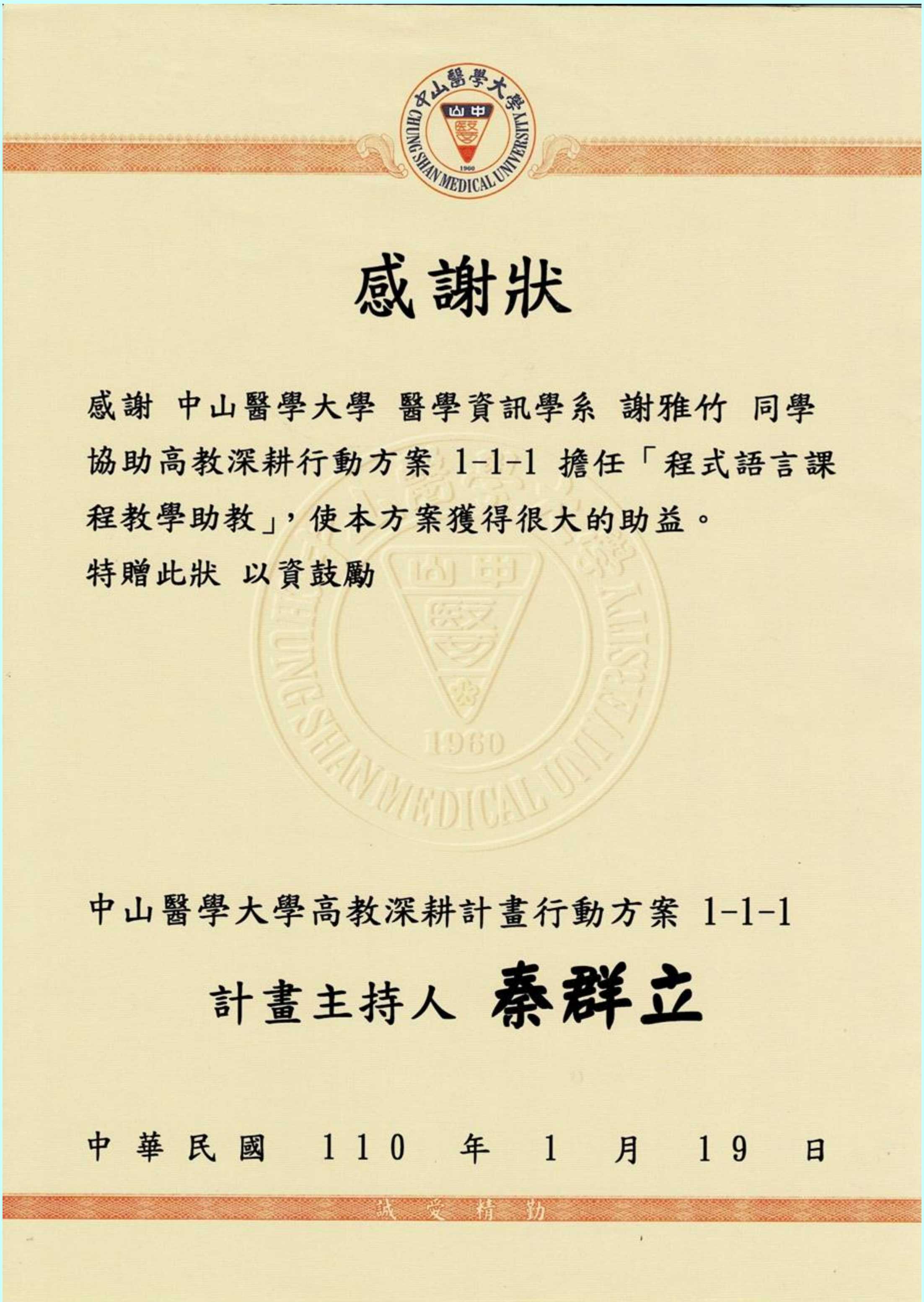
詹寶珠
總計畫主持人 詹寶珠教授
Prof.Pau-Choo Chung,
Chief Investigator

張傳育
子計畫主持人 張傳育教授
Prof.Chuan-Yu Chang,
Principal Investigator

高教深耕行動方案

擔任程式語言助教，輔助老師教學並幫助同學學習

時間： 2019/09~2020/02、2020/10~2021/02



中山醫學大學 CHUNG SHAN MEDICAL UNIVERSITY 1960

感謝狀

感謝 中山醫學大學 醫學資訊學系 謝雅竹 同學

協助高教深耕行動方案 1-1-1 擔任「程式語言課程教學助教」，使本方案獲得很大的助益。

特贈此狀 以資鼓勵

中山醫學大學高教深耕計畫行動方案 1-1-1

計畫主持人 秦群立

中 華 民 國 110 年 1 月 19 日

誠 愛 精 勤

「智慧醫療 AI 應用技術人才養成班」之智慧醫療 AI 技術應用

擔任 AI 專題實作課程助教，輔助老師教學並幫助同學學習

時間： 2021/07/20~2021/07/26

工業技術研究院
Industrial Technology Research Institute

Certificate

培 訓 證 書

感 謝 狀

茲感謝 謝雅竹 小姐 於 110 年 7 月 20 日至 7 月 26 日，擔任勞發署『產業新尖兵試辦計畫』補助課程「智慧醫療 AI 應用技術人才養成班」之智慧醫療 AI 技術應用-專題實作課程助教，令學員受益良多，頒發此狀，代表衷心的感謝。

工研院產業學院 執行長
中華民國 110 年 07 月 26 日

ITRI College Executive Director
July 26, 2021

ITRI College

二、程式語言與實作系統

程式語言能力	實作案例
程式語言程式設計：Python **深度學習框架**：Tensorflow、Keras、Pytorch	✓ 基於深度學習辨識血糖及血壓機之七段顯示器系統 ✓ 結合 TensorRT Pose 之訓練報告能力系統 ✓ 結合 BeautyGAN 之妝容模擬系統
嵌入式硬體與行動軟體設計實作：Arduino-C、Android、JAVA、SQLite **嵌入式硬體**：Arduino、Raspberry Pi、Arduino Jetson Nano、Arduino Jetson TX2、Arduino Jetson Xavier NX	✓ 「得心應手」心絞痛預測及偵測系統 ✓ 「聲之形」聽障輔助行車安全系統 ✓ 「一鏡到底」穿搭建議系統 ✓ 「明鏡照形」妝容模擬及化妝輔助系統
網際網路程式設計：PHP、MySQL、HTML、JavaScript、CSS	✓ 「孩好有您」年長者照護系統 ✓ 「We can help」聽障輔助學習及溝通系統 ✓ 中華民國模糊學會--2020 International Conference on Fuzzy Theory and Its Applications 暨中華民國第二十八屆模糊理論及其應用研討會網站
巨集語言：Visual Basic for Applications (VBA)	✓ 中山醫學大學視光學系實習成績計算系統 ✓ 中山醫學大學視光學系選填志願系統

三、大學期間讀過的SCI論文

Deep Learning for Sensor-based Activity Recognition: A Survey			
收錄期刊	*Pattern Recognition Letters*	期刊等級	**SCIE**
Impact Factor (2020)	3.756	Rank (2020)	Q2
全文連結	雲端連結		

Pattern Recognition Letters 119 (2019) 3–11

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Deep learning for sensor-based activity recognition: A survey

Jindong Wang[a,b], Yiqiang Chen[a,b,*], Shuji Hao[c], Xiaohui Peng[a,b], Lisha Hu[a,b]

[a] Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[b] University of Chinese Academy of Sciences, Beijing, China
[c] Institute of High Performance Computing, A*STAR, Singapore

ARTICLE INFO

Article history:
Available online 21 February 2018

Keywords:
Deep learning
Activity recognition
Pattern recognition
Pervasive computing

ABSTRACT

Sensor-based activity recognition seeks the profound high-level knowledge about human activities from multitudes of low-level sensor readings. Conventional pattern recognition approaches have made tremendous progress in the past years. However, those methods often heavily rely on heuristic hand-crafted feature extraction, which could hinder their generalization performance. Additionally, existing methods are undermined for unsupervised and incremental learning tasks. Recently, the recent advancement of deep learning makes it possible to perform automatic high-level feature extraction thus achieves promising performance in many areas. Since then, deep learning based methods have been widely adopted for the sensor-based activity recognition tasks. This paper surveys the recent advance of deep learning based sensor-based activity recognition. We summarize existing literature from three aspects: sensor modality, deep model, and application. We also present detailed insights on existing work and propose grand challenges for future research.

1. Introduction

Human activity recognition (HAR) plays an important role in people's daily life for its competence in learning profound high-level knowledge about human activity from raw sensor inputs. Successful HAR applications include home behavior analysis [61], video surveillance [50], gait analysis [21], and gesture recognition [31]. There are mainly two types of HAR: *video-based* HAR and *sensor-based* HAR [13]. Video-based HAR analyzes videos or images containing human motions from the camera, while sensor-based HAR focuses on the motion data from smart sensors such as an accelerometer, gyroscope, Bluetooth, sound sensors and so on. Due to the thriving development of sensor technology and pervasive computing, sensor-based HAR is becoming more popular and widely used with privacy well protected. Therefore, in this paper, our main focus is on sensor-based HAR.

HAR can be treated as a typical pattern recognition (PR) problem. Conventional PR approaches have made tremendous progress on HAR by adopting machine learning algorithms such as decision tree, support vector machine, naive Bayes, and hidden Markov models [34]. It is no wonder that in some controlled environments where there are only a few labeled data or certain domain knowledge is required (e.g. some disease issues), conventional PR methods are fully capable of achieving satisfying results. However, in most daily HAR tasks, those methods may heavily rely on heuristic hand-crafted feature extraction, which is usually limited by human domain knowledge [5]. Furthermore, only shallow features can be learned by those approaches [66], leading to undermined performance for unsupervised and incremental tasks. Due to those limitations, the performances of conventional PR methods are restricted regarding classification accuracy and model generalization.

Recent years have witnessed the fast development and advancement of deep learning, which achieves unparalleled performance in many areas such as visual object recognition, natural language processing, and logic reasoning [35]. Different from traditional PR methods, deep learning can largely relieve the effort on designing features and can learn much more high-level and meaningful features by training an end-to-end neural network. In addition, the deep network structure is more feasible to perform unsupervised and incremental learning. Therefore, deep learning is an ideal approach for HAR and has been widely explored in existing work [3,33,47].

Although some surveys have been conducted in deep learning [5,35,58] and HAR [7,34], respectively, there has been no specific survey focusing on the intersections of these two areas. To our best knowledge, this is the *first* article to present the recent advance on deep learning based HAR. We hope this survey can provide a helpful summary of existing work, and present potential future research directions.

* Corresponding author.
E-mail address: yqchen@ict.ac.cn (Y. Chen).

https://doi.org/10.1016/j.patrec.2018.02.010

2020 JOURNAL IMPACT FACTOR

3.756

View calculation

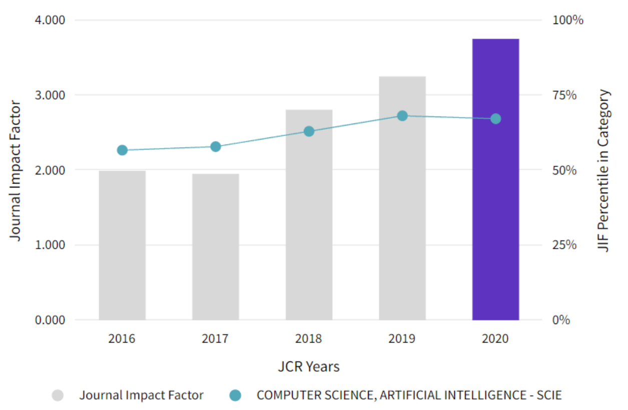
JOURNAL IMPACT FACTOR WITHOUT SELF CITATIONS

3.422

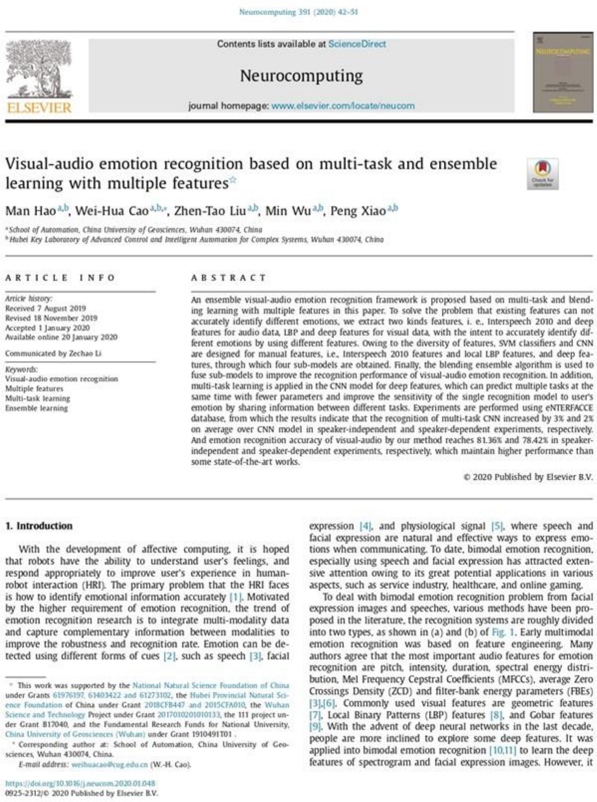
View calculation

Journal Impact Factor Trend 2020

Export



Visual-audio emotion recognition based on multi-task and ensemble learning with multiple features			
收錄期刊	*Neurocomputing*	期刊等級	**SCIE**
Impact Factor (2020)	5.719	Rank (2020)	Q1
全文連結	雲端連結		


Neurocomputing 391 (2020) 42–51

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Visual-audio emotion recognition based on multi-task and ensemble learning with multiple features[☆]

Man Hao[a,b], Wei-Hua Cao[a,b,*], Zhen-Tao Liu[a,b], Min Wu[a,b], Peng Xiao[a,b]

[a] School of Automation, China University of Geosciences, Wuhan 430074, China

[b] Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China

ARTICLE INFO

Article history:
Received 7 August 2019
Revised 18 November 2019
Accepted 1 January 2020
Available online 20 January 2020

Communicated by Zechao Li

Keywords:
Visual-audio emotion recognition
Multiple features
Multi-task learning
Ensemble learning

ABSTRACT

An ensemble visual-audio emotion recognition framework is proposed based on multi-task and blending learning with multiple features in this paper. To solve the problem that existing features can not accurately identify different emotions, we extract two kinds features, i. e., Interspeech 2010 and deep features for audio data, LBP and deep features for visual data, with the intent to accurately identify different emotions by using different features. Owing to the diversity of features, SVM classifiers and CNN are designed for manual features, i.e., Interspeech 2010 features and local LBP features, and deep features, through which four sub-models are obtained. Finally, the blending ensemble algorithm is used to fuse sub-models to improve the recognition performance of visual-audio emotion recognition. In addition, multi-task learning is applied in the CNN model for deep features, which can predict multiple tasks at the same time with fewer parameters and improve the sensitivity of the single recognition model to user's emotion by sharing information between different tasks. Experiments are performed using eNTERFACCE database, from which the results indicate that the recognition of multi-task CNN increased by 3% and 2% on average over CNN model in speaker-independent and speaker-dependent experiments, respectively. And emotion recognition accuracy of visual-audio by our method reaches 81.36% and 78.42% in speaker-independent and speaker-dependent experiments, respectively, which maintain higher performance than some state-of-the-art works.

1. Introduction

With the development of affective computing, it is hoped that robots have the ability to understand user's feelings, and respond appropriately to improve user's experience in human-robot interaction (HRI). The primary problem that the HRI faces is how to identify emotional information accurately [1]. Motivated by the higher requirement of emotion recognition, the trend of emotion recognition research is to integrate multi-modality data and capture complementary information between modalities to improve the robustness and recognition rate. Emotion can be detected using different forms of cues [2], such as speech [3], facial expression [4], and physiological signal [5], where speech and facial expression are natural and effective ways to express emotions when communicating. To date, bimodal emotion recognition, especially using speech and facial expression has attracted extensive attention owing to its great potential applications in various aspects, such as service industry, healthcare, and online gaming.

To deal with bimodal emotion recognition problem from facial expression images and speeches, various methods have been proposed in the literature, the recognition systems are roughly divided into two types, as shown in (a) and (b) of Fig. 1. Early multimodal emotion recognition was based on feature engineering. Many authors agree that the most important audio features for emotion recognition are pitch, intensity, duration, spectral energy distribution, Mel Frequency Cepstral Coefficients (MFCCs), average Zero Crossings Density (ZCD) and filter-bank energy parameters (FBEs) [3],[6]. Commonly used visual features are geometric features [7], Local Binary Patterns (LBP) features [8], and Gobar features [9]. With the advent of deep neural networks in the last decade, people are more inclined to explore some deep features. It was applied into bimodal emotion recognition [10,11] to learn the deep features of spectrogram and facial expression images. However, it

☆ This work was supported by the National Natural Science Foundation of China under Grants 61976197, 61403422 and 61273102, the Hubei Provincial Natural Science Foundation of China under Grant 2018CFB447 and 2015CFA010, the Wuhan Science and Technology Project under Grant 2017010201010133, the 111 project under Grant B17040, and the Fundamental Research Funds for National University, China University of Geosciences (Wuhan) under Grant 1910491T01 .

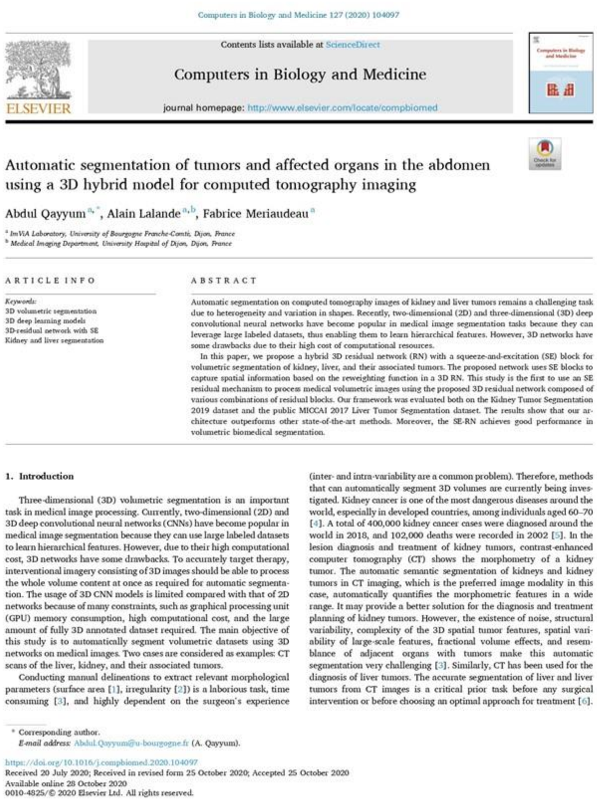
* Corresponding author at: School of Automation, China University of Geosciences, Wuhan 430074, China.

E-mail address: weihuacao@cug.edu.cn (W.-H. Cao).

https://doi.org/10.1016/j.neucom.2020.01.048
0925-2312/© 2020 Published by Elsevier B.V.

Automatic segmentation of tumors and affected organs in the abdomen using a 3D hybrid model for computed tomography imaging

收錄期刊	*Computers in Biology and Medicine*	期刊等級	**SCIE**
Impact Factor (2020)	4.589	Rank (2020)	Q2
全文連結	雲端連結		



Computers in Biology and Medicine 127 (2020) 104097

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Automatic segmentation of tumors and affected organs in the abdomen using a 3D hybrid model for computed tomography imaging

Abdul Qayyum [a,*], Alain Lalande [a,b], Fabrice Meriaudeau [a]

[a] ImViA Laboratory, University of Bourgogne Franche-Comté, Dijon, France
[b] Medical Imaging Department, University Hospital of Dijon, Dijon, France

ARTICLE INFO

Keywords:
3D volumetric segmentation
3D deep learning models
3D-residual network with SE
Kidney and liver segmentation

ABSTRACT

Automatic segmentation on computed tomography images of kidney and liver tumors remains a challenging task due to heterogeneity and variation in shapes. Recently, two-dimensional (2D) and three-dimensional (3D) deep convolutional neural networks have become popular in medical image segmentation tasks because they can leverage large labeled datasets, thus enabling them to learn hierarchical features. However, 3D networks have some drawbacks due to their high cost of computational resources.

In this paper, we propose a hybrid 3D residual network (RN) with a squeeze-and-excitation (SE) block for volumetric segmentation of kidney, liver, and their associated tumors. The proposed network uses SE blocks to capture spatial information based on the reweighting function in a 3D RN. This study is the first to use an SE residual mechanism to process medical volumetric images using the proposed 3D residual network composed of various combinations of residual blocks. Our framework was evaluated both on the Kidney Tumor Segmentation 2019 dataset and the public MICCAI 2017 Liver Tumor Segmentation dataset. The results show that our architecture outperforms other state-of-the-art methods. Moreover, the SE-RN achieves good performance in volumetric biomedical segmentation.

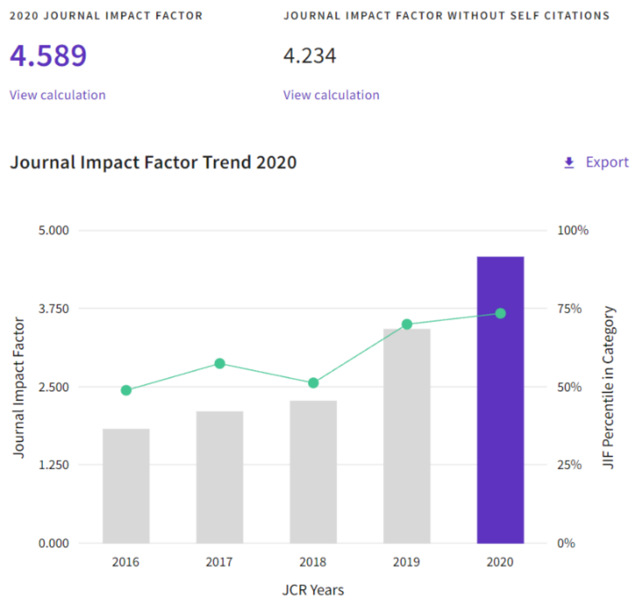
1. Introduction

Three-dimensional (3D) volumetric segmentation is an important task in medical image processing. Currently, two-dimensional (2D) and 3D deep convolutional neural networks (CNNs) have become popular in medical image segmentation because they can use large labeled datasets to learn hierarchical features. However, due to their high computational cost, 3D networks have some drawbacks. To accurately target therapy, interventional imagery consisting of 3D images should be able to process the whole volume content at once as required for automatic segmentation. The usage of 3D CNN models is limited compared with that of 2D networks because of many constraints, such as graphical processing unit (GPU) memory consumption, high computational cost, and the large amount of fully 3D annotated dataset required. The main objective of this study is to automatically segment volumetric datasets using 3D networks on medical images. Two cases are considered as examples: CT scans of the liver, kidney, and their associated tumors.

Conducting manual delineations to extract relevant morphological parameters (surface area [1], irregularity [2]) is a laborious task, time consuming [3], and highly dependent on the surgeon's experience (inter- and intra-variability are a common problem). Therefore, methods that can automatically segment 3D volumes are currently being investigated. Kidney cancer is one of the most dangerous diseases around the world, especially in developed countries, among individuals aged 60–70 [4]. A total of 400,000 kidney cancer cases were diagnosed around the world in 2018, and 102,000 deaths were recorded in 2002 [5]. In the lesion diagnosis and treatment of kidney tumors, contrast-enhanced computer tomography (CT) shows the morphometry of a kidney tumor. The automatic semantic segmentation of kidneys and kidney tumors in CT imaging, which is the preferred image modality in this case, automatically quantifies the morphometric features in a wide range. It may provide a better solution for the diagnosis and treatment planning of kidney tumors. However, the existence of noise, structural variability, complexity of the 3D spatial tumor features, spatial variability of large-scale features, fractional volume effects, and resemblance of adjacent organs with tumors make this automatic segmentation very challenging [3]. Similarly, CT has been used for the diagnosis of liver tumors. The accurate segmentation of liver and liver tumors from CT images is a critical prior task before any surgical intervention or before choosing an optimal approach for treatment [6].

* Corresponding author.
E-mail address: Abdul.Qayyum@u-bourgogne.fr (A. Qayyum).

https://doi.org/10.1016/j.compbiomed.2020.104097
Received 20 July 2020; Received in revised form 25 October 2020; Accepted 25 October 2020
Available online 28 October 2020



Deep Learning-Based Gait Recognition Using Smartphones in the Wild			
收錄期刊	*IEEE Transactions on Information Forensics and Security*	期刊等級	**SCIE**
Impact Factor (2020)	7.178	Rank (2020)	Q1
全文連結	雲端連結		

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, 2020 1

Deep Learning-Based Gait Recognition Using Smartphones in the Wild

Qin Zou, Yanling Wang, Qian Wang, Yi Zhao, Qingquan Li

https://github.com/qinnzou/Gait-Recognition-Using-Smartphones

Abstract—Compared to other biometrics, gait is difficult to conceal and has the advantage of being unobtrusive. Inertial sensors, such as accelerometers and gyroscopes, are often used to capture gait dynamics. These inertial sensors are commonly integrated into smartphones and are widely used by the average person, which makes gait data convenient and inexpensive to collect. In this paper, we study gait recognition using smartphones in the wild. In contrast to traditional methods, which often require a person to walk along a specified road and/or at a normal walking speed, the proposed method collects inertial gait data under unconstrained conditions without knowing when, where, and how the user walks. To obtain good person identification and authentication performance, deep-learning techniques are presented to learn and model the gait biometrics based on walking data. Specifically, a hybrid deep neural network is proposed for robust gait feature representation, where features in the space and time domains are successively abstracted by a convolutional neural network and a recurrent neural network. In the experiments, two datasets collected by smartphones for a total of 118 subjects are used for evaluations. The experiments show that the proposed method achieves higher than 93.5% and 93.7% accuracies in person identification and authentication, respectively.

Index Terms—Gait recognition, inertial sensor, person identification, convolutional neural network, recurrent neural network.

I. Introduction

BIOMETRICS refers to the automatic identification of a person based on his or her physiological or behavioral characteristics. With the increasing demand for person identification and verification in the era of big data and artificial intelligence, the research and development of biometric systems have attracted broad attention from both academia and industry. Many biometrics, e.g., the fingerprint, iris, face, and voice etc., have been implemented commercially. Some of these biometrics are obtrusive to users as they require the cooperation of users to collect the data. For example, users are asked to place a finger on a device to have their fingerprints captured or to look at a camera close enough to have their irises imaged. In such cases, a user may feel offended and easily realizes that his/her identity is being checked. Meanwhile, some biometrics are easily forged and attacked. For example, face recognition can be cheated by using an image or a video of the target face [1], [2]. As a result, less obtrusive and more robust biometrics are currently in great demand.

Unobtrusiveness is especially important for a biometric system that must work in a discrete manner, e.g., to recognize the identity of a person but not to let him/her know he/she is being identified. Among various biometrics, gait not only satisfies the requirement of being unobtrusive but also is more difficult to conceal [3]–[6].

Gait biometric involves identifying a person based on his/her walking characteristics. Generally, gait recognition can be performed on two types of data: a sequence of images (e.g., from a video), or an inertial gait time series generated by inertial sensors. If gait images or inertial gait time series can be captured, we can perform gait recognition and hence person identification.

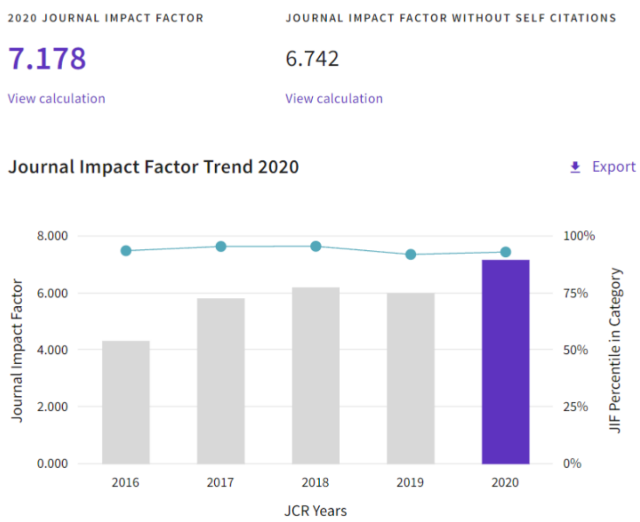
For vision-based methods, Osaka University provided the very early and effective work [7], where the gait data is collected on thousands of subjects. With the development of advanced semantic segmentation algorithms [8], [9], the segmentation of a walking person from an arbitrary background has become possible. Even in cluttered backgrounds, accurate silhouette can be extracted [10], which leads to the successful gait recognition in complex environments [11]. Besides, some studies bypass the use of silhouette images [12]–[14], e.g., using pose estimation.

Inertia-based methods, inertial sensors, such as accelerometers and gyroscopes, are used to record the inertial data generated by the movement of a walking body. These inertial data capture the gait dynamics in a general manner and have been shown to be useful for extracting walking patterns [15]. In the past decade, multiple methods for inertia-based gait recognition have been developed [16]–[20]. However, most of them require inertia sensors to be fastened to specific joints of the human body, which is inconvenient for the gait-data collection.

In real-world application scenarios, the vision-based gait recognition is often used when closed circuit televisions (CCTVs) are installed around the target subjects, which may be unaware of the observation. Inertia-based approaches, however, can be adopted even without the existence of CCTVs, e.g., electronic shackles are put on a prisoner for supervisory control.

As known, many advanced inertial sensors, including accelerometers and gyroscopes, are commonly integrated into smartphones nowadays [21], [22]. So it is very convenient and cheap to collect the interial gait data, and these benefits have inspired a number of solutions to adopt smartphones for gait recognitions [23]–[25]. Smartphone-based gait recognition has many demands, e.g., person identification and authentication. To achieve person identification with high

Q. Zou and Y. Zhao are with the School of Computer Science, Wuhan University, Wuhan 430072, P.R. China (E-mails: qzou@whu.edu.cn, yizhaowhu@whu.edu.cn).
Y. Wang and Q. Wang are with the School of Cyber Science and Engineering, Wuhan University, Wuhan 430079, P.R. China (E-mails: yanling@whu.edu.cn, qianwang@whu.edu.cn).
Q. Li is with Shenzhen Key Laboratory of Spatial Smart Sensing and Service, Shenzhen University, Guangdong 518060, P.R. China (E-mail: liqq@szu.edu.cn).



四、論文發表

(以下論文本文 請參閱附錄一)

◆ SCI Paper

期刊	*Measurement*
論文題目	**Based on LCNet Model and Multi-Feature to Evaluate Gait Data**
論文作者	I Miao Chen, Pin Yu Yeh, Ting Chi Chang, **Ya Chu Hsieh**, Chiun Li Chin*
論文狀態	Under Review

◆ EI Paper

期刊	*Biomedical Engineering: Applications, Basis and Communications*
論文題目	**Calcification clusters and lesions analysis in mammogram using multi-architecture deep learning algorithms**
論文作者	Hao-Hung Tsai, Chia-Shin Wei, **Ya-Chu Hsieh**, I-Miao Chen, Pin-Yu Yeh, Darren Shih, Chiun-Li Chin*
論文狀態	Accepted

◆ Journal Paper

期刊	*Chun Shan Medical Journal (CSMJ)*
論文題目	**GAN-based SSD Segmentation Algorithm to Assist the Character Recognition of Seven-Segment Display Digits**
論文作者	Jung-Mao Lu, Chiun-Li Chin, I-Miao Chen, **Ya-Chu Hsieh**, Ting-Chi Chang, Pin-Yu Yeh*
論文狀態	Accepted

◆ Conference Paper

研討會	*2020 International Conference on Fuzzy Theory and Its Applications (**iFUZZY 2020**)*
論文題目	**Combining VGG16, Mask R-CNN and Inception V3 to identify the benign and malignant of breast microcalcification clusters**
論文作者	**Ya-Chu Hsieh**, Chiun-Li Chin, Chia-Shin Wei, I-Miao Chen, Pin-Yu Yeh, Ru-Jiun Tseng
口頭發表	2020/11/05

研討會	*2020 International Conference on Fuzzy Theory and Its Applications (**iFUZZY 2020**)*
論文題目	**Based on Mask R-CNN Tooth Position Labeling And Periodontal Disease Identification**
論文作者	Chiun-Li Chin, Ming-Chen Hsu, I-Miao Chen, Pin-Yu Yeh, Ting-Ya Chang, **Ya-Chu Hsieh**
口頭發表	2020/11/05

研討會	*2020 International Automatic Control Conference (**CACS 2020**)*
論文題目	**The analysis of high-risk group of angina recurrence using smart health-box and deep learning algorithm**
論文作者	Chiun-Li Chin, Ting-Ya Chang, Pin-Yu Yeh, **Ya-Chu Hsieh**, I-Miao Chen
口頭發表	2020/11/07

研討會	*第十七屆全國電子設計創意競賽暨學術研討會*
論文題目	**結合空拍機與 TensorRT Pose 之訓練報告能力系統**
論文作者	秦群立、陳怡妙、葉品郁、**謝雅竹**、張婷淇、賴彥名、丁敬訓
論文狀態	Accepted

研討會	*第十七屆全國電子設計創意競賽暨學術研討會*
論文題目	**基於 DL-RMS 多重架構驗證群聚鈣化點之特徵抽取、篩選及分類方法**
論文作者	秦群立、陳怡妙、葉品郁、**謝雅竹**、張婷淇、陳婷、李杰祐
論文狀態	Accepted

◆ 論文發表合影紀錄

✓ iFUZZY2020、第十七屆全國電子設計創意競賽暨學術研討會

五、科技部大專生研究計畫

計畫名稱	基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類輔助診斷系統
研究學生	謝雅竹
指導教授	徐麗蘋 老師
計畫編號	110-2813-C-040-096-E
研究期間	2021/07/01~2022/02/28

六、 大學競賽優異表現

競賽總覽

主辦單位	競賽名稱及組別	名次
中華民國模糊學會	「2020 International Conference on Fuzzy Theory and Its Applications」 Best Conference Paper Award	**Second Place**
教育部 & 經濟部	「2020 第 25 屆大專校院資訊應用服務創新競賽」 資訊應用組八(IP8-04)	**第一名 (7.69%)**
	「2020 第 25 屆大專校院資訊應用服務創新競賽」 Open Data 創意應用開發組(Open Data-13)	**第一名 (7.69%)**
	「2020 第 25 屆大專校院資訊應用服務創新競賽」 資訊應用組六(IP6-09)	**第三名 (23.07%)**
	「2020 第 25 屆大專校院資訊應用服務創新競賽」 教育開放資料組(EDUOD-06)	佳作 (30.76%)
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 資訊應用組七(IP7-04)	**第三名 (23.07%)**
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 資訊應用組六(IP6-02)	**第三名 (23.07%)**
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 大桃園資料競技場創新應用組(TY Data Arena-07)	佳作 (30.76%)
	「2020 第 24 屆大專校院資訊應用服務創新競賽」 資訊應用組十(IP10-12)	**人氣獎 (7.6%)**
教育部	「2020 全國大專校院智慧創新暨跨域整合創作競賽」 物聯網組	**跨域整合特別獎 (4.54%)**
	「第二屆智慧專題聯網專題實作競賽」 智慧醫療或智慧農業應用	**金牌 (6.67%)**
		銀牌 (13.33%)
高雄科技大學	「2021 第十七屆全國電子設計創意競賽暨學術研討會」 學術研討組-通訊領域	**優秀論文獎**
	「2021 第十七屆全國電子設計創意競賽暨學術研討會」 學術研討組-醫電領域	**最佳論文獎**
	「2020 第十六屆全國電子設計創意競賽」 大專資通類	**第二名 (6.06%)**
	「2020 第十六屆全國電子設計創意競賽」 大專電子類	佳作 (5.79%)

中山 醫學大學	「2019 中山醫學大學全國大專校院創新、創意及創業競賽」 創新科技組	第二名 (28.57%) 佳作 (57.14%)

競賽	**2020 International Conference on Fuzzy Theory and Its Applications**
組別	**Best Conference Paper Award**
題目	**「Dentition Labeling in Panoramic Radiographs Using Generative Adversarial Network」**
名次	**Second Place**


2020 International Conference on Fuzzy Theory and Its Applications

BEST PAPER AWARD

Second Place

Presented to

Chiun-Li Chin, Ming-Chen Hsu, I-Miao Chen, Pin-Yu Yeh, Ting-Ya Chang, Ya-Chu Hsieh

For paper:

Based on Mask R-CNN Tooth Position Marking And Periodontal Disease Identification

Awarded: Nov. 7, 2020

General Chair of iFUZZY 2020

TFSA

President of TFSA

競賽	2020 第 25 屆大專校院資訊應用服務創新競賽
組別	資訊應用組八(IP8-04)
題目	「生衣訊號」
名次	第一名(7.69%)

組 別：IP8資訊應用組
名 次：第一名
作 品：生衣訊號
學 校：中山醫學大學
作 者：謝雅竹、張婷淇、陳怡妙、劉俊圻、陳婷、丁敬訓、賴彥名、李峻瑞

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，「2020第25屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2020"
The First Place is presented to
Authors: Ya-Chu Hsieh, Ting-Chi Chang, I-Miao Chen,
Chun-Chi Liu, Ting Chen , TING CHING-HSUN,
Yan-Ming Lai, Jyun-Ruei Lee
Chung Shan Medical University
for the creation of The smart life jacket for swimmers

經濟部工業局 局長 呂正華
Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau,
Ministry of Economic Affairs, R.O.C.

中華民國資訊管理學會 理事長 廖則竣
Llao, Che-Chen
President
Chinese Society for Information Management

競賽	2020 第 25 屆大專校院資訊應用服務創新競賽
組別	Open Data 創意應用開發組(Open Data-13)
題目	「基於 Mask R-CNN 之評估聲帶病灶指標產生系統」
名次	第一名(7.69%)


第25屆 2020
大專校院資訊應用服務創新競賽
InnoServe Awards
獎狀

組 別：Open Data創意應用開發組
名 次：第一名
作 品：基於MaskR-CNN之評估聲帶病灶指標產生系統
學 校：中山醫學大學
作 者：葉品郁、陳怡妙、謝雅竹、張婷淇、林俊丞、劉俊圻、李杰祐、林嘉俊

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，行政院科技會報辦公室共同主辦之「2020第25屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2020"
The First Place is presented to
Authors: Pin-Yu Yeh, I-Miao Chen, Ya-Chu Hsieh, Ting-Chi Chang, Jun-Cheng Lin, Chun-Chi Liu, Chieh-Yu Li, Chia-Chun Lin
Chung Shan Medical University
for the creation of The system of vocal cord lesion index generation based on Mask R-CNN

競賽團隊資訊

呂正華 Richard Leu
經濟部工業局 局長 呂正華
Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau,
Ministry of Economic Affairs, R.O.C.

廖則竣
中華民國資訊管理學會 理事長 廖則竣
Liao, Che-Chen
President
Chinese Society for Information Management

競賽	2020 第 25 屆大專校院資訊應用服務創新競賽
組別	資訊應用組六(IP6-09)
題目	「基於 Mask R-CNN 之評估聲帶病灶指標產生系統」
名次	第三名(23.07%)

第25屆 2020

大專校院資訊應用服務創新競賽

頒獎字號(Certificate No.) 獎字第109CSIM獎字第0177

InnoServe

Awards

獎狀

組 別：IP6資訊應用組

名 次：第三名

作 品：基於MaskR-CNN之評估聲帶病灶指標產生系統

學 校：中山醫學大學

作 者：葉品郁、陳怡妙、謝雅竹、張婷淇、林俊丞、劉俊圻、李杰祐、林嘉俊

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，「2020第25屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2020"

The Third Place is presented to

Authors: Pin-Yu Yeh, I-Miao Chen, Ya-Chu Hsieh,
Ting-Chi Chang, Jun-Cheng Lin, Chun-Chi Liu,
Chieh-Yu Li, Chia-Chun Lin

Chung Shan Medical University

for the creation of The system of vocal cord lesion index generation based on Mask R-CNN

競賽國際資訊

呂正華 Richard Leu

經濟部工業局 局長 呂正華

Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau,
Ministry of Economic Affairs, R.O.C.

廖則竣

中華民國資訊管理學會 理事長 廖則竣

Liao, Che-Chen
President
Chinese Society for Information Management

競賽	2020 第 25 屆大專校院資訊應用服務創新競賽
組別	教育開放資料組(EDUOD-06)
題目	「We can help」
名次	佳作(30.76%)

第25屆 2020
大專校院資訊應用服務創新競賽

頒獎字號(Certificate No.) 獎字第109CSIM獎字第0703

InnoServe Awards
獎狀

組　別：教育開放資料組
名　次：佳作
作　品：We can help
學　校：中山醫學大學
作　者：陳怡妙、葉品郁、謝雅竹、張婷淇、林俊丞、賴彥名、林嘉俊、陳昱翔

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，「2020第25屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2020"
The Excellent Work is presented to
Authors: I-Miao Chen, Pin-Yu Yeh, Ya-Chu Hsieh,
Ting-Chi Chang, Jun-Cheng Lin, Yan-Ming Lai,
Chia-Chun Lin, Yu-Hsiang Chen
Chung Shan Medical University
for the creation of We can help

競賽團隊資訊

呂正華 Richard Leu
經濟部工業局 局長 呂正華
Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau,
Ministry of Economic Affairs, R.O.C.

廖則竣
中華民國資訊管理學會 理事長 廖則竣
Liao, Che-Chen
President
Chinese Society for Information Management

競賽	2020 第 24 屆大專校院資訊應用服務創新競賽
組別	資訊應用組七(IP7-04)
題目	「牙位標記及根管治療判斷」
名次	第三名(23.07%)

2019第24屆
大專校院資訊應用服務創新競賽
頒獎字號(Certificate No.) 獎字第108CSIM獎字第0194

InnoServe
Awards
獎狀

組　別：IP7資訊應用組
名　次：第三名
作　品：牙位標記及根管治療判斷
學　校：中山醫學大學
作　者：許銘真、魏嘉芯、劉宇捷、林永隆、林若瑋、莊庭碩、陳怡妙、謝雅竹

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，「2019第24屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2019"
The Third Place is presented to
Authors: Ming-Chen Hsu, Chia-Shin Wei, Yu-Chieh Liu,
Yong-Long Lin, Jo-Wei Lin, Ting-Shou Chuang,
I-Miao Chen, Ya-Chu Hsieh
Chung Shan Medical University
for the creation of dentition labeling and root canal identification

呂正華 Richard Leu
經濟部工業局 局長 呂正華
Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau,
Ministry of Economic Affairs, R.O.C.

廖則竣
中華民國資訊管理學會 理事長 廖則竣
Llao, Che-Chen
President
Chinese Society for Information Management

競賽	**2020 第 24 屆大專校院資訊應用服務創新競賽**
組別	**資訊應用組六(IP6-02)**
題目	**「不礙坐，愛座」**
名次	**第三名(23.07%)**

2019第24屆
大專校院資訊應用服務創新競賽

頒獎字號(Certificate No.) 獎字第108CSIM獎字第0172

InnoServe
Awards
獎狀

組 別：IP6資訊應用組
名 次：第三名
作 品：不礙坐，愛座
學 校：中山醫學大學
作 者：林永隆、劉宇捷、魏嘉芯、莊庭碩、吳述蓉、謝雅竹、張婷雅、許栢維

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，「2019第24屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2019"
The Third Place is presented to
Authors: Yong-Long Lin, Yu-Chieh Liu, Chia-Shin Wei,
Ting-Shou Chuang, Su-Jung Wu, Ya-Chu Hsieh,
Ting-Ya Chang, Po-Wei Hsu
Chung Shan Medical University
for the creation of AiZo

呂正華 Richard Leu
經濟部工業局 局長 呂正華
Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau,
Ministry of Economic Affairs, R.O.C.

廖則竣
中華民國資訊管理學會 理事長 廖則竣
Llao, Che-Chen
President
Chinese Society for Information Management

競賽	**2020 第 24 屆大專校院資訊應用服務創新競賽**
組別	**大桃園資料競技場創新應用組(TY Data Arena-07)**
題目	**「不礙坐，愛座」**
名次	**佳作(30.76%)**

2019第24屆
大專校院資訊應用服務創新競賽
頒獎字號(Certificate No.) 獎字第108CSIM獎字第0685
InnoServe
Awards
獎狀

組 別：大桃園資料競技場創新應用組
名 次：佳作
作 品：不礙坐，愛座
學 校：中山醫學大學
作 者：林永隆、劉宇捷、魏嘉芯、莊庭碩、吳述蓉、謝雅竹、張婷雅、許栢維

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，桃園市政府資訊科技局共同主辦之「2019第24屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2019"
The Excellent Work is presented to
Authors: Yong-Long Lin, Yu-Chieh Liu, Chia-Shin Wei, Ting-Shou Chuang, Su-Jung Wu, Ya-Chu Hsieh, Ting-Ya Chang, Po-Wei Hsu
Chung Shan Medical University
for the creation of AiZo

呂正華 Richard Leu
經濟部工業局 局長 呂正華
Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau, Ministry of Economic Affairs, R.O.C.

廖則竣
中華民國資訊管理學會 理事長 廖則竣
Llao, Che-Chen
President
Chinese Society for Information Management

陳尚榮
桃園市政府資訊科技局 局長 陳尚榮
Kenny Chen
Director-General
Department of Information Technology, Taoyuan City Government

競賽	2020 第 24 屆大專校院資訊應用服務創新競賽
組別	資訊應用組十(IP10-12)
題目	「交通安全你我他」
名次	人氣獎(7.6%)

2019第24屆
大專校院資訊應用服務創新競賽

頒獎字號(Certificate No.) 獎字第108CSIM獎字第0970

InnoServe
Awards
獎狀

組　別：資訊應用組
名　次：人氣獎
作　品：交通安全你我他
學　校：中山醫學大學
作　者：劉宇捷、林永隆、莊庭碩、林若瑋、吳述蓉、謝雅竹、許栢維、許銘真

參加經濟部工業局、教育部資訊及科技教育司及中華民國資訊管理學會主辦，「2019第24屆大專校院資訊應用服務創新競賽」，表現傑出，特頒此狀，以資獎勵。

Certificate of Award

"International ICT Innovative Services Awards 2019"
The Best Popularity Award is presented to
Authors: Yu-Chieh Liu, Yong-Long Lin, Ting-Shou Chuang,
Jo-Wei Lin, Su-Jung Wu, Ya-Chu Hsieh,
Po-Wei Hsu, Ming-Chen Hsu
Chung Shan Medical University
for the creation of traffic safety

呂正華 Richard Leu
經濟部工業局 局長 呂正華
Dr. Jang-Hwa Leu
Director General
Industrial Development Bureau,
Ministry of Economic Affairs, R.O.C.

廖則竣
中華民國資訊管理學會 理事長 廖則竣
Llao, Che-Chen
President
Chinese Society for Information Management

競賽	2020 全國大專校院智慧創新暨跨域整合創作競賽
組別	物聯網組
題目	「明鏡照形」
名次	跨域整合特別獎(4.54%)


臺教資（二）字第1090162401號

教育部獎狀
Ministry of Education
2020 Intelligent Innovation and Interdisciplinary Creation Contest

中山醫學大學醫學資訊學系 謝雅竹 葉品郁 丁敬訓 陳婷
作品：明鏡照形
參加2020年全國大專校院智慧創新暨跨域整合創作競賽
經評定為 物聯網組 跨域整合特別獎
特頒此狀 以資鼓勵

This certificate is awarded to Ya-Chu Hsieh, Pin-Yu Yeh, Jing-Xun Ding and Ting Chen from the Department of Medical Informatics at Chung Shan Medical University for achieving the Special Award for Cross-disciplinary Integration in the Internet of Things Division of the 2020 National Intelligent Innovation and Interdisciplinary Creation Contest in Universities and Colleges.

部長 潘文忠
Wen-Chung Pan
Wen-Chung Pan
Minister of Education
Republic of China (Taiwan)

中 華 民 國 109 年 12 月 11 日

競賽	第二屆智慧專題聯網專題實作競賽
組別	智慧醫療或智慧農業應用
題目	「基於 GAN 深度學習演算法之牙位標記與牙齒重疊判斷系統」
名次	金牌


臺教資（二）字第1090117822號

教育部獎狀
Ministry of Education
Intelligent IOT Project Competition Award

中山醫學大學 醫學資訊學系 許銘真 劉宇捷 陳怡妙 謝雅竹 林俊丞 丁敬訓 賴彥名
榮獲第二屆智慧聯網專題實作競賽 智慧醫療或智慧農業應用 金牌
特頒此狀 以資鼓勵

This Gold Prize Award is presented to Ming-Chen Hsu, Yu-Chieh Liu, I-Miao Chen, Ya-Chu Hsieh, Jyun-Cheng Lin, Jing-Xun Ding and Yan-Ming Lai from Department of Medical Informatics at Chung Shan Medical University for excellent achievement in the IoT Application for Smart Healthcare/Agriculture of the 2nd Intelligent IOT Project Competition.

部長 潘文忠

Wen-Chung Pan
Minister of Education
Republic of China (Taiwan)

中 華 民 國 109 年 8 月 25 日

競賽	**第二屆智慧專題聯網專題實作競賽**
組別	**智慧醫療或智慧農業應用**
題目	**「BI-CNN 偵測婦女乳房群聚鈣化點及預測鈣化點良惡系統」**
名次	**銀牌**

臺教資（二）字第1090117822號

教育部獎狀

Ministry of Education
Intelligent IOT Project Competition Award

中山醫學大學 醫學資訊學系 魏嘉芯 林永隆 葉品郁 李杰祐 謝雅竹 張婷雅 劉俊圻
榮獲第二屆智慧聯網專題實作競賽 智慧醫療或智慧農業應用 銀牌
特頒此狀 以資鼓勵

This Silver Prize Award is presented to Chia-Shin Wei, Yong-Long Lin, Pin-Yu Yeh, Chieh-Yu Li, Ya-Chu Hsieh, Ting-Ya Chang and Chun-Chi Liu from Department of Medical Informatics at Chung Shan Medical University for excellent achievement in the IoT Application for Smart Healthcare/Agriculture of the 2nd Intelligent IOT Project Competition.

部長 潘文忠

Wen-Chung Pan
Minister of Education
Republic of China (Taiwan)

中 華 民 國 109 年 8 月 25 日

競賽	2021 第十七屆全國電子設計創意競賽暨學術研討會
組別	學術研討組-通訊領域
題目	「結合空拍機與 TensorRT Pose 之訓練報告能力系統」
名次	優秀論文獎

國立高雄科技大學
獎狀

秦群立、陳怡妙、葉品郁、謝雅竹
張婷淇、賴彥名、丁敬訓
以作品「結合空拍機與 TensorRT Pose 之訓練報告能力系統」(編號:B06)
參加 2021 第十七屆全國電子設計創意競賽暨學術研討會
榮獲 優秀論文獎

校長 楊慶煜

中 華 民 國 110 年 3 月 27 日

競賽	**2021 第十七屆全國電子設計創意競賽暨學術研討會**
組別	**學術研討組-醫電領域**
題目	**「基於 DL-RMS 多重架構驗證群聚鈣化點之特徵抽取、篩選及分類方法」**
名次	**最佳論文獎**

國立高雄科技大學

獎狀

陳怡妙、葉品郁、謝雅竹、張婷淇、陳婷、李杰祐

以作品「基於 **DL-RMS** 多重架構驗證群聚鈣化點之特徵抽取、篩選及分類方法」

(編號: D07)

參加 2021 第十七屆全國電子設計創意競賽暨學術研討會

榮獲 最佳論文獎

特頒獎狀，以茲鼓勵。

電子工程系主任 楊素華

中 華 民 國 110 年 3 月 27 日

競賽	2020 第十六屆全國電子設計創意競賽
組別	大專資通類
題目	「改鞋歸正」
名次	第二名(6.06%)

國立高雄科技大學
獎狀

秦群立老師指導
陳怡妙、葉品郁、謝雅竹、張婷雅同學
以作品「改鞋歸正」(編號:UI07)
參加 2020 第十六屆全國電子設計創意競賽
榮獲 亞軍
特頒獎狀，以茲鼓勵。

主持人 楊素華

中華民國 109 年 4 月 11 日

競賽	2020 第十六屆全國電子設計創意競賽
組別	大專電子類
題目	「安枕無憂」
名次	佳作(5.79%)

國立高雄科技大學
獎狀

秦群立老師指導
張婷雅、謝雅竹、葉品郁、陳怡妙同學
以作品「安枕無憂」(編號:UE19)
參加 2020 第十六屆全國電子設計創意競賽
榮獲 佳作
特頒獎狀，以茲鼓勵。

主持人 楊素華

國立高雄科技大學電子工程系

中 華 民 國 1 0 9 年 4 月 1 1 日

競賽	**2019 中山醫學大學全國大專校院創新、創意及創業競賽**
組別	**創新科技組**
題目	**「一鏡到底」**
名次	**第二名(28.57%)**

BEST

中山醫學大學全國大專院校創新、創意及創業競賽

創新科技組

第二名

參賽學校：中山醫學大學
團隊主題：一鏡到底
團隊名稱：一鏡到底
團隊成員：謝雅竹、張婷雅、吳逑蓉、陳怡妙、葉品郁
指導老師：秦群立

參加 中山醫學大學全國大專院校創新、創意及創業競賽
表現優異獲選【第二名】

特頒此狀，以資鼓勵。

校長 呂克桓

中山醫學大學

中華民國一〇八年十二月六日

競賽	**2019 中山醫學大學全國大專校院創新、創意及創業競賽**
組別	創新科技組
題目	「基於深度學習辨識血糖及血壓機之七段顯示器」
名次	佳作(57.14%)

七、大學學業成績優異表現

大學學業成績排名

學期		成績	全系排名	全系累積排名
大三	下學期	94.68	第 5 名	**第 8 名**
	上學期	94.21	第 2 名	第 8 名
大二	下學期	91.48	第 8 名	第 10 名
	上學期	90.56	第 12 名	第 14 名
大一	下學期	86.25	第 17 名	第 16 名
	上學期	86.5	第 15 名	第 15 名

榮獲成績優異獎金證明


中山醫學大學

獎　狀

中山醫學獎字第 1102101114-289 號

查 醫學資訊學系三年級

謝雅竹 同學於 109 學年度第一學期敦品勵學成績表現優異殊堪嘉許榮獲本校績優獎學金特頒此狀以資鼓勵

校長 黃 建 寧

中華民國 110 年 05 月 24 日

大學各科成績排名

年級	課程名稱	學分	成績	排名(名次/修課人數)
三	醫用行動軟體設計	3	99	2/43
	優質網路健康照護	2	97	2/43
	醫院資訊系統實務	2	93	9/42
	實習	3	96	2/41
	電子病歷實務	2	95	3/45
	醫用資訊系統分析與設計	3	97	1/46
	大數據管理系統	3	88	17/45
	圖形識別概論	3	99	1/16
	深度學習	2	98	3/34
	離散數學	3	98	1/48
二	醫用程式語言	3	95	12/48
	醫學影像處理	3	99	1/23
	進階程式語言	2	96	24/44
	醫學影像傳輸系統	2	91	17/100
	醫用電子學	2	99	1/16
	臨床醫學網路暨資訊安全	3	99	1/43
	機率與統計	3	79	17/50
	資料結構	3	88	25/45
	作業系統	3	84	31/51
	生醫訊號處理概論	3	100	1/21
一	物件導向程式設計	3	99	1/52
	線性代數	2	70	26/64
	.NET 程式設計	3	98	6/46
	微積分	2	92	12/50
	計算機組織與結構	3	81	20/50
	醫學資訊概論	3	91	17/56
	資訊倫理	2	90	14/45
	管理數學	2	84	9/24

八、 專題研究

中　山　醫　學　大　學

醫 學 資 訊 學 系

畢 業 專 題 文 件

基於 AIR 多特徵抽取演算法與統計特徵篩選之

惡性乳房鈣化點分類輔助診斷系統(目錄頁碼要改)

專題編號：

專題學生：謝雅竹

指導教授：秦群立 博士

中 華 民 國 一 百 一 十 年 九 月

基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類輔助診斷系統

專題學生：謝雅竹

指導教授：秦群立

摘 要

醫師在使用乳房攝影影像診斷病人病灶時，會先判斷影像中是否有鈣化點群聚的地方，若有則會依據鈣化點的形狀、大小及分布來初步判定鈣化點的良惡，當初步判定群聚鈣化點為惡性時，會建議患者接受穿刺切片檢測做進一步的確認，然而穿刺切片檢測為侵入式的檢測，易造成病患傷口疼痛，因此許多患者接受度不高，而導致錯失治療的時機。另外，乳房的惡性鈣化點可進一步分為三大類：原位癌、前期癌及癌症，這三類治療方法皆有所差異，醫師若能夠詳細了解病患的病況，即可針對病患的狀況，給予出合適的治療方案。因此，本研究與中山醫學大學附設醫院的醫學影像部醫師合作，提出「基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類預測系統」，透過乳房穿刺影像及醫囑得知鈣化點群聚的位置及其詳細的惡性的類別，將這些資料作為訓練的目標資訊，並且利用 Gathering CNN 偵測乳房攝影影像中鈣化點群聚處，接著將群聚鈣化點處的影像輸入至 Mask R-CNN 切割出鈣化點的區域，針對每一張鈣化點的影像，運用影像組學(Radiomics)特徵抽取法及深度學習的自動編碼器(AutoEncoder)抽取出惡性鈣化點的特徵後，使用獨立樣本 t 檢定、ANOVA 及 MANOVA 方法進行最佳特徵選取，以找出能代表鈣化點特徵的數值，最後透過深度神經網路(Deep Neural Network, DNN)、支援向量機(Support Vector Machine, SVM)及隨機森林法(Random Forest)將惡性鈣化點分類為原位癌、前期癌或癌症。SVM、Random Forest 及 DNN 分類惡性鈣化點的準確率分別為 90.56%、91.82%及 93.71%，均高於 90%，證明本研究提出的特徵抽取方法與最佳特徵篩選法有助於提升模型的分類準確率，並能輔助醫師診斷，提供更多的資訊作為參考依據，也希望透過上述方法，使病患不須接受穿刺切片檢測，也能得知自身的病況，盡早接受治療。

關鍵字：乳房攝影影像、影像組學特徵抽取法、深度學習、最佳特徵選取、惡性鈣化點分類

Classification aided diagnosis system in malignant microcalcification clusters based on AIR multi-feature extraction algorithm and statistical feature selection

Abstract

When using mammograms to identify patient's lesions, radiologists will first determine whether there are microcalcification clusters in the image. If there are, radiologists will preliminarily determine the benign and malignant of the microcalcification clusters based on the shape, size and distribution of the microcalcifications in the clusters. When it is preliminarily determined that the microcalcifications clusters are malignant, the patient will be recommended to undergo needle localization surgical biopsy for further confirmation. However, the needle localization surgical biopsy is an invasive detection, which is easy to cause pain in the patient's wound, so many patients have lower acceptance of this test, which leads to treatment opportunities are missed. In addition, the malignant microcalcification clusters can be further divided into three categories: carcinoma in situ, pre-carcinoma and cancer. The treatment methods of these three types of diseases are different. If radiologists can understand the patient's condition in detail, he can give an appropriate treatment plan based on the patient's condition. Thus, this research is supported by doctors in department of Medical Imaging, Chung Shan Medical University Hospital, and we propose "Classification aided diagnosis system in malignant microcalcification clusters based on AIR multi-feature extraction algorithm and statistical feature selection." The system learns the location of microcalcification clusters and their detailed malignant types through images of needle localization surgical biopsy and doctor's medical orders, and using these data as training target information. Moreover, we use Gathering CNN to detect the microcalcification clusters in the mammograms, and then input them into the Mask R-CNN to segment the microcalcifications. For each image of the microcalcifications clusters, radiomics feature extraction method and AutoEncoder are used to extract the features of the malignant microcalcifications clusters, and then the independent sample t test, ANOVA and MANOVA methods are used for optimal feature selection to find the values that can represent the characteristics of the microcalcifications clusters. Finally, the malignant microcalcification clusters are classified as carcinoma in situ, pre-carcinoma, or cancer through deep neural network (DNN), support vector machine (SVM) and Random Forest. The accuracy of SVM, Random Forest and DNN to classify malignant microcalcification clusters are 90.56%, 91.82% and 93.71%, respectively, which are all higher than 90%, proving that the feature extraction method and the optimal feature selection method proposed in this research can help improve the classification performance of the model, and can provide radiologists with more information to diagnosis. It is hoped that through the above methods, patients do not need to undergo needle localization surgical biopsy, but can also know their condition and receive treatment as soon as possible.

Keywords: mammograms, radiomics feature extraction method, deep learning, optimal feature selection, malignant microcalcification clusters classification

目錄

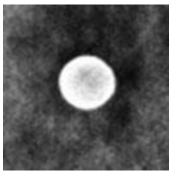
圖目錄

表目錄

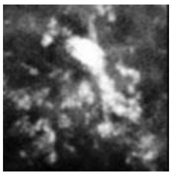
第一章　緒論

1.1 研究動機

乳癌為台灣女性好發率第一名之癌症，衛生署每 2 年補助 45-69 歲的婦女做 1 次的乳房攝影檢查，可見其重視。而目前最主要的乳房篩檢方法為乳房攝影，由醫師觀察乳房影像中微鈣化點的形狀、大小及分布來初步判定微鈣化點為良性或惡性，為了實際了解醫師判定乳房影像的狀況及需求，本研究與中山醫學大學附設醫院的醫學影像部醫師合作，實際訪問醫師後得知，良性鈣化點形狀上多為圓形，分布較為均勻或對稱，且在影像中的大小通常偏大；惡性鈣化點形狀上通常具有不規則的外型，分布較為密集，且在影像中的大小通常偏小，如圖 1.1.1 所示。



(a)



(b)

圖 1.1.1、(a)為良性鈣化點，(b)為惡性鈣化點

當醫師初步判定鈣化點為惡性時，會通知患者接受穿刺切片檢測，以便進一步確認該鈣化點的良惡，若為惡性，可進一步將其分為三類：原位癌、前期癌及癌症。然而，穿刺切片檢測可能會**造成病患傷口疼痛**，且病患往往需**花費一段時間考慮**或者**前往別家醫院檢測**後才願意接受穿刺切片檢測，過程中**耗費許多時間**，可能**導致病情延誤，無法及時的給予治療**。

1.2 研究目的

現今人工智慧發展蓬勃，可將機器擬作人眼甚至超越人眼來進行影像判讀，搭配影像組學，從病灶中快速抽取大量特徵進行分析，並運用統計學的方法進行特徵篩選，透過此方式可分析隱含於影像後的問題，找出真正影響分類惡性乳房鈣化點的特徵，為臨床診斷提供更多思路，因此本研究透過上述方法開發出「基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類預測系統」，期望透過影像組學及統計學的方法找出對於**分類乳房鈣化點有影響力的特徵**，再使用深度學習的方法將其**分類為原位癌**、**前期癌**或**癌症**，透過本研究所提出的方法，可**避免患者**因難以接受穿刺切片檢測，而**延誤治療**的時機，醫師也可透過本系統從乳房攝影影像中**快速判讀病患的病情**，且本系統透過**大量的資料分析**，提供醫師**可靠的數據結果，可加速診療流程，並即時給予**患者**相對應的治療**。

1.3 章節安排

本論文共分為五個章節，各章節內容簡述如下：

第一章主要介紹研究動機及目的

第二章為相關文獻的探討與回顧

第三章說明本論文所提出的方法及系統功能架構與流程

第四章本論文之實驗結果、系統介面圖展示與效能評估

第五章本論文之結論

第二章　文獻探討與回顧

現今影像組學、統計學及深度學習的方法廣泛應用於醫學影像分析及辨識，本研究參考多篇文獻，取其特長並與多篇論文所提出的想法結合，以達成本研究之目的，本研究將參考的文獻分為四大類，以下將一一進行介紹。

2.1 提取醫學影像特徵之相關研究

在 2017 年，Laura Macías-García 學者等人提出利用自動編碼器(AutoEncoder)深度學習演算法，對原始資料集中經過免疫組織化學染色的組織切片數據進行預處理[1]，以降低數據中雜訊的影響並提高數據質量，獲得具有價值的乳腺癌基因組特徵。此研究證明使用自動編碼器可從醫學影像中獲得具有價值的特徵。

在 2019 年，Losurdo 學者等人提出利用 Radiomics 從 CESM (Contrast-Enhanced Spectral Mammography)影像中提取特徵並分類乳腺之良惡[2]，此研究證明 Radiomics 所提取的紋理特徵有助於區分良性乳腺疾病或是惡性乳腺癌。

在 2019 年，Ibrahim 學者等人針對 Radiomics 應用於臨床醫學領域所面臨的挑戰，以及在臨床決策支持系統中的應用潛力進行回顧與統整[3]，提出使用 Radiomics 提取的特徵未必能獲得高準確率結果的觀點，因此須針對特徵依賴性和相關性進行詳盡的分析。

在 2019 年，Ijurra 學者等人使用 Radiomics 特徵抽取法獲取大腦腫瘤的特徵，再運用 CNN 進行腦腫瘤嚴重程度的分類[4]，在此研究結果中提到，其運算速度提高了 7205%，相較於傳統的方法速度更快，減少更多的運算時間。

在 2019 年，Bizzego 學者等人欲分析出頭頸部腫瘤的分期，使用 Radiomics 及 CNN 抽取頭頸部腫瘤 CT 及 PET 影像的特徵，並將抽取出的特徵分為三大類，分別為由 Radiomics 抽取出的特徵、由 CNN 抽取出的特徵及融合了 Radiomics 及 CNN 抽取出的特徵[5]，再分別運用 ANOVA 進行特徵篩選，結果顯示融合了 Radiomics 及 CNN 抽取出的特徵效果最好，在分辨頭頸部腫了的分期上，準確率高達了 96.5%。

2.2 統計學的方法應用於特徵篩選的相關研究

在 2017 年，Ling 學者等人分別將 global shape descriptor、Hu moment invariant 和 Fourier descriptor 用於抽取脊柱的 MRI 影像中胸、腰和骨的特徵，再運用 ANOVA 選擇特徵[6]，結果證明將 ANOVA 作為特徵選擇的方法可有效選出重要的特徵。

在 2020 年，Daisy Das 學者等人提出將統計學的 MANOVA 方法用於選擇活體組織切片影像中特徵的研究[7]，利用 MANOVA 方法檢驗兒童髓母細胞瘤(Childhood Medulloblastoma, CMB)的活體組織切片中，是否存在具有統計學意義的特徵，以便後續對正常樣本及其各種亞型中的 CMB 進行分類。此研究實際比較進行特徵選擇與未進行特徵選擇前後的分類結果，證明經過特徵選擇步驟再利用 SVM 方法進行兒童髓母細胞瘤分類的準確率相對較高，因此 MANOVA 的統計方法可以有效地作為機器學習的特徵選擇方法。

在 2020 年，Caballo 學者等人提出利用 Unet 深度學習演算法從乳房 CT 影像中分割出良性與惡性腫瘤，再運用 Radiomics 從乳房 CT 影像中抽取大量的特徵，最後利用統計學中的 MANOVA 分析，同時比較 2 個或以上的應變數[8]，此研究主要用於評估分割的腫瘤影像與 Radiomics 所抽取出的特徵之間的差異性，以篩選出主要的特徵。

由於本研究目的為分辨乳房攝影影像中的微小鈣化點，因此本研究參考上述使用統計學方法達成特徵篩選的功能，使用獨立樣本 t 檢定、ANOVA 及 MANOVA 從大量惡性鈣化點特徵中，找出具有代表性的微小鈣化點特徵。

2.3 深度學習的方法應用於醫學影像分類的相關研究

2018 年，Gaurav Makwana 學者等人提出了一種用於乳房鈣化點良惡分類的方法，先運用 GLCM 抽取統計特徵，再運用 SVM 依照抽取的統計特徵進行分類[9]，其分類良性及惡性鈣化點的準確率皆高於 85%以上，另外，在 2019 年，為了將乳房 X 光影像分類為正常與異常，R. Vijayarajeswari 學者等人先使用霍夫變換辨別出乳房 X 光影像中的特徵，再運用 SVM 學習乳房 X 光影像中的特徵並進行分類[10]，從結果可看出，此方法可有效分類異常的乳房 X 光影像，由上述文獻可推論，SVM 在分類醫學影像上皆有不錯的成果，且展現其優秀的泛化能力。

在 2018 年，Zemouri 學者等人提出基於 DNN 的架構，用於預測 Oncotype DX (ODX)乳腺癌檢測管腔 B 型乳癌的復發風險程度[11]，並在測試結果有良好的表現。

醫學數據分類是非常複雜且具有挑戰性的，在 2019 年，Alam 學者等人提出基於特徵排名與特徵選擇策略的醫學數據分類方法[12]，透過此方法所採用的隨機森林分類器將能區分資料集中排名較高

的特徵，以預測出具有類似特徵的疾病。此研究透過在 10 個不同疾病資料集上進行實驗，以證實在臨床案例中，隨機森林分類器在具有相似特徵的疾病分類方面有良好的表現。

由於本研究目的是將惡性乳房鈣化點分類為原位癌、前期癌或癌症，因此本研究參考上述文獻的方法，將 DNN、SVM 及 Random Forest 用於學習惡性乳房鈣化點的特徵並進行分類。

綜合上述文獻回顧，目前尚未有將影像組學及深度學習運用於抽取乳房鈣化點之特徵並進行惡性分類：惡性乳房鈣化點分類為原位癌、前期癌及癌症的研究，但影像組學已廣泛被使用在抽取大量醫學影像的特徵，且透過影像組學所抽取出的特徵，可觀察出更多內含於影像中的特徵，因此本研究希望實現上述文獻所提出的想法，並取其特長，並搭配自行開發的 GUI 介面，建置「基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類預測系統」，以快速輔助醫師判讀乳房影像，及時給出相對應的治療方案。

第三章　研究方法

3.1 系統功能架構

本研究利用 Radiomics、統計學、機器學習與深度學習方法建置出「基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類預測系統」，其中 AIR(AI Radiomics)分為 Radiomics 及 AutoEncoder 特徵抽取法抽取惡性乳房鈣化點特徵；獨立樣本 t 檢定、ANOVA 及 MANOVA 篩選出對於惡性鈣化點分類具有影響力的特徵和 DNN、SVM 及 Random Forest 將乳房鈣化點分類為不同惡性類別，惡性類別分別為原位癌、前期癌及癌症，計畫流程圖如圖 3.1.1 所示。

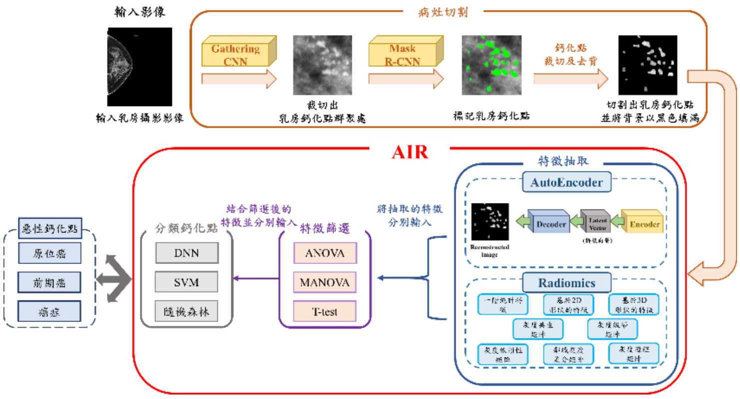
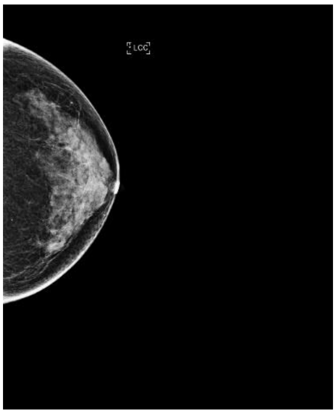


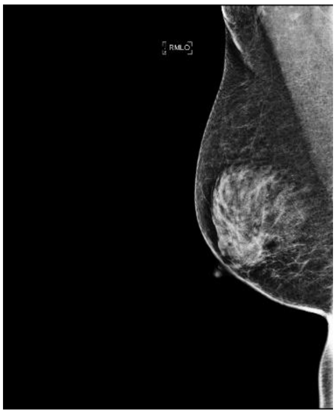
圖 3.1.1、系統流程圖

3.2 輸入乳房攝影影像

本研究與中山醫學大學附設醫院(Chung Shan Medical University Hospital)醫學影像部合作，由專業醫師提供 1586 張乳房攝影影像作為影像資料集，其中包含每位病患的乳房上下側影像(CC view)，及左右側影像(MLO view)，如圖 3.2.1 所示。本研究將此資料集中 80%的影像資料作為訓練影像，而 20%的影像資料作為測試影像。



(a)



(b)

圖 3.2.1、(a)為乳房上下側影像(CC view)，(b)為乳房左右側影像(MLO view)

3.3 使用 Gathering CNN 偵測鈣化點群聚處

Gathering CNN 可分為兩部分，分別為全卷積神經網路(Fully Convolutional Neural Networks, FCN)及 ISODATA(Iterative Self-Organizing Data Analysis Techniques Algorithm)分類法。

為了自動找出乳房攝影影像中的鈣化點群聚處，本研究會先透過醫師提供的細針定位影像得知確切的鈣化點群聚位置，以作為模型分類乳房鈣化點群聚與非群聚影像的依據，將乳房攝影影像以128×128 的大小擷取出鈣化點群聚處。接著為了因應不同設備拍攝影像導致影像的對比度不佳，以及乳房鈣化點的輪廓不明顯而影響模型辨識結果的問題，本研究將乳房攝影影像進行銳利化處理，如圖3.3.1 所示，使鈣化點的邊緣及輪廓更加的明顯。此外，除了包含鈣化點群聚處的影像之外，資料集中還須具有非群聚處的影像，方能精準區分鈣化點群聚處與非群聚處，因此以擷取出鈣化點群聚處的影像數量為標準，再以相同數量擷取影像中非群聚處，如：乳腺、乳頭及標籤等，接著將影像及其群聚鈣化點座標範圍輸入 FCN 模型進行分類群聚與非群聚的訓練，以產生數個候選框(Candidate Box)。最後，由於 ISODATA 分群法擁有自動調整群數的能力，適合用於未知群數的乳房攝影影像，因此本研究使用 ISODATA 對 FCN 所產生的候選框進行分群，並將相同群聚的候選框融合，以找出真正的群聚鈣化點數量及其在影像中的確切位置。

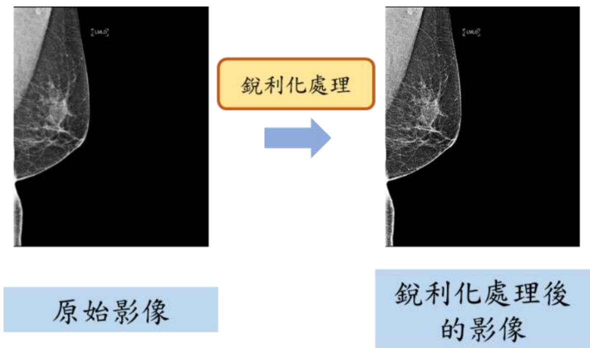


圖 3.3.1、將乳房影像經銳利化處理之示意圖

3.4 使用 Mask R-CNN 切割出鈣化點並分類良惡

在切割出鈣化點群聚處影像後，本研究會將 Gathering CNN 偵測與切割出的群聚鈣化點影像作為 Mask R-CNN 的訓練與測試的目標影像，如圖 3.4.1 (a)所示，接著標記群聚鈣化點影像中的惡性鈣化點特徵，作為 Mask R-CNN 訓練時的目標影像，如圖 3.4.1 (b)所示，最後當 Mask R-CNN 從群聚鈣化點影像中將惡性鈣化點切割出來後，為避免背景干擾導致良惡判斷失準，本研究將背景的部分以黑色填滿，如圖 3.4.1 (c)所示。

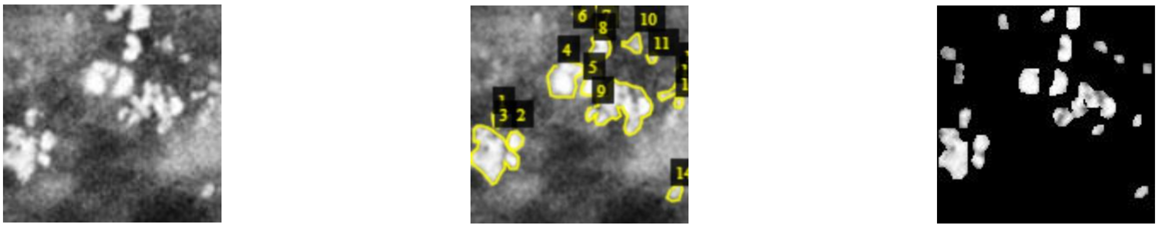


圖 3.4.1、(a)為乳房鈣化點群聚處影像，(b)為 Mask R-CNN 訓練時的目標影像，(c)為將除了鈣化點以外的部分以黑色填滿之影像

3.5 運用 Radiomics 及 AutoEncoder 特徵抽取法抽取惡性鈣化點特徵

在找出乳房鈣化點群聚處後，本研究分別運用 Radiomics 與自動編碼器(AutoEncoder)抽取影像中惡性鈣化點的特徵。Radiomics 是高通量的特徵抽取方法，可從每張鈣化點影像中快速提取大量的資訊，將感興趣區域(Region of Interest, RoI)影像轉為多筆具有高分辨率的特徵空間數據，可反映潛在病理及生理特徵，常應用於疾病檢測、腫瘤分型與評估疾病預後等方向，因此現今已有多篇研究使用 Radiomics 特徵抽取法從醫學影像中抽取大量特徵如腫瘤與乳腺癌特徵等[1-3]。AutoEncoder 是一種非監督式學習演算法，常用於降低高維度資料及移除影像中的雜訊，並保留原始資料中的重要特徵，且在文獻探討部份提到預先使用 AutoEncoder 對影像進行預處理將能有效提高分類或預測模型的準確率[5-6]。

Radiomics 抽取的特徵可分為 8 類，分別為一階統計特徵、基於 3D 形狀的特徵、基於 2D 形狀的特徵、灰度共生矩陣(Gray Level Co-Occurrence Matrix)特徵、灰度遊程矩陣(Gray Level Run-Length Matrix)特徵、灰度級帶矩陣(Gray Level Size Zone Matrix)特徵、灰度依賴性矩陣(Gray Level Dependence Matrix)特徵與鄰域灰度差分矩陣(Neighbor Gray Tone Difference Matrix)，並將獲取的 8 類特徵皆作為描述鈣化點的定量特徵，以獲取鈣化點群聚處的鈣化點資訊，為後續的鈣化點分類提供更多思路。

本研究除了運用 Radiomics 方法大量獲取乳房鈣化點特徵外，還使用 AutoEncoder 深度學習方法自動提取乳房鈣化點群聚處影像的特徵，AutoEncoder 架構中包含編碼器(Encoder)和解碼器(Decoder)兩部份，其架構圖如圖 3.5.1 所示，編碼器與解碼器主要分別進行影像壓縮與解壓縮的動作，將高維度的原始群聚鈣化點影像經過編碼器後，會以低維度的特徵向量保留輸入影像中具代表性的鈣化點特徵，方便解碼器重建出相近於原始影像的輸出影像。

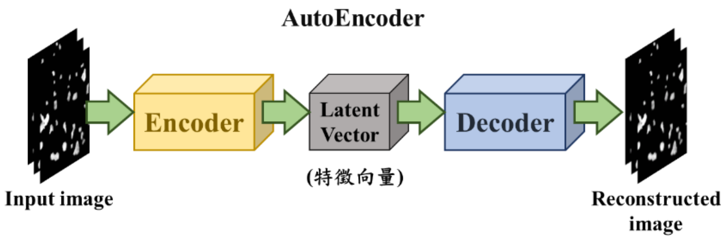


圖 3.5.1、AutoEncoder 架構圖

再透過訓練資料集中的大量群聚鈣化點影像進行神經網路的訓練，以獲得輸入影像的特徵向量。訓練完成的 AutoEncoder 模型能從群聚鈣化點影像中自動進行特徵抽取與選擇，而影像組學無需訓練階段便能抽取影像中的大量特徵，但一般須透過特徵工程的方法來選擇出具有意義的鈣化點特徵。本研究結合影像組學與深度學習方法進行特徵抽取，以確保後續的特徵選擇或分類乳房惡性乳房鈣化點中原位癌、前期癌或癌症的結果具有穩定與可靠性，並獲得更多存在乳房鈣化點群聚處影像中的資訊。

3.6 使用獨立樣本 t 檢定、ANOVA 及 MANOVA 篩選特徵

由於 Radiomics 特徵抽取法與 AutoEncoder 模型抽取的大量鈣化點特徵中，並非所有特徵皆能用於將惡性鈣化點分類為三類的代表性特徵，因此本研究為了找出 Radiomics 與 AutoEncoder 中最具代表性的特徵，分別使用獨立樣本 t 檢定、單因子獨立變異數分析(Analysis of variance, ANOVA)及多變量變異數分析(Multivariate analysis of variance, MANOVA)進行最佳特徵選取，除去影像中不相關

(irrelevant)或是冗餘(redundant)特徵，以獲得真正影響惡性鈣化點分類的特徵。

獨立樣本 t 檢定是醫學領域常用的統計學方法，常用於比較兩組特徵的平均數是否呈顯著差異，因此本研究分別從 Radiomics 與 AutoEncoder 抽取出的多類別特徵中，透過排列組合的方式各選擇一類特徵，再透過統計學的假設檢定進行兩特徵類別的比較。在假設檢定中H_0是指虛無假設，而H_1則是指對立假設，H_0表示抽取的特徵與 Radiomics 或 AutoEncoder 的特徵類別之間無差異，反之則表示有差異，接著透過公式(1)進行統計量 t 值的計算。

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} \times \left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \tag{1}$$

其中，$\overline{\mathrm{X}_1}$與$\overline{\mathrm{X}_2}$分別代表兩組樣本的平均數，n_1與n_2代表兩組樣本的數量，而S_1和S_2分別代表兩組樣本的標準差，可用來判斷兩特徵類別之間是否有顯著差異。在獨立樣本 t 檢定中，顯著水準是作為假設檢定判斷的依據，用來表示觸犯第一類型錯誤(Type I Error)機率，通常以 α 表示，可利用邦佛洛尼校正(Bonferroni correction)方法修正 α。在統計學與許多研究領域通常認為 α 為 0.05，是指假設H_0為真，但檢定結果卻為拒絕H_0的機率。而 p 值則是用來衡量整體樣本資料拒絕H_0的機率，表示在分佈常態分布下，大於等於以及小於等於 t 值的機率密度值(probability density value)，當 p 值越小，統計意義越大。最後自由度(degree of freedom, df)是指各特徵類別的所有特徵中有多少特徵可以用來評估特徵類別的變異數，可透過公式(2)計算 Radiomics 與 AutoEncoder 兩者特徵類別的自由度大小，並依據自由度與 p 值查詢「t 值表」，以找出 t 值的臨界值，判斷 Radiomics 與 AutoEncoder 兩組特徵類別之間是否具有顯著差異。

$$df = n_1 + n_2 - 2 \tag{2}$$

接著本研究還會採用 ANOVA 與 MANOVA 方法進行鈣化點特徵選取。ANOVA 主要用於比較兩組以上的特徵類別間是否存在平均數差異，而 MANOVA 是 ANOVA 的延伸，可用於控制整體標準，瞭解多組特徵類別在兩個以上的應變數之間的相關性。兩者的差異在於 MANOVA 可以同時比較兩個以上的應變數，而 ANOVA 一次僅能分析一種應變數。

ANOVA 包含總離均差平方和(Sum of squares Total, SST)，組內離均差平方和(Sum of Squares Within, SSW)與組間離均差平方和(Sum of Squares Between, SSB)三部分，SST 表示整體樣本在應變項的差異狀況，如公式(3)，其可細分為 SSW 與 SSB 兩部分，SSW 是用於衡量 Radiomics 與 AutoEncoder 的每個特徵類別內部樣本之間的差異，一般認為是隨機誤差如公式(4)，而 SSB 則是用於衡量不同特徵類別之間的差異，如公式(5)。

$$SST=\sum(X_i-\overline{X})^2=SSW+SSB \tag{3}$$

$$SSW=\sum(X_i-\overline{X_k})^2 \tag{4}$$

$$SSB=\sum n_k\,(\overline{X_k}-\overline{X})^2 \tag{5}$$

其中總特徵數為 n，各特徵類別的特徵數為n_k，X_i是指總特徵中的第 i 個特徵，$\overline{X_k}$是指各特徵類別的平均數，而$\bar{X}$是總特徵值的平均數，比較 SSB 與 SSW 的比值是否高於臨界值，若數值高於臨界值則表示樣本特徵對於特徵類別有一定的區分度。而界定臨界值則要用到 ANOVA 檢定的統計量 F，如公式(6)所示。

$$F=\frac{MSB}{MSE}=\frac{SSB/(k-1)}{SSW/(n-k)} \tag{6}$$

其中 n 為總特徵數，k 為特徵類別數， MSB 與 MSE 是分別由 SSB 與 SSW 除以其自由度計算而得，而 SSW 和 SSB 的自由度分別為$n-k$和$k-1$，之後再對照「F 分布表」，依設定的顯著水準決定臨界值，判斷 Radiomics 與 AutoEncoder 各組特徵類別間是否呈顯著差異。此外，在大部分的統計推論研究中常會超過一個應變數，且應變數之間容易存在中等相關，因此僅採用 ANOVA 分析易違反 Type I Error，代表應該修正 α 方能獲得正確的顯著水準。本研究為了確保所選擇的特徵能正確將惡性鈣化點分類為原位癌、前期癌及癌症，還利用 MANOVA 評估 Radiomics 所獲得的大量特徵，MANOVA 與 ANOVA 的運算邏輯相似，是利用總體、組內與組間平方和與交叉相乘矩陣(Sum of Square and Cross Product)表示整體特徵在應變項的差異狀況，如公式(7)、(8)與(9)所示。

$$T=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij}-\overline{y}\right)\left(y_{ij}-\overline{y}\right)' \tag{7}$$

$$W=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij}-\overline{y_i}\right)\left(y_{ij}-\overline{y_i}\right)' \tag{8}$$

$$B=\sum_{i=1}^{k}(\overline{y_i}-\overline{y})(\overline{y_i}-\overline{y})' \tag{9}$$

其中 k 表示特徵類別數，n_i代表各特徵類別的特徵數，y_{ij}表示第 i 組特徵類別中的第 j 項，$\overline{y_i}$為各特徵類別之平均數，而$\overline{y}$為總特徵值的平均數，接著同樣透過 F 檢定找出臨界值，判斷各特徵類別之間是否存在顯著差異，以統計學所獲得的數據為基礎，作為乳房鈣化點特徵選擇的依據，如公式(10)所示。

$$F=\frac{MSA}{MSE}=\frac{B/(k-1)}{W/(n-k)}=\frac{n-k}{k-1}W^{-1}B \tag{10}$$

3.7 採用 DNN、SVM 及 Random Forest 分類惡性鈣化點

為了驗證獨立樣本 t 檢定、ANOVA 及 MANOVA 所篩選的特徵是否能夠用來精確的將惡性乳房鈣化點分類為原位癌、前期癌或癌症，本研究採用 DNN、SVM 及 Random Forest，三種常用於醫學影像分類的方法來進行檢驗，其中 DNN 內部的神經網路層可分為輸入層，隱藏層和輸出層三類，且包含前向傳遞與反向傳遞兩階段，前向傳遞階段是從輸入層開始逐層往後計算，直到獲得鈣化點良惡的預測值。而反向傳遞階段會先設定目標函數來衡量實際與期望輸出值之間的誤差，再透過梯度下降法進行權重值的更新與訓練，將目標函數最小化，達到分類惡性鈣化點為原位癌、前期癌或癌症之目的；SVM 是一種二元分類器，當樣本數量不多時，分類的準確率高，且泛化能力強，適合用於乳房攝影影像中惡性微鈣化點的分類；Random Forest 可用於處理高維度的數據，適合用於處理經 Radiomics 及 AutoEncoder 所抽取出的多特徵的數據，本研究預計透過上述 3 種分類演算法來分類惡性乳房鈣化點為原位癌、前期癌或癌症，若分類結果良好，可證明獨立樣本 t 檢定、ANOVA 及 MANOVA 所篩選出的特徵為惡性乳房鈣化點分類之有效特徵。

第四章　實驗結果

在實驗結果中，首先會先介紹系統所使用到的開發環境與工具，接著會搭配系統的操作界面來詳細介紹系統的操作流程，最後會針對結果進行討論。

4.1 開發環境與工具

本研究每次訓練及測試的實驗環境皆使用相同的硬體配置，其規格為：16GB RAM、 i9-9900X CPU 和 NVIDIA Quadro RTX5000。本研究使用 Python 程式語言開發深度學習模型，並利用 Python 內建的 GUI 函式庫 Tkinter 進行本系統的圖形介面設計實作。

4.2 系統實作介面及操作流程

本系統僅提供醫學影像部的醫事人員操作使用，為保護病患隱私，在使用此系統前，均需先登入核對身分，方可使用，「基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類輔助診斷系統」登入頁如圖 4.2.1 所示。

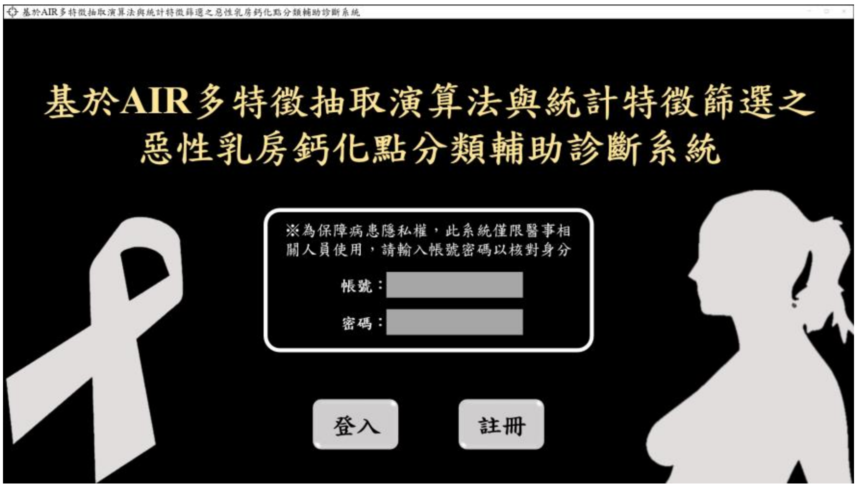


圖 4.2.1、系統登入頁

登入系統後即可進入群聚分析的功能頁面，左上方為工具列，可回上頁、回首頁或是查看系統簡介，右上方為醫師及病患資訊，醫師在看診時可點選「選擇影像」，輸入欲分析病患的乳房攝影影像，如圖 4.2.2 所示。

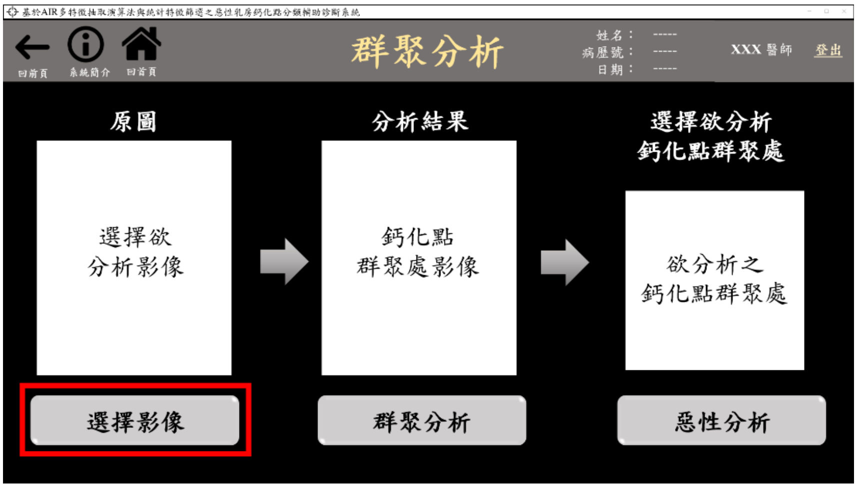


圖 4.2.2、輸入欲分析病患的乳房攝影影像

當醫師將病患的乳房攝影影像輸入至系統後，系統會讀取影像的標頭檔，並將病患的資訊呈現於系統的右上角，接著醫師可點選「群聚分析」，即可開始分析乳房攝影影像中群聚鈣化點處，如圖 4.2.3 所示。

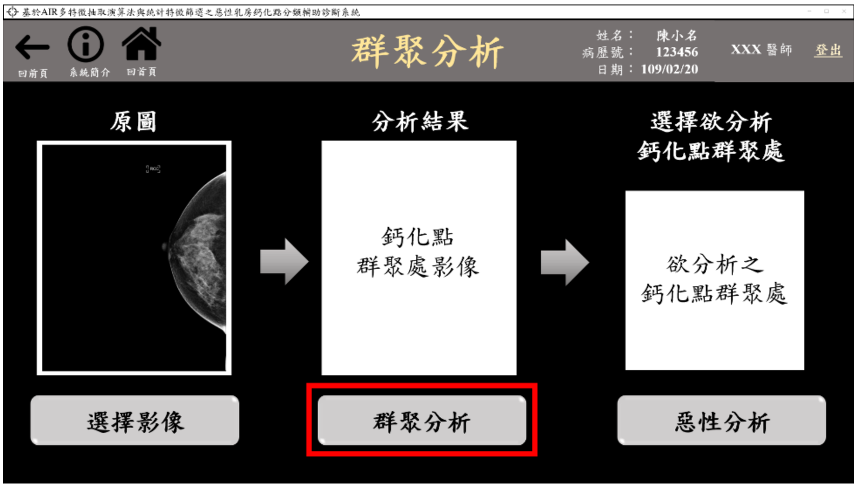


圖 4.2.3、點選群聚分析以找出乳房攝影影像中鈣化點群聚處

當系統分析出鈣化點群聚處後，會利用綠色候選框將其框選出來，並在分析結果中呈現，醫師若需儲存分析結果影像時，可點選分析結果右側的下載按鈕，即會將分析結果影像儲存至醫生電腦，接下來，醫師可利用滑鼠點選欲進行惡性群聚鈣化點分析之群聚鈣化點的候選框，如圖 4.2.4 所示。

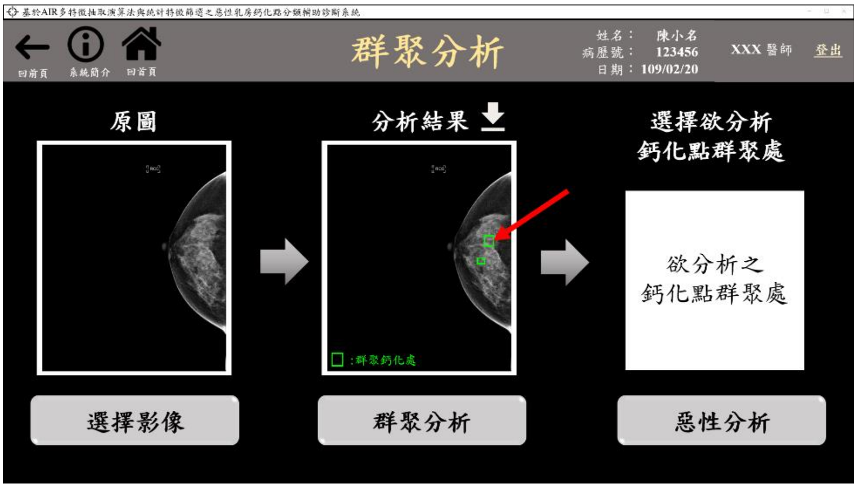


圖 4.2.4、點選欲分析之鈣化點群聚處候選框

醫師點選完欲進行惡性群聚鈣化點分析之群聚鈣化點的候選框後，系統會將醫師點選的群聚鈣化點影像呈現於系統的右側，如圖 4.2.5 所示，若醫師確認選擇的候選框無誤後，可點選「惡性分析」。

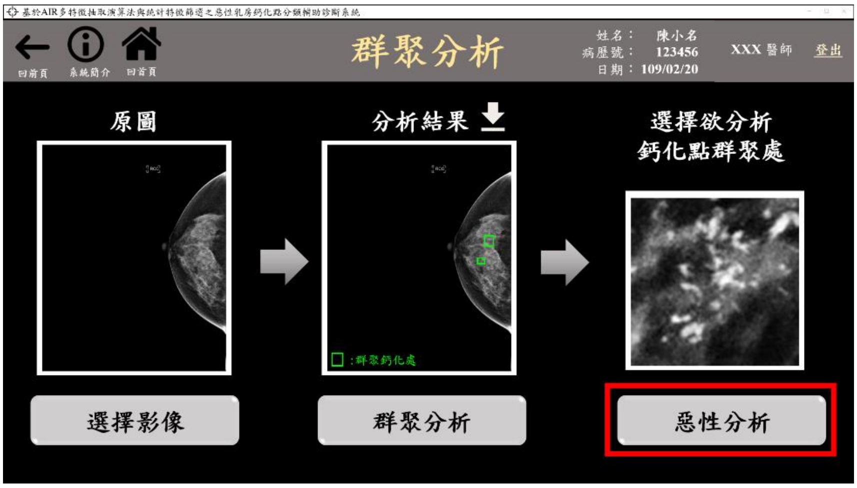


圖 4.2.5、點選惡性分析，分析群聚鈣化點之惡性類別

在醫師點選「惡性分析」後，系統會將鈣化點以外的部分去除，形成去背影像，同樣地，若醫師需儲存去背影像，可點選分析結果右側的下載按鈕，將影像儲存至醫師電腦，另外，分析結果分為四種，分別為非惡性、原位癌、前期癌及癌症，系統會將分析結果紀錄於右方醫囑文字框內，如圖 4.2.6 所示，醫師也可對該文字框進行醫囑的編輯與儲存。

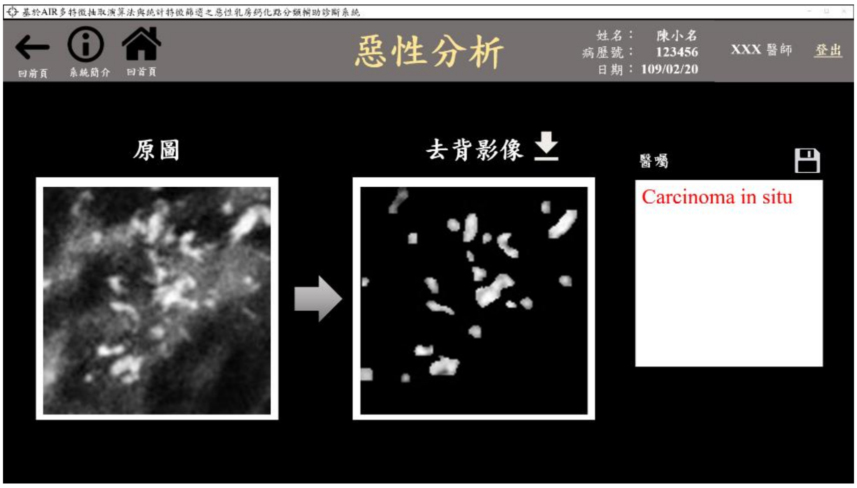


圖 4.2.6、系統將惡性分析結果紀錄於右方醫囑文字欄

4.3 實驗結果與討論

本研究分別進行了兩個實驗，實驗一是利用 Mask R-CNN 去背的群聚鈣化點影像分別訓練 SVM、Random Forest 及 DNN 分類惡性乳房鈣化點，而實驗二是將 Mask R-CNN 去背的群聚鈣化點影像透過本系統提出的特徵抽取法及最佳特徵篩選法篩選後的特徵來分別訓練 SVM、Random Forest 及 DNN 分類惡性乳房鈣化點，表 4.3.1 為經過實驗一的三個模型測試結果，可看出準確率皆不高於 90%。

另外，本研究所採用的三個模型分別有其限制，如 SVM 在訓練樣本數量大且計算量大時，會有訓練效果不佳的問題；隨機森林雖然可以處理高維度特徵資料，但卻須使用大量的決策樹來進行分類，提高了計算的成本；DNN 是全連接的網路，因此下一層的神經元均能和上一層的神經元形成連接，但這此方式容易導致參數數量過多，計算量龐大的問題。

表 4.3.1、實驗一的三個模型分別於群聚鈣化點惡性分析的表現

	Accuracy	Precision	Sensitivity	Specificity
SVM	87.42%	86.75%	88.89%	85.90%
Random Forest	88.68%	86.39%	91.82%	85.53%
DNN	89.31%	87.57%	91.93%	86.62%

實驗二的結果如表 4.3.2 所示，SVM、Random Forest 及 DNN 分類惡性乳房鈣化點的準確率分別為 90.56%、91.82%及 93.71%，敏感度分別為 91.19%、92.50%及 94.37%，特異度分別為 89.93%、91.13%及 93.03%，精準度分別為 90.06%、91.35%及 93.20%，其中準確率皆高於 90%，由此可證明本研究提出的特徵抽取法及最佳特徵篩選法除了有效降低計算量外，還能篩選出重要的特徵以利提升模型的分類準確率。

表 4.3.2、實驗二的三個模型分別於群聚鈣化點惡性分析的表現

	Accuracy	Precision	Sensitivity	Specificity
SVM	90.56%	90.06%	91.19%	89.93%
Random Forest	91.82%	91.35%	92.50%	91.13%
DNN	93.71%	93.20%	94.37%	93.03%

第五章　結論

本研究以分類乳房惡性鈣化點為原位癌、前期癌或癌症及篩選出影響分類的具代表性特徵為目標，透過實際訪問醫學影像部醫師，了解期望改善之處，並參考多篇文獻後開發出「基於 AIR 多特徵抽取演算法與統計特徵篩選之惡性乳房鈣化點分類預測系統」，此系統運用影像組學及統計學的方法找出對於分類乳房鈣化點有影響力的特徵，再使用深度學習的方法將其分類為原位癌、前期癌或癌症。本研究分別將未經過特徵抽取與選擇的群聚鈣化點去背影像以及經過 AIR 架構篩選出的特徵值輸入至 SVM、Random Forest 及 DNN 進行惡性鈣化點的分類，其中直接將群聚鈣化點去背影像進行分類的 SVM、Random Forest 及 DNN 模型準確率分別為 87.42%、88.68%及 89.31%，而利用 AIR 架構篩選出的特徵值分類的 SVM、Random Forest 及 DNN 模型準確率分別為 90.56%、91.82%及 93.71%。另外，GUI 介面的設計還參考醫院內醫療輔助系統的介面後開發出利於醫師操作的 GUI 介面，此系統可輔助醫師進行乳房攝影影像的判讀及診斷，提供更多的資訊作為判斷的依據，同時病患也可避免接受侵入式的檢測，既能達到提高民眾就診意願，還能早期發現乳癌，再搭配適當的治療方法即可提高乳癌患者的存活率。

參考文獻

[1] L. Macías-García, J. M. Luna-Romera, J. García-Gutiérrez, M. Martínez-Ballesteros, J. C. Riquelme-Santos et al., “A study of the suitability of autoencoders for preprocessing data in breast cancer experimentation,” Journal of Biomedical Informatics, vol. 72, pp. 33-44, 2017.

[2] L. Losurdo, A. Fanizzi, M. Teresa Basile, R. Bellotti, U. Bottigli et al., “Radiomics Analysis on Contrast-Enhanced Spectral Mammography Images for Breast Cancer Diagnosis: A Pilot Study,” Entropy, vol. 21, pp. 11, 2019.

[3] A. Ibrahim, M. Vallières, H. Woodruff, S. Primakov, M. Beheshti et al., “Radiomics Analysis for Clinical Decision Support in Nuclear Medicine,” Seminars in Nuclear Medicine, vol. 45, pp. 438-449, 2019.

[4] O. M. Ijurra, Y. Jiao, Q. Wang, “ConvRadiomics: Convolutional Radiomics Feature Extraction Toolkit,” 2019 International Conference on Medical Imaging Physics and Engineering (ICMIPE), pp. 1-5, 2019.

[5] A. Bizzego, N. Bussola, D. Salvalai, M. Chierici, “Integrating deep and radiomics features in cancer bioimaging,” 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp. 1-8, 2019.

[6] C. S. Ling, W. Mimi Diyana W. Zaki, A. Hussain, W. Siti Halimatul Munirah W. Ahmad, “Shape Based Image Retrieval System for MRI Spine,” 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI), pp. 1-6, 2017.

[7] D. Das, L. B. Mahanta, S. Ahmed, B. K. Baishya, “A study on MANOVA as an efective feature reduction technique in classifcation of childhood medulloblastoma and its subtypes,” Network Modeling Analysis in Health Informatics and Bioinformatics, vol. 9, pp. 16, 2020.

[8] M. Caballo, D. R. Pangallo, R. M. Mann, L. Sechopoulos, “Deep learning-based segmentation of breast masses in dedicated breast CT imaging: Radiomic feature stability between radiologists and artificial intelligence,” Computers in Biology and Medicine, vol. 118, 2020.

[9] G. Makwana, L. Gupta, “Identification of Micro-calcification in Mammogram for Breast Cancer Analysis using SVM Classifier,” International Journal of Engineering & Technology, vol. 7, pp. 29-32, 2018.

[10] R. Vijayarajeswari, P. Parthasarathy, S.Vivekanandan, A. A. Basha, “Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform,” Measurement, vol. 146, pp. 800-805, 2019.

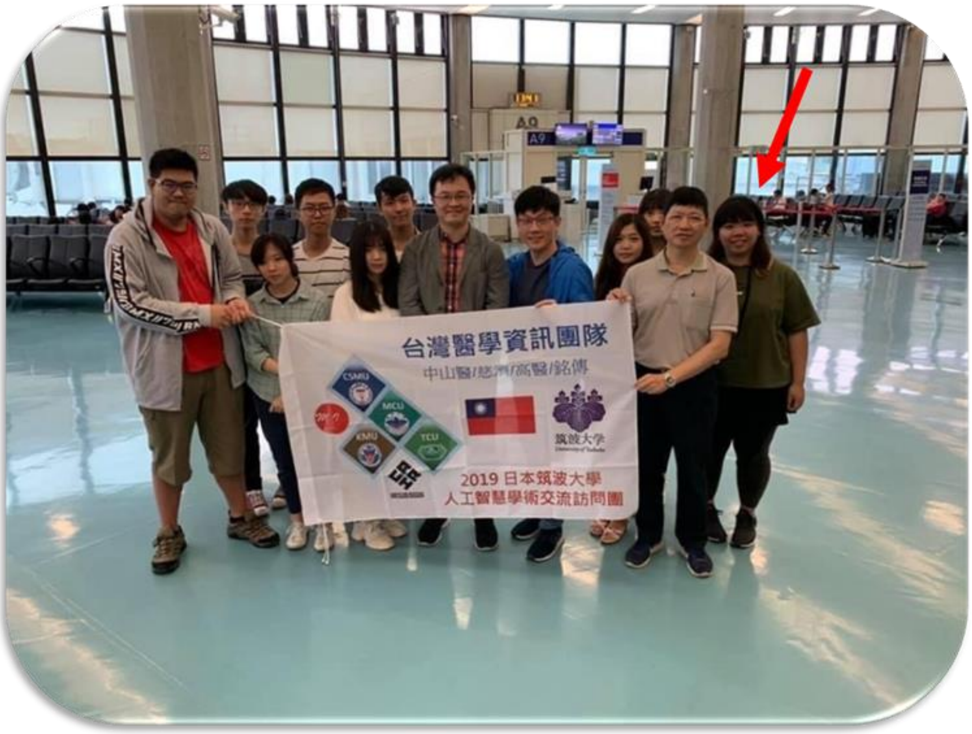
[11] R. Zemouri, N. Omri, B. Morello, C. Devalland, L. Arnould et al., “Constructive Deep Neural Network for Breast Cancer Diagnosis,” IFAC-PapersOnLine, vol. 51, pp. 98-103, 2018.

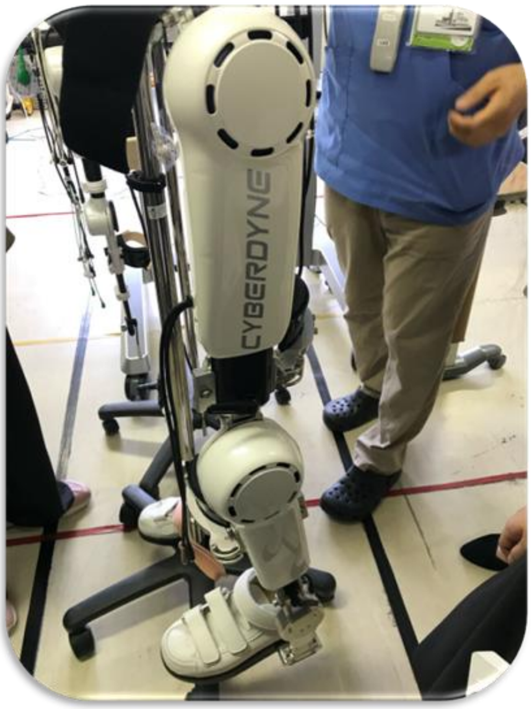
[12] M. Z. Alam, M. S. Rahman, M. S. Rahman, “A Random Forest based predictor for medical data classification using feature ranking,” Informatics in Medicine Unlocked, vol. 15, 2019.

九、 課外活動

筑波大學參訪

交流日期	2019/06/24~2019/06/27





中科智慧機器人自造基地參訪	
參訪日期	2020/07/01

玖、附件

附錄目錄

附錄一、論文本文

Biomedical Engineering: Applications, Basis and Communications

Calcification clusters and lesions analysis in mammogram using multi-architecture deep learning algorithms

Hao-Hung Tsai(蔡鎬鴻)[1, 2, 3], Chia-Shin Wei(魏嘉芯)[4], Ya-Chu Hsieh(謝雅竹)[4], I-Miao Chen(陳怡妙)[4], Pin-Yu Yeh(葉品郁)[4], Darren Shih(施敦壬)[5] and Chiun-Li Chin(秦群立)[4,*]

[1] Department of Medical Imaging, Chung Shan Medical University Hospital, Taichung, Taiwan

[2] Department of Medical Imaging in School of Medicine, Chung Shan Medical University, Taichung, Taiwan

[3] Department of Medical Imaging and Radiological Sciences, Chung Shan Medical University, Taichung, Taiwan

ABSTRACT

Today, radiologists observe a mammogram to determine whether breast tissue is normal. However, calcifications on the mammogram are so small that sometimes radiologists cannot locate them without a magnified observation to make a judgment. If clusters formed by malignant calcifications are found, the patient should undergo a needle localization surgical biopsy to determine whether the calcification cluster is benign or malignant. However, a needle localization surgical biopsy is an invasive examination. This invasive examination leaves scars, causes pain, and makes the patient feel uncomfortable and unwilling to receive an immediate biopsy, resulting in a delay in treatment time. The researcher cooperated with a medical radiologist to analyze calcification clusters and lesions, employing a mammogram using a multi-architecture deep learning algorithm to solve these problems. The features of the location of the cluster and its benign or malignant status are collected from the needle localization surgical biopsy images and medical order and are used as the target training data in this study. This study adopts the steps of a radiologist examination. First, VGG16 is used to locate calcification clusters on the mammogram, and then the Mask R-CNN model is used to find micro-calcifications in the cluster to remove background interference. Finally, an Inception V3 model is used to analyze whether the calcification cluster is benign or malignant. The prediction precision rates of VGG16, Mask R-CNN, and Inception V3 in this study are 93.63%, 99.76%, and 88.89%, respectively, proving that they can effectively assist radiologists and help patients avoid undergoing a needle localization surgical biopsy.

***Keywords*:** Mammogram, malignant calcification clusters, VGG16, Mask R-CNN, Inception V3.

1. Introduction

Some studies indicate that more than 90% of breast patients have been diagnosed with breast cancer without metastasis,1 and for nonmetastatic patients, the goal of treatment is to eradicate the tumor and prevent a recurrence. The study results found that a

*Corresponding author: Chiun-Li Chin, Medical Informatics, Department of Health Care Management, Chung Shan Medical University, Taichung, Taiwan. E-mail: ernestli@csmu.edu.tw

regular mammography once every two years can reduce breast cancer mortality by 41% and reduce the incidence of advanced breast cancer by 30%. The calcifications in the mammogram are the primary basis for judging whether breast cancer exists.2,3,4 Malignant calcification clusters can easily turn into breast cancer.

In addition, using calcifications to diagnose breast cancer is more accurate than using masses.5 Therefore, radiologists first observe whether calcification clusters occur in the mammogram. If more than three calcifications appear in the image within a 5×5 cm area, it is a cluster area. Then, based on the information regarding the shape, size, arrangement, and brightness of the calcifications in the area, the cluster can be diagnosed as malignant or benign for the subsequent diagnosis. However, for radiologists, calcifications are very small in mammograms. Unless the images are enlarged, they cannot be directly identified with the naked eye, and it takes considerable time to diagnose.

In addition, although computer-aided diagnosis (CAD) assists in locating calcification clusters, no system currently exists on the market that can directly identify calcification clusters as benign or malignant. Whether the calcification clusters are benign or malignant still depends on an invasive needle localization surgical biopsy to detect the condition. However, according to a survey by the Health Promotion Administration, approximately 14.4% of women are unwilling to undergo detection due to the discomfort and pain during the mammography process, leading to delays in the optimum treatment period.

This study establishes a set of methods to imitate the radiologist consultation procedure. First, calcification clusters are located accurately using the mammogram and are further identified as benign or malignant to assist radiologists during an examination. Moreover, this method reduces the pain and fear of patients by avoiding the need for a needle localization surgical biopsy, increasing the willingness of patients to receive treatment.

Deep learning has been gradually emphasized and applied, and studies on the application of deep learning to analyze images have been proposed. This study uses deep learning methods to extract features from numerous images, summarizing and organizing them to predict the results. The following sections focus on using deep learning to classify images and segment objects in images and their use in a related study on mammograms concerning the ideas put forward in the literature to strengthen their practices or reduce their shortcomings to make the method more complete and accurate.

1.1. Using Deep Learning to Classify Images

In recent years, many researchers have used deep learning algorithms to classify medical images. In 2020, Lee et al.6 proposed using VGG167 to classify coronavirus disease 2019 (COVID-19), pneumonia, and normal chest X-ray images. In 2020, Sukegawa et al.8 used VGG16 and VGG19 to classify the types of implants in panoramic images. The results reveal that VGG16 exhibits a better classification result. Nkwentsha et al.9 focused on the classification of X-ray images in 2020. They used the ImageCLEF 2009 dataset, including 193 categories, such as the abdomen, hand, clavicle, and cephalogram, and classified them using Inception V3.10 In 2018, Li et al.11 proposed using Inception V3 to classify colorectal cancer lymph nodes in magnetic resonance imaging as benign or malignant, and the final accuracy rate was 94.4%. Their study also mentioned using ImageNet data and that pretraining can improve the generalizability of the model. The above studies have proved that VGG16 and Inception V3 have been widely used to analyze medical images and have good performance.

In 2019, Guan et al.12 used VGG16 to classify fragmented cytological images as papillary thyroid carcinoma (PTC) or benign nodules, and they compared it with the classification results of AlexNet and Inception V3. The study pointed out that the convolutional layer of VGG16 is deeper than that of AlexNet, so the feature extraction ability is better. Inception V3 uses different sizes of convolutional kernels, so it is more suitable for targets with different sizes in images. After a comprehensive evaluation, VGG16 is the most advantageous in classifying fragmented cytological images. In this study, referring to the study method, VGG16 was used to classify the clustered and nonclustered calcifications in mammograms, and Inception V3 was used to classify calcification clusters as benign or malignant.

1.2. Using Deep Learning to Segment Objects in Images

Many researchers have used deep learning algorithms to segment specific areas in medical images. The Mask R-CNN[13] is widely used to segment medical images and has a good segmenting effect[14,15]. In 2018, Liu et al.[16] mentioned that the shape of lung nodules is complex with no clear outline. Previously, the method of detecting lung nodules could only locate the center point, but the size of the lung nodules is an important diagnostic criterion, so Mask R-CNN is used to segment only the outline of the lung nodules in computed tomography (CT) images. In 2020, Hu et al.[17] further combined the Mask R-CNN with K-means and changed the last layer of the Mask R-CNN to an unsupervised method.

In 2014, Luis et al.[18] proposed the scale-invariant feature transform (SIFT) and K-means classification to determine abnormal regions in mammograms. However, SIFT may find too many feature points at different scales and is susceptible to noise interference. The authors of this study also found that the muscle brightness in the X-ray image affected the recognition results and caused K-means to output incorrect results when clustering. This study used deep learning to improve the above situation and make the prediction results more accurate.

In 2018, Ribli et al.[19] used the Faster R-CNN to detect abnormal regions in mammograms, using ImageNet to pretrain VGG16. The region proposal network (RPN) was used to obtain the candidate box of the abnormal area and further classify and adjust the position of the candidate box. This study referred to the study method of adding the radiologists' professional knowledge and needle localization surgical biopsy images and using the Mask R-CNN to further detect each calcification.

In 2019, Cai et al.[20] proposed traditional methods, such as morphology and the maximum connection area, to detect calcification clusters in mammograms. They used the AlexNet deep learning algorithm to classify clusters as benign or malignant, aiming to effectively assist the radiologist's diagnosis of benign or malignant, similar to the motivation and purpose of this study. However, their method of detecting clustered calcifications uses traditional methods instead of deep learning algorithms to learn and train large amounts of data; thus, there is still room for improved accuracy. In addition, the radiologists' diagnosis experience is not considered, and there are more than three calcifications in the 5×5 cm of the mammogram, which are calcification clusters. This method results in interpretive differences when used by radiologists and reduces the reference to deep learning methods.

In 2019, Alqudah et al.[21] built a CAD system to analyze mammograms. After using traditional image processing methods to denoise, enhance contrast, and segment the mass area, the probabilistic neural network was used to classify the mass into one normal and six abnormal categories (calcification, circumscribed, spiculated, ill-defined, architectural distortion, and asymmetry). Then, the support vector machine (SVM) was used to classify the abnormality into benign or malignant according to the degree of abnormality. This study expects to detect breast cancer precursors at an earlier stage, so the detection target of this study changed from the mass in the image to the calcification cluster. Referring to the analysis process, the abnormal area is first segmented from the mammogram. The method of segmenting abnormal areas also changed from traditional methods to deep learning algorithms with higher accuracy. The follow-up analysis of the benign or malignant calcification clusters also changed to a more suitable deep learning algorithm.

In 2020, Min et al.[22] proposed a multiscale morphological sifter to convert mammograms into pseudocolor mammograms, extract blocky patterns, and suppress the background to convert the grayscale image into a color image with special meaning so that the Mask R-CNN can segment a fibrocystic breast. This study referred to the idea of using the Mask R-CNN to segment the tiny calcifications in the mammogram and remove the complex background.

A comprehensive review of the above literature indicates that VGG16, Mask R-CNN, and Inception V3 can achieve good results in medical imaging. This study is based on the ideas proposed in the above literature, with further improvements and optimization, along with a self-built graphical user interface (GUI), the calcification cluster and lesion analysis in mammogram using multi-architecture deep learning algorithms. The contributions of this study are as follows:

(1) The deep learning algorithms automatically analyze the mammogram. First, VGG16 detects clustered calcifications in the mammogram. Then, the Mask R-CNN detects calcification clusters and further removes the background. Finally, with the background removed, Inception V3 classifies the images of the calcification cluster as benign or malignant.

(2) The location of the breast lesion of the patient can be found without a needle localization surgical biopsy, reducing the psychological burden of patients and eliminating the pain caused by invasive treatment.

(3) Numerous data and a robust neural network analyze reliable data to improve patient identification, speeding up patient consideration time and allowing them to receive treatment as soon as possible.

(4) Quick analysis and precise identification provide radiologists with reference information and make radiologists more confident in the diagnosis. It does not easily miss tiny calcifications, effectively assisting radiologists in diagnosis and speeding up the examination and diagnosis time.

2. Materials and methods

In this study, the multi-architecture deep learning algorithm comprises three architectures: VGG16, Mask R-CNN, and Inception V3. First, VGG16 was used to locate cluster areas in mammograms. Then, all cluster images were used to train the Mask R-CNN to extract the individual calcification area from each cluster image. The image of each calcification point was removed from the background. Finally, Inception V3 was used to classify the benign or malignant calcification images obtained in the previous step. The flowchart of this study is illustrated in Fig. 1. The characteristics of datasets are described in Section 2.1, and the detailed individual architecture used in this study is explained in Sections 2.2 to 2.4.

2.1. Datasets

In this study, the researcher cooperated with radiologists, who provided professional knowledge and mammograms. These images include the craniocaudal (CC) view image depicted in Fig. 2(a) and the mediolateral oblique (MLO) view image depicted in Fig. 2(b) of each patient's breast. In addition, the needle localization surgical biopsy image is the X-ray image taken after the locating needle punctures the calcification clusters before the patient takes needle localization surgical biopsy, as illustrated in Fig. 2(c). This image was used as the basis for locating the calcification cluster in the mammogram corresponding to the CC and MLO views of the patient, framing the calcification cluster areas and segmenting them into the training data for the calcification cluster analysis. Our dataset has a total of 29,208 images. They can be divided into two types: clusters of calcifications and no clusters of calcifications. The size of each image in the dataset is 128×128. In VGG16 cluster detection model, 23366 images are used as the training dataset and 5842 images are used as the testing dataset. In Mask R-CNN calcification detection model, 14604 calcification cluster images are used to perform training and testing. Among them, 11683 calcification cluster images are used as the training dataset, and 2921 calcification cluster images are used as the testing dataset. The number of images used as the training dataset and testing dataset of the Inception V3 benign and malignant identification model is the same as the Mask R-CNN calcification detection model, but the images only retain calcification points.

2.2. Using VGG16 to Locate Calcification Clusters in Mammograms

In this study, the clusters and nonclusters in mammograms were cropped to a size of 128×128. The cluster image is presented in Fig. 3(a). The cropping standard is the needle localization surgical biopsy image provided by the radiologist. The localization needle in the needle localization surgical biopsy image was used to inform the radiologist of the location of the calcification cluster. Therefore, clusters are the positions of the locating needle in the needle localization surgical biopsy image. Nonclusters are the mammae, nipple, tag, and breast edge in the image, as displayed in Fig. 3 from (b) to (d).

After cropping the images, the clustered and nonclustered images were input into VGG16 to train the VGG16 cluster detection model. In addition, VGG16 contains 13 convolutional layers to extract features from mammograms, 5 pooling layers to reduce the size of the feature maps, and three fully connected layers to output the results as clustering or nonclustering.

In the testing, a sliding window of 128×128 was used to crop the mammogram starting from the upper left corner by moving 50 pixels from left to right and top to bottom. The cut image was input into the VGG16 cluster detection model, which predicted that the cluster was in the cluster candidate box.

In this study, the intersection over union (IoU) was used to screen cluster candidate boxes to determine the exact location of calcification clusters in mammograms. The mammograms before and after screening are presented in Fig. 4. In addition, IoU uses the area of two candidate boxes to calculate the degree of overlap between them. The calculation result is a value in the interval from 0 to 1. A value closer to 1 represents a greater overlap between the two candidate boxes. After several tests, the overlap between the two candidate boxes was too low when the IoU was lower than 0.6, which means that the area where the candidate box was located has a low probability of being a calcification cluster. Therefore, these candidate boxes were deleted, leaving candidate boxes with an IoU higher than 0.6. The calculation formula is as follows (1):

$$IoU = \frac{AO(Area\ of\ Overlap)}{AU(Area\ of\ Union)} = \frac{Area\ of\ Candidate\ Box1\ \cap\ Area\ of\ Candidate\ Box2}{Area\ of\ Candidate\ Box1\ \cup\ Area\ of\ Candidate\ Box2}. \tag{1}$$

2.3. Using Mask R-CNN to Locate Calcifications in the Cluster and Remove the Background

In this study, the images classified by VGG16 as clusters were labeled with an open-source VGG image annotator23 tool. Individual calcifications on clusters were labeled one by one according to the appearance of the calcification, as displayed in Fig. 5(b). In this study, professional radiologists were asked to help confirm whether the labeling was correct. After labeling, the file name and coordinates of the labeling were saved as a JSON file, which was used as the training data for the Mask R-CNN.

After detecting the cluster calcifications, the clusters were segmented and then the Mask R-CNN learned the shape and brightness of a single calcification. Then, the calcifications were segmented in the image, as presented in Fig. 5(c). The method includes a backbone composed of ResNet101 and the feature pyramid network (FPN)24 to extract the features of calcifications in the image. Then, the RPN25 uses anchor boxes of different aspect ratios and sizes to scan the entire image to detect objects in the image. Then, the region of interest align13 uses bilinear interpolation to fine-tune the region of interest and fully connected layers to output the object coordinates and classes. The mask on the image is the calcification area masked by fully convolutional networks.

The Mask R-CNN locates all calcifications in the cluster and a self-built algorithm removes the background, as illustrated in Fig. 5(d). The reason for removing the background is that radiologists stated that the brightness of the mammae in the mammogram is relatively low, which easily affects the analysis of calcification clusters as benign or malignant. Therefore, they proposed that the interface would provide images of the calcification clusters with only the calcifications. In addition, experiments have confirmed that the idea of a radiologist removing the background can effectively improve the accuracy of classifying calcification cluster images as benign or malignant.

2.4. Using Inception V3 to Classify as Benign or Malignant

We used the medical record provided by the Department of Medical Imaging at Chung Shan Medical University Hospital to obtain the needle localization surgical biopsy results for each patient. Professional radiologists divided the results into benign and malignant and used them as training data for the identification model.

In this study, Inception V3 was used to extract the features of benign and malignant calcification clusters to establish a benign and malignant identification model. The convolutional layer of Inception V3 uses filters of different sizes to extract features of different dimensions, from calcifications to clusters, to increase recognition accuracy. In addition, Inception V3 uses the inception module to increase the width of the architecture to speed up training time. Inception V3 not only has high recognition accuracy, but it can also reduce the computing time. Unlike other deep learning classification models, Inception V3 has a considerable advantage, so this study applies it to classify benign and malignant calcification clusters.

3. Experimental results

3.1. Confusion Matrix of the Mask R-CNN Calcification Detection Model and Easy-to-Understand Graphical User Interface

In this study, the confusion matrix of the Mask R-CNN calcification detection model considers the labeling status of all calcifications, including the position and contour. If they are all labeled correctly, as depicted in Fig. 6(a), then this type of image is defined as a true positive. If the background is incorrectly labeled as a calcification, as presented in Fig. 6(b), it is defined as a false positive. If the calcification is misclassified as the background and is not labeled, as in Fig. 6(c), it is defined as a false negative.

In this study, a simple and easy-to-understand GUI was built for radiologists to use. The system is only used by medical personnel to protect patient privacy, and they must enter their account and password to verify their identity before use. After login, the radiologist clicks the “Select Image” button to select the image to be analyzed. To analyze the calcification clusters in the mammogram, the radiologist clicks the “Cluster Analysis” button, and the analysis result is displayed in the interface, as illustrated in Fig. 7. If the radiologist wants to further analyze the benign or malignant conditions of the calcification clusters, he or she can click the “Benign or Malignant” button. The system detects the calcifications in the image, removes the background for the radiologist’s reference, and records the benign or malignant results in the medical order on the left of the interface, as depicted in Fig. 8.

3.2. *Radiologists Judge the Calcification Cluster Type*

This study uses a GUI for radiologists to observe cluster images and asks radiologists to determine whether the cluster images are benign or malignant based on their experience. The interface is divided into three stages to prevent radiologists from affecting the test results due to fatigue. The test time for each stage is two weeks apart, and the image sequence is adjusted to prevent the radiologist from choosing the answer based on the memory of the previous stage. The first stage interface of the

radiologist test is presented in Fig. 9(a). The interface provides cluster images cropped from mammograms. After the radiologist judges the status of the cluster images, he or she clicks the “Benign calcification clusters,” “Malignant calcification clusters,” or ”Unrecognized” button on the right side of the interface. In the second stage, the radiologist can observe the calcification cluster images after removing the background. The second stage of the interface of the radiologist test is displayed in Fig. 9(b). In the third stage, the original image and the image with the background removed are displayed in the interface to determine the cluster status. The third stage of the interface of the radiologist test is in Fig. 9(c).

4. Discussion

4.1. *Performance of the Testing Phase*

In this study, the accuracy, precision, sensitivity, and specificity calculated for the VGG16 cluster detection model, Mask R-CNN calcification detection model, and Inception V3 benign and malignant identification model are higher than 85%. The table of comprehensive results is listed in Table 1. In conclusion, the three architectures effectively analyzed the mammograms.

The accuracy, precision, sensitivity, and specificity of the confusion matrix calculated by VGG16 in the test phase are 90.11%, 93.63%, 85.65%, and 94.40%, respectively. In addition, VGG16 has an excellent effect on classifying the cluster and noncluster calcification images, but the sensitivity is 85.65%; thus, room exists for improvement. The sensitivity is expected to improve if more calcification cluster images of different conditions are added during training.

The precision and sensitivity of the Mask R-CNN in segmenting calcifications are 99.76% and 99.28%, respectively. These two metrics verify that the Mask R-CNN has a high ability to segment calcifications. Only images with too dense calcifications or blurred contours are incorrectly labeled.

The Mask R-CNN segments the calcifications in the cluster and then removes the background. Inception V3 classifies the images of calcification clusters with removed backgrounds as benign or malignant. The classification accuracy, precision, sensitivity, and specificity are 93.76%, 88.89%, 95.60%, and 92.64%, respectively. Removing the background has a good effect on the calcification cluster images of Inception V3 classification.

4.2. *Radiologist Test Results*

The radiologist’s accuracy in judging calcification clusters is 46.88%, and 10.75% of the images cannot be evaluated empirically. The above experimental results demonstrate that radiologists have a limited ability to judge benign and malignant statuses of calcification clusters. Even if they have seen images of countless patients, the judgment results are still not ideal. Therefore, radiologists still rely on needle localization surgical biopsy to determine the clustering status after finding calcification clusters.

The radiologist judges the benign or malignant calcification cluster image after removing the background with an accuracy of 58.28%. Compared with the first stage, removing the complex background, such as mammae and tissue, can assist the radiologist in judging the condition of the calcification cluster more accurately.

The above three-stage test results indicate that using the Mask R-CNN to locate calcifications and remove the cluster image background can improve the accuracy of the radiologist's judgment. However, the radiologist's judgment accuracy still does not exceed 60%. In this study, the accuracy of using Inception V3 to classify calcification clusters as benign or malignant can be as high as 93.76%, providing radiologists with accurate results and reducing the misjudgment rate.

Providing radiologists with calcification cluster images and the image after removing the background to judge between benign and malignant statuses results in an accuracy of 60%. Compared with the judgment result of the second stage, which only removes the background image, the reference and comparison of the original image and the removed background image can effectively improve the accuracy of the radiologist's judgment.

4.3. Comparison with Other Related Studies

This study was compared with the method proposed by Cai et al.20 to analyze malignant calcification clusters in mammograms and compare the benign and malignant analysis results of the two methods. The comparison is in Table 2. The results demonstrate that the results of this study are superior in accuracy, sensitivity, and specificity. This study uses deep learning to detect calcification clusters in mammograms and then uses Inception V3 to classify them as benign or malignant, which is better than using traditional methods is to locate the suspected lesion area and using AlexNet to classify them. Due to the differences in the overall process and method selection, this study has more advantages.

5. Conclusion

This study proposes the use of three deep learning architecture methods to analyze mammograms. These deep learning algorithms include using VGG16 to classify clustering and nonclustering calcifications, using the Mask R-CNN to detect calcifications in clusters and Inception V3 to classify calcification clusters as benign or malignant. The training materials are prepared by referring to the needle localization surgical biopsy images and medical orders. In this study, the radiologist's viewpoint is specifically considered through interviews with professional radiologists to understand the procedures for consultation and the standards for observing mammograms. Following the radiologist's procedures, calcification clusters were first identified and classified as benign or malignant. The contribution of this study is to accurately analyze the patient's breast condition to assist the radiologist in diagnosis, providing an easy-to-operate graphical user interface for the radiologist to view. The malignant calcification cluster can be detected in the mammogram without requiring a needle localization surgical biopsy. In the future, it is hoped that after the detection of a malignant calcification cluster, the type of malignancy can be further predicted.

6. Acknowledgments

The research was supported by department of Medical Imaging, Chung Shan Medical University Hospital (CSH-2013-A-020).

References

1. Waks AG, Winer EP, Breast cancer diagnosis: a review, *JAMA Network*, **321**(3):288, 2019.
2. Baker R, Rogers K D, Shepherd N, Stone N, New relationships between breast microcalcifications and cancer, British journal of cancer, **103**(7), pp. 1034-1039, 2010.
3. Melekoodappattu, J. G., Subbian, P. S., A hybridized ELM for automatic micro calcification detection in mammogram images based on multi-scale features, Journal of medical systems, **43**(7), pp. 1-12, 2019.
4. O'Grady, S., Morgan, M. P., Microcalcifications in breast cancer: From pathophysiology to diagnosis and prognosis, Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, **1869**(2), pp. 310-320, 2018.
5. Wang, J., Yang, X., Cai, H., Tan, W., Jin, C., Li, L., Discrimination of breast cancer with microcalcifications on mammography by deep learning, Scientific reports, **6**(1), pp. 1-9, 2016.
6. Lee KS, Kim JY, Jeon ET, Choi WS, Kim NH, Lee KY, Evaluation of Scalability and Degree of Fine-Tuning of Deep Convolutional Neural Networks for COVID-19 Screening on Chest X-ray Images Using Explainable Deep-Learning Algorithm, *National Library of Medicine*, **10**(4):213, 2020.
7. Simonyan K, Zisserman A, Very Deep Convolutional Networks for Large-Scale Image Recognition, *Cornell University*, arXiv:1409.1556, 2014.
8. Sukegawa S, Yoshii K, Hara T, Yamashita K, Nakano K, Yamamoto N, Nagatsuka H, Furuki Y, Deep Neural Networks for Dental Implant System Classification, *Biomolecules*, **10**(7):984, 2020.
9. Nkwentsha X, Hounkanrin A, Nicolls F, Automatic classification of medical X-ray images with convolutional neural networks, *2020 International SAUPEC/RobMech/PRASA Conference*, pp. 1-4, 2020.
10. Szegedy C, Vincent V, Sergey L, Jonathon S, Zbigniew W, Rethinking the inception architecture for computer vision, *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, pp. 2818-2826, 2016.
11. Li J, Wang P, Li Y, Zhao Y, Liu X, Luan K, Transfer learning of pre-trained Inceptionv3 model for colorectal cancer lymph node metastasis classification, *Proc IEEE Conf Mechatronics and Automation*, pp. 1650-1654, 2018.
12. Guan Q, Wang Y, Ping B, Li D, Du J, Qin Y, Lu H, Wan X, Xiang J, Deep Convolutional Neural Network VGG-16 Model for Differential Diagnosing of Papillary Thyroid Carcinomas in Cytological Images: A Pilot Study, *J Cancer*, **10**(20):4876-4882, 2019.
13. He K, Gkioxari G, Dollár P, Girshick R, Mask R-CNN, *Proc IEEE Comput Soc Conf Comput Vis*, pp. 2961-2969, 2017.
14. Shu JH, Nian FD, Yu MH, Li X, An Improved Mask R-CNN Model for Multiorgan Segmentation, *Mathematical Problems in Engineering*, **2020**:8351725, 2020.
15. Wessel J, Heinrich MP, Berg JV, Franz A, Saalbach A, Sequential rib labeling and segmentation in Chest X-Ray using Mask R-CNN, *arXiv: Image and Video Processing*, **1908**:08329, 2019.
16. Liu M, Dong J, Dong X, Yu H, Qi L, Segmentation of lung nodule in CT images based on Mask R-CNN, *2018 9th international conference on awareness science and technology (iCAST)*, pp. 1-6, 2018.
17. Hu Q, de F. Souza L F, Holanda GB, Alves SSA, Dos S Silva FH, Han T, Rebouças Filho PP, An effective approach for CT lung segmentation using mask region-based convolutional neural networks, *Artificial Intelligence in Medicine*, **103**:101792, 2020.

18. Salazar-Licea LA, Mendoza C, Aceves MA, Pedraza JC, Automatic Segmentation of Mammograms Using a Scale-Invariant Feature Transform and K-Means Clustering algorithm, *2014 11th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, pp. 1-6, 2014.
19. Ribli D, Horváth A, Unger Z, Pollner P, Csabai I, Detecting and classifying lesions in mammograms with deep learning. *Scientific Reports*, **8**:4165, 2018.
20. Cai H, Huang Q, Rong W, Song Y, Li J, Wang J, Chen J, Li L, Breast Microcalcification Diagnosis Using Deep Convolutional Neural Network from Digital Mammograms, *Computational and Mathematical Methods in Medicine*, **2019**:2717454, 2019.
21. Alqudah, A M, Algharib, H M, Algharib, A M, Algharib, H M, Computer aided diagnosis system for automatic two stages classification of breast mass in digital mammogram images, Biomedical Engineering: Applications, Basis and Communications, **31**(01):1950007, 2019.
22. Min H, Wilson D, Huang Y, Kelly S, Crozier S, Bradley A, Chandra S, Fully automatic computer-aided mass detection and segmentation via pseudo-color mammograms and Mask R-CNN, *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1111-1115, 2020.
23. VGG image annotator (VIA), Available at http://www.robots.ox.ac.uk/~vgg/software/via/, 2016.
24. Lin T Y, Dollar P, Girshick R, He K, Hariharan B, Belongie S, Feature pyramid networks for object detection, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 936-944, 2017.
25. Ren S, He K, Girshick R, Sun J, Faster R-CNN: towards real-time object detection with region proposal networks. In NIPS, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 91-99, 2015.

Figures and Tables

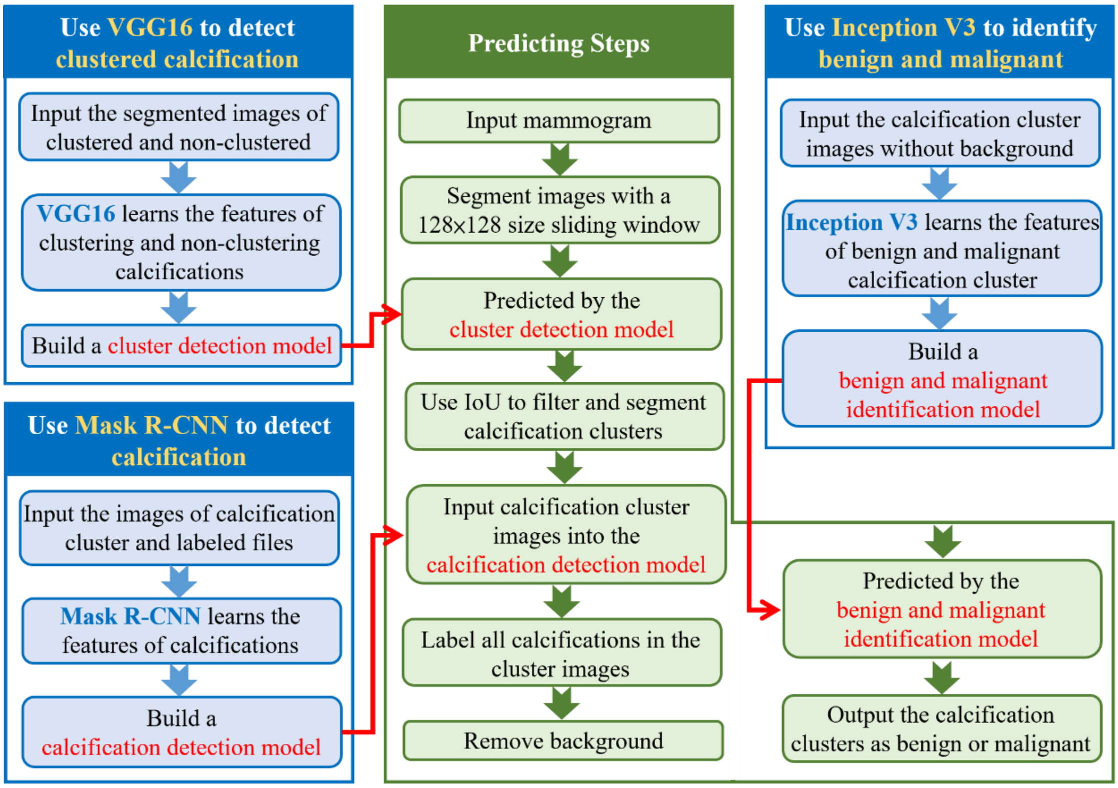
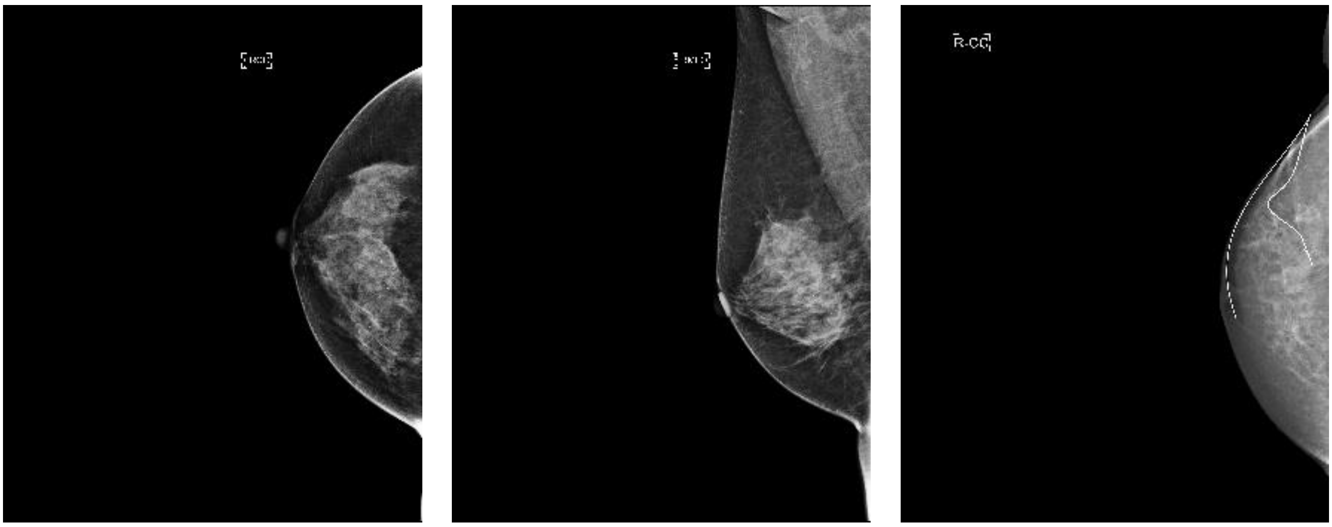
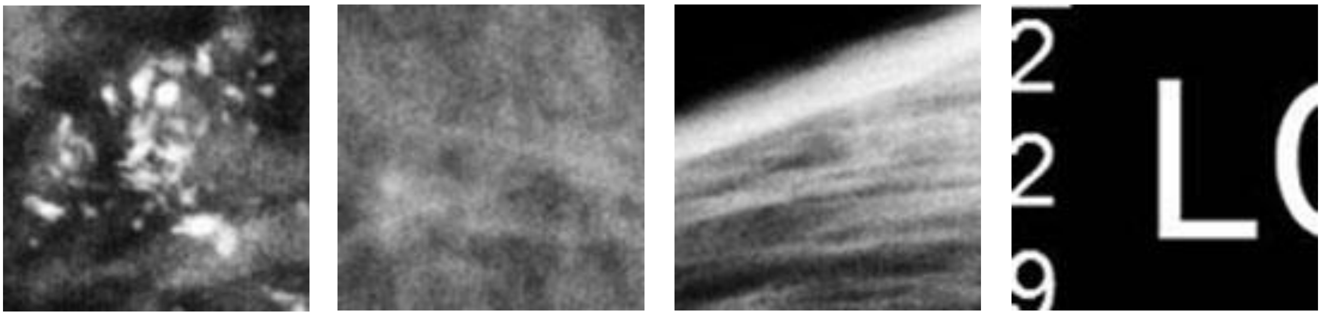


Fig. 1 The flow chart of our proposed method.



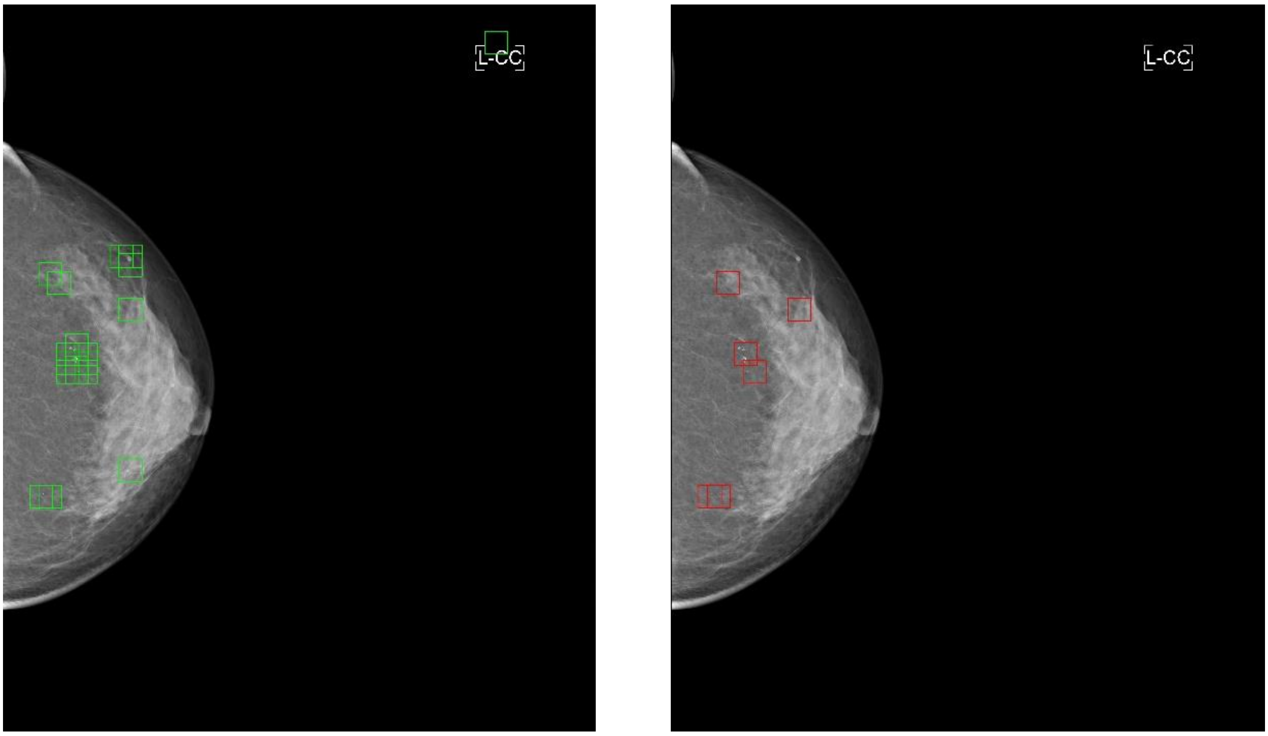
(a) (b) (c)

Fig. 2 From (a) to (c) are CC view, MLO view, and needle localization surgical biopsy image.



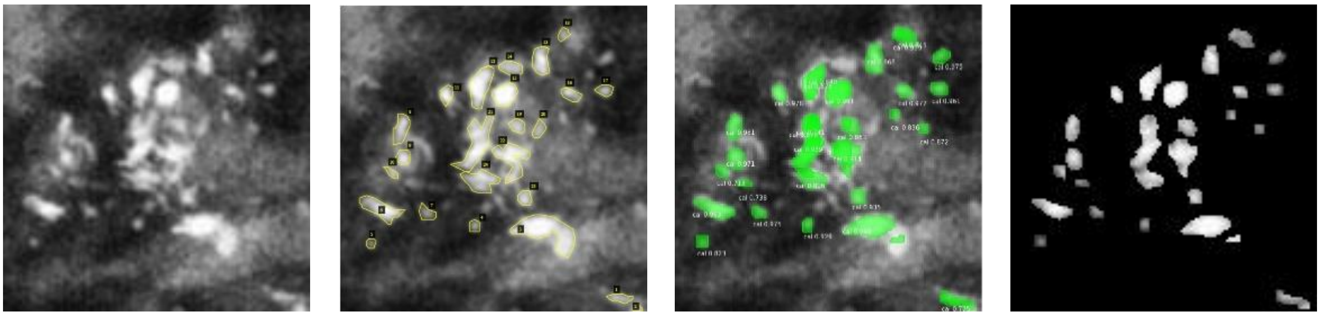
(a) (b) (c) (d)

Fig. 3 (a) shows the cluster image, from (b) to (d) show non-cluster images, including: (b) is the mammae, (c) is the edge of breast, and (d) shows the tag in the mammogram.



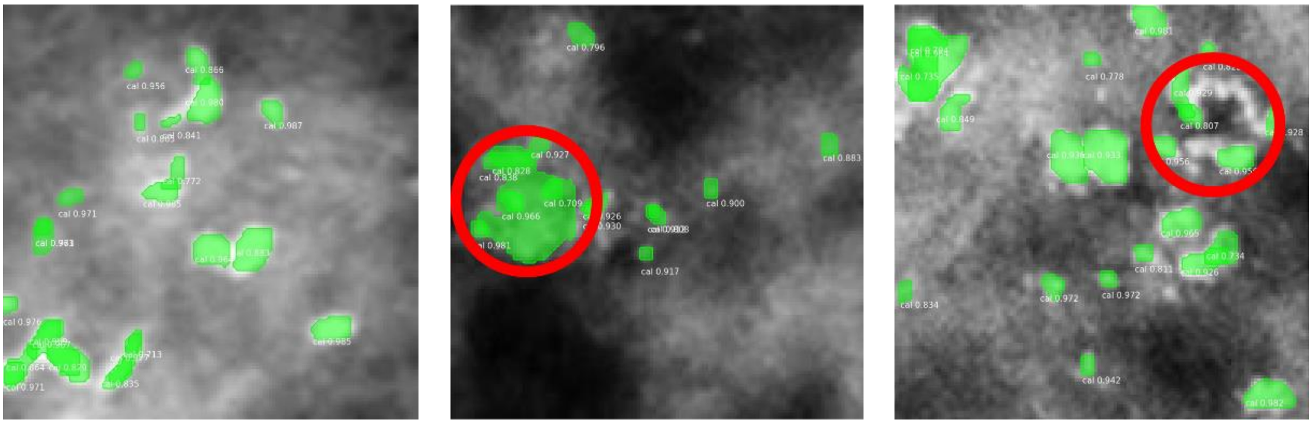
(a) (b)

Fig. 4 (a) shows the cluster candidate boxes on the mammogram, and (b) shows the result of using IoU to filter the cluster boxes on the mammogram.



(a) (b) (c) (d)

Fig. 5 (a) shows the image of calcification cluster, (b) shows an image labeled by VIA, (c) shows an image labeled by Mask R-CNN, and (d) shows the image with the background removed.



(a) (b) (c)

Fig. 6 (a) shows the correct image labeled by Mask R-CNN, which is defined as True Positive (TP). (b) shows the image of Mask R-CNN which incorrectly labeled the background as calcification, which is defined as False Positive (FP). (c) shows the image of Mask R-CNN which misclassified the calcification as the background without labeling the image, which is defined as False Negative (FN).

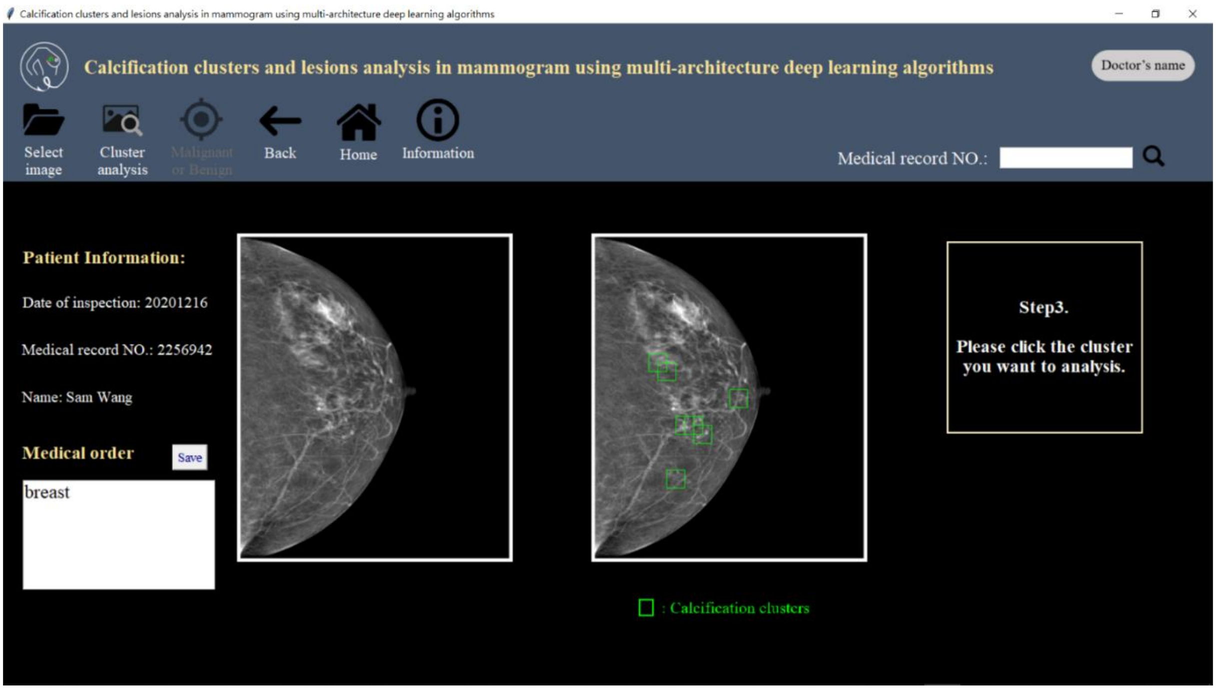


Fig. 7 Analysis result of calcification clusters in our developed system.

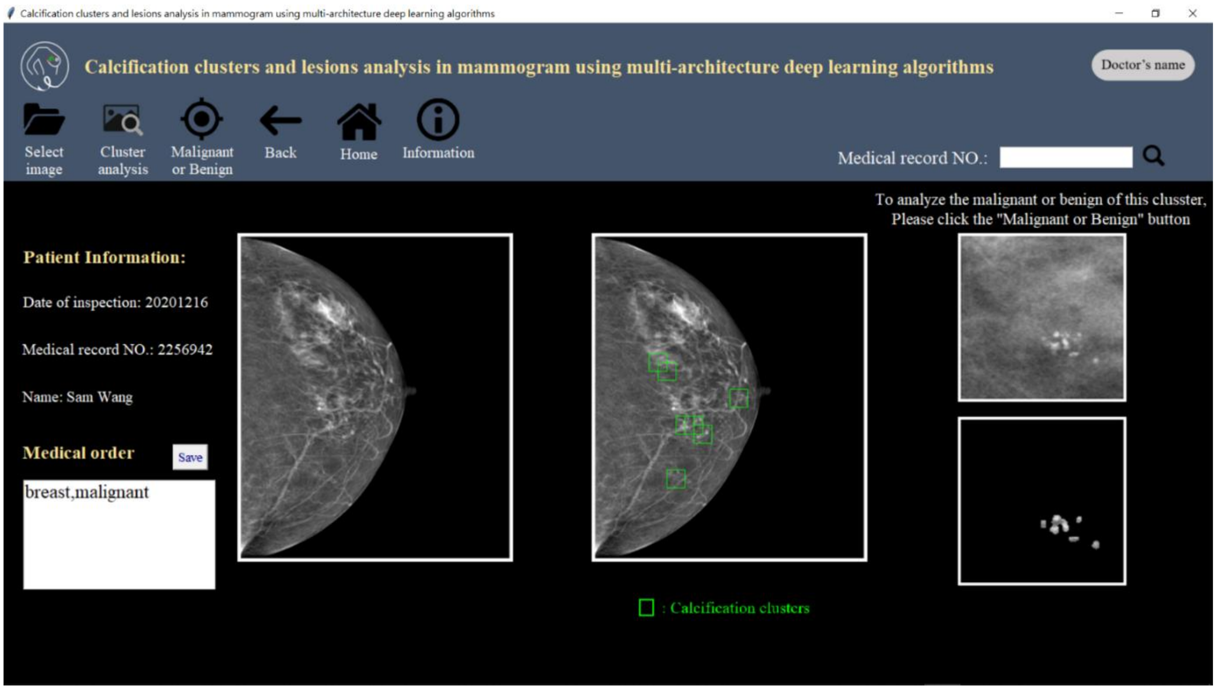
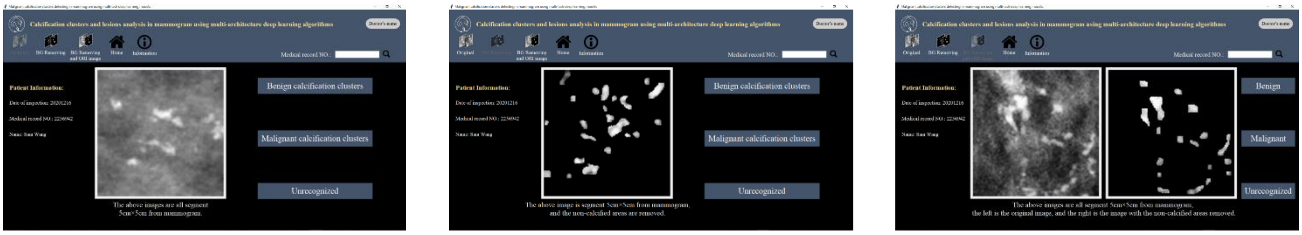


Fig. 8 Benign and malignant analysis result and background removal result.



(a) (b) (c)

Fig. 9 (a) shows the first stage interface of the radiologist test. (b) shows the second stage interface of the radiologist test. (c) shows the third stage interface of the radiologist test.

TABLE 1. THE TABLE OF THE COMBINED RESULTS OF THE THREE ARCHITECTURES

	Accuracy	Precision	Sensitivity	Specificity
Cluster detection model	90.11%	93.63%	85.65%	94.40%
Calcification detection model	-	99.76%	99.28%	-
Benign and malignant identification model	93.76%	88.89%	95.60%	92.64%

TABLE 2. RESULTS OF TESTING PHASE WITH RELATED STUDY

Model	Accuracy	Precision	Sensitivity	Specificity
AlexNet[20]	88.59%	89.32%	88.43%	86.89%
Benign and malignant identification model	93.76%	88.89%	95.60%	92.64%

Chun Shan Medical Journal (CSMJ)

GAN-based SSD Segmentation Algorithm to Assist the Character Recognition of Seven-Segment Display Digits

Jung-Mao Lu(呂榮茂)[1], Chiun-Li Chin(秦群立)[1], I-Miao Chen(陳怡妙)[1],
Ya-Chu Hsieh(謝雅竹)[1], Ting-Chi Chang(張婷淇)[1],Pin-Yu Yeh(葉品郁)[1*]

Abstract：

According to the research, the population of chronic diseases is increasing year by year. Therefore, the demand for measuring vital signs using sphygmomanometers, blood glucose meters and lipid profile machines is also increasing. However, the traditional manual recordingmethod may cause transcription errors and most of the devices on the market that have storageor transmission functions are too expensive. In addition, these medical devices use seven- segment display digits to display the measurement results. The seven-segment display digit contains the discontinuous field and quite different from the printed numbers that may easily cause poor Optical Character Recognition(OCR). Therefore, we propose using GAN to label the discontinuous fields of seven-segment display digits area and using the GAN-based SSD(Seven Segment Digits) Segmentation algorithm to automatically adjust the direction of the images. Then, use a famous OCR tool to recognize the digit value in the image of seven- segment display digits and automatically record the value to complete the monitoring of the systolic blood pressure, diastolic blood pressure and pulse. The experiment result shows that the relevant performance indicators for evaluating trained/tested model are used. Finally, the accuracy of the test stage is 94.5%, the precision is 98.4%, the sensitivity is 90.9%, and the specificity is 98.4%. The recognition accuracy rate is 99%. According to the above data, the method proposed in this paper can effectively improve the success rate of the seven-segment display digits recognition.

Keywords : seven-segment display digits, OCR, GAN, GAN-based SSD Segmentation Algorithm, digits recognition

[1] Department of Medical Informatics, Chung Shan Medical University

- Corresponding Author: Pin-Yu Yeh(pinky88911@gmail.com)

I. Introduction：

After patient was diagnosed with chronic disease, doctor will ask the patient to measure and record the data of their vital sign with a sphygmomanometer or blood glucose meter for follow-up treatment. Patient at home usually use handwriting to record the data display on themeasure instrument. However, this way may cause clerical error, which will affect the follow-up treatment. At present, there are many sphygmomanometers and blood glucose meters with smart transmission function are sold on the market, but they are expensive, and the operation steps are too complicated for the elderly to follow, which is inconvenient to use. Most of the sphygmomanometer and blood glucose meter use seven-segment display digits to display the measured values [1]. If the seven-segment display digits are recognized by general OCR, the recognition rate will be poor due to discontinuities between the numbers [2-3], as shown in Figure 1. Thus, we hope to set up a system which can automatically segment and identify seven-segment display digits. Meanwhile, if the captured image is skewed, it will be automatically corrected and the value in the image will be recognized to provide a basis for physician to conduct clinical diagnosis.

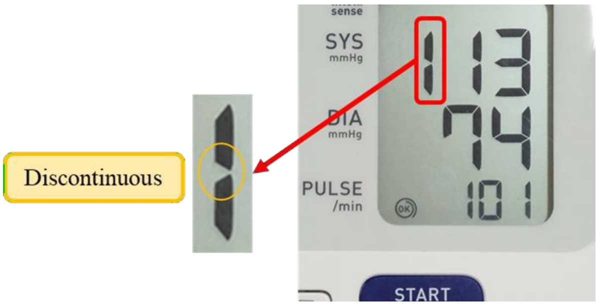


Figure 1. The discontinuous field in seven-segment display digits

II. Related works：

In order to complete the identification of in the seven-segment display digits, this paper conducted a review and discussion of related literature. It can be divided into two parts, namelythe traditional image recognition method, and deep learning method.

1. The traditional image recognition method applied to the digits recognition of the seven-segment display

Because the colored seven-segment display digits image with complex background is difficult to detect and recognize numbers or characters, Popayorm scholars et. al. [4] proposeda predefined HSV color slicing technique to accurately find the areas of digits and characters in the red LED seven-segment display digits image and segment them out. The proposed method can separate the foreground and background of the image to improve the identificationperformance of image of digits and characters of the red LED seven-segment display.

In order to identify the numbers on the digital meter, Kanagarathinam scholars et. al. [5] used the seven-segment display digits images in the “YUVA EB” dataset containing various tilt resolutions and brightness as the developed dataset. Then they used the MSER method fortext detection, and then used OCR for digits recognition. The accuracy of the proposed methodis higher than 93%, and it has a good performance in seven-segment display digits. Moreover,most elderly people or patients with chronic disease often use medical equipment such as bloodpressure monitors and blood glucose meters to measure and record their blood pressure and blood glucose value. Tsiktsiris scholars et. al. [6] have proposed the method of accelerating optical character recognition of seven-segment display digits in order to achieve the purpose of automatically recording the health of the elderly. The proposed method can accelerate the speed of digits recognition, besides the accuracy of OCR can be as high as 96.22%. Therefore,this paper use OCR method to identify the segmented numbers in the blood pressure monitor image.

2. Deep learning method applied to digits recognition

Deep learning has been widely used in the fields of image processing and computer visionto label images and identify the numbers or characters within images. In terms of image labeling, Li [7] proposed a medical image labeling model based on Generative Adversarial Network (GAN), which can be used to reduce the task of manually labeling ground truth imageand speed up the time of image labeling to reduce the time and cost of manually labeling mammography images. Therefore, we refer to the architecture of this paper and apply it to labelseven-segment display digits and characters.

In order to successfully identify the material required for steelmaking—the identificationnumber (BIN) on the billet, Koo proposed a unified deep learning network framework [8], which can detect and identify BIN at the same time. When the image is rotated, the BIN can still be successfully recognized. Hence, we refer to this concept to collect seven-segment display digits with different angles, and apply it to the image recognition of the seven-segmentdisplay digits. Kalyan [9] has proposed an Android application that can use Inception V3 deeplearning model for digits recognition of seven-segment display digits on mobile devices such as smart phones. In order to effectively identify banknote serial numbers with low pixels and complex backgrounds, Wang [10] has developed a smaller Dilat-Net. This network does not use convolutional layers and pooling layers but extracts image features through expanded convolution. Therefore, it can avoid losing the pixel information of the serial number. After that, the adaptive quantization training method is used to determine the percentage of each quantization, and the accuracy can be increased to more than 99%. This method can quickly
and effectively identify the banknote serial number image.

3. GAN deep learning algorithm applied to digits recognition

Recently, due to the excellent performance of GAN in image generation, many scholars were interested in it and make further optimizations. Zhuang et al. [11] proposed to use their improved GAN to regenerate the handwritten text from low-resolution to high-resolution images. In this paper, the details of the text are clearer, and the accuracy of text recognition was improved. Besides, Zhong et al. [12] proposed a DCGAN-based house digits recognition architecture, which used DCGAN to blur the house numbers in street view images to varying degrees to enhance the diversity of the training data set. From the above papers, it can be seenthat GAN is often used in data augmentation or data pre-processing in the related research of deep learning applied to digits recognition, and it can obtain better results.

Based on the above literature, we found that there are few papers on the recognition of seven-segment display digits currently, and many seven-segment display digits images captured with smart phones will have many problems to be solved for recognition such as tilt,insufficient brightness, uneven brightness, etc. Therefore, this paper first proposed to rotate andadjust the brightness of the original blood pressure monitor image captured by the smart phonesto increase the diversity of the data, and then use the GAN deep learning method to label the value of the seven-segment display digits image. After that, the RGB to HSV color space conversion method will be used to segment the values in the blood pressure monitor image byindividually setting the HSV range by setting the HSV range separately. Finally, in terms of number and character recognition, this paper uses the OCR method with excellent recognition results for digits recognition.

The rest of the paper is organized as follows. In Section 3, we will introduce the proposed methodology. Next, in Section 4, we will describe and analyze the experimental results. Finally,in Section 5, we will conclude the presented research works and discuss the impact and future of the project.

III. Method：

In this paper, we will first label the systolic blood pressure, diastolic blood pressure and pulse in the sphygmomanometer image that taken from smart phone as red, green and blue respectively. And then use these images as the ground truth when training GAN to label the digits in the image of sphygmomanometer. Next, we extract the area containing the seven- segment display digits in the labeled image and pre-process the extracted image. Final, we input them into OCR to identify the digits. The flowchart of our proposed method is shown inFigure 2. The method we proposed can be divided into four parts “Input the image of sphygmomanometer”, “Label the image of sphygmomanometer”, “Image pre-processing” and“Optical character recognition”. The detailed description will be explained below.

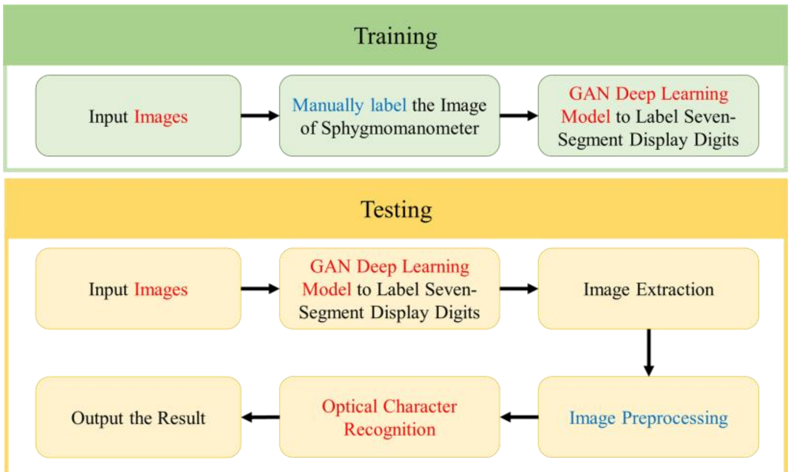


Figure 2. Flowchart of our proposed method

1. Input Images：

In this paper, we collected 1103 images of sphygmomanometer with a smartphone and used them as training dataset. Each image in this dataset contained the digits of systolic bloodpressure, diastolic blood pressure and pulse, where the digits are the seven-segment display digits. In order to increase the diversity of images, we adjusted the angle and brightness of theimages in the training dataset. In angle adjustment part, we rotate the image by 0 degrees, 90 degrees, 180 degrees and 270 degrees. In brightness adjustment part, we lighten the image by
30% and darken them by 30% respectively. After data augmentation, there are 13236 images in the training dataset. In this paper, the 90% of the images that size is 256×256 are used for training and 10% of the images are used for testing.

(a)

(b)

(c)

(d)

(e)

(f)

Figure 3. The image in the training dataset, (a) shows the original image, (b) shows the image with a 30% increase in brightness, (c) shows the image with a 30% decrease in brightness, (d) shows the image rotated 90 degrees to the right, (e) shows the image rotated 180 degrees to the right and (f) shows the image rotated 270 degrees

2. Label the Image of Sphygmomanometer：

2.1 Manually Label the image of sphygmomanometer

In order to correctly identify and label the seven-segment display digits of the patient's systolic, diastolic, and pulse from the image, this paper use the image editing software to labelin this paper. Filling up the discontinuous parts of the seven-segment display digits, and label the systolic blood pressure, diastolic blood pressure and pulse numbers as red, green, and bluerespectively. It is shown in Figure 4. Next, we input the labeled images as ground truth imagefor training.

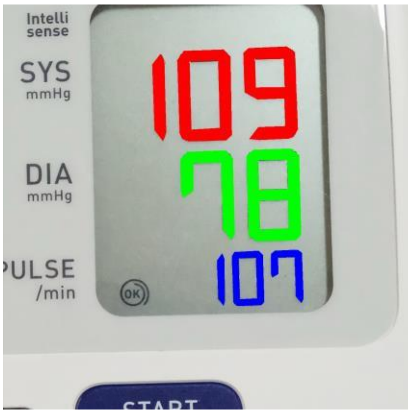


Figure 4. Schematic diagram of labeled seven-segment display digits

2.2 GAN-based labeled seven-segment digits model

In order to improve the accuracy of subsequent digits recognition, this paper use RTGAN deep learning model [13] to label seven-segment display digits. RTGAN can be used to label the contours of different organs in brain CT images. The range of brightness value in each organof brain CT image is different, such as the discontinuity in the seven-segment display digits, so it can be effectively used in this paper.

2.3 Loss function

In this paper, the multiple loss functions are adopted so that the GAN labeling model canaccurately label the seven-segment display digits in the image and assume that the goal of eachloss function is to minimize the total loss function in equation (11). The total loss function consists of five parts, namely Resemble Loss, Competitive Loss, LD Loss, Evaluative Loss and L1 Loss. The function of each loss function will be introduced in detail in the following.

3.2.1 Resemble Loss

The main purpose of Resemble Loss is to measure the similarity between the generator labeled images and the manually labeled images. First, we input the image labeled by generator and the manually labeled image into the VGG-19 network, and thenpass the generated feature maps through the ReLU activation function to generate new feature maps, and then compare the differences of each feature maps through Resemble Loss(1). Resemble Loss is defined as:

$$L_{Resemble} = \frac{1}{HWB}\sum_{j}\sum_{a,b}\sqrt{\left(\tau_{\lambda_t}^{\emptyset_t}\left(K(X_{a,b}) - (Y_{a,b})\right)\right)^2} \tag{1}$$

where H and W represents height and width of the images and B represents batch size, $K(X)$ represents images labeled by the generator and Y represents manually labeling images, a and b are the pixels of X and Y dimensions. $\tau_{\lambda_t}^{\emptyset_t}$ represents the feature map in λ_tth layer of $\emptyset_t$th block in the pre-trained VGG-19 network, and we defined $\emptyset_t \epsilon 4$ and $\lambda_t \epsilon 5$.

3.2.2 Competitive Loss

The main purpose of Competitive Loss is to evaluate the adversarial probability distribution of between the generator and the discriminator in order to minimize the difference between the GAN labeled image and the manually labeled image. In this paper, we use Competitive Loss equation (2) to compare the labeled image produced by generator with the manually labeled image. When $(1 - P_i) and \left(\frac{P_i}{1-P_i}\right)$ are equal, Competitive Loss will equal to zero, that is, the discriminator will regard the labeled image produced by generator and the manually labeled image as the same image. If Adversarial Loss is not equal to zero, GAN will iterate through each training, update the parameters to minimize the loss of cross entropy and make the labeled images produced by the generator as close as possible to the manually labeled images. Competitive Loss equation (2) is defined as:

$$L_{Competitive} = -\sum_{i=1}^{n} S_i\, log(1 - P_i) + S_i\, log\left(\frac{P_i}{1 - P_i}\right) \tag{2}$$

where P_i reprensents the probability that the discriminator identifies the real image, S represents the discriminator.

3.2.3 LD Loss

The main purpose of LD Loss is to compare whether the feature value weight setting of the generator labeled image and the manually labeled image can make the similarity between the two higher. In this paper, LD Loss calculated as the Euclidean distance between the labeled image produced by the generator and the manually labeled image, and we use Laplacian of Gaussian ($LoG(\cdot)$) to detect the edge of the image, which is equivalent to adding a convolutional layer to CNNs with LoG kernel. LD Loss equation (3) is defined as:

$$L_{LD} = \sum_{i}\sqrt{\left(T(K(X_i))\right)^2 - \left(T(Y_i)\right)^2} \tag{3}$$

where $T(K(X_i))$ and $T(Y_i)$ are enhanced images of X and Y:

$$T\left(K(X_{(i,j)})\right) = \sum_{a,b} LoG(a,b) \times X(i+a, j+b) \tag{4}$$

LoG algorithm (5) is as follows, and defines σ=1.4:

$$LoG(a,b) = \frac{\partial^2 I(a,b)}{\partial a^2} + \frac{\partial^2 I(a,b)}{\partial b^2} \tag{5}$$

3.2.4 Evaluative loss

In this paper, we use two image quality evaluation metrics to evaluate similarity between labeled and raw images, namely MS-SSIM and PSNR. MS-SSIM equation (6) is used to estimate the overall similarity between two images x and y.

The GAN model in [13] uses SSIM Loss to evaluate the pros and cons of the training model. However, SSIM is a single-scale algorithm, which is more suitable for images where the shooting distance is fixed or does not change much, such as the brain CT images used in this paper. Therefore, we refer to the [15] and change SSIM to Multi-Scale SSIM (MS-SSIM). MS-SSIM is scale-invariant and more suitable for the image of seven- segment display digits used in the paper, which include images with various shooting distances and rotated images. MS-SSIM Loss is defined as:

$$L_{MSSSIM} = [l_M(X,Y)]^{\alpha^M} \times \prod_{j=1}^{M} [c_j(X,Y) \times s_j(X,Y)]^{\alpha_j} \tag{6}$$

Where l_M represents the function used to obtain the average of the grayscale values of the input image, c_j represents the function used to obtain the standard of the grayscale values of the input image and s_j represents the function used to normalize the grayscale values of the image. PSNR equation (7) is defined by Mean Square Error (MSE), which can find different pixel values between the labeled image produced by generator and the manually labeled image, where N_{MAX} represents the maximum possible pixel value in the image; in equation (8), Na and Nb represent the width and height of the input image, respectively.

$$L_{PSNR} = -\sum_i 20 \cdot log\,(\frac{N_{MAX}}{\sqrt{MSE(K(x),Y)}}) \tag{7}$$

$$MSE(K(X),Y) = \frac{\sum_{a=1}^{N_a}\sum_{b=1}^{N_b}\big(K(X)(a,b) - Y(a,b)\big)^2}{N_a N_b} \tag{8}$$

In this paper, the PSNR and MS-SSIM are combined to obtain Evaluative Loss Function. After many experiments, we found that the best results can be obtained when SSIM = 1.5 and PSNR = 1. Evaluative Loss equation (9) is defined as:

$$L_{Evaluative} = 1.5 \cdot L_{MSSSIM} + 1 \cdot L_{PSNR} \tag{9}$$

3.2.5 L1 Loss

In [14], the effectiveness of multiple Loss Function and the combination of different Loss Function are analyzed. They found that MS_SSIM only retains the brightness of the image, and it is easy to cause chromatic aberration in the image. But, L1 Loss retains color information. Therefore, this paper use L1 Loss to make up for the shortcomings of MS_SSIM to obtain better results. L1 Loss is defined as follows:

$$L_{L1} = \sum_{i=1}^{n} |Y - T_\theta(X)| \tag{10}$$

3.2.6 Total Loss

This paper integrates the Resemble Loss, Competitive Loss, LD Loss, Evaluative Loss and L1 Loss into Total Loss(L_{Total}). The calculation method of Total Loss is shown in equation (11).

$$L_{Total} = 1 \cdot L_{Resemble} + 1 \cdot L_{Competitive} + 5 \cdot L_{LD} + 1 \cdot L_{Evaluative} + 1 \cdot L_{L1} \tag{11}$$

3. Image Preprocessing

3.1 Image Extraction

In order to extract the digits of red color representing the systolic blood pressure, the digits of green color representing the diastolic blood pressure and the digits of blue color representing the pulse pressure in the labeled image, this paper convert the RGB image into the HSV image and take out the value of hue in the image as the basis for distinguishing three sets of values. According to the range of the hue of red color, green color, and blue color, separately distinguishes the values of hue of three groups of number in the image as so as to extract the values of the systolic blood pressure, diastolic blood pressure and pulse pressure in the image. Moreover, the region without value is filled with white color. The result is shown in Figure 5, the number on the seven-segment display is extracted from the image.

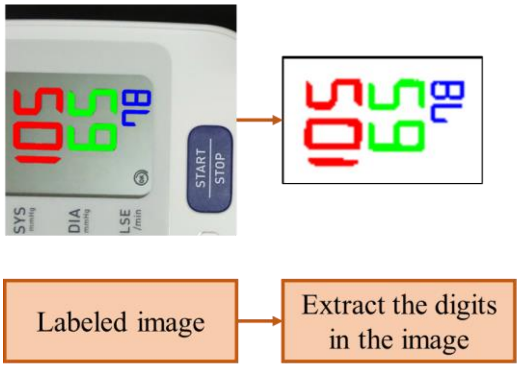


Figure 5. Schematic diagram of extracting digits in the seven-segment display digits image

3.2 GAN-based SSD Segmentation Algorithm

Since various angles are generated when taking the each of sphygmomanometer image, this will cause the digits in the image to be skewed. The skewed images are shown in Figure 6 (a) and (b). In order to improve the accuracy of recognizing digits on the seven-segment display, the angle of inclination of numbers in the image must be the same. Therefore, the image will be rotated with three stages in this paper to ensure the angle of each number in the image to be the same. The detailed method will be described as the follows.

(a) (b)

Figure 6. (a) and (b) show seven-segment display digits with skewed

3.2.1 Vertical Correction

This paper will perform vertical correction for image containing three sets of values. First, the image will be rotated clockwise. The maximum angle of rotation is 180 degrees. After each rotation, the X coordinate value of the non-white pixel in the image is obtained by vertical projection. Next the width of the non-white area can be calculated using equation (12). When the width of the non-white region is the minimum, it means that the angle of the groups of number in the image can be arranged vertically. The result is shown in Figure 7 (a) and (b).

(a) (b)

Figure 7. Schematic diagram of number arranged vertically,
(a) shows the image with arrangement vertically and the direction is not reverse,
(b) shows the image with arrangement vertically and the direction is reverse

$$W = Max(x_c) - Min(x_c) \tag{12}$$

In equation (12), x_c is defined as the X coordinate of the non-white pixel in the image, and W is defined as the difference between the maximum and minimum of x_c, which represents the width of the non-white region after the image is rotated.

3.2.2 Orientation Correction

As can be seen from Figure 7, the number in the seven-segment display digits image may be upside down. In order to improve this situation, the image will be rotated through equation (13) and (14). Then, calculating the value of the image before and after the rotation can be obtained using equation (15). If the value before the rotation is greater than the value after rotation, it means that the image is not upside down. On the contrary, it means that the image is upside down. In this case, the image will be corrected by inversion.

$$x = \cos(180)\,x - \sin(180)y \tag{13}$$

$$y = \sin(180)\,x + cos\,(180)y \tag{14}$$

$$V = \Sigma(y_{RT} - y_{BD})^2(\Sigma X_{Ri})^2 + \Sigma(y_{BT} - y_{RD})^2(\Sigma X_{Bi})^2 \tag{15}$$

In the equation of (13), (14) and (15), x represents X coordinate of each pixel in the image which rotated by 180 degrees, y represents Y coordinate of each pixel in the image which rotated by 180 degrees, V represents an index used to determine whether the image is upside down, y_{RT} represents the maximum of Y coordinate of all red pixels in the image, y_{BD} represents the minimum of Y coordinate of all blue pixels in the image, y_{BT} represents the maximum of Y coordinate of all blue pixels in the image, y_{RD} represents the minimum of Y coordinate of all red pixels in the image, X_{Ri} represents the sum of the red pixel in the ith column of the image, and X_{Bi} represents the sum of the blue pixel in the ith column of the image.

3.2.3 Fine Tuning

In order to make the subsequent recognition process more accurate, we will rotate the seven-segment display digits image for the third time to confirm that the seven-segment display digits image can obtain the best rotation angle. Therefore, this paper will use the equation (16) and (17) to rotate the image by 30 degrees left and right. Then, the histogram of the three sets of values after horizontal projection is calculated by equation (18), and uses equation (19) to determine whether there are two horizontal line between the three sets of values. If the horizontal line can be found, it will be recorded as 1, otherwise, it will be recorded as 0. Finally, we will use the result of equation (20) to compare the area of circumscribed rectangle formed by the three sets of values. If the area of the rectangle is the smallest, it means that this angle is the best rotation angle.

$$x_j' = \cos(j)\,x_1 - \sin(j)y_1\,, -30 \le j \le 30 \tag{16}$$

$$y_j' = \sin(j)\,x_1 + cos\,(j)y_1\,, -30 \le j \le 30 \tag{17}$$

$$P_{ij} = \Sigma X_{ij}' \tag{18}$$

$$L = \begin{cases} 0, P_i > 0 \\ 1, P_i = 0 \end{cases} \tag{19}$$

$$Area = (Max(x_i') - Min(x_i'))(Max(y_i') - Min(y_i')) \tag{20}$$

In the equation of (16) to (20), x_j' and y_j' represents the x-coordinate and y-coordinate after X rotates j degrees; P_{ij} represents the histogram of the row i after image rotates j degrees; L is used to record whether the horizontal line can be found in the row i; $Area$ represents the smallest area of the circumscribed rectangle formed by the three sets of values.

3.3 Image Inpainting

The image of seven-segment display digits captured by smartphone may have the situation of overexposure or underexposure due to exposure or insufficient light, it leads to the unsatisfactory labeling results of the GAN deep learning model, such as incomplete labeling or incorrect color labeling. These problems may affect the effect when recognize the seven-segment display digits. In order to improve the accuracy of recognizing the value of the seven-segment display digits, this paper automatically correct the orientation of the captured digits and then superimpose them with the original image to fill in the incompletely labeled seven-segment display digits image. The result is shown in Figure 8.

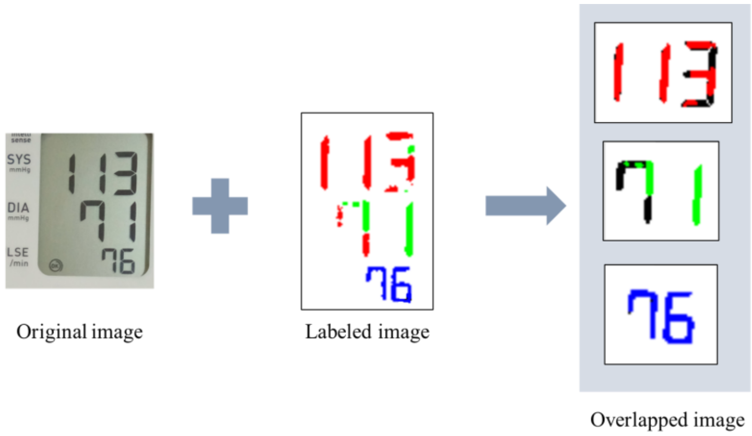


Figure 8. Schematic diagram of the overlapped image

4. Optical Character Recognition

In order to recognize the seven-segment display digits in the image, we adopt Optical Character Recognition (OCR) which was trained by numerous and diverse text images [16] as our recognition method in this paper. Its training images also includes seven-segment display digits, Chinese characters, punctuation and so on. Hence, it is superior to traditional OCR in recognition of seven-segment display digits image. First, the image will be binarized and input it into Tesseract for recognition. Then, it will obtain the position of each object in the image through vertical projection and horizontal projection to segment the text, digit and punctuation in the image and then erode each object to preserve the important information and remove noise or other data to avoid misjudgment during recognition. Next, it will extract the features of the objects in the image by statistical methods and compare with the characters in the database. Finally, the output will be the characters that have the highest similarity.

IV. Result and Discussion

1. Result

In this paper, our proposed method was trained on a personal computer of Windows operating system with 16GB RAM, i78750HQ CPU and NVIDIA GTX 1080Ti GPU for labeling seven-segment digits image. Among them, there are 11912 images for training datasets and 1324 images for testing datasets.

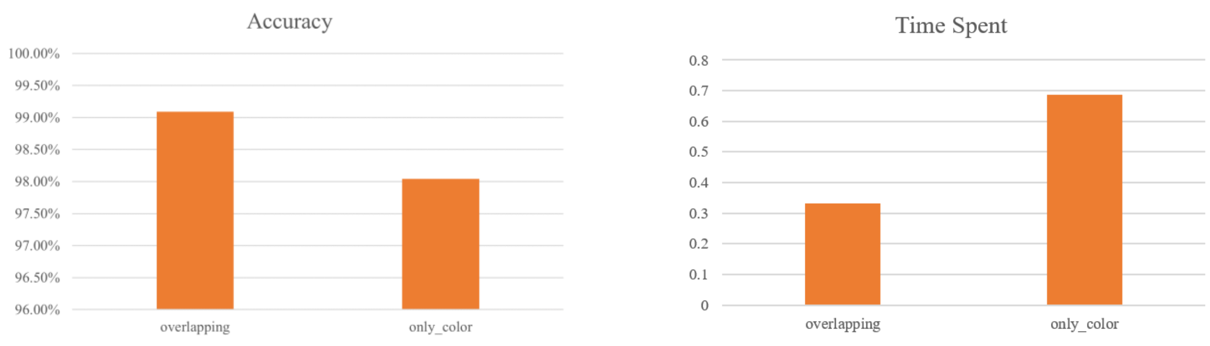
In the experimental results, it will be divided into two parts, including the result of GAN model labeling and recognizing the digit in the seven-segment display image respectively. To improve the accuracy of recognition, this paper will superimpose the image which labeled by GAN model with the original image and compare the accuracy of recognition before and after overlap.

In order to evaluate the effectiveness of GAN applied to the labeling of seven-segment display digits image, we calculate the value of confusion matrix as the evaluating index of experimental result in this paper. In the case of 300987 epochs and 2 batch sizes, the test accuracy rate is 94.5% in the evaluation index calculated from the confusion matrix of labeling seven-segment display digits image with GAN, which can prove that GAN model can label digit in the seven-segment display image accurately, and the labeled color in the image is correct. The precision of testing is 98.5%. It shows that the seven-segment digits of position and color labeled by GAN model are mostly correct. In addition, the sensitivity of testing is 90.9%. It indicates that GAN model can label number correctly. Finally, the specificity of testing is 98.4%. It indicates that in the region that is actually a non-seven-segment display digit image, the proportion of correctly labeled as high as 90%. Thus, it can be seen that GAN labels the digits on seven-segment display image having quite good result which proves that our proposed method can label digit accurately in the image and fill in the discontinuity in the seven-segment display digits image.

Figure 9 shows the results of GAN labeling the digits on seven-segment display digits image. The upper row is the original image and the labeled image with better result and the lower row is the original image and the labeled image with poor result. It can be found from the image that when the brightness of the image is too high or too low, the result of GAN labeling the digits on seven-segment display image may be poor. For example, the systolic blood pressure that should be labeled as red may be mixed with green and the diastolic blood pressure that should be labeled as green may be mixed with red which are shown in the circled place in Figure 9.

Figure 9. Results of GAN labeling the digits on seven-segment display image

This paper calculates the seven-segment display digits recognition accuracy rate in this paper. The recognition accuracy rate of the labeled image that are not overlapped with the original image is 98.04%. On the contrary, the digits recognition rate of superimposing the original image reaches 99.09%. In addition, if the image is not superimposed with the original image, the recognition time is about 0.7 seconds, and the recognition time for superimposing with the original image can be reduced to 0.3 seconds, and organize the above data into a bar graph of the recognition accuracy of whether it overlaps with the original image, as shown in Figure 10. Therefore, it is proved that superimposing the GAN-labeled image with the original image can effectively improve the accuracy of OCR digits recognition and it also can reduce the subsequent recognition time.



(a) (b)

Figure 10. The bar graph comparing whether the image overlapping with original image, (a) shows the bar graph of accuracy, (b) shows the bar graph of time spent

2. Discussion

In addition to using the GAN deep learning algorithm to label seven-segment display digits image, this paper also compares with the method using in the paper of Mask R-CNN [17]. It can be seen that Mask R-CNN has a good performance in segmenting and labeling different objects from the experimental results. Therefore, this paper uses this method to label the numbers in the image of blood pressure monitor, and the result of using Mask R-CNN and our proposed method to label seven-segment numbers are as shown in Figure 11. As it can be seen from Figure 11, Mask R-CNN usually shows the deviation of the position of the mask from the actual position of the seven-segment digits. As the row 4 from Figure 11, the edge of the mask generated by Mask R-CNN cannot completely fit the edge of the actual digits. It is guessed that the possible reason is the discontinuity in the seven-segment display digits image, which affects Mask R-CNN hard to label the edge of seven-segment display digits image. Compared with GAN deep learning algorithm, the labeling effect still needs to be improved.

Figure 11. Comparing the result image of Mask R-CNN and GAN,
left shows the result image of GAN, right shows the result image of Mask R-CNN

We also compare the experimental results of Mask R-CNN with the results of seven-segment display digits labeled by GAN by calculating the relevant indicators of the confusion matrix. Table 1 shows the evaluation indicators calculating by confusion matrix of GAN and Mask R-CNN, and it can be seen from the line chart in Figure 12 that the accuracy, precision, sensitivity, and specificity of using GAN model to label images of blood pressure monitor are better than those of Mask R-CNN model does, which proves that the method proposed in this paper can effectively and correctly label the seven-segment display digits from the image.

Table 1. Confusion matrix of testing of GAN and Mask R-CNN

Confusion Matrix of Testing	Testing result			
	Accuracy	Predision	Sensitivity	Specificity
GAN	94.51%	98.49%	90.9%	98.4%
Mask R-CNN	87.50%	90.47%	86.36%	88.88%

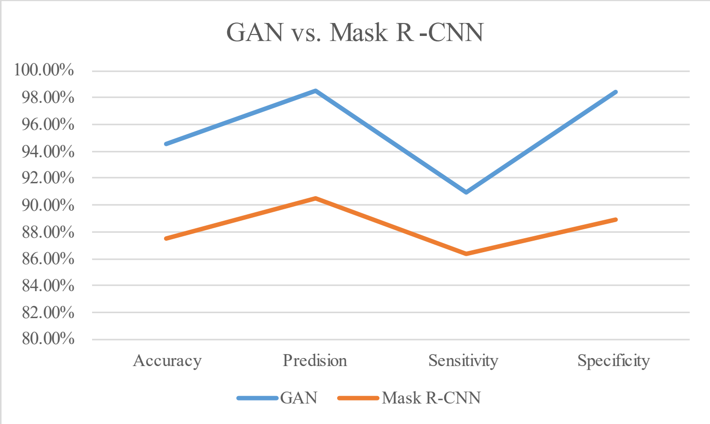


Figure 12. The line chart of the testing result of GAN and Mask R-CNN

V. Conclusion

As the population of patients with chronic diseases increases year by year, the number of people who use medical equipment to measure vital signs has gradually increased. Nowadays, medical equipment often uses seven-segment display digits to display measurement results. However, due to the discontinuity of the seven-segment display digits, patients cannot use the existing identification tools for identification and recording. Therefore, this paper proposes to use the GAN deep learning method to automatically label seven-segment display digits to completely cut-out part and automatically correct the slope angle of the image according to the GAN-based SSD Segmentation Algorithm. Then this paper uses OCR to identify the labeled image to increase the success rate of digit recognition. In the future, we hope to improve and reduce the steps of image pre-processing to reduce the time spent on overall recognition.

VI. Reference

[1] E. Finnegan, M. Villarroel, C. Velardo, and L. Tarassenko: Automated method for detecting and reading seven-segment digits from images of blood glucose metres and blood pressure monitors. Journal of Medical Engineering & Technology 2019; 43(6): 341-355.

[2] M. Namysl, and I Konya: Efficient, Lexicon-Free OCR using Deep Learning. Computer Vision and Pattern Recognition 2019: 295-301.

[3] Q. Ye, and D. Doermann: Text detection and recognition in imagery: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 2015; 37(7): 1480-1500.

[4] S. Popayorm, T. Titijaroonroj, T. Phoka, and W. Massagram: Seven Segment Display Detection and Recognition using Predefined HSV Color Slicing Technique. 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE) 2019: 224-229.

[5] K. Kanagarathinam, and K. Sekar: Text detection and recognition in raw image dataset of seven segment digital energy meter display. Energy Reports 2019; 5: 842-852.

[6] D. Tsiktsiris, K. Kechagias, M. Dasygenis, and P. Angelidis: Accelerated Seven Segment Optical Character Recognition Algorithm. IEEE 2019 Panhellenic Conference on Electronics and Telecommunications (PACET) 2019: 1-5.

[7] J. Li, and Y. Hong: Label Generation System Based on Generative Adversarial Network for Medical Image. Proceedings of the 2nd International Conference on Artificial Intelligence and Pattern

Recognition 2019: 78-82.

[8] G. Koo, J. Yun, H. Choi, and S. Kim: Unified deep neural networks for end-to-end recognition of multi-oriented billet identification number. Expert Systems with Applications 2021; 168.

[9] V. Kalyan, M. Katti, and S. Pai: Application of Convolutional Neural Networks in Mobile Devices for Inferring Readings from Medical Apparatus. International Journal of Research and Scientific Innovation (IJRSI) 2017; 4: 49-52.

[10] F. Wang, H. Zhua, W. Li, and K. Li: A hybrid convolution network for serial number recognition on banknotes. Information Sciences 2020; 512: 952-963.

[11] Q. Zhuang, K. Huang, Q. F. Wang, J. Xiao, and R. Zhang: Generative adversarial classifier for handwriting characters super-resolution. Pattern Recognition 2020; 107.

[12] J. Zhong, J. Gao, R. Chen, and J. Li: Digital recognition of street view house numbers based on DCGAN. International Conference on Image and Graphics Processing 2019: 19-22.

[13] Ya. Hsieh, H. Tseng, C. Chin, and Y. Shao: Based on DICOM RT Structure and Multiple Loss Function Deep Learning Algorithm in Organ Segmentation of Head and Neck Image. Future Trends in Biomedical and Health Informatics and Cybersecurity in Medical Devices 2019: 428-435.

[14] H. Zhao, O. Gallo, I. Frosio, J. Kautz: Loss Functions for Image Restoration With Neural Networks. IEEE Transactions on Computational Imaging 2017; 3.

[15] D. Bang, H. Shim: Improved Training of Generative Adversarial Networks Using Representative Features. Proceedings of the 35th International Conference on Machine Learning 2018; 80.

[16] E. Zacharias, M. Teuchler, and B. Bernier: Image Processing Based Scene-Text Detection and Recognition with Tesseract. Computer Vision and Pattern Recognition 2020; 2004.

[17] K. He, G. Gkioxari, P. Dollár, R. Girshick: Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV) 2017: 2980-2988.

輔助七段顯示數字之字元辨識

呂榮茂 [1] 秦群立 [1] 陳怡妙 [1] 謝雅竹 [1] 張婷淇 [1] 葉品郁 [1*]

摘要：

根據研究顯示患有慢性疾病的人口逐年增加，因此使用血壓計、血糖計及血脂機等醫療器材測量相關生命表徵的需求也越來越高，然而傳統的手動紀錄方式可能造成謄寫錯誤，且目前市面上具有儲存或傳輸功能的器材，價格大多過於昂貴，此外，這些器材大多使用七段顯示數字作為呈現數據的方式，然後這樣的數字中有不連續的狀況發生，與印刷體的數字有較大的差異，容易造成光學字元辨識(Optical Character Recognition, OCR)辨識效果不佳，因此本論文提出利用 GAN 標記七段顯示數字不連續的區域，並利用 GAN-based SSD(Seven Segment Digits) Segmentation 演算法自動調整影像的方向，接著透過著名的 OCR 辨識工具辨識標記及校正後的七段顯示數字影像，以此自動紀錄病患的收縮壓、舒張壓及脈搏。在實驗結果的部分，我們計算 GAN 標記影像之混淆矩陣的相關指標，其中測試階段的準確率為 94.5%、精準度為 98.4%、敏感度為 90.9%以及特異度為 98.4%，而辨識數字的準確率則有 99%，由上述數據可知，本論文提出之方法可有效提升七段顯示數字的辨識成功率。

關鍵字：七段顯示數字、OCR、GAN、GAN-based SSD Segmentation 演算法、數字辨識

[1] 中山醫學大學 醫學資訊學系碩士班
[2] 中山醫學大學 醫學資訊學系
[*] 通訊作者：葉品郁
地址：台中市南區建國北路一段 110 號
電話：+886-961-575-886
信箱：pinky88911@gmail.com

2020 International Conference on Fuzzy Theory and Its Applications (iFUZZY2020)

Combining VGG16, Mask R-CNN and Inception V3 to identify the benign and malignant of breast microcalcification clusters

1st Ya-Chu Hsieh
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
as26872076@gmail.com

2nd Chiun-Li Chin
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
ernestli@csmu.edu.tw

3rd Chia-Shin Wei
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
s0658028@gm.csmu.edu.tw

4th I-Miao Chen
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
loveyou890121@gmail.com

5th Pin-Yu Yeh
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
pinky888911@gmail.com

6th Ru-Jiun Tseng
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
a0903031822@gmail.com

***Abstract*—Breast cancer has the highest incidence among women in Taiwan, which is mainly screened by mammography. When the doctor observes the mammogram and initially judges that the patient has malignant microcalcification (MC) clusters, the patient must undergo needle localization surgical biopsy. However, needle localization surgical biopsy makes the patient painful, and the color of breast tissue and MCs are all white, which makes it difficult for doctors to judge where MCs are clustered immediately. Thus, we use VGG16 to find out breast MC clusters from the image. Moreover, we use Mask R-CNN to find MCs from the clusters to remove the noise from the background. Finally, we use Inception V3 to identify the benign and malignant of MC clusters. The accuracy of the cluster classification, MCs labeling and benign and malignant analysis are 93%, 95% and 91%. Furthermore, the precision, specificity and sensitivity of our proposed methods are about 87%, 89% and 90%, respectively. It proved that our system can effectively assist doctors in diagnosing and reduce the burden on patients and medical personnel.**

***Keywords*—mammography, microcalcification (MC), VGG16, Mask R-CNN, Inception V3**

I. Introduction

When diagnosing, doctors will first find out the place of breast MC clusters, and according to the size, distribution and the shape of the MCs to determine whether they are clusters of malignant breast MCs [1-2]. When the doctor judges that there are clusters of malignant MCs, the patient needs to undergo needle localization surgical biopsy to further confirm that the clusters of MCs are benign or malignant. However, this process takes a lot of time, and the features of breast tissue and MCs are similar, and they are all white in X-ray images. Thus, it is difficult for doctors to observe clusters of MCs from mammography images immediately, and the images must be magnified to judge. Moreover, needle localization surgical biopsy is an invasive test method, which may cause the breast location of patient painful and reduce the willingness of the patient to receive the detection as soon as possible, causing the patient to miss the best time for treatment. In this paper, we use VGG16 [3-4] to find out the place of MC clusters from the place of mammography image, and Mask R-CNN [5-6] to find all MCs from the MC clusters. Finally, we use Inception V3 [7] to classify the benign and malignant from MC cluster. We use the above three methods to achieve the goal of automatically analyzing mammography images to find the malignant MCs from clusters. Furthermore, there are many people have used deep learning methods in their papers to mark breast MCs, but they still cannot classify benign and malignant MCs cluster in mammography images. Thus, we use deep learning algorithms to simulate visit steps of doctors, analyze and diagnose quickly, and precisely to provide doctors with reference information, which makes doctors make a diagnosis more confidently. Moreover, patients do not need to undergo needle localization surgical biopsy and doctors can also find out the location of lesion, reducing the psychological burden on patients and eliminating invasive treatment steps.

II. Related Works

Breast cancer is one of the most common diseases among women. Patients usually need to undergo mammography screening and needle localization surgical biopsy before so that the doctor can make a diagnosis. However, needle localization surgical biopsy is an invasive detection method, which may leave wounds and pain on the patient, and this also reduces the willingness of patient to receive the detection, leading to delays in treatment time. With the advance of deep learning methods, there are many people applying deep learning in various fields for classification or segmentation in their papers. Thus, we refer to many references in this paper and use deep learning technology to detect clusters of malignant MCs in mammography images. It is hoped that it can actually help doctors speed up the time of diagnosis and treatment, and help doctors formulate the most appropriate treatment course.

Zhao et al. [8] proposed to use Multi-Scale VGG16 architecture to divide lung nodules into benign and malignant in lung CT images in order to improve the traditional diagnosis method of lung nodules, to prevent radiologists from misjudgment due to fatigue and other human factors. It proves that VGG16 has a high accuracy rate in classifying benign and malignant lung nodules, and it also proves that VGG16 can accurately classify medical images in this paper.

Qinhua Hu et al. [9] proposed an automatic segmentation method for lungs in CT images using Mask R-CNN, and combined with Support Vectors Machine (SVM), K-means and Gaussian Mixture Models (GMMs), which produced the best results for lung segmentation. In this paper, it has proved that using Mask R-CNN in medical image segmentation has a very good performance.

Dong et al. [10] have proposed a cell identification algorithm that combines Inception v3 and manual feature extraction. It is mainly used to classify cervical cancer cells in cell images. It is hoped that deep learning techniques can be used to improve the manual inspection of cervical cancer cells and reduce the risk of human negligence. It proves that Inception v3 improves the accuracy of classification and identification of cervical cancer cells in this paper. It also proves that Inception v3 can extract more subtle features from medical images, effectively improving the accuracy of medical image classification.

Based on the above literature, we use VGG16 to classify mammography images into clustered MCs and non-clustered MCs images in this paper, and then input the clustered MCs images into the Mask R-CNN model to segment the MC clusters in the image. Finally, we use Inception V3 to classify the benign and malignant of clustered MCs on the breast image.

III. METHODS and Results

In this paper, we collaborate with professional doctors from the Medical Imaging Department, we have built a system that uses VGG16, Mask R-CNN and Inception V3 to detect cluster MCs in the breast and analyze the benign and malignant of cluster MCs. First, we will learn the clusters of calcification points according to the fine needle positioning images and mammography images provided by the doctor, and segment clusters and non-clusters, as the training datasets of the VGG16 model, and submit the classified result images to the professional doctor to manually mark the clusters, and then the system will use the Mask R-CNN model to extract the features of clustered MCs in the mammography image, and detect the possible locations of the clusters of MCs in the mammography image, and output the labeled mammography image, and then remove the background outside the breast MCs through a self-built algorithm. Finally, the Inception V3 model is used to classify the clusters of MCs in the mammography image labeled by Mask R-CNN to classify the clusters of benign or malignant MCs. Fig. 1. shows the flowchart of our proposed system.

3.1 Input Mammography Image

This paper is supported by the professional doctors of the affiliated hospital of Chung-Shan Medical University, and the doctors in the Medical Imaging Department provide mammography images and professional knowledge. There are 1600 images in the datasets, including CC view and MLO view. After deducting the images with artifacts, there are 1586 images. The 80% of the data is used for model training, and the 20% of the data is used to evaluate the accuracy of the model.

3.2 Resize and Sharpen Image

First, we will extract the clustered and non-clustered images in a size of 128×128. Next, in order to avoid the problem of poor image contrast caused by images taken by different equipment, leading to the contours of breast MCs are not obvious, which easily affects the recognition effect of the model, resulting in low accuracy of model recognition. Thus, we will sharpen mammograms to make the edge contours of MCs more obvious in this paper.

3.3 VGG16 Taxonomic Cluster Calcification

VGG16 is composed of thirteen convolutional layers, five pooling layers and three fully connected layers. In order to automatically find the clusters of MCs in the mammography image, we will first learn the clusters of MCs according to the fine needle positioning image provided by the doctor in this paper, and compare the mammography images to extract the clusters with a size of 128×128. Then, capturing the non-clustered areas in the image with the same amount, such as mammary, nipples, and tags, and then input them into VGG16 model for classification and non-clustering training. Fig. 2. shows the result image that marking the clusters of MCs.

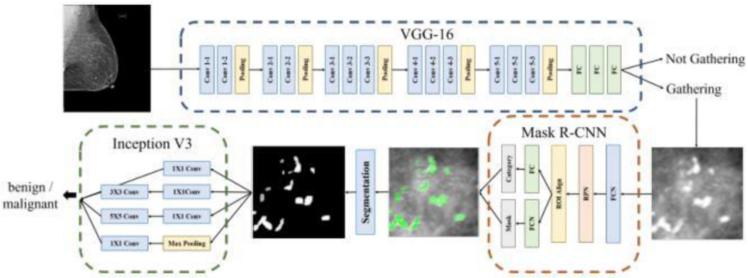


Fig. 1. The flowchart of our proposed system

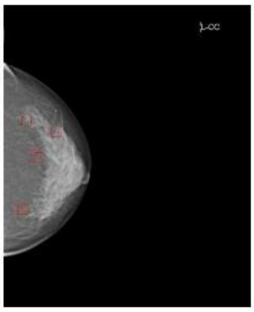
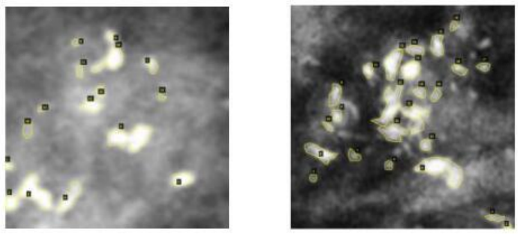


Fig. 2. Marking result of the clusters of MCs through VGG16

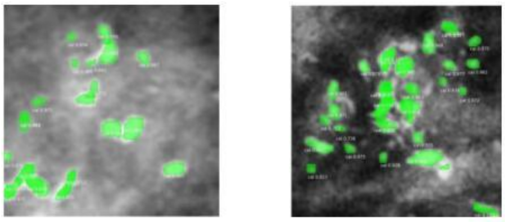
3.4 Mask R-CNN Figure Out All MCs From The Cluster

We use Mask R-CNN model to figure out all MCs from clusters. We will ask professional doctors to mark images classified as clustered MCs by VGG16 according to the contours of MCs which is shown in Fig. 3, and use the images labeled by the doctor as Ground Truth during Mask R-CNN training. Mask R-CNN contains two parts, namely before obtaining Proposal and after obtaining Proposal, and the processing before obtaining Proposal can be divided into Backbone and RPN (Region Proposal Network). Backbone is mainly composed of ResNet101 and FPN (Feature Pyramid Networks), which is used to extract and output feature maps of clustered MCs in mammography images. And we input the feature map and the mammography image into the RPN to detect the possible locations of clusters of MCs in the image. Then we input the feature map and the mammography image into the RPN to detect the possible locations of clusters of MCs in the image. Then, the locations where MCs may be detected and the feature maps extracted from the mammography images are adjusted in RoI (Region of Interesting) Align, and the adjusted results are input to the fully connected layer for classification and the positioning of the target. Next, Mask R-CNN will output an image marking the clusters of MCs in the mammography image, which is shown in Fig. 4. Finally, we will take out the clusters of MCs in the mammography image, and set the background to black to reduce the effect of the background complexity. The image with black background is shown in Fig. 5.



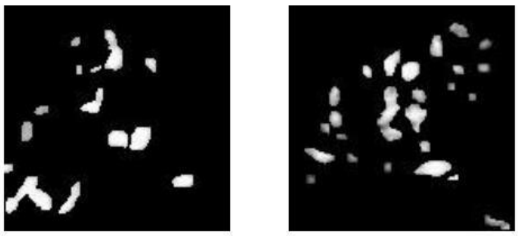
(a) (b)

Fig. 3. The manually labeled image of clustered MCs labeled by doctor : (a) is the image of benign cluster of MCs, (b) is the image of malignant cluster of MCs.



(a) (b)

Fig. 4. The image of Mask R-CNN labels all MCs from the images selected by VGG16 and classified as cluster MC: (a) is the image of benign cluster of MCs, (b) is the image of malignant cluster of MCs.



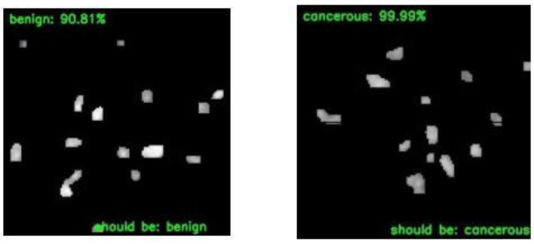
(a) (b)

Fig. 5. The image with black background: (a) is the image of benign cluster of MCs, (b) is the image of malignant cluster of MCs.

3.5 Inception V3 Judge Benign or Malignant Clusters

We use the Inception V3 model to classify the clusters MCs in mammography images labeled by Mask R-CNN, and the classification is based on the fact that the results of the needle localization surgical biopsy of each cluster are known to be benign or malignant clusters according to the content of the order of the doctor, and this is the Ground Truth during

Inception V3 training. Inception V3 is composed of three sizes of convolutional layers and one pooling layer, which is used to extract clusters and more features in the breast MC cluster image. Finally, we use fully connected layer to classify the clusters of MCs in mammography images as benign or malignant, and compare them with the order of doctor to confirm whether the model training achieves the goal of this paper . Fig. 6. shows the result image of benign and malignant analysis through Inception V3.



(a) (b)

Fig. 6. The result image of benign and malignant analysis : (a) is the image of benign cluster of MCs, (b) is the image of malignant cluster of MCs.

3.6 Result

The accuracy, precision, specificity, and sensitivity calculated by the confusion matrix during the testing in this paper are used as the evaluation indicators for the three models in this system, Table 1. shows the confusion matrix of VGG16 taxonomic cluster calcification. We can see from Table 1. that VGG16 has an accuracy rate of 93.3%, and Table 2. shows the confusion matrix of Mask R-CNN figure out all MCs from the cluster. In addition, we can see from Table 2. that Mask R-CNN has an accuracy rate of 95.5%. Then, Table 3. shows the confusion matrix of Inception V3 judge benign or malignant clusters. Moreover, we can see from Table 3. that Inception V3 has an accuracy rate of 91.4%. Furthermore, the precision, specificity and sensitivity of our proposed methods are about 87%, 89% and 90%, respectively. This proves that our research is credible.

TABLE 1. CONFUSION MATRIX OF VGG16 TAXONOMIC CLUSTER CALCIFICATION

Confusion Matrix of Testing		**True Condition**	
		Positive	***Negative***
Predicted Outcome	Positive	154	11
	Negative	10	142

TABLE 2. CONFUSION MATRIX OF MASK R-CNN FIGURE OUT ALL MCS FROM THE CLUSTER

Confusion Matrix of Testing		**True Condition**	
		Positive	***Negative***
Predicted Outcome	Positive	162	8
	Negative	6	141

TABLE 3. CONFUSION MATRIX OF INCEPTION V3 JUDGE BENIGN OR MALIGNANT CLUSTERS

Confusion Matrix of Testing		**True Condition**	
		Positive	***Negative***
Predicted Outcome	Positive	190	9
	Negative	18	100

IV. CONCLUSION

In this paper, we use VGG16 to find out the area of MC clusters, and use Mask R-CNN to find out all MCs from clusters, and then use Inception V3 to analyze the benign and malignant of the MC cluster to achieve an automated, intelligent and highly reliable medical assistance system to assist doctors in clinical diagnosis. In the future, it is hoped that we can further judge the stage of breast cancer as pre-cancer, carcinoma in situ or cancer by the clusters of MCs judged to be malignant.

REFERENCES

[1] M. Cheng, W. Tong, J Lou, J Liang, and F. Pan et. al., "Value of Contrast-enhanced Ultrasound in the Diagnosis of Breast US-BI-RADS 3 and 4 Lesions with Calcifications," Clinical Radiology, 2020.

[2] A. Ben Lakhdar, M. Daudon, M. C. Mathieu, A. Kellum, and C. Balleyguier et. al., "Underlining the Complexity of the Structural and Chemical Characteristics of Ectopic Calcifications in Breast Tissues through FE-SEM and μFTIR Spectroscopy," Comptes Rendus Chimie, Vol. 19, pp. 1610-1624, 2016.

[3] H. Qassim, A. Verma, and D. Feinzimer, "Compressed Residual-VGG16 CNN Model for Big Data Places Image Recognition," 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), 2018.

[4] Y. Chen, X. Li, D Jin, and Y. Shen, "Classification of a Small-data-set Thyroid Nodules Based on Supplementary Feature Layer Improved VGG16," 2020 39th Chinese Control Conference (CCC), 2020.

[5] A. S. Paste, and S. Chickerur, "Analysis of Instance Segmentation using Mask RCNN,"Instrumentation and Control Technologies (ICICICT), 2019.

[6] R. Anantharaman, M. Velazquez, and Y. Lee, "Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases," 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2018.

[7] Y. Mednikov, S. Nehemia, B. Zheng, O. Benzaquen, and D. Lederman, "Transfer Representation Learning using Inception-V3 for the Detection of Masses in Mammography," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018.

[8] D. Zhao, D. Zhu, J. Lu, and Y. Luo et. al., "Synthetic Medical Images Using F&BGAN for Improved Lung Nodules Classification by MultiScale VGG16," symmetry, Vol. 10, 2018.

[9] Q. Hu, L. F. D. F. Souza, G. B. Holanda, S.S.A. Alves, and F. H. D. S. Silva, "An Effective Approach for CT Lung Segmentation Using Mask Region-Based Convolutional Neural Networks," Artificial Intelligence In Medicine, 2020.

[10] N. Dong, L. Zhao, C. H. Wu, and J. F. Chang, "Inception v3 based cervical cell classification combined with artificially extracted features," Applied Soft Computing Journal, Vol. 93, 2020.

2020 International Conference on Fuzzy Theory and Its Applications (iFUZZY2020)

Based on Mask R-CNN Tooth Position Marking And Periodontal Disease Identification

*Note: Sub-titles are not captured in Xplore and should not be used

1st Chiun-Li Chin
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
ernestli@csmu.edu.tw

2nd Ming-Chen Hsu
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
exo050282@gmail.com

3rd I-Miao Chen
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
loveyou890121@gmail.com

4th Pin-Yu Yeh
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
pinky888911@gmail.com

5th Ting-Ya Chang
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
amy222024@gmail.com

6th Ya-Chu Hsieh
Medical Information
Chung Shan Medical Unversity
Taichung, Taiwan
as26872076@gmail.com

Abstract*—There are 90% population that suffer from periodontal disease in Taiwan. When identifying the diseased tooth, dentists will identify the number of the tooth with FDI world dental federation notation, and then use periodontal probe to assess whether the patient has periodontal disease and record it. However, it may take much time. In this paper, we propose a trained Mask R-CNN model to provide doctors more information from panoramic X-ray image when diagnosing. The accuracy rate of the model can reach 92% in the testing, which can help dentists label tooth position and identify periodontal disease.

***Keywords*—periodontal disease, FDI world dental federation notation, Mask R-CNN, panoramic X-ray image, mark tooth position**

I. Introduction

There are 20-50% population suffer from periodontal disease in the worldwide [1-2], and up to the 90% of Taiwanese people [3] suffer from periodontal disease. At the beginning of periodontal disease, it has the symptom of having gingivitis which will lead to alveolar bone and other periodontal tissues in the gums be damaged. While performing a diagnose of periodontal disease, dentists will interpret the tooth situation due to the panoramic X-ray image of patient. Dentists record diagnosis result with the tooth serial number of FDI world dental federation notation to indicate the tooth position and apply for health insurance based on the result. However, the diagnosis result from the doctor will be affected by the different growing conditions of each person's teeth and manual diagnosis may lead to misidentification due to factors such as fatigue, personal emotions or inexperience. In addition, dentists will use periodontal probe to verify the depth of periodontal pocket and bleeding gums when detecting periodontal disease in order to understand the loss of dentary around all teeth. But, this detection method often causes pain and takes longer time.

Since the dentist takes a lot of time on image diagnosis when visiting, in order to lessen the burden of dentists and reduce the procedure of diagnosis, we use Mask R-CNN to mark tooth position and the three areas for periodontal disease identification in this paper. To build a deep learning model that can be used to label tooth position and the three areas for periodontal disease identification, we trained Mask R-CNN to learn lots of labeled images. Applying it to dental diagnosis, dentists do not need to waste much time and energy on diagnosing images to realize the oral symptoms of the patient, which can also help dentists with no experience to diagnose, reduce time and the chance of misdiagnosis.

The rest of this paper is organized of the following section. Section two is going to review the traditional approaches of deep learning on labeling tooth position and periodontal disease. Section three is going to describe the algorithms of automatic tooth position labeling and identifying periodontal disease which are proposed in this paper and the experimental results. Finally, it is going to describe conclusions and recommendations for the future work of this paper.

II. Related Works

In oral diseases, periodontal disease is one of the common disease of many people. In traditional diagnosis, dentists used the way of artificial tooth marking, but it consumed too much time. With the advance of technology, images had been used to identify tooth position and periodontal in many papers. Thus, we refer to many papers and proposes the deep learning method to identify tooth position and periodontal disease which can help dentists diagnose quickly.

Lin et al. proposed one method to automatically measure the degree of alveolar bone loss [5]. By calculating ratio of the distance between the enamel and the crown position and the distance between the enamel and the root, it helped the dentist to quickly know the severity of periodontal disease. However, when the image is blurred, it will cause inaccurate position of enamel, crown and root.

In recent years, deep learning methods have been gradually applied to tooth position marking. Zhang et al. proposed an improved CNN deep learning network architecture [6] which was mainly used for detecting tooth position automatically in periapical image. In addition, it obtained quite good results in a small number of datasets to prove that using deep learning methods to detect tooth X-ray images is extremely feasible.

Lee et al. developed a deep network architecture based on CNN [7] which can analyze the size of clinical attachment loss (CAL) in the periapical image, as a reference for diagnosis and the prediction of damaged periodontal teeth. However, this architecture were only learning about the information of periapical images in this paper, but not every patient took periapical images. Therefore, this method was more inconvenient in practice than using panoramic X-ray images for analysis. Furthermore, Krois et al. developed a network architecture based on CNN [8] and applied it to panoramic X-ray images to detect the degree of periodontal bone loss and quickly know the current degree of periodontal disease. However, the image must be cropped in this paper, which reduces the accuracy and also limits the development of other diseases in the future.

Chen et al. proposed to use the Faster R-CNN architecture for tooth detection and tooth numbering in tooth apical X-ray(periapical) image. By using the algorithm which was developed by them to filter the overlapping frame of the same tooth being repeatedly framed, and build a network model for detecting missing teeth to improve accuracy [9]. However, it treated two adjacent teeth that were not sound as the same tooth while utilizing this method.

Based on the above related works, we found that no scholars had used Mask R-CNN in periodontal labeling, also there had no scholars simultaneously recognized periodontal disease and teeth position. Therefore, we label the tooth position on the panoramic X-ray image to assist the doctors to diagnose in this paper.

III. Methods and Results

In order to implement the automatic tooth position identification and diagnosis of periodontal disease, we use Mask R-CNN to label tooth for getting tooth position and label the three areas in panoramic X-ray image for periodontal disease identification. Our system can be divided into five parts, including inputting panoramic X-ray image, resizing and sharpening image, labeling panoramic X-ray image, teeth labeling model and periodontal disease identification, through the above five parts to achieve the final purpose of this paper. The flowchart of our proposed system is shown in Fig. 1.

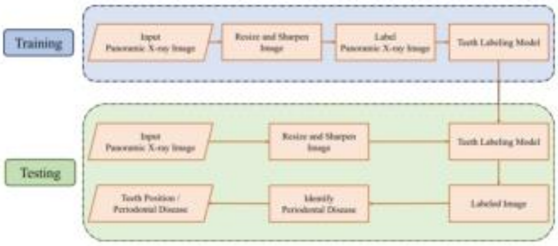


Fig. 1. Flowchart for research approach presented in the paper

3.1 Input Panoramic X-ray Image

This paper is supported by the dentists of the affiliated hospital of Chung-Shan Medical University, and the doctors of the department of stomatology provided panoramic X-ray images and professional knowledges. The participants of this paper are aimed at adults, and each panoramic X-ray image must contain 32 teeth which does not contain deciduous teeth, missing teeth and contains X-ray images of periodontal disease. We use this dataset for model training and testing in this paper. The 80% of the data is used for model training, and the 20% of the data is used to evaluate the accuracy of the model.

3.2 Resize and Sharpen Image

Before analyzing the image, the collected panoramic X-ray images will be resized to the image with a width of 1024. After that, in order to reduce taking a distinct brightness image on different devices which lead to poor contrast, inconspicuous tooth contour and poor recognition. Hence, we will sharpen the image in order to make tooth edge and contour more clearly visible in this paper.

3.3 Label Panoramic X-ray image

In order to mark tooth position and identify periodontal disease, we let professional dentists to assist in marking the original images in this paper. In the part of tooth position marking, we divide the panoramic dental X-ray image into four quadrants with FDI world dental federation notation. Marking the position of the teeth with two digits. The first digit represents the quadrant where it is located, the second digit represents the location of the teeth in this quadrant, and giving numbers from one to eight from the inside out. In the part of periodontal disease identification, the panoramic X-ray image is divided into the root and gingival area of the upper teeth, the area between the cemento-enamel junction of the upper and lower teeth and the root and gum area of the lower teeth. We will treat these marked images as the target images when training-Mask R-CNN.

3.4 Teeth Labeling Model

After resizing and sharpening the image, we utilize Mask R-CNN to automatically mark the tooth position and the three areas for periodontal disease identification in this paper. First, we use the images marked by dentists as a training dataset, and then input the images into Mask R-CNN. Moreover, Mask R-CNN uses Feature Pyramid Network to extract the features in the tooth image, and uses Region Proposal Network to obtain the region of interest, then classifies and marks the region by Fully Connected Layer and Fully Convolutional Network.

3.5 Identify Periodontal Disease

After Mask R-CNN has marked the three areas for identifying periodontal disease, we will automatically identify the periodontal disease on the marked images in this paper. Since periodontal disease can lead to cement loss, we determine whether there is a gap between each area to know whether there is cement loss, and calculates the size of the gap to further know the loss of cement and identify the severity of periodontal disease in this paper.

3.6 Result

Fig. 2. shows the result image that marking the three areas which are the regional of marking the periodontal disease area. It can be found from Fig. 2. (a) and (b) that Mask R-CNN can mark accurately whether it is an image of periodontal disease. However, it can be found in Fig. 2. (c) and (d) when the wisdom tooth on both sides of upper and lower teeth in the image do not grow out of the gums, the marking effect is poor.

It is guessed that when the wisdom tooth grows out of the gums, the position of the cementoenamel junction and other teeth are too different. As a result, it occurs an error when marking with Mask R-CNN. Furthermore, it may also occur an error when the image is not clear.

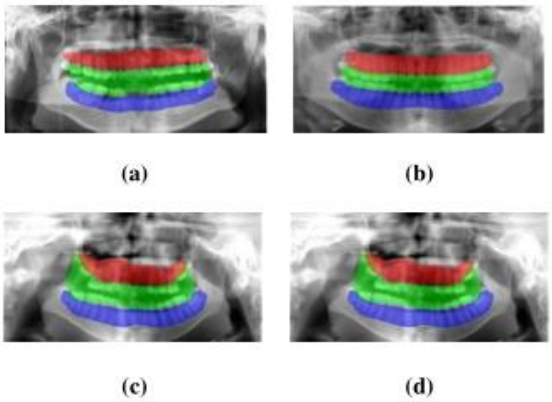


Fig. 2. Marking results of the areas for periodontal disease identification: (a) the image of patient with periodontal disease, (b) the image of healthy patient, (c) the image of upper and lower rows of wisdom teeth that have not grown out of the gums and (d) the blurred image

In order to evaluate the accuracy of the model, we use accuracy, precision, and recall calculated from the confusion matrix as evaluation indicators to identify the effectiveness of model training. It can be shown in Table 1, the accuracy of Mask R-CNN can reach 99.4%, the sensitivity is 98.9% and the specificity is 84.4%.

Table 1. Confusion matrix of the regional marks for periodontal disease identification

Confusion Matrix of Testing		True Condition	
		Positive	Negative
Predicted Outcome	Positive	46	4
	Negative	10	21

The testing results of tooth position marking in this paper are shown in Fig. 3. It can be found in Fig. 2. (a) and (b) that whether the teeth are arranged neatly or not, it does not affect the performance of Mask R-CNN on the tooth position mark.

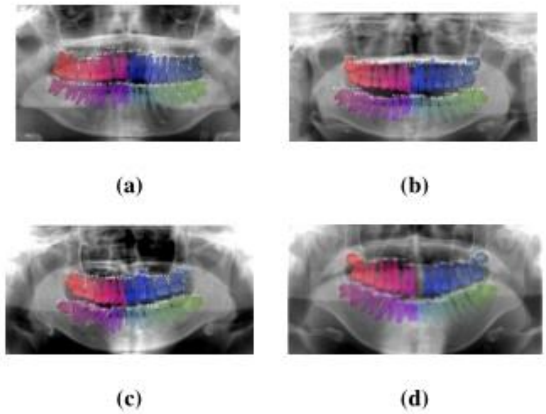


Fig. 3. The result images of tooth position marking: (a) the image of the patient with the teeth arranged irregular, (b) the image of the patient with the teeth arranged neatly, (c) the image of the blurred tooth edges and (d) the blurred image

The confusion matrix of tooth position marking is shown in Table 2. In Table 2, we know the accuracy of Mask R-CNN in tooth position marking can reach 95%, the sensitivity is 94.3% and the specificity is 87.2%.

Table 2. Confusion matrix of tooth position marking

Confusion Matrix of Testing		True Condition	
		Positive	Negative
Predicted Outcome	Positive	628	4
	Negative	7	21

IV. Conclusion

The numerous people suffer from periodontal disease around the world. During the diagnostic process of periodontal disease, dentist must spend much time on diagnosis. In order to assist dentist reducing diagnostic time, we use Mask R-CNN and the self-developed deep learning algorithm in this paper to build the dental diagnosis system that can mark tooth position automatically and diagnose periodontal disease. Compared the marked results with the diagnosis results of professional dentists, the accuracy rate can reach 95%, which can effectively assist dentists in clinical diagnosis. In the future, it is expected to add the detection of other diseases to increase the range of the application of this dental diagnosis system, providing more information for dentists. In the paper, we utilize Mask R-CNN combined with the self-developed deep learning algorithm to build a tooth position marking and periodontal disease diagnosis system, which effectively assists dentists in rapid and correct clinical diagnosis.

ACKNOWLEDGMENT

The authors would like to thank to Chung-Shan Medical University Hospital in Taichung, Taiwan for providing the panoramic X-ray images and Dr. Chang for marking the ground truth. The authors would also like to thank to the anonymous reviewers for providing with the suggestion of teeth loss measurement standard. The paper was partially supported by Computer of Science and Technology.

REFERENCES

[1] H. A. Nazir, "Prevalence of periodontal disease, its association with systemic disease and prevention," International Journal of Health Sciences, pp. 72-80, 2017.

[2] G. A. Nuno, L. P. Fátima, R. Ricardo, and F. R. João ,et. al., "Toll-like receptor 9 gene in Periodontal Disease – A promising biomarker," Gene, pp. 207-211, 2019.

[3] C. L. Chan, H. J. You, H. J. Lian and C. H. Huang, "Patients receiving comprehensive periodontal treatment have better clinical outcomes than patients receiving conventional periodontal treatment," Journal of the Formosan Medical Association, pp. 152-162, 2016.

[4] M. Y. Chung, M. K. Lee, J. Hong, and S. U. Park, et. al., "Pose-aware instance segmentation framework from cone beam CT images for tooth segmentation," Computers in Biology and Medicine, Vol. 120, 2020.

[5] P. L. Lin, P. Y. Huang, and P. W. Huang "Automatic methods for alveolar bone loss degree measurement in periodontitis periapical radiographs," Computer Methods and Programs in Biomedicine, pp. 1-11, 2017.

[6] K. L. Zhang, J. Wu, H. Chen and P. J. Lyu, "An effective teeth recognition method using label tree with cascade network structure," Computerized Medical Imaging and Graphics, Vol. 68, pp. 61-70, 2018.

[7] J. H. Lee, D. H. Kim, S. N. Jeong and S. H. Choi, "Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm," Journal of Periodontal and Implant Science, pp. 114-123, 2018.

[8] J. Krois, T. Ekert, L. Meinhold and T. Golla, et. al., "Deep Learning for the Radiographic Detection of Periodontal Bone Loss," Scientific Report, 2019.

[9] H. Chen, K. L. Zhang, P. J. Lyu, and H. Li, et. al., "A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films," Science Report, Vol. 9, No. 1, pp. 38-40, 2019.

2020 International Automatic Control Conference (CACS 2020)

The analysis of high-risk group of angina recurrence using smart health-box and deep learning algorithm

Chiun-Li Chin, Ting-Ya Chang, Pin-Yu Yeh, Ya-Chu Hsieh, I-Miao Chen

Abstract*—The cause of recurrence or worsening of angina is closely related to the living environment and habits of the patient. In order to know whether patient is often exposed to the environment that can cause angina. We collect the data of the surrounding environment factors of the user through a smart health-box with many sensors which is proposed in this paper. Next, we convert collected data into an image and extract features from the image by CNN to identify whether the user is a high-risk group of angina. Finally, the accuracy of our method in the test stage can reach 94%, and the sensitivity and specificity of our proposed method are about 95% and 89%, which can effectively remind patients whether the living environment is helpful for the recovery of angina.

***Keywords*—angina, smart health-box, CNN**

I. Introduction

If patients with angina are exposed to sudden changes in weather for a long time or always smoking or drinking can easily worsen the condition of angina. In order to know whether the environment that patients lived cause the symptom of angina. We collect the data of temperature, carbon monoxide concentration and alcohol concentration with various sensors in smart health-box in this paper, and convert the collected data into an image, and then extract the features by CNN to identify whether the patient is a high-risk group of angina recurrence, which can remind the patient whether the current living environment is beneficial to cure angina.

II. Related Works

Sohn et al.[1] mentioned that smoking, drinking and sudden weather changing are risk factors of angina. Moreover, Jiang et al.[2] proposed to input the shortened signal of the ECG to CNN architecture for classification. However, it had low accuracy. Pinzón-Arenas et al.[3] converted the detected hand EMG signal into a feature map and input it into CNN architecture to analyze gesture and it was feasible. Based on the literature above, we find that there are not any scholars converting the detected signal into an image and inputting it into CNN architecture to classify. Thus, we use CNN as a method to identify whether the patient is a high-risk group of angina recurrence.

Chiun-Li Chin is with the Chung-Shan Medical University, Taichung, Taiwan (e-mail: ernestli@csmu.ed.tw).

Ting-Ya Chang is with the Chung-Shan Medical University, Taichung, Taiwan (e-mail: amy222024 @gmail.com).

Pin-Yu Yeh is with the Chung-Shan Medical University, Taichung, Taiwan (phone: 0961575886; e-mail: pinky88911@gmail.com).

Ya-Chu Hsieh is with the Chung-Shan Medical University, Taichung, Taiwan (e-mail: as26872076@gmail.com).

I-Miao Chen is with the Chung-Shan Medical University, Taichung, Taiwan (e-mail: loveyou890121@gmail.com).

III. Methods and Results

This paper is supported by Chung Shan Medical University Hospital, and the participants of this research are people with no angina who live normally and people with angina who often expose to the environment that can cause angina. Fig. 1. shows the flowchart of our proposed system.

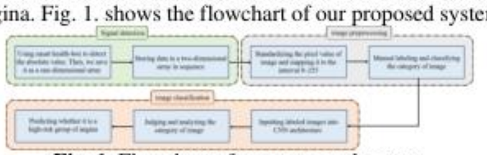


Fig. 1. Flowchart of our proposed system

The smart health-box that we proposed is composed of an alcohol sensor, a temperature sensor, a carbon monoxide sensor, and a Bluetooth module. Each sensor in the smart health-box can detect the values of environment factors. Then, the Bluetooth module will return the collected values to the system and our system will store them in a one dimensional array every 5 minutes. Next, we calculate the difference from the previous value and store six values in a two-dimensional array. Therefore, every collected data will produce a 6×288 image and the x-axis of the image represents the values of different sensors, and the y-axis represents the data detected at different times. We map them to the range of 0-255 through standardized formulas respectively, and then save them back into the 2D array. Finally, we label the category of the image in advance, and then input the image into a CNN model to train it learn how to identify whether it is a high-risk group of angina recurrence. During testing, the accuracy of angina classification is 94%, the sensitivity is 95%, and the specificity is 89% in this paper. According to the result above, it proves that the proposed method can identify the risk factors of angina around the environment precisely.

IV. Conclusion

In order to know the data of risk factors around the environment which may cause angina. We collect the data by smart health-box and sample signals of the sensors. Then, using a CNN architecture to achieve our goal. Finally, the accuracy of our proposed method can reach 94%. We expect to provide a more perfect system in the future.

References

[1] S. M. Sohn, B. G. Choi, S. Y. Choi, and J. K. Byun, et. al., "Impact of alcohol drinking on acetylcholine-induced coronary artery spasm in Korean populations," Atherosclerosis, pp. 163-169, 2018.

[2] J. Jiang, H. Zhang, D. Pi, and Chenglong Dai, "A novel multi-module neural network system for imbalanced heartbeats classification," Expert Systems with Applications : X, 2019.

[3] J. O. Pinzón-Arenas, R. Jiménez Moreno, and A. Rubiano, " Percentage estimation of muscular activity of the forearm by means of EMG signals based on the gesture recognized using CNN," Sensing and Bio-Sensing Research, 2020.