

天主教輔仁大學資訊管理學系碩士論文

指導教授：盧浩鈞 教授

深度學習應用於乳癌偵測

**Deep Learning Applied To Breast Cancer
Detection**



研究生：王建智撰

中華民國 109 年 7 月

論文題目：深度學習應用於乳癌偵測

校（院）系所組別：輔仁大學資訊管理學系碩士班

研 究 生：王建智 **指導教授：**盧浩鈞 論文頁數：38

關 鍵 詞：乳癌分類、卷積神經網路、深度學習、區域卷積神經網路

論文提要內容：

近幾年來，女性死亡癌症第一名一直是乳癌，透過乳房攝影檢測癌前病變，並經由醫師判斷是否得病、是否為高危險族群，有助於提早治療與提高治癒率，伴隨著醫療人力不足與經驗傳承的時間有限，具有醫學影像判讀能力的醫師正逐漸減少，近年來電腦視覺研究成熟，本研究期待著透過醫師與電腦的合作，能夠及早發現病人病灶給與早期治療。

本研究蒐集近 7000 筆的標記資料，希望透過 Faster RCNN 尋找病灶，尋求實際系統化的可行性。



Title of thesis: Deep Learning Applied To Breast Cancer Detection

Name of Institute: Department of Information Management , College of Management ,
Fu Jen Catholic University.

Name of Student: Wang,Chien-Chih Advisor:Lu,Hao-Chun Total Pages:38

Keywords: Breast cancer detection, Convolution Neural Network, Deep learning 、RCNN

Abstract:

In recent years, the first rank of the top ten causes of death among women has been breast cancer. The detection of precancerous lesions through mammography, and the determination by physicians of whether the disease is at risk or not, can help to improve the early treatment and cure rate.

With the shortage of medical manpower and the limited time of experience passing on, the number of doctors with the ability to interpret medical images is gradually decreasing. In recent years, computer vision research has become mature. This study expects that through the cooperation between doctors and computers, patients' lesions can be found as early as possible and given early treatment.

This study collected nearly 7,000 markers, hoping to find lesions through Faster RCNN and seek for the feasibility of actual systematization.

謝詞

終於完成論文了，首先感謝我的指導老師 盧浩鈞教授，在論文撰寫期間不斷雕琢我的研究架構以及給予我許多研究建議與研究資源，至今日得以完成這篇論文。同時也感謝 台北醫學大學 羅爾維助理教授給予許多醫學影像方面的專業諮詢和 輔仁大學統計資訊學系 邵皓強副教授擔任口試委員期間給予論文上許多寶貴的建議，使本研究更加完善。

於這 2 年研究所過程中，感謝同實驗室的 家瑋、嘉豪，陪我一起在論文的研讀、以及程式執行的研究，祝福大家都能鵬程萬里。感謝一路上給予過我任何協助的系上秘書與師長們，也謝謝我的家人一直支持我。

碩士二年級的下半年，由於疫情的爆發，使得求學過程倍感艱辛，同時也加深了我對於「人體是很脆弱的」認知，人類即使號稱是萬物之靈，面對疾病與癌症也應謙卑、戒慎恐懼，希望，我們能以知識的力量戰勝疾病與癌症。



王建智 謹致

中華民國一零九年七月

目錄

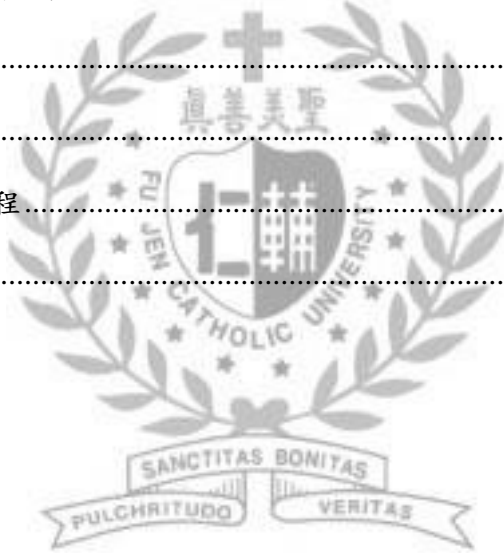
表次	v
圖次	vi
第壹章 緒論	1
第一節 研究背景與動機	1
第二節 研究目的	4
第三節 研究流程	5
第貳章 文獻探討	7
第一節 深度學習應用於乳房攝影相關文獻	7
第二節 區域卷積神經網路(Regions with CNN features)	11
第三節 以區域卷積神經網路為基礎的網路	12
第參章 研究方法	15
第一節 研究流程	15
第三節 乳癌影像資料	21
第四節 資料切割方法	22
第五節 模型架構	22
第六節 公開資料集 COCO 與 KITTI	25
第肆章 實驗結果	27
第一節 實驗流程	27
第二節 實驗數據	28
第伍章 結論	31
參考文獻	33
附錄甲 實驗混淆矩陣	37

表次

表 1- 1 乳癌 BI-RADS 等級分類(資料來源: Sickles et al., 2013).....	3
表 2- 1 現有 Faster RCNN 的研究比較(資料來源:本研究整理).....	9
表 3- 1 乳癌角度(CC View、MLO View,資料來源:本研究數據資料).....	18
表 3- 2 乳癌等級影像(資料來源:本研究數據資料).....	19
表 3- 3 乳癌等級資料分布(資料來源:本研究資料整理).....	21
表 3- 4 labellmg 標註等級數量(資料來源:本研究資料整理).....	21
表 3- 5 實驗 1 原始神經網路架構(資料來源: Lu et al. 2019).....	23
表 4- 1 資料前處理方式(資料來源: Lu et al.,2019).....	27
表 4- 2 資料原始數量與資料增強數量(資料來源:本研究整理).....	27
表 4- 3 Ensemble CNN 實驗結果(資料來源:本研究實驗資料).....	29
表 4- 4 CNN 實驗結果(資料來源:Lu et al,2019).....	29
表 4- 5 圈選比較(資料來源:本研究實驗整理).....	30
表 附錄- 1 Method a 混淆矩陣(資料來源:本研究實驗數據).....	37
表 附錄- 2 Method b 混淆矩陣(資料來源:本研究實驗數據).....	37
表 附錄- 3 Method c 混淆矩陣(資料來源:本研究實驗數據).....	38
表 附錄- 4 Method d 混淆矩陣(資料來源:本研究實驗數據).....	38

圖次

圖 1- 1 女性乳癌死亡率統計(資料來源:衛生福利部, 2019).....	1
圖 1- 2 研究流程.....	5
圖 2- 1 RCNN 示意圖(來源: Girshick et al., 2014)	11
圖 2- 2 Fast RCNN 示意圖(來源: Girshick,2015)	13
圖 2- 3 Faster RCNN 示意圖(資料來源: Ren et al., 2015).....	14
圖 3- 1 實驗 1 流程圖	16
圖 3- 2 實驗 2 實驗流程圖	17
圖 3- 3 資料分割法	22
圖 3- 4 資料分割說明	24
圖 3- 5 子模型訓練流程	25
圖 3- 6 實驗 2 示意圖	25



第壹章 緒論

第一節 研究背景與動機

衛生福利部(2019)發布 2018 年我國人民的死因統計結果，因為惡性腫瘤死亡的人數為 4 萬 8784 人，占有所有死亡人數的 28.2%，其中依照死亡率排名的癌症依序為氣管、支氣管和肺癌、肝和肝內膽管癌、結腸直腸和肛門癌、女性乳癌、口腔癌、前列腺(攝護腺)癌、胃癌、胰臟癌、食道癌、子宮頸及部位未明示子宮癌，此份排名與同機構於前年發布的統計結果相同，而乳癌也以每 10 萬人口 20.4%的死亡率連續蟬連女性死因第一名。在整理該單位近幾年發布的死因統計之後，我們可以發現，女性乳癌的死亡率趨勢如圖 1-1，呈現上升趨勢。

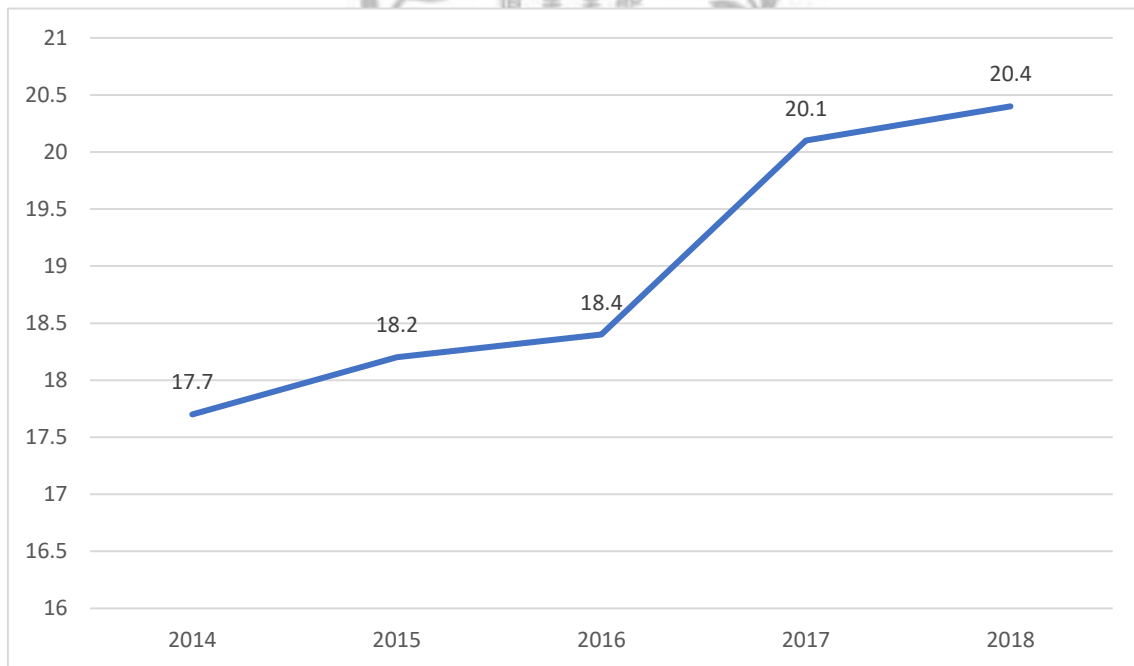


圖 1- 1 女性乳癌死亡率統計(資料來源:衛生福利部, 2019)

Nelson et al. (2016), Bleyer et al.(2016) 於各自的研究證明，乳房攝影(Mammography)可以幫助病患提早發現病症，並接受醫師的治療，進而使病患的存活率提升、死亡率下降。而乳房攝影也是除了觸診、乳房超音波外，病患可以選擇

的篩檢方式，觸診為醫師或者病患以手部接觸乳房，檢測乳房是否有不當之硬塊，此檢測方式所檢測的通常為已具相當規模的腫塊，也因為腫塊形成的時間已久，通常發現時，病狀已十分嚴重，所以通常醫師會建議以另外兩種方式檢查。一為乳房超音波，醫師透過超音波儀器掃描病患，並透過回傳內容判斷是否有異狀，但是腫塊與其他組織容易被認定是相同物質，因此也不容易及早發現；乳房攝影，相對於前者，雖檢查方式會讓病患稍微不適，不過可以檢查臨床檢查不到的微型鈣化點，之後再配合其他檢查，可以讓病患提早得知病況，並即早接受治療。

乳癌等級的判別可以依照美國放射協會(American College of Radiology, ACR)提出的乳房影像與資料分析系統(Breast Image Reporting and Data-analyzing System, BI-RADS)(Sickles et al., 2013)，此系統發布時定義了乳癌的各種級別與建議相對的醫療處理，大幅降低各地醫師對於乳癌病症的分類差異，也讓彼此溝通能夠更具效益，目前的 BI-RADS 為 2013 年公告的第五版本，其等級如表 1-1 陳述：



表 1- 1 乳癌 BI-RADS 等級分類(資料來源: Sickles et al., 2013)

Numeric Code	說明
0	此等級說明 Mammogram 資訊不夠完整(Incomplete)，不足讓醫師判讀是否為良性或惡性。推薦病患應重新攝影或以其他方式，如乳房超音波等，來重新給醫師判讀。
1	正常(Benign)，乳房影像沒有檢視到有腫塊、鈣化或是其他的異常。推薦病人依照醫師指示定期健康檢查。
2	正常(Benign)，乳房影像有鈣化的纖維瘤、淋巴結、人造植入物(如隆乳植入物)等經醫師確認與乳癌無相關內容。推薦病人依照醫師指示定期健康檢查。
3	可能為良性(Probably Benign)，乳房影像具有未確定之鈣化、局部不對稱的組織、群聚的鈣化點。推薦病人應追蹤檢查，增加檢查頻率(6 個月)。此級別若經醫師確認可疑的鈣化與相關組織對人體無害，可判讀為良性(2%可能被診斷為惡性)，若惡化則可能被歸納於後面三種等級。
4	懷疑為惡性(Suspicious)，此等級可再細分為:4A 懷疑(2% to 10%可能被診斷為惡性)、4B 懷疑(10% to 50%可能被診斷為惡性)、4C 懷疑(50% to 95%可能被診斷為惡性)，原則上會推薦加入其他篩檢方式，如:乳房超音波、採集活人檢體檢驗，若有疑慮可以採取切除。
5	高度懷疑(Hightly Suspicious)，此等級有 95%以上會被診斷為惡性，推薦採集活人檢體檢驗，可接受切除手術。
6	確診為惡性腫瘤，推薦病患儘速依照醫師指示進行手術或者化學藥物治療。

自從電腦發明之後，電腦輔助診斷(Computer Aid Detection/Diagnosis,CAD)，即被醫學影像領域的研究者視為輔助診斷病患的良好工具，隨著硬體技術的提升與相關演算法的研究，使得電腦判讀不再淪為幻想，從 2012 年開始，ImageNet 競賽在 Krizhevsky, A. et al.(2012) 提出的卷積神經網路 (Convolution Neural Network,CNN)結合顯示卡的運算大幅降低了判讀的錯誤率，開始了電腦影像判讀研究的加速時代，近幾年各式類神經提出百花齊放，現在我國醫學界人才嚴重不足，若能讓深度學習模型協助年輕放射師對乳癌影像做出精確判讀，不僅可以降低年輕放射師的壓力，也可以使病人獲得黃金治療期，進而降低醫療費用支出與家庭照顧的潛在壓力。

因此，本研究期望以神經網路模型判斷乳癌影像的 BI-RADS 等級，並且透過區域卷積神經網路(Regions with CNN features ,RCNN)圈選病灶，以利醫師對於 BI-RADS 代碼為 3、4、5 的病患進行精準治療。

第二節 研究目的

根據本章第一節，本研究將使用台灣北部地區某間醫院提供的乳癌影像進行以下研究：

1. 嘗試將訓練良好的 CNN 模型配合相關的分散式學習來建立醫學影像的分類系統，透過各項醫學、統計指標來評估系統的好壞。
2. 導入預先訓練好的 Faster RCNN 訓練醫師圈選的資料，並嘗試以訓練好的 Faster RCNN 模型圈選病灶，並評估目前取得的研究資料資料的對於電腦判讀的限制、與醫師人力圈選的差異。

第三節 研究流程

本研究流程如下，首先對本研究題目做探討並說明，參考相關深度學習預測乳癌之文獻並整理。建立神經網路模型判斷乳癌影像的 BI-RADS 等級、，並透過 Kappa 指標、混淆矩陣(Confusion matrix)，評估模型優劣，另外評估 Faster RCNN 圈選的病灶位置與專業放射醫師圈選的差異，最後提出結論。如圖 1-2 所示：

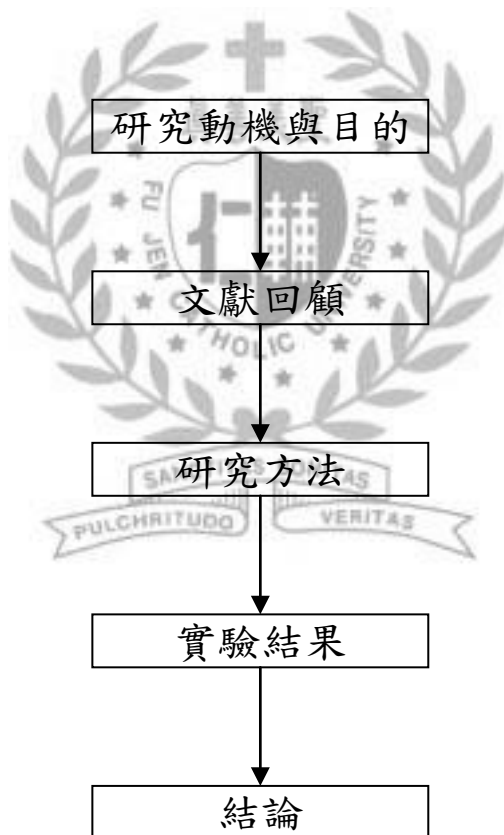


圖 1-2 研究流程



第貳章 文獻探討

第一節 深度學習應用於乳房攝影相關文獻

目前為止用於影像分析最成功的模型方法為卷積神經網路(Litjens et al., 2017)，隨著判讀越來越準確，醫師開始注重另一個面向。也就是，能否讓電腦幫忙圈選病灶。以利後續針對病灶進行精準治療，本節將探討適用於乳癌影像分析的神經網路架構。

Akselrod-Ballin, A. et al.(2016)分析數個大型醫學中心的資料，解析度~4K*3K 影像張數 850 張(包含 BI-RAIDS 2-5)，資料事先分割多餘的胸骨等影響電腦判讀的部分，接著使用 Fuzzy Logistic Approach 偵測出乳房纖維腺(FibroGlandular)，由於硬體限制，Akselrod-Ballin, A. et al.使用重疊(Overlap)的方式建立 800*800 像素的圖片作為資料增強(Data Augmentation)，使用修正的 Faster RCNN 進行特徵萃取與判斷，透過 Region Proposal Network(RPN)將圖片畫出疑似病灶作為感興趣區域(Region Of Interest,ROI)，並用 VGG-16 作為判讀的 CNN 架構。其 RPN 部分的參數使用 ImageNet 訓練好的模型，實驗評估模型的指標分為 ROC 曲線下面積之 AUC 與準確率(Accuracy);實驗分為兩個模型，一為良性(BI-RADS 2)與惡性(BI-RADS 3-5)的資料集;另一模型為良性(BI-RADS 2)與惡性(BI-RADS 4-5)的資料集。前者模型 AUC 為 0.6，Accuracy 為 0.78;後者 AUC 為 0.72，Accuracy 為 0.77，並於實驗後使用公開資料集 DDSM-BCRP、INBreast 資料集進行影像分類;並確認較他人先前研究的分類模型具有更高的 AUC ROC 與 True Positive Rate (TPR)。

Akselrod-Ballin, A. et al.(2017) 透過輸入腫塊的 Bounding Box(BB)座標、大小、圖片作為訓練資料，使用 3 階層 Faster RCNN 分別預測腫塊(mass)、判別腫塊的良性(Benign)與惡性(Malignant)，資料處理仍然使用 OVERLAP 增加資料量、並帶入 ImageNet 預先訓練好的模型作為 RPN，並使用修改過的 VGG 作為 CNN 架構，最

後將三個作用的 CNN 架構串聯，作為 Bounding Box 的產出與預測結果，同時為了符合修改的 VGG 架構，圖層處理成 224×224 像素，使用的衡量指標為 TPR、False Positives Per Image(FPI)，並分別陳述分類預測與腫塊預測位置的相關指標;此研究資料集為 INBreast 與大型醫院資料 3500 張，並將預測結果{2,3,4,5}依照 BI-RADS 分為良性與惡性，測試結果顯示於大型醫院資料檢測的腫塊 TPR0.93、0.56FPI，用於 INBreast 資料判別得到 TPR0.9@1FPI;另一方面檢測腫塊良、惡性的數值為 TPR0.96 \pm 0.03@1.2FPI、TPR0.87 \pm 0.14@0.8FPI;經實驗證實 FPI 較同作者於去年研究有顯著的降低，且有助於細小腫塊的判定。

Ribli, D. et al.(2018)嘗試將 Faster RCNN 導入電腦輔助偵測系統，主要架構仍然是以 ImageNet 預訓練的 RPN 結合 VGG16 並將分辨分為良性、惡性，訓練使用資料為 DDSM 數據集 2620 張與 Semmelweis University in Budapest 提供的 847 張 FFDM 數據資料，在測試方面使用 INBreast 資料集進行測試，Dezso et al.將所有資料標準化成 0-255;使用的衡量指標為 ROC 曲線下面積之 AUC，於測試實驗中得到的 AUC 為 0.95，可分辨 INBreast 中 90%資料的惡性組織，並達到每張 0.3 的假陽性結果(FPI);於 Digital Mammography DREAM Challenge 亦得到 0.85 的 AUC，結論為此分類器略優於當時的 CAD 系統，並且具有圈選病灶的能力。

Nath, A. M. et al.(2019)以 Faster RCNN、YOLOv3 與 Retinanet 進行 Mammogram 判讀，有別於 YOLOv2 的 Darknet-19 架構，YOLOv3 改變架構，成為一個 106 層的完全卷積架構，可以捕獲細小顆粒; Retinanet 以金字塔網路(Pyramid Network)與焦點損失作為(Focus loss)基礎，可以捕獲大小不一的腫塊;主要實驗不以辨識良性惡性，而是以精準抓取多少個腫塊組織為目的，使用的資料集有 DDSM 數據集、INBreast 資料集、Indian AIIMS 資料集，使用 Transfer Learning from Chest 論文提出的模型作為預訓練模型;衡量指標以 AUC ROC 與 FROC 作為衡量指標，INBreast

於 Faster RCNN、Retinanet 皆達到 0.94 AUC ROC, YOLO 雖 AUC 不像前兩者數值高，但於細小腫塊判讀數量優於前兩者。

表 2- 1 現有 Faster RCNN 的研究比較(資料來源:本研究整理)

作者	圖片區域	影像前處理	分類	評估方法	遷移學習
Akselrod-Ballin, A. et al. (2016)	整張	fuzzy logistic approach 資料增強	正常、良性、惡性	Accuracy、AUC ROC	有
Akselrod-Ballin, A. et al. (2017)	整張	資料增強	良性、惡性	TPR、FPI	有
Ribli, D. et al. (2018)	整張	資料增強	正常、良性、惡性	AUC ROC FPI, FROC	有
Nath, A. M. et al. (2019)	整張	資料增強	腫塊數量	ROC AUC FROC	有



第二節 區域卷積神經網路(Regions with CNN features)

深度學習在 Hinton 率領的團隊於 2012 年 ImageNet 大賽提出卷積神經網路之後蓬勃發展，後續的研究如 Google 的 VGG、Microsoft 的 ResNet 等分支，由於前述領域研究成熟與進步，區域性的影像辨識也開始受人重視，區域卷積神經網路 (Regions with CNN features, RCNN)，是 Girshick et al.(2014)提出的。主要的概念是模型讀入影像之後，選取待選區域(Region Proposal)，換句話說就是從圖片中框處各種範圍，包含欲判斷物件的領域，接著將這些大小不一的待選區域統一調整成一定大小的影像(以發表論文為例，就是將待選區域調整成 227×227 的解析大小)，透過 CNN 從這些調整過的待選區域提取特徵，最後再導入分類層如 SVM 模型進行分類。其過程如圖 2-1:

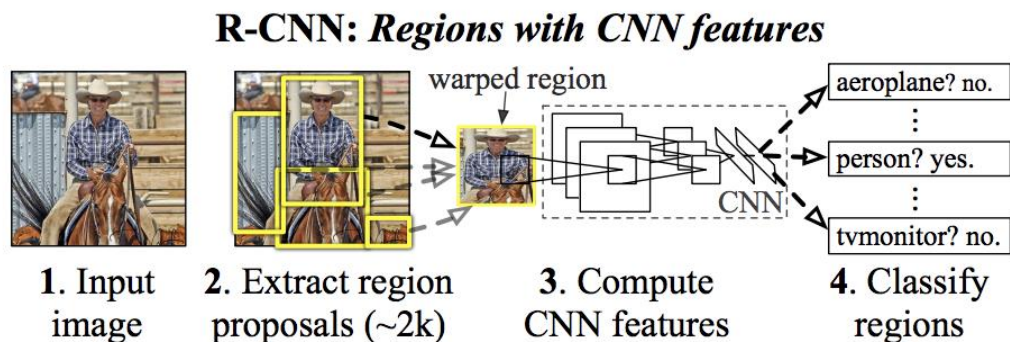


圖 2- 1 RCNN 示意圖(來源: Girshick et al., 2014)

1. 待選區域(Region Proposal):

Girshick et al.(2014)在 RCNN 的研究與 Girshick.(2015)改良的 Fast RCNN 皆使用 Selective Search 對原始影像進行分割，每次放入都需設定待選區域數量，待選定多個起始點，透過計算起始點周圍顏色、紋理相似性，將相似性高的納入區域加以合併，最後生成多個待選區域。

2. 特徵萃取(Feature Extraction)

Girshick et al. 先將提取的待選區域轉換成 227×227 pixel 的大小，接著將待選區域放入 CNN 進行特徵萃取。

3. 分類判別(Classified)

當 CNN 將特徵萃取出來後，將這些特徵放入 SVM 模型產生預測結果與區域偏移值。

4. 由於影像在選取待選區域時可能切割到物體，所以我們必須透過 SVM 計算出來的圈選區域迴歸偏移值(Bounding Box Regression)對圈選區域進行修正，讓待選區域能夠圈選出物體的正確位置。

RCNN 的出現讓我們除了可以判讀物體，也同時可以鎖定物體具體的座標，於一張照片同時判定多個物體變得更容易。

RCNN 的缺點為：

圖片產生待選區域的方法是透過 Selective Search 進行選取，因為沒有任何學習的過程，所以每次選取都要耗費許多時間計算選取區域，且待選區域不一定適合做分類。

第三節 以區域卷積神經網路為基礎的網路

Girshick (2015) 提出感興趣區域(Region of Interest, ROI)，將 SVM 分類改以 CNN 進行判斷，以 softmax 作為物件類別預測、Bounding Box Regressor 作為圈選區域修正，Girshick 將方法命名為快速區域卷積神經網路(Fast Region-based Convolutional Network, Fast RCNN)

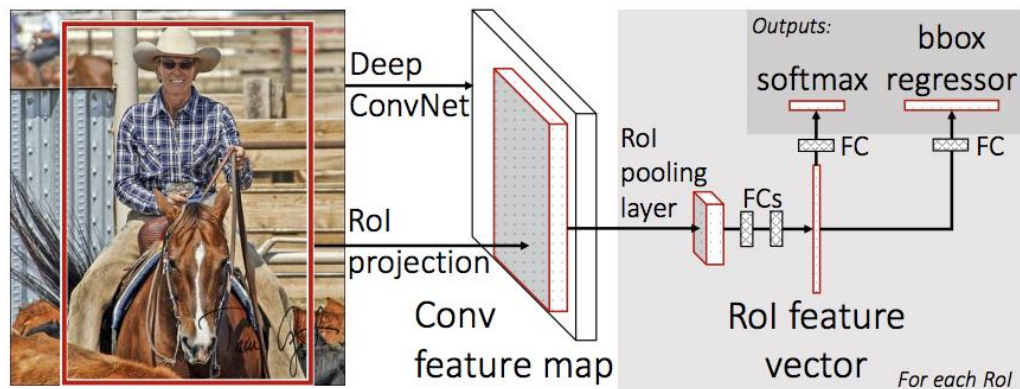


圖 2- 2 Fast RCNN 示意圖(來源: Girshick,2015)

改良部分:

1. 感興趣區域,作者提出 Region Of Interest Pooling 的架構,透過 Max Pooling 的方式將大小不一的候選區域縮減成特定大小的方格
2. RCNN 是透過 CNN 萃取特徵圖(Feature Map),再放入 SVM 產生預測並回傳迴歸偏移值, Fast RCNN 以神經網路運算可以同步計算迴歸偏移值、產生分類結果;由於不用分別進行兩種模型參數的調整訓練,因此訓練速度與判斷速度加快。

隔年 Ren et al.(2016)將尋找待選區域任務也合併到整個神經網路架構內,自此,從候選區域的選取、候選區域的特徵萃取、分類、選定區域終於合為一個類神經架構,此方法被稱為 Faster RCNN 。

待選區域網路(Region Proposal Network,RPN), Fast RCNN 模型的待選區域是透過 Selective Search 進行選取,其計算待選區域的演算過於耗時,隨著待選區域要求數量的增加,執行時間也會變長,因此,Girshick 提出 RPN,建立一個 ConvNet 接取前 ConvNet 產生出的 Feature Maps, RPN 會先由使用者定義 Anchor Boxes 與 Intersection-over-Union(IoU),定義不同大小的 Anchor Boxes 篩選相對應大小的物

件圈選區域，IOU 則是判斷 Anchor Boxes 內的圖像是物體或者是背景，電腦可以進行深度學習或者使用訓練好的模型進行遷移式學習，因此大幅降低了候選區域的選取與訓練時間。

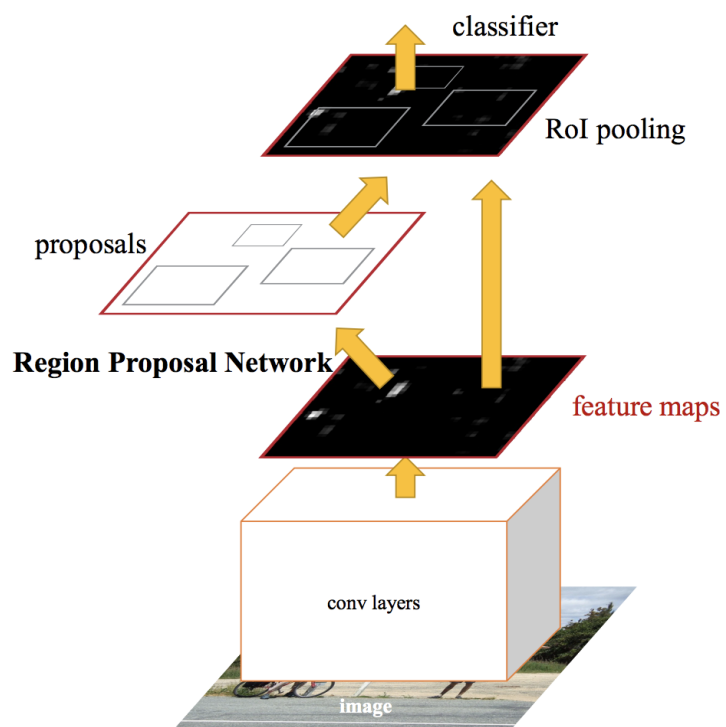


圖 2- 3 Faster RCNN 示意圖(資料來源: Ren et al., 2015)

第參章 研究方法

第一節 研究流程

本研究乳癌 X 光片使用台灣某醫院之放射資料，所取得之乳房攝影影像角度包含有 Bilateral Cranio-Caudal View(CC)，意即由上往下照；Bilateral Mediolateral-Oblique View(MLO)，意即由側邊照，左右乳房皆有，Superlateral to Inferomedial Oblique，照射方式為 MLO 之反向。上述之角度如表 3-1 所示。本實驗所取用乳癌攝影影像 BI-RADS 等級包含 1、2、3、4、5，如表 3-2 所示。乳癌影像原始大小為 2294*1914，從醫院所獲取之檔案為去識別化的個人資料。Lu et al. (2019) 在研究透過中值濾波器、維納濾波器、對比限制自適性直方圖等針對乳癌影像去除雜訊處理，於實驗中證實方法有助於影像判讀，本研究也將於實驗 1 進行影像處理。

本研究將分為兩個實驗：

實驗 1 會先對影像進行中值濾波器、維納濾波器、對比限制自適性直方圖等進行前處理，接著進行水平翻轉與垂直翻轉等資料擴增。就現有 CNN 模型進行超參數調整、改善模型架構，建立更加可信的模型架構。

實驗 2 嘗試導入現有基礎的 RCNN 架構進行預測，導入預先訓練好的 Faster RCNN 模型加以擬合乳癌實驗資料，讓模型分類出乳癌可能所屬的 BI-RADS 等級，流程如圖 3-1、圖 3-2 所示：



圖 3-1 實驗 1 流程圖

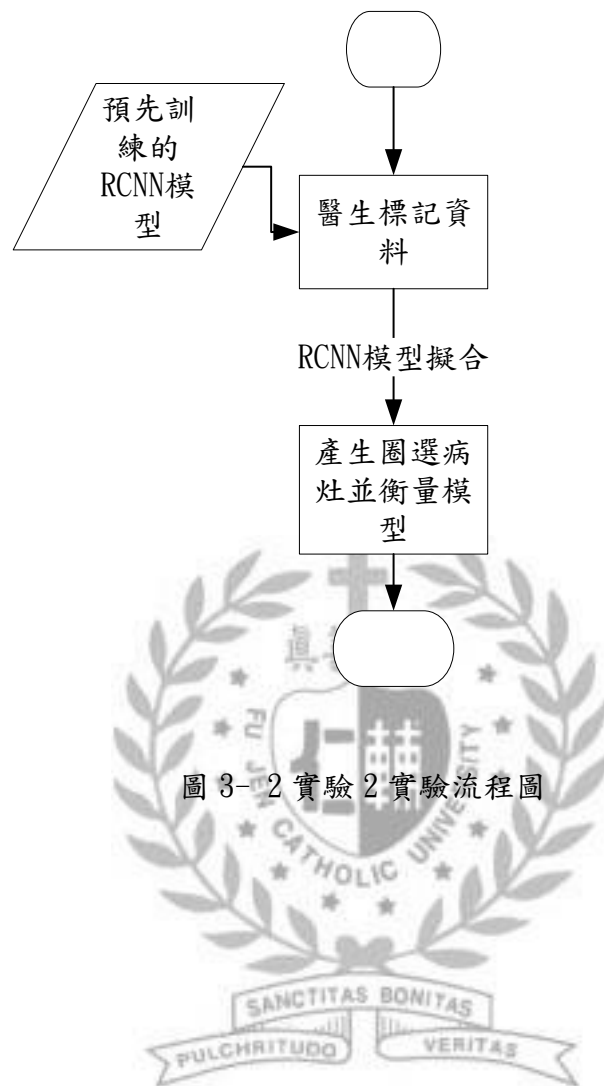


圖 3-2 實驗 2 實驗流程圖

表 3- 1 乳癌角度(CC View、MLO View,資料來源:本研究數據資料)

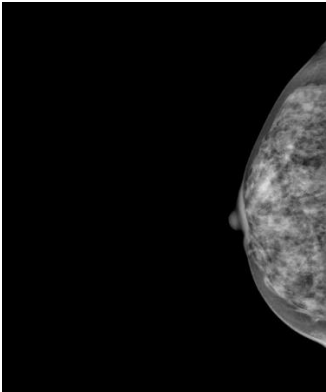
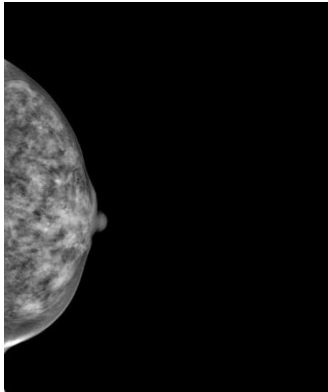
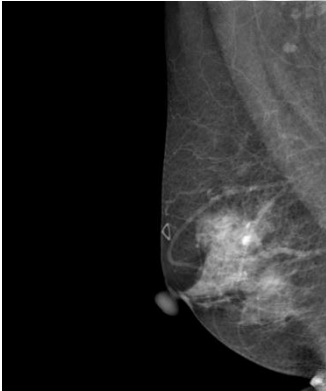

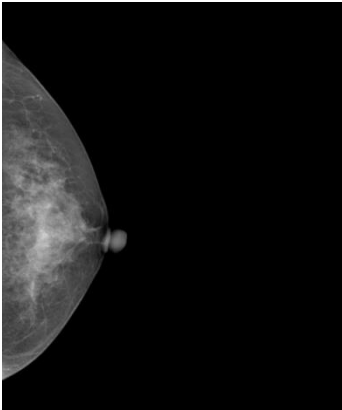
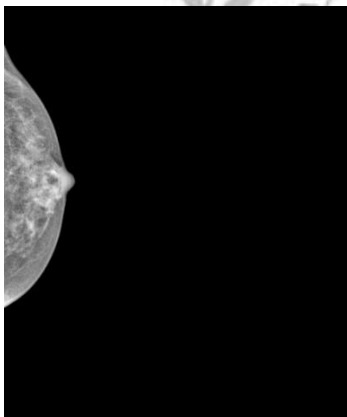
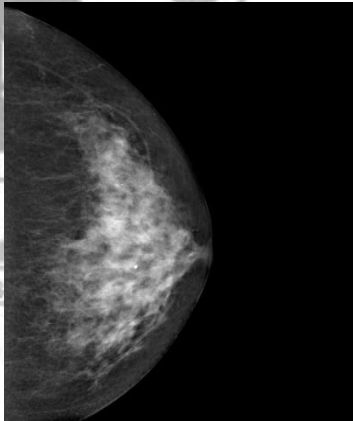
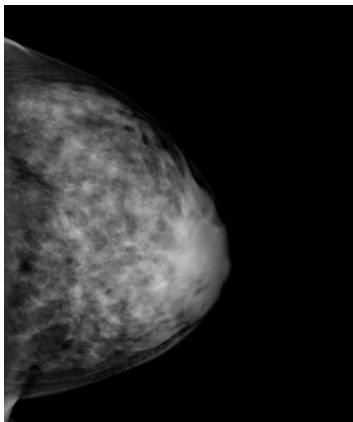
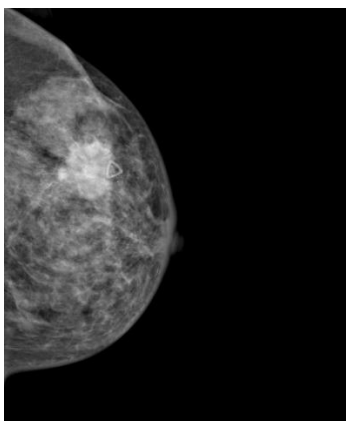
Right CC view	Left CC view
	
Right MLO view	Left MLO view
	

表 3-2 乳癌等級影像(資料來源:本研究數據資料)

BI-RADS1	
	
BI-RADS2	BI-RADS3
	
BI-RADS4	BI-RADS5
	

第二節 影像前處理方法

一、 中值濾波器(Median Filter):

中值濾波器又被稱作中值模糊(Median Blur)是種非線性的影像前處理技術，常使用於去除影像中的雜訊，在消除雜訊的過程中，中值濾波器能夠透過使用者對參數的設定保留影像邊緣避免模糊，讓處理過後的影像與原始影像可以儘量保留雜訊外的影像資訊。

二、 維納濾波(Wiener Filter):

維納(氏)濾波器是由 Norbert Wiener 提出的濾波器，使用統計分析技巧，對影像及雜訊之頻譜進行運算並抑制雜訊，讓輸出與期望輸出之平方差(MSE)達到最小值。

三、 對比限制自適性直方圖等化(Contrast Limited Adaptive Histogram Equalization, CLAHE):

CLAHE 由 Pizer et al. (1990)所提出，目的在於改善自適應直方圖均衡化(Adaptive histogram Equalization, AHE)方法所造成區域過度強化問題。當影像區域過度強化時，會使欲分析影像的影像特徵產生不必要的雜訊。而直方圖等化受到累積分布函數(Cumulative Distribution Function, CDF) 的影響，因此對比限制自適性直方圖等化在計算累積分布函數前就先裁剪直方圖來抑制雜訊，其切割下來的部分將會被平均重新分配，如果重新分配之後又超出限制高度則會被裁剪並再重新分配。

上述前處理皆為常見、且應用於電腦視覺的前處理方法，但資料視覺的研究也有一部分學者認為，透過前處理會導致影像中，人眼所無法辨識的資料特徵遭受破壞，反而使電腦無法做出更有效的判讀，因此在做前處理時，應適當衡量資料前處理的必要性。

第三節 乳癌影像資料

本研究共蒐集到 7325 筆乳癌影像，每張影像大小為 2294*1914。

於實驗 1，扣除部分過度曝光、醫師認定無法使用的影像，共使用 1770 位病患資料。詳細等級數量如表 3-3 所示：

表 3-3 乳癌等級資料分布(資料來源:本研究資料整理)

等級	人數	張數
1	598	2392
2	599	2397
3	478	1914
4	70	518
5	25	104

實驗 2，會請專業醫師針對蒐集到的乳癌圖像使用 labelImg 標註病灶，用於製作 TFRecord，此次實驗共計標註 2907 張圖片、6519 個標註點，其中 BI-RADS 0 此次不用於實驗，詳細標註乳癌等級數量如表 3-4 所示：

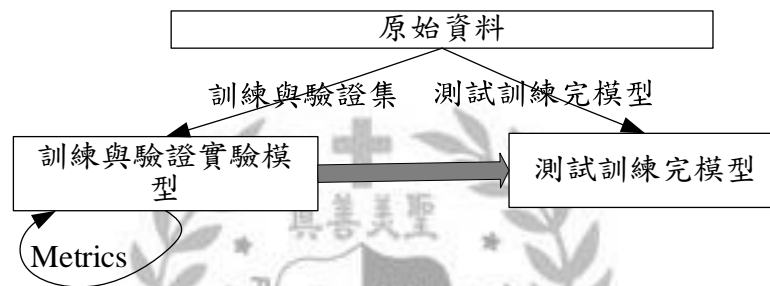
表 3-4 labelImg 標註等級數量(資料來源:本研究資料整理)

等級	標註點數量
BI-RADS 0	371
BI-RADS 2	4081
BI-RADS 3	1837
BI-RADS 4	169
BI-RADS 5	61

第四節 資料切割方法

本研究的實驗會依照是否有做資料擴增將分割方法分為兩種，若資料集沒有做資料擴增，則將資料集依照指定比例分割;若資料經過資料擴增，則將原始資料與擴增的資料做混合，再依比例分割(見圖 3-3)，分割期間皆兼顧各類別資料比例。

分割法 1: 資料不含資料擴增



分割法 2: 資料包含擴增後資料

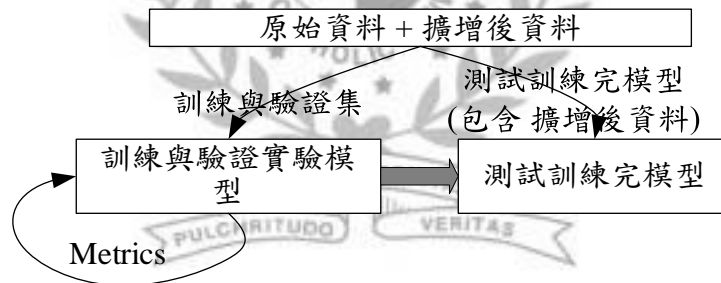


圖 3-3 資料分割法

第五節 模型架構

本研究將進行兩種方向的實驗，實驗 1 為改良 CNN 模型研究，實驗 2 為應用 FASTER RCNN 於醫學影像研究。

實驗 1 修改 Lu et al. (2019) 提出的模型，原始神經網路網路包含 7 層卷積層，

3 層池化層，並使用批次正規化 (BN)取代舊有的 Dropout 層，激活函數使用 relu、並於最後連接層採用 softmax 作為激活函數用於生成各分類機率，在優化器使用 adadelata。其原始模型模型，容易出現區域最佳解模型，因此本研究將提出改良方式與 Lu et al.(2019) 的實驗數據做分析。

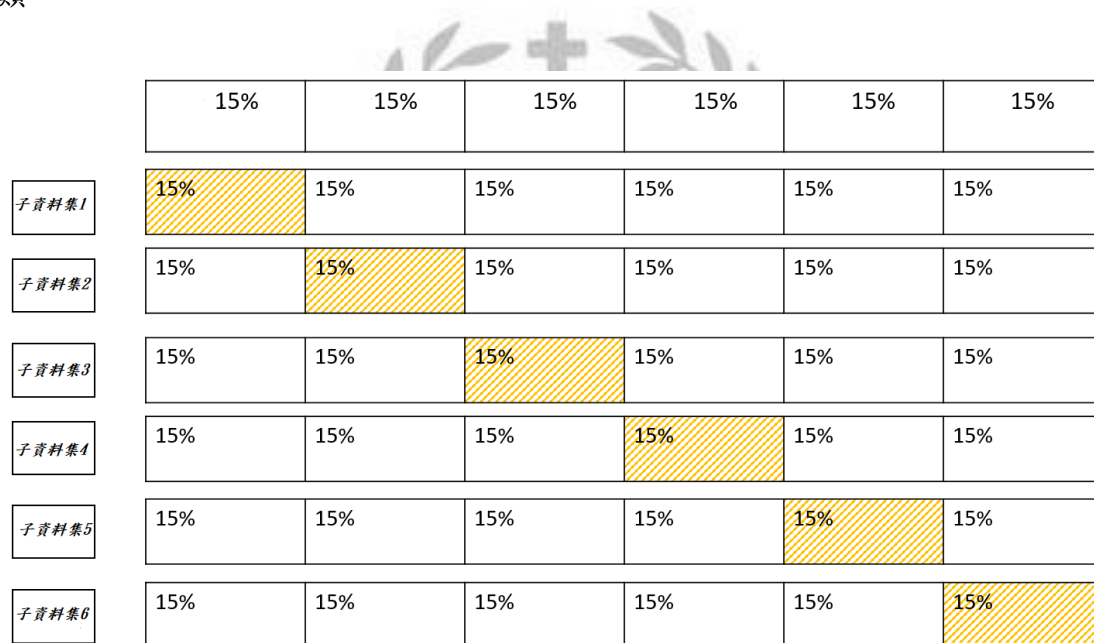
表 3-5 實驗 1 原始神經網路架構(資料來源: Lu et al. 2019)

Layer	Filter	Filter Size	Strides
Convolution (Batch Normalization)	32	5	2
Max Pooling		3	2
Convolution (Batch Normalization)	64	3	2
Convolution	64	3	1
Convolution	64	3	1
Max Pooling		2	2
Convolution (Batch Normalization)	96	3	1
Convolution	96	3	1
Convolution	96	3	1
Max Pooling		2	2
Global Average Pooling			
Fully Connected	1024(nodes)		

Fully Connected	512(nodes)
Fully Connected	5 (nodes)

Ensemble CNN 概念:

將原始資料集以 90:10 比例分割為訓練集與測試集後，將訓練集資料均分為六等分(資料分割參照圖 3-4)，並以不同的六分之五等分用來訓練原始神經網路架構，以各自的六分之一等分進行子模型的驗證，分別訓練產生六個子模型(訓練過程如圖 3-5 所示)，最後透過子模型預測的結果乘計算權重，對於乳癌影像進行投票分類。



	15%	15%	15%	15%	15%	15%
子資料集1	15%	15%	15%	15%	15%	15%
子資料集2	15%	15%	15%	15%	15%	15%
子資料集3	15%	15%	15%	15%	15%	15%
子資料集4	15%	15%	15%	15%	15%	15%
子資料集5	15%	15%	15%	15%	15%	15%
子資料集6	15%	15%	15%	15%	15%	15%

圖 3-4 資料分割說明

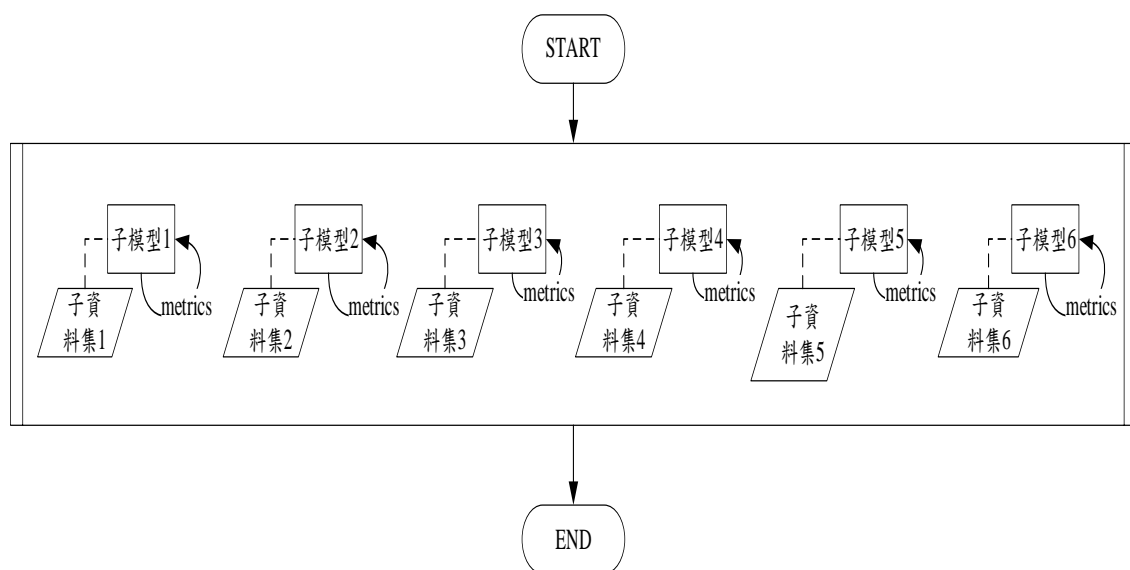


圖 3-5 子模型訓練流程

實驗 2 以 Google 的 Object Detection API 導入預先訓練的 Faster RCNN 擬合研究資料，輸出類別與相對應的圈選區域(Bounding Box,BBox)。

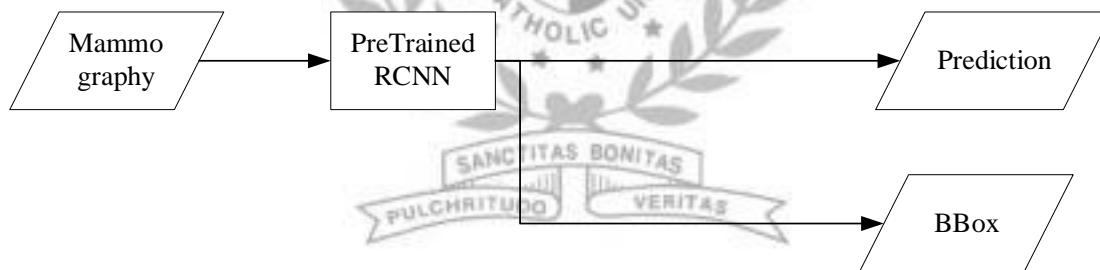


圖 3-6 實驗 2 示意圖

本研究使用的 Faster RCNN 模型，是基於 Resnet101 為底的 Faster RCNN 模型，使用的 CNN 架構採用 Resnet101 是因為希望透過層數較多的 Resnet101 萃取更多的特徵;預先訓練資料集為 KITTI 資料集，採用目的為，使用該模型重新訓練自製資料集效果優於使用公開資料集 COCO 預先訓練的 Faster RCNN。

第六節 公開資料集 COCO 與 KITTI

COCO 全名為 Common Objects in Context，COCO 資料集為資訊公司微軟所

公開的資料集，資料集項目包含物件分割(Object Segmentation)、上下文辨識(Recognition in Context)等，全資料集超過 33 萬張圖片，其中有 20 多萬張圖片經過標記，內容涵蓋多種物體與實例，並且逐年更新，是個學習電腦視覺的知名資料集

KITTI 則是由德國學府 Karlsruhe Institute of Technology 以及汽車公司 Toyota Technological Institute at Chicago 聯合使用的公開資料集，其用途原本適用於自動駕駛車影像判讀研究。



第肆章 實驗結果

第一節 實驗流程

實驗 1 將實作 4 個方法與 Lu et al.(2019)的研究作比較，資料的前處理以前述研究提出的實驗方法 Method1-4 作為資料分割前的影像前處理方式(詳見表 4-1)，相關參數設置如下優化器為 Adadelata、學習率為 1.5、epochs 為 100、batch size 為 8。因各方法在資料筆數上不盡相同，本研究選用 Kappa 與幾何平均數(G-mean)作為評估方法指標。Kappa 與 G-mean 皆可用在不平衡多類別方法時之評估指標(Alejo et al, 2008)，並搭配混淆矩陣所計算之指標作為評估方法優劣指標

表 4-1 資料前處理方式(資料來源: Lu et al.,2019)

Method (Num)	Data Augmentation	CLAHE	Median Filter	Wiener Filter
(1)	No	No	No	No
(2)	Yes	No	No	No
(3)	Yes	Yes	Yes	No
(4)	Yes	Yes	No	Yes

表 4-2 資料原始數量與資料增強數量(資料來源:本研究整理)

BI-RADS	Number	Data Augmentation
1	2392	2140
2	2397	2112
3	1914	1776
4	518	3239
5	104	822
Total	7325	10079

實驗 2，實驗開始前，將聘請專業醫師標註乳癌影像病灶，計算切割資料比例後，將資料集分別轉換出訓練集與測試集的 TFRecord，透過 Object Detection API 設定 config，本實驗將 num_step 設定為 35000，模型最佳化設定沿用 API config 原始參數，於訓練後輸入乳癌影像，並比較輸出結果與醫師圈選的差異。

第二節 實驗數據

實驗 1，為了方便與 Lu et al.(2019)的模型比較，本論文的實驗以 a-d 對照 Lu et al.(2019)的 method1-4，其實驗數據分別如表 4-3、表 4-4 所示，比較 method a、1，實驗結果發現，Ensemble CNN 於此實驗對照組並不如預期優異，推測原因為該資料集沒有 Data Augmentation，對於原本資料量只有 104 張的 BI-RADS 5，Ensemble CNN 將原始資料分到各子訓練集的資料為 78 張驗證資料為 16 張，測試集資料更是只有 10 張，BI-RADS 5 在訓練資料中相對稀少，使得此模型對於第五類的分類能力接近於 0，故而導致成效不彰。

比較 Method b、2 發現，對資料集進行 Data Augmentation 之後，因 BI-RADS 5 樣本增加，因此 CNN 開始對 BI-RADS 5 有分類效果，同時 Ensemble CNN 分類 KAPPA、G-mean 已開始優於單一 CNN。比較 Method c、3 發現中 Ensemble CNN G-mean 0.88 優於原本單一 CNN 的 G-Mean0.71；從 method d、4 亦發現有進行 Wiener filter 與 Data Augmentation 取得相對於 method1 也有改善的效果。

同時，於實驗進行中曾發現，各別訓練 4 方法的子模型時，偶會出現區域最佳解的弱分類器，但由於計算權重時會以子模型的驗證準確率作調整，因此，弱分類器並不會影響整體判讀結果。

由表 4-3、4-4 總結，可以發現，Ensemble CNN 訓練多個 CNN 使得強分類器可以正常發揮分類功能，區域最佳解的弱分類器降低對影像判讀的影響；同時也看到 Ensemble CNN 不適合用於資料量少的資料集，因為切割子訓練集會使得原本資料量低的類別資料訓練樣本更少，造成弱分類器的大量產生。大量的資料集，或者

進行 Data Augmentation 確實可以幫助 Ensemble CNN 穩定生成多個強分類器。

表 4- 3 Ensemble CNN 實驗結果(資料來源:本研究實驗資料)

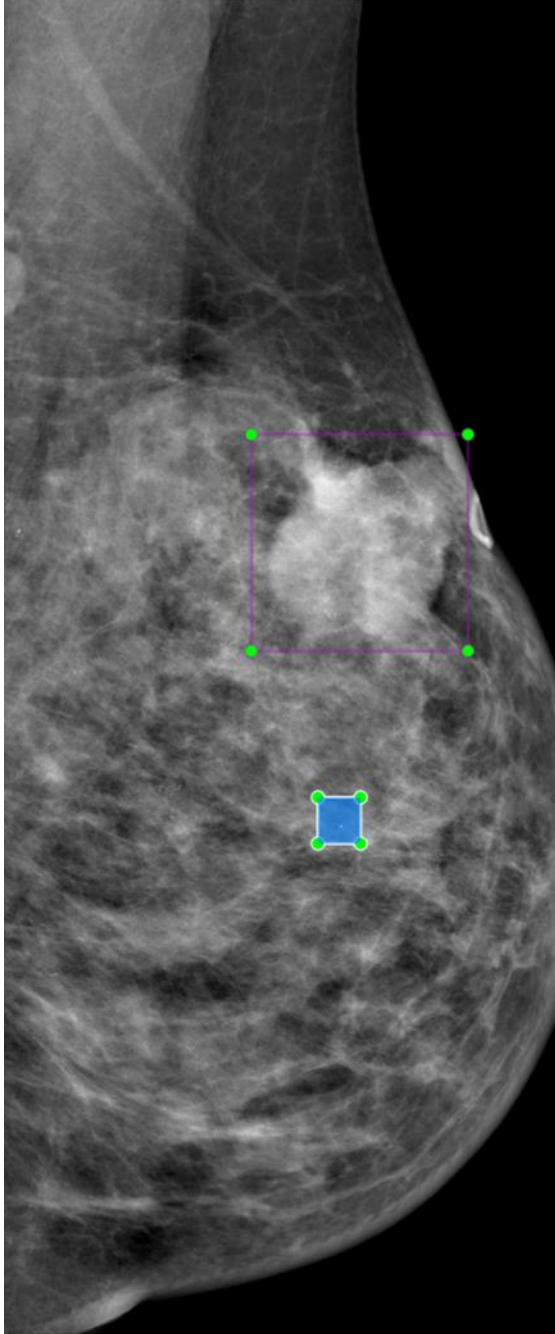
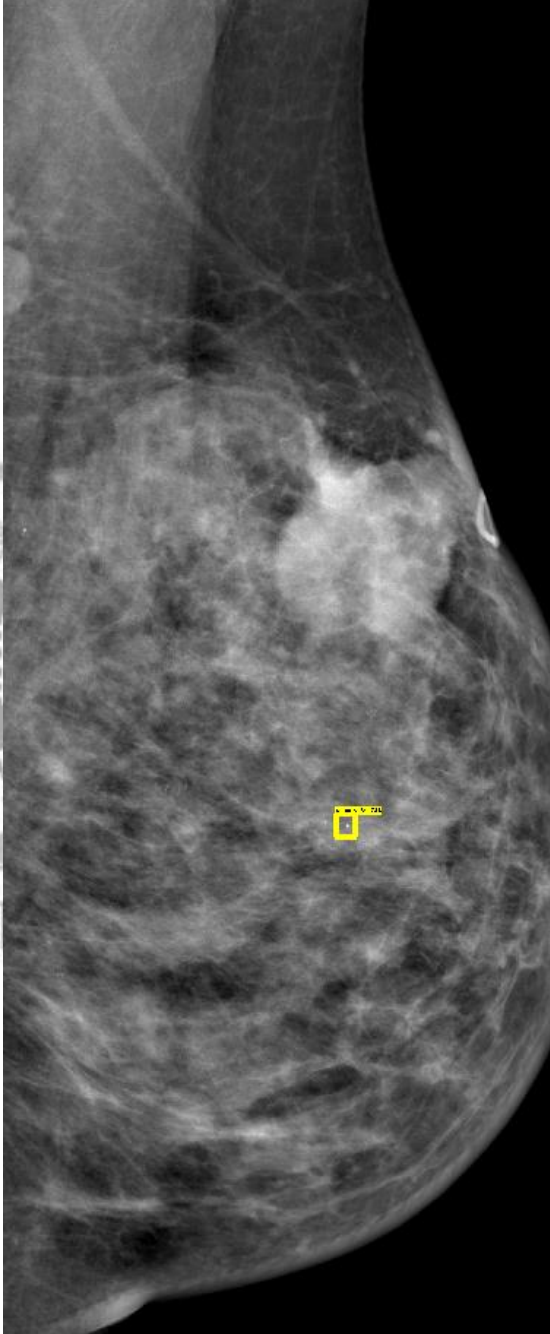
Method	F1-score weighted	Macro-precision	Macro-recall	Kappa	G-mean
a	0.50	None	0.50	0.29	0
b	0.77	0.83	0.78	0.71	0.79
c	0.87	0.90	0.87	0.83	0.88
d	0.65	0.70	0.66	0.56	0.65

表 4- 4 CNN 實驗結果(資料來源:Lu et al.2019)

Method	F1 score weighted	Macro-precision	Macro-recall	Kappa	G-Mean
(1)	0.6	0.53	0.5	0.43	0.46
(2)	0.69	0.72	0.65	0.59	0.62
(3)	0.7	0.69	0.72	0.61	0.71
(4)	0.69	0.77	0.65	0.59	0.61

實驗 2 產生的圈選結果如表 4-5 所示，表格左側是醫師圈選的區域，醫師於此張範例圈選了一個 BI-RADS 2、一個 BI-RADS 4，而 Faster RCNN 在定義閾值為 0.7 狀況下，僅圈出 BI-RADS 2，無法正確圈選出 BI-RADS 4，實驗嘗試將圈選閾值降至 0.1，該模型才圈選另一肉眼可見的大腫塊，但判定該腫塊為 BI-RADS 3，與醫師判定 BI-RADS 4 相異，對於此模型研判閾值不夠高導致圈選區域不足與分類不夠精確是因為樣本數量過少且類別資料不平衡，此訓練資料中 BI-RADS 2、3 資料相對較多，其他分類相對較少，若要改善應提升相對應分類的訓練資料。同時，於實驗過程中也發現，此模型會圈選出許多醫師遺漏圈選的 BoundingBox，可以證實 Faster RCNN 確實具備輔助放射師醫師圈選病灶的能力。

表 4-5 圈選比較(資料來源:本研究實驗整理)

醫師圈選	Faster RCNN 圈選
 <p>This panel shows a mammogram image with two manual selection boxes. The first box is a purple rectangle with green corner markers, highlighting a large, irregularly shaped, high-density mass in the upper central region of the breast. The second box is a small blue square with green corner markers, located in the lower central region of the breast.</p>	 <p>This panel shows the same mammogram image with a single automated selection box. It is a small yellow rectangle with black corner markers, located in the lower central region of the breast, corresponding to the location of the second manual box in the left panel.</p>

第伍章 結論

深度學習應用於醫學在此次 COVID 19 疫情中大放異彩，不論是判讀肺部 X 光或是判讀乳房攝影圖像都仰賴研究人員修改模型，提高判讀準確率，本研究於此研究發現：

1. Ensemble CNN 可有效降低單一模型陷入區域最佳解導致模型不具足夠的判讀能力的情形;而使用該方法須注意訓練資料集應具有一定數量、且類別資料應該要盡可能的平衡。
2. Pretrained model 有助於降低訓練 Faster RCNN 的時間。
3. Faster RCNN 確實可以圈選出許多醫師遺漏的病灶，但仍須提升準確率，供醫師對病情進行判讀。

此研究仍有許多可以改善之處，例如修正標籤規範、此次研究於資料整理階段大量篩選出異常標籤，未來應與醫師溝通統一標籤規範，讓醫師標註的資料可以納入訓練;同時也應儘量取得更多數據資料，平衡數據集各類數量，讓 Faster RCNN 可以充分學習，提高精準度，期望此篇論文可做為日後研究基石，在未來完成一個高分辨率的模型，讓病人可以提早獲得醫師處置與治療。



參考文獻

1. 衛生福利部 (2019) 。 死因統計 。 取自: <https://dep.mohw.gov.tw/DOS/np-1776-113.html> 。
2. Akselrod-Ballin, A., Karlinsky, L., Alpert, S., Hasoul, S., Ben-Ari, R., & Barkan, E. (2016). A region based convolutional network for tumor detection and classification in breast mammography. In *Deep learning and data labeling for medical applications* (pp. 197-205). Springer, Cham.
3. Akselrod-Ballin, A., Karlinsky, L., Hazan, A., Bakalo, R., Horesh, A. B., Shoshan, Y., & Barkan, E. (2017). Deep learning for automatic detection of abnormal findings in breast mammography. In *Deep learning in medical image analysis and multimodal learning for clinical decision support* (pp. 321-329). Springer, Cham.
4. Alejo, R., Sotoca, J. M., & Casañ, G. A. (2008, September). An empirical study for the multi-class imbalance problem with neural networks. In *Iberoamerican Congress on Pattern Recognition* (pp. 479-486). Springer, Berlin, Heidelberg.
5. Bleyer, A., Baines, C., & Miller, A. B. (2016). Impact of screening mammography on breast cancer mortality. *International journal of cancer*, 138(8), 2003-2012.
6. Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).
7. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
8. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing*

systems (pp. 1097-1105).

9. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42, 60-88
10. Lu, H. C., Loh, E. W., & Huang, S. C. (2019, May). The Classification of Mammogram Using Convolutional Neural Network with Specific Image Preprocessing for Breast Cancer Detection. In *2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)* (pp. 9-12). IEEE.
11. Nath, A. M. S. S. (2019). *BREAST CANCER DETECTION IN MAMMOGRAMS USING DEEP NEURAL NETWORKS* (Doctoral dissertation, Indraprastha Institute of Information Technology New Delhi)..
12. Nelson, H. D., Cantor, A., Humphrey, L., Fu, R., Pappas, M., Daeges, M., & Griffin, J. (2016). Screening for breast cancer: a systematic review to update the 2009 US Preventive Services Task Force Recommendation.
13. Pizer, S. M., Johnston, R. E., Ericksen, J. P., Yankaskas, B. C., & Muller, K. E. (1990, May). Contrast-limited adaptive histogram equalization: speed and effectiveness. In [1990] *Proceedings of the First Conference on Visualization in Biomedical Computing* (pp. 337-345). IEEE.
14. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (pp. 91-99).
15. Ribli, D., Horváth, A., Unger, Z., Pollner, P., & Csabai, I. (2018). Detecting and classifying lesions in mammograms with deep learning. *Scientific reports*, 8(1), 1-7.
16. Sickles, E. A., D'Orsi, C. J., Bassett, L. W., Appleton, C. M., Berg, W. A., &

Burnside, E. S. (2013). ACR BI-RADS® Atlas, Breast imaging reporting and data system. *Reston, VA: American College of Radiology.*





附錄甲 實驗混淆矩陣

表 附錄- 1 Method a 混淆矩陣(資料來源:本研究實驗數據)

Actual\Predict	1	2	3	4	5	Sum
1	160	33	43	3	0	239
2	130	55	53	2	0	240
3	49	17	123	2	0	191
4	7	0	19	26	0	52
5	0	0	2	8	0	10
Sum	346	105	240	41	0	732
Recall	0.67	0.23	0.64	0.5	0.0	
Precision	0.46	0.52	0.51	0.63	None	
F1 score	0.55	0.32	0.57	0.56	0.0	

表 附錄- 2 Method b 混淆矩陣(資料來源:本研究實驗數據)

Actual\Predict	1	2	3	4	5	Sum
1	374	60	11	8	0	453
2	139	280	23	9	0	451
3	60	73	231	4	0	368
4	0	0	0	376	0	376
5	0	0	2	2	89	93
Sum	573	413	267	399	89	1741
Recall	0.83	0.62	0.63	1.0	0.96	
Precision	0.65	0.68	0.87	0.94	1.0	
F1 score	0.73	0.65	0.73	0.97	0.98	

表 附錄- 3 Method c 混淆矩陣(資料來源:本研究實驗數據)

Actual\Predict	1	2	3	4	5	Sum
1	428	8	12	5	0	453
2	121	282	44	4	0	451
3	13	14	337	4	0	368
4	0	1	0	374	0	375
5	0	1	0	0	91	92
Sum	562	306	393	387	91	1739
Recall	0.66	0.6	0.69	0.98	0.69	
Precision	0.76	0.92	0.86	0.97	1	
F1 score	0.94	0.63	0.92	1	0.99	

表 附錄- 4 Method d 混淆矩陣(資料來源:本研究實驗數據)

Actual\Predict	1	2	3	4	5	Sum
1	329	87	29	8	0	453
2	249	134	37	28	3	451
3	61	47	234	21	5	368
4	0	2	0	374	0	376
5	1	1	1	11	79	93
Sum	640	271	301	442	87	1741
Recall	0.73	0.30	0.64	0.99	0.85	
Precision	0.51	0.49	0.78	0.85	0.91	
F1 score	0.60	0.37	0.70	0.91	0.88	