

科技部補助
大專學生研究計畫研究成果報告

計 畫 名 稱	： 基於BI-RADS分級的良/惡與半惡性分類深度學習演算法 之乳房鈣化點辨識系統
------------	--

執行計畫學生：蘇瑞慈

學生計畫編號：MOST 107-2813-C-040-004-E

研究期間：107年07月01日至108年02月28日止，計8個月

指導教授：賴慶祥

處理方式：本計畫涉及專利或其他智慧財產權，2年後可公開查詢

執行單位：中山醫學大學醫學資訊學系

中華民國 108年03月29日

(一) 摘要

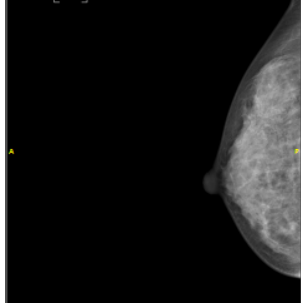
乳癌為我國婦女發生率第 1 位之癌症，死亡率也高居第四位，同時也是十大癌症中上升速度最快者，而國人偏好油炸飲食、晚婚生育和停經後肥胖等因素，導致乳癌以每年百分之三到五的發生率持續上升。乳癌可經由篩檢提早發現，透過篩檢偵測出乳房鈣化點或微小腫瘤，然而，有些微小鈣化點型態仍無法明確判斷屬良性或惡性，此時便需作切片化驗，醫師須透過切片檢查及自身經驗才能確認是否為惡性鈣化點。目前診斷乳房鈣化點的方式是以人工進行影像判讀，但此種方式會因每位醫師判讀經驗的差異而有不同的結果，可能導致治療時間延遲。

為了解決以上的問題，本計畫提出了一套結合 BI-RADS 分級與具有良/惡與半惡性分類的乳房鈣化點輔助診斷系統，BI-RADS 分級 0、分級 1、分級 2 與分級 6 屬良/惡性，而分級 3~分級 5 屬半惡性，由於大多數研究僅探討乳房鈣化點的良惡辨識，極少數的研究提及半惡性，醫師也較難辨識，因此本計畫藉由深度學習演算法解決此問題，利用具良/惡與半惡性自動標記之深度學習演算法進行訓練，以得知乳房鈣化點的良惡。本計畫透過 Deep Learning 深度學習演算法結合乳房攝影報告與資料分析系統 BI-RADS 的鈣化點分級，不僅辨識乳房鈣化點的良惡，也能判定出半惡性，減少乳癌患者做乳房相關檢測的費用與等待結果的時間，輔助醫師在最短時間內做出更準確的診斷，並提高乳房鈣化點良惡辨識的正確率。

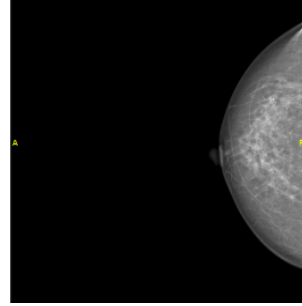
(二) 研究動機與研究問題

根據國民健康署癌症登記統計資料顯示，乳癌為我國婦女發生率第 1 位之癌症，同時也是十大癌症中上升比例最高者，發生高峰約在 45~69 歲之間，每年有逾萬位婦女罹患乳癌，近 2,000 名婦女死於乳癌，相當於每天有 31 位婦女被診斷罹患乳癌、6 位婦女因乳癌而失去寶貴性命[1]。乳癌可經由篩檢提早發現，透過篩檢偵測出乳房鈣化點或微小腫瘤[2]，使病人能夠盡早接受治療，日後可提升病人的存活率，大幅減少因乳癌而死亡的人數。

「乳房鈣化」是指乳房組織內的鈣化物以鈣鹽形式存在，而鈣化現象會發生於壞死組織、發炎組織、疤痕組織、良性乳房腫瘤或是癌組織中，因此鈣化本身即是一種現象，並非指一特定之疾病。分辨良性或惡性乳房鈣化，可從乳房 X 光影像中鈣化點的形狀、分佈情形及是否合併鈣化點周圍乳腺組織密度變化等多項因素加以分析。一般「良性乳房鈣化點」，形狀多為圓形、爆米花狀、茶杯形，分佈上也比較均勻或對稱，如圖 1(a)所示；而「惡性乳房鈣化點」有其特殊型態，形狀上較不規則、大小不一，分佈上常呈現局部聚集或沿著乳管形成線狀分佈，有時合併腫塊之影像存在，如圖 1(b)所示。



(a)



(b)

圖 1、(a)乳房鈣化點良性影像，(b)乳房鈣化點惡性影像

目前診斷乳房鈣化點狀態的方式是以人工進行影像判讀，但此種方式會因每位醫師判讀經驗的差異而有不同的結果，可能導致治療時間延遲。此外，雖然上述特徵可以幫助我們偵測出早期乳癌或原位癌(第 0 期乳癌)[3]，然而，有些微小鈣化點型態仍無法明確判斷屬良性或惡性，也就是所謂「半惡性」，如圖 2 所示，此時便需作切片化驗，醫師須透過切片檢查及自身經驗才能確認是否為惡性鈣化點，且開刀必會伴隨著一定的風險。

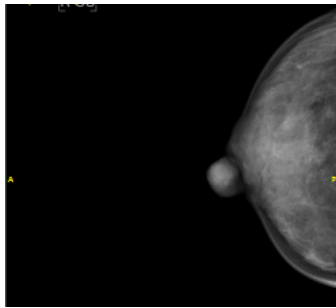


圖 2、乳房鈣化點半惡性影像

隨著醫療技術的進步，許多學者紛紛投入發展以醫學影像進行判讀的電腦輔助診斷系統，而本計畫為了使醫學影像判讀上的準確率能大幅提高，因此提出以深度學習技術為架構，發展出一套結合 BI-RADS 分級與具有良/惡與半惡性分類的乳房鈣化點輔助診斷系統，讓醫生除了依靠臨床經驗之外，也能透過參考本系統更精準的分辨乳房鈣化點的良惡，不僅提高辨識的正確率，同時縮短診斷的時間，減少醫師因疲勞而有誤判的情況發生。

(三) 文獻回顧與探討

近年來乳癌發生率逐年上升，許多女性患者不了解乳癌篩檢的重要性，導致到醫求診時已經是晚期的疾病。若能輔助醫生更準確且快速的判斷乳房鈣化點之良惡，並盡早治療，必能將低乳癌的發生率。此外，雖然至今已有許多分析乳房鈣化點良惡的論文被提出，但針對辨識乳房鈣化點半惡性的方法仍極為少數。若能參考多篇文獻做為基礎，並發展出判斷乳房鈣化點良惡的方法，將更具意義。

2018 年 Juan Wang 等人提出一個關於內容敏感的深度神經網路[4]，目的在偵測局部特徵及其周圍組織背景。深度神經分類器經過訓練後自動抽取鈣化點的

相關特徵內容，該方法在 292 張乳房攝影影像上進行分析，使用自由響應接收者操作特徵曲線(Free-response Receiver Operating Characteristic, FROC)在單個鈣化點和群聚鈣化點上評估準確度，結果表示，與兩個基於鈣化點偵測的方法相比，該方法實現更顯著的 FROC 曲線。此研究由兩個子網路組成，一個用於抽取鈣化點的局部影像特徵，另一個用於表徵其周圍的組織背景，結果表明，在鈣化點偵測中結合影像周圍資訊可確實提高辨識準確率。

2016 年 Thijs Kooi 等學者發表了一篇利用大尺度深度學習辨識乳房攝影病變的論文[5]，此研究提及傳統電腦輔助偵測系統(Computer-Aided Detection, CAD)，但基於 CAD 系統對於乳房病變的判定較不準確，因此提出利用深度學習之技術檢測乳房病變，首先針對乳房攝影的影像進行影像擴增，改善資料缺乏並防止過度擬合(overfitting)的問題，網路架構則是利用 VGG 架構的縮小版本，並使用小批量隨機梯度漸降法(minibatch Stochastic Gradient Descent, minibatch SGD)進行參數最佳化。此研究提出辨識乳房病變的演算法，辨識所得出的結果與放射科醫師判定的病變相似，但並未提出如何辨識乳房鈣化點半惡性。

2015 年 Nijad Al-Najdawi 等學者發表了有關於乳房攝影影像之增強、切割與分類的論文[6]，首先針對乳房攝影的影像進行去雜訊，接著採用三種方法增強影像：對比有限自適應直方圖均化(Contrast Limited Adaptive Histogram Equalization, CLAHE)、中值濾波器(Median filtering)以及高斯平滑化(Gaussian smoothing)。接下來利用 Canny 進行邊緣偵測，切割出影像中感興趣的區域(Region of Interest, ROI)，最後將切割的影像分類，首先計算腫塊的似圓性(Roundness)並得到一值 Metric，接著設定 Metric 的閾值，將腫塊分為四類：良性、較可能為良性、較可能為惡性以及惡性。此研究提出結合 CLAHE 和 Median filtering 的影像增強方法，所切割出的區域跟放射科醫師判斷的區域較為接近，但其特徵只採用似圓性進行分類，方法過於單一，因此準確性有待商榷。

2014 年 Rahimeh Rouhi 等學者提出有關良性和惡性乳房腫瘤分類的論文[7]，使用的方法為區域成長法(Region Growing)和 CNN 深度學習，並將兩者的結果進行比較。研究裡面提及將影像透過區域成長法進行分類並自動切割，透過訓練的人工神經網絡(Artificial Neural Network, ANN)獲得自適應閾值(adaptive threshold)；在另一方法中，分割是由類神經網絡(CNN)執行，其參數由基因演算法(Genetic Algorithm, GA)所決定，從分段腫瘤中提取密度，紋理和形狀特徵，基因演算法用於從提取的特徵集中選擇適當的特徵。最後，使用不同的分類器，例如隨機森林(random forest)、單純貝氏分類器(naive bayes classifier)、SVM 和 KNN，評估所提出方法的性能，本計畫表示在腫瘤分類中，CNN 所得的效果優於區域成長法。

而在 2017 年，Hae Kyoung Jung 與其他學者提出了應用 BI-RADS 乳房攝影影像對乳房病變的評估[8]，本計畫使用 BI-RADS 乳房攝影影像報告並對 DBCT 乳腺病變進行評估，總共有 34 個病變，包含 24 個(70.6%)腫塊、7 個(20.6%)鈣化和 3 個(8.8%)結構畸形，其中有 8 個(23.5%)病變是惡性的，26 個(76.5%)是良

性的，接著使用廣義估計方程式和 Fleiss kappa(κ)統計量來計算其診斷性能，以區分良性和惡性病變以及評估觀察者間變異性。此計畫發現有放射科醫師使用 BI-RADS 治療 DBCT 病變的結果整體是高性能的。

綜合以上的文獻回顧，本計畫發現大部分的文獻僅探討乳房鈣化點的良性及惡性，極少數的研究提及半惡性，而 BI-RADS 為美國放射醫學學會所發展的系統，相信此系統對乳房的檢測有很大的可信度，因此本計畫將利用由附設醫院所獲得的乳房攝影之影像與 DDSM 影像並結合深度學習(Deep Learning)演算法進行分析，將影像輸入至循環生成對抗網路(Cycle-Consistent Generative Adversarial Network, CycleGAN)自動標記乳房鈣化點良性、惡性及半惡性。

(四) 研究方法及步驟

本計畫所提出之系統流程如圖 3 所示，主要可分為「輸入原始之乳房攝影影像/輸入醫師標記之乳房攝影影像」、「良惡與半惡性自動標記深度學習演算法」及「結果」，以下將針對各步驟詳細說明。

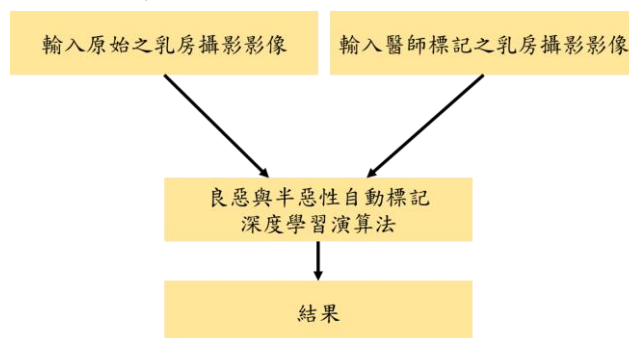


圖 3、系統流程圖

1. 輸入原始之乳房攝影影像/輸入醫師標記之乳房攝影影像

此步驟包含輸入原始影像以及醫師標記後影像，醫師需針對乳房攝影影像進行手動標記，而本計畫也開發介面供醫師進行標記，如圖 4 所示。

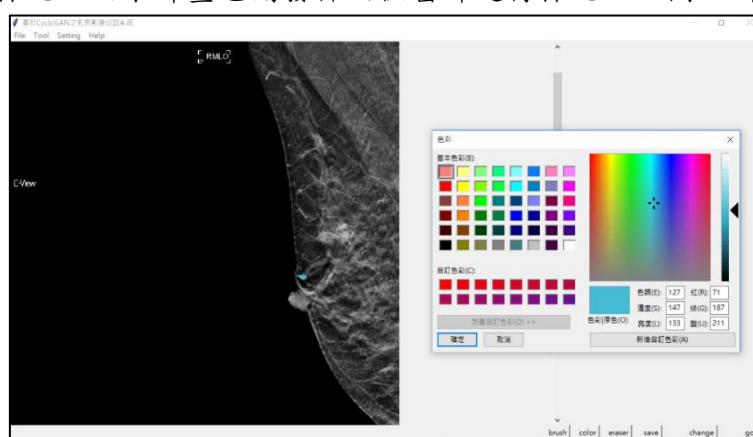
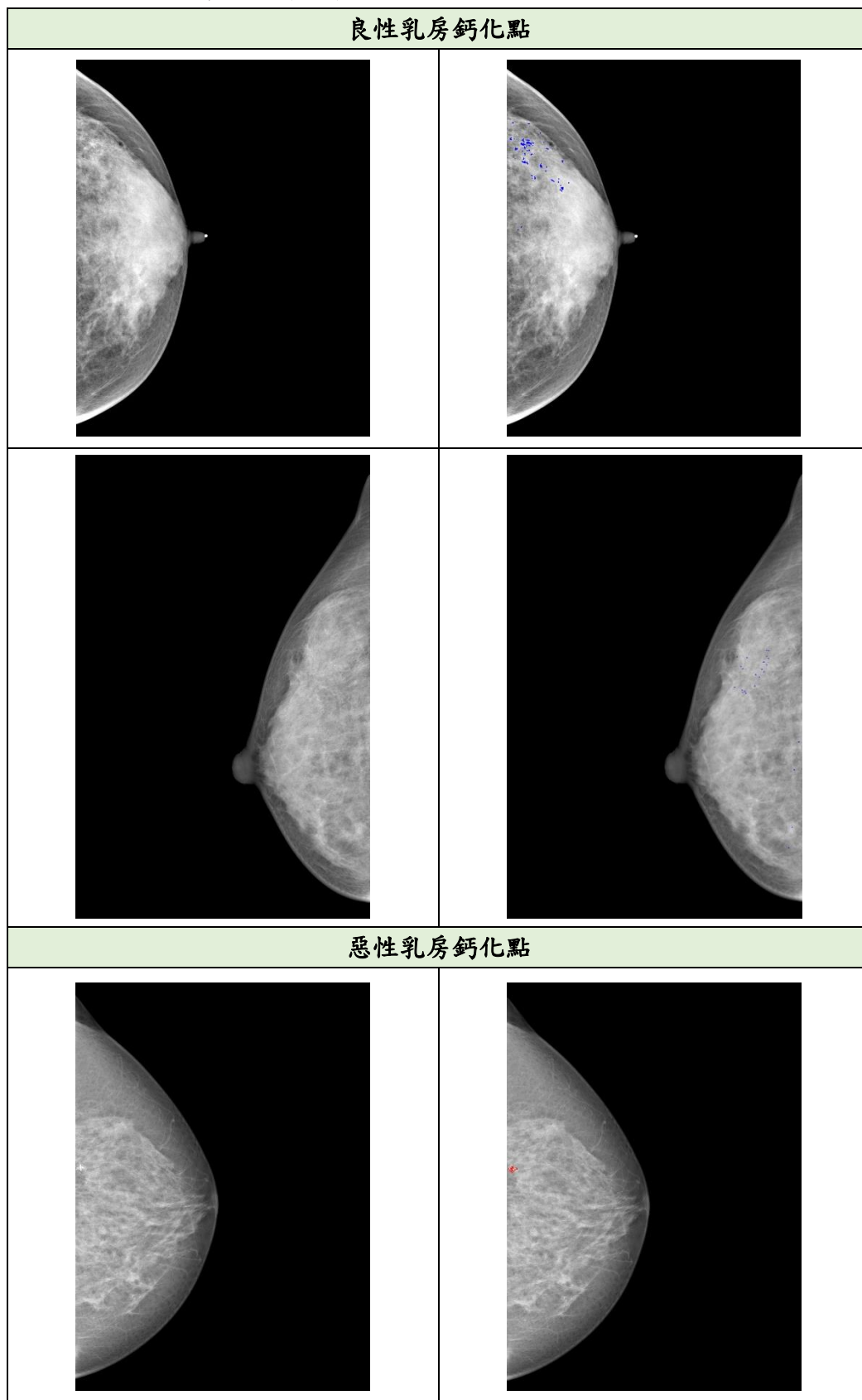


圖 4、醫師標記影像之介面

首先，開啟一張乳房攝影影像並輸入至系統，系統會產生筆刷、筆刷顏色選擇、筆刷粗細大小等按鈕，醫師可直接透過此介面在乳房攝影影像上進行乳房鈣化點標記，並將標記後的影像存入資料集中，作為深度學習演算法

訓練之影像，如圖 5 所示，其中藍色標記、紅色標記、綠色標記分別表示為良性、惡性及半惡性。而本計畫在過程中持續與醫學影像放射科醫師進行討論，藉以提高本計畫判斷的準確性。



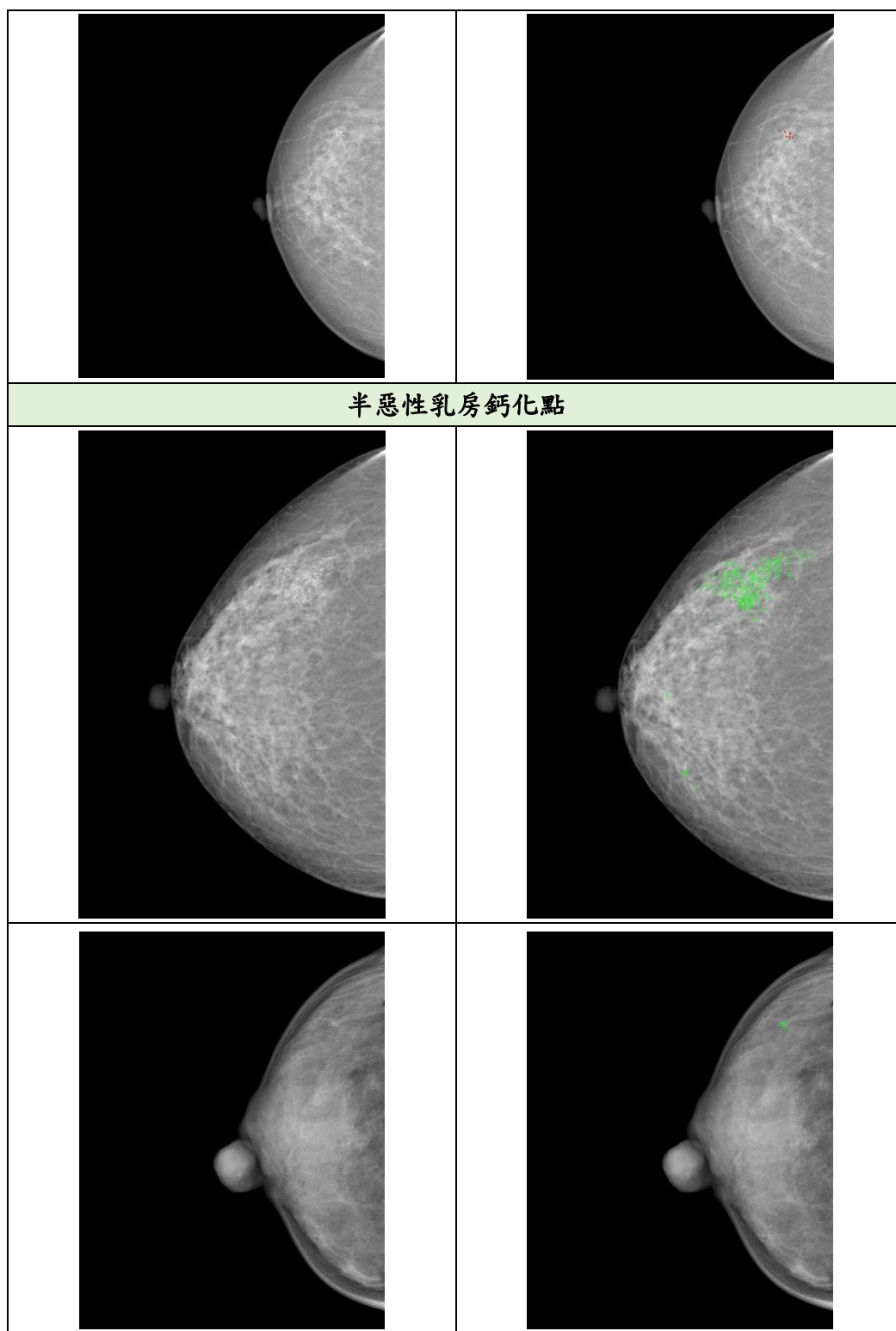


圖 5、乳房攝影影像(左欄為原始影像、右欄為醫師標記後影像)

2. 良惡與半惡性自動標記深度學習演算法

本計畫選擇以 Deep Learning 深度學習[9-11]演算法做為分類器，並將以上特徵經由此分類器進行訓練。以往的技術當中，大部分都是選擇 Support Vector Machine(SVM)做為分類器[12]，但是當測試的資料數目過大時，就需要花費較多的時間。而 Deep Learning 深度學習是機器學習研究中的新領域，

其動機在於模擬人腦進行分析學習的神經網路，藉由模仿人腦的機制來解釋數據，例如訊號、聲音及影像等。

本計畫使用 CycleGAN 深度學習演算法[13]，針對乳房攝影影像中的鈣化點進行標記，以輔助醫師快速辨別乳房攝影影像，便能擬定合適的治療方針，有助於病人的治療。

CycleGAN 即是 GAN 所衍生的分支，是一種在沒有配對範例資料的情況下進行學習的方法，將影像從 domain X 轉換為 domain Y，目標是學習映射 $G: X \rightarrow Y$ ，使得來自 $G(X)$ 的影像與使用對抗性損失的目標域 Y 無法區分。CycleGAN 可以讓兩個不同類別的圖片互相轉換，有別於傳統單向生成的 GAN，CycleGAN 是雙向生成且透過兩個鏡像的 GAN 所組成，如圖 6 所示，要將圖片進行轉換，首先需要有一個生成器 A，輸入一張 domain X 中的影像會生成 domain Y 的影像，並透過判別器來判別其是否符合 domain Y 的影像。此時，還需要有另外一個生成器 B，能讓 domain Y 的影像再轉回 domain X，如此一來可避免生成器 A 忽視輸入之圖片，產生符合判別器標準的 domain Y 的影像。最後，因生成器 B 也需要有判別器來判別產生的影像是否符合 domain X 的影像，所以需要有鏡像的生成對抗網路，如此循環下去，以提升整體網路產生圖片之準確度。

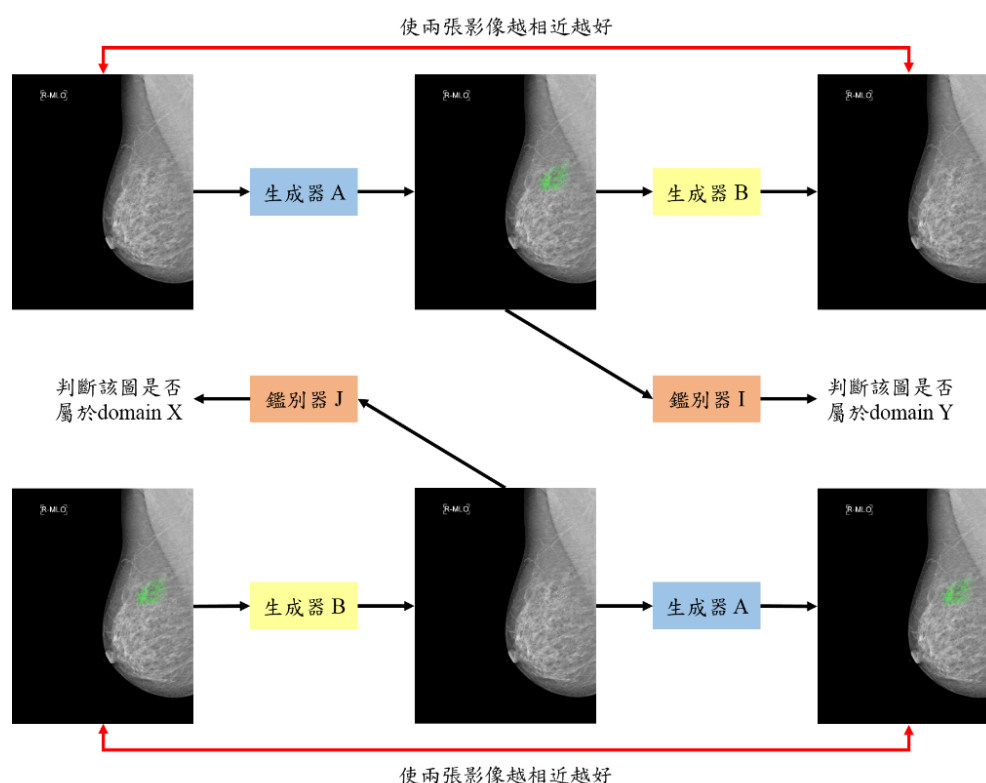


圖 6、CycleGAN 訓練流程示意圖

另外，本計畫也結合「乳房攝影報告與資料分析系統」(Breast Imaging Reporting and Data System, BI-RADS)的分類等級[14]，藉由分析乳房攝影上的特定鈣化型態，可將攝影結果判別成幾個不同的類別，BI-RADS 的分級

含 0 到 6 共 7 個等級，所代表的意義簡單說明如下：分級 0 是「未完整的影像」，意指單靠乳房攝影並不能夠做到準確判定。在此情形下，需借助其他檢查例如乳房超音波，彌補乳房攝影對緻密乳房敏感度較低的缺點，或對較密集的鈣化群區域做乳房攝影的局部放大影像，重複判定。分級 1 為「正常」，是指完全沒看到鈣化或腫瘤。分級 2 為「良性」，是指一些良性的發現，例如纖維腺瘤、明顯良性鈣化或正常淋巴結等。分級 3 為「可能良性」，比起分級 2 多了一點不確定性，但還沒立刻切片化驗必要，而是以 6 個月的重複攝影追蹤，來比較鈣化前後差異，並重新判斷其好壞，惡性鈣化點的機率低於 2%。分級 4 為「懷疑」，意即有惡性的疑慮而需考慮做切片檢查，才不會延遲乳癌的確診與治療，惡性的機率為 2%~95%。而到了分級 5 為「惡性」，有超過 95%的機率確定會成為惡性腫瘤，必須馬上進行治療。分級 6 則是指該個案已確診為「惡性腫瘤」，影像檢查是為了做治療前評估或治療後追蹤。而我們將 BI-RADS 分級 0~分級 2 表示為「良性」、BI-RADS 分級 3~分級 5 表示為「半惡性」、BI-RADS 分級 6 表示為「惡性」。

(五) 結果

1. 乳房攝影影像資料集

本計畫的影像來源為中山醫學大學附設醫院(Chung Shan Medical University Hospital)醫學影像放射科醫師所提供的乳房攝影影像，以及乳房攝影數位資料集(Digital Database for Screening Mammography, DDSM)。前者的影像資料集為 2010 年 9 月至 2015 年 1 月，總共 32 位病患的影像，每個個案皆會拍攝乳房上下側照像(Cranio-Caudal view, CC view) 及左右側照像(Mediolateral-Oblique view, MLO view)，而在輸入影像前會先進行影像去連結的動作，因此不會有病人隱私洩漏的問題；後者則是由麻省總醫院(Massachusetts General Hospital)、南佛羅里達大學(University of South Florida)和桑迪亞國家實驗室(Sandia National Laboratories, SNL)共同參與的合作項目，是乳房攝影影像分析研究團體使用的資源，它包含正常、良性病例、惡性病例以及經過驗證的病理資訊。

2. 開發工具與環境

在實作方面，本計畫透過 NVIDIA GPU 加速深度神經網路的訓練，GPU 含有數千個更小型、高效率的核心，其加速運算效能高且能有效處理平行運算，負責運算密集的工作進而減輕 CPU 的負擔，可在極短的時間內處理大量的資料，學習速度提高 10 至 100 倍以上。

本計畫在 Ubuntu 16.04 作業系統搭配 Intel i7-3960X CPU 與 NVIDIA GTX 2080 的顯示卡，運用 Anaconda3 建立 Tensorflow 深度學習框架，以 Python 程式語言進行 CycleGAN 深度學習架構的撰寫及網路訓練，並使用 Python 的 GUI 套件 Tkinter 來完成標記介面的設計與開發，Tkinter 具有跨平台的能力可以符合多種不同的作業環境，並支持多種圖形資料庫的使用。

3. 實驗結果

將原始乳房攝影影像和醫師標記後影像輸入至 CycleGAN 深度學習演算法，本計畫的實驗可分為實驗 1 與實驗 2，兩者實驗不同之處為影像前處理的流程，以下將針對兩者實驗進行說明。

實驗 1：由於 CycleGAN 訓練時輸入影像為 256×256 ，因此乳房攝影影像在醫師標記後需調整為 256×256 ，以便進行後續的辨識與分析，實驗 1 的影像前處理流程如圖 7 所示，而 CycleGAN 迭代次數與輸出影像如圖 8 所示，透過圖 8 可以看出，輸出影像並未標記出乳房鈣化點。

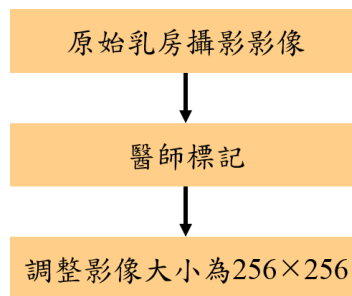


圖 7、實驗 1 之影像前處理流程

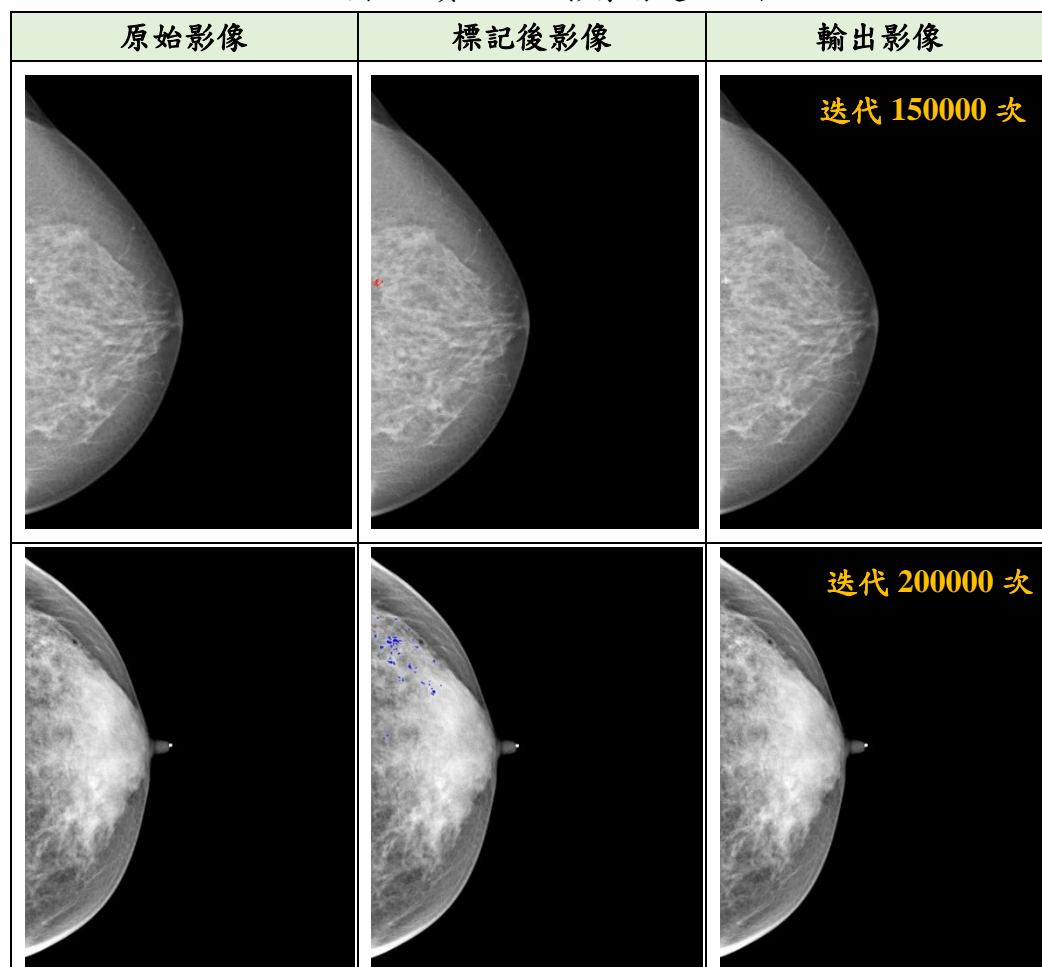


圖 8、實驗 1 之迭代次數與輸出影像

實驗 2：相較於實驗 1，實驗 2 的影像前處理流程如圖 9 所示，而 CycleGAN 迭代次數與輸出影像如圖 10 所示，透過圖 10 可以看出，輸出影像已有標記出鈣化點位置，不過相較於醫師標記的影像，標記之顏色仍有些許落差。未來期望持續與中山醫學大學放射科醫師合作，以便獲取更多的實驗影像，希望藉由以上方法有效提升系統準確率。

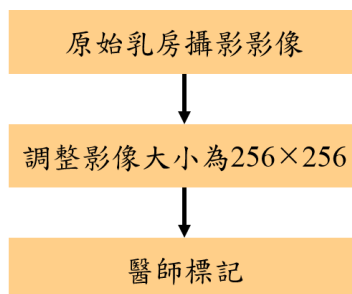


圖 9、實驗 2 之影像前處理流程

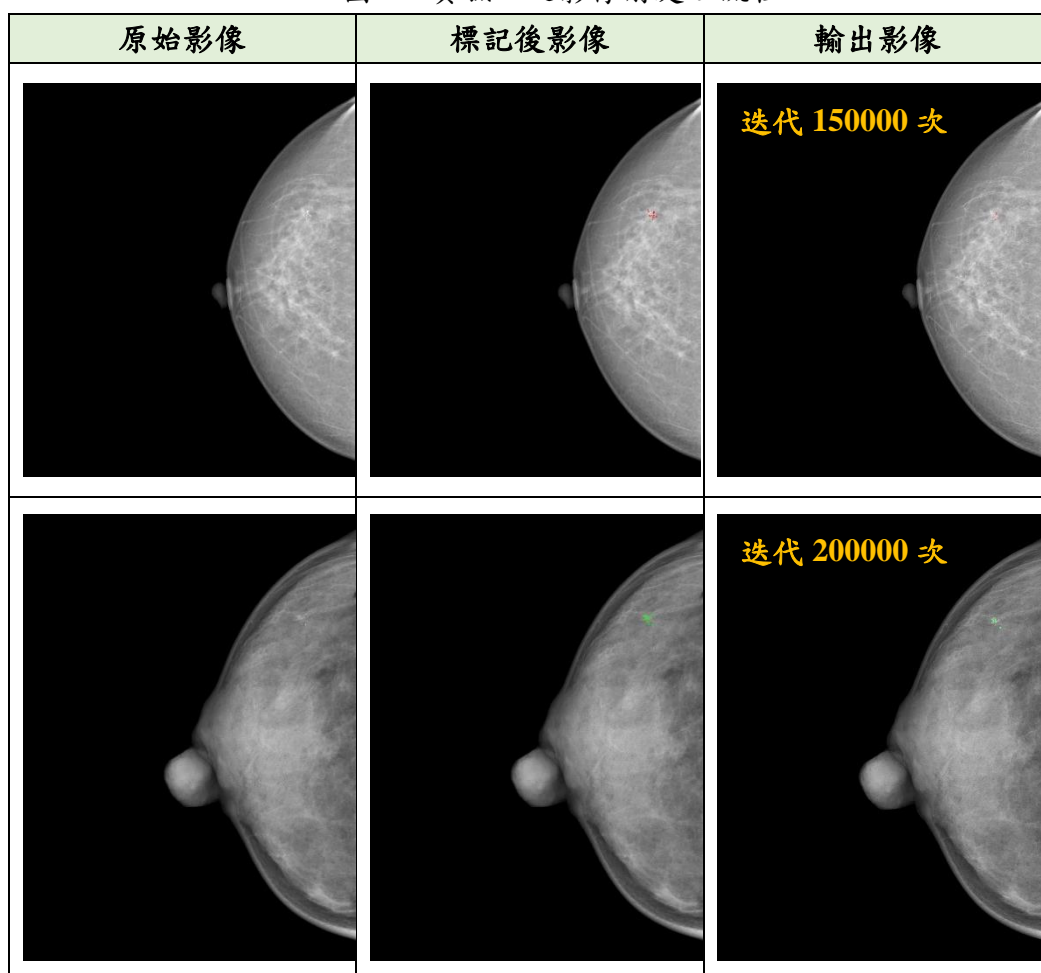


圖 10、實驗 2 之迭代次數與輸出影像

(六) 結論

乳房鈣化點的判讀十分仰賴有經驗的放射科醫師的專業知識，而目前醫師的診斷方法會依據鈣化的大小、形狀，以及鈣化點聚集的情況來判斷是良性或是惡

性，若無法立即判斷是良性或惡性，則會依病患的不同情況建議進一步的檢查，幫助確診，例如：半惡性，但是長時間下來醫師便容易感到疲倦，容易發生誤判的狀況。

本計畫以近年來人工智慧主流的深度學習為架構，提出了一套結合 BI-RADS 分級與具有良/惡與半惡性分類的乳房鈣化點輔助診斷系統，如此能提升醫師診斷的速度，做為開立處方的依據，並且可以減少患者做其他乳房相關檢測的費用以及等待報告結果的時間。

(七) 參考文獻

- [1] DeSantis C et al. "Breast cancer statistics, 2013," CA: A Cancer Journal for Clinicians, vol. 64, pp. 52-62, 2014.
- [2] A. S. Coates et al. "Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015," Annals of Oncology, vol. 26, pp. 1533-1546, 2015.
- [3] Shadi Aminololama-Shakeri et al. "Differentiation of ductal carcinoma in-situ from benign micro-calcifications by dedicated breast computed tomography," European Journal of Radiology, vol. 85, pp. 297-303, 2016.
- [4] Juan Wang and Yongyi Yang, "A context-sensitive deep learning approach for microcalcification detection in mammograms," Pattern Recognition, pp. 12-22, 2018.
- [5] Thijs Kooi et al. "Large scale deep learning for computer aided detection of mammographic lesions," Medical Image Analysis, vol. 35, pp. 303-312, 2016.
- [6] Nijad Al-Najdawi, Mariam Biltawi and Sara Tedmori, "Mammogram image visual enhancement, mass segmentation and classification," Applied Soft Computing, vol. 35, pp. 175-185, 2015.
- [7] Rahimeh Rouhi et al. "Benign and malignant breast tumors classification based on region growing and CNN segmentation," Expert Systems with Applications, vol. 42, pp. 990-1002, 2015.
- [8] Hae Kyoung Jung et al. "Potential Use of American College of Radiology BI-RADS Mammography Atlas for Reporting and Assessing Lesions Detected on Dedicated Breast CT Imaging: Preliminary Study," Academic Radiology, vol. 24, pp. 1395-1401, 2017.
- [9] Guoxing Wu et al. "Regional deep learning model for visual tracking," Neurocomputing, vol. 175, pp. 310-323, 2016.
- [10] B. Chandra and Rajesh K. Sharma, "Fast learning in Deep Neural Networks," Neurocomputing, vol. 171, pp. 1205-1215, 2016.
- [11] Vamsi K. Ithapu et al. "Imaging-based enrichment criteria using deep learning algorithms for efficient clinical trials in mild cognitive impairment," Alzheimer's

- & Dementia, vol. 11, pp. 1489-1499, 2015.
- [12] Maxine Tan et al. "Optimization of breast mass classification using sequential forward floating selection (SFFS) and a support vector machine (SVM) model," International Journal of Computer Assisted Radiology and Surgery, vol. 9, pp 1005-1020, 2014.
- [13] Jun-Yan Zhu et al. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," IEEE International Conference on Computer Vision, 2018.
- [14] Mona El Khoury et al. "Breast imaging reporting and data system (BI-RADS) lexicon for breast MRI: Interobserver variability in the description and assignment of BI-RADS category," European Journal of Radiology, vol. 84, pp. 71-76, 2015.