

探討不同演算核心應於醫學影像品管之研究

Quality use of artificial intelligence in medical images

劉國英¹，呂鎮中¹，黃詠暉²，陳泰賓²，呂南翰¹

¹義大醫療財團法人義大癌治療醫院影像醫學部

²義守大學醫學影像暨放射科學系

研究目的

探討使用不同演算核心建立辨識醫學影像方向模型之研究，期望可由此模型自動辨識醫囑與臨床影像不符之事件，本研究以手部X光為範例，尋找高準確率之模型，並期盼未來可將其結果擴展使用於其他部位之醫學影像。

材料與方法

使用公開資料庫 MURA(musculoskeletal radiographs)之部分手掌影像200張進行模型建立，193張影像進行模型驗證。研究中使用SeNet154、efficientNet-b7、DenseNet201、MobileNet V2、ResNet101、Vgg11、Vgg11bn、Vgg16、Vgg16bn、Vgg19和Vgg19bn，共11種不同的演算核心進行結果比較。除不同演算核心外其餘參數設定皆相同，模型訓練參數如表1所示。

結果

Vgg11、Vgg16、Vgg16bn和Vgg19bn訓練時的模型準確率皆為100%；將所建立的模型使用驗證集資料193張影像進行推論，所獲得之準確率最高為69%，可見訓練時的模型準確率並不同於實際準確率。而演算核心SeNet154其訓練與推論準確率最為接近，分別是98與97%，其模型訓練時間也是最長，為328分鐘。不同演算核心其模型與推論準確率以及訓練時間如表2所示。

表1、模型訓練所使用之設定參數

Batch size	Training epoch	Learning rate	Weight decay	Momentum	Validation data
32	100	0.000001-1	0-0.1	0-0.99	0.2

表2、不同演算核心之模型與推論準確率以及訓練時間

演算核心	模型準確率 (%)	推論準確率 (%)	訓練時間 (mins)
SeNet154	98	97	328
efficientNet-b7	65	51	119
DenseNet201	90	54	120
MobileNet V2	75	51	81
ResNet101	88	49	203
Vgg11	100	51	258
Vgg11bn	98	54	124
Vgg16	100	49	230
Vgg16bn	100	69	196
Vgg19	80	64	108
Vgg19bn	100	68	231

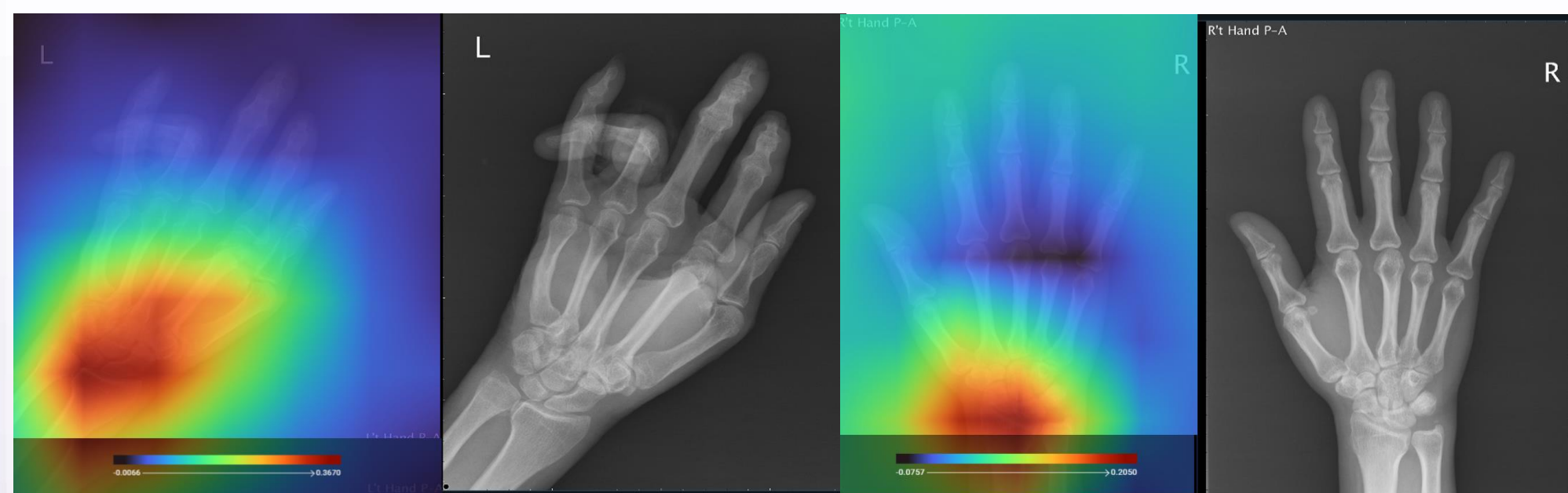


圖1、熱圖影像。由SeNet154所訓練出來的模型，驗證影像所生成之熱圖(Heat map)，左手(左圖)顯示的熱區值最高為0.367，右手(右圖)顯示的熱區值最高為0.205，且皆為手腕處。推估模型應該是依照手腕骨頭型態進行左右手的辨識，非用手掌之型態，故模型之準確率較高。

討論

雖然使用不同演算核心所建立之模型準確率皆超過75%，更有4種核心可達到準確率100%；但將所建立之模型進行推論驗證，發現模型準確率皆無法超過70%。而使用演算核心SeNet154的模型，其推論可達到97%的高準確率，與模型訓練時的準確率98%並無差異，在本研究少量資料集訓練的結果最為優異。

關鍵字: 醫學影像品管、演算核心、SeNet154

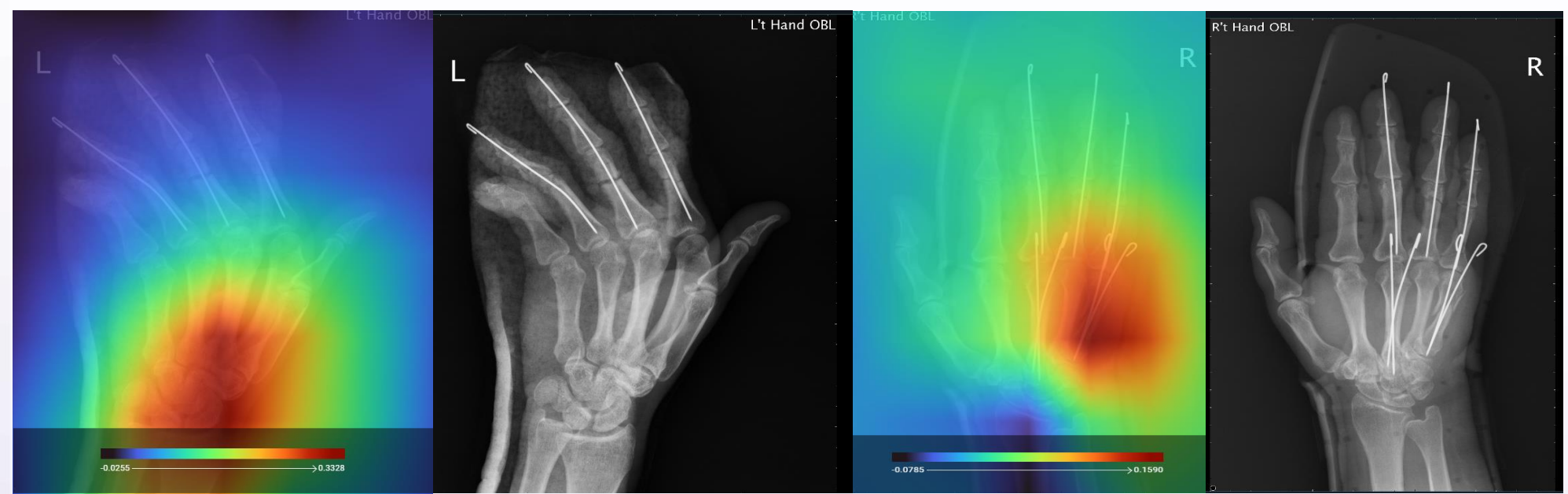


圖2、熱圖影像。由SeNet154所訓練出來的模型，驗證有人工植入物影像所生成之熱圖(Heat map)，左手(左圖)顯示的熱區值最高為0.3328，右手(右圖)顯示的熱區值最高為0.159，且位於掌骨的部分，模型判斷信度值偏低，模型顯示左手與右手的比例為54比上46，判定左手，為錯誤之判斷。

結論

模型訓練結果的準確度並無法代表模型真正的效度，據研究結果顯示，建議使用另外驗證之資料集進行模型效度評估。如圖1熱圖結果所示，推估模型主要是經由手腕進行左右邊之辨識，即使是手掌骨折之影像，模型仍可準確辨識，即便有人工植入物，如圖2左圖，模型亦可正確辨識；如果熱區為掌骨區，如圖2右圖，模型則辨識錯誤。本研究雖無使用巨量資料集，但其結果應可提供欲建立醫學影像辨識模型研究者參考。