深度學習於醫學影像分析—Deep learning in medical image analysis
Fall 2024

Homework 5, due on 2024/12/30

1. The purpose of this homework is to use deep learning methods to **mark abnormal parts in chest X-ray images and annotate possible chest diseases reflected by the abnormal parts**. This homework uses the chest X-ray images provided by the E-Da Hospital with disease labels and location bounding boxes for deep learning model building.

   These chest X-ray images are either from disease-free subjects (normal), or may contain multiple sites of abnormalities (主動脈硬鈣化 (aortic_atherosclerosis_calcification)、主動脈彎曲(aortic_curvature)、肺野浸潤增加(lung_field_infiltration)、胸椎退化性關節病變 (degenerative_joint_disease_of_the_thoracic_spine)、脊椎側彎(scoliosis)、肺尖肋膜增厚(intercostal_pleural_thickening)、心臟肥大(cardiac_hypertrophy)) in one image.

2. The zipped file "**hwk05_data.zip**", which can be downloaded from e3 (https://e3.nycu.edu.tw/) under "Homework 5", contains all datasets for this homework.

   All training data are in the folder "**train**". Under the training data folder, there are 8 sub-folders for 8 different disease categories. Within each category sub-folder, there are two sub-sub-folders "**image**" and "**mark**", which contain original chest X-ray images (.dcm files) and bounding boxes annotating the disease locations (.jpg files), respectively. File names of training chest X-ray images and their corresponding bounding boxes can be found in the file "**train.csv**". Please read the file "**README.pdf**" first for the meaning of each data item in "**train.csv**".

   Original chest X-ray images of test data can be found under the folder "**test/image**", and the file "**test.csv**" has the file names of test data.

3. This homework aims to develop predictive models to **use bounding boxes to label abnormal parts and annotate possible chest diseases reflected by the abnormal parts**. You are ask to:

a. Image normalization: When display the image information of our DICOM files, we can see that “Pixel Intensity Relationship” is “LOG” and “Photometric Interpretation” is “MONOCHROME1”, which is equivalent to “Presentation LUT Shape” is “INVERSE”. The former represents the logarithmic proportion of the relationship between the pixel sample values and the intensity of the X-ray beam. The latter represents the presentation look up table (LUT) transformation inverting the pixel values into P-Values. From these two pieces of information, it can be stated that our dataset should use logarithmic transformation to reverse the fundamental original pixel values back. We thus need to do the following “**intensity log-transformation**”:

Let $x[i]$ is the ith pixel values of the X-ray image, and N is the number of pixels. $iMax$ is defined as the sum of the Window Center (WC) and half of the Window Width (WW). $iMin$ is determined by the difference between WC and half of WW. “BitsStored” is the number of bits stored for each pixel sample.

Input: x

for $i = 0, \cdots, N-1$ **do**

if $x[i] < iMin$, **then** $x[i] = iMin$

if $x[i] > iMax$, **then** $x[i] = iMax$

$$z[i] = -\log\left(\frac{1 + \boldsymbol{x}[i]}{2^{\text{BitsStored}}}\right)$$

Output: z

Chest X-ray images often contain something that is less relevant to our detection, such as right or left chest marks. We can find that after logarithmic transformation, those marks “L” font in the image will be overexposed compared with the overall picture. Thus, we need to adjust the contrast for the image to optimize the structures we consider helpful. The following “**simplest color balance algorithm**” can do this:

Let v_{min} and v_{max} be the saturation extrema, which can be seen as quantiles of the pixel values distribution.

Input: z

for $i = 0, \cdots, N-1$ **do**

$$c[i] = \frac{z[i] - v_{min}}{v_{max} - v_{min}}$$

if $c[i] < 0$, **then** $c[i] = 0$

if $c[i] > 1$, **then** $c[i] = 1$

Output: c

Please perform above intensity log-transformation and simplest color balance algorithm with $v_{min} = 0$ and $v_{max} = 2.5$ for all your DICOM images. These normalized images are the used for the following analyses.

b. Image resizing: You can resize your images to a smaller size (e.g., 512×512), which can save memory and accelerate the training of a deep neural network but probably reduce your prediction accuracy. You are not required to do image resizing; please make your own decision.

c. Use **deep learning models** (for example: faster R-CNN, mask R-CNN, RetinaNet, YOLO, and so on) **with transfer learning** for chest disease object detection.

4. For each deep learning model, load test chest X-ray images into your prediction model to generate their predicted **location bounding boxes** and **corresponding disease categories**.

5. Collect all predicted results from these deep learning models, and upload the result from the one of the models you choose.

   Please also upload the Python codes of your solution (saved as a .ipynb file) and its compiled html file.

6. 作業評分標準與繳交格式：

Code (50%)

- 不限制使用的模型
- 需繳交三個檔案，分別為：
  - 學號.ipynb (5%)
  - 學號.html (5%)
  - 學號.csv (40%)

繳交格式錯誤一律以 0 分計算，請同學多加留意。

- 學號.csv 以 **IoU = 0.5 下的 mAP** 進行評分，評分方式為：
  - mAP >= strong baseline (0.2)：40 分
  - simple baseline (0.15) <= mAP < strong baseline (0.2)：25 分
  - mAP < simple baseline (0.15)：10 分

  其中 simple baseline 和 strong baseline 可能視同學們最後的繳交結果進行下調。
- 學號.csv 格式及注意事項：
  - 依序包含 7 個 column，分別為：
    - ID：病患的 ID
    - category：該 bounding box 的所屬類別的英文名稱
    - score：該 bounding box 的 confidence score
    - xmin：該 bounding box 左上角的 x 座標除以 W
    - xmax：該 bounding box 右下角的 x 座標除以 W
    - ymin：該 bounding box 左上角的 y 座標除以 H
    - ymax：該 bounding box 右下角的 y 座標除以 H

    其中 W 為放入模型的影像寬度，H 為放入模型的影像高度。
  - ID 順序請和 test.csv 中的 ID 一欄相同
  - 若該 ID 所對應的影像預測出的類別為 normal，則不須寫入檔案中
  - 一個 row 只能包含一個 bounding box 的相關資訊，若一張影像預測出多個 bounding box，請寫入不同的 row 中
  - 檔案範例如下：

	A	B	C	D	E	F	G
1	ID	category	score	xmin	xmax	ymin	ymax
2	TDR02_20161123_145314	cardiac_hypertrophy	0.41471	0.1391	0.7434	0.3489	0.89534
3	TDR02_20161123_145314	scoliosis	0.11895	0.34866	0.4583	0.2535	0.3247
4	TDR01_20171106_111727	lung_field_infiltration	0.33468	0.3469	0.8355	0.1112	0.67324
5	TDR01_20180510_090210	lung_field_infiltration	0.93562	0.27905	0.4896	0.5861	0.9345

Report (50%)

這份作業 report 將以「計劃書面報告」的形式繳交，內容詳述分析的：問題(目的為何？想預測或估計什麼？重要性？)、資料來源(公用資料庫下載、實驗室資料或自行搜集？)與背景、分析流程(分析方法、模型、視覺化手法(如：Saliency maps, Grad-CAM)、…)、結果(新發現、與觀眾溝通、視覺化)。書面報告的格式，請參閱所附檔案：“report_example_1.pdf”,

“report_example_2.pdf”, “report_example_3.pdf”。