Final Project：Disease Prediction

Requirement：

Choose one of the following tasks to complete. Only one !　Design a model and complete the experiment. Finally, two files need to be submitted:

  Experiment Report (Use the report template)

  Code File (zip package)

(These two documents must be named after their student ID when submitting, for example：*Report+ 20183290375.docx ,　Code+ 20183290375.zip*)

Project Receiving Mailbox：shuchao_shanty@126.com

Submission deadline: End of the semester(The specific time will be notified later)

Invasive Ductal Carcinoma (IDC) Classification

Content: According to IDC Histology Image DataSet to predict whether IDC is positive or negative. The dataset download address ↓

link1：https://drive.google.com/file/d/1l3F9W_csc8ZcJxnu7F4vA9nz_SnpRnm5/view?usp=sharing

link2：https://www.kaggle.com/paultimothymooney/breast-histopathology-images

Dataset Description

The original dataset consisted of 162 whole mount slide images of Breast Cancer (BCa) specimens scanned at 40x. From that, 277,524 patches of size 50 x 50 were extracted (198,738 IDC negative and 78,786 IDC positive).

Each patch's file name is of the format:

u_xX_yY_classC.png — > example 10253_idx5_x1351_y1101_class0.png

Where **u** is the patient ID (10253_idx5), **X** is the x-coordinate of where this patch was cropped from, **Y** is the y-coordinate of where this patch was cropped from, and **C** indicates the class where 0 is non-IDC and 1 is IDC.

The data and traning/test set partitions are located here (1.6G).
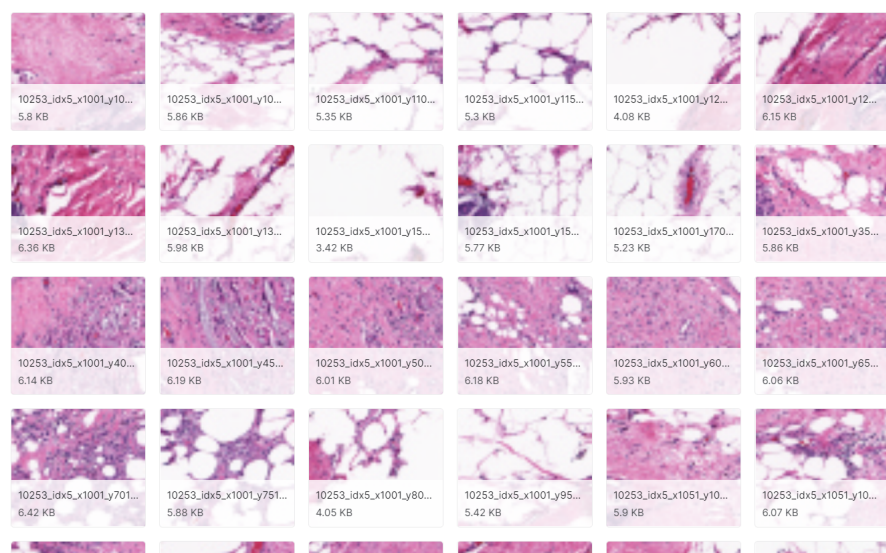
Data Explorer
1.49 GB

- 10253
  - 0
  - 1
- 10254
- 10255
- 10256
- 10257
- 10258
- 10259
- 10260
- 10261
- 10262
- 10264
- 10268
- 10269
- 10272
- 10273
- 10274
- 10275
- 10276
- 10277
- 10278
- 10279
- 10282
- 10285

Summary
- 278k files

< 0 (479 files)

10253_idx5_x1001_y10... 5.8 KB	10253_idx5_x1001_y10... 5.86 KB	10253_idx5_x1001_y110... 5.35 KB	10253_idx5_x1001_y115... 5.3 KB	10253_idx5_x1001_y12... 4.08 KB	10253_idx5_x1001_y12... 6.15 KB
10253_idx5_x1001_y13... 6.36 KB	10253_idx5_x1001_y13... 5.98 KB	10253_idx5_x1001_y15... 3.42 KB	10253_idx5_x1001_y15... 5.77 KB	10253_idx5_x1001_y170... 5.23 KB	10253_idx5_x1001_y35... 5.86 KB
10253_idx5_x1001_y40... 6.14 KB	10253_idx5_x1001_y45... 6.19 KB	10253_idx5_x1001_y50... 6.01 KB	10253_idx5_x1001_y55... 6.18 KB	10253_idx5_x1001_y60... 5.93 KB	10253_idx5_x1001_y65... 6.06 KB
10253_idx5_x1001_y701... 6.42 KB	10253_idx5_x1001_y751... 5.88 KB	10253_idx5_x1001_y80... 4.05 KB	10253_idx5_x1001_y95... 5.42 KB	10253_idx5_x1051_y10... 5.9 KB	10253_idx5_x1051_y10... 6.07 KB

Classification model detect pneumonia

Content: According to Chest X-Ray Images (Pneumonia) DataSet to predict whether lungs is normal or pneumonia. The dataset download address ↓

link3： https://drive.google.com/file/d/17vv7KGOIpzUJ70VgqV3DeHJiIIU1DOL5/view?usp=sharing

link4： https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

Dataset Description

The dataset is organized into 3 folders (train, test, val) and contains subfolders for each image category (Pneumonia/Normal). There are 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal).

Chest X-ray images (anterior-posterior) were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou. All chest X-ray imaging was performed as part of patients' routine clinical care.

For the analysis of chest x-ray images, all chest radiographs were initially screened for quality control by removing all low quality or unreadable scans. The diagnoses for the images were then graded by two expert physicians before being cleared for training the AI system. In order to account for any grading errors, the evaluation set was also checked by a third expert.

Data Explorer

1.15 GB

- ▼ 🗀 chest_xray
  - ▼ 🗀 test
    - ▶ 🗀 NORMAL
    - ▶ 🗀 PNEUMONIA
  - ▼ 🗀 train
    - ▶ 🗀 NORMAL
    - ▶ 🗀 PNEUMONIA
  - ▼ 🗀 val
    - ▶ 🗀 NORMAL
    - ▶ 🗀 PNEUMONIA

Summary

- ▶ 🗀 5856 files

< **NORMAL** (1341 files)

IM-0128-0001.jpeg 391.26 KB	IM-0129-0001.jpeg 410.42 KB	IM-0131-0001.jpeg 447.35 KB	IM-0133-0001.jpeg 375.91 KB	IM-0135-0001.jpeg 706.94 KB	IM-0137-0001.jpeg 401.86 KB
IM-0140-0001.jpeg 488.89 KB	IM-0141-0001.jpeg 646.23 KB	IM-0143-0001.jpeg 1.01 MB	IM-0145-0001.jpeg 493.92 KB	IM-0147-0001.jpeg 1.01 MB	IM-0149-0001.jpeg 432.37 KB
IM-0151-0001.jpeg 414.04 KB	IM-0152-0001.jpeg 469.94 KB	IM-0154-0001.jpeg 579.19 KB	IM-0156-0001.jpeg 478.01 KB	IM-0158-0001.jpeg 209.28 KB	IM-0160-0001.jpeg 586.51 KB
IM-0162-0001.jpeg 666.63 KB	IM-0164-0001.jpeg 253.45 KB	IM-0166-0001.jpeg 831.78 KB	IM-0168-0001.jpeg 469.34 KB	IM-0170-0001.jpeg 824.52 KB	IM-0172-0001.jpeg 754.27 KB

This preview shows 30 out of 1341 items. Load more