Project - Medical Report Generation using X-Ray Images

Shanun Randev (G47103805)

Natural Language Processing

Professor - Dr. Amir Jafari

INTRODUCTION

In modern healthcare, radiology plays a crucial role in diagnosing and managing a wide array of medical conditions. Chest X-rays, in particular, are among the most frequently used diagnostic tools for detecting abnormalities such as Pneumonia, Hernia and Cardiomegaly. The motivation behind this project is to create a tool that supports radiologists in their demanding roles by automating the generation of preliminary radiology reports. By leveraging advanced computer vision and large language models, the system aims to draft accurate and detailed reports from chest X-ray images, serving as a starting point for radiologists to review and refine.

The individual report begins with an Introduction that outlines the motivation and scope of the project, followed by a detailed description of the dataset, highlighting its structure and features. The Methodology section explains the technical approach, including preprocessing steps, models used, and training processes. My Contributions section explains my contributions to the project.

DESCRIPTION OF THE DATASET The MIMIC-CXR is a publicly available chest X-ray dataset for chest radiography research.It comprises 15,000 chest X-ray images in Dicom format and their associated radiology reports in XML format. The dataset has the following key features: ● Image File Path: Location or link to the corresponding chest X-ray image. ● Findings: A textual description of abnormalities or observations made by the radiologist. ● Impression: A concise summary of the radiologist's primary conclusions. The dataset has 14 labels corresponding to common chest X-ray pathologies. The pathology labels include Atelectasis, Cardiomegaly, Consolidation, Edema, Enlarged Cardiomediastinum, Fracture, Lung Lesion, Lung Opacity, Pleural Effusion, Pleural Other, Pneumonia, Pneumothorax, Support Devices, and No Finding.

METHODOLOGY

1. **Data Collection and Preprocessing:** Initial steps involved converting DICOM images to PNG format, extracting relevant information, preparing datasets for further analysis and performing text and image transformations.
2. **Label Extraction**: The next stage utilized ChexBERT to generate multi-label classifications from the associated radiology reports, enabling the identification of relevant medical conditions in each image.
3. **Model Architecture Design**: During this stage, different combinations of image encoders, alignment models, and language models were experimented with to determine the most effective architecture for processing medical image data and generating accurate reports. 4.
4. **Training and Evaluation**: Once the model architecture was defined, the models were trained using the preprocessed data. Performance was closely monitored and evaluated using key metrics such as ROUGE-L to assess the quality of the model's output.

Contributions to the Project - I led the development the dataset,pytroch dataset class, data cleaning and processing of images.

-> Extracted the FINDINGS and IMPRESSIONS from .XML report files and mapped them with the corresponding image_id

	dicom_path	png_path	dicom_id	findings	impressions
0	/home/ubuntu/nlp_project/Code/physionet.org/fi...	/home/ubuntu/nlp_project/Code/physionet.org/fi...	e084de3b-be89b11e-20fe3f9f-9c8d8dfe-4cfd202c.dcm	The cardiac, mediastinal and hilar contours ar...	No acute cardiopulmonary abnormality.
1	/home/ubuntu/nlp_project/Code/physionet.org/fi...	/home/ubuntu/nlp_project/Code/physionet.org/fi...	2a2277a9-b0ded155-c0de8eb9-c124d10e-82c5caab.dcm	The cardiac, mediastinal and hilar contours ar...	No acute cardiopulmonary abnormality.
2	/home/ubuntu/nlp_project/Code/physionet.org/fi...	/home/ubuntu/nlp_project/Code/physionet.org/fi...	ea030e7a-2e3b1346-bc518786-7a8fd698-f673b44c.dcm	The lungs are clear of focal consolidation, pl...	No acute cardiopulmonary process.
3	/home/ubuntu/nlp_project/Code/physionet.org/fi...	/home/ubuntu/nlp_project/Code/physionet.org/fi...	174413ec-4ec4c1f7-34ea26b7-c5f994f8-79ef1962.dcm	There is no focal consolidation, pleural effus...	No acute cardiopulmonary process.
4	/home/ubuntu/nlp_project/Code/physionet.org/fi...	/home/ubuntu/nlp_project/Code/physionet.org/fi...	02aa804e-bde0afdd-112c0b34-7bc16630-4e384014.dcm	There is no focal consolidation, pleural effus...	No acute cardiopulmonary process.

-> Wrote a Bash script to download images from a public portal,and thereby increasing the dataset size for training.

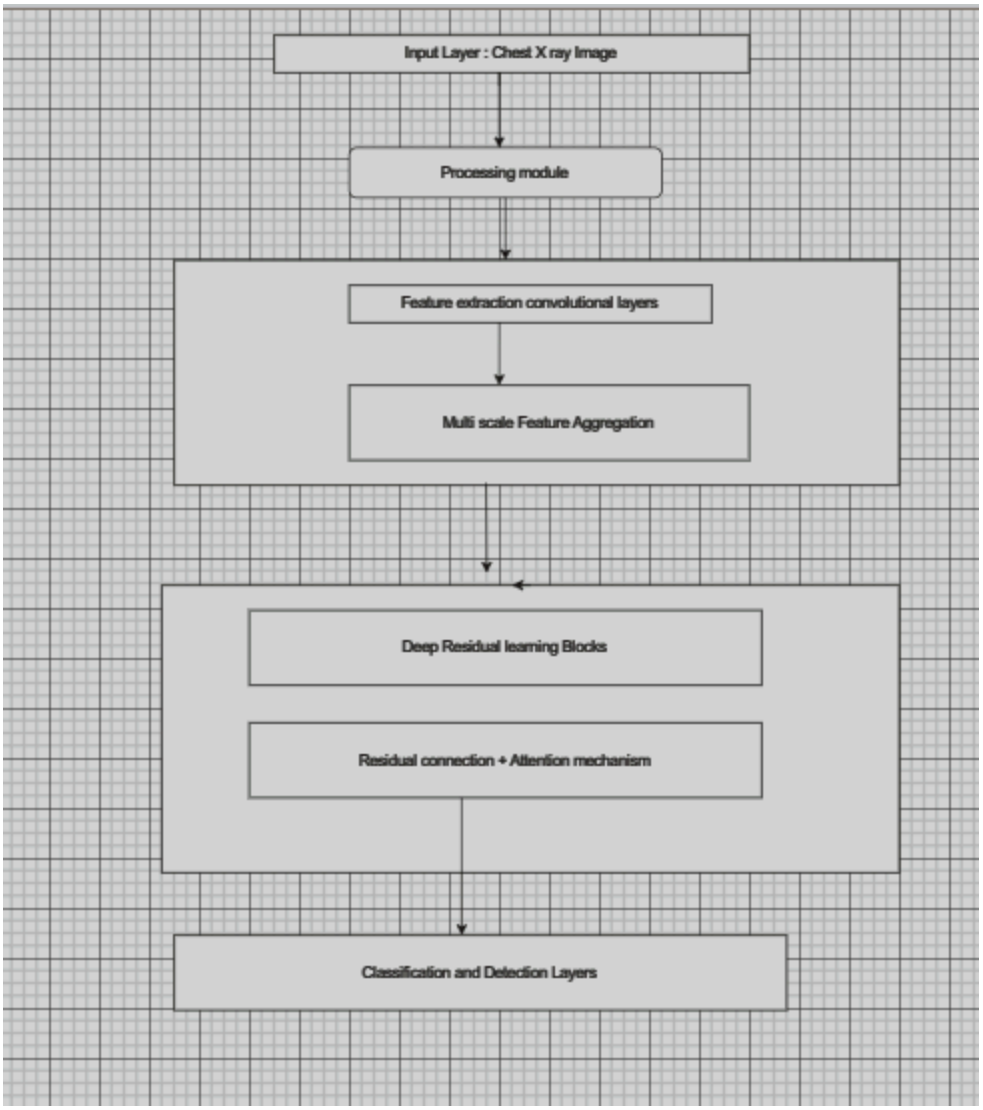
```
# Base URL
base_url="https://physionet.org/files/mimic-cxr/2.1.0/files/p15"

# Credentials
username="saniyas28"
password="Saniya@280398"

# Loop through each folder and download
for folder in "${folders[@]}"; do
  echo "Downloading folder: $folder"
  wget -r -N -c -np --user="$username" --password="$password" "$base_url/$folder/"
  if [ $? -eq 0 ]; then
    echo "Successfully downloaded $folder"
  else
    echo "Failed to download $folder"
  fi
done

echo "Download completed."
```

->Developed the logic to convert dcm to png format and concatenate them with the old 7k png images increasing the dataset size to around 15k.

```
    # Function to convert DICOM to PNG
    def convert_dicom_to_png(dicom_path, output_path):
        try:
            # Read the DICOM file
            dicom = pydicom.dcmread(dicom_path)
            # Get pixel array
            pixel_array = dicom.pixel_array
            # Normalize pixel values to 0-255
            pixel_array = ((pixel_array - pixel_array.min()) / (pixel_array.max() - pixel_array.min()) * 255).astype(
                np.uint8)
            # Save as PNG
            image = Image.fromarray(pixel_array)
            image.save(output_path)
        except Exception as e:
            print(f"Error converting {dicom_path} to PNG: {e}")
```

-> Led the development of inference code for BioMed Clip model and CXR mate model, and testing of BLIP2 model on evaluation sets.

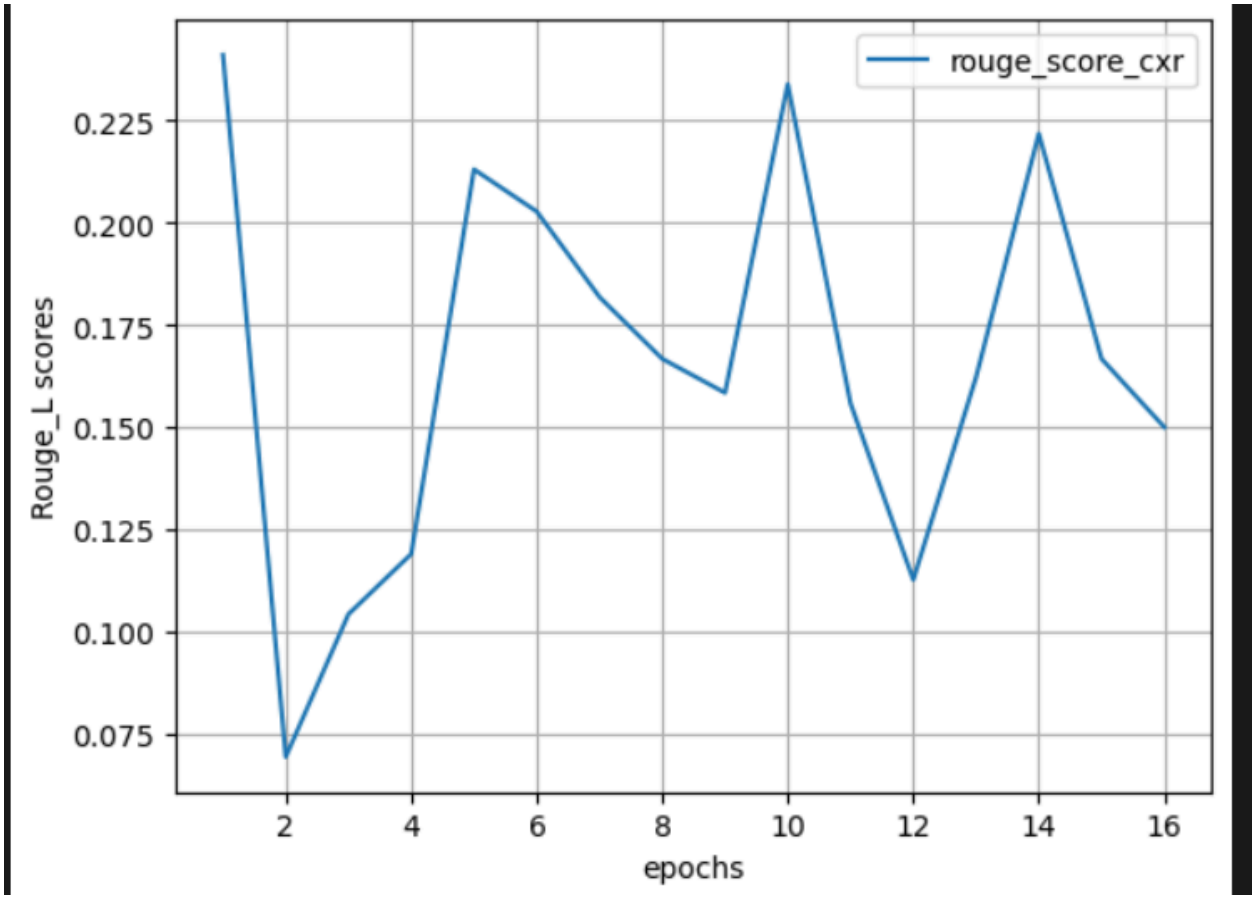
CXR MATE model



1. n the CXR pipeline, filtering of images is done based on valid medical findings.Then we apply some image augmentations and transformations.

2. The module supports 14 condition labels which we predict using the ChexNet model. In the training loop we load CXR Mate model and tokenizer and the medical report generator.The training loop implements image feature extraction, report generation and checkpoint saving for inference

3. Rouge L score analysis for cxr mate



-> Built a custom densenet121 for feature extraction which formed the base for the ChexNet module used for extracting labels from images.

- Implemented a feature extractor that uses DenseNet-121's convolutional layers up to its second last layer.
- Applies **average pooling** to the features to generate compact feature representations.
- Key Components :

  get_model() :

  1. Loads DenseNet-121 pre-trained on ImageNet.
  2. Extracts convolutional layers and the classifier's input dimensions
  3. Returns the convolutional layers as a sequential model and an average pooling layer.

  forward() :

  1. Takes an input image tensor.Passes it through the model and applies avgpooling

2. Load_and_preprocess_image - Loads an image from the specified path using PIL. Preprocessing of the image :

   ->Resizes to 224x224
   ->Converts to a tensor and normalizes it (ImageNet mean and std).
   ->Adds a batch dimension for model compatibility.