# CODE DESCRIPTION

## Medical Visual Question Answering (MVQA)

Medical Visual Question Answering (MVQA) is a multidisciplinary field combining computer vision and natural language processing to interpret visual data and respond to related questions in natural language. The MVQA focuses on analyzing medical images (e.g., X-rays, CT scans, MRI scans) and providing relevant answers to queries posed by medical practitioners or researchers. This capability aims to assist in diagnosis, decision-making, and streamlining workflows in clinical environments.

## Fetch required inputs from the datasets

The MVQA dataset is divided into three subsets: training, validation, and test. Each subset consists of medical images paired with corresponding Question-Answer Pairs (QA-Pairs). The preparation of these inputs and computation of their required values is facilitated by the ***data_utils.py*** script.

Two key issues arise within the QA-Pairs and images:

1. **Imbalanced Labels**: Certain labels have disproportionately large or small sample sizes.
2. **Insufficient Samples**: Some answers have very few associated samples, affecting representativeness.

These challenges are addressed through two processes:

1. **Improvement of QA-Pairs**: Balancing and augmenting QA-Pair labels to ensure better distribution.
2. **Improvement of Medical Images**: Enhancing the dataset by increasing the sample representativeness of underrepresented labels.

Finally, the improved QA-Pairs and medical images are concatenated based on their image_id, resulting in an enhanced and balanced dataset ready for model training and evaluation.

## Improvement of QA-Pairs

This module addresses the two primary issues in the dataset through the following steps:

**Resolving the First Issue (Imbalanced Labels)**

1. **Computation of Question Counts**: Calculate the number of questions associated with each label.
2. **Hard-Sample Threshold Calculation**: Determine a threshold value to identify underrepresented labels (hard samples).
3. **Selection of an Average Value**: Choose an appropriate average value to guide balancing efforts.

**Resolving the Second Issue (Insufficient Samples)**

New samples are generated for underrepresented labels based on the computed hard-sample threshold, ensuring a more balanced distribution across the dataset.

These steps collectively improve the quality and representativeness of the QA-Pairs and associated labels, contributing to a more robust dataset for model training and evaluation.

# Improvement of Medical Images

This module addresses the dataset's key issues through the following approaches:

## Resolving the First Issue (Imbalanced Labels)

1. **Image Count per Label**: Calculate the number of images associated with each label to identify imbalances.
2. **Hard-Sample Threshold**: Determine a threshold value to identify underrepresented labels.
3. **Average Value Selection**: Select an appropriate average value to guide the balancing process.

## Resolving the Second Issue (Insufficient Samples)

The second issue is addressed using:

1. **Mixup**: A data augmentation technique that generates new samples by combining pairs of images and their labels, enhancing diversity.
2. **Label Smoothing**: Regularizing the labels to reduce overconfidence in the model and improve generalization.

These methods collectively enhance the dataset's balance and quality, ensuring a more effective training process.

# MVQA Model Creation

From the improved dataset, features are extracted and combined to build the model.

1. **Feature Extraction**
    - **Image Features**: Medical images are preprocessed, and features are extracted using VGGNet. These features are stored in separate files for training, validation, and testing: ***train_features.pkl, valid_features.pkl,*** and ***test_features.pkl.***
    - **Text Features**: Word embeddings for QA-pairs are generated using GloVe vectors (stored in the "Glove" folder). Using these embeddings, LSTM generates a sequence of words for the answers based on the timestamps.
2. **Feature Combination**

    The extracted image features and text features are concatenated using LSTM to form a unified representation.

3. **Model Generation and Prediction**

A model is trained using the concatenated feature vectors. This trained model is then used to predict answers for the test set.

4. **Validation**
   The predicted answers are evaluated using quantitative metrics to measure the model's performance.

## Layerwise Relevance Propagation eXplainable Artificial Intelligence (LRP XAI)

The reasoning behind the generated answers is visualized using Layer-wise Relevance Propagation (LRP) Explainable AI (XAI). The process involves:

1. **Heatmap Generation**
   - Gradients are computed between the most contributing layer in the model and the final output layer.
   - These gradients are used to create a heatmap that highlights the regions in the input contributing most to the model's decision.
2. **Superimposed Images**
   - The heatmap is overlaid on the input image to visualize significant regions, resulting in superimposed images.
   - Examples of these visualizations for samples from the Path-VQA dataset under various scenarios are shown in ***Figs. 1 and 2.***

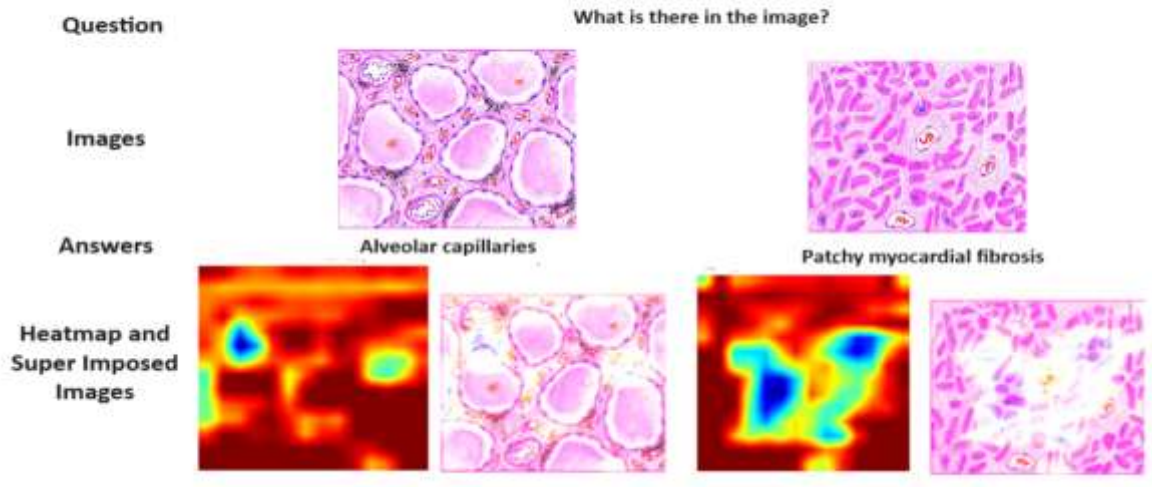


**Fig 1. Visualization of heatmaps and super imposed images for different samples**

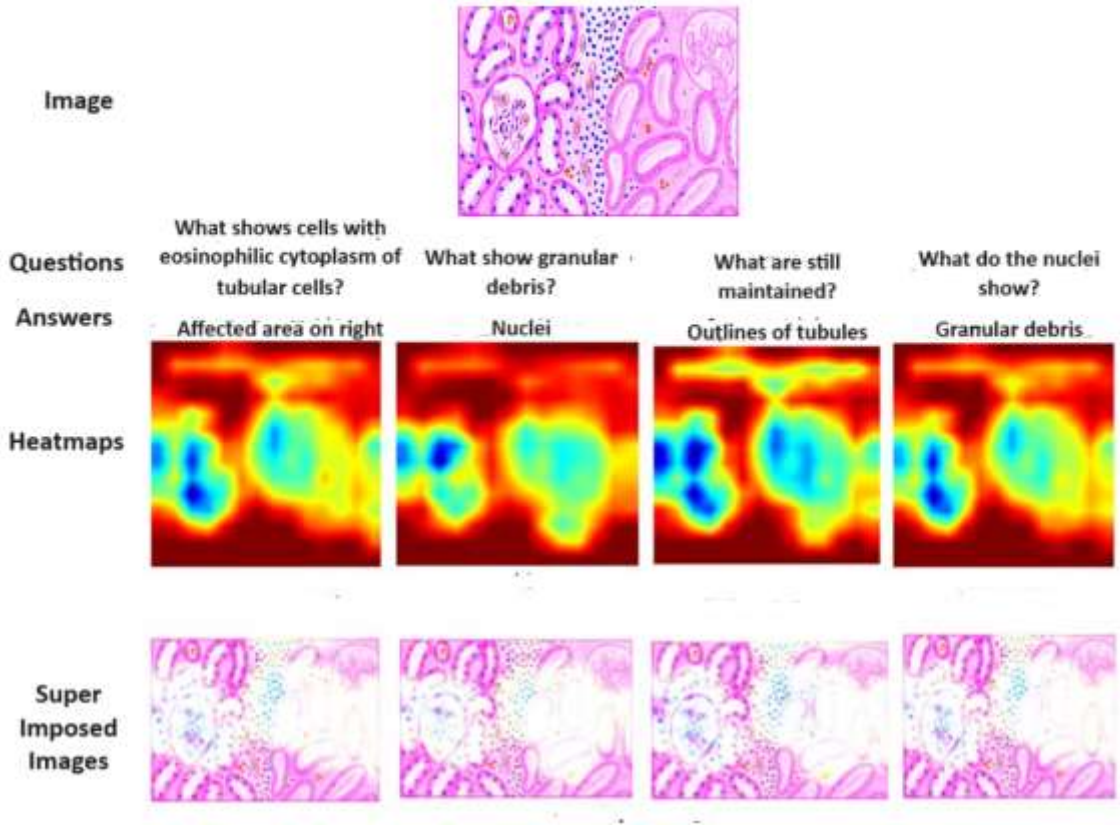


**Fig 2. Visualization of heatmaps and super imposed images for four different QA pairs**

3. **Effect of Augmentation and Visualization**

    - The correctly and incorrectly classified samples from five MVQA datasets are analyzed.
    - ***Figs. 3 to 7*** showcase the impact of data augmentation and visualization techniques on answer prediction accuracy.

This visualization approach provides insights into the model's decision-making process, enhancing interpretability and understanding of its predictions.

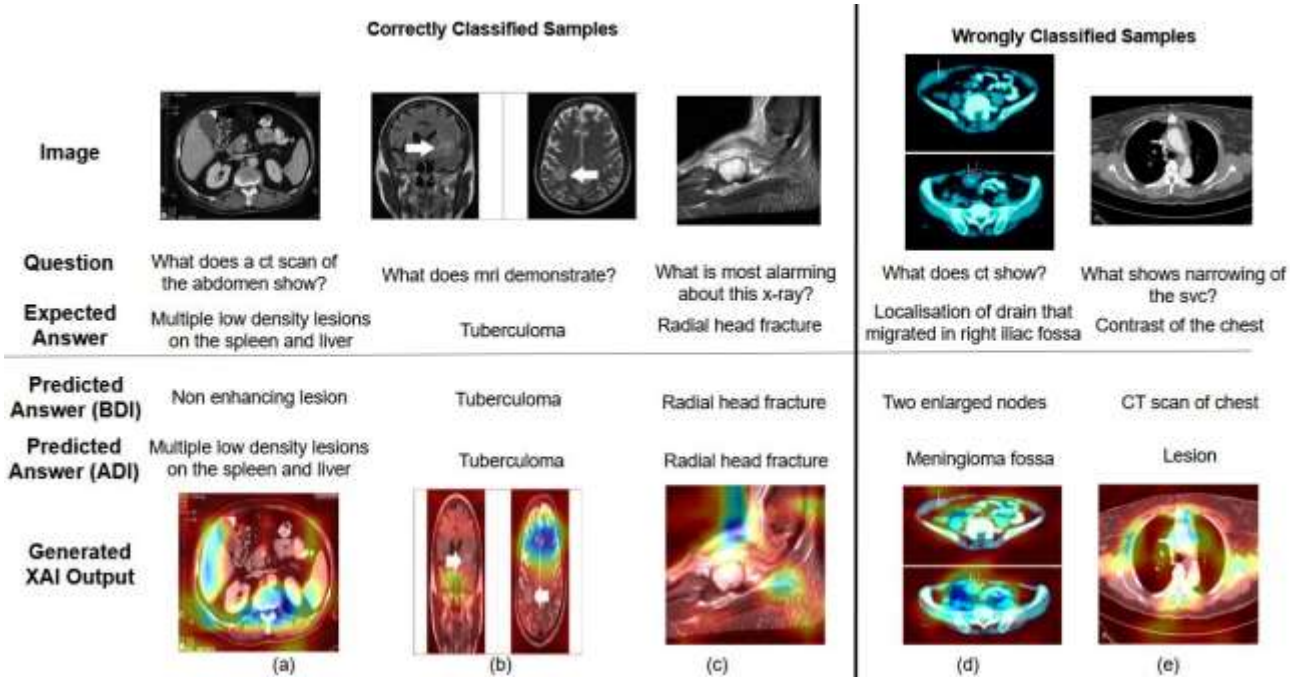


**Fig 3. Visualization of samples from VQA-MED 2018 dataset w.r.t predicted answer**

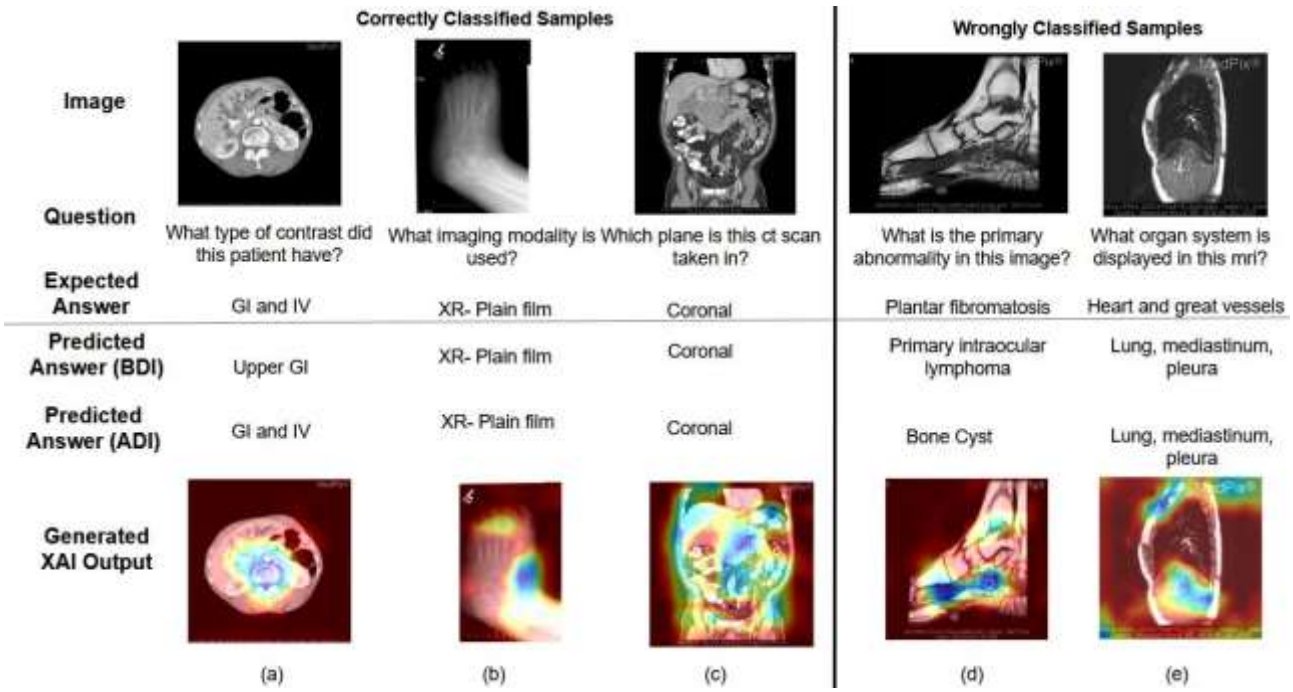


**Fig 4. Visualization of samples from VQA-MED 2019 dataset w.r.t predicted answer**

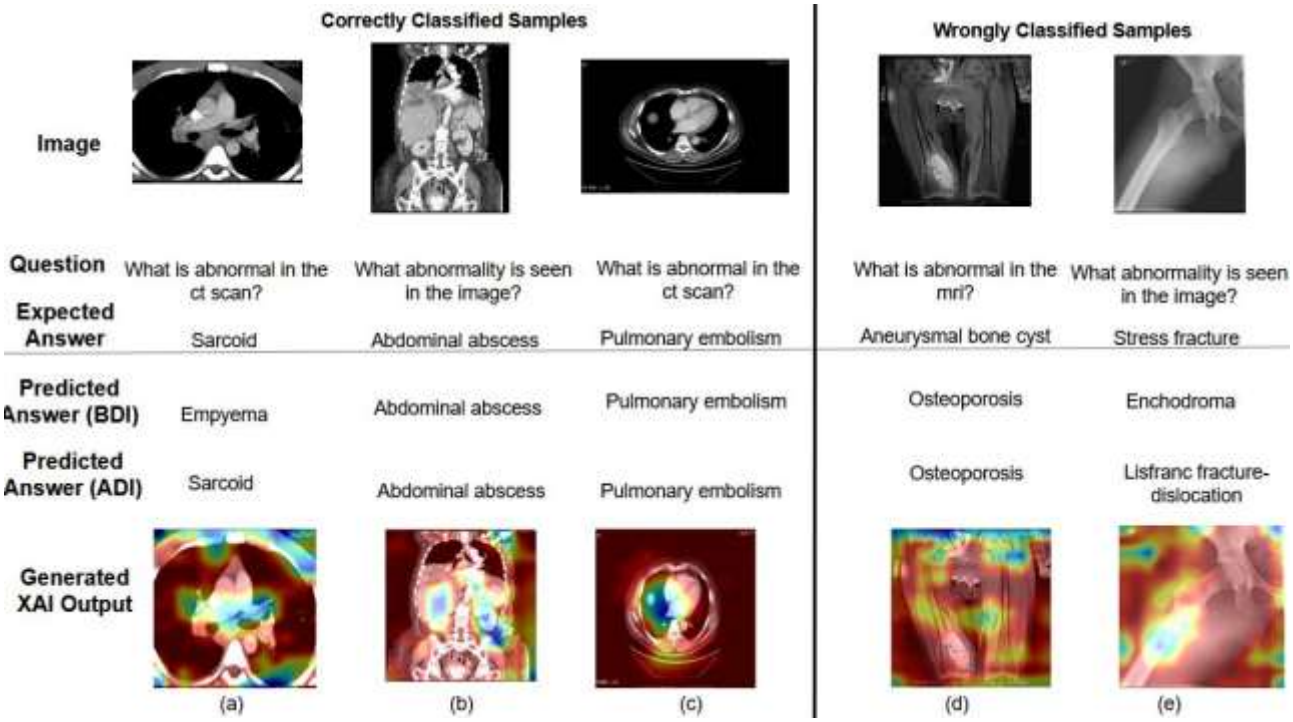


**Fig 5. Visualization of samples from VQA-MED 2020 dataset w.r.t predicted answer**

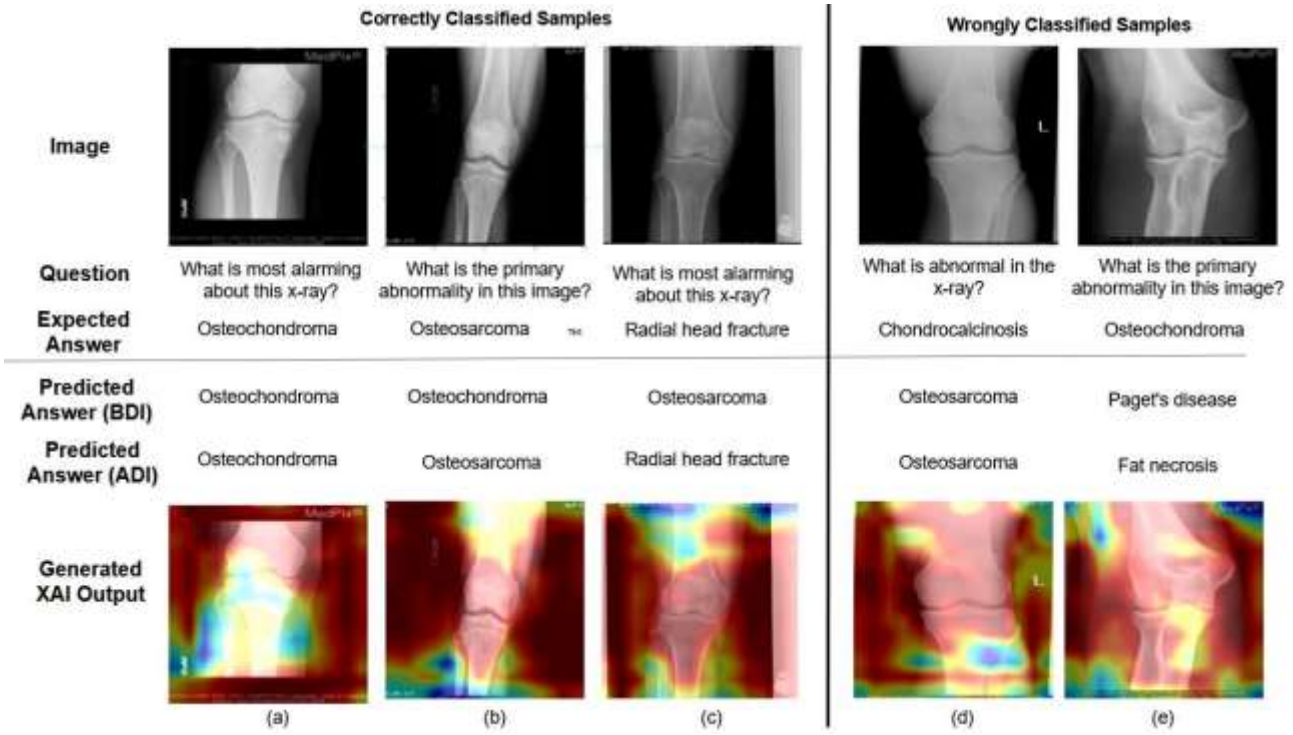


**Fig 6. Visualization of samples from VQA-MED 2021 dataset w.r.t predicted answer**

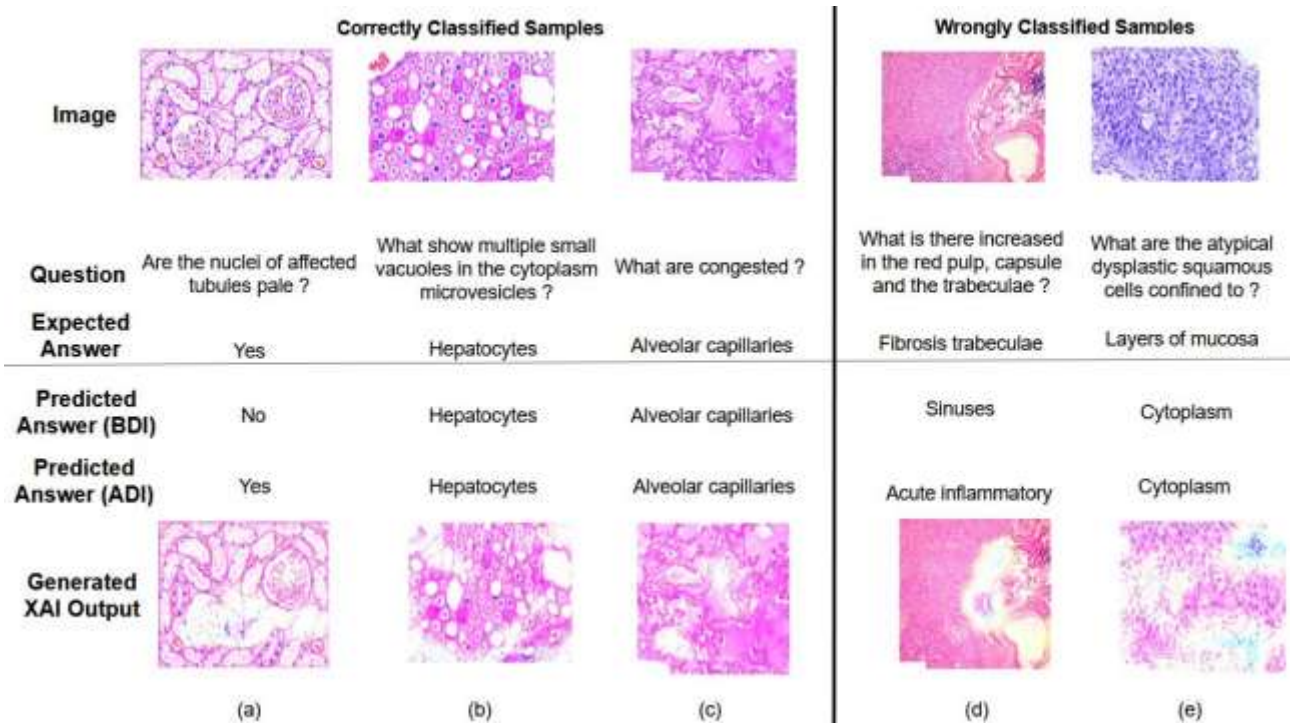


**Fig 7. Visualization of samples from Path-VQA dataset w.r.t predicted answer**