**Visual Questioning Answering for Medical Field**

**ICRL-1005-22**

# **Introduction**

The aim of this project is to detect the disease and symptoms related human beings after taking image where someone feel some problem. After getting image, this system let us know about the disease in textual form.

# **Dataset (VQA-Med-2018)**

The QA pairs were generated from captions by a semi-automatic approach. First, a rule-based question generation (QG) system4 automatically generated possible QA pairs by sentence simplification, answer phrase identification, question generation, and candidate questions ranking. Then, two expert human annotators (including one expert in clinical medicine) manually checked all generated QA pairs in two passes. Respectively, one pass ensures semantic correctness, and another ensures the clinical relevance to associated medical images.

# **Preprocessing**

The preprocessing steps of the proposed project are following:

There was no preprocessing of dataset was involved in this stage.

# **Model Training**

* We have used EfficientNet for images. EfficientNet is a convolutional neural network architecture and scaling method that uniformly scales all dimensions of depth/width/resolution using a compound coefficient. BERT (Bidirectional Encoder Representations from Transformers) for textual Q & A..

# **Results**

* Got 90% accuracy value for validation samples

# **AUTHOR CONTRIBUTIONS**

Hamza Mukhtar conceived the original idea. Adil Ansari developed the theory and performed the computations. Hamza Mukhtar verified the analytical methods. Muhammad Adil optimize the model for deployment. All authors discussed the results and contributed to the final report.