ISSN: 2320-5407

ISSN NO. 2320-5407

International Journal of Advanced Research

Publisher's Name: Jana Publication and Research LLP

www.journalijar.com

REVIEWER'S REPORT

Manuscript No.: IJAR- 52851 Date: 18-07-2025

Title:

"Transformer Decoder for Chest X-ray Image Captioning Using Deep Feature Extraction"

Recommendation:
Accept as it is ..... **YES** .....
Accept after minor revision**...**
Accept after major revision .....
Do not accept (*Reasons below*) .....

Rating	Excel.	Good	Fair	Poor
Originality			YES	
Techn. Quality			YES	
Clarity			YES	
Significance			YES	

Reviewer Name: Gulnawaz Gani

Comments for Publication

This paper contributes a novel hybrid deep learning architecture combining Vision Transformer and CheXNet with a Transformer decoder for automated chest X-ray image captioning, aiming to assist in medical diagnosis.

Reviewer's Comment / Report

- This paper proposes a hybrid Vision Transformer (ViT) and CheXNet approach with a Transformer decoder for generating medical reports from chest X-ray images.
- While the integration of these models for image-to-text generation is innovative, the paper could benefit from a more detailed quantitative analysis of its performance against a wider range of state-of-the-art models beyond just accuracy, such as clinical metrics like recall and precision for specific pathologies.
- The dataset description, while mentioning public sources, lacks specific details on the distribution of different disease classes, which is crucial for evaluating model robustness.
- Furthermore, the discussion on limitations is brief; a deeper exploration of potential biases in the dataset or challenges in handling rare disease cases would strengthen the review.
- Finally, while attention maps are shown, a more thorough qualitative analysis of cases where the model's attention aligns or misaligns with clinical findings would provide further insights.