

REVIEWER'S REPORT

Manuscript No.: **IJAR- 52851**

Date: 17-07-2025

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

Recommendation:

Accept as it is

Accept after minor revision✓.....

Accept after major revision.....

Do not accept (*Reasons below*)

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

Reviewer Name: **Sudhanshu Sekhar Tripathy**

Date: 17-07-2025

Reviewer's Comment for Publication.

(To be published with the manuscript in the journal)

The reviewer is requested to provide a brief comment (3-4 lines) highlighting the significance, strengths, or key insights of the manuscript. This comment will be Displayed in the journal publication alongside with the reviewer's name.

Reviewer's Comment for Publication

The manuscript presents a well-structured and timely study on the application of a Transformer Decoder architecture for automated chest X-ray image captioning, leveraging deep feature extraction from Vision Transformer (ViT) and CheXNet models. The proposed hybrid approach demonstrates a thoughtful integration of visual and clinical feature representations, aiming to support radiologists with consistent, automated reporting, particularly in post-COVID diagnostic scenarios.

The authors have clearly described the dataset, model architecture, and evaluation strategy. The use of BLEU score and comparison with baseline models further adds to the validity of the study. The work contributes meaningfully to the growing field of medical image captioning and aligns with current research trends in deep learning for healthcare.

That said, a **few minor formatting and presentation issues should be addressed**. These include section numbering, consistent caption placement for tables and figures, citation formatting, and improved textual flow in dense sections. Additionally, a visual summary (workflow diagram or architecture block) would enhance clarity for readers unfamiliar with the full model pipeline.

REVIEWER'S REPORT

Once these minor revisions are implemented, the manuscript will be suitable for publication and will serve as a valuable resource for researchers and practitioners working in medical imaging, NLP, and AI-driven diagnostic systems.

Detailed Reviewer's Report

Recommendation: **Accept after minor revision.**

Comments & Suggestions for Improvement

1. Scope & Relevance:

- The manuscript addresses a highly relevant and timely problem in medical imaging: automated captioning of chest X-ray images using deep learning.
- The integration of ViT (Vision Transformer) and CheXNet for feature extraction, followed by a Transformer Decoder for report generation, demonstrates solid technical depth.
- The focus on post-COVID diagnostic automation adds real-world importance and aligns with clinical AI advancements.
- The study contributes to improving diagnostic speed, consistency, and support in radiology, making it both practically and academically valuable.

2. Structure & Technical Presentation:

- The manuscript contains all essential sections (Abstract, Introduction, Related Work, Methodology, Results, Conclusion), but section numbering is missing and should be consistently applied.

The manuscript currently lacks consistent section numbering, which can hinder navigation and reduce professional readability. For clarity and academic structure, each major section and its sub-sections should be numbered systematically.

Let say an Examples for Correct Numbering Style:

- **1. Introduction**
- **2. Related Work**
- **3. Proposed Methodology**
 - 3.1 Dataset Description
 - 3.2 Feature Extraction (ViT, CheXNet)
 - 3.3 Transformer Decoder Architecture

REVIEWER'S REPORT

- 4. Experimental Results
 - 4.1 Evaluation Metrics
 - 4.2 Comparative Analysis
- 5. Conclusion and Future Work

This numbering structure allows readers to follow the flow more easily and makes cross-referencing between sections simpler (e.g., “As discussed in Section 3.2...”).

Please revise the manuscript to include consistent numbering like the above throughout all sections and sub-sections.

- All table captions are missing or incorrectly placed (Table 1, Table 2, Table 3 and Table 4). Follow academic standards: **figure captions below figures, table captions above tables.**
- While the figure captions are correctly placed below the figures as per standard academic formatting, it is noted that the table (1,2,3,4) captions are incorrectly positioned below the tables.
- According to widely accepted formatting guidelines, **table captions must be placed above the table**, clearly describing its content before the reader views the data.

Example correction:

[Correct] Table 1,2,3 and 4: Performance metrics of the proposed model (**placed above the table**)

Kindly revise the manuscript to ensure that all **table captions** follow this format for consistency and professionalism.

- Figures 2, 3, and 4 should be better referenced and explained in the text to improve interpretability.
- The experimental design and dataset sources are well explained, but visual workflow diagrams (e.g., for the overall model pipeline) would improve clarity.
- Consider reformatting dense paragraphs (**especially in Sections 3 and 4**) for readability and coherence.

REVIEWER'S REPORT

3. References & Citations

Upon reviewing the **References** section, it is noted that while approximately **20 citations are listed**, several of them are **not referenced anywhere in the main text**. This creates inconsistency between the **bibliography** and **in-text citations**, which may confuse readers and affect the academic integrity of the work.

To maintain clarity and adhere to academic standards, please ensure that:

- Every citation listed in the References section is **cited at least once** in the manuscript body.
- In-text citations should be clearly indicated using the appropriate format (e.g., [3], [5], [7]).
- If any references are not directly discussed or used in the text, they should be **removed from the reference list** to avoid redundancy.

4. Language & Style

- The manuscript is generally well-written with clear technical explanations. However, minor grammatical errors, informal expressions, and inconsistent phrasing are present. Improving sentence structure, maintaining academic tone, and refining transitions will enhance readability and the overall professional quality of the work.

Note: Kindly implement the above minor corrections to uphold the academic integrity and formatting standards expected by the journal. These improvements will enhance the manuscript's clarity, coherence, and overall presentation quality.