

REVIEWER'S REPORT

Manuscript No.: **IJAR-52461**

Date: 26/06/25

Title: Image Story: Enhanced Cognitive Visual Narrative

Recommendation:

Accept as it is

Accept after minor revision.....

Accept after major revision

Do not accept (*Reasons below*)

Reviewer Name: Dr Gulnawaz

The manuscript introduces the **Enhanced Cognitive Visual Narrative System (ECVNS)** — an advanced multimodal AI framework tailored for automated visual storytelling. This system effectively combines multiple cutting-edge deep learning architectures, including:

OWLv2 for object identification,

BLIP for generating captions and handling visual Q&A tasks,

CLIP for assessing emotional context, and

ViLT for scene-level comprehension.

The integration of these technologies positions ECVNS at the forefront of innovation in the field of multimodal artificial intelligence.

The study holds promise for a broad range of practical applications, notably in digital content development, inclusive technology, educational tools, and interactive media.

While the core framework is strong and the innovation noteworthy, the manuscript would benefit from a few refinements. Specifically, the clarity regarding sample sizes, accuracy of statistical reporting, and overall linguistic polish should be improved to enhance its readability and scholarly value.