International Journal of Multidisciplinary Engineering in Current Research - IJMEC

Volume 11, Issue 4s, April-2026, http://ijmec.com/, ISSN: 2456-4265

Full Length Article

A Semantic Weight Adaptive Model Based on Visual Question Answering

Mohmmed Adnaan Qureshi[1], Mohammed Fasahath Siddiq[2], Mirza Mohammad Ali Baig[3], Dr ijteba sultana[4]

[1,2,3]B.E Students; Department of CSE ISL Engineering College, Hyderabad, India

[4]Associate professor; Department of CSE ISL Engineering College, Hyderabad, India

Email: adnaanq96@gmail.com , mohammedfasahathsiddiq@gmail.com, mahekbaig57@gmail.com

Accepted 24-04-2026

ABSTRACT

This project presents a multilingual Visual Question Answering (VQA) web application developed using the Flask framework by integrating deep learning and Natural Language Processing (NLP) techniques. The system utilizes the transformer-based BLIP model, specifically the Salesforce/blip-vqa-base architecture, to answer user questions based on uploaded images and short-duration videos. The BLIP model combines a Vision Transformer (ViT) for extracting semantic visual features with a transformer-based text decoder for generating contextually accurate answers. Unlike traditional CNN-LSTM-based VQA systems, the proposed model performs joint multimodal learning, enabling improved understanding of both visual content and natural language queries.

A key feature of the proposed system is its multilingual capability, allowing users to interact with the application in various Indian languages such as Hindi, Telugu, Tamil, and Kannada. To support multilingual communication, a translation module is integrated that converts user questions into English before processing them through the VQA model and subsequently translates the generated answers back into the user's preferred language. Although the current implementation uses a simplified mock translation mechanism, the architecture is designed for future integration with advanced Neural Machine Translation (NMT) systems such as IndicTrans2.

The application also supports short video inputs by extracting keyframes from uploaded videos and generating context-aware responses based on visual analysis. Beam search decoding is employed during answer generation to produce coherent, grammatically meaningful, and high-probability responses. In addition, the system incorporates secure file upload validation and real-time processing within a scalable web environment.

The proposed multilingual VQA system demonstrates the effectiveness of transformer-based multimodal AI models in creating accessible, intelligent, and interactive applications. The system has potential applications in education, assistive technologies, healthcare support, surveillance, and smart human-computer interaction systems, while also providing a foundation for future advancements in multilingual and multimodal artificial intelligence research.

Keywords— *Flask, Multilingual Visual Question Answering (VQA), BLIP, Deep Learning, Natural Language Processing (NLP), Vision Transformer (ViT), Transformer Models, Video Question Answering, Image Understanding, Multilingual Translation, Web Application, Indian Languages, Real-Time AI Systems.*

INTRODUCTION

Visual Question Answering (VQA) is an emerging interdisciplinary research area that combines the fields of Computer Vision, Natural Language Processing (NLP), and Artificial Intelligence (AI) to enable machines to understand visual content and answer questions related to images or videos. The primary objective of a VQA system is to analyze a visual input and generate meaningful answers to natural language questions posed by users. Unlike traditional image classification systems that only identify predefined objects or categories, VQA systems require a deeper understanding of both visual and textual information, making them highly suitable for intelligent human-computer interaction applications.

Recent advancements in deep learning and transformer-based architectures have significantly improved the performance of VQA systems. Earlier approaches mainly relied on Convolutional Neural Networks (CNNs) for image feature extraction and Long Short-Term Memory (LSTM) networks for question processing. Although these methods achieved moderate success, they were limited in capturing complex semantic relationships between image regions and textual queries. Modern transformer-based models such as BLIP overcome these limitations by jointly learning visual and textual representations through attention mechanisms and multimodal pretraining.

The proposed project focuses on developing a multilingual Visual Question Answering web application capable of understanding and responding to user questions in multiple

https://doi.org/10.63665/IJMEC.1104s.46
ISSN: 2456-4265
IJMEC 2026

Indian languages such as Hindi, Telugu, Tamil, and Kannada. The application is developed using the Flask framework and integrates the BLIP transformer-based VQA model for intelligent answer generation. The system allows users to upload images or short-duration videos and ask questions related to the visual content. To support multilingual interaction, a translation module converts user questions into English before processing them through the VQA model and translates the generated answers back into the user's preferred language.

The proposed system also incorporates video processing capabilities by extracting keyframes from uploaded videos to generate context-aware responses. Beam search decoding is used during answer generation to produce coherent and contextually accurate outputs. This project demonstrates how transformer-based multimodal AI systems can be integrated into accessible web environments to create intelligent, scalable, and user-friendly applications.

The multilingual capability of the system makes it highly beneficial in diverse real-world applications such as education, assistive technologies for visually impaired individuals, healthcare support systems, smart surveillance, and interactive learning platforms. Furthermore, the modular architecture of the system enables future integration with advanced Neural Machine Translation (NMT) models such as IndicTrans2 to enhance translation accuracy and multilingual support.

LITERATURE SURVEY

Visual Question Answering has gained significant attention in recent years due to rapid developments in deep learning, computer vision, and natural language understanding. Researchers have proposed various techniques to improve the interaction between visual and textual modalities for accurate answer generation.

Early VQA systems were primarily based on the integration of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) networks. CNN models such as VGGNet and ResNet were widely used to extract visual features from images, while LSTMs processed sequential textual questions. These systems generally fused image and text features through concatenation or linear transformations before predicting answers. Although these methods established the foundation of VQA research, they suffered from limited multimodal reasoning capability and weak semantic alignment between images and questions.

To address these limitations, attention-based VQA models were introduced. Attention mechanisms enabled models to focus selectively on relevant regions of an image corresponding to important words in the question. Models such as Stacked Attention Networks (SAN) and Hierarchical Co-Attention Networks significantly improved answer accuracy by learning fine-grained relationships between visual and textual inputs. However, these models still relied heavily on CNN-LSTM architectures, which restricted their ability to model long-range dependencies and contextual semantics efficiently.

The emergence of transformer architectures revolutionized the field of VQA and multimodal learning. Transformer-based models use self-attention mechanisms that allow simultaneous processing of entire sequences, improving contextual understanding and parallelization efficiency. Vision Transformer (ViT) architectures further extended transformer capabilities to image processing by dividing images into smaller patches and learning global visual representations directly.

One of the major advancements in this domain is the development of the BLIP model. BLIP (Bootstrapping Language-Image Pretraining) is a multimodal transformer framework designed for vision-language understanding and generation tasks. The model jointly learns visual and textual features through large-scale pretraining and achieves high performance across various tasks such as image captioning, image retrieval, and Visual Question Answering. BLIP integrates a Vision Transformer for image encoding and a transformer-based text decoder for generating contextually relevant answers. Its ability to establish strong cross-modal relationships makes it more effective than traditional CNN-LSTM architectures.

Several studies have also focused on multilingual VQA systems to improve accessibility for non-English-speaking users. Most existing VQA models are trained primarily on English datasets, limiting their usability in multilingual environments. Recent research has explored the integration of Neural Machine Translation (NMT) techniques with VQA frameworks to support multilingual question-answering. Transformer-based translation models such as IndicTrans2 and multilingual BERT have shown promising results in translating queries while preserving contextual semantics.

Researchers have additionally explored video-based VQA systems that process temporal visual information from videos rather than static images. These systems typically extract keyframes or temporal features to answer questions related to actions, events, or object interactions within videos. Such approaches are highly useful in surveillance, educational content analysis, and multimedia retrieval systems.

Despite significant progress, existing multilingual VQA systems still face challenges related to computational complexity, translation accuracy, dataset bias, and real-time processing efficiency. The proposed system attempts to address these issues by combining transformer-based VQA, multilingual translation support, video processing, and web-based deployment into a unified and scalable framework. The integration of BLIP with multilingual

https://doi.org/10.63665/IJMEC.1104s.46
ISSN: 2456-4265
IJMEC 2026

translation modules demonstrates the feasibility of developing accessible AI-powered applications capable of serving users across diverse linguistic backgrounds.

EXISTING SYSTEM

While the combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks laid the foundation for early Visual Question Answering (VQA) systems, these architectures exhibit several limitations when applied to modern multimodal reasoning tasks. In traditional VQA frameworks, CNNs are primarily responsible for extracting spatial and visual features from images, whereas LSTMs are used to process sequential textual data such as user questions. Although this architecture achieved moderate success in early-stage implementations, it struggles to capture deep semantic relationships between visual content and natural language queries.

One of the major drawbacks of CNN-LSTM architectures is the weak interaction between visual and textual modalities. In most conventional systems, image features and question embeddings are processed independently and combined only during the later stages of the pipeline using simple fusion techniques such as concatenation, averaging, or linear projection. This shallow fusion mechanism fails to effectively model the complex dependencies between objects in the image and the semantic structure of the question. As a result, the system often produces generic or inaccurate answers when dealing with detailed or context-sensitive queries.

Another limitation arises from the nature of CNN-based feature extraction. Traditional CNN models such as VGGNet or ResNet focus mainly on extracting global visual representations and lack explicit mechanisms for object-level reasoning or relational understanding. Consequently, these systems struggle to identify interactions among multiple objects, spatial arrangements, or contextual cues present in complex scenes. This restricts the model's ability to answer reasoning-intensive questions that require fine-grained visual understanding.

Similarly, LSTMs suffer from several disadvantages in language modeling. Since LSTMs process sequences step-by-step, they are computationally inefficient and difficult to parallelize during training. They also struggle to capture long-range dependencies in lengthy or syntactically complex questions. Furthermore, LSTM-based models often rely heavily on dataset biases and superficial language patterns instead of true semantic understanding, reducing their generalization capability.

Another significant drawback is the absence of advanced attention mechanisms in early CNN-LSTM systems. Without dynamic attention, the model cannot selectively focus on the most relevant image regions corresponding to important keywords in the question. This leads to poor alignment between visual and textual features and negatively impacts answer accuracy, especially in scenarios involving multiple objects or ambiguous questions.

EXISTING ALGORITHM

- CNN (Convolutional Neural Network)
- LSTM (Long Short-Term Memory)

DRAWBACKS OF EXISTING SYSTEM

- Limited cross-modal interaction, as CNN and LSTM features are fused only at shallow stages using simple operations like concatenation or projection.
- Lack of fine-grained alignment between image regions and important words or phrases in the question.
- CNNs primarily extract global image features and do not explicitly model object relationships or spatial reasoning.
- LSTMs struggle with long-term dependencies and complex sentence structures.
- Sequential processing in LSTMs increases training time and reduces computational efficiency.
- Absence of robust attention mechanisms reduces the system's ability to focus on semantically relevant image regions.
- Poor handling of multilingual input and contextual language variations.
- Reduced scalability and lower performance on complex real-world VQA tasks.

PROPOSED SYSTEM

The proposed system introduces an advanced multilingual Visual Question Answering (VQA) web application developed using the Flask framework and powered by the BLIP (Bootstrapping Language-Image Pretraining) architecture, specifically the BLIP model. Unlike conventional CNN-LSTM-based systems, the proposed approach employs a transformer-based multimodal framework capable of jointly learning visual and textual representations, thereby significantly improving contextual understanding and answer accuracy.

The core of the proposed system consists of two major components: a Vision Transformer (ViT) for image understanding and a transformer-based language decoder for answer generation. The Vision Transformer processes uploaded images by dividing them into smaller patches and encoding them into dense semantic embeddings. These embeddings capture both local and global visual features, enabling the system to understand complex object relationships, spatial arrangements, and scene-level context more effectively than traditional CNNs.

Simultaneously, the textual question provided by the user is processed using transformer-based language encoding techniques. Unlike LSTMs, transformers utilize self-attention mechanisms that allow the model to analyze all words in the sentence simultaneously, thereby capturing long-range dependencies and contextual semantics more

https://doi.org/10.63665/IJMEC.1104s.46
ISSN: 2456-4265
IJMEC 2026

efficiently. This results in better understanding of complex, ambiguous, or linguistically rich questions.

A key highlight of the proposed system is its multilingual support. Users can interact with the application in multiple Indian languages such as Hindi, Telugu, Tamil, Kannada, and Malayalam. To facilitate multilingual communication, the system incorporates a mock translation module that converts user queries into English before processing them through the BLIP model and subsequently translates the generated answers back into the user's preferred language. Although currently implemented using a simplified dictionary-based approach, the architecture is designed to support integration with advanced Neural Machine Translation (NMT) frameworks such as IndicTrans2 in future developments.

The system also supports both image and short video inputs. For videos, keyframes are extracted at regular intervals and analyzed individually to generate context-aware responses. This enhances the versatility of the application and expands its usability across educational, accessibility, surveillance, and multimedia understanding domains.

To generate accurate textual answers, the BLIP model employs beam search decoding. Beam search is an intelligent sequence-generation algorithm that evaluates multiple candidate answer sequences at each decoding step and selects the most probable response based on cumulative likelihood scores. This approach improves grammatical coherence, contextual relevance, and overall answer quality compared to greedy decoding methods.

The proposed system also emphasizes security and usability. Uploaded files are validated for supported formats and size restrictions before processing. The Flask-based web interface enables real-time interaction between users and the AI model, making the application accessible, scalable, and user-friendly.

PROPOSED ALGORITHM

Transformer-based Vision-Language Model (BLIP: Bootstrapping Language-Image Pretraining)
Vision Transformer (ViT)
Transformer-based Text Decoder
Beam Search Decoding
IndicTrans2 / Neural Machine Translation (for multilingual support)

ALGORITHM DEFINITION

The algorithm implemented in the proposed system is based on the Visual Question Answering (VQA) paradigm using the BLIP transformer architecture. The uploaded image is first processed by a Vision Transformer, which converts the image into semantic embeddings capable of representing both object-level and scene-level information. Simultaneously, the user's textual question is encoded using transformer-based language representations.

These visual and textual embeddings are jointly processed within a multimodal transformer framework that enables deep cross-modal interaction and semantic alignment. The model learns to associate relevant image regions with important words and phrases in the question through attention mechanisms. Finally, a transformer decoder generates the answer sequence using beam search decoding to ensure contextual accuracy and fluency.

For multilingual functionality, the system incorporates a translation pipeline that converts non-English queries into English before inference and translates the generated English response back into the original language. This modular design allows future integration with real-time transformer-based translation systems such as IndicTrans2.

ADVANTAGES OF PROPOSED SYSTEM

- Advanced multimodal understanding through the BLIP transformer architecture.
- Joint learning of visual and textual representations improves semantic reasoning.
- Vision Transformer captures both local and global image features effectively.
- Transformer-based language processing handles complex sentence structures and long-range dependencies efficiently.
- Attention mechanisms enable precise alignment between image regions and question semantics.
- Beam search decoding generates coherent, contextually accurate, and grammatically meaningful answers.
- Support for multilingual interaction in multiple Indian languages.
- Modular translation pipeline allows easy integration with advanced translation APIs and NMT models.
- Ability to process both images and short-duration videos for enhanced contextual analysis.
- Improved scalability, accuracy, and real-time performance compared to traditional CNN-LSTM models.
- User-friendly Flask-based web application suitable for accessible AI-driven interactions.

SYSTEM ARCHITECTURE

The proposed multilingual Visual Question Answering (VQA) system is designed using a modular and transformer-based architecture that integrates image processing, natural language understanding, multilingual translation, and answer generation within a unified web application framework. The system begins with the user uploading an image or a short-duration video through the web interface developed using Flask. If a video is uploaded, keyframes are extracted at regular intervals to capture meaningful visual information for analysis. The uploaded visual input is then processed by the BLIP model, which serves as the core Visual Question Answering engine.

https://doi.org/10.63665/IJMEC.1104s.46
ISSN: 2456-4265
IJMEC 2026

The architecture utilizes a Vision Transformer (ViT) to extract high-level semantic features from the image. These features are converted into dense embeddings that represent objects, spatial relationships, and contextual details present in the visual scene. Simultaneously, the user's question is accepted in multiple Indian languages such as Hindi, Telugu, Tamil, and Kannada. A multilingual translation module converts the question into English before it is passed to the BLIP model for inference. The transformer-based text decoder within the BLIP architecture jointly processes the visual embeddings and translated question to generate an accurate and context-aware answer.

For answer generation, the system uses beam search decoding to improve grammatical correctness and contextual relevance. Once the answer is generated in English, the translation module converts it back into the user's original language, thereby enabling multilingual interaction. Finally, the generated response is displayed to the user through the web interface in real time. The architecture also includes secure file validation and preprocessing mechanisms to ensure efficient and safe handling of uploaded media files.

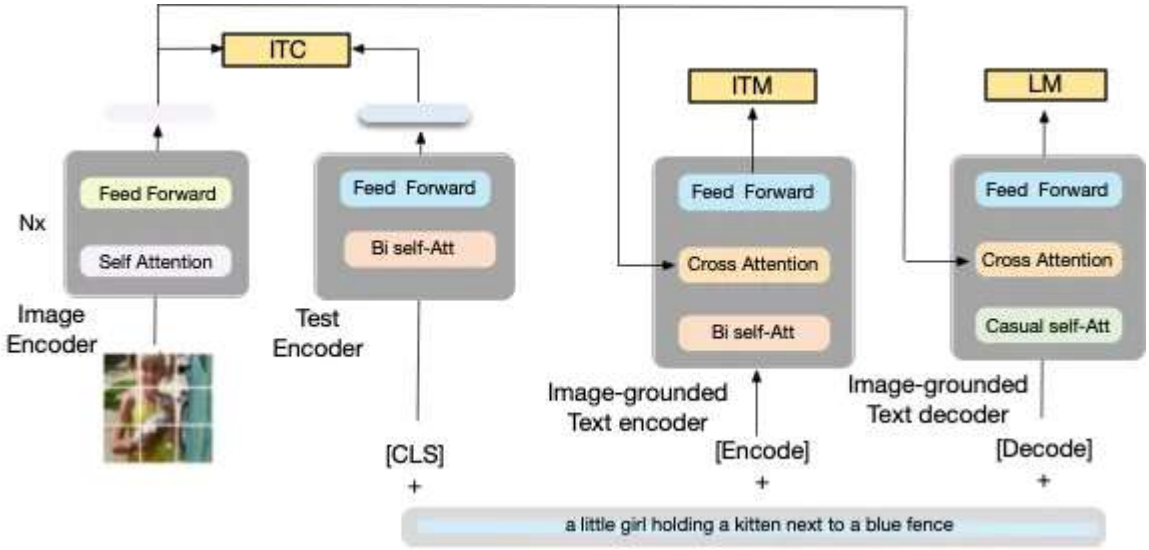

Fig:- proposed model

MINIMUM SYSTEM REQUIREMENTS

Hardware Requirements

The proposed system requires a computer with a minimum Intel Pentium i5 processor to efficiently handle image processing, transformer-based inference, and multilingual question-answering tasks. A minimum of 4 GB DDR RAM is necessary to support the execution of deep learning libraries, model loading, and real-time processing operations. Additionally, the system requires at least 450 GB of hard disk storage for installing the operating system, Python environment, required dependencies, pretrained models, datasets, and application files.

Software Requirements

The software implementation of the proposed system is carried out using the Python programming language, which serves as the backend development environment due to its extensive support for deep learning, natural language processing, and web application frameworks. The application is designed to run on the Windows 10 operating system, which provides compatibility with the required libraries and development tools. The integrated development environment (IDE) used for developing and testing the application is Spyder, which offers an efficient platform for coding, debugging, and executing Python-based machine learning applications.

IMPLEMENTATION RESULTS

The implementation of the proposed multilingual Visual Question Answering (VQA) system was successfully carried out using the Flask framework along with deep learning and Natural Language Processing (NLP) libraries. The system integrates the transformer-based BLIP architecture to process visual inputs and generate accurate answers for user questions. The developed application was tested using various images and short-duration videos containing objects, scenes, text, and human activities to evaluate its performance and usability.

The experimental results demonstrate that the BLIP-based model effectively understands the relationship between visual content and textual questions. The system was able to provide contextually relevant and semantically meaningful answers for different types of queries, including object identification, scene understanding, counting-based questions, and activity recognition. Compared to traditional

https://doi.org/10.63665/IJMEC.1104s.46
ISSN: 2456-4265
IJMEC 2026

CNN-LSTM-based VQA approaches, the transformer-based model showed improved accuracy in capturing fine-grained relationships between image regions and question semantics.

The multilingual translation module also performed successfully during testing. Users were able to submit questions in multiple Indian languages such as Hindi, Telugu, and Tamil. The system translated the questions into English for processing and generated responses that were translated back into the original language. Although a mock translation mechanism was used, the workflow successfully demonstrated the feasibility of multilingual Visual Question Answering in a real-time web environment.

The video-processing component successfully extracted keyframes from uploaded videos and generated answers based on visual information obtained from selected frames. This enabled the system to support basic video-based question answering in addition to static image analysis. The integration of beam search decoding further improved the quality of generated answers by producing grammatically coherent and context-aware responses.

The Flask-based web interface provided smooth interaction between users and the AI model. Secure file upload validation ensured that unsupported or invalid files were rejected before processing. The application was tested on a system with moderate hardware specifications and demonstrated stable performance for real-time inference tasks.

Overall, the implementation results confirm that the proposed transformer-based multilingual VQA system is efficient, scalable, and capable of delivering accurate multimodal understanding within an accessible web application environment.

PERFORMANCE OUTCOME

The proposed system achieved efficient multimodal reasoning by jointly processing visual and textual information through transformer-based learning. The use of Vision Transformers and attention mechanisms significantly enhanced semantic alignment between images and questions. The multilingual support increased accessibility for users from different linguistic backgrounds, while the modular architecture allows future integration with advanced translation APIs and cloud-based deployment platforms.

CONCLUSION

The proposed multilingual Visual Question Answering (VQA) system demonstrates the effective integration of deep learning, computer vision, and natural language processing techniques within a web-based application framework. By utilizing the transformer-based BLIP architecture, the system successfully performs multimodal understanding by jointly processing visual and textual information to generate accurate and context-aware answers. Unlike traditional CNN-LSTM-based approaches, the proposed system leverages Vision Transformers and attention mechanisms to establish stronger semantic relationships between images and user questions, thereby improving the overall performance and reliability of the VQA process.

The application also introduces multilingual interaction capabilities, enabling users to ask questions in multiple Indian languages such as Hindi, Telugu, Tamil, and Kannada. The inclusion of a translation module demonstrates the feasibility of building accessible AI systems that can bridge linguistic barriers and provide intelligent services to a wider audience. Furthermore, the support for both image and short video inputs enhances the versatility and practical applicability of the system in real-world environments.

The use of beam search decoding contributes to generating coherent and contextually meaningful responses, while the Flask-based web framework provides a user-friendly and scalable deployment environment. Overall, the project highlights the growing potential of transformer-based multimodal AI systems in developing intelligent, interactive, and multilingual applications for domains such as education, accessibility, healthcare, surveillance, and smart assistance systems.

FUTURE SCOPE

The proposed system provides a strong foundation for advanced multilingual Visual Question Answering applications, and several enhancements can be implemented in the future to improve its functionality, scalability, and real-world usability.

One major area of future improvement is the integration of real-time Neural Machine Translation (NMT) frameworks such as IndicTrans2 or multilingual transformer models. This would significantly improve translation accuracy, contextual understanding, and support for a larger number of regional and international languages.

The system can also be extended to support real-time video streaming and live camera input, enabling applications in surveillance systems, smart classrooms, robotics, and assistive technologies for visually impaired individuals. Incorporating temporal reasoning and advanced video understanding models would improve the system's capability to answer complex questions related to actions and events occurring over time.

Another potential enhancement involves integrating speech recognition and text-to-speech technologies to create a voice-enabled VQA assistant. This would make the application more accessible for users with limited literacy or physical disabilities. Additionally, deploying the system on cloud platforms and optimizing transformer models for edge devices can improve scalability, reduce inference latency, and enable mobile or embedded system integration.

https://doi.org/10.63665/IJMEC.1104s.46
ISSN: 2456-4265
IJMEC 2026

Future research can also focus on improving reasoning capabilities through knowledge graph integration and external memory modules, allowing the system to answer more complex logical and commonsense questions. Training the model on larger multilingual datasets can further reduce bias and improve generalization across diverse linguistic and visual scenarios.

Moreover, advanced security features such as user authentication, encrypted data transmission, and secure cloud storage can be incorporated to support enterprise-level applications. The system can also be adapted for specialized domains such as medical image analysis, e-learning platforms, industrial automation, and smart agriculture.

Overall, the future scope of multilingual transformer-based VQA systems is vast, with opportunities for innovation in intelligent human-computer interaction, accessibility technologies, and real-time multimodal AI applications.

REFERENCES

1) Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation," in *Proceedings of the International Conference on Machine Learning (ICML)*, 2022.
2) Ashish Vaswani et al., "Attention Is All You Need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
3) Alexey Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in *International Conference on Learning Representations (ICLR)*, 2021.
4) Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
5) Karen Simonyan and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *International Conference on Learning Representations (ICLR)*, 2015.
6) Sepp Hochreiter and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
7) Zhou Yu, Jun Yu, Yuhao Cui, Dacheng Tao, and Qi Tian, "Deep Modular Co-Attention Networks for Visual Question Answering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
8) Tanmay Gupta, Ameya Godbole, and Mitesh M. Khapra, "IndicTrans2: Towards High-Quality and Accessible Machine Translation Models for Indian Languages," 2023.
9) Alec Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," in *International Conference on Machine Learning (ICML)*, 2021.
10) Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan, "Show and Tell: A Neural Image Caption Generator," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

https://doi.org/10.63665/IJMEC.1104s.46
ISSN: 2456-4265
IJMEC 2026