**Title of your invention**

**Visual Question Answering (VQA) System Using Satellite Images**

**Type of Invention**

Software-Based AI System integrating Image Processing and Natural Language Understanding for Remote Sensing.

**Brief Description of your invention**

The invention proposes a Visual Question Answering system tailored for satellite images. It allows users to upload satellite images and ask natural language questions, receiving AI-generated answers based on the image content. It merges convolutional neural networks for image interpretation and transformer-based models for language understanding.

**Objective of your invention**

To develop a real-time, intelligent, and scalable AI system that interprets satellite images using natural language queries to aid in urban planning, disaster response, environmental monitoring, and land use classification.

**How to use the invention**

* Upload a satellite image via the web interface.
* Enter a text-based query (e.g., “What type of land cover is shown?”).
* The system processes the image and question, returning a text-based answer with visual highlights.
* Results are displayed in under 3 seconds.

**Problem your invention is solving**

Manual interpretation of satellite images is time-consuming, requires domain expertise, and is not scalable. This system democratizes geospatial intelligence through automated, AI-based interpretation.

**Purpose and object of Invention**

To automate satellite image understanding for non-experts using AI, improving accessibility to geospatial insights for decision-makers in real time.

**Discuss potential commercial application of the invention**

* Urban planning tools for municipalities
* Disaster management dashboards for relief agencies
* Environmental monitoring platforms for research organizations
* Integration into GIS-based SaaS products

**Provide any additional material**

* Project Report titled “Visual Question Answering with Satellite Images”
* Screenshots of system interface and model architecture
* Results tables showing 92.4% accuracy with <3s response time
* Preprint research paper under peer-review (if applicable)

**Abstract**

The invention is a Visual Question Answering system designed for satellite imagery that leverages CNNs and transformer-based models. It enables users to ask natural language questions about satellite images and receive relevant, accurate answers in real time. Achieving a response accuracy of 92.4% with under 3 seconds of latency, it supports use cases in disaster response, land classification, and urban monitoring.

**Summary of the invention**

This system integrates:

* CNNs for extracting visual features from satellite imagery
* Transformer NLP models (like BERT) for question understanding
* A multimodal fusion module to combine text and image features
* A web-based UI for real-time interaction
* Backend optimization for fast inference
* It is scalable, secure, and applicable across several domains involving geospatial intelligence.

**Detail description of invention with methodology**

* **Image Preprocessing**: Normalization, segmentation, and noise reduction
* **Feature Extraction**: CNN-based model trained on 50,000+ images
* **Question Encoding**: Tokenization and embedding using BERT
* **Multimodal Fusion**: Aligns spatial and linguistic data
* **Inference**: Generates answers and overlays on images
* **Deployment**: React-based frontend and Flask backend; cloud-hosted
* **Performance**: Tested with 92.4% accuracy and <3s average response time

**Applicant and inventor details:**

| **Name** | **Roll Number** | **Email** |
| --- | --- | --- |
| Sarthak Sharma | 2100290120147 | sarthak.2125cs1081@kiet.edu |
| Vinayak Mishra | 21002090120190 | vinayak.2125cs1091@kiet. |
| Rishikesh | 2200290129015 | rishikesh.2125cs1222@kiet.edu |
|  |  |  |

**Supervisor**:

Dr. Gaurav Dubey, Professor, Computer Science Department