

✔ Project Requirements for visionGuard

1. Software Requirements:

Programming Language: Python (preferred for ML/AI)

Frameworks & Libraries:

TensorFlow / Keras → Model building (captioning & segmentation)

OpenCV → Image processing & video streaming

NumPy, Matplotlib → Data handling & visualization

Flask → Web-based deployment interface

2. Hardware Requirements:

Processor: Intel i5 or higher (or equivalent AMD/Ryzen)

GPU: NVIDIA GPU (e.g., GTX 1660 / RTX series) for faster training (optional but recommended)

RAM: Minimum 8 GB (16 GB recommended for training large datasets)

Storage: At least 50 GB (datasets + trained models + logs)

Camera/Video Source: For live video streaming and real-time testing

3. Dataset Requirements:

Captioning Dataset:

MS COCO (330,000+ images, 5 captions per image)

Segmentation Datasets:

Cityscapes (urban traffic scenes, pixel-level labeling)

ADE20K (broad scene segmentation dataset)

Preprocessing:

Image resizing & normalization

Data augmentation (flipping, rotation, etc.)

Caption tokenization & sequence padding

4. Functional Requirements:

Ability to detect and segment traffic-relevant objects (vehicles, pedestrians, signals, roads, etc.)

Generate short, safety-focused captions for detected hazards

Perform real-time video streaming with integrated captioning + segmentation

Handle unseen traffic scenarios with generalization

5. Non- Functional Requirements:

Performance: Low latency in video streaming

Scalability: Can be extended to larger datasets and new environments

Reliability: Consistent hazard detection and caption generation

Usability: Simple interface for testing (Flask web app)