RGB And IR Fusion Data Collection

1. Dataset Collection

For the RGB–IR image fusion project, the quality of the dataset is critical. The dataset must contain paired and spatially aligned RGB and infrared images to ensure that corresponding features from both modalities represent the same scene. Commonly used sources include:

- Public Benchmark Datasets:
  - TNO Image Fusion Dataset: Contains various day/night military and surveillance scenes with registered RGB and IR pairs.
  - MSRS (Multi-Spectral Road Scene Dataset): Designed for autonomous driving scenarios with high-quality aligned RGB–IR frames.
  - FLIR Thermal Dataset: Provides thermal images along with RGB data for object detection tasks.
- Custom Data Collection (if required):
  - Capturing RGB and IR data simultaneously using a dual-sensor camera setup.
  - Ensuring hardware synchronization to prevent motion-related misalignments.
  - Using calibration targets (checkerboards, thermal-visible markers) for accurate spatial registration.

2. Preprocessing Steps

Preprocessing ensures that the input data is normalized, aligned, and formatted for the fusion model. The following operations are applied:

a) Geometric Alignment (Image Registration)

RGB and IR sensors often have different resolutions, fields of view, and optical distortions. Before fusion:

- Apply feature-based registration (SIFT, ORB, SURF) or deep-learning-based registration (VoxelMorph).
- Perform affine or homography transformations to match spatial coordinates.

b) Resolution Matching

- Rescale both images to a fixed size (e.g., 256×256 or 512×512).
- Use bicubic interpolation for resizing to minimize loss of detail.

c) Intensity Normalization

- Normalize pixel values to [0,1] or standardize (zero mean, unit variance).

- For IR images, apply histogram equalization or contrast-limited adaptive histogram equalization (CLAHE) to enhance thermal detail.

d) Channel Formatting

- RGB images: 3 channels (R, G, B).
- IR images: single channel → replicated or stacked with RGB to create a 4D voxel input (H×W×4).

e) Data Augmentation

To improve model generalization:

- Random rotations (±10–15°), flips, and small translations.
- Adding Gaussian noise or blurring to simulate real-world distortions.
- Adjusting brightness and contrast for robustness across lighting conditions.

f) Dataset Splitting

- Training Set: 70%
- Validation Set: 15%
- Test Set: 15%

2.3 Prompt Design for RGB–IR Medical Image Fusion Project

In a medical dataset context, prompts are not "user-facing text inputs" like in a chatbot but **structured model directives and parameter settings** that guide the fusion process, particularly when using deep-learning frameworks or generative AI models. In research reporting, you typically define:

1. Purpose of Prompt Design

For medical image fusion (e.g., combining RGB endoscopic images with IR thermal scans):

- The **prompt** specifies **what features to preserve** from each modality (e.g., anatomical color vs. thermal contrast).
- It constrains the model so that **diagnostically important structures** are not lost during fusion.
- It formalizes **fusion rules** (e.g., "retain all edges from IR, preserve color texture from RGB").

2. Prompt Components

Component	Example for Medical Fusion
Modality Roles	"Preserve fine anatomical color texture from RGB; emphasize vascular heat patterns from IR."

Component	Example for Medical Fusion
Region of Interest (ROI)	"Focus fusion on the lesion area (x,y coordinates) while maintaining overall organ structure."
Structural Priority	"Retain boundaries and edges from IR for sharpness; avoid over-smoothing."
Contrast Management	"Increase visibility of low-intensity IR regions while maintaining RGB natural appearance."
Noise Handling	"Suppress IR sensor noise but do not blur diagnostically relevant hot-spots."

3. Example Prompt for Model Training (Pseudo)

yaml

CopyEdit

```yaml
fusion_prompt:
  objective: "Fuse RGB and IR medical images to enhance vascular and thermal anomalies while preserving anatomical color fidelity."
  constraints:
    - preserve_edges: true
    - maintain_color_naturalness: true
    - avoid_texture_loss: true
  roi:
    - coordinates: [100:300, 200:400]  # region containing lesion
  weights:
    rgb_detail_weight: 0.6
    ir_contrast_weight: 0.4
  noise_suppression: "mild"
```

2.4 Gradio UI Development and System Setup

1. Purpose of Gradio UI in RGB–IR Fusion Project

A Gradio-based user interface allows non-technical users (e.g., clinicians or researchers) to:

- Upload paired RGB and IR medical images.

- Perform fusion using the trained model.
- Visualize the original and fused outputs side by side.
- Download or analyze fused images without writing code.
- **Libraries:**
  - torch, torchvision (for the fusion model)
  - opencv-python (image preprocessing)
  - numpy, matplotlib (data handling & visualization)
  - gradio (UI development)