Summary of "CameraNet: A Two-Stage Framework for Effective Camera ISP Learning"

The paper proposes **CameraNet**, a two-stage CNN framework for camera image signal processing (ISP). Traditional ISP pipelines use cascaded modules for tasks like denoising, demosaicking, and color enhancement, but they often suffer from error accumulation and poor performance in challenging scenarios. CameraNet divides ISP tasks into two groups: **restoration** (e.g., denoising, demosaicking) and **enhancement** (e.g., tone mapping, color adjustment). The framework uses two subnetworks (Restore-Net and Enhance-Net) trained with intermediate and final ground truths. Experiments on three datasets demonstrate superior performance over existing methods.

Dataset Handling and Processing

The paper evaluates CameraNet on three datasets: **HDR+**, **SID**, and **FiveK**. Below are details on their creation, processing, and subsets:

1. HDR+ Dataset

- **Focus**: Burst denoising and detail enhancement.
- **Subset**: Nexus 6P subset (665 training scenes, 240 testing scenes).
- **Input**: Single raw image (reference frame from burst alignment).
- **Ground Truths**:
  - **Restoration (G_r)**: Generated using DCraw for demosaicking, white balance, and color conversion on fused raw images.
  - **Enhancement (G_o)**: Provided sRGB images processed by the HDR+ algorithm.
- **Sampling**: Testing data sampled based on ISO values to reflect noise levels.

2. SID Dataset

- **Focus**: Low-light denoising.
- **Subset**: Sony A7S2 subset (181 training scenes, 50 testing scenes).
- **Input**: Noisy short-exposure raw images.
- **Ground Truths**:
  - **Restoration (G_r)**: Generated from long-exposure raw images using DCraw.
  - **Enhancement (G_o)**: Created using Photoshop's auto-enhancement tool.
- **Sampling**: Testing data sampled based on ISO values.

3. FiveK Dataset

- **Focus**: Manual color/style retouching.
- **Subset**: Nikon D700 subset (500 training images, 150 testing images).
- **Input**: Raw images.
- **Ground Truths**:

- **Restoration (G_r)**: Generated using DCraw on raw images.
- **Enhancement (G_o)**: Expert-C's manually retouched images from Lightroom.
- **Sampling**: Testing data uniformly sampled.

Key Dataset Processing Steps

1. **Ground Truth Creation**:

   - **Restoration (G_r)**: Tools like DCraw and Adobe Camera Raw were used for tasks like demosaicking, denoising, and white balance.
   - **Enhancement (G_o)**: Adobe Lightroom or Photoshop auto-enhancement tools were used for color/style adjustments.
   - For HDR+, burst raw images were fused to reduce noise before restoration.

2. **Data Splits**:

   - Training/testing splits are camera-specific (e.g., Sony A7S2 for SID, Nikon D700 for FiveK).
   - Testing sets were sampled to reflect real-world conditions (e.g., ISO distributions).

3. **Cross-Dataset Testing**:

   - Models trained on one dataset (e.g., FiveK) were tested on another (e.g., HDR+), revealing sensor dependency in Restore-Net.

Limitations and Notes

- **Overfitting**: Observed in SID and FiveK due to limited training data.
- **Sensor Dependency**: Restoration tasks heavily depend on sensor characteristics, requiring retraining for new cameras.
- **Enhancement Flexibility**: Enhance-Net can be fine-tuned for different styles (e.g., nighttime, portrait).

This structured approach to dataset handling and ground truth generation ensures effective training of the two-stage CameraNet framework.