Reproducibility Documentation

Random Seed Implementation

All random operations used fixed seeds:

- Data splitting: `seed=42`
- Data augmentation: `seed=42`

Data Splitting Code

```
import numpy as np
from sklearn.model_selection import train_test_split

# ===== FIXED SEEDS ===== #
np.random.seed(42)  # Global seed

# Load all 1000 original images
originals = [f"IMG_{i:03d}.jpg" for i in range(1, 1001)]

# Step 1: Split original images (8:1:1)
train_val, test = train_test_split(originals, test_size=300, random_state=42)
train, val = train_test_split(train_val, test_size=300, random_state=42)

# Step 2: Generate augmentations ONLY for training set
train_aug_flip = [f"{img}_flip.jpg" for img in train]  # Horizontal flip
train_aug_crop = [f"{img}_crop.jpg" for img in train]  # Random crop

# Final sets
train_set = train + train_aug_flip + train_aug_crop
val_set = val
test_set = test
```