Model 1

Number of epochs: 2

```
Epoch 1/2 [Training]: 100%|██████████| 157/157 [11:25<00:00,  4.37s/it,
avg_loss=0.205]
Epoch 1/2 [Validation]: 100%|██████████| 34/34 [30:23<00:00, 53.62s/it,
avg_loss=0.0358]
Best model saved!
Epoch 1/2 - Train Loss: 0.2051, Val Loss: 0.0358, Time: 2509.46s
Learning Rate: 0.001

Epoch 2/2 [Training]: 100%|██████████| 157/157 [06:29<00:00,  2.48s/it,
avg_loss=0.0205]
Epoch 2/2 [Validation]: 100%|██████████| 34/34 [01:22<00:00,  2.43s/it,
avg_loss=0.0122]
Best model saved!
Epoch 2/2 - Train Loss: 0.0205, Val Loss: 0.0122, Time: 472.77s
Learning Rate: 0.001
```

Evaluation Metrics:

```
=== Model Evaluation on Test Set ===
IoU per class:
  Mean: 0.3861 | Std: 0.3960

Dice (F1) per class:
  Mean: 0.4389 | Std: 0.4171

Precision per class:
  Mean: 0.5972 | Std: 0.4416

Recall per class:
  Mean: 0.3989 | Std: 0.4055
```

Model 2

Results:

Number of Epochs: 6

```
Epoch 1/6 [Training]: 100%|██████████| 157/157 [2:11:18<00:00, 50.18s/it,
avg_loss=0.149]
Epoch 1/6 [Validation]: 100%|██████████| 34/34 [29:36<00:00, 52.26s/it,
avg_loss=0.0335]
Best model saved!
Epoch 1/6 - Train Loss: 0.1490, Val Loss: 0.0335, Time: 9664.82s
Learning Rate: 0.001

Epoch 2/6 [Training]: 100%|██████████| 157/157 [06:32<00:00,  2.50s/it,
avg_loss=0.0189]
Epoch 2/6 [Validation]: 100%|██████████| 34/34 [01:21<00:00,  2.41s/it,
avg_loss=0.0126]
Best model saved!
Epoch 2/6 - Train Loss: 0.0189, Val Loss: 0.0126, Time: 474.87s
Learning Rate: 0.001

Epoch 3/6 [Training]: 100%|██████████| 157/157 [07:03<00:00,  2.70s/it,
avg_loss=0.00937]
Epoch 3/6 [Validation]: 100%|██████████| 34/34 [01:38<00:00,  2.91s/it,
avg_loss=0.00716]
Best model saved!
Epoch 3/6 - Train Loss: 0.0094, Val Loss: 0.0072, Time: 522.99s
Learning Rate: 0.001

Epoch 4/6 [Training]: 100%|██████████| 157/157 [06:33<00:00,  2.51s/it,
avg_loss=0.00655]
Epoch 4/6 [Validation]: 100%|██████████| 34/34 [01:22<00:00,  2.41s/it,
avg_loss=0.00561]
Best model saved!
Epoch 4/6 - Train Loss: 0.0066, Val Loss: 0.0056, Time: 476.36s
Learning Rate: 0.001

Epoch 5/6 [Training]: 100%|██████████| 157/157 [06:29<00:00,  2.48s/it,
avg_loss=0.00532]
Epoch 5/6 [Validation]: 100%|██████████| 34/34 [01:23<00:00,  2.44s/it,
avg_loss=0.00483]
Best model saved!
Epoch 5/6 - Train Loss: 0.0053, Val Loss: 0.0048, Time: 472.94s
Learning Rate: 0.001
```
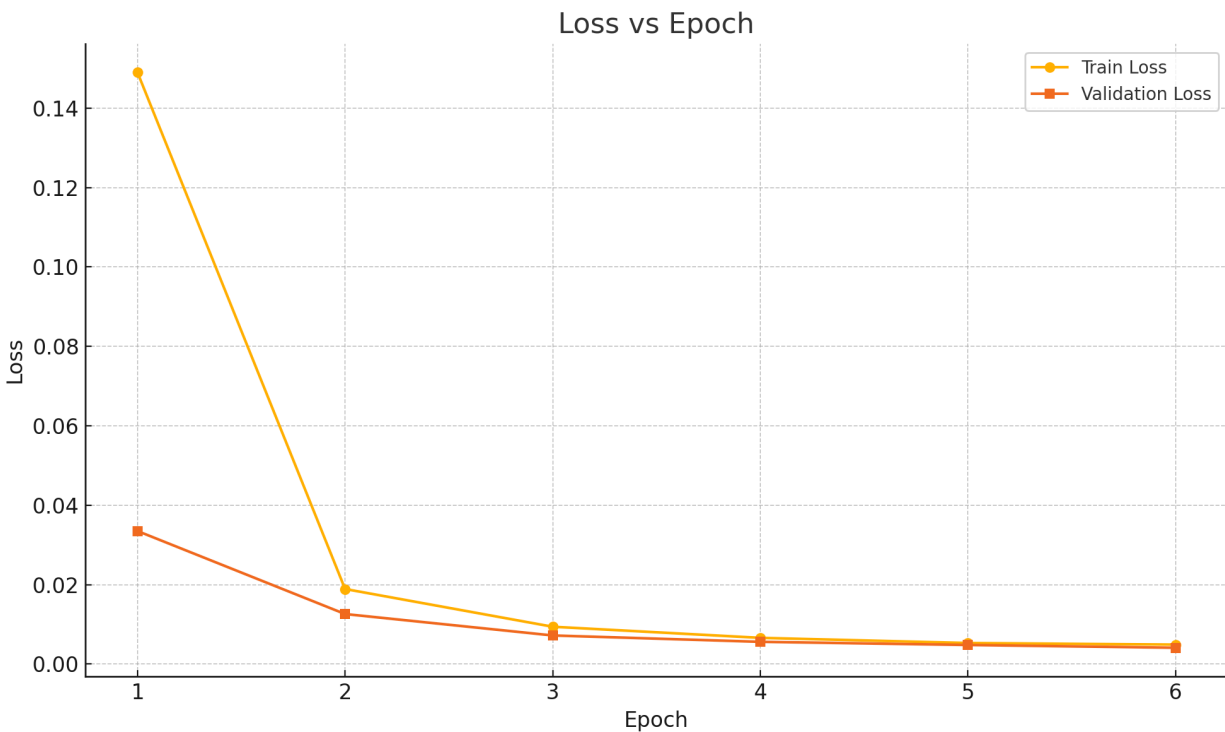
```
Epoch 6/6 [Training]: 100%|████████████| 157/157 [06:28<00:00,  2.47s/it,
avg_loss=0.00489]
Epoch 6/6 [Validation]: 100%|████████████| 34/34 [01:22<00:00,  2.42s/it,
avg_loss=0.00406]
Best model saved!
Epoch 6/6 - Train Loss: 0.0049, Val Loss: 0.0041, Time: 471.16s
Learning Rate: 0.001
```

Loss vs Epoch graph:

Observations:

1. A sharp drop in both training and validation loss from epoch 1 to 2 — the model quickly learned meaningful features.
2. Continued gradual improvement across all epochs — a great sign of stable convergence.
3. No overfitting was observed. Validation loss follows training loss closely.

Evaluation Metrics:

```
=== Model Evaluation on Test Set ===
IoU per class:
  Mean: 0.4894 | Std: 0.3807

Dice (F1) per class:
  Mean: 0.5558 | Std: 0.4023

Precision per class:
  Mean: 0.6038 | Std: 0.4244

Recall per class:
  Mean: 0.5225 | Std: 0.3921
```

Interpretations:
1. IoU of ~0.49 and Dice of ~0.56 indicate that the model reasonably overlaps the ground truth masks but still leaves room for improvement.
2. High standard deviations (~0.38–0.42) suggest performance inconsistency across classes — the model likely **performs better on some defect types than others**.
3. Precision > Recall implies:
    The model is conservative in predictions — it prefers to avoid false positives.
    However, this comes at the cost of missing some actual defects (false negatives).

Evaluation Metrics of each class:

```
Per-class Metrics:
Class 0(No defect,Defect_Class0)
  Precision: 0.9334
  Recall:    0.9932
  F1-score:  0.9624
  IoU:       0.9275

Class 1(Swelling,Defect_Class5)
  Precision: 0.0000
  Recall:    0.0000
  F1-score:  0.0000
  IoU:       0.0000

Class 2(Spatter,Defect_Class8)
  Precision: 0.7821
  Recall:    0.7186
  F1-score:  0.7490
  IoU:       0.5987
```

```
Class 3(Misprint, Defect_Class9)
  Precision: 0.0000
  Recall:    0.0000
  F1-score:  0.0000
  IoU:       0.0000

Class 4(Over-Melting, Defect_Class10)
  Precision: 0.0000
  Recall:    0.0000
  F1-score:  0.0000
  IoU:       0.0000

Class 5(Under-Melting, Defect_Class11)
  Precision: 0.0000
  Recall:    0.0000
  F1-score:  0.0000
  IoU:       0.0000
```

After Melting Defect Detection

Original - layer_54.png	Ground Truth	Prediction

Original - layer_56.png	Ground Truth	Prediction

Original - layer_74.png	Ground Truth	Prediction

Original - layer_77.png	Ground Truth	Prediction

Original - layer_83.png	Ground Truth	Prediction