```
ckpt/gmflow_things-e9887eda.pth

GMFlow(
  (backbone): CNNEncoder(
    (conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
    (norm1): InstanceNorm2d(64, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
    (relu1): ReLU(inplace=True)
    (layer1): Sequential(
      (0): ResidualBlock(
        (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (relu): ReLU(inplace=True)
        (norm1): InstanceNorm2d(64, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm2): InstanceNorm2d(64, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
      )
      (1): ResidualBlock(
        (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (relu): ReLU(inplace=True)
        (norm1): InstanceNorm2d(64, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm2): InstanceNorm2d(64, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
      )
    )
    (layer2): Sequential(
      (0): ResidualBlock(
        (conv1): Conv2d(64, 96, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1),
bias=False)
        (conv2): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (relu): ReLU(inplace=True)
        (norm1): InstanceNorm2d(96, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm2): InstanceNorm2d(96, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm3): InstanceNorm2d(96, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (downsample): Sequential(
          (0): Conv2d(64, 96, kernel_size=(1, 1), stride=(2, 2))
          (1): InstanceNorm2d(96, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        )
      )
      (1): ResidualBlock(
        (conv1): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (conv2): Conv2d(96, 96, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (relu): ReLU(inplace=True)
        (norm1): InstanceNorm2d(96, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm2): InstanceNorm2d(96, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
      )
    )
    (layer3): Sequential(
      (0): ResidualBlock(
        (conv1): Conv2d(96, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1),
```

```
bias=False)
        (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (relu): ReLU(inplace=True)
        (norm1): InstanceNorm2d(128, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm2): InstanceNorm2d(128, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm3): InstanceNorm2d(128, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (downsample): Sequential(
          (0): Conv2d(96, 128, kernel_size=(1, 1), stride=(2, 2))
          (1): InstanceNorm2d(128, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        )
      )
      (1): ResidualBlock(
        (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
        (relu): ReLU(inplace=True)
        (norm1): InstanceNorm2d(128, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
        (norm2): InstanceNorm2d(128, eps=1e-05, momentum=0.1, affine=False,
track_running_stats=False)
      )
    )
    (conv2): Conv2d(128, 128, kernel_size=(1, 1), stride=(1, 1))
  )
  (transformer): FeatureTransformer(
    (layers): ModuleList(
      (0): TransformerBlock(
        (self_attn): TransformerLayer(
          (q_proj): Linear(in_features=128, out_features=128, bias=False)
          (k_proj): Linear(in_features=128, out_features=128, bias=False)
          (v_proj): Linear(in_features=128, out_features=128, bias=False)
          (merge): Linear(in_features=128, out_features=128, bias=False)
          (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
        )
        (cross_attn_ffn): TransformerLayer(
          (q_proj): Linear(in_features=128, out_features=128, bias=False)
          (k_proj): Linear(in_features=128, out_features=128, bias=False)
          (v_proj): Linear(in_features=128, out_features=128, bias=False)
          (merge): Linear(in_features=128, out_features=128, bias=False)
          (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
          (mlp): Sequential(
            (0): Linear(in_features=256, out_features=1024, bias=False)
            (1): GELU(approximate=none)
            (2): Linear(in_features=1024, out_features=128, bias=False)
          )
          (norm2): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
        )
      )
      (1): TransformerBlock(
        (self_attn): TransformerLayer(
          (q_proj): Linear(in_features=128, out_features=128, bias=False)
          (k_proj): Linear(in_features=128, out_features=128, bias=False)
          (v_proj): Linear(in_features=128, out_features=128, bias=False)
          (merge): Linear(in_features=128, out_features=128, bias=False)
          (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
        )
        (cross_attn_ffn): TransformerLayer(
          (q_proj): Linear(in_features=128, out_features=128, bias=False)
          (k_proj): Linear(in_features=128, out_features=128, bias=False)
          (v_proj): Linear(in_features=128, out_features=128, bias=False)
```

```
      (merge): Linear(in_features=128, out_features=128, bias=False)
      (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
      (mlp): Sequential(
        (0): Linear(in_features=256, out_features=1024, bias=False)
        (1): GELU(approximate=none)
        (2): Linear(in_features=1024, out_features=128, bias=False)
      )
      (norm2): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
    )
  )
  (2): TransformerBlock(
    (self_attn): TransformerLayer(
      (q_proj): Linear(in_features=128, out_features=128, bias=False)
      (k_proj): Linear(in_features=128, out_features=128, bias=False)
      (v_proj): Linear(in_features=128, out_features=128, bias=False)
      (merge): Linear(in_features=128, out_features=128, bias=False)
      (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
    )
    (cross_attn_ffn): TransformerLayer(
      (q_proj): Linear(in_features=128, out_features=128, bias=False)
      (k_proj): Linear(in_features=128, out_features=128, bias=False)
      (v_proj): Linear(in_features=128, out_features=128, bias=False)
      (merge): Linear(in_features=128, out_features=128, bias=False)
      (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
      (mlp): Sequential(
        (0): Linear(in_features=256, out_features=1024, bias=False)
        (1): GELU(approximate=none)
        (2): Linear(in_features=1024, out_features=128, bias=False)
      )
      (norm2): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
    )
  )
  (3): TransformerBlock(
    (self_attn): TransformerLayer(
      (q_proj): Linear(in_features=128, out_features=128, bias=False)
      (k_proj): Linear(in_features=128, out_features=128, bias=False)
      (v_proj): Linear(in_features=128, out_features=128, bias=False)
      (merge): Linear(in_features=128, out_features=128, bias=False)
      (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
    )
    (cross_attn_ffn): TransformerLayer(
      (q_proj): Linear(in_features=128, out_features=128, bias=False)
      (k_proj): Linear(in_features=128, out_features=128, bias=False)
      (v_proj): Linear(in_features=128, out_features=128, bias=False)
      (merge): Linear(in_features=128, out_features=128, bias=False)
      (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
      (mlp): Sequential(
        (0): Linear(in_features=256, out_features=1024, bias=False)
        (1): GELU(approximate=none)
        (2): Linear(in_features=1024, out_features=128, bias=False)
      )
      (norm2): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
    )
  )
  (4): TransformerBlock(
    (self_attn): TransformerLayer(
      (q_proj): Linear(in_features=128, out_features=128, bias=False)
      (k_proj): Linear(in_features=128, out_features=128, bias=False)
      (v_proj): Linear(in_features=128, out_features=128, bias=False)
      (merge): Linear(in_features=128, out_features=128, bias=False)
      (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
    )
    (cross_attn_ffn): TransformerLayer(
      (q_proj): Linear(in_features=128, out_features=128, bias=False)
      (k_proj): Linear(in_features=128, out_features=128, bias=False)
      (v_proj): Linear(in_features=128, out_features=128, bias=False)
```

```
          (merge): Linear(in_features=128, out_features=128, bias=False)
          (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
          (mlp): Sequential(
            (0): Linear(in_features=256, out_features=1024, bias=False)
            (1): GELU(approximate=none)
            (2): Linear(in_features=1024, out_features=128, bias=False)
          )
          (norm2): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
        )
      )
      (5): TransformerBlock(
        (self_attn): TransformerLayer(
          (q_proj): Linear(in_features=128, out_features=128, bias=False)
          (k_proj): Linear(in_features=128, out_features=128, bias=False)
          (v_proj): Linear(in_features=128, out_features=128, bias=False)
          (merge): Linear(in_features=128, out_features=128, bias=False)
          (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
        )
        (cross_attn_ffn): TransformerLayer(
          (q_proj): Linear(in_features=128, out_features=128, bias=False)
          (k_proj): Linear(in_features=128, out_features=128, bias=False)
          (v_proj): Linear(in_features=128, out_features=128, bias=False)
          (merge): Linear(in_features=128, out_features=128, bias=False)
          (norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
          (mlp): Sequential(
            (0): Linear(in_features=256, out_features=1024, bias=False)
            (1): GELU(approximate=none)
            (2): Linear(in_features=1024, out_features=128, bias=False)
          )
          (norm2): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
        )
      )
    )
  )
  (feature_flow_attn): FeatureFlowAttention(
    (q_proj): Linear(in_features=128, out_features=128, bias=True)
    (k_proj): Linear(in_features=128, out_features=128, bias=True)
  )
  (upsampler): Sequential(
    (0): Conv2d(130, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): Conv2d(256, 576, kernel_size=(1, 1), stride=(1, 1))
  )
)

{'EFTs_Car100_epe': 44.3183, 'EFTs_Car200_epe': 41.73856, 'EFTs_Car2000_epe': 41.250183,
'final': 42.43568}
```