

# Draft\_Final Project\_YL

April 7, 2024

## 1 Introduction

### 1.1 Background of the Problem

**Context:** Advances in deep learning have enhanced automated medical image analysis, but existing techniques face high computational requirements and performance drops with reduced batch sizes or training epochs.

**Problem Type:** The paper addresses issues in medical image analysis, particularly focusing on self-supervised learning approaches for processing label-free images efficiently.

**Importance:** Solving this problem is crucial because data labeling in medical imaging is expensive and time-consuming, and often data is scarce, especially for emerging diseases like certain autoimmune conditions.

**Difficulty:** Challenges include minimal data availability, the need for domain-specific knowledge for labeling, patient privacy issues, and an incomplete understanding of diseases.

**State-of-the-Art Methods and Effectiveness** Current Techniques: The prevailing methods in medical image analysis primarily rely on self-supervised learning frameworks utilizing either Convolutional Neural Networks (CNNs) or Transformers. These techniques are heavily dependent on extensive datasets and large batch sizes.

Performance Challenges: A notable limitation of these existing approaches is their significant reduction in performance when the conditions of large datasets and batch sizes are not met. This issue becomes more pronounced with constrained computational resources.

Computational Demands: Current state-of-the-art methods require considerable computational power, which poses a barrier to their application, especially in settings with limited resources. Such extensive computational requirements limit the practical accessibility of these advanced techniques.

### 1.2 Paper Explanation

**Proposal:** The paper introduces Cross Architectural Self-Supervision (CASS), a novel method combining CNNs and Transformers in a self-supervised learning setting. It addresses the challenges of limited data and computational resources in medical image analysis.

**Innovations:** CASS leverages both CNN and Transformer architectures simultaneously, improving robustness to changes in batch size and training epochs and reducing computational requirements.

**Performance:** Demonstrated improvements across four medical datasets in terms of F1 Score and Recall, using less labeled data and significantly less training time compared to existing methods.

**Contribution:** CASS represents a significant step in self-supervised learning for medical image analysis, especially beneficial for emerging diseases with limited data. It stands out in efficiency, effectiveness, and adaptability to resource constraints.

## 2 Scope of Reproducibility - Hypotheses to be Tested:

**\*\* Hypotheses 1 Reproduced CASS-trained models outperform existing self-supervised learning methods in terms of accuracy and efficiency on healthcare tasks shown in the paper, i.e., disease cell classification, brain tumor classification, and skin lesion classification.** Hypotheses 2 **\*\* Reproduced CASS demonstrates greater robustness to variations in batch size and pretraining epochs compared to current methods**

## 3 Methodology:

### 3.1 Data

**Data descriptions** Autoimmune Diseases Biopsy Slides Dataset: This dataset includes 198 TIFF images from muscle biopsies of dermatomyositis patients. These slides are stained with different proteins to help diagnose dermatomyositis, a type of autoimmune disease. The dataset involves multi-label classification for different cell classes, such as TFH-1, TFH-217, TFH-Like B cells, and others. The images are consistent in size, measuring 352 by 469 pixels in RGB format .

Dermofit Dataset: Comprising 1300 normal RGB images captured indoors with an SLR camera and ring lightning, this dataset categorizes images into 10 classes associated with skin lesions and conditions. The images vary in size, ranging from 205×205 to 1020×1020 pixels, with no two images being the same size. The dataset’s primary task is multi-class classification .

Brain Tumor MRI Dataset: This dataset includes 7022 images of human brain MRIs, classified into four categories: glioma, meningioma, no tumor, and pituitary. The images vary in size from 512×512 to 219×234 pixels. The dataset’s source is a combination of different datasets and includes 5712 images for training and 1310 for testing .

Using Brain Tumor MRI Dataset for a demo in this report, switch the name to run experiments on other datasets.

```
[1]: import torch
import torch.nn as nn
import torch.optim as optim
```

```

import torchvision.transforms as transforms
from torch.utils.data import Dataset, DataLoader, random_split
from torch.utils.data import DataLoader
from torchvision.models import resnet18
from torchvision.datasets import ImageFolder
from torchvision.utils import make_grid
import torchvision.transforms.functional as F
import numpy as np
import matplotlib.pyplot as plt

import os
import pytorch_lightning as pl
import pandas as pd
import timm
from tqdm import tqdm
from PIL import Image
from sklearn.model_selection import KFold
import math

from pytorch_lightning import Trainer, seed_everything
from pytorch_lightning.loggers import CSVLogger
from pytorch_lightning.callbacks import ModelCheckpoint, EarlyStopping
from torchcontrib.optim import SWA
from torchmetrics import Metric
from torch.utils.tensorboard import SummaryWriter

device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

```

```

/home/jw/anaconda3/envs/ss13/lib/python3.8/site-packages/tqdm/auto.py:21:
TqdmWarning: IProgress not found. Please update jupyter and ipywidgets. See
https://ipywidgets.readthedocs.io/en/stable/user_install.html
from .autonotebook import tqdm as notebook_tqdm

```

```

[2]: # Data preprocessing and augmentation
data_transforms = transforms.Compose([
    transforms.Resize((224, 224)),
    transforms.RandomHorizontalFlip(),
    transforms.RandomRotation(10),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])
])

train_dataset = ImageFolder('brain_tumor/Training', transform=data_transforms)
train_loader = DataLoader(train_dataset, batch_size=16, shuffle=True)

val_dataset = ImageFolder('brain_tumor/Testing', transform=data_transforms)

```

```

val_loader = DataLoader(val_dataset, batch_size=16, shuffle=False)

# Load a batch of images and labels for visualization
data_iter = iter(train_loader)
images, labels = next(data_iter)

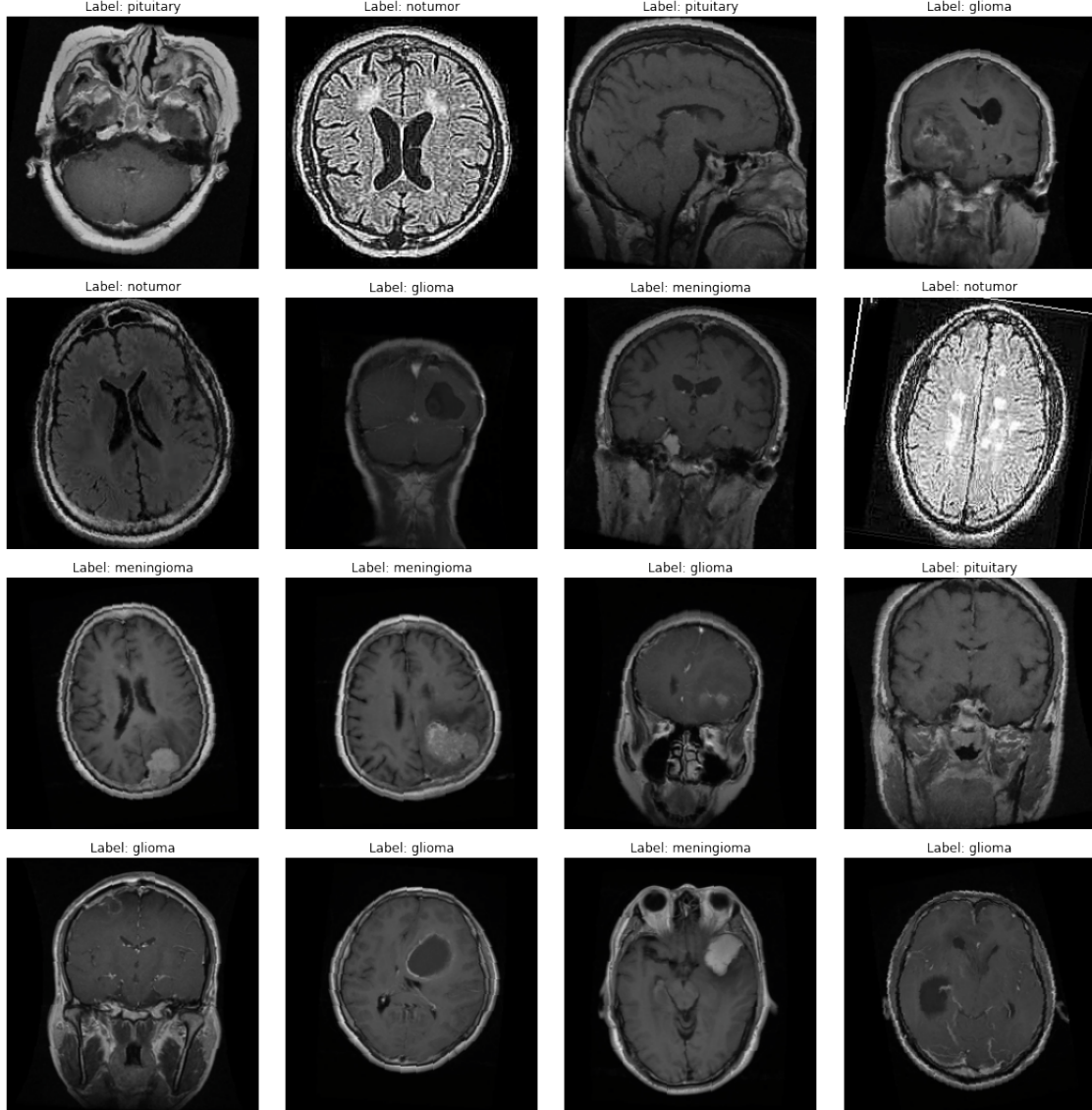
# Convert images to numpy arrays and denormalize
mean = np.array([0.485, 0.456, 0.406])
std = np.array([0.229, 0.224, 0.225])
images = (images.numpy().transpose((0, 2, 3, 1)) * std + mean).clip(0, 1)

# Create a grid of images
num_images = len(images)
rows = int(np.ceil(num_images / 4))
fig, axes = plt.subplots(rows, 4, figsize=(15, 15))

# Plot images with labels
for i, ax in enumerate(axes.flat):
    if i < num_images:
        ax.imshow(images[i])
        ax.set_title(f'Label: {train_dataset.classes[labels[i]]}')
        ax.axis('off')

plt.tight_layout()
plt.show()

```



### 3.2 Model

**Model Architecture** Layers and Types: CASS employs a dual architecture comprising a Convolutional Neural Network (CNN) and a Vision Transformer (ViT). Specific examples mentioned are ResNet-50 for the CNN and ViT Base/16 for the Transformer. Activation Functions: The paper does not explicitly mention the activation functions used in the model architectures, but standard practices for ResNet and ViT typically involve ReLU and GELU activations, respectively. ##### Training Objectives Loss Function: The model uses a cosine similarity-based loss function and focal loss, specifically designed for comparing the logits outputs from the CNN and ViT. ##### Optimizer: The training employs Adam optimizer with a learning rate of 1e-3 for both the CNN and ViT, along with stochastic weight averaging (SWA) for optimization.

**Weight of Each Loss Term:** Details about the weight of each loss term are not explicitly mentioned. However, the loss is computed as the mean value of all elements in the tensor derived from the cosine similarity calculation between the outputs of the CNN and ViT. ##### Others Pretraining: CASS is based on self-supervised learning, and it is mentioned that the models were trained from ImageNet initialization for 100 epochs.

**Training Process:** The paper describes the training process, mentioning that they use the same set of augmentations in self-supervised training and also detail the hyper-parameters for both the self-supervised and supervised training phases.

```
[3]: """
Define Focal-Loss
"""

class FocalLoss(nn.Module):
    """
    The focal loss for fighting against class-imbalance
    """
    def __init__(self, alpha=1, gamma=2):
        super(FocalLoss, self).__init__()
        self.alpha = alpha
        self.gamma = gamma
        self.epsilon = 1e-12 # prevent training from Nan-loss error
        self.cls_weights = torch.tensor([CFG.cls_weight], dtype=torch.float,
    ↪requires_grad=False, device=CFG.device)

    def forward(self, logits, target):
        """
        logits & target should be tensors with shape [batch_size, num_classes]
        """
        probs = torch.sigmoid(logits)
        one_subtract_probs = 1.0 - probs
        # add epsilon
        probs_new = probs + self.epsilon
        one_subtract_probs_new = one_subtract_probs + self.epsilon
        # calculate focal loss
        log_pt = target * torch.log(probs_new) + (1.0 - target) * torch.
    ↪log(one_subtract_probs_new)
        pt = torch.exp(log_pt)
        focal_loss = -1.0 * (self.alpha * (1 - pt) ** self.gamma) * log_pt
        focal_loss = focal_loss * self.cls_weights
        return torch.mean(focal_loss)
```

```
[4]: class CFG:
    # label_num2str = {'0': 'glioma', '1': 'meningioma', '2': 'notumor', '3':
    ↪'pituitary'}
    label_num2str = {0: 'glioma', 1: 'meningioma', 2: 'notumor', 3: 'pituitary'}
```

```

label_str2num = {'glioma': '0', 'meningioma': '1', 'notumor': '2', 'pituitary': '3'}
fl_alpha = 1.0 # alpha of focal_loss
fl_gamma = 2.0 # gamma of focal_loss
cls_weight = [0.5, 0.5, 0.5, 0.5]
cnn_name='resnet50'
vit_name='vit_base_patch16_224'
seed = 77
num_classes = 4
batch_size = 16
t_max = 16
lr = 1e-3
min_lr = 1e-6
n_fold = 6
num_workers = 8
gpu_idx = 0
device = torch.device(f'cuda:{gpu_idx}' if torch.cuda.is_available() else 'cpu')
gpu_list = [gpu_idx]
CNN = True
VIT = False

cfg=CFG()

```

```

[5]: model_cnn = timm.create_model(cfg.cnn_name, pretrained=True)
model_vit = timm.create_model(cfg.vit_name, pretrained=True)
model_cnn.to(device)
model_vit.to(device)

```

```

[5]: VisionTransformer(
  (patch_embed): PatchEmbed(
    (proj): Conv2d(3, 768, kernel_size=(16, 16), stride=(16, 16))
    (norm): Identity()
  )
  (pos_drop): Dropout(p=0.0, inplace=False)
  (patch_drop): Identity()
  (norm_pre): Identity()
  (blocks): Sequential(
    (0): Block(
      (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
      (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
      )
    )
  )
)

```

```

)
(ls1): Identity()
(drop_path1): Identity()
(norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
(mlp): Mlp(
  (fc1): Linear(in_features=768, out_features=3072, bias=True)
  (act): GELU(approximate='none')
  (drop1): Dropout(p=0.0, inplace=False)
  (norm): Identity()
  (fc2): Linear(in_features=3072, out_features=768, bias=True)
  (drop2): Dropout(p=0.0, inplace=False)
)
(ls2): Identity()
(drop_path2): Identity()
)
(1): Block(
  (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
  (attn): Attention(
    (qkv): Linear(in_features=768, out_features=2304, bias=True)
    (q_norm): Identity()
    (k_norm): Identity()
    (attn_drop): Dropout(p=0.0, inplace=False)
    (proj): Linear(in_features=768, out_features=768, bias=True)
    (proj_drop): Dropout(p=0.0, inplace=False)
  )
  (ls1): Identity()
  (drop_path1): Identity()
  (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
  (mlp): Mlp(
    (fc1): Linear(in_features=768, out_features=3072, bias=True)
    (act): GELU(approximate='none')
    (drop1): Dropout(p=0.0, inplace=False)
    (norm): Identity()
    (fc2): Linear(in_features=3072, out_features=768, bias=True)
    (drop2): Dropout(p=0.0, inplace=False)
  )
  (ls2): Identity()
  (drop_path2): Identity()
)
(2): Block(
  (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
  (attn): Attention(
    (qkv): Linear(in_features=768, out_features=2304, bias=True)
    (q_norm): Identity()
    (k_norm): Identity()
    (attn_drop): Dropout(p=0.0, inplace=False)
    (proj): Linear(in_features=768, out_features=768, bias=True)

```



```

        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(3): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(4): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)

```

```

        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(5): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(6): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()

```

```

        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(7): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(8): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()

```

```

        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(9): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(10): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)

```

```

        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
(11): Block(
    (norm1): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (attn): Attention(
        (qkv): Linear(in_features=768, out_features=2304, bias=True)
        (q_norm): Identity()
        (k_norm): Identity()
        (attn_drop): Dropout(p=0.0, inplace=False)
        (proj): Linear(in_features=768, out_features=768, bias=True)
        (proj_drop): Dropout(p=0.0, inplace=False)
    )
    (ls1): Identity()
    (drop_path1): Identity()
    (norm2): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
    (mlp): Mlp(
        (fc1): Linear(in_features=768, out_features=3072, bias=True)
        (act): GELU(approximate='none')
        (drop1): Dropout(p=0.0, inplace=False)
        (norm): Identity()
        (fc2): Linear(in_features=3072, out_features=768, bias=True)
        (drop2): Dropout(p=0.0, inplace=False)
    )
    (ls2): Identity()
    (drop_path2): Identity()
)
)
(norm): LayerNorm((768,), eps=1e-06, elementwise_affine=True)
(fc_norm): Identity()

```

```

(head_drop): Dropout(p=0.0, inplace=False)
(head): Linear(in_features=768, out_features=1000, bias=True)
)

```

### 3.3 Training

**Computational requirements and GPU Utilization:** The model was trained on a single NVIDIA RTX8000 GPU, which significantly facilitated a reduced training time. We use NVIDIA RXT2080ti to reproduce all the experiments.

**Training Time Efficiency:** The paper highlights that CASS took substantially less time compared to the DINO method for self-supervised training. For example, on the Autoimmune Diseases Biopsy Slides dataset, CASS required only 21 minutes compared to DINO’s 1 hour 13 minutes.

```

[6]: def
    ssl_train_model(train_loader,model_vit,criterion_vit,optimizer_vit,scheduler_vit,model_cnn,
    writer = SummaryWriter()
    phase = 'train'
    model_cnn.train()
    model_vit.train()
    f1_score_cnn=0
    f1_score_vit=0
    for i in tqdm(range(num_epochs)):
        with torch.set_grad_enabled(phase == 'train'):
            for img,_ in tqdm(train_loader):
                f1_score_cnn=0
                f1_score_vit=0
                img = img.to(device)
                pred_vit = model_vit(img)
                pred_cnn = model_cnn(img)
                model_sim_loss=loss_fn(pred_vit,pred_cnn)
                loss = model_sim_loss.mean()
                loss.backward()
                optimizer_cnn.step()
                optimizer_vit.step()
                scheduler_cnn.step()
                scheduler_vit.step()
            print('For -',i,'Loss:',loss)
            writer.add_scalar("Self-Supervised Loss/train", loss, i)
    writer.flush()

```

```

[7]: optimizer_cnn = SWA(torch.optim.Adam(model_cnn.parameters(), lr= 1e-3))
optimizer_vit = SWA(torch.optim.Adam(model_vit.parameters(), lr= 1e-3))
scheduler_cnn = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer_cnn,

```

```

T_max=16,
    ↪eta_min=1e-6)
scheduler_vit = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer_vit,
T_max=16,
    ↪eta_min=1e-6)

fl_alpha = 1.0 # alpha of focal_loss
fl_gamma = 2.0 # gamma of focal_loss
cls_weight = [0.5, 0.5, 0.5, 0.5]
criterion_vit = FocalLoss(fl_alpha, fl_gamma)
criterion_cnn = FocalLoss(fl_alpha, fl_gamma)

```

```

[8]: def loss_fn(x, y):
      x = torch.nn.functional.normalize(x, dim=-1, p=2)
      y = torch.nn.functional.normalize(y, dim=-1, p=2)
      return 2 - 2 * (x * y).sum(dim=-1)

```

```

[9]: ssl_train_model(train_loader,model_vit,criterion_vit,optimizer_vit,scheduler_vit,model_cnn,criterion_cnn)
      #Saving SSL Models
      print('Saving Cov-T')

      torch.save(model_cnn,'./cass-r50.pt')
      torch.save(model_vit,'./cass-vit.pt')

```

```

0%|          | 0/10 [00:00<?, ?it/s]
0%|          | 0/357 [00:00<?, ?it/s]
0%|          | 1/357 [00:01<08:08, 1.37s/it]
1%|          | 2/357 [00:01<04:05, 1.45it/s]
1%|          | 3/357 [00:01<02:57, 1.99it/s]
1%|          | 4/357 [00:02<02:26, 2.42it/s]
1%|          | 5/357 [00:02<02:08, 2.73it/s]
2%|          | 6/357 [00:02<01:58, 2.97it/s]
2%|          | 7/357 [00:02<01:52, 3.12it/s]
2%|          | 8/357 [00:03<01:48, 3.23it/s]
3%|          | 9/357 [00:03<01:46, 3.26it/s]
3%|          | 10/357 [00:03<01:47, 3.23it/s]
3%|          | 11/357 [00:04<01:46, 3.24it/s]
3%|          | 12/357 [00:04<01:46, 3.24it/s]
4%|          | 13/357 [00:04<01:46, 3.24it/s]
4%|          | 14/357 [00:05<02:19, 2.46it/s]
4%|          | 15/357 [00:05<02:07, 2.68it/s]
4%|          | 16/357 [00:06<02:00, 2.83it/s]
5%|          | 17/357 [00:06<01:54, 2.96it/s]
5%|          | 18/357 [00:06<01:48, 3.13it/s]

```

5%	19/357	[00:06<01:43,	3.26it/s]
6%	20/357	[00:07<01:40,	3.36it/s]
6%	21/357	[00:07<01:38,	3.42it/s]
6%	22/357	[00:07<01:36,	3.48it/s]
6%	23/357	[00:08<01:34,	3.52it/s]
7%	24/357	[00:08<01:33,	3.55it/s]
7%	25/357	[00:08<01:33,	3.56it/s]
7%	26/357	[00:08<01:32,	3.58it/s]
8%	27/357	[00:09<01:31,	3.59it/s]
8%	28/357	[00:09<01:31,	3.59it/s]
8%	29/357	[00:09<01:31,	3.59it/s]
8%	30/357	[00:09<01:31,	3.59it/s]
9%	31/357	[00:10<01:30,	3.58it/s]
9%	32/357	[00:10<01:30,	3.59it/s]
9%	33/357	[00:10<01:30,	3.60it/s]
10%	34/357	[00:11<01:29,	3.60it/s]
10%	35/357	[00:11<01:29,	3.61it/s]
10%	36/357	[00:11<01:29,	3.60it/s]
10%	37/357	[00:11<01:28,	3.60it/s]
11%	38/357	[00:12<01:28,	3.61it/s]
11%	39/357	[00:12<01:28,	3.61it/s]
11%	40/357	[00:12<01:27,	3.61it/s]
11%	41/357	[00:13<01:28,	3.56it/s]
12%	42/357	[00:13<01:27,	3.58it/s]
12%	43/357	[00:13<01:27,	3.59it/s]
12%	44/357	[00:13<01:26,	3.60it/s]
13%	45/357	[00:14<01:26,	3.60it/s]
13%	46/357	[00:14<01:26,	3.61it/s]
13%	47/357	[00:14<01:27,	3.55it/s]
13%	48/357	[00:14<01:26,	3.56it/s]
14%	49/357	[00:15<01:26,	3.58it/s]
14%	50/357	[00:15<01:25,	3.59it/s]
14%	51/357	[00:15<01:25,	3.59it/s]
15%	52/357	[00:16<01:24,	3.60it/s]
15%	53/357	[00:16<01:24,	3.59it/s]
15%	54/357	[00:16<01:24,	3.59it/s]
15%	55/357	[00:16<01:24,	3.58it/s]
16%	56/357	[00:17<01:23,	3.59it/s]
16%	57/357	[00:17<01:23,	3.60it/s]
16%	58/357	[00:17<01:22,	3.60it/s]
17%	59/357	[00:18<01:22,	3.60it/s]
17%	60/357	[00:18<01:22,	3.61it/s]
17%	61/357	[00:18<01:22,	3.61it/s]
17%	62/357	[00:18<01:21,	3.60it/s]
18%	63/357	[00:19<01:21,	3.60it/s]
18%	64/357	[00:19<01:21,	3.59it/s]
18%	65/357	[00:19<01:21,	3.58it/s]
18%	66/357	[00:19<01:21,	3.59it/s]



19%	67/357	[00:20<01:20,	3.60it/s]
19%	68/357	[00:20<01:20,	3.60it/s]
19%	69/357	[00:20<01:19,	3.60it/s]
20%	70/357	[00:21<01:19,	3.60it/s]
20%	71/357	[00:21<01:19,	3.61it/s]
20%	72/357	[00:21<01:18,	3.61it/s]
20%	73/357	[00:21<01:18,	3.60it/s]
21%	74/357	[00:22<01:18,	3.61it/s]
21%	75/357	[00:22<01:18,	3.60it/s]
21%	76/357	[00:22<01:18,	3.59it/s]
22%	77/357	[00:23<01:17,	3.59it/s]
22%	78/357	[00:23<01:17,	3.59it/s]
22%	79/357	[00:23<01:17,	3.59it/s]
22%	80/357	[00:24<01:38,	2.82it/s]
23%	81/357	[00:24<01:31,	3.02it/s]
23%	82/357	[00:24<01:26,	3.17it/s]
23%	83/357	[00:24<01:23,	3.28it/s]
24%	84/357	[00:25<01:20,	3.38it/s]
24%	85/357	[00:25<01:19,	3.44it/s]
24%	86/357	[00:25<01:17,	3.48it/s]
24%	87/357	[00:26<01:16,	3.52it/s]
25%	88/357	[00:26<01:15,	3.54it/s]
25%	89/357	[00:26<01:15,	3.56it/s]
25%	90/357	[00:26<01:14,	3.57it/s]
25%	91/357	[00:27<01:14,	3.57it/s]
26%	92/357	[00:27<01:14,	3.58it/s]
26%	93/357	[00:27<01:13,	3.59it/s]
26%	94/357	[00:28<01:13,	3.59it/s]
27%	95/357	[00:28<01:13,	3.58it/s]
27%	96/357	[00:28<01:12,	3.59it/s]
27%	97/357	[00:28<01:12,	3.59it/s]
27%	98/357	[00:29<01:12,	3.59it/s]
28%	99/357	[00:29<01:11,	3.60it/s]
28%	100/357	[00:29<01:11,	3.60it/s]
28%	101/357	[00:29<01:11,	3.60it/s]
29%	102/357	[00:30<01:10,	3.60it/s]
29%	103/357	[00:30<01:10,	3.60it/s]
29%	104/357	[00:30<01:10,	3.60it/s]
29%	105/357	[00:31<01:11,	3.54it/s]
30%	106/357	[00:31<01:10,	3.55it/s]
30%	107/357	[00:31<01:09,	3.58it/s]
30%	108/357	[00:31<01:09,	3.60it/s]
31%	109/357	[00:32<01:08,	3.61it/s]
31%	110/357	[00:32<01:08,	3.62it/s]
31%	111/357	[00:32<01:07,	3.62it/s]
31%	112/357	[00:33<01:07,	3.63it/s]
32%	113/357	[00:33<01:07,	3.63it/s]
32%	114/357	[00:33<01:06,	3.63it/s]

32%	115/357	[00:33<01:06,	3.63it/s]
32%	116/357	[00:34<01:06,	3.63it/s]
33%	117/357	[00:34<01:06,	3.63it/s]
33%	118/357	[00:34<01:05,	3.63it/s]
33%	119/357	[00:34<01:05,	3.63it/s]
34%	120/357	[00:35<01:05,	3.63it/s]
34%	121/357	[00:35<01:04,	3.63it/s]
34%	122/357	[00:35<01:04,	3.63it/s]
34%	123/357	[00:36<01:04,	3.63it/s]
35%	124/357	[00:36<01:04,	3.63it/s]
35%	125/357	[00:36<01:03,	3.63it/s]
35%	126/357	[00:36<01:03,	3.63it/s]
36%	127/357	[00:37<01:03,	3.63it/s]
36%	128/357	[00:37<01:03,	3.63it/s]
36%	129/357	[00:37<01:02,	3.63it/s]
36%	130/357	[00:37<01:02,	3.63it/s]
37%	131/357	[00:38<01:02,	3.64it/s]
37%	132/357	[00:38<01:01,	3.63it/s]
37%	133/357	[00:38<01:01,	3.64it/s]
38%	134/357	[00:39<01:01,	3.63it/s]
38%	135/357	[00:39<01:01,	3.63it/s]
38%	136/357	[00:39<01:00,	3.64it/s]
38%	137/357	[00:39<01:00,	3.63it/s]
39%	138/357	[00:40<01:00,	3.64it/s]
39%	139/357	[00:40<00:59,	3.63it/s]
39%	140/357	[00:40<00:59,	3.64it/s]
39%	141/357	[00:40<00:59,	3.64it/s]
40%	142/357	[00:41<00:59,	3.64it/s]
40%	143/357	[00:41<00:58,	3.64it/s]
40%	144/357	[00:41<00:58,	3.64it/s]
41%	145/357	[00:42<00:58,	3.64it/s]
41%	146/357	[00:42<00:58,	3.63it/s]
41%	147/357	[00:42<00:57,	3.64it/s]
41%	148/357	[00:42<00:57,	3.63it/s]
42%	149/357	[00:43<00:57,	3.63it/s]
42%	150/357	[00:43<00:56,	3.63it/s]
42%	151/357	[00:43<00:56,	3.63it/s]
43%	152/357	[00:44<00:56,	3.64it/s]
43%	153/357	[00:44<00:56,	3.64it/s]
43%	154/357	[00:44<00:55,	3.64it/s]
43%	155/357	[00:44<00:55,	3.63it/s]
44%	156/357	[00:45<00:55,	3.64it/s]
44%	157/357	[00:45<00:55,	3.64it/s]
44%	158/357	[00:45<00:54,	3.64it/s]
45%	159/357	[00:45<00:54,	3.64it/s]
45%	160/357	[00:46<00:54,	3.64it/s]
45%	161/357	[00:46<00:53,	3.64it/s]
45%	162/357	[00:46<00:53,	3.64it/s]

46%	163/357	[00:47<00:53,	3.64it/s]
46%	164/357	[00:47<00:53,	3.64it/s]
46%	165/357	[00:47<00:52,	3.64it/s]
46%	166/357	[00:47<00:52,	3.63it/s]
47%	167/357	[00:48<00:52,	3.63it/s]
47%	168/357	[00:48<00:51,	3.64it/s]
47%	169/357	[00:48<00:51,	3.63it/s]
48%	170/357	[00:49<00:53,	3.51it/s]
48%	171/357	[00:49<00:52,	3.55it/s]
48%	172/357	[00:49<00:51,	3.58it/s]
48%	173/357	[00:49<00:51,	3.60it/s]
49%	174/357	[00:50<00:50,	3.61it/s]
49%	175/357	[00:50<00:50,	3.62it/s]
49%	176/357	[00:50<00:49,	3.62it/s]
50%	177/357	[00:51<00:56,	3.21it/s]
50%	178/357	[00:51<00:58,	3.08it/s]
50%	179/357	[00:51<00:59,	3.00it/s]
50%	180/357	[00:52<00:57,	3.05it/s]
51%	181/357	[00:52<00:57,	3.04it/s]
51%	182/357	[00:52<00:56,	3.10it/s]
51%	183/357	[00:53<00:56,	3.10it/s]
52%	184/357	[00:53<00:53,	3.23it/s]
52%	185/357	[00:53<00:51,	3.34it/s]
52%	186/357	[00:53<00:49,	3.43it/s]
52%	187/357	[00:54<00:48,	3.49it/s]
53%	188/357	[00:54<00:48,	3.48it/s]
53%	189/357	[00:54<00:48,	3.49it/s]
53%	190/357	[00:54<00:47,	3.52it/s]
54%	191/357	[00:55<00:47,	3.52it/s]
54%	192/357	[00:55<00:47,	3.50it/s]
54%	193/357	[00:55<00:48,	3.39it/s]
54%	194/357	[00:56<00:48,	3.35it/s]
55%	195/357	[00:56<00:49,	3.28it/s]
55%	196/357	[00:56<00:50,	3.22it/s]
55%	197/357	[00:57<00:52,	3.06it/s]
55%	198/357	[00:57<01:11,	2.22it/s]
56%	199/357	[00:58<01:05,	2.43it/s]
56%	200/357	[00:58<01:00,	2.60it/s]
56%	201/357	[00:58<00:56,	2.76it/s]
57%	202/357	[00:59<00:53,	2.89it/s]
57%	203/357	[00:59<00:51,	2.97it/s]
57%	204/357	[00:59<00:50,	3.03it/s]
57%	205/357	[01:00<00:49,	3.09it/s]
58%	206/357	[01:00<00:48,	3.12it/s]
58%	207/357	[01:00<00:47,	3.15it/s]
58%	208/357	[01:01<00:46,	3.18it/s]
59%	209/357	[01:01<00:46,	3.20it/s]
59%	210/357	[01:01<00:46,	3.19it/s]

59%	211/357	[01:01<00:45,	3.21it/s]
59%	212/357	[01:02<00:45,	3.21it/s]
60%	213/357	[01:02<00:44,	3.22it/s]
60%	214/357	[01:02<00:44,	3.21it/s]
60%	215/357	[01:03<00:44,	3.19it/s]
61%	216/357	[01:03<00:44,	3.20it/s]
61%	217/357	[01:03<00:46,	3.02it/s]
61%	218/357	[01:04<00:45,	3.04it/s]
61%	219/357	[01:04<00:45,	3.06it/s]
62%	220/357	[01:04<00:44,	3.10it/s]
62%	221/357	[01:05<00:43,	3.13it/s]
62%	222/357	[01:05<00:42,	3.15it/s]
62%	223/357	[01:05<00:42,	3.16it/s]
63%	224/357	[01:06<00:44,	2.99it/s]
63%	225/357	[01:06<00:43,	3.02it/s]
63%	226/357	[01:06<00:43,	3.04it/s]
64%	227/357	[01:07<00:43,	2.99it/s]
64%	228/357	[01:07<00:42,	3.06it/s]
64%	229/357	[01:07<00:41,	3.10it/s]
64%	230/357	[01:08<00:41,	3.08it/s]
65%	231/357	[01:08<00:40,	3.12it/s]
65%	232/357	[01:08<00:39,	3.15it/s]
65%	233/357	[01:09<00:39,	3.14it/s]
66%	234/357	[01:09<00:39,	3.08it/s]
66%	235/357	[01:09<00:39,	3.12it/s]
66%	236/357	[01:10<00:38,	3.13it/s]
66%	237/357	[01:10<00:38,	3.15it/s]
67%	238/357	[01:10<00:38,	3.12it/s]
67%	239/357	[01:11<00:37,	3.11it/s]
67%	240/357	[01:11<00:38,	3.05it/s]
68%	241/357	[01:11<00:40,	2.88it/s]
68%	242/357	[01:12<00:45,	2.52it/s]
68%	243/357	[01:12<00:42,	2.69it/s]
68%	244/357	[01:13<00:43,	2.59it/s]
69%	245/357	[01:13<00:41,	2.69it/s]
69%	246/357	[01:13<00:40,	2.76it/s]
69%	247/357	[01:14<00:38,	2.86it/s]
69%	248/357	[01:14<00:39,	2.79it/s]
70%	249/357	[01:14<00:38,	2.84it/s]
70%	250/357	[01:15<00:36,	2.94it/s]
70%	251/357	[01:15<00:36,	2.90it/s]
71%	252/357	[01:15<00:35,	2.99it/s]
71%	253/357	[01:16<00:34,	3.01it/s]
71%	254/357	[01:16<00:33,	3.08it/s]
71%	255/357	[01:16<00:33,	3.01it/s]
72%	256/357	[01:17<00:33,	3.06it/s]
72%	257/357	[01:17<00:32,	3.10it/s]
72%	258/357	[01:17<00:31,	3.13it/s]

73%		259/357	[01:17<00:32,	3.04it/s]
73%		260/357	[01:18<00:32,	2.96it/s]
73%		261/357	[01:18<00:31,	3.03it/s]
73%		262/357	[01:19<00:33,	2.84it/s]
74%		263/357	[01:19<00:33,	2.77it/s]
74%		264/357	[01:19<00:32,	2.89it/s]
74%		265/357	[01:20<00:31,	2.93it/s]
75%		266/357	[01:20<00:30,	3.02it/s]
75%		267/357	[01:20<00:29,	3.06it/s]
75%		268/357	[01:21<00:28,	3.09it/s]
75%		269/357	[01:21<00:28,	3.14it/s]
76%		270/357	[01:21<00:27,	3.16it/s]
76%		271/357	[01:22<00:29,	2.87it/s]
76%		272/357	[01:22<00:28,	2.94it/s]
76%		273/357	[01:22<00:27,	3.01it/s]
77%		274/357	[01:23<00:27,	3.06it/s]
77%		275/357	[01:23<00:26,	3.11it/s]
77%		276/357	[01:23<00:25,	3.14it/s]
78%		277/357	[01:23<00:25,	3.15it/s]
78%		278/357	[01:24<00:25,	3.16it/s]
78%		279/357	[01:24<00:31,	2.45it/s]
78%		280/357	[01:25<00:33,	2.33it/s]
79%		281/357	[01:25<00:30,	2.47it/s]
79%		282/357	[01:26<00:29,	2.58it/s]
79%		283/357	[01:26<00:27,	2.72it/s]
80%		284/357	[01:26<00:25,	2.84it/s]
80%		285/357	[01:27<00:24,	2.88it/s]
80%		286/357	[01:27<00:24,	2.92it/s]
80%		287/357	[01:27<00:23,	2.98it/s]
81%		288/357	[01:27<00:22,	3.05it/s]
81%		289/357	[01:28<00:22,	2.99it/s]
81%		290/357	[01:28<00:21,	3.06it/s]
82%		291/357	[01:29<00:24,	2.75it/s]
82%		292/357	[01:29<00:22,	2.88it/s]
82%		293/357	[01:29<00:21,	2.95it/s]
82%		294/357	[01:30<00:21,	2.97it/s]
83%		295/357	[01:30<00:25,	2.43it/s]
83%		296/357	[01:31<00:27,	2.24it/s]
83%		297/357	[01:31<00:24,	2.41it/s]
83%		298/357	[01:31<00:24,	2.44it/s]
84%		299/357	[01:33<00:36,	1.60it/s]
84%		300/357	[01:33<00:36,	1.56it/s]
84%		301/357	[01:34<00:30,	1.83it/s]
85%		302/357	[01:34<00:25,	2.15it/s]
85%		303/357	[01:34<00:22,	2.41it/s]
85%		304/357	[01:34<00:20,	2.62it/s]
85%		305/357	[01:35<00:18,	2.83it/s]
86%		306/357	[01:35<00:17,	2.98it/s]

86%		307/357	[01:35<00:16,	3.08it/s]
86%		308/357	[01:36<00:15,	3.14it/s]
87%		309/357	[01:36<00:16,	2.94it/s]
87%		310/357	[01:36<00:15,	3.00it/s]
87%		311/357	[01:37<00:14,	3.11it/s]
87%		312/357	[01:37<00:14,	3.19it/s]
88%		313/357	[01:37<00:13,	3.20it/s]
88%		314/357	[01:38<00:13,	3.25it/s]
88%		315/357	[01:38<00:12,	3.31it/s]
89%		316/357	[01:38<00:12,	3.26it/s]
89%		317/357	[01:38<00:12,	3.30it/s]
89%		318/357	[01:39<00:11,	3.38it/s]
89%		319/357	[01:39<00:11,	3.43it/s]
90%		320/357	[01:39<00:10,	3.47it/s]
90%		321/357	[01:40<00:10,	3.44it/s]
90%		322/357	[01:40<00:10,	3.42it/s]
90%		323/357	[01:40<00:09,	3.43it/s]
91%		324/357	[01:40<00:09,	3.38it/s]
91%		325/357	[01:41<00:09,	3.39it/s]
91%		326/357	[01:41<00:09,	3.44it/s]
92%		327/357	[01:41<00:08,	3.49it/s]
92%		328/357	[01:42<00:08,	3.50it/s]
92%		329/357	[01:42<00:08,	3.25it/s]
92%		330/357	[01:42<00:08,	3.23it/s]
93%		331/357	[01:43<00:07,	3.25it/s]
93%		332/357	[01:43<00:07,	3.31it/s]
93%		333/357	[01:43<00:07,	3.34it/s]
94%		334/357	[01:43<00:06,	3.36it/s]
94%		335/357	[01:44<00:06,	3.35it/s]
94%		336/357	[01:44<00:06,	3.37it/s]
94%		337/357	[01:44<00:05,	3.38it/s]
95%		338/357	[01:45<00:05,	3.36it/s]
95%		339/357	[01:45<00:05,	3.35it/s]
95%		340/357	[01:45<00:05,	3.31it/s]
96%		341/357	[01:46<00:04,	3.30it/s]
96%		342/357	[01:46<00:04,	3.34it/s]
96%		343/357	[01:46<00:04,	3.33it/s]
96%		344/357	[01:46<00:03,	3.32it/s]
97%		345/357	[01:47<00:03,	3.29it/s]
97%		346/357	[01:47<00:03,	3.23it/s]
97%		347/357	[01:48<00:03,	2.85it/s]
97%		348/357	[01:48<00:03,	2.93it/s]
98%		349/357	[01:48<00:02,	2.96it/s]
98%		350/357	[01:49<00:02,	2.55it/s]
98%		351/357	[01:49<00:02,	2.20it/s]
99%		352/357	[01:50<00:02,	2.27it/s]
99%		353/357	[01:50<00:01,	2.50it/s]
99%		354/357	[01:50<00:01,	2.69it/s]

```

99%|      | 355/357 [01:51<00:00, 2.84it/s]
100%|     | 356/357 [01:51<00:00, 2.97it/s]
100%|     | 357/357 [01:51<00:00, 3.20it/s]
10%|      | 1/10 [01:51<16:47, 111.89s/it]

```

```
For - 0 Loss: tensor(0.0024, device='cuda:0', grad_fn=<MeanBackward0>)
```

```

0%|      | 0/357 [00:00<?, ?it/s]
0%|      | 1/357 [00:00<01:39, 3.57it/s]
1%|      | 2/357 [00:00<01:38, 3.62it/s]
1%|      | 3/357 [00:00<01:37, 3.63it/s]
1%|      | 4/357 [00:01<01:36, 3.64it/s]
1%|      | 5/357 [00:01<01:36, 3.65it/s]
2%|      | 6/357 [00:01<01:36, 3.65it/s]
2%|      | 7/357 [00:01<01:36, 3.64it/s]
2%|      | 8/357 [00:02<01:39, 3.50it/s]
3%|      | 9/357 [00:02<01:42, 3.39it/s]
3%|      | 10/357 [00:02<01:42, 3.37it/s]
3%|      | 11/357 [00:03<01:42, 3.36it/s]
3%|      | 12/357 [00:03<01:45, 3.28it/s]
4%|      | 13/357 [00:03<01:47, 3.19it/s]
4%|      | 14/357 [00:04<01:48, 3.15it/s]
4%|      | 15/357 [00:04<01:49, 3.13it/s]
4%|      | 16/357 [00:05<02:46, 2.05it/s]
5%|      | 17/357 [00:05<02:42, 2.09it/s]
5%|      | 18/357 [00:06<02:25, 2.32it/s]
5%|      | 19/357 [00:06<02:13, 2.53it/s]
6%|      | 20/357 [00:06<02:05, 2.68it/s]
6%|      | 21/357 [00:07<01:59, 2.82it/s]
6%|      | 22/357 [00:07<01:54, 2.93it/s]
6%|      | 23/357 [00:07<01:50, 3.02it/s]
7%|      | 24/357 [00:07<01:48, 3.08it/s]
7%|      | 25/357 [00:08<01:46, 3.12it/s]
7%|      | 26/357 [00:08<01:44, 3.16it/s]
8%|      | 27/357 [00:08<01:43, 3.19it/s]
8%|      | 28/357 [00:09<01:42, 3.22it/s]
8%|      | 29/357 [00:09<01:41, 3.23it/s]
8%|      | 30/357 [00:09<01:40, 3.24it/s]
9%|      | 31/357 [00:10<01:40, 3.24it/s]
9%|      | 32/357 [00:10<01:40, 3.24it/s]
9%|      | 33/357 [00:10<01:40, 3.23it/s]
10%|     | 34/357 [00:11<01:40, 3.21it/s]
10%|     | 35/357 [00:11<01:40, 3.19it/s]
10%|     | 36/357 [00:11<01:40, 3.20it/s]
10%|     | 37/357 [00:12<01:39, 3.20it/s]
11%|     | 38/357 [00:12<01:39, 3.20it/s]
11%|     | 39/357 [00:12<01:39, 3.21it/s]
11%|     | 40/357 [00:12<01:38, 3.22it/s]

```

11%		41/357	[00:13<01:38,	3.21it/s]
12%		42/357	[00:13<01:37,	3.22it/s]
12%		43/357	[00:13<01:37,	3.21it/s]
12%		44/357	[00:14<01:37,	3.22it/s]
13%		45/357	[00:14<01:37,	3.21it/s]
13%		46/357	[00:14<01:36,	3.22it/s]
13%		47/357	[00:15<01:36,	3.22it/s]
13%		48/357	[00:15<01:36,	3.21it/s]
14%		49/357	[00:15<01:35,	3.21it/s]
14%		50/357	[00:16<01:35,	3.21it/s]
14%		51/357	[00:16<01:35,	3.20it/s]
15%		52/357	[00:16<01:35,	3.21it/s]
15%		53/357	[00:16<01:34,	3.21it/s]
15%		54/357	[00:17<01:34,	3.21it/s]
15%		55/357	[00:17<01:33,	3.22it/s]
16%		56/357	[00:17<01:33,	3.21it/s]
16%		57/357	[00:18<01:33,	3.22it/s]
16%		58/357	[00:18<01:32,	3.22it/s]
17%		59/357	[00:18<01:32,	3.22it/s]
17%		60/357	[00:19<01:32,	3.22it/s]
17%		61/357	[00:19<01:32,	3.19it/s]
17%		62/357	[00:19<01:33,	3.15it/s]
18%		63/357	[00:20<01:33,	3.16it/s]
18%		64/357	[00:20<01:34,	3.09it/s]
18%		65/357	[00:20<01:30,	3.22it/s]
18%		66/357	[00:21<01:27,	3.33it/s]
19%		67/357	[00:21<01:24,	3.42it/s]
19%		68/357	[00:21<01:22,	3.49it/s]
19%		69/357	[00:21<01:21,	3.54it/s]
20%		70/357	[00:22<01:20,	3.57it/s]
20%		71/357	[00:22<01:19,	3.60it/s]
20%		72/357	[00:22<01:19,	3.59it/s]
20%		73/357	[00:22<01:19,	3.56it/s]
21%		74/357	[00:23<01:20,	3.53it/s]
21%		75/357	[00:23<01:20,	3.51it/s]
21%		76/357	[00:23<01:19,	3.53it/s]
22%		77/357	[00:24<01:18,	3.57it/s]
22%		78/357	[00:24<01:17,	3.59it/s]
22%		79/357	[00:24<01:16,	3.61it/s]
22%		80/357	[00:24<01:16,	3.61it/s]
23%		81/357	[00:25<01:16,	3.62it/s]
23%		82/357	[00:25<01:15,	3.63it/s]
23%		83/357	[00:25<01:15,	3.63it/s]
24%		84/357	[00:26<01:15,	3.63it/s]
24%		85/357	[00:26<01:14,	3.63it/s]
24%		86/357	[00:26<01:14,	3.64it/s]
24%		87/357	[00:26<01:14,	3.63it/s]
25%		88/357	[00:27<01:13,	3.64it/s]



25%	89/357	[00:27<01:13,	3.64it/s]
25%	90/357	[00:27<01:13,	3.64it/s]
25%	91/357	[00:27<01:13,	3.64it/s]
26%	92/357	[00:28<01:12,	3.65it/s]
26%	93/357	[00:28<01:12,	3.65it/s]
26%	94/357	[00:28<01:12,	3.65it/s]
27%	95/357	[00:29<01:11,	3.65it/s]
27%	96/357	[00:29<01:11,	3.65it/s]
27%	97/357	[00:29<01:11,	3.65it/s]
27%	98/357	[00:29<01:11,	3.65it/s]
28%	99/357	[00:30<01:10,	3.65it/s]
28%	100/357	[00:30<01:10,	3.65it/s]
28%	101/357	[00:30<01:10,	3.65it/s]
29%	102/357	[00:30<01:09,	3.65it/s]
29%	103/357	[00:31<01:09,	3.65it/s]
29%	104/357	[00:31<01:09,	3.65it/s]
29%	105/357	[00:31<01:09,	3.65it/s]
30%	106/357	[00:32<01:08,	3.65it/s]
30%	107/357	[00:32<01:08,	3.65it/s]
30%	108/357	[00:32<01:08,	3.65it/s]
31%	109/357	[00:32<01:07,	3.65it/s]
31%	110/357	[00:33<01:12,	3.40it/s]
31%	111/357	[00:33<01:10,	3.47it/s]
31%	112/357	[00:33<01:09,	3.53it/s]
32%	113/357	[00:34<01:08,	3.56it/s]
32%	114/357	[00:34<01:07,	3.59it/s]
32%	115/357	[00:34<01:07,	3.61it/s]
32%	116/357	[00:34<01:06,	3.62it/s]
33%	117/357	[00:35<01:06,	3.63it/s]
33%	118/357	[00:35<01:05,	3.64it/s]
33%	119/357	[00:35<01:05,	3.64it/s]
34%	120/357	[00:35<01:04,	3.65it/s]
34%	121/357	[00:36<01:04,	3.65it/s]
34%	122/357	[00:36<01:04,	3.65it/s]
34%	123/357	[00:36<01:04,	3.65it/s]
35%	124/357	[00:37<01:03,	3.65it/s]
35%	125/357	[00:37<01:04,	3.62it/s]
35%	126/357	[00:37<01:03,	3.63it/s]
36%	127/357	[00:37<01:03,	3.64it/s]
36%	128/357	[00:38<01:02,	3.64it/s]
36%	129/357	[00:38<01:02,	3.65it/s]
36%	130/357	[00:38<01:02,	3.65it/s]
37%	131/357	[00:38<01:01,	3.65it/s]
37%	132/357	[00:39<01:01,	3.65it/s]
37%	133/357	[00:39<01:01,	3.65it/s]
38%	134/357	[00:39<01:01,	3.65it/s]
38%	135/357	[00:40<01:00,	3.65it/s]
38%	136/357	[00:40<01:00,	3.65it/s]

38%	137/357	[00:40<01:00,	3.65it/s]
39%	138/357	[00:40<01:00,	3.65it/s]
39%	139/357	[00:41<00:59,	3.65it/s]
39%	140/357	[00:41<00:59,	3.65it/s]
39%	141/357	[00:41<00:59,	3.66it/s]
40%	142/357	[00:41<00:58,	3.65it/s]
40%	143/357	[00:42<00:58,	3.65it/s]
40%	144/357	[00:42<00:58,	3.64it/s]
41%	145/357	[00:42<00:58,	3.64it/s]
41%	146/357	[00:43<00:57,	3.64it/s]
41%	147/357	[00:43<00:57,	3.64it/s]
41%	148/357	[00:43<00:57,	3.64it/s]
42%	149/357	[00:43<00:57,	3.64it/s]
42%	150/357	[00:44<00:56,	3.65it/s]
42%	151/357	[00:44<00:56,	3.65it/s]
43%	152/357	[00:44<00:56,	3.65it/s]
43%	153/357	[00:44<00:55,	3.65it/s]
43%	154/357	[00:45<00:55,	3.65it/s]
43%	155/357	[00:45<00:55,	3.65it/s]
44%	156/357	[00:45<00:55,	3.65it/s]
44%	157/357	[00:46<00:54,	3.65it/s]
44%	158/357	[00:46<00:54,	3.65it/s]
45%	159/357	[00:46<00:54,	3.65it/s]
45%	160/357	[00:46<00:53,	3.65it/s]
45%	161/357	[00:47<00:53,	3.64it/s]
45%	162/357	[00:47<00:53,	3.64it/s]
46%	163/357	[00:47<00:53,	3.64it/s]
46%	164/357	[00:48<00:52,	3.64it/s]
46%	165/357	[00:48<00:52,	3.64it/s]
46%	166/357	[00:48<00:52,	3.65it/s]
47%	167/357	[00:48<00:52,	3.65it/s]
47%	168/357	[00:49<00:51,	3.65it/s]
47%	169/357	[00:49<00:51,	3.65it/s]
48%	170/357	[00:49<00:51,	3.65it/s]
48%	171/357	[00:49<00:51,	3.65it/s]
48%	172/357	[00:50<01:31,	2.02it/s]
48%	173/357	[00:51<01:18,	2.34it/s]
49%	174/357	[00:51<01:09,	2.62it/s]
49%	175/357	[00:51<01:03,	2.86it/s]
49%	176/357	[00:52<00:59,	3.06it/s]
50%	177/357	[00:52<00:55,	3.22it/s]
50%	178/357	[00:52<00:53,	3.34it/s]
50%	179/357	[00:52<00:51,	3.43it/s]
50%	180/357	[00:53<00:50,	3.49it/s]
51%	181/357	[00:53<00:49,	3.54it/s]
51%	182/357	[00:53<00:49,	3.50it/s]
51%	183/357	[00:53<00:49,	3.55it/s]
52%	184/357	[00:54<00:48,	3.58it/s]

52%	185/357	[00:54<00:47,	3.60it/s]
52%	186/357	[00:54<00:47,	3.61it/s]
52%	187/357	[00:55<00:46,	3.62it/s]
53%	188/357	[00:55<00:46,	3.63it/s]
53%	189/357	[00:55<00:46,	3.64it/s]
53%	190/357	[00:55<00:46,	3.63it/s]
54%	191/357	[00:56<00:45,	3.63it/s]
54%	192/357	[00:56<00:45,	3.59it/s]
54%	193/357	[00:56<00:45,	3.61it/s]
54%	194/357	[00:57<00:48,	3.38it/s]
55%	195/357	[00:57<00:46,	3.46it/s]
55%	196/357	[00:57<00:46,	3.50it/s]
55%	197/357	[00:57<00:45,	3.54it/s]
55%	198/357	[00:58<00:44,	3.58it/s]
56%	199/357	[00:58<00:43,	3.60it/s]
56%	200/357	[00:58<00:43,	3.62it/s]
56%	201/357	[00:58<00:44,	3.53it/s]
57%	202/357	[00:59<00:43,	3.56it/s]
57%	203/357	[00:59<00:42,	3.59it/s]
57%	204/357	[00:59<00:42,	3.61it/s]
57%	205/357	[01:00<00:41,	3.62it/s]
58%	206/357	[01:00<00:41,	3.64it/s]
58%	207/357	[01:00<00:41,	3.64it/s]
58%	208/357	[01:00<00:40,	3.64it/s]
59%	209/357	[01:01<00:40,	3.65it/s]
59%	210/357	[01:01<00:40,	3.65it/s]
59%	211/357	[01:01<00:39,	3.65it/s]
59%	212/357	[01:02<00:39,	3.65it/s]
60%	213/357	[01:02<00:39,	3.65it/s]
60%	214/357	[01:02<00:39,	3.65it/s]
60%	215/357	[01:02<00:38,	3.65it/s]
61%	216/357	[01:03<00:38,	3.65it/s]
61%	217/357	[01:03<00:38,	3.65it/s]
61%	218/357	[01:03<00:38,	3.66it/s]
61%	219/357	[01:03<00:37,	3.65it/s]
62%	220/357	[01:04<00:37,	3.65it/s]
62%	221/357	[01:04<00:37,	3.65it/s]
62%	222/357	[01:04<00:36,	3.65it/s]
62%	223/357	[01:05<00:36,	3.65it/s]
63%	224/357	[01:05<00:36,	3.65it/s]
63%	225/357	[01:05<00:36,	3.65it/s]
63%	226/357	[01:05<00:36,	3.62it/s]
64%	227/357	[01:06<00:35,	3.63it/s]
64%	228/357	[01:06<00:35,	3.63it/s]
64%	229/357	[01:06<00:35,	3.63it/s]
64%	230/357	[01:06<00:34,	3.64it/s]
65%	231/357	[01:07<00:34,	3.65it/s]
65%	232/357	[01:07<00:34,	3.65it/s]

65%		233/357	[01:07<00:33,	3.65it/s]
66%		234/357	[01:08<00:33,	3.66it/s]
66%		235/357	[01:08<00:33,	3.65it/s]
66%		236/357	[01:08<00:33,	3.60it/s]
66%		237/357	[01:08<00:33,	3.60it/s]
67%		238/357	[01:09<00:33,	3.52it/s]
67%		239/357	[01:09<00:33,	3.56it/s]
67%		240/357	[01:09<00:32,	3.59it/s]
68%		241/357	[01:10<00:32,	3.59it/s]
68%		242/357	[01:10<00:31,	3.61it/s]
68%		243/357	[01:10<00:31,	3.61it/s]
68%		244/357	[01:10<00:31,	3.61it/s]
69%		245/357	[01:11<00:30,	3.62it/s]
69%		246/357	[01:11<00:30,	3.61it/s]
69%		247/357	[01:11<00:30,	3.62it/s]
69%		248/357	[01:11<00:30,	3.63it/s]
70%		249/357	[01:12<00:29,	3.64it/s]
70%		250/357	[01:12<00:29,	3.64it/s]
70%		251/357	[01:12<00:29,	3.65it/s]
71%		252/357	[01:13<00:28,	3.65it/s]
71%		253/357	[01:13<00:28,	3.65it/s]
71%		254/357	[01:13<00:28,	3.65it/s]
71%		255/357	[01:13<00:27,	3.65it/s]
72%		256/357	[01:14<00:27,	3.63it/s]
72%		257/357	[01:14<00:27,	3.63it/s]
72%		258/357	[01:14<00:27,	3.63it/s]
73%		259/357	[01:14<00:26,	3.64it/s]
73%		260/357	[01:15<00:26,	3.64it/s]
73%		261/357	[01:15<00:26,	3.62it/s]
73%		262/357	[01:15<00:26,	3.63it/s]
74%		263/357	[01:16<00:25,	3.64it/s]
74%		264/357	[01:16<00:25,	3.64it/s]
74%		265/357	[01:16<00:25,	3.64it/s]
75%		266/357	[01:16<00:25,	3.64it/s]
75%		267/357	[01:17<00:25,	3.57it/s]
75%		268/357	[01:17<00:24,	3.60it/s]
75%		269/357	[01:17<00:24,	3.61it/s]
76%		270/357	[01:17<00:24,	3.62it/s]
76%		271/357	[01:18<00:23,	3.59it/s]
76%		272/357	[01:18<00:23,	3.61it/s]
76%		273/357	[01:18<00:23,	3.60it/s]
77%		274/357	[01:19<00:22,	3.62it/s]
77%		275/357	[01:19<00:22,	3.61it/s]
77%		276/357	[01:19<00:22,	3.58it/s]
78%		277/357	[01:19<00:22,	3.60it/s]
78%		278/357	[01:20<00:21,	3.62it/s]
78%		279/357	[01:20<00:21,	3.60it/s]
78%		280/357	[01:20<00:21,	3.59it/s]

79%	281/357	[01:21<00:22,	3.44it/s]
79%	282/357	[01:21<00:21,	3.49it/s]
79%	283/357	[01:21<00:21,	3.51it/s]
80%	284/357	[01:21<00:20,	3.54it/s]
80%	285/357	[01:22<00:20,	3.52it/s]
80%	286/357	[01:22<00:20,	3.52it/s]
80%	287/357	[01:22<00:19,	3.53it/s]
81%	288/357	[01:23<00:19,	3.56it/s]
81%	289/357	[01:23<00:19,	3.53it/s]
81%	290/357	[01:23<00:18,	3.55it/s]
82%	291/357	[01:23<00:18,	3.58it/s]
82%	292/357	[01:24<00:18,	3.57it/s]
82%	293/357	[01:24<00:17,	3.58it/s]
82%	294/357	[01:24<00:18,	3.49it/s]
83%	295/357	[01:25<00:17,	3.45it/s]
83%	296/357	[01:25<00:17,	3.39it/s]
83%	297/357	[01:25<00:17,	3.41it/s]
83%	298/357	[01:25<00:17,	3.41it/s]
84%	299/357	[01:26<00:17,	3.40it/s]
84%	300/357	[01:26<00:16,	3.40it/s]
84%	301/357	[01:26<00:16,	3.42it/s]
85%	302/357	[01:27<00:16,	3.36it/s]
85%	303/357	[01:27<00:15,	3.40it/s]
85%	304/357	[01:27<00:15,	3.40it/s]
85%	305/357	[01:28<00:15,	3.40it/s]
86%	306/357	[01:28<00:14,	3.47it/s]
86%	307/357	[01:28<00:14,	3.48it/s]
86%	308/357	[01:28<00:14,	3.48it/s]
87%	309/357	[01:29<00:13,	3.45it/s]
87%	310/357	[01:29<00:13,	3.45it/s]
87%	311/357	[01:29<00:13,	3.48it/s]
87%	312/357	[01:30<00:13,	3.40it/s]
88%	313/357	[01:30<00:12,	3.41it/s]
88%	314/357	[01:30<00:12,	3.41it/s]
88%	315/357	[01:30<00:12,	3.41it/s]
89%	316/357	[01:31<00:12,	3.36it/s]
89%	317/357	[01:31<00:11,	3.35it/s]
89%	318/357	[01:31<00:11,	3.38it/s]
89%	319/357	[01:32<00:11,	3.40it/s]
90%	320/357	[01:32<00:10,	3.41it/s]
90%	321/357	[01:32<00:10,	3.48it/s]
90%	322/357	[01:32<00:10,	3.48it/s]
90%	323/357	[01:33<00:09,	3.47it/s]
91%	324/357	[01:33<00:09,	3.49it/s]
91%	325/357	[01:33<00:09,	3.48it/s]
91%	326/357	[01:34<00:08,	3.53it/s]
92%	327/357	[01:34<00:08,	3.51it/s]
92%	328/357	[01:34<00:08,	3.46it/s]

```

92%|      | 329/357 [01:34<00:08,  3.44it/s]
92%|      | 330/357 [01:35<00:07,  3.39it/s]
93%|      | 331/357 [01:35<00:07,  3.30it/s]
93%|      | 332/357 [01:36<00:12,  1.94it/s]
93%|      | 333/357 [01:36<00:10,  2.20it/s]
94%|      | 334/357 [01:37<00:09,  2.46it/s]
94%|      | 335/357 [01:37<00:08,  2.69it/s]
94%|      | 336/357 [01:37<00:07,  2.88it/s]
94%|      | 337/357 [01:38<00:06,  2.99it/s]
95%|      | 338/357 [01:38<00:06,  3.11it/s]
95%|      | 339/357 [01:38<00:05,  3.18it/s]
95%|      | 340/357 [01:38<00:05,  3.23it/s]
96%|      | 341/357 [01:39<00:04,  3.27it/s]
96%|      | 342/357 [01:39<00:04,  3.27it/s]
96%|      | 343/357 [01:39<00:04,  3.31it/s]
96%|      | 344/357 [01:40<00:03,  3.29it/s]
97%|      | 345/357 [01:40<00:03,  3.27it/s]
97%|      | 346/357 [01:40<00:03,  3.30it/s]
97%|      | 347/357 [01:41<00:03,  3.29it/s]
97%|      | 348/357 [01:41<00:02,  3.27it/s]
98%|      | 349/357 [01:41<00:02,  3.28it/s]
98%|      | 350/357 [01:42<00:02,  3.25it/s]
98%|      | 351/357 [01:42<00:01,  3.24it/s]
99%|      | 352/357 [01:42<00:01,  3.25it/s]
99%|      | 353/357 [01:42<00:01,  3.27it/s]
99%|      | 354/357 [01:43<00:00,  3.28it/s]
99%|      | 355/357 [01:43<00:00,  3.28it/s]
100%|     | 356/357 [01:43<00:00,  3.27it/s]
100%|     | 357/357 [01:44<00:00,  3.43it/s]
20%|      | 2/10 [03:36<14:19, 107.45s/it]

```

For - 1 Loss: tensor(0.0021, device='cuda:0', grad\_fn=<MeanBackward0>)

```

0%|      | 0/357 [00:00<?,  ?it/s]
0%|      | 1/357 [00:00<01:38,  3.63it/s]
1%|      | 2/357 [00:00<01:37,  3.65it/s]
1%|      | 3/357 [00:00<01:37,  3.65it/s]
1%|      | 4/357 [00:01<01:36,  3.65it/s]
1%|      | 5/357 [00:01<01:36,  3.65it/s]
2%|      | 6/357 [00:01<01:35,  3.66it/s]
2%|      | 7/357 [00:01<01:35,  3.66it/s]
2%|      | 8/357 [00:02<01:35,  3.66it/s]
3%|      | 9/357 [00:02<01:35,  3.66it/s]
3%|      | 10/357 [00:02<01:34,  3.66it/s]
3%|      | 11/357 [00:03<01:34,  3.66it/s]
3%|      | 12/357 [00:03<01:34,  3.66it/s]
4%|      | 13/357 [00:03<01:33,  3.66it/s]
4%|      | 14/357 [00:03<01:33,  3.66it/s]

```

4%	15/357	[00:04<01:33,	3.66it/s]
4%	16/357	[00:04<01:33,	3.66it/s]
5%	17/357	[00:04<01:32,	3.66it/s]
5%	18/357	[00:04<01:32,	3.66it/s]
5%	19/357	[00:05<01:32,	3.66it/s]
6%	20/357	[00:05<01:31,	3.66it/s]
6%	21/357	[00:05<01:31,	3.66it/s]
6%	22/357	[00:06<01:31,	3.66it/s]
6%	23/357	[00:06<01:31,	3.66it/s]
7%	24/357	[00:06<01:31,	3.66it/s]
7%	25/357	[00:06<01:30,	3.66it/s]
7%	26/357	[00:07<01:30,	3.66it/s]
8%	27/357	[00:07<01:30,	3.66it/s]
8%	28/357	[00:07<01:29,	3.66it/s]
8%	29/357	[00:07<01:29,	3.66it/s]
8%	30/357	[00:08<01:29,	3.66it/s]
9%	31/357	[00:08<01:28,	3.66it/s]
9%	32/357	[00:08<01:28,	3.67it/s]
9%	33/357	[00:09<01:28,	3.66it/s]
10%	34/357	[00:09<01:28,	3.66it/s]
10%	35/357	[00:09<01:27,	3.66it/s]
10%	36/357	[00:09<01:27,	3.66it/s]
10%	37/357	[00:10<01:27,	3.66it/s]
11%	38/357	[00:10<01:27,	3.66it/s]
11%	39/357	[00:10<01:26,	3.66it/s]
11%	40/357	[00:10<01:26,	3.66it/s]
11%	41/357	[00:11<01:26,	3.66it/s]
12%	42/357	[00:11<01:26,	3.66it/s]
12%	43/357	[00:11<01:25,	3.66it/s]
12%	44/357	[00:12<01:25,	3.66it/s]
13%	45/357	[00:12<01:25,	3.66it/s]
13%	46/357	[00:12<01:24,	3.66it/s]
13%	47/357	[00:12<01:24,	3.66it/s]
13%	48/357	[00:13<01:24,	3.66it/s]
14%	49/357	[00:13<01:24,	3.66it/s]
14%	50/357	[00:13<01:23,	3.66it/s]
14%	51/357	[00:13<01:23,	3.66it/s]
15%	52/357	[00:14<01:23,	3.66it/s]
15%	53/357	[00:14<01:23,	3.66it/s]
15%	54/357	[00:14<01:22,	3.66it/s]
15%	55/357	[00:15<01:22,	3.66it/s]
16%	56/357	[00:15<01:22,	3.66it/s]
16%	57/357	[00:15<01:21,	3.66it/s]
16%	58/357	[00:15<01:21,	3.66it/s]
17%	59/357	[00:16<01:21,	3.66it/s]
17%	60/357	[00:16<01:21,	3.66it/s]
17%	61/357	[00:16<01:20,	3.66it/s]
17%	62/357	[00:16<01:20,	3.66it/s]

18%	63/357	[00:17<01:20,	3.66it/s]
18%	64/357	[00:17<01:19,	3.67it/s]
18%	65/357	[00:17<01:19,	3.66it/s]
18%	66/357	[00:18<01:19,	3.66it/s]
19%	67/357	[00:18<01:19,	3.67it/s]
19%	68/357	[00:18<01:18,	3.66it/s]
19%	69/357	[00:18<01:20,	3.56it/s]
20%	70/357	[00:19<01:19,	3.59it/s]
20%	71/357	[00:19<01:19,	3.62it/s]
20%	72/357	[00:19<01:18,	3.63it/s]
20%	73/357	[00:19<01:17,	3.64it/s]
21%	74/357	[00:20<01:17,	3.65it/s]
21%	75/357	[00:20<01:17,	3.65it/s]
21%	76/357	[00:20<01:16,	3.66it/s]
22%	77/357	[00:21<01:16,	3.66it/s]
22%	78/357	[00:21<01:16,	3.66it/s]
22%	79/357	[00:21<01:15,	3.66it/s]
22%	80/357	[00:21<01:15,	3.67it/s]
23%	81/357	[00:22<01:15,	3.66it/s]
23%	82/357	[00:22<01:15,	3.66it/s]
23%	83/357	[00:22<01:14,	3.67it/s]
24%	84/357	[00:22<01:14,	3.66it/s]
24%	85/357	[00:23<01:14,	3.67it/s]
24%	86/357	[00:23<01:13,	3.67it/s]
24%	87/357	[00:23<01:13,	3.66it/s]
25%	88/357	[00:24<01:13,	3.66it/s]
25%	89/357	[00:24<01:14,	3.60it/s]
25%	90/357	[00:24<01:13,	3.62it/s]
25%	91/357	[00:24<01:13,	3.63it/s]
26%	92/357	[00:25<01:12,	3.64it/s]
26%	93/357	[00:25<01:12,	3.65it/s]
26%	94/357	[00:25<01:12,	3.65it/s]
27%	95/357	[00:25<01:11,	3.66it/s]
27%	96/357	[00:26<01:11,	3.66it/s]
27%	97/357	[00:26<01:11,	3.66it/s]
27%	98/357	[00:26<01:10,	3.66it/s]
28%	99/357	[00:27<01:10,	3.66it/s]
28%	100/357	[00:27<01:10,	3.66it/s]
28%	101/357	[00:27<01:09,	3.66it/s]
29%	102/357	[00:27<01:09,	3.66it/s]
29%	103/357	[00:28<01:09,	3.66it/s]
29%	104/357	[00:28<01:09,	3.66it/s]
29%	105/357	[00:28<01:08,	3.66it/s]
30%	106/357	[00:28<01:08,	3.66it/s]
30%	107/357	[00:29<01:08,	3.66it/s]
30%	108/357	[00:29<01:07,	3.66it/s]
31%	109/357	[00:29<01:07,	3.66it/s]
31%	110/357	[00:30<01:07,	3.66it/s]



31%	111/357	[00:30<01:07,	3.66it/s]
31%	112/357	[00:30<01:06,	3.67it/s]
32%	113/357	[00:30<01:06,	3.66it/s]
32%	114/357	[00:31<01:06,	3.66it/s]
32%	115/357	[00:31<01:06,	3.66it/s]
32%	116/357	[00:31<01:05,	3.66it/s]
33%	117/357	[00:31<01:05,	3.66it/s]
33%	118/357	[00:32<01:05,	3.66it/s]
33%	119/357	[00:32<01:04,	3.66it/s]
34%	120/357	[00:32<01:04,	3.66it/s]
34%	121/357	[00:33<01:04,	3.66it/s]
34%	122/357	[00:33<01:04,	3.66it/s]
34%	123/357	[00:33<01:03,	3.66it/s]
35%	124/357	[00:33<01:03,	3.67it/s]
35%	125/357	[00:34<01:03,	3.66it/s]
35%	126/357	[00:34<01:03,	3.67it/s]
36%	127/357	[00:34<01:02,	3.66it/s]
36%	128/357	[00:34<01:02,	3.66it/s]
36%	129/357	[00:35<01:02,	3.66it/s]
36%	130/357	[00:35<01:02,	3.64it/s]
37%	131/357	[00:35<01:01,	3.65it/s]
37%	132/357	[00:36<01:01,	3.65it/s]
37%	133/357	[00:36<01:01,	3.65it/s]
38%	134/357	[00:36<01:01,	3.65it/s]
38%	135/357	[00:36<01:00,	3.66it/s]
38%	136/357	[00:37<01:00,	3.66it/s]
38%	137/357	[00:37<01:00,	3.66it/s]
39%	138/357	[00:37<00:59,	3.66it/s]
39%	139/357	[00:38<00:59,	3.67it/s]
39%	140/357	[00:38<00:59,	3.66it/s]
39%	141/357	[00:38<00:58,	3.66it/s]
40%	142/357	[00:38<00:58,	3.66it/s]
40%	143/357	[00:39<00:58,	3.66it/s]
40%	144/357	[00:39<00:58,	3.66it/s]
41%	145/357	[00:39<00:57,	3.66it/s]
41%	146/357	[00:39<00:57,	3.66it/s]
41%	147/357	[00:40<00:57,	3.66it/s]
41%	148/357	[00:40<00:57,	3.66it/s]
42%	149/357	[00:40<00:56,	3.66it/s]
42%	150/357	[00:41<00:56,	3.66it/s]
42%	151/357	[00:41<00:56,	3.66it/s]
43%	152/357	[00:41<00:55,	3.66it/s]
43%	153/357	[00:41<00:55,	3.66it/s]
43%	154/357	[00:42<00:55,	3.66it/s]
43%	155/357	[00:42<00:55,	3.66it/s]
44%	156/357	[00:42<00:55,	3.65it/s]
44%	157/357	[00:42<00:54,	3.66it/s]
44%	158/357	[00:43<00:54,	3.66it/s]

45%	159/357	[00:43<00:54,	3.64it/s]
45%	160/357	[00:43<00:54,	3.63it/s]
45%	161/357	[00:44<00:53,	3.64it/s]
45%	162/357	[00:44<00:53,	3.64it/s]
46%	163/357	[00:44<00:53,	3.65it/s]
46%	164/357	[00:44<00:52,	3.65it/s]
46%	165/357	[00:45<00:52,	3.66it/s]
46%	166/357	[00:45<00:52,	3.66it/s]
47%	167/357	[00:45<00:51,	3.66it/s]
47%	168/357	[00:45<00:51,	3.67it/s]
47%	169/357	[00:46<00:51,	3.66it/s]
48%	170/357	[00:46<00:51,	3.66it/s]
48%	171/357	[00:46<00:50,	3.66it/s]
48%	172/357	[00:47<00:50,	3.66it/s]
48%	173/357	[00:47<00:50,	3.66it/s]
49%	174/357	[00:47<00:50,	3.66it/s]
49%	175/357	[00:47<00:49,	3.66it/s]
49%	176/357	[00:48<00:49,	3.66it/s]
50%	177/357	[00:48<00:49,	3.66it/s]
50%	178/357	[00:48<00:48,	3.66it/s]
50%	179/357	[00:48<00:48,	3.66it/s]
50%	180/357	[00:49<00:48,	3.66it/s]
51%	181/357	[00:49<00:48,	3.66it/s]
51%	182/357	[00:49<00:48,	3.63it/s]
51%	183/357	[00:50<00:47,	3.64it/s]
52%	184/357	[00:50<00:47,	3.65it/s]
52%	185/357	[00:50<00:47,	3.65it/s]
52%	186/357	[00:50<00:46,	3.65it/s]
52%	187/357	[00:51<00:46,	3.66it/s]
53%	188/357	[00:51<00:46,	3.66it/s]
53%	189/357	[00:51<00:45,	3.66it/s]
53%	190/357	[00:51<00:45,	3.66it/s]
54%	191/357	[00:52<00:45,	3.67it/s]
54%	192/357	[00:52<00:45,	3.66it/s]
54%	193/357	[00:52<00:44,	3.67it/s]
54%	194/357	[00:53<00:44,	3.66it/s]
55%	195/357	[00:53<00:44,	3.66it/s]
55%	196/357	[00:53<00:43,	3.66it/s]
55%	197/357	[00:53<00:43,	3.66it/s]
55%	198/357	[00:54<00:43,	3.66it/s]
56%	199/357	[00:54<00:43,	3.66it/s]
56%	200/357	[00:54<00:46,	3.38it/s]
56%	201/357	[00:55<00:45,	3.46it/s]
57%	202/357	[00:55<00:44,	3.52it/s]
57%	203/357	[00:55<00:43,	3.56it/s]
57%	204/357	[00:55<00:42,	3.59it/s]
57%	205/357	[00:56<00:42,	3.61it/s]
58%	206/357	[00:56<00:41,	3.62it/s]

58%	207/357	[00:56<00:41,	3.64it/s]
58%	208/357	[00:56<00:40,	3.65it/s]
59%	209/357	[00:57<00:40,	3.65it/s]
59%	210/357	[00:57<00:40,	3.65it/s]
59%	211/357	[00:57<00:39,	3.66it/s]
59%	212/357	[00:58<00:39,	3.66it/s]
60%	213/357	[00:58<00:39,	3.66it/s]
60%	214/357	[00:58<00:39,	3.66it/s]
60%	215/357	[00:58<00:38,	3.66it/s]
61%	216/357	[00:59<00:38,	3.66it/s]
61%	217/357	[00:59<00:38,	3.66it/s]
61%	218/357	[00:59<00:37,	3.66it/s]
61%	219/357	[00:59<00:37,	3.66it/s]
62%	220/357	[01:00<00:37,	3.66it/s]
62%	221/357	[01:00<00:37,	3.67it/s]
62%	222/357	[01:00<00:36,	3.67it/s]
62%	223/357	[01:01<00:36,	3.67it/s]
63%	224/357	[01:01<00:36,	3.67it/s]
63%	225/357	[01:01<00:35,	3.67it/s]
63%	226/357	[01:01<00:35,	3.65it/s]
64%	227/357	[01:02<00:35,	3.65it/s]
64%	228/357	[01:02<00:35,	3.66it/s]
64%	229/357	[01:02<00:34,	3.66it/s]
64%	230/357	[01:02<00:34,	3.66it/s]
65%	231/357	[01:03<00:34,	3.66it/s]
65%	232/357	[01:03<00:34,	3.66it/s]
65%	233/357	[01:03<00:33,	3.67it/s]
66%	234/357	[01:04<00:33,	3.67it/s]
66%	235/357	[01:04<00:33,	3.65it/s]
66%	236/357	[01:04<00:33,	3.62it/s]
66%	237/357	[01:04<00:32,	3.64it/s]
67%	238/357	[01:05<00:32,	3.64it/s]
67%	239/357	[01:05<00:32,	3.65it/s]
67%	240/357	[01:05<00:31,	3.66it/s]
68%	241/357	[01:05<00:31,	3.66it/s]
68%	242/357	[01:06<00:31,	3.63it/s]
68%	243/357	[01:06<00:31,	3.64it/s]
68%	244/357	[01:06<00:30,	3.65it/s]
69%	245/357	[01:07<00:30,	3.65it/s]
69%	246/357	[01:07<00:30,	3.66it/s]
69%	247/357	[01:07<00:30,	3.66it/s]
69%	248/357	[01:08<00:53,	2.04it/s]
70%	249/357	[01:08<00:45,	2.35it/s]
70%	250/357	[01:09<00:40,	2.64it/s]
70%	251/357	[01:09<00:37,	2.85it/s]
71%	252/357	[01:09<00:34,	3.04it/s]
71%	253/357	[01:09<00:32,	3.20it/s]
71%	254/357	[01:10<00:30,	3.33it/s]

71%		255/357	[01:10<00:29,	3.42it/s]
72%		256/357	[01:10<00:28,	3.49it/s]
72%		257/357	[01:11<00:28,	3.54it/s]
72%		258/357	[01:11<00:27,	3.58it/s]
73%		259/357	[01:11<00:27,	3.58it/s]
73%		260/357	[01:11<00:26,	3.61it/s]
73%		261/357	[01:12<00:26,	3.63it/s]
73%		262/357	[01:12<00:26,	3.54it/s]
74%		263/357	[01:12<00:26,	3.57it/s]
74%		264/357	[01:13<00:25,	3.60it/s]
74%		265/357	[01:13<00:25,	3.62it/s]
75%		266/357	[01:13<00:25,	3.53it/s]
75%		267/357	[01:13<00:25,	3.57it/s]
75%		268/357	[01:14<00:24,	3.60it/s]
75%		269/357	[01:14<00:24,	3.62it/s]
76%		270/357	[01:14<00:24,	3.57it/s]
76%		271/357	[01:14<00:24,	3.58it/s]
76%		272/357	[01:15<00:23,	3.61it/s]
76%		273/357	[01:15<00:23,	3.60it/s]
77%		274/357	[01:15<00:23,	3.60it/s]
77%		275/357	[01:16<00:22,	3.62it/s]
77%		276/357	[01:16<00:22,	3.59it/s]
78%		277/357	[01:16<00:22,	3.58it/s]
78%		278/357	[01:16<00:22,	3.53it/s]
78%		279/357	[01:17<00:22,	3.53it/s]
78%		280/357	[01:17<00:21,	3.55it/s]
79%		281/357	[01:17<00:21,	3.58it/s]
79%		282/357	[01:18<00:21,	3.56it/s]
79%		283/357	[01:18<00:20,	3.56it/s]
80%		284/357	[01:18<00:20,	3.56it/s]
80%		285/357	[01:18<00:20,	3.51it/s]
80%		286/357	[01:19<00:19,	3.55it/s]
80%		287/357	[01:19<00:19,	3.51it/s]
81%		288/357	[01:19<00:19,	3.51it/s]
81%		289/357	[01:20<00:19,	3.53it/s]
81%		290/357	[01:20<00:18,	3.54it/s]
82%		291/357	[01:20<00:18,	3.55it/s]
82%		292/357	[01:20<00:18,	3.57it/s]
82%		293/357	[01:21<00:17,	3.58it/s]
82%		294/357	[01:21<00:17,	3.59it/s]
83%		295/357	[01:21<00:17,	3.55it/s]
83%		296/357	[01:22<00:18,	3.34it/s]
83%		297/357	[01:22<00:17,	3.38it/s]
83%		298/357	[01:22<00:17,	3.44it/s]
84%		299/357	[01:22<00:16,	3.44it/s]
84%		300/357	[01:23<00:16,	3.49it/s]
84%		301/357	[01:23<00:15,	3.50it/s]
85%		302/357	[01:23<00:15,	3.51it/s]

85%		303/357	[01:24<00:15,	3.49it/s]
85%		304/357	[01:24<00:15,	3.51it/s]
85%		305/357	[01:24<00:14,	3.50it/s]
86%		306/357	[01:24<00:14,	3.49it/s]
86%		307/357	[01:25<00:14,	3.48it/s]
86%		308/357	[01:25<00:14,	3.45it/s]
87%		309/357	[01:25<00:13,	3.47it/s]
87%		310/357	[01:26<00:13,	3.46it/s]
87%		311/357	[01:26<00:13,	3.43it/s]
87%		312/357	[01:26<00:13,	3.44it/s]
88%		313/357	[01:26<00:12,	3.46it/s]
88%		314/357	[01:27<00:12,	3.45it/s]
88%		315/357	[01:27<00:12,	3.46it/s]
89%		316/357	[01:27<00:11,	3.46it/s]
89%		317/357	[01:28<00:11,	3.49it/s]
89%		318/357	[01:28<00:11,	3.45it/s]
89%		319/357	[01:28<00:10,	3.46it/s]
90%		320/357	[01:28<00:10,	3.43it/s]
90%		321/357	[01:29<00:10,	3.42it/s]
90%		322/357	[01:29<00:10,	3.43it/s]
90%		323/357	[01:29<00:09,	3.45it/s]
91%		324/357	[01:30<00:09,	3.47it/s]
91%		325/357	[01:30<00:09,	3.49it/s]
91%		326/357	[01:30<00:09,	3.44it/s]
92%		327/357	[01:30<00:08,	3.43it/s]
92%		328/357	[01:31<00:08,	3.43it/s]
92%		329/357	[01:31<00:08,	3.44it/s]
92%		330/357	[01:31<00:07,	3.40it/s]
93%		331/357	[01:32<00:07,	3.39it/s]
93%		332/357	[01:32<00:07,	3.41it/s]
93%		333/357	[01:32<00:06,	3.46it/s]
94%		334/357	[01:33<00:06,	3.41it/s]
94%		335/357	[01:33<00:06,	3.40it/s]
94%		336/357	[01:33<00:06,	3.39it/s]
94%		337/357	[01:33<00:05,	3.36it/s]
95%		338/357	[01:34<00:05,	3.33it/s]
95%		339/357	[01:34<00:05,	3.34it/s]
95%		340/357	[01:34<00:05,	3.34it/s]
96%		341/357	[01:35<00:04,	3.34it/s]
96%		342/357	[01:35<00:04,	3.15it/s]
96%		343/357	[01:35<00:04,	3.19it/s]
96%		344/357	[01:36<00:04,	3.23it/s]
97%		345/357	[01:36<00:03,	3.24it/s]
97%		346/357	[01:36<00:03,	3.24it/s]
97%		347/357	[01:37<00:03,	3.24it/s]
97%		348/357	[01:37<00:02,	3.24it/s]
98%		349/357	[01:37<00:02,	3.28it/s]
98%		350/357	[01:37<00:02,	3.30it/s]

```

98%|      | 351/357 [01:38<00:01, 3.27it/s]
99%|      | 352/357 [01:38<00:01, 3.29it/s]
99%|      | 353/357 [01:38<00:01, 3.29it/s]
99%|      | 354/357 [01:39<00:00, 3.29it/s]
99%|      | 355/357 [01:39<00:00, 3.30it/s]
100%|     | 356/357 [01:39<00:00, 3.30it/s]
100%|     | 357/357 [01:40<00:00, 3.57it/s]
30%|      | 3/10 [05:16<12:09, 104.16s/it]

```

For - 2 Loss: tensor(0.0014, device='cuda:0', grad\_fn=<MeanBackward0>)

```

0%|      | 0/357 [00:00<?, ?it/s]
0%|      | 1/357 [00:00<01:36, 3.70it/s]
1%|      | 2/357 [00:00<01:36, 3.67it/s]
1%|      | 3/357 [00:00<01:36, 3.68it/s]
1%|      | 4/357 [00:01<01:36, 3.67it/s]
1%|      | 5/357 [00:01<01:35, 3.67it/s]
2%|      | 6/357 [00:01<01:35, 3.68it/s]
2%|      | 7/357 [00:01<01:35, 3.68it/s]
2%|      | 8/357 [00:02<01:34, 3.67it/s]
3%|      | 9/357 [00:02<01:35, 3.66it/s]
3%|      | 10/357 [00:02<01:34, 3.67it/s]
3%|      | 11/357 [00:02<01:34, 3.67it/s]
3%|      | 12/357 [00:03<01:33, 3.67it/s]
4%|      | 13/357 [00:03<01:33, 3.68it/s]
4%|      | 14/357 [00:03<01:33, 3.67it/s]
4%|      | 15/357 [00:04<01:33, 3.68it/s]
4%|      | 16/357 [00:04<01:32, 3.68it/s]
5%|      | 17/357 [00:04<01:32, 3.67it/s]
5%|      | 18/357 [00:04<01:32, 3.68it/s]
5%|      | 19/357 [00:05<01:31, 3.68it/s]
6%|      | 20/357 [00:05<01:31, 3.68it/s]
6%|      | 21/357 [00:05<01:31, 3.68it/s]
6%|      | 22/357 [00:05<01:31, 3.68it/s]
6%|      | 23/357 [00:06<01:30, 3.68it/s]
7%|      | 24/357 [00:06<01:30, 3.68it/s]
7%|      | 25/357 [00:06<01:30, 3.68it/s]
7%|      | 26/357 [00:07<01:30, 3.67it/s]
8%|      | 27/357 [00:07<01:29, 3.68it/s]
8%|      | 28/357 [00:07<01:29, 3.68it/s]
8%|      | 29/357 [00:07<01:29, 3.67it/s]
8%|      | 30/357 [00:08<01:28, 3.68it/s]
9%|      | 31/357 [00:08<01:28, 3.68it/s]
9%|      | 32/357 [00:08<01:28, 3.68it/s]
9%|      | 33/357 [00:08<01:28, 3.68it/s]
10%|     | 34/357 [00:09<01:27, 3.68it/s]
10%|     | 35/357 [00:09<01:27, 3.67it/s]
10%|     | 36/357 [00:09<01:27, 3.67it/s]

```

10%	37/357	[00:10<01:27,	3.67it/s]
11%	38/357	[00:10<01:26,	3.67it/s]
11%	39/357	[00:10<01:26,	3.67it/s]
11%	40/357	[00:10<01:26,	3.67it/s]
11%	41/357	[00:11<01:26,	3.67it/s]
12%	42/357	[00:11<01:25,	3.67it/s]
12%	43/357	[00:11<01:25,	3.67it/s]
12%	44/357	[00:11<01:25,	3.67it/s]
13%	45/357	[00:12<01:24,	3.67it/s]
13%	46/357	[00:12<01:24,	3.67it/s]
13%	47/357	[00:12<01:24,	3.67it/s]
13%	48/357	[00:13<01:24,	3.68it/s]
14%	49/357	[00:13<01:23,	3.68it/s]
14%	50/357	[00:13<01:23,	3.68it/s]
14%	51/357	[00:13<01:23,	3.68it/s]
15%	52/357	[00:14<01:23,	3.67it/s]
15%	53/357	[00:14<01:22,	3.67it/s]
15%	54/357	[00:14<01:22,	3.67it/s]
15%	55/357	[00:14<01:22,	3.67it/s]
16%	56/357	[00:15<01:21,	3.67it/s]
16%	57/357	[00:15<01:21,	3.67it/s]
16%	58/357	[00:15<01:21,	3.68it/s]
17%	59/357	[00:16<01:21,	3.68it/s]
17%	60/357	[00:16<01:20,	3.67it/s]
17%	61/357	[00:16<01:21,	3.64it/s]
17%	62/357	[00:16<01:20,	3.65it/s]
18%	63/357	[00:17<01:20,	3.65it/s]
18%	64/357	[00:17<01:20,	3.66it/s]
18%	65/357	[00:17<01:19,	3.65it/s]
18%	66/357	[00:17<01:19,	3.66it/s]
19%	67/357	[00:18<01:19,	3.66it/s]
19%	68/357	[00:18<01:18,	3.67it/s]
19%	69/357	[00:18<01:18,	3.67it/s]
20%	70/357	[00:19<01:18,	3.67it/s]
20%	71/357	[00:19<01:17,	3.67it/s]
20%	72/357	[00:19<01:17,	3.67it/s]
20%	73/357	[00:19<01:17,	3.67it/s]
21%	74/357	[00:20<01:17,	3.67it/s]
21%	75/357	[00:20<01:16,	3.67it/s]
21%	76/357	[00:20<01:16,	3.68it/s]
22%	77/357	[00:20<01:16,	3.68it/s]
22%	78/357	[00:21<01:15,	3.67it/s]
22%	79/357	[00:21<01:15,	3.68it/s]
22%	80/357	[00:21<01:15,	3.68it/s]
23%	81/357	[00:22<01:15,	3.68it/s]
23%	82/357	[00:22<01:14,	3.68it/s]
23%	83/357	[00:22<01:14,	3.67it/s]
24%	84/357	[00:22<01:14,	3.68it/s]

24%	85/357	[00:23<01:13,	3.68it/s]
24%	86/357	[00:23<01:13,	3.68it/s]
24%	87/357	[00:23<01:13,	3.67it/s]
25%	88/357	[00:23<01:13,	3.68it/s]
25%	89/357	[00:24<01:12,	3.68it/s]
25%	90/357	[00:24<01:12,	3.68it/s]
25%	91/357	[00:24<01:12,	3.68it/s]
26%	92/357	[00:25<01:12,	3.68it/s]
26%	93/357	[00:25<01:11,	3.68it/s]
26%	94/357	[00:25<01:11,	3.68it/s]
27%	95/357	[00:25<01:11,	3.68it/s]
27%	96/357	[00:26<01:10,	3.68it/s]
27%	97/357	[00:26<01:10,	3.68it/s]
27%	98/357	[00:26<01:10,	3.68it/s]
28%	99/357	[00:26<01:10,	3.68it/s]
28%	100/357	[00:27<01:09,	3.68it/s]
28%	101/357	[00:27<01:09,	3.68it/s]
29%	102/357	[00:27<01:09,	3.68it/s]
29%	103/357	[00:28<01:09,	3.68it/s]
29%	104/357	[00:28<01:08,	3.68it/s]
29%	105/357	[00:28<01:08,	3.68it/s]
30%	106/357	[00:28<01:08,	3.68it/s]
30%	107/357	[00:29<01:07,	3.68it/s]
30%	108/357	[00:29<01:07,	3.68it/s]
31%	109/357	[00:29<01:07,	3.68it/s]
31%	110/357	[00:29<01:07,	3.68it/s]
31%	111/357	[00:30<01:06,	3.68it/s]
31%	112/357	[00:30<01:06,	3.68it/s]
32%	113/357	[00:30<01:06,	3.68it/s]
32%	114/357	[00:31<01:06,	3.67it/s]
32%	115/357	[00:31<01:05,	3.68it/s]
32%	116/357	[00:31<01:05,	3.67it/s]
33%	117/357	[00:31<01:05,	3.68it/s]
33%	118/357	[00:32<01:04,	3.68it/s]
33%	119/357	[00:32<01:04,	3.68it/s]
34%	120/357	[00:32<01:04,	3.67it/s]
34%	121/357	[00:32<01:04,	3.67it/s]
34%	122/357	[00:33<01:03,	3.67it/s]
34%	123/357	[00:33<01:03,	3.68it/s]
35%	124/357	[00:33<01:03,	3.68it/s]
35%	125/357	[00:34<01:03,	3.68it/s]
35%	126/357	[00:34<01:02,	3.67it/s]
36%	127/357	[00:35<01:53,	2.03it/s]
36%	128/357	[00:35<01:37,	2.34it/s]
36%	129/357	[00:35<01:26,	2.63it/s]
36%	130/357	[00:36<01:18,	2.87it/s]
37%	131/357	[00:36<01:13,	3.07it/s]
37%	132/357	[00:36<01:09,	3.23it/s]



37%	133/357	[00:36<01:06,	3.35it/s]
38%	134/357	[00:37<01:04,	3.44it/s]
38%	135/357	[00:37<01:03,	3.51it/s]
38%	136/357	[00:37<01:02,	3.56it/s]
38%	137/357	[00:38<01:01,	3.59it/s]
39%	138/357	[00:38<01:00,	3.62it/s]
39%	139/357	[00:38<00:59,	3.64it/s]
39%	140/357	[00:38<00:59,	3.65it/s]
39%	141/357	[00:39<00:59,	3.66it/s]
40%	142/357	[00:39<00:58,	3.66it/s]
40%	143/357	[00:39<00:58,	3.67it/s]
40%	144/357	[00:39<00:58,	3.67it/s]
41%	145/357	[00:40<00:57,	3.67it/s]
41%	146/357	[00:40<00:57,	3.67it/s]
41%	147/357	[00:40<00:57,	3.68it/s]
41%	148/357	[00:41<00:56,	3.68it/s]
42%	149/357	[00:41<00:56,	3.68it/s]
42%	150/357	[00:41<00:56,	3.68it/s]
42%	151/357	[00:41<00:56,	3.67it/s]
43%	152/357	[00:42<00:56,	3.66it/s]
43%	153/357	[00:42<00:55,	3.66it/s]
43%	154/357	[00:42<00:55,	3.67it/s]
43%	155/357	[00:42<00:55,	3.67it/s]
44%	156/357	[00:43<00:54,	3.67it/s]
44%	157/357	[00:43<00:54,	3.68it/s]
44%	158/357	[00:43<00:54,	3.67it/s]
45%	159/357	[00:44<00:53,	3.68it/s]
45%	160/357	[00:44<00:53,	3.68it/s]
45%	161/357	[00:44<00:53,	3.68it/s]
45%	162/357	[00:44<00:53,	3.68it/s]
46%	163/357	[00:45<00:52,	3.68it/s]
46%	164/357	[00:45<00:52,	3.68it/s]
46%	165/357	[00:45<00:52,	3.68it/s]
46%	166/357	[00:45<00:51,	3.68it/s]
47%	167/357	[00:46<00:51,	3.68it/s]
47%	168/357	[00:46<00:51,	3.68it/s]
47%	169/357	[00:46<00:51,	3.68it/s]
48%	170/357	[00:47<00:50,	3.68it/s]
48%	171/357	[00:47<00:50,	3.67it/s]
48%	172/357	[00:47<00:50,	3.67it/s]
48%	173/357	[00:47<00:50,	3.67it/s]
49%	174/357	[00:48<00:49,	3.67it/s]
49%	175/357	[00:48<00:49,	3.67it/s]
49%	176/357	[00:48<00:49,	3.67it/s]
50%	177/357	[00:48<00:49,	3.67it/s]
50%	178/357	[00:49<00:48,	3.67it/s]
50%	179/357	[00:49<00:48,	3.67it/s]
50%	180/357	[00:49<00:48,	3.67it/s]

51%	181/357	[00:50<00:47,	3.67it/s]
51%	182/357	[00:50<00:47,	3.67it/s]
51%	183/357	[00:50<00:47,	3.67it/s]
52%	184/357	[00:50<00:47,	3.67it/s]
52%	185/357	[00:51<00:46,	3.67it/s]
52%	186/357	[00:51<00:46,	3.67it/s]
52%	187/357	[00:51<00:46,	3.67it/s]
53%	188/357	[00:51<00:46,	3.67it/s]
53%	189/357	[00:52<00:45,	3.68it/s]
53%	190/357	[00:52<00:45,	3.67it/s]
54%	191/357	[00:52<00:45,	3.67it/s]
54%	192/357	[00:52<00:44,	3.67it/s]
54%	193/357	[00:53<00:44,	3.67it/s]
54%	194/357	[00:53<00:44,	3.67it/s]
55%	195/357	[00:53<00:44,	3.68it/s]
55%	196/357	[00:54<00:46,	3.48it/s]
55%	197/357	[00:54<00:45,	3.53it/s]
55%	198/357	[00:54<00:44,	3.57it/s]
56%	199/357	[00:54<00:43,	3.60it/s]
56%	200/357	[00:55<00:43,	3.62it/s]
56%	201/357	[00:55<00:42,	3.64it/s]
57%	202/357	[00:55<00:42,	3.65it/s]
57%	203/357	[00:56<00:42,	3.66it/s]
57%	204/357	[00:56<00:41,	3.67it/s]
57%	205/357	[00:56<00:41,	3.66it/s]
58%	206/357	[00:56<00:41,	3.66it/s]
58%	207/357	[00:57<00:40,	3.67it/s]
58%	208/357	[00:57<00:40,	3.67it/s]
59%	209/357	[00:57<00:40,	3.67it/s]
59%	210/357	[00:57<00:39,	3.68it/s]
59%	211/357	[00:58<00:39,	3.68it/s]
59%	212/357	[00:58<00:39,	3.67it/s]
60%	213/357	[00:58<00:39,	3.67it/s]
60%	214/357	[00:59<00:38,	3.68it/s]
60%	215/357	[00:59<00:38,	3.68it/s]
61%	216/357	[00:59<00:38,	3.68it/s]
61%	217/357	[00:59<00:38,	3.68it/s]
61%	218/357	[01:00<00:37,	3.68it/s]
61%	219/357	[01:00<00:37,	3.68it/s]
62%	220/357	[01:00<00:37,	3.68it/s]
62%	221/357	[01:00<00:36,	3.68it/s]
62%	222/357	[01:01<00:36,	3.68it/s]
62%	223/357	[01:01<00:36,	3.68it/s]
63%	224/357	[01:01<00:36,	3.68it/s]
63%	225/357	[01:02<00:35,	3.68it/s]
63%	226/357	[01:02<00:35,	3.67it/s]
64%	227/357	[01:02<00:35,	3.67it/s]
64%	228/357	[01:02<00:35,	3.67it/s]

64%		229/357	[01:03<00:34,	3.67it/s]
64%		230/357	[01:03<00:34,	3.66it/s]
65%		231/357	[01:03<00:34,	3.61it/s]
65%		232/357	[01:03<00:34,	3.63it/s]
65%		233/357	[01:04<00:34,	3.64it/s]
66%		234/357	[01:04<00:33,	3.65it/s]
66%		235/357	[01:04<00:33,	3.66it/s]
66%		236/357	[01:05<00:33,	3.66it/s]
66%		237/357	[01:05<00:33,	3.59it/s]
67%		238/357	[01:05<00:32,	3.61it/s]
67%		239/357	[01:05<00:32,	3.63it/s]
67%		240/357	[01:06<00:32,	3.65it/s]
68%		241/357	[01:06<00:31,	3.65it/s]
68%		242/357	[01:06<00:31,	3.66it/s]
68%		243/357	[01:06<00:32,	3.54it/s]
68%		244/357	[01:07<00:31,	3.57it/s]
69%		245/357	[01:07<00:31,	3.60it/s]
69%		246/357	[01:07<00:30,	3.62it/s]
69%		247/357	[01:08<00:30,	3.64it/s]
69%		248/357	[01:08<00:29,	3.63it/s]
70%		249/357	[01:08<00:29,	3.65it/s]
70%		250/357	[01:08<00:29,	3.66it/s]
70%		251/357	[01:09<00:29,	3.65it/s]
71%		252/357	[01:09<00:28,	3.65it/s]
71%		253/357	[01:09<00:28,	3.66it/s]
71%		254/357	[01:10<00:28,	3.58it/s]
71%		255/357	[01:10<00:28,	3.60it/s]
72%		256/357	[01:10<00:28,	3.59it/s]
72%		257/357	[01:10<00:27,	3.62it/s]
72%		258/357	[01:11<00:27,	3.63it/s]
73%		259/357	[01:11<00:26,	3.65it/s]
73%		260/357	[01:11<00:26,	3.66it/s]
73%		261/357	[01:11<00:26,	3.64it/s]
73%		262/357	[01:12<00:26,	3.56it/s]
74%		263/357	[01:12<00:26,	3.57it/s]
74%		264/357	[01:12<00:25,	3.60it/s]
74%		265/357	[01:13<00:25,	3.62it/s]
75%		266/357	[01:13<00:25,	3.63it/s]
75%		267/357	[01:13<00:24,	3.64it/s]
75%		268/357	[01:13<00:24,	3.64it/s]
75%		269/357	[01:14<00:24,	3.57it/s]
76%		270/357	[01:14<00:24,	3.59it/s]
76%		271/357	[01:14<00:23,	3.61it/s]
76%		272/357	[01:15<00:23,	3.58it/s]
76%		273/357	[01:15<00:23,	3.54it/s]
77%		274/357	[01:15<00:23,	3.54it/s]
77%		275/357	[01:15<00:23,	3.54it/s]
77%		276/357	[01:16<00:22,	3.57it/s]

78%	277/357	[01:16<00:22,	3.56it/s]
78%	278/357	[01:16<00:22,	3.53it/s]
78%	279/357	[01:16<00:21,	3.56it/s]
78%	280/357	[01:17<00:21,	3.56it/s]
79%	281/357	[01:17<00:21,	3.59it/s]
79%	282/357	[01:17<00:20,	3.59it/s]
79%	283/357	[01:18<00:20,	3.58it/s]
80%	284/357	[01:18<00:20,	3.59it/s]
80%	285/357	[01:18<00:20,	3.59it/s]
80%	286/357	[01:18<00:19,	3.58it/s]
80%	287/357	[01:19<00:20,	3.36it/s]
81%	288/357	[01:19<00:20,	3.41it/s]
81%	289/357	[01:19<00:19,	3.45it/s]
81%	290/357	[01:20<00:19,	3.48it/s]
82%	291/357	[01:20<00:18,	3.52it/s]
82%	292/357	[01:20<00:19,	3.38it/s]
82%	293/357	[01:21<00:18,	3.41it/s]
82%	294/357	[01:21<00:18,	3.43it/s]
83%	295/357	[01:21<00:17,	3.49it/s]
83%	296/357	[01:21<00:17,	3.51it/s]
83%	297/357	[01:22<00:16,	3.54it/s]
83%	298/357	[01:22<00:16,	3.56it/s]
84%	299/357	[01:22<00:16,	3.56it/s]
84%	300/357	[01:22<00:16,	3.53it/s]
84%	301/357	[01:23<00:16,	3.50it/s]
85%	302/357	[01:23<00:15,	3.51it/s]
85%	303/357	[01:23<00:15,	3.49it/s]
85%	304/357	[01:24<00:15,	3.50it/s]
85%	305/357	[01:24<00:14,	3.49it/s]
86%	306/357	[01:24<00:14,	3.45it/s]
86%	307/357	[01:24<00:14,	3.46it/s]
86%	308/357	[01:25<00:14,	3.43it/s]
87%	309/357	[01:25<00:13,	3.45it/s]
87%	310/357	[01:25<00:13,	3.48it/s]
87%	311/357	[01:26<00:13,	3.43it/s]
87%	312/357	[01:26<00:13,	3.42it/s]
88%	313/357	[01:26<00:12,	3.44it/s]
88%	314/357	[01:27<00:12,	3.42it/s]
88%	315/357	[01:27<00:12,	3.42it/s]
89%	316/357	[01:27<00:11,	3.44it/s]
89%	317/357	[01:27<00:11,	3.51it/s]
89%	318/357	[01:28<00:11,	3.49it/s]
89%	319/357	[01:29<00:20,	1.82it/s]
90%	320/357	[01:29<00:17,	2.10it/s]
90%	321/357	[01:29<00:15,	2.39it/s]
90%	322/357	[01:30<00:13,	2.63it/s]
90%	323/357	[01:30<00:12,	2.82it/s]
91%	324/357	[01:30<00:11,	2.94it/s]

```

91%|      | 325/357 [01:31<00:10,  3.06it/s]
91%|      | 326/357 [01:31<00:09,  3.18it/s]
92%|      | 327/357 [01:31<00:09,  3.24it/s]
92%|      | 328/357 [01:31<00:08,  3.27it/s]
92%|      | 329/357 [01:32<00:08,  3.22it/s]
92%|      | 330/357 [01:32<00:08,  3.26it/s]
93%|      | 331/357 [01:32<00:07,  3.31it/s]
93%|      | 332/357 [01:33<00:07,  3.33it/s]
93%|      | 333/357 [01:33<00:07,  3.31it/s]
94%|      | 334/357 [01:33<00:06,  3.31it/s]
94%|      | 335/357 [01:34<00:06,  3.32it/s]
94%|      | 336/357 [01:34<00:06,  3.32it/s]
94%|      | 337/357 [01:34<00:06,  3.31it/s]
95%|      | 338/357 [01:35<00:05,  3.28it/s]
95%|      | 339/357 [01:35<00:05,  3.30it/s]
95%|      | 340/357 [01:35<00:04,  3.40it/s]
96%|      | 341/357 [01:35<00:04,  3.39it/s]
96%|      | 342/357 [01:36<00:04,  3.34it/s]
96%|      | 343/357 [01:36<00:04,  3.33it/s]
96%|      | 344/357 [01:36<00:03,  3.30it/s]
97%|      | 345/357 [01:37<00:03,  3.31it/s]
97%|      | 346/357 [01:37<00:03,  3.28it/s]
97%|      | 347/357 [01:37<00:03,  3.29it/s]
97%|      | 348/357 [01:38<00:02,  3.29it/s]
98%|      | 349/357 [01:38<00:02,  3.30it/s]
98%|      | 350/357 [01:38<00:02,  3.26it/s]
98%|      | 351/357 [01:38<00:01,  3.27it/s]
99%|      | 352/357 [01:39<00:01,  3.27it/s]
99%|      | 353/357 [01:39<00:01,  3.27it/s]
99%|      | 354/357 [01:39<00:00,  3.29it/s]
99%|      | 355/357 [01:40<00:00,  3.25it/s]
100%|     | 356/357 [01:40<00:00,  3.25it/s]
100%|     | 357/357 [01:40<00:00,  3.54it/s]
40%|      | 4/10 [06:57<10:17, 102.91s/it]

```

For - 3 Loss: tensor(0.0013, device='cuda:0', grad\_fn=<MeanBackward0>)

```

0%|      | 0/357 [00:00<?, ?it/s]
0%|      | 1/357 [00:00<01:43,  3.44it/s]
1%|      | 2/357 [00:00<01:39,  3.57it/s]
1%|      | 3/357 [00:00<01:37,  3.62it/s]
1%|      | 4/357 [00:01<01:37,  3.64it/s]
1%|      | 5/357 [00:01<01:36,  3.66it/s]
2%|      | 6/357 [00:01<01:35,  3.66it/s]
2%|      | 7/357 [00:01<01:35,  3.67it/s]
2%|      | 8/357 [00:02<01:35,  3.67it/s]
3%|      | 9/357 [00:02<01:34,  3.67it/s]
3%|      | 10/357 [00:02<01:34,  3.68it/s]

```

3%	11/357	[00:03<01:34,	3.68it/s]
3%	12/357	[00:03<01:33,	3.68it/s]
4%	13/357	[00:03<01:33,	3.68it/s]
4%	14/357	[00:03<01:33,	3.68it/s]
4%	15/357	[00:04<01:32,	3.68it/s]
4%	16/357	[00:04<01:32,	3.68it/s]
5%	17/357	[00:04<01:32,	3.68it/s]
5%	18/357	[00:04<01:32,	3.68it/s]
5%	19/357	[00:05<01:31,	3.68it/s]
6%	20/357	[00:05<01:31,	3.68it/s]
6%	21/357	[00:05<01:31,	3.68it/s]
6%	22/357	[00:05<01:30,	3.68it/s]
6%	23/357	[00:06<01:30,	3.68it/s]
7%	24/357	[00:06<01:30,	3.68it/s]
7%	25/357	[00:06<01:30,	3.68it/s]
7%	26/357	[00:07<01:29,	3.68it/s]
8%	27/357	[00:07<01:29,	3.68it/s]
8%	28/357	[00:07<01:29,	3.68it/s]
8%	29/357	[00:07<01:29,	3.68it/s]
8%	30/357	[00:08<01:28,	3.68it/s]
9%	31/357	[00:08<01:28,	3.68it/s]
9%	32/357	[00:08<01:28,	3.68it/s]
9%	33/357	[00:08<01:27,	3.68it/s]
10%	34/357	[00:09<01:27,	3.68it/s]
10%	35/357	[00:09<01:27,	3.68it/s]
10%	36/357	[00:09<01:27,	3.68it/s]
10%	37/357	[00:10<01:26,	3.68it/s]
11%	38/357	[00:10<01:26,	3.68it/s]
11%	39/357	[00:10<01:26,	3.68it/s]
11%	40/357	[00:10<01:26,	3.68it/s]
11%	41/357	[00:11<01:25,	3.68it/s]
12%	42/357	[00:11<01:25,	3.68it/s]
12%	43/357	[00:11<01:25,	3.68it/s]
12%	44/357	[00:11<01:25,	3.68it/s]
13%	45/357	[00:12<01:24,	3.68it/s]
13%	46/357	[00:12<01:24,	3.68it/s]
13%	47/357	[00:12<01:24,	3.68it/s]
13%	48/357	[00:13<01:23,	3.68it/s]
14%	49/357	[00:13<01:23,	3.68it/s]
14%	50/357	[00:13<01:23,	3.68it/s]
14%	51/357	[00:13<01:23,	3.68it/s]
15%	52/357	[00:14<01:22,	3.68it/s]
15%	53/357	[00:14<01:22,	3.68it/s]
15%	54/357	[00:14<01:22,	3.68it/s]
15%	55/357	[00:14<01:22,	3.68it/s]
16%	56/357	[00:15<01:21,	3.68it/s]
16%	57/357	[00:15<01:21,	3.68it/s]
16%	58/357	[00:15<01:21,	3.68it/s]

17%	59/357	[00:16<01:21,	3.66it/s]
17%	60/357	[00:16<01:21,	3.66it/s]
17%	61/357	[00:16<01:20,	3.67it/s]
17%	62/357	[00:16<01:20,	3.67it/s]
18%	63/357	[00:17<01:20,	3.67it/s]
18%	64/357	[00:17<01:19,	3.68it/s]
18%	65/357	[00:17<01:19,	3.68it/s]
18%	66/357	[00:17<01:19,	3.68it/s]
19%	67/357	[00:18<01:18,	3.68it/s]
19%	68/357	[00:18<01:18,	3.68it/s]
19%	69/357	[00:18<01:18,	3.68it/s]
20%	70/357	[00:19<01:18,	3.68it/s]
20%	71/357	[00:19<01:17,	3.68it/s]
20%	72/357	[00:19<01:17,	3.66it/s]
20%	73/357	[00:19<01:17,	3.67it/s]
21%	74/357	[00:20<01:17,	3.67it/s]
21%	75/357	[00:20<01:16,	3.67it/s]
21%	76/357	[00:20<01:16,	3.68it/s]
22%	77/357	[00:20<01:16,	3.68it/s]
22%	78/357	[00:21<01:15,	3.68it/s]
22%	79/357	[00:21<01:21,	3.42it/s]
22%	80/357	[00:21<01:19,	3.49it/s]
23%	81/357	[00:22<01:17,	3.55it/s]
23%	82/357	[00:22<01:16,	3.59it/s]
23%	83/357	[00:22<01:15,	3.61it/s]
24%	84/357	[00:22<01:15,	3.63it/s]
24%	85/357	[00:23<01:14,	3.65it/s]
24%	86/357	[00:23<01:14,	3.66it/s]
24%	87/357	[00:23<01:13,	3.67it/s]
25%	88/357	[00:24<01:13,	3.67it/s]
25%	89/357	[00:24<01:12,	3.67it/s]
25%	90/357	[00:24<01:12,	3.68it/s]
25%	91/357	[00:24<01:12,	3.68it/s]
26%	92/357	[00:25<01:12,	3.67it/s]
26%	93/357	[00:25<01:11,	3.68it/s]
26%	94/357	[00:25<01:11,	3.68it/s]
27%	95/357	[00:25<01:11,	3.68it/s]
27%	96/357	[00:26<01:10,	3.68it/s]
27%	97/357	[00:26<01:10,	3.67it/s]
27%	98/357	[00:26<01:10,	3.67it/s]
28%	99/357	[00:27<01:10,	3.68it/s]
28%	100/357	[00:27<01:09,	3.68it/s]
28%	101/357	[00:27<01:09,	3.68it/s]
29%	102/357	[00:27<01:09,	3.68it/s]
29%	103/357	[00:28<01:09,	3.68it/s]
29%	104/357	[00:28<01:08,	3.68it/s]
29%	105/357	[00:28<01:08,	3.68it/s]
30%	106/357	[00:28<01:08,	3.68it/s]

30%	107/357	[00:29<01:08,	3.68it/s]
30%	108/357	[00:29<01:07,	3.68it/s]
31%	109/357	[00:29<01:07,	3.68it/s]
31%	110/357	[00:29<01:07,	3.68it/s]
31%	111/357	[00:30<01:06,	3.68it/s]
31%	112/357	[00:30<01:06,	3.68it/s]
32%	113/357	[00:30<01:06,	3.68it/s]
32%	114/357	[00:31<01:06,	3.68it/s]
32%	115/357	[00:31<01:05,	3.68it/s]
32%	116/357	[00:31<01:05,	3.68it/s]
33%	117/357	[00:31<01:05,	3.68it/s]
33%	118/357	[00:32<01:04,	3.68it/s]
33%	119/357	[00:32<01:04,	3.68it/s]
34%	120/357	[00:32<01:04,	3.68it/s]
34%	121/357	[00:32<01:04,	3.68it/s]
34%	122/357	[00:33<01:03,	3.68it/s]
34%	123/357	[00:33<01:03,	3.68it/s]
35%	124/357	[00:33<01:03,	3.67it/s]
35%	125/357	[00:34<01:03,	3.67it/s]
35%	126/357	[00:34<01:02,	3.68it/s]
36%	127/357	[00:34<01:02,	3.68it/s]
36%	128/357	[00:34<01:02,	3.68it/s]
36%	129/357	[00:35<01:02,	3.68it/s]
36%	130/357	[00:35<01:01,	3.68it/s]
37%	131/357	[00:35<01:01,	3.68it/s]
37%	132/357	[00:35<01:01,	3.68it/s]
37%	133/357	[00:36<01:00,	3.68it/s]
38%	134/357	[00:36<01:00,	3.68it/s]
38%	135/357	[00:36<01:00,	3.68it/s]
38%	136/357	[00:37<01:00,	3.68it/s]
38%	137/357	[00:37<00:59,	3.68it/s]
39%	138/357	[00:37<00:59,	3.68it/s]
39%	139/357	[00:37<00:59,	3.68it/s]
39%	140/357	[00:38<00:58,	3.68it/s]
39%	141/357	[00:38<00:58,	3.68it/s]
40%	142/357	[00:38<00:58,	3.68it/s]
40%	143/357	[00:38<00:58,	3.68it/s]
40%	144/357	[00:39<00:57,	3.68it/s]
41%	145/357	[00:39<00:58,	3.65it/s]
41%	146/357	[00:39<00:57,	3.66it/s]
41%	147/357	[00:40<00:57,	3.67it/s]
41%	148/357	[00:40<00:56,	3.67it/s]
42%	149/357	[00:40<00:56,	3.68it/s]
42%	150/357	[00:40<00:56,	3.67it/s]
42%	151/357	[00:41<00:55,	3.68it/s]
43%	152/357	[00:41<00:55,	3.68it/s]
43%	153/357	[00:41<00:55,	3.68it/s]
43%	154/357	[00:41<00:55,	3.68it/s]



43%	155/357	[00:42<00:54,	3.68it/s]
44%	156/357	[00:42<00:54,	3.68it/s]
44%	157/357	[00:42<00:54,	3.68it/s]
44%	158/357	[00:43<00:54,	3.68it/s]
45%	159/357	[00:43<00:53,	3.68it/s]
45%	160/357	[00:43<00:53,	3.68it/s]
45%	161/357	[00:43<00:53,	3.68it/s]
45%	162/357	[00:44<00:52,	3.68it/s]
46%	163/357	[00:44<00:52,	3.68it/s]
46%	164/357	[00:44<00:52,	3.68it/s]
46%	165/357	[00:44<00:52,	3.68it/s]
46%	166/357	[00:45<00:51,	3.68it/s]
47%	167/357	[00:45<00:51,	3.68it/s]
47%	168/357	[00:45<00:51,	3.68it/s]
47%	169/357	[00:46<00:51,	3.68it/s]
48%	170/357	[00:46<00:50,	3.68it/s]
48%	171/357	[00:46<00:50,	3.68it/s]
48%	172/357	[00:46<00:50,	3.68it/s]
48%	173/357	[00:47<00:50,	3.68it/s]
49%	174/357	[00:47<00:49,	3.67it/s]
49%	175/357	[00:47<00:49,	3.68it/s]
49%	176/357	[00:47<00:49,	3.68it/s]
50%	177/357	[00:48<00:49,	3.63it/s]
50%	178/357	[00:48<00:49,	3.62it/s]
50%	179/357	[00:48<00:48,	3.64it/s]
50%	180/357	[00:49<00:48,	3.65it/s]
51%	181/357	[00:49<00:48,	3.66it/s]
51%	182/357	[00:49<00:47,	3.66it/s]
51%	183/357	[00:49<00:47,	3.67it/s]
52%	184/357	[00:50<00:47,	3.67it/s]
52%	185/357	[00:50<00:46,	3.67it/s]
52%	186/357	[00:50<00:46,	3.67it/s]
52%	187/357	[00:50<00:46,	3.63it/s]
53%	188/357	[00:51<00:46,	3.64it/s]
53%	189/357	[00:51<00:48,	3.47it/s]
53%	190/357	[00:51<00:47,	3.53it/s]
54%	191/357	[00:52<00:46,	3.57it/s]
54%	192/357	[00:52<00:45,	3.59it/s]
54%	193/357	[00:52<00:45,	3.62it/s]
54%	194/357	[00:52<00:44,	3.63it/s]
55%	195/357	[00:53<00:44,	3.64it/s]
55%	196/357	[00:53<00:44,	3.65it/s]
55%	197/357	[00:53<00:47,	3.40it/s]
55%	198/357	[00:54<00:45,	3.48it/s]
56%	199/357	[00:54<00:44,	3.54it/s]
56%	200/357	[00:54<00:43,	3.58it/s]
56%	201/357	[00:54<00:43,	3.61it/s]
57%	202/357	[00:55<00:42,	3.63it/s]

57%	203/357	[00:55<00:42,	3.64it/s]
57%	204/357	[00:55<00:41,	3.65it/s]
57%	205/357	[00:55<00:41,	3.66it/s]
58%	206/357	[00:56<00:41,	3.66it/s]
58%	207/357	[00:56<00:40,	3.66it/s]
58%	208/357	[00:56<00:40,	3.67it/s]
59%	209/357	[00:57<00:40,	3.67it/s]
59%	210/357	[00:57<00:40,	3.67it/s]
59%	211/357	[00:57<00:39,	3.67it/s]
59%	212/357	[00:57<00:39,	3.67it/s]
60%	213/357	[00:58<00:39,	3.67it/s]
60%	214/357	[00:58<00:38,	3.67it/s]
60%	215/357	[00:58<00:38,	3.67it/s]
61%	216/357	[00:58<00:38,	3.67it/s]
61%	217/357	[00:59<00:38,	3.68it/s]
61%	218/357	[00:59<00:37,	3.68it/s]
61%	219/357	[00:59<00:37,	3.68it/s]
62%	220/357	[01:00<00:37,	3.68it/s]
62%	221/357	[01:00<00:36,	3.68it/s]
62%	222/357	[01:00<00:36,	3.67it/s]
62%	223/357	[01:00<00:36,	3.67it/s]
63%	224/357	[01:01<00:36,	3.67it/s]
63%	225/357	[01:01<00:35,	3.68it/s]
63%	226/357	[01:01<00:35,	3.66it/s]
64%	227/357	[01:01<00:35,	3.66it/s]
64%	228/357	[01:02<00:35,	3.67it/s]
64%	229/357	[01:02<00:34,	3.67it/s]
64%	230/357	[01:02<00:34,	3.66it/s]
65%	231/357	[01:03<00:34,	3.66it/s]
65%	232/357	[01:03<00:56,	2.21it/s]
65%	233/357	[01:04<00:51,	2.41it/s]
66%	234/357	[01:04<00:45,	2.68it/s]
66%	235/357	[01:04<00:41,	2.92it/s]
66%	236/357	[01:05<00:40,	2.96it/s]
66%	237/357	[01:05<00:38,	3.14it/s]
67%	238/357	[01:05<00:37,	3.21it/s]
67%	239/357	[01:05<00:35,	3.34it/s]
67%	240/357	[01:06<00:34,	3.43it/s]
68%	241/357	[01:06<00:33,	3.51it/s]
68%	242/357	[01:06<00:32,	3.56it/s]
68%	243/357	[01:07<00:31,	3.59it/s]
68%	244/357	[01:07<00:31,	3.62it/s]
69%	245/357	[01:07<00:30,	3.64it/s]
69%	246/357	[01:07<00:30,	3.65it/s]
69%	247/357	[01:08<00:30,	3.66it/s]
69%	248/357	[01:08<00:29,	3.66it/s]
70%	249/357	[01:08<00:29,	3.67it/s]
70%	250/357	[01:08<00:29,	3.67it/s]

70%	251/357	[01:09<00:28,	3.68it/s]
71%	252/357	[01:09<00:28,	3.62it/s]
71%	253/357	[01:09<00:28,	3.63it/s]
71%	254/357	[01:10<00:28,	3.65it/s]
71%	255/357	[01:10<00:27,	3.65it/s]
72%	256/357	[01:10<00:27,	3.65it/s]
72%	257/357	[01:10<00:27,	3.63it/s]
72%	258/357	[01:11<00:27,	3.64it/s]
73%	259/357	[01:11<00:26,	3.64it/s]
73%	260/357	[01:11<00:27,	3.57it/s]
73%	261/357	[01:12<00:26,	3.60it/s]
73%	262/357	[01:12<00:26,	3.60it/s]
74%	263/357	[01:12<00:26,	3.61it/s]
74%	264/357	[01:12<00:25,	3.58it/s]
74%	265/357	[01:13<00:25,	3.60it/s]
75%	266/357	[01:13<00:25,	3.60it/s]
75%	267/357	[01:13<00:25,	3.56it/s]
75%	268/357	[01:13<00:25,	3.56it/s]
75%	269/357	[01:14<00:24,	3.59it/s]
76%	270/357	[01:14<00:24,	3.62it/s]
76%	271/357	[01:14<00:23,	3.60it/s]
76%	272/357	[01:15<00:23,	3.60it/s]
76%	273/357	[01:15<00:23,	3.59it/s]
77%	274/357	[01:15<00:23,	3.54it/s]
77%	275/357	[01:15<00:23,	3.54it/s]
77%	276/357	[01:16<00:22,	3.53it/s]
78%	277/357	[01:16<00:22,	3.53it/s]
78%	278/357	[01:16<00:22,	3.52it/s]
78%	279/357	[01:17<00:22,	3.54it/s]
78%	280/357	[01:17<00:21,	3.58it/s]
79%	281/357	[01:17<00:21,	3.56it/s]
79%	282/357	[01:17<00:21,	3.54it/s]
79%	283/357	[01:18<00:20,	3.55it/s]
80%	284/357	[01:18<00:20,	3.55it/s]
80%	285/357	[01:18<00:20,	3.51it/s]
80%	286/357	[01:19<00:20,	3.53it/s]
80%	287/357	[01:19<00:19,	3.56it/s]
81%	288/357	[01:19<00:19,	3.51it/s]
81%	289/357	[01:19<00:19,	3.52it/s]
81%	290/357	[01:20<00:18,	3.54it/s]
82%	291/357	[01:20<00:18,	3.53it/s]
82%	292/357	[01:20<00:18,	3.56it/s]
82%	293/357	[01:21<00:18,	3.54it/s]
82%	294/357	[01:21<00:17,	3.51it/s]
83%	295/357	[01:21<00:17,	3.47it/s]
83%	296/357	[01:21<00:17,	3.49it/s]
83%	297/357	[01:22<00:17,	3.52it/s]
83%	298/357	[01:22<00:16,	3.49it/s]

84%		299/357	[01:22<00:16,	3.47it/s]
84%		300/357	[01:23<00:16,	3.53it/s]
84%		301/357	[01:23<00:15,	3.51it/s]
85%		302/357	[01:23<00:15,	3.52it/s]
85%		303/357	[01:23<00:15,	3.51it/s]
85%		304/357	[01:24<00:15,	3.53it/s]
85%		305/357	[01:24<00:14,	3.47it/s]
86%		306/357	[01:24<00:14,	3.47it/s]
86%		307/357	[01:25<00:14,	3.49it/s]
86%		308/357	[01:25<00:14,	3.38it/s]
87%		309/357	[01:25<00:14,	3.38it/s]
87%		310/357	[01:25<00:13,	3.39it/s]
87%		311/357	[01:26<00:13,	3.39it/s]
87%		312/357	[01:26<00:13,	3.39it/s]
88%		313/357	[01:26<00:12,	3.42it/s]
88%		314/357	[01:27<00:12,	3.42it/s]
88%		315/357	[01:27<00:12,	3.42it/s]
89%		316/357	[01:27<00:12,	3.40it/s]
89%		317/357	[01:27<00:11,	3.41it/s]
89%		318/357	[01:28<00:11,	3.42it/s]
89%		319/357	[01:28<00:11,	3.33it/s]
90%		320/357	[01:28<00:11,	3.34it/s]
90%		321/357	[01:29<00:10,	3.36it/s]
90%		322/357	[01:29<00:10,	3.34it/s]
90%		323/357	[01:29<00:10,	3.38it/s]
91%		324/357	[01:30<00:10,	3.15it/s]
91%		325/357	[01:30<00:09,	3.22it/s]
91%		326/357	[01:30<00:09,	3.27it/s]
92%		327/357	[01:31<00:09,	3.31it/s]
92%		328/357	[01:31<00:08,	3.32it/s]
92%		329/357	[01:31<00:08,	3.36it/s]
92%		330/357	[01:31<00:08,	3.35it/s]
93%		331/357	[01:32<00:07,	3.37it/s]
93%		332/357	[01:32<00:07,	3.40it/s]
93%		333/357	[01:32<00:07,	3.39it/s]
94%		334/357	[01:33<00:06,	3.38it/s]
94%		335/357	[01:33<00:06,	3.37it/s]
94%		336/357	[01:33<00:06,	3.33it/s]
94%		337/357	[01:33<00:06,	3.33it/s]
95%		338/357	[01:34<00:05,	3.33it/s]
95%		339/357	[01:34<00:05,	3.32it/s]
95%		340/357	[01:34<00:05,	3.35it/s]
96%		341/357	[01:35<00:04,	3.37it/s]
96%		342/357	[01:35<00:04,	3.35it/s]
96%		343/357	[01:35<00:04,	3.35it/s]
96%		344/357	[01:36<00:03,	3.32it/s]
97%		345/357	[01:36<00:03,	3.30it/s]
97%		346/357	[01:36<00:03,	3.25it/s]

```

97%|      | 347/357 [01:37<00:03,  3.25it/s]
97%|      | 348/357 [01:37<00:02,  3.27it/s]
98%|      | 349/357 [01:37<00:02,  3.28it/s]
98%|      | 350/357 [01:37<00:02,  3.20it/s]
98%|      | 351/357 [01:38<00:01,  3.22it/s]
99%|      | 352/357 [01:38<00:01,  3.24it/s]
99%|      | 353/357 [01:38<00:01,  3.28it/s]
99%|      | 354/357 [01:39<00:00,  3.26it/s]
99%|      | 355/357 [01:39<00:00,  3.29it/s]
100%|     | 356/357 [01:39<00:00,  3.26it/s]
100%|     | 357/357 [01:40<00:00,  3.57it/s]
50%|      |  5/10 [08:37<08:29, 101.95s/it]

```

For - 4 Loss: tensor(0.0008, device='cuda:0', grad\_fn=<MeanBackward0>)

```

0%|      |  0/357 [00:00<?,  ?it/s]
0%|      |  1/357 [00:00<01:36,  3.70it/s]
1%|      |  2/357 [00:00<01:36,  3.69it/s]
1%|      |  3/357 [00:00<01:35,  3.69it/s]
1%|      |  4/357 [00:01<01:36,  3.68it/s]
1%|      |  5/357 [00:01<01:35,  3.68it/s]
2%|      |  6/357 [00:01<01:35,  3.68it/s]
2%|      |  7/357 [00:01<01:35,  3.68it/s]
2%|      |  8/357 [00:02<01:34,  3.68it/s]
3%|      |  9/357 [00:02<01:34,  3.68it/s]
3%|      | 10/357 [00:02<01:34,  3.68it/s]
3%|      | 11/357 [00:02<01:33,  3.68it/s]
3%|      | 12/357 [00:03<01:33,  3.68it/s]
4%|      | 13/357 [00:03<01:33,  3.69it/s]
4%|      | 14/357 [00:03<01:33,  3.68it/s]
4%|      | 15/357 [00:04<01:32,  3.68it/s]
4%|      | 16/357 [00:04<01:32,  3.68it/s]
5%|      | 17/357 [00:04<01:32,  3.68it/s]
5%|      | 18/357 [00:04<01:32,  3.68it/s]
5%|      | 19/357 [00:05<01:31,  3.68it/s]
6%|      | 20/357 [00:05<01:31,  3.68it/s]
6%|      | 21/357 [00:05<01:31,  3.68it/s]
6%|      | 22/357 [00:05<01:31,  3.68it/s]
6%|      | 23/357 [00:06<01:30,  3.68it/s]
7%|      | 24/357 [00:06<01:30,  3.68it/s]
7%|      | 25/357 [00:06<01:30,  3.68it/s]
7%|      | 26/357 [00:07<01:29,  3.68it/s]
8%|      | 27/357 [00:07<01:29,  3.68it/s]
8%|      | 28/357 [00:07<01:29,  3.68it/s]
8%|      | 29/357 [00:07<01:29,  3.68it/s]
8%|      | 30/357 [00:08<01:28,  3.68it/s]
9%|      | 31/357 [00:08<01:28,  3.68it/s]
9%|      | 32/357 [00:08<01:28,  3.68it/s]

```

9%	33/357	[00:08<01:27,	3.68it/s]
10%	34/357	[00:09<01:27,	3.68it/s]
10%	35/357	[00:09<01:27,	3.68it/s]
10%	36/357	[00:09<01:27,	3.68it/s]
10%	37/357	[00:10<01:27,	3.68it/s]
11%	38/357	[00:10<01:26,	3.68it/s]
11%	39/357	[00:10<01:26,	3.68it/s]
11%	40/357	[00:10<01:26,	3.68it/s]
11%	41/357	[00:11<01:25,	3.68it/s]
12%	42/357	[00:11<01:25,	3.68it/s]
12%	43/357	[00:11<01:25,	3.68it/s]
12%	44/357	[00:11<01:25,	3.68it/s]
13%	45/357	[00:12<01:24,	3.68it/s]
13%	46/357	[00:12<01:24,	3.68it/s]
13%	47/357	[00:12<01:24,	3.68it/s]
13%	48/357	[00:13<01:23,	3.68it/s]
14%	49/357	[00:13<01:23,	3.68it/s]
14%	50/357	[00:13<01:23,	3.68it/s]
14%	51/357	[00:13<01:23,	3.68it/s]
15%	52/357	[00:14<01:22,	3.68it/s]
15%	53/357	[00:14<01:22,	3.68it/s]
15%	54/357	[00:14<01:22,	3.68it/s]
15%	55/357	[00:14<01:22,	3.68it/s]
16%	56/357	[00:15<01:21,	3.68it/s]
16%	57/357	[00:15<01:21,	3.68it/s]
16%	58/357	[00:15<01:21,	3.68it/s]
17%	59/357	[00:16<01:20,	3.68it/s]
17%	60/357	[00:16<01:20,	3.68it/s]
17%	61/357	[00:16<01:20,	3.68it/s]
17%	62/357	[00:16<01:20,	3.68it/s]
18%	63/357	[00:17<01:19,	3.68it/s]
18%	64/357	[00:17<01:19,	3.68it/s]
18%	65/357	[00:17<01:19,	3.68it/s]
18%	66/357	[00:17<01:19,	3.68it/s]
19%	67/357	[00:18<01:18,	3.68it/s]
19%	68/357	[00:18<01:18,	3.68it/s]
19%	69/357	[00:18<01:18,	3.68it/s]
20%	70/357	[00:19<01:18,	3.68it/s]
20%	71/357	[00:19<01:17,	3.68it/s]
20%	72/357	[00:19<01:17,	3.68it/s]
20%	73/357	[00:19<01:17,	3.68it/s]
21%	74/357	[00:20<01:16,	3.68it/s]
21%	75/357	[00:20<01:16,	3.68it/s]
21%	76/357	[00:20<01:16,	3.68it/s]
22%	77/357	[00:20<01:16,	3.68it/s]
22%	78/357	[00:21<01:15,	3.68it/s]
22%	79/357	[00:21<01:15,	3.68it/s]
22%	80/357	[00:21<01:15,	3.68it/s]

23%	81/357	[00:22<01:14,	3.68it/s]
23%	82/357	[00:22<01:14,	3.68it/s]
23%	83/357	[00:22<01:14,	3.68it/s]
24%	84/357	[00:22<01:14,	3.68it/s]
24%	85/357	[00:23<01:13,	3.69it/s]
24%	86/357	[00:23<01:13,	3.68it/s]
24%	87/357	[00:23<01:13,	3.68it/s]
25%	88/357	[00:23<01:13,	3.68it/s]
25%	89/357	[00:24<01:12,	3.69it/s]
25%	90/357	[00:24<01:12,	3.68it/s]
25%	91/357	[00:24<01:12,	3.68it/s]
26%	92/357	[00:24<01:11,	3.68it/s]
26%	93/357	[00:25<01:11,	3.68it/s]
26%	94/357	[00:25<01:11,	3.68it/s]
27%	95/357	[00:25<01:11,	3.68it/s]
27%	96/357	[00:26<01:10,	3.68it/s]
27%	97/357	[00:26<01:10,	3.68it/s]
27%	98/357	[00:26<01:10,	3.68it/s]
28%	99/357	[00:27<02:13,	1.93it/s]
28%	100/357	[00:27<01:53,	2.26it/s]
28%	101/357	[00:28<01:40,	2.55it/s]
29%	102/357	[00:28<01:30,	2.81it/s]
29%	103/357	[00:28<01:23,	3.03it/s]
29%	104/357	[00:29<01:19,	3.20it/s]
29%	105/357	[00:29<01:15,	3.34it/s]
30%	106/357	[00:29<01:13,	3.43it/s]
30%	107/357	[00:29<01:11,	3.51it/s]
30%	108/357	[00:30<01:09,	3.56it/s]
31%	109/357	[00:30<01:10,	3.54it/s]
31%	110/357	[00:30<01:08,	3.59it/s]
31%	111/357	[00:31<01:10,	3.51it/s]
31%	112/357	[00:31<01:32,	2.65it/s]
32%	113/357	[00:31<01:27,	2.80it/s]
32%	114/357	[00:32<01:23,	2.91it/s]
32%	115/357	[00:32<01:21,	2.99it/s]
32%	116/357	[00:32<01:18,	3.07it/s]
33%	117/357	[00:33<01:17,	3.11it/s]
33%	118/357	[00:33<01:16,	3.14it/s]
33%	119/357	[00:33<01:15,	3.16it/s]
34%	120/357	[00:34<01:14,	3.18it/s]
34%	121/357	[00:34<01:13,	3.20it/s]
34%	122/357	[00:34<01:12,	3.22it/s]
34%	123/357	[00:35<01:12,	3.23it/s]
35%	124/357	[00:35<01:12,	3.23it/s]
35%	125/357	[00:35<01:11,	3.24it/s]
35%	126/357	[00:35<01:11,	3.24it/s]
36%	127/357	[00:36<01:10,	3.25it/s]
36%	128/357	[00:36<01:10,	3.24it/s]

36%	129/357	[00:36<01:10,	3.24it/s]
36%	130/357	[00:37<01:09,	3.25it/s]
37%	131/357	[00:37<01:09,	3.25it/s]
37%	132/357	[00:37<01:09,	3.26it/s]
37%	133/357	[00:38<01:09,	3.24it/s]
38%	134/357	[00:38<01:08,	3.25it/s]
38%	135/357	[00:38<01:08,	3.25it/s]
38%	136/357	[00:39<01:08,	3.25it/s]
38%	137/357	[00:39<01:07,	3.25it/s]
39%	138/357	[00:39<01:07,	3.25it/s]
39%	139/357	[00:39<01:06,	3.26it/s]
39%	140/357	[00:40<01:06,	3.25it/s]
39%	141/357	[00:40<01:06,	3.25it/s]
40%	142/357	[00:40<01:06,	3.25it/s]
40%	143/357	[00:41<01:05,	3.25it/s]
40%	144/357	[00:41<01:05,	3.26it/s]
41%	145/357	[00:41<01:04,	3.26it/s]
41%	146/357	[00:42<01:04,	3.26it/s]
41%	147/357	[00:42<01:04,	3.26it/s]
41%	148/357	[00:42<01:04,	3.26it/s]
42%	149/357	[00:43<01:03,	3.25it/s]
42%	150/357	[00:43<01:03,	3.26it/s]
42%	151/357	[00:43<01:03,	3.26it/s]
43%	152/357	[00:43<01:02,	3.26it/s]
43%	153/357	[00:44<01:02,	3.26it/s]
43%	154/357	[00:44<01:02,	3.26it/s]
43%	155/357	[00:44<01:02,	3.25it/s]
44%	156/357	[00:45<01:01,	3.25it/s]
44%	157/357	[00:45<01:01,	3.25it/s]
44%	158/357	[00:45<01:01,	3.25it/s]
45%	159/357	[00:46<01:01,	3.23it/s]
45%	160/357	[00:46<01:00,	3.23it/s]
45%	161/357	[00:46<01:00,	3.25it/s]
45%	162/357	[00:47<00:59,	3.25it/s]
46%	163/357	[00:47<00:59,	3.26it/s]
46%	164/357	[00:47<00:59,	3.26it/s]
46%	165/357	[00:47<00:59,	3.25it/s]
46%	166/357	[00:48<00:59,	3.20it/s]
47%	167/357	[00:48<00:59,	3.22it/s]
47%	168/357	[00:48<00:58,	3.23it/s]
47%	169/357	[00:49<00:58,	3.24it/s]
48%	170/357	[00:49<00:57,	3.25it/s]
48%	171/357	[00:49<00:57,	3.25it/s]
48%	172/357	[00:50<00:57,	3.23it/s]
48%	173/357	[00:50<00:56,	3.24it/s]
49%	174/357	[00:50<00:56,	3.24it/s]
49%	175/357	[00:51<00:56,	3.24it/s]
49%	176/357	[00:51<00:55,	3.25it/s]



50%	177/357	[00:51<00:55,	3.24it/s]
50%	178/357	[00:51<00:55,	3.25it/s]
50%	179/357	[00:52<00:54,	3.24it/s]
50%	180/357	[00:52<00:54,	3.24it/s]
51%	181/357	[00:52<00:54,	3.25it/s]
51%	182/357	[00:53<00:54,	3.24it/s]
51%	183/357	[00:53<00:53,	3.24it/s]
52%	184/357	[00:53<00:53,	3.25it/s]
52%	185/357	[00:54<00:52,	3.25it/s]
52%	186/357	[00:54<00:52,	3.26it/s]
52%	187/357	[00:54<00:52,	3.24it/s]
53%	188/357	[00:55<00:52,	3.24it/s]
53%	189/357	[00:55<00:51,	3.24it/s]
53%	190/357	[00:55<00:51,	3.25it/s]
54%	191/357	[00:56<00:54,	3.03it/s]
54%	192/357	[00:56<00:53,	3.09it/s]
54%	193/357	[00:56<00:52,	3.15it/s]
54%	194/357	[00:56<00:51,	3.18it/s]
55%	195/357	[00:57<00:50,	3.20it/s]
55%	196/357	[00:57<00:50,	3.18it/s]
55%	197/357	[00:57<00:50,	3.20it/s]
55%	198/357	[00:58<00:49,	3.19it/s]
56%	199/357	[00:58<00:49,	3.17it/s]
56%	200/357	[00:58<00:49,	3.19it/s]
56%	201/357	[00:59<00:48,	3.20it/s]
57%	202/357	[00:59<00:48,	3.23it/s]
57%	203/357	[00:59<00:47,	3.23it/s]
57%	204/357	[01:00<00:47,	3.23it/s]
57%	205/357	[01:00<00:47,	3.23it/s]
58%	206/357	[01:00<00:46,	3.24it/s]
58%	207/357	[01:00<00:46,	3.24it/s]
58%	208/357	[01:01<00:45,	3.25it/s]
59%	209/357	[01:01<00:45,	3.26it/s]
59%	210/357	[01:01<00:45,	3.25it/s]
59%	211/357	[01:02<00:45,	3.23it/s]
59%	212/357	[01:02<00:44,	3.23it/s]
60%	213/357	[01:02<00:44,	3.23it/s]
60%	214/357	[01:03<00:44,	3.23it/s]
60%	215/357	[01:03<00:43,	3.24it/s]
61%	216/357	[01:03<00:43,	3.25it/s]
61%	217/357	[01:04<00:43,	3.24it/s]
61%	218/357	[01:04<00:42,	3.25it/s]
61%	219/357	[01:04<00:42,	3.25it/s]
62%	220/357	[01:04<00:42,	3.24it/s]
62%	221/357	[01:05<00:42,	3.24it/s]
62%	222/357	[01:05<00:41,	3.25it/s]
62%	223/357	[01:05<00:41,	3.24it/s]
63%	224/357	[01:06<00:41,	3.18it/s]

63%		225/357	[01:06<00:41,	3.15it/s]
63%		226/357	[01:06<00:41,	3.18it/s]
64%		227/357	[01:07<00:40,	3.20it/s]
64%		228/357	[01:07<00:40,	3.22it/s]
64%		229/357	[01:07<00:39,	3.22it/s]
64%		230/357	[01:08<00:39,	3.24it/s]
65%		231/357	[01:08<00:40,	3.10it/s]
65%		232/357	[01:08<00:39,	3.16it/s]
65%		233/357	[01:09<00:38,	3.19it/s]
66%		234/357	[01:09<00:38,	3.20it/s]
66%		235/357	[01:09<00:37,	3.22it/s]
66%		236/357	[01:09<00:37,	3.22it/s]
66%		237/357	[01:10<00:37,	3.18it/s]
67%		238/357	[01:10<00:37,	3.20it/s]
67%		239/357	[01:10<00:36,	3.23it/s]
67%		240/357	[01:11<00:37,	3.10it/s]
68%		241/357	[01:11<00:36,	3.15it/s]
68%		242/357	[01:11<00:36,	3.18it/s]
68%		243/357	[01:12<00:35,	3.20it/s]
68%		244/357	[01:12<00:35,	3.21it/s]
69%		245/357	[01:12<00:35,	3.18it/s]
69%		246/357	[01:13<00:34,	3.19it/s]
69%		247/357	[01:13<00:34,	3.21it/s]
69%		248/357	[01:13<00:33,	3.22it/s]
70%		249/357	[01:14<00:33,	3.18it/s]
70%		250/357	[01:14<00:33,	3.18it/s]
70%		251/357	[01:14<00:33,	3.20it/s]
71%		252/357	[01:15<00:33,	3.17it/s]
71%		253/357	[01:15<00:32,	3.20it/s]
71%		254/357	[01:15<00:32,	3.21it/s]
71%		255/357	[01:15<00:31,	3.21it/s]
72%		256/357	[01:16<00:31,	3.21it/s]
72%		257/357	[01:16<00:31,	3.16it/s]
72%		258/357	[01:16<00:31,	3.16it/s]
73%		259/357	[01:17<00:30,	3.16it/s]
73%		260/357	[01:17<00:30,	3.18it/s]
73%		261/357	[01:17<00:30,	3.19it/s]
73%		262/357	[01:18<00:29,	3.20it/s]
74%		263/357	[01:18<00:29,	3.18it/s]
74%		264/357	[01:18<00:29,	3.17it/s]
74%		265/357	[01:19<00:28,	3.18it/s]
75%		266/357	[01:19<00:28,	3.20it/s]
75%		267/357	[01:19<00:27,	3.22it/s]
75%		268/357	[01:20<00:27,	3.22it/s]
75%		269/357	[01:20<00:27,	3.15it/s]
76%		270/357	[01:20<00:27,	3.15it/s]
76%		271/357	[01:20<00:27,	3.18it/s]
76%		272/357	[01:21<00:26,	3.17it/s]

76%	273/357	[01:21<00:26,	3.18it/s]
77%	274/357	[01:21<00:25,	3.20it/s]
77%	275/357	[01:22<00:25,	3.21it/s]
77%	276/357	[01:22<00:25,	3.15it/s]
78%	277/357	[01:22<00:25,	3.16it/s]
78%	278/357	[01:23<00:25,	3.11it/s]
78%	279/357	[01:23<00:25,	3.06it/s]
78%	280/357	[01:23<00:24,	3.11it/s]
79%	281/357	[01:24<00:24,	3.14it/s]
79%	282/357	[01:24<00:23,	3.13it/s]
79%	283/357	[01:24<00:24,	3.07it/s]
80%	284/357	[01:25<00:23,	3.07it/s]
80%	285/357	[01:25<00:23,	3.10it/s]
80%	286/357	[01:25<00:22,	3.16it/s]
80%	287/357	[01:26<00:22,	3.13it/s]
81%	288/357	[01:26<00:22,	3.11it/s]
81%	289/357	[01:26<00:22,	3.08it/s]
81%	290/357	[01:27<00:22,	3.01it/s]
82%	291/357	[01:27<00:21,	3.02it/s]
82%	292/357	[01:27<00:21,	3.08it/s]
82%	293/357	[01:28<00:20,	3.11it/s]
82%	294/357	[01:28<00:20,	3.10it/s]
83%	295/357	[01:28<00:19,	3.11it/s]
83%	296/357	[01:29<00:19,	3.11it/s]
83%	297/357	[01:29<00:19,	3.09it/s]
83%	298/357	[01:29<00:19,	3.07it/s]
84%	299/357	[01:30<00:18,	3.06it/s]
84%	300/357	[01:30<00:19,	2.99it/s]
84%	301/357	[01:30<00:18,	3.00it/s]
85%	302/357	[01:31<00:18,	2.94it/s]
85%	303/357	[01:31<00:27,	1.99it/s]
85%	304/357	[01:32<00:25,	2.10it/s]
85%	305/357	[01:32<00:22,	2.29it/s]
86%	306/357	[01:33<00:20,	2.52it/s]
86%	307/357	[01:33<00:18,	2.66it/s]
86%	308/357	[01:33<00:17,	2.79it/s]
87%	309/357	[01:33<00:16,	2.86it/s]
87%	310/357	[01:34<00:16,	2.93it/s]
87%	311/357	[01:34<00:16,	2.87it/s]
87%	312/357	[01:34<00:15,	2.97it/s]
88%	313/357	[01:35<00:14,	3.02it/s]
88%	314/357	[01:35<00:14,	3.03it/s]
88%	315/357	[01:35<00:13,	3.03it/s]
89%	316/357	[01:36<00:13,	3.05it/s]
89%	317/357	[01:36<00:13,	3.07it/s]
89%	318/357	[01:36<00:13,	2.95it/s]
89%	319/357	[01:37<00:13,	2.89it/s]
90%	320/357	[01:37<00:12,	2.88it/s]

90%	321/357	[01:38<00:12,	2.86it/s]
90%	322/357	[01:38<00:11,	3.01it/s]
90%	323/357	[01:38<00:10,	3.11it/s]
91%	324/357	[01:38<00:10,	3.20it/s]
91%	325/357	[01:39<00:10,	3.18it/s]
91%	326/357	[01:39<00:09,	3.21it/s]
92%	327/357	[01:39<00:09,	3.24it/s]
92%	328/357	[01:40<00:09,	3.21it/s]
92%	329/357	[01:40<00:08,	3.25it/s]
92%	330/357	[01:40<00:08,	3.25it/s]
93%	331/357	[01:41<00:07,	3.26it/s]
93%	332/357	[01:41<00:07,	3.26it/s]
93%	333/357	[01:41<00:07,	3.24it/s]
94%	334/357	[01:41<00:07,	3.24it/s]
94%	335/357	[01:42<00:06,	3.36it/s]
94%	336/357	[01:42<00:06,	3.25it/s]
94%	337/357	[01:43<00:06,	2.97it/s]
95%	338/357	[01:43<00:06,	3.06it/s]
95%	339/357	[01:43<00:05,	3.22it/s]
95%	340/357	[01:43<00:05,	3.25it/s]
96%	341/357	[01:44<00:04,	3.27it/s]
96%	342/357	[01:44<00:04,	3.38it/s]
96%	343/357	[01:44<00:04,	3.47it/s]
96%	344/357	[01:45<00:03,	3.42it/s]
97%	345/357	[01:45<00:03,	3.38it/s]
97%	346/357	[01:45<00:03,	3.46it/s]
97%	347/357	[01:45<00:02,	3.41it/s]
97%	348/357	[01:46<00:02,	3.36it/s]
98%	349/357	[01:46<00:02,	3.33it/s]
98%	350/357	[01:46<00:02,	3.26it/s]
98%	351/357	[01:47<00:01,	3.04it/s]
99%	352/357	[01:47<00:01,	3.11it/s]
99%	353/357	[01:47<00:01,	3.17it/s]
99%	354/357	[01:48<00:00,	3.03it/s]
99%	355/357	[01:48<00:00,	3.06it/s]
100%	356/357	[01:48<00:00,	3.11it/s]
100%	357/357	[01:49<00:00,	3.27it/s]
60%	6/10	[10:27<06:57,	104.46s/it]

For - 5 Loss: tensor(0.0006, device='cuda:0', grad\_fn=<MeanBackward0>)

0%	0/357	[00:00<?, ?it/s]	
0%	1/357	[00:00<01:37,	3.65it/s]
1%	2/357	[00:00<01:36,	3.68it/s]
1%	3/357	[00:00<01:35,	3.69it/s]
1%	4/357	[00:01<01:35,	3.69it/s]
1%	5/357	[00:01<01:35,	3.70it/s]
2%	6/357	[00:01<01:34,	3.70it/s]

2%	7/357	[00:01<01:34,	3.70it/s]
2%	8/357	[00:02<01:34,	3.70it/s]
3%	9/357	[00:02<01:33,	3.70it/s]
3%	10/357	[00:02<01:33,	3.70it/s]
3%	11/357	[00:02<01:33,	3.70it/s]
3%	12/357	[00:03<01:33,	3.70it/s]
4%	13/357	[00:03<01:32,	3.70it/s]
4%	14/357	[00:03<01:32,	3.71it/s]
4%	15/357	[00:04<01:32,	3.70it/s]
4%	16/357	[00:04<01:32,	3.70it/s]
5%	17/357	[00:04<01:31,	3.70it/s]
5%	18/357	[00:04<01:31,	3.70it/s]
5%	19/357	[00:05<01:31,	3.70it/s]
6%	20/357	[00:05<01:31,	3.70it/s]
6%	21/357	[00:05<01:30,	3.70it/s]
6%	22/357	[00:05<01:30,	3.70it/s]
6%	23/357	[00:06<01:30,	3.70it/s]
7%	24/357	[00:06<01:30,	3.70it/s]
7%	25/357	[00:06<01:29,	3.70it/s]
7%	26/357	[00:07<01:29,	3.70it/s]
8%	27/357	[00:07<01:29,	3.70it/s]
8%	28/357	[00:07<01:28,	3.70it/s]
8%	29/357	[00:07<01:28,	3.70it/s]
8%	30/357	[00:08<01:28,	3.70it/s]
9%	31/357	[00:08<01:27,	3.71it/s]
9%	32/357	[00:08<01:27,	3.70it/s]
9%	33/357	[00:08<01:27,	3.70it/s]
10%	34/357	[00:09<01:27,	3.70it/s]
10%	35/357	[00:09<01:26,	3.70it/s]
10%	36/357	[00:09<01:26,	3.70it/s]
10%	37/357	[00:09<01:26,	3.70it/s]
11%	38/357	[00:10<01:26,	3.70it/s]
11%	39/357	[00:10<01:25,	3.70it/s]
11%	40/357	[00:10<01:25,	3.70it/s]
11%	41/357	[00:11<01:25,	3.70it/s]
12%	42/357	[00:11<01:25,	3.70it/s]
12%	43/357	[00:11<01:24,	3.70it/s]
12%	44/357	[00:11<01:24,	3.70it/s]
13%	45/357	[00:12<01:24,	3.70it/s]
13%	46/357	[00:12<01:24,	3.70it/s]
13%	47/357	[00:12<01:23,	3.70it/s]
13%	48/357	[00:12<01:23,	3.70it/s]
14%	49/357	[00:13<01:23,	3.70it/s]
14%	50/357	[00:13<01:22,	3.70it/s]
14%	51/357	[00:13<01:22,	3.70it/s]
15%	52/357	[00:14<01:22,	3.70it/s]
15%	53/357	[00:14<01:22,	3.70it/s]
15%	54/357	[00:14<01:21,	3.70it/s]

15%	55/357	[00:14<01:21,	3.70it/s]
16%	56/357	[00:15<01:21,	3.70it/s]
16%	57/357	[00:15<01:21,	3.69it/s]
16%	58/357	[00:15<01:20,	3.70it/s]
17%	59/357	[00:15<01:24,	3.52it/s]
17%	60/357	[00:16<01:23,	3.57it/s]
17%	61/357	[00:16<01:22,	3.61it/s]
17%	62/357	[00:16<01:21,	3.63it/s]
18%	63/357	[00:17<01:20,	3.65it/s]
18%	64/357	[00:17<01:19,	3.66it/s]
18%	65/357	[00:17<01:19,	3.68it/s]
18%	66/357	[00:17<01:19,	3.68it/s]
19%	67/357	[00:18<01:18,	3.69it/s]
19%	68/357	[00:18<01:18,	3.69it/s]
19%	69/357	[00:18<01:18,	3.69it/s]
20%	70/357	[00:18<01:17,	3.69it/s]
20%	71/357	[00:19<01:17,	3.69it/s]
20%	72/357	[00:19<01:17,	3.69it/s]
20%	73/357	[00:19<01:16,	3.69it/s]
21%	74/357	[00:20<01:16,	3.69it/s]
21%	75/357	[00:20<01:16,	3.69it/s]
21%	76/357	[00:20<01:16,	3.69it/s]
22%	77/357	[00:20<01:15,	3.69it/s]
22%	78/357	[00:21<01:15,	3.69it/s]
22%	79/357	[00:21<01:15,	3.69it/s]
22%	80/357	[00:21<01:14,	3.70it/s]
23%	81/357	[00:21<01:14,	3.69it/s]
23%	82/357	[00:22<01:14,	3.69it/s]
23%	83/357	[00:22<01:14,	3.69it/s]
24%	84/357	[00:22<01:13,	3.69it/s]
24%	85/357	[00:23<01:13,	3.69it/s]
24%	86/357	[00:23<01:13,	3.69it/s]
24%	87/357	[00:23<01:13,	3.69it/s]
25%	88/357	[00:23<01:12,	3.69it/s]
25%	89/357	[00:24<01:12,	3.69it/s]
25%	90/357	[00:24<01:12,	3.69it/s]
25%	91/357	[00:24<01:12,	3.69it/s]
26%	92/357	[00:24<01:11,	3.69it/s]
26%	93/357	[00:25<01:11,	3.69it/s]
26%	94/357	[00:25<01:11,	3.68it/s]
27%	95/357	[00:25<01:11,	3.69it/s]
27%	96/357	[00:26<01:10,	3.69it/s]
27%	97/357	[00:26<01:10,	3.68it/s]
27%	98/357	[00:26<01:10,	3.68it/s]
28%	99/357	[00:26<01:10,	3.68it/s]
28%	100/357	[00:27<01:09,	3.68it/s]
28%	101/357	[00:27<01:09,	3.69it/s]
29%	102/357	[00:27<01:09,	3.69it/s]

29%	103/357	[00:27<01:08,	3.69it/s]
29%	104/357	[00:28<01:08,	3.69it/s]
29%	105/357	[00:28<01:08,	3.69it/s]
30%	106/357	[00:28<01:08,	3.69it/s]
30%	107/357	[00:28<01:07,	3.69it/s]
30%	108/357	[00:29<01:07,	3.69it/s]
31%	109/357	[00:29<01:07,	3.69it/s]
31%	110/357	[00:29<01:06,	3.69it/s]
31%	111/357	[00:30<01:06,	3.69it/s]
31%	112/357	[00:30<01:06,	3.69it/s]
32%	113/357	[00:30<01:06,	3.69it/s]
32%	114/357	[00:30<01:05,	3.69it/s]
32%	115/357	[00:31<01:05,	3.69it/s]
32%	116/357	[00:31<01:05,	3.69it/s]
33%	117/357	[00:31<01:05,	3.69it/s]
33%	118/357	[00:31<01:04,	3.69it/s]
33%	119/357	[00:32<01:04,	3.69it/s]
34%	120/357	[00:32<01:04,	3.69it/s]
34%	121/357	[00:32<01:03,	3.69it/s]
34%	122/357	[00:33<01:03,	3.69it/s]
34%	123/357	[00:33<01:03,	3.69it/s]
35%	124/357	[00:33<01:03,	3.69it/s]
35%	125/357	[00:33<01:02,	3.69it/s]
35%	126/357	[00:34<01:02,	3.69it/s]
36%	127/357	[00:34<01:02,	3.69it/s]
36%	128/357	[00:34<01:02,	3.69it/s]
36%	129/357	[00:34<01:01,	3.69it/s]
36%	130/357	[00:35<01:01,	3.69it/s]
37%	131/357	[00:35<01:01,	3.69it/s]
37%	132/357	[00:35<01:01,	3.69it/s]
37%	133/357	[00:36<01:00,	3.69it/s]
38%	134/357	[00:36<01:00,	3.69it/s]
38%	135/357	[00:36<01:00,	3.69it/s]
38%	136/357	[00:36<00:59,	3.69it/s]
38%	137/357	[00:37<00:59,	3.69it/s]
39%	138/357	[00:37<00:59,	3.69it/s]
39%	139/357	[00:37<00:59,	3.69it/s]
39%	140/357	[00:37<00:58,	3.69it/s]
39%	141/357	[00:38<00:58,	3.69it/s]
40%	142/357	[00:38<00:58,	3.69it/s]
40%	143/357	[00:38<00:57,	3.69it/s]
40%	144/357	[00:39<00:57,	3.69it/s]
41%	145/357	[00:39<00:57,	3.69it/s]
41%	146/357	[00:39<00:57,	3.69it/s]
41%	147/357	[00:39<00:56,	3.69it/s]
41%	148/357	[00:40<00:56,	3.69it/s]
42%	149/357	[00:40<00:56,	3.69it/s]
42%	150/357	[00:40<00:56,	3.69it/s]

42%	151/357	[00:40<00:55,	3.69it/s]
43%	152/357	[00:41<00:55,	3.69it/s]
43%	153/357	[00:41<00:55,	3.69it/s]
43%	154/357	[00:41<00:55,	3.69it/s]
43%	155/357	[00:42<00:54,	3.69it/s]
44%	156/357	[00:42<00:54,	3.69it/s]
44%	157/357	[00:42<00:54,	3.69it/s]
44%	158/357	[00:42<00:53,	3.69it/s]
45%	159/357	[00:43<00:53,	3.69it/s]
45%	160/357	[00:43<00:53,	3.69it/s]
45%	161/357	[00:43<00:53,	3.69it/s]
45%	162/357	[00:43<00:52,	3.68it/s]
46%	163/357	[00:44<00:52,	3.69it/s]
46%	164/357	[00:44<00:52,	3.69it/s]
46%	165/357	[00:44<00:52,	3.68it/s]
46%	166/357	[00:44<00:51,	3.68it/s]
47%	167/357	[00:45<00:51,	3.69it/s]
47%	168/357	[00:45<00:51,	3.67it/s]
47%	169/357	[00:45<00:51,	3.68it/s]
48%	170/357	[00:46<00:50,	3.68it/s]
48%	171/357	[00:46<00:50,	3.68it/s]
48%	172/357	[00:46<00:50,	3.68it/s]
48%	173/357	[00:46<00:49,	3.68it/s]
49%	174/357	[00:47<00:49,	3.69it/s]
49%	175/357	[00:47<00:49,	3.68it/s]
49%	176/357	[00:47<00:49,	3.68it/s]
50%	177/357	[00:47<00:48,	3.68it/s]
50%	178/357	[00:48<00:48,	3.69it/s]
50%	179/357	[00:48<00:48,	3.69it/s]
50%	180/357	[00:48<00:48,	3.68it/s]
51%	181/357	[00:49<00:47,	3.69it/s]
51%	182/357	[00:49<00:47,	3.69it/s]
51%	183/357	[00:49<00:47,	3.69it/s]
52%	184/357	[00:49<00:46,	3.69it/s]
52%	185/357	[00:50<00:46,	3.69it/s]
52%	186/357	[00:50<00:46,	3.69it/s]
52%	187/357	[00:50<00:46,	3.69it/s]
53%	188/357	[00:50<00:45,	3.69it/s]
53%	189/357	[00:51<00:45,	3.69it/s]
53%	190/357	[00:51<00:45,	3.69it/s]
54%	191/357	[00:51<00:45,	3.68it/s]
54%	192/357	[00:52<00:45,	3.65it/s]
54%	193/357	[00:52<00:44,	3.67it/s]
54%	194/357	[00:52<00:44,	3.67it/s]
55%	195/357	[00:52<00:44,	3.64it/s]
55%	196/357	[00:53<00:44,	3.65it/s]
55%	197/357	[00:53<00:43,	3.66it/s]
55%	198/357	[00:53<00:43,	3.67it/s]



56%	199/357	[00:53<00:43,	3.67it/s]
56%	200/357	[00:54<00:42,	3.68it/s]
56%	201/357	[00:54<00:42,	3.68it/s]
57%	202/357	[00:54<00:42,	3.68it/s]
57%	203/357	[00:55<00:41,	3.68it/s]
57%	204/357	[00:55<00:41,	3.68it/s]
57%	205/357	[00:55<00:41,	3.68it/s]
58%	206/357	[00:55<00:40,	3.69it/s]
58%	207/357	[00:56<00:40,	3.68it/s]
58%	208/357	[00:56<00:40,	3.68it/s]
59%	209/357	[00:56<00:40,	3.63it/s]
59%	210/357	[00:56<00:40,	3.64it/s]
59%	211/357	[00:57<00:39,	3.65it/s]
59%	212/357	[00:57<00:39,	3.66it/s]
60%	213/357	[00:57<00:39,	3.65it/s]
60%	214/357	[00:58<00:39,	3.66it/s]
60%	215/357	[00:58<00:38,	3.67it/s]
61%	216/357	[00:59<01:10,	2.00it/s]
61%	217/357	[00:59<01:00,	2.32it/s]
61%	218/357	[00:59<00:53,	2.61it/s]
61%	219/357	[01:00<00:48,	2.86it/s]
62%	220/357	[01:00<00:44,	3.07it/s]
62%	221/357	[01:00<00:42,	3.23it/s]
62%	222/357	[01:00<00:40,	3.36it/s]
62%	223/357	[01:01<00:38,	3.45it/s]
63%	224/357	[01:01<00:37,	3.52it/s]
63%	225/357	[01:01<00:36,	3.57it/s]
63%	226/357	[01:02<00:36,	3.61it/s]
64%	227/357	[01:02<00:35,	3.63it/s]
64%	228/357	[01:02<00:35,	3.64it/s]
64%	229/357	[01:02<00:35,	3.66it/s]
64%	230/357	[01:03<00:34,	3.66it/s]
65%	231/357	[01:03<00:34,	3.67it/s]
65%	232/357	[01:03<00:36,	3.44it/s]
65%	233/357	[01:04<00:35,	3.51it/s]
66%	234/357	[01:04<00:34,	3.56it/s]
66%	235/357	[01:04<00:33,	3.59it/s]
66%	236/357	[01:04<00:33,	3.62it/s]
66%	237/357	[01:05<00:32,	3.64it/s]
67%	238/357	[01:05<00:32,	3.66it/s]
67%	239/357	[01:05<00:32,	3.62it/s]
67%	240/357	[01:05<00:32,	3.64it/s]
68%	241/357	[01:06<00:31,	3.65it/s]
68%	242/357	[01:06<00:31,	3.67it/s]
68%	243/357	[01:06<00:31,	3.67it/s]
68%	244/357	[01:07<00:30,	3.67it/s]
69%	245/357	[01:07<00:30,	3.68it/s]
69%	246/357	[01:07<00:30,	3.68it/s]

69%		247/357	[01:07<00:30,	3.60it/s]
69%		248/357	[01:08<00:30,	3.57it/s]
70%		249/357	[01:08<00:30,	3.56it/s]
70%		250/357	[01:08<00:29,	3.60it/s]
70%		251/357	[01:08<00:29,	3.62it/s]
71%		252/357	[01:09<00:28,	3.64it/s]
71%		253/357	[01:09<00:28,	3.63it/s]
71%		254/357	[01:09<00:28,	3.63it/s]
71%		255/357	[01:10<00:28,	3.59it/s]
72%		256/357	[01:10<00:27,	3.61it/s]
72%		257/357	[01:10<00:27,	3.63it/s]
72%		258/357	[01:10<00:27,	3.65it/s]
73%		259/357	[01:11<00:26,	3.64it/s]
73%		260/357	[01:11<00:26,	3.64it/s]
73%		261/357	[01:11<00:26,	3.63it/s]
73%		262/357	[01:11<00:26,	3.65it/s]
74%		263/357	[01:12<00:25,	3.66it/s]
74%		264/357	[01:12<00:25,	3.66it/s]
74%		265/357	[01:12<00:25,	3.65it/s]
75%		266/357	[01:13<00:24,	3.64it/s]
75%		267/357	[01:13<00:24,	3.66it/s]
75%		268/357	[01:13<00:24,	3.67it/s]
75%		269/357	[01:13<00:24,	3.66it/s]
76%		270/357	[01:14<00:23,	3.63it/s]
76%		271/357	[01:14<00:23,	3.60it/s]
76%		272/357	[01:14<00:23,	3.58it/s]
76%		273/357	[01:15<00:23,	3.55it/s]
77%		274/357	[01:15<00:23,	3.59it/s]
77%		275/357	[01:15<00:22,	3.58it/s]
77%		276/357	[01:15<00:22,	3.57it/s]
78%		277/357	[01:16<00:22,	3.56it/s]
78%		278/357	[01:16<00:22,	3.57it/s]
78%		279/357	[01:16<00:21,	3.58it/s]
78%		280/357	[01:17<00:22,	3.49it/s]
79%		281/357	[01:17<00:21,	3.53it/s]
79%		282/357	[01:17<00:21,	3.55it/s]
79%		283/357	[01:17<00:21,	3.40it/s]
80%		284/357	[01:18<00:21,	3.44it/s]
80%		285/357	[01:18<00:20,	3.47it/s]
80%		286/357	[01:18<00:20,	3.51it/s]
80%		287/357	[01:19<00:19,	3.52it/s]
81%		288/357	[01:19<00:19,	3.55it/s]
81%		289/357	[01:19<00:19,	3.57it/s]
81%		290/357	[01:19<00:18,	3.54it/s]
82%		291/357	[01:20<00:18,	3.58it/s]
82%		292/357	[01:20<00:18,	3.55it/s]
82%		293/357	[01:20<00:18,	3.53it/s]
82%		294/357	[01:20<00:17,	3.51it/s]

83%		295/357	[01:21<00:17,	3.53it/s]
83%		296/357	[01:21<00:17,	3.51it/s]
83%		297/357	[01:21<00:17,	3.52it/s]
83%		298/357	[01:22<00:16,	3.53it/s]
84%		299/357	[01:22<00:16,	3.54it/s]
84%		300/357	[01:22<00:16,	3.48it/s]
84%		301/357	[01:22<00:16,	3.49it/s]
85%		302/357	[01:23<00:15,	3.50it/s]
85%		303/357	[01:23<00:15,	3.49it/s]
85%		304/357	[01:23<00:15,	3.50it/s]
85%		305/357	[01:24<00:14,	3.54it/s]
86%		306/357	[01:24<00:14,	3.54it/s]
86%		307/357	[01:24<00:14,	3.53it/s]
86%		308/357	[01:24<00:13,	3.52it/s]
87%		309/357	[01:25<00:13,	3.49it/s]
87%		310/357	[01:25<00:13,	3.47it/s]
87%		311/357	[01:25<00:13,	3.48it/s]
87%		312/357	[01:26<00:12,	3.48it/s]
88%		313/357	[01:26<00:12,	3.46it/s]
88%		314/357	[01:26<00:12,	3.44it/s]
88%		315/357	[01:27<00:12,	3.40it/s]
89%		316/357	[01:27<00:11,	3.42it/s]
89%		317/357	[01:27<00:11,	3.40it/s]
89%		318/357	[01:27<00:11,	3.32it/s]
89%		319/357	[01:28<00:11,	3.32it/s]
90%		320/357	[01:28<00:11,	3.33it/s]
90%		321/357	[01:28<00:10,	3.37it/s]
90%		322/357	[01:29<00:10,	3.39it/s]
90%		323/357	[01:29<00:10,	3.36it/s]
91%		324/357	[01:29<00:09,	3.39it/s]
91%		325/357	[01:29<00:09,	3.39it/s]
91%		326/357	[01:30<00:09,	3.40it/s]
92%		327/357	[01:30<00:08,	3.37it/s]
92%		328/357	[01:30<00:08,	3.23it/s]
92%		329/357	[01:31<00:08,	3.28it/s]
92%		330/357	[01:31<00:08,	3.32it/s]
93%		331/357	[01:31<00:07,	3.34it/s]
93%		332/357	[01:32<00:07,	3.34it/s]
93%		333/357	[01:32<00:07,	3.31it/s]
94%		334/357	[01:32<00:06,	3.32it/s]
94%		335/357	[01:33<00:06,	3.34it/s]
94%		336/357	[01:33<00:06,	3.37it/s]
94%		337/357	[01:33<00:05,	3.33it/s]
95%		338/357	[01:33<00:05,	3.34it/s]
95%		339/357	[01:34<00:05,	3.36it/s]
95%		340/357	[01:34<00:05,	3.35it/s]
96%		341/357	[01:34<00:04,	3.34it/s]
96%		342/357	[01:35<00:04,	3.34it/s]

```

96%|      | 343/357 [01:35<00:04,  3.34it/s]
96%|      | 344/357 [01:35<00:03,  3.34it/s]
97%|      | 345/357 [01:35<00:03,  3.36it/s]
97%|      | 346/357 [01:36<00:03,  3.34it/s]
97%|      | 347/357 [01:36<00:03,  3.33it/s]
97%|      | 348/357 [01:36<00:02,  3.30it/s]
98%|      | 349/357 [01:37<00:02,  3.26it/s]
98%|      | 350/357 [01:37<00:02,  3.26it/s]
98%|      | 351/357 [01:37<00:01,  3.26it/s]
99%|      | 352/357 [01:38<00:01,  3.27it/s]
99%|      | 353/357 [01:38<00:01,  3.28it/s]
99%|      | 354/357 [01:38<00:00,  3.27it/s]
99%|      | 355/357 [01:39<00:00,  3.25it/s]
100%|     | 356/357 [01:39<00:00,  3.26it/s]
100%|     | 357/357 [01:39<00:00,  3.58it/s]
70%|      | 7/10 [12:06<05:08, 102.95s/it]

```

For - 6 Loss: tensor(0.0004, device='cuda:0', grad\_fn=<MeanBackward0>)

```

0%|      | 0/357 [00:00<?, ?it/s]
0%|      | 1/357 [00:00<01:42,  3.48it/s]
1%|      | 2/357 [00:00<01:38,  3.60it/s]
1%|      | 3/357 [00:00<01:37,  3.64it/s]
1%|      | 4/357 [00:01<01:36,  3.65it/s]
1%|      | 5/357 [00:01<01:36,  3.67it/s]
2%|      | 6/357 [00:01<01:35,  3.67it/s]
2%|      | 7/357 [00:01<01:35,  3.68it/s]
2%|      | 8/357 [00:02<01:34,  3.68it/s]
3%|      | 9/357 [00:02<01:34,  3.68it/s]
3%|      | 10/357 [00:02<01:34,  3.68it/s]
3%|      | 11/357 [00:03<01:33,  3.68it/s]
3%|      | 12/357 [00:03<01:33,  3.68it/s]
4%|      | 13/357 [00:03<01:33,  3.68it/s]
4%|      | 14/357 [00:03<01:33,  3.68it/s]
4%|      | 15/357 [00:04<01:32,  3.69it/s]
4%|      | 16/357 [00:04<01:32,  3.69it/s]
5%|      | 17/357 [00:04<01:32,  3.69it/s]
5%|      | 18/357 [00:04<01:31,  3.69it/s]
5%|      | 19/357 [00:05<01:31,  3.69it/s]
6%|      | 20/357 [00:05<01:31,  3.69it/s]
6%|      | 21/357 [00:05<01:31,  3.69it/s]
6%|      | 22/357 [00:05<01:30,  3.69it/s]
6%|      | 23/357 [00:06<01:30,  3.69it/s]
7%|      | 24/357 [00:06<01:30,  3.68it/s]
7%|      | 25/357 [00:06<01:30,  3.69it/s]
7%|      | 26/357 [00:07<01:29,  3.69it/s]
8%|      | 27/357 [00:07<01:29,  3.69it/s]
8%|      | 28/357 [00:07<01:29,  3.69it/s]

```

8%	29/357	[00:07<01:28,	3.69it/s]
8%	30/357	[00:08<01:28,	3.69it/s]
9%	31/357	[00:08<01:28,	3.68it/s]
9%	32/357	[00:08<01:28,	3.68it/s]
9%	33/357	[00:08<01:27,	3.68it/s]
10%	34/357	[00:09<01:27,	3.69it/s]
10%	35/357	[00:09<01:27,	3.69it/s]
10%	36/357	[00:09<01:27,	3.69it/s]
10%	37/357	[00:10<01:26,	3.69it/s]
11%	38/357	[00:10<01:26,	3.69it/s]
11%	39/357	[00:10<01:26,	3.68it/s]
11%	40/357	[00:10<01:26,	3.68it/s]
11%	41/357	[00:11<01:25,	3.69it/s]
12%	42/357	[00:11<01:25,	3.69it/s]
12%	43/357	[00:11<01:25,	3.69it/s]
12%	44/357	[00:11<01:25,	3.67it/s]
13%	45/357	[00:12<01:24,	3.67it/s]
13%	46/357	[00:12<01:24,	3.68it/s]
13%	47/357	[00:12<01:24,	3.68it/s]
13%	48/357	[00:13<01:23,	3.68it/s]
14%	49/357	[00:13<01:23,	3.69it/s]
14%	50/357	[00:13<01:23,	3.67it/s]
14%	51/357	[00:13<01:23,	3.68it/s]
15%	52/357	[00:14<01:22,	3.68it/s]
15%	53/357	[00:14<01:22,	3.69it/s]
15%	54/357	[00:14<01:22,	3.69it/s]
15%	55/357	[00:14<01:21,	3.69it/s]
16%	56/357	[00:15<01:21,	3.69it/s]
16%	57/357	[00:15<01:21,	3.69it/s]
16%	58/357	[00:15<01:20,	3.69it/s]
17%	59/357	[00:16<01:20,	3.69it/s]
17%	60/357	[00:16<01:20,	3.69it/s]
17%	61/357	[00:16<01:20,	3.69it/s]
17%	62/357	[00:16<01:20,	3.68it/s]
18%	63/357	[00:17<01:19,	3.68it/s]
18%	64/357	[00:17<01:19,	3.68it/s]
18%	65/357	[00:18<02:29,	1.95it/s]
18%	66/357	[00:18<02:07,	2.28it/s]
19%	67/357	[00:18<01:52,	2.57it/s]
19%	68/357	[00:19<01:42,	2.83it/s]
19%	69/357	[00:19<01:34,	3.04it/s]
20%	70/357	[00:19<01:29,	3.21it/s]
20%	71/357	[00:20<01:25,	3.34it/s]
20%	72/357	[00:20<01:22,	3.44it/s]
20%	73/357	[00:20<01:20,	3.51it/s]
21%	74/357	[00:20<01:19,	3.56it/s]
21%	75/357	[00:21<01:18,	3.60it/s]
21%	76/357	[00:21<01:17,	3.62it/s]

22%	77/357	[00:21<01:16,	3.65it/s]
22%	78/357	[00:21<01:16,	3.66it/s]
22%	79/357	[00:22<01:15,	3.66it/s]
22%	80/357	[00:22<01:15,	3.67it/s]
23%	81/357	[00:22<01:15,	3.68it/s]
23%	82/357	[00:23<01:14,	3.68it/s]
23%	83/357	[00:23<01:14,	3.68it/s]
24%	84/357	[00:23<01:14,	3.68it/s]
24%	85/357	[00:23<01:13,	3.68it/s]
24%	86/357	[00:24<01:13,	3.68it/s]
24%	87/357	[00:24<01:13,	3.68it/s]
25%	88/357	[00:24<01:13,	3.68it/s]
25%	89/357	[00:24<01:12,	3.69it/s]
25%	90/357	[00:25<01:12,	3.69it/s]
25%	91/357	[00:25<01:12,	3.69it/s]
26%	92/357	[00:25<01:11,	3.69it/s]
26%	93/357	[00:26<01:11,	3.68it/s]
26%	94/357	[00:26<01:11,	3.68it/s]
27%	95/357	[00:26<01:11,	3.68it/s]
27%	96/357	[00:26<01:10,	3.68it/s]
27%	97/357	[00:27<01:10,	3.68it/s]
27%	98/357	[00:27<01:10,	3.68it/s]
28%	99/357	[00:27<01:10,	3.68it/s]
28%	100/357	[00:27<01:09,	3.68it/s]
28%	101/357	[00:28<01:09,	3.68it/s]
29%	102/357	[00:28<01:09,	3.68it/s]
29%	103/357	[00:28<01:09,	3.68it/s]
29%	104/357	[00:29<01:08,	3.68it/s]
29%	105/357	[00:29<01:08,	3.68it/s]
30%	106/357	[00:29<01:08,	3.68it/s]
30%	107/357	[00:29<01:07,	3.68it/s]
30%	108/357	[00:30<01:07,	3.68it/s]
31%	109/357	[00:30<01:07,	3.69it/s]
31%	110/357	[00:30<01:07,	3.69it/s]
31%	111/357	[00:30<01:06,	3.69it/s]
31%	112/357	[00:31<01:06,	3.68it/s]
32%	113/357	[00:31<01:06,	3.68it/s]
32%	114/357	[00:31<01:05,	3.69it/s]
32%	115/357	[00:32<01:05,	3.69it/s]
32%	116/357	[00:32<01:05,	3.68it/s]
33%	117/357	[00:32<01:05,	3.69it/s]
33%	118/357	[00:32<01:04,	3.69it/s]
33%	119/357	[00:33<01:04,	3.68it/s]
34%	120/357	[00:33<01:04,	3.68it/s]
34%	121/357	[00:33<01:04,	3.68it/s]
34%	122/357	[00:33<01:03,	3.68it/s]
34%	123/357	[00:34<01:03,	3.68it/s]
35%	124/357	[00:34<01:03,	3.68it/s]

35%	125/357	[00:34<01:03,	3.68it/s]
35%	126/357	[00:35<01:02,	3.68it/s]
36%	127/357	[00:35<01:02,	3.68it/s]
36%	128/357	[00:35<01:02,	3.68it/s]
36%	129/357	[00:35<01:01,	3.68it/s]
36%	130/357	[00:36<01:01,	3.68it/s]
37%	131/357	[00:36<01:01,	3.68it/s]
37%	132/357	[00:36<01:01,	3.68it/s]
37%	133/357	[00:36<01:00,	3.68it/s]
38%	134/357	[00:37<01:00,	3.68it/s]
38%	135/357	[00:37<01:00,	3.68it/s]
38%	136/357	[00:37<00:59,	3.69it/s]
38%	137/357	[00:37<00:59,	3.69it/s]
39%	138/357	[00:38<00:59,	3.68it/s]
39%	139/357	[00:38<00:59,	3.68it/s]
39%	140/357	[00:38<00:58,	3.68it/s]
39%	141/357	[00:39<00:58,	3.68it/s]
40%	142/357	[00:39<01:19,	2.69it/s]
40%	143/357	[00:39<01:13,	2.93it/s]
40%	144/357	[00:40<01:08,	3.12it/s]
41%	145/357	[00:40<01:04,	3.27it/s]
41%	146/357	[00:40<01:02,	3.38it/s]
41%	147/357	[00:41<01:00,	3.47it/s]
41%	148/357	[00:41<00:59,	3.52it/s]
42%	149/357	[00:41<00:59,	3.51it/s]
42%	150/357	[00:41<00:58,	3.56it/s]
42%	151/357	[00:42<00:57,	3.60it/s]
43%	152/357	[00:42<00:56,	3.62it/s]
43%	153/357	[00:42<00:56,	3.64it/s]
43%	154/357	[00:42<00:55,	3.65it/s]
43%	155/357	[00:43<00:55,	3.66it/s]
44%	156/357	[00:43<00:54,	3.67it/s]
44%	157/357	[00:43<00:54,	3.67it/s]
44%	158/357	[00:44<00:54,	3.68it/s]
45%	159/357	[00:44<00:53,	3.68it/s]
45%	160/357	[00:44<00:53,	3.68it/s]
45%	161/357	[00:44<00:53,	3.68it/s]
45%	162/357	[00:45<00:52,	3.68it/s]
46%	163/357	[00:45<00:52,	3.68it/s]
46%	164/357	[00:45<00:52,	3.68it/s]
46%	165/357	[00:45<00:52,	3.69it/s]
46%	166/357	[00:46<00:51,	3.68it/s]
47%	167/357	[00:46<00:51,	3.69it/s]
47%	168/357	[00:46<00:51,	3.68it/s]
47%	169/357	[00:47<00:51,	3.69it/s]
48%	170/357	[00:47<00:50,	3.69it/s]
48%	171/357	[00:47<00:50,	3.69it/s]
48%	172/357	[00:47<00:50,	3.68it/s]

48%	173/357	[00:48<00:49,	3.68it/s]
49%	174/357	[00:48<00:49,	3.69it/s]
49%	175/357	[00:48<00:49,	3.68it/s]
49%	176/357	[00:48<00:49,	3.68it/s]
50%	177/357	[00:49<00:48,	3.68it/s]
50%	178/357	[00:49<00:48,	3.68it/s]
50%	179/357	[00:49<00:48,	3.69it/s]
50%	180/357	[00:50<00:48,	3.68it/s]
51%	181/357	[00:50<00:47,	3.68it/s]
51%	182/357	[00:50<00:47,	3.68it/s]
51%	183/357	[00:50<00:47,	3.68it/s]
52%	184/357	[00:51<00:46,	3.68it/s]
52%	185/357	[00:51<00:46,	3.68it/s]
52%	186/357	[00:51<00:46,	3.68it/s]
52%	187/357	[00:51<00:46,	3.68it/s]
53%	188/357	[00:52<00:45,	3.68it/s]
53%	189/357	[00:52<00:45,	3.68it/s]
53%	190/357	[00:52<00:45,	3.68it/s]
54%	191/357	[00:53<00:45,	3.68it/s]
54%	192/357	[00:53<00:44,	3.68it/s]
54%	193/357	[00:53<00:44,	3.68it/s]
54%	194/357	[00:53<00:45,	3.62it/s]
55%	195/357	[00:54<00:47,	3.40it/s]
55%	196/357	[00:54<00:46,	3.46it/s]
55%	197/357	[00:54<00:45,	3.52it/s]
55%	198/357	[00:54<00:44,	3.57it/s]
56%	199/357	[00:55<00:43,	3.60it/s]
56%	200/357	[00:55<00:43,	3.62it/s]
56%	201/357	[00:55<00:42,	3.64it/s]
57%	202/357	[00:56<00:42,	3.65it/s]
57%	203/357	[00:56<00:42,	3.66it/s]
57%	204/357	[00:56<00:41,	3.66it/s]
57%	205/357	[00:56<00:41,	3.66it/s]
58%	206/357	[00:57<00:41,	3.67it/s]
58%	207/357	[00:57<00:40,	3.67it/s]
58%	208/357	[00:57<00:40,	3.68it/s]
59%	209/357	[00:57<00:40,	3.68it/s]
59%	210/357	[00:58<00:39,	3.68it/s]
59%	211/357	[00:58<00:39,	3.66it/s]
59%	212/357	[00:58<00:39,	3.66it/s]
60%	213/357	[00:59<00:39,	3.67it/s]
60%	214/357	[00:59<00:38,	3.67it/s]
60%	215/357	[00:59<00:38,	3.68it/s]
61%	216/357	[00:59<00:38,	3.67it/s]
61%	217/357	[01:00<00:39,	3.56it/s]
61%	218/357	[01:00<00:38,	3.59it/s]
61%	219/357	[01:00<00:38,	3.62it/s]
62%	220/357	[01:01<00:37,	3.64it/s]



62%		221/357	[01:01<00:37,	3.65it/s]
62%		222/357	[01:01<00:36,	3.66it/s]
62%		223/357	[01:01<00:36,	3.67it/s]
63%		224/357	[01:02<00:36,	3.67it/s]
63%		225/357	[01:02<00:35,	3.67it/s]
63%		226/357	[01:02<00:35,	3.68it/s]
64%		227/357	[01:02<00:35,	3.68it/s]
64%		228/357	[01:03<00:35,	3.68it/s]
64%		229/357	[01:03<00:34,	3.68it/s]
64%		230/357	[01:03<00:34,	3.68it/s]
65%		231/357	[01:03<00:34,	3.68it/s]
65%		232/357	[01:04<00:33,	3.68it/s]
65%		233/357	[01:04<00:33,	3.68it/s]
66%		234/357	[01:04<00:33,	3.68it/s]
66%		235/357	[01:05<00:33,	3.68it/s]
66%		236/357	[01:05<00:32,	3.68it/s]
66%		237/357	[01:05<00:32,	3.68it/s]
67%		238/357	[01:05<00:32,	3.68it/s]
67%		239/357	[01:06<00:32,	3.65it/s]
67%		240/357	[01:06<00:34,	3.42it/s]
68%		241/357	[01:06<00:33,	3.48it/s]
68%		242/357	[01:07<00:32,	3.53it/s]
68%		243/357	[01:07<00:31,	3.57it/s]
68%		244/357	[01:07<00:31,	3.61it/s]
69%		245/357	[01:07<00:30,	3.62it/s]
69%		246/357	[01:08<00:30,	3.62it/s]
69%		247/357	[01:08<00:30,	3.63it/s]
69%		248/357	[01:08<00:30,	3.63it/s]
70%		249/357	[01:08<00:29,	3.63it/s]
70%		250/357	[01:09<00:29,	3.61it/s]
70%		251/357	[01:09<00:29,	3.63it/s]
71%		252/357	[01:09<00:28,	3.64it/s]
71%		253/357	[01:10<00:28,	3.63it/s]
71%		254/357	[01:10<00:28,	3.64it/s]
71%		255/357	[01:10<00:28,	3.62it/s]
72%		256/357	[01:10<00:27,	3.62it/s]
72%		257/357	[01:11<00:27,	3.62it/s]
72%		258/357	[01:11<00:27,	3.63it/s]
73%		259/357	[01:11<00:26,	3.64it/s]
73%		260/357	[01:12<00:26,	3.64it/s]
73%		261/357	[01:12<00:26,	3.62it/s]
73%		262/357	[01:12<00:26,	3.60it/s]
74%		263/357	[01:12<00:26,	3.59it/s]
74%		264/357	[01:13<00:25,	3.59it/s]
74%		265/357	[01:13<00:25,	3.59it/s]
75%		266/357	[01:13<00:25,	3.56it/s]
75%		267/357	[01:13<00:25,	3.55it/s]
75%		268/357	[01:14<00:25,	3.54it/s]

75%	269/357	[01:14<00:24,	3.53it/s]
76%	270/357	[01:14<00:24,	3.57it/s]
76%	271/357	[01:15<00:24,	3.56it/s]
76%	272/357	[01:15<00:23,	3.57it/s]
76%	273/357	[01:15<00:23,	3.56it/s]
77%	274/357	[01:15<00:23,	3.59it/s]
77%	275/357	[01:16<00:23,	3.56it/s]
77%	276/357	[01:16<00:23,	3.52it/s]
78%	277/357	[01:16<00:22,	3.53it/s]
78%	278/357	[01:17<00:22,	3.56it/s]
78%	279/357	[01:17<00:21,	3.55it/s]
78%	280/357	[01:17<00:21,	3.53it/s]
79%	281/357	[01:17<00:21,	3.52it/s]
79%	282/357	[01:18<00:21,	3.51it/s]
79%	283/357	[01:18<00:21,	3.51it/s]
80%	284/357	[01:18<00:20,	3.48it/s]
80%	285/357	[01:19<00:21,	3.38it/s]
80%	286/357	[01:19<00:20,	3.42it/s]
80%	287/357	[01:19<00:20,	3.45it/s]
81%	288/357	[01:19<00:19,	3.49it/s]
81%	289/357	[01:20<00:19,	3.50it/s]
81%	290/357	[01:20<00:19,	3.48it/s]
82%	291/357	[01:20<00:18,	3.50it/s]
82%	292/357	[01:21<00:18,	3.52it/s]
82%	293/357	[01:21<00:18,	3.52it/s]
82%	294/357	[01:21<00:18,	3.46it/s]
83%	295/357	[01:21<00:18,	3.44it/s]
83%	296/357	[01:22<00:17,	3.46it/s]
83%	297/357	[01:22<00:17,	3.48it/s]
83%	298/357	[01:22<00:16,	3.48it/s]
84%	299/357	[01:23<00:16,	3.49it/s]
84%	300/357	[01:23<00:16,	3.49it/s]
84%	301/357	[01:23<00:16,	3.49it/s]
85%	302/357	[01:23<00:15,	3.46it/s]
85%	303/357	[01:24<00:15,	3.42it/s]
85%	304/357	[01:24<00:15,	3.43it/s]
85%	305/357	[01:24<00:15,	3.43it/s]
86%	306/357	[01:25<00:14,	3.44it/s]
86%	307/357	[01:25<00:14,	3.42it/s]
86%	308/357	[01:25<00:14,	3.40it/s]
87%	309/357	[01:26<00:13,	3.43it/s]
87%	310/357	[01:26<00:13,	3.43it/s]
87%	311/357	[01:26<00:13,	3.44it/s]
87%	312/357	[01:26<00:13,	3.46it/s]
88%	313/357	[01:27<00:12,	3.44it/s]
88%	314/357	[01:27<00:12,	3.42it/s]
88%	315/357	[01:27<00:12,	3.40it/s]
89%	316/357	[01:28<00:12,	3.39it/s]

```

89%|      | 317/357 [01:28<00:11,  3.35it/s]
89%|      | 318/357 [01:28<00:11,  3.32it/s]
89%|      | 319/357 [01:28<00:11,  3.32it/s]
90%|      | 320/357 [01:30<00:20,  1.82it/s]
90%|      | 321/357 [01:30<00:17,  2.11it/s]
90%|      | 322/357 [01:30<00:15,  2.33it/s]
90%|      | 323/357 [01:31<00:13,  2.58it/s]
91%|      | 324/357 [01:31<00:12,  2.75it/s]
91%|      | 325/357 [01:31<00:10,  2.93it/s]
91%|      | 326/357 [01:31<00:10,  3.06it/s]
92%|      | 327/357 [01:32<00:09,  3.16it/s]
92%|      | 328/357 [01:32<00:08,  3.25it/s]
92%|      | 329/357 [01:32<00:08,  3.30it/s]
92%|      | 330/357 [01:33<00:08,  3.32it/s]
93%|      | 331/357 [01:33<00:07,  3.32it/s]
93%|      | 332/357 [01:33<00:07,  3.31it/s]
93%|      | 333/357 [01:33<00:07,  3.34it/s]
94%|      | 334/357 [01:34<00:06,  3.38it/s]
94%|      | 335/357 [01:34<00:06,  3.36it/s]
94%|      | 336/357 [01:34<00:06,  3.32it/s]
94%|      | 337/357 [01:35<00:05,  3.34it/s]
95%|      | 338/357 [01:35<00:05,  3.35it/s]
95%|      | 339/357 [01:35<00:05,  3.34it/s]
95%|      | 340/357 [01:36<00:05,  2.98it/s]
96%|      | 341/357 [01:36<00:05,  3.07it/s]
96%|      | 342/357 [01:36<00:04,  3.12it/s]
96%|      | 343/357 [01:37<00:04,  3.18it/s]
96%|      | 344/357 [01:37<00:04,  3.23it/s]
97%|      | 345/357 [01:37<00:03,  3.25it/s]
97%|      | 346/357 [01:38<00:03,  3.22it/s]
97%|      | 347/357 [01:38<00:03,  3.25it/s]
97%|      | 348/357 [01:38<00:02,  3.23it/s]
98%|      | 349/357 [01:38<00:02,  3.20it/s]
98%|      | 350/357 [01:39<00:02,  3.21it/s]
98%|      | 351/357 [01:39<00:01,  3.20it/s]
99%|      | 352/357 [01:39<00:01,  3.23it/s]
99%|      | 353/357 [01:40<00:01,  3.23it/s]
99%|      | 354/357 [01:40<00:00,  3.23it/s]
99%|      | 355/357 [01:40<00:00,  3.25it/s]
100%|     | 356/357 [01:41<00:00,  3.23it/s]
100%|     | 357/357 [01:41<00:00,  3.52it/s]
80%|      | 8/10 [13:48<03:25, 102.55s/it]

```

For - 7 Loss: tensor(0.0003, device='cuda:0', grad\_fn=<MeanBackward0>)

```

0%|      | 0/357 [00:00<?, ?it/s]
0%|      | 1/357 [00:00<01:40,  3.54it/s]
1%|      | 2/357 [00:00<01:37,  3.62it/s]

```

1%	3/357 [00:00<01:36, 3.65it/s]
1%	4/357 [00:01<01:36, 3.66it/s]
1%	5/357 [00:01<01:35, 3.67it/s]
2%	6/357 [00:01<01:35, 3.68it/s]
2%	7/357 [00:01<01:35, 3.68it/s]
2%	8/357 [00:02<01:34, 3.68it/s]
3%	9/357 [00:02<01:34, 3.68it/s]
3%	10/357 [00:02<01:34, 3.69it/s]
3%	11/357 [00:02<01:33, 3.69it/s]
3%	12/357 [00:03<01:33, 3.69it/s]
4%	13/357 [00:03<01:33, 3.68it/s]
4%	14/357 [00:03<01:33, 3.69it/s]
4%	15/357 [00:04<01:32, 3.68it/s]
4%	16/357 [00:04<01:32, 3.68it/s]
5%	17/357 [00:04<01:32, 3.69it/s]
5%	18/357 [00:04<01:31, 3.69it/s]
5%	19/357 [00:05<01:31, 3.69it/s]
6%	20/357 [00:05<01:31, 3.68it/s]
6%	21/357 [00:05<01:31, 3.68it/s]
6%	22/357 [00:05<01:30, 3.68it/s]
6%	23/357 [00:06<01:30, 3.68it/s]
7%	24/357 [00:06<01:30, 3.68it/s]
7%	25/357 [00:06<01:30, 3.68it/s]
7%	26/357 [00:07<01:29, 3.69it/s]
8%	27/357 [00:07<01:29, 3.69it/s]
8%	28/357 [00:07<01:29, 3.68it/s]
8%	29/357 [00:07<01:29, 3.68it/s]
8%	30/357 [00:08<01:28, 3.68it/s]
9%	31/357 [00:08<01:28, 3.69it/s]
9%	32/357 [00:08<01:28, 3.68it/s]
9%	33/357 [00:08<01:27, 3.69it/s]
10%	34/357 [00:09<01:27, 3.68it/s]
10%	35/357 [00:09<01:27, 3.69it/s]
10%	36/357 [00:09<01:27, 3.68it/s]
10%	37/357 [00:10<01:26, 3.68it/s]
11%	38/357 [00:10<01:26, 3.68it/s]
11%	39/357 [00:10<01:26, 3.69it/s]
11%	40/357 [00:10<01:26, 3.68it/s]
11%	41/357 [00:11<01:25, 3.68it/s]
12%	42/357 [00:11<01:25, 3.69it/s]
12%	43/357 [00:11<01:25, 3.68it/s]
12%	44/357 [00:11<01:24, 3.69it/s]
13%	45/357 [00:12<01:24, 3.69it/s]
13%	46/357 [00:12<01:25, 3.65it/s]
13%	47/357 [00:12<01:24, 3.66it/s]
13%	48/357 [00:13<01:24, 3.67it/s]
14%	49/357 [00:13<01:23, 3.67it/s]
14%	50/357 [00:13<01:23, 3.68it/s]

14%		51/357	[00:13<01:23,	3.68it/s]
15%		52/357	[00:14<01:22,	3.68it/s]
15%		53/357	[00:14<01:22,	3.68it/s]
15%		54/357	[00:14<01:22,	3.69it/s]
15%		55/357	[00:14<01:21,	3.68it/s]
16%		56/357	[00:15<01:21,	3.68it/s]
16%		57/357	[00:15<01:21,	3.69it/s]
16%		58/357	[00:15<01:21,	3.68it/s]
17%		59/357	[00:16<01:20,	3.69it/s]
17%		60/357	[00:16<01:20,	3.69it/s]
17%		61/357	[00:16<01:20,	3.68it/s]
17%		62/357	[00:16<01:20,	3.69it/s]
18%		63/357	[00:17<01:19,	3.69it/s]
18%		64/357	[00:17<01:19,	3.69it/s]
18%		65/357	[00:17<01:19,	3.68it/s]
18%		66/357	[00:17<01:18,	3.69it/s]
19%		67/357	[00:18<01:18,	3.68it/s]
19%		68/357	[00:18<01:18,	3.68it/s]
19%		69/357	[00:18<01:18,	3.68it/s]
20%		70/357	[00:19<01:17,	3.68it/s]
20%		71/357	[00:19<01:17,	3.68it/s]
20%		72/357	[00:19<01:17,	3.68it/s]
20%		73/357	[00:19<01:17,	3.68it/s]
21%		74/357	[00:20<01:16,	3.68it/s]
21%		75/357	[00:20<01:16,	3.68it/s]
21%		76/357	[00:20<01:16,	3.68it/s]
22%		77/357	[00:20<01:16,	3.68it/s]
22%		78/357	[00:21<01:15,	3.68it/s]
22%		79/357	[00:21<01:15,	3.68it/s]
22%		80/357	[00:21<01:15,	3.68it/s]
23%		81/357	[00:22<01:14,	3.68it/s]
23%		82/357	[00:22<01:14,	3.69it/s]
23%		83/357	[00:22<01:14,	3.69it/s]
24%		84/357	[00:22<01:14,	3.68it/s]
24%		85/357	[00:23<01:13,	3.68it/s]
24%		86/357	[00:23<01:13,	3.69it/s]
24%		87/357	[00:23<01:13,	3.69it/s]
25%		88/357	[00:23<01:12,	3.69it/s]
25%		89/357	[00:24<01:12,	3.69it/s]
25%		90/357	[00:24<01:12,	3.69it/s]
25%		91/357	[00:24<01:12,	3.69it/s]
26%		92/357	[00:24<01:11,	3.68it/s]
26%		93/357	[00:25<01:11,	3.68it/s]
26%		94/357	[00:25<01:11,	3.68it/s]
27%		95/357	[00:25<01:11,	3.68it/s]
27%		96/357	[00:26<01:10,	3.68it/s]
27%		97/357	[00:26<01:10,	3.68it/s]
27%		98/357	[00:26<01:10,	3.68it/s]

28%	99/357	[00:26<01:10,	3.68it/s]
28%	100/357	[00:27<01:09,	3.68it/s]
28%	101/357	[00:27<01:09,	3.68it/s]
29%	102/357	[00:27<01:09,	3.68it/s]
29%	103/357	[00:27<01:09,	3.68it/s]
29%	104/357	[00:28<01:08,	3.68it/s]
29%	105/357	[00:28<01:08,	3.68it/s]
30%	106/357	[00:28<01:08,	3.68it/s]
30%	107/357	[00:29<01:07,	3.68it/s]
30%	108/357	[00:29<01:07,	3.68it/s]
31%	109/357	[00:29<01:07,	3.68it/s]
31%	110/357	[00:29<01:07,	3.68it/s]
31%	111/357	[00:30<01:07,	3.67it/s]
31%	112/357	[00:30<01:08,	3.59it/s]
32%	113/357	[00:30<01:07,	3.62it/s]
32%	114/357	[00:30<01:06,	3.64it/s]
32%	115/357	[00:31<01:06,	3.65it/s]
32%	116/357	[00:31<01:05,	3.66it/s]
33%	117/357	[00:31<01:05,	3.65it/s]
33%	118/357	[00:32<01:05,	3.64it/s]
33%	119/357	[00:32<01:05,	3.65it/s]
34%	120/357	[00:32<01:04,	3.66it/s]
34%	121/357	[00:32<01:04,	3.67it/s]
34%	122/357	[00:33<01:04,	3.67it/s]
34%	123/357	[00:33<01:03,	3.68it/s]
35%	124/357	[00:33<01:03,	3.68it/s]
35%	125/357	[00:33<01:03,	3.68it/s]
35%	126/357	[00:34<01:02,	3.68it/s]
36%	127/357	[00:34<01:02,	3.68it/s]
36%	128/357	[00:34<01:02,	3.68it/s]
36%	129/357	[00:35<01:02,	3.68it/s]
36%	130/357	[00:35<01:01,	3.67it/s]
37%	131/357	[00:35<01:01,	3.67it/s]
37%	132/357	[00:35<01:01,	3.68it/s]
37%	133/357	[00:36<01:00,	3.68it/s]
38%	134/357	[00:36<01:00,	3.68it/s]
38%	135/357	[00:36<01:00,	3.68it/s]
38%	136/357	[00:36<01:00,	3.68it/s]
38%	137/357	[00:37<00:59,	3.68it/s]
39%	138/357	[00:37<00:59,	3.68it/s]
39%	139/357	[00:37<00:59,	3.68it/s]
39%	140/357	[00:38<00:58,	3.68it/s]
39%	141/357	[00:38<00:58,	3.68it/s]
40%	142/357	[00:38<00:58,	3.68it/s]
40%	143/357	[00:38<00:58,	3.68it/s]
40%	144/357	[00:39<00:57,	3.68it/s]
41%	145/357	[00:39<00:57,	3.68it/s]
41%	146/357	[00:39<00:58,	3.59it/s]

41%	147/357	[00:39<00:58,	3.62it/s]
41%	148/357	[00:40<00:57,	3.64it/s]
42%	149/357	[00:40<00:57,	3.65it/s]
42%	150/357	[00:40<00:56,	3.66it/s]
42%	151/357	[00:41<00:56,	3.66it/s]
43%	152/357	[00:41<00:55,	3.67it/s]
43%	153/357	[00:41<00:55,	3.67it/s]
43%	154/357	[00:41<00:55,	3.68it/s]
43%	155/357	[00:42<00:54,	3.68it/s]
44%	156/357	[00:42<00:54,	3.68it/s]
44%	157/357	[00:42<00:54,	3.68it/s]
44%	158/357	[00:42<00:54,	3.68it/s]
45%	159/357	[00:43<00:53,	3.68it/s]
45%	160/357	[00:43<00:53,	3.68it/s]
45%	161/357	[00:43<00:53,	3.68it/s]
45%	162/357	[00:44<00:53,	3.68it/s]
46%	163/357	[00:44<00:52,	3.68it/s]
46%	164/357	[00:44<00:52,	3.68it/s]
46%	165/357	[00:44<00:52,	3.68it/s]
46%	166/357	[00:45<00:51,	3.68it/s]
47%	167/357	[00:45<00:51,	3.68it/s]
47%	168/357	[00:45<00:51,	3.68it/s]
47%	169/357	[00:45<00:51,	3.68it/s]
48%	170/357	[00:46<00:50,	3.68it/s]
48%	171/357	[00:46<00:52,	3.51it/s]
48%	172/357	[00:46<00:57,	3.21it/s]
48%	173/357	[00:47<00:57,	3.20it/s]
49%	174/357	[00:47<00:57,	3.21it/s]
49%	175/357	[00:47<00:56,	3.22it/s]
49%	176/357	[00:48<00:56,	3.22it/s]
50%	177/357	[00:48<00:56,	3.21it/s]
50%	178/357	[00:48<00:55,	3.21it/s]
50%	179/357	[00:49<00:55,	3.21it/s]
50%	180/357	[00:49<00:55,	3.21it/s]
51%	181/357	[00:49<00:54,	3.23it/s]
51%	182/357	[00:50<00:54,	3.22it/s]
51%	183/357	[00:50<00:54,	3.21it/s]
52%	184/357	[00:50<00:54,	3.19it/s]
52%	185/357	[00:50<00:54,	3.16it/s]
52%	186/357	[00:51<00:55,	3.10it/s]
52%	187/357	[00:51<00:52,	3.23it/s]
53%	188/357	[00:51<00:50,	3.34it/s]
53%	189/357	[00:52<00:48,	3.44it/s]
53%	190/357	[00:52<00:47,	3.51it/s]
54%	191/357	[00:52<00:46,	3.56it/s]
54%	192/357	[00:52<00:45,	3.60it/s]
54%	193/357	[00:53<00:45,	3.62it/s]
54%	194/357	[00:53<00:45,	3.62it/s]

55%	195/357	[00:53<00:45,	3.58it/s]
55%	196/357	[00:54<00:45,	3.56it/s]
55%	197/357	[00:54<00:45,	3.55it/s]
55%	198/357	[00:54<00:44,	3.55it/s]
56%	199/357	[00:54<00:44,	3.59it/s]
56%	200/357	[00:55<00:43,	3.62it/s]
56%	201/357	[00:55<00:42,	3.64it/s]
57%	202/357	[00:55<00:42,	3.65it/s]
57%	203/357	[00:56<00:42,	3.66it/s]
57%	204/357	[00:56<00:41,	3.67it/s]
57%	205/357	[00:56<00:41,	3.68it/s]
58%	206/357	[00:56<00:41,	3.68it/s]
58%	207/357	[00:57<00:40,	3.68it/s]
58%	208/357	[00:57<00:40,	3.68it/s]
59%	209/357	[00:57<00:40,	3.68it/s]
59%	210/357	[00:57<00:39,	3.68it/s]
59%	211/357	[00:58<00:39,	3.68it/s]
59%	212/357	[00:58<00:39,	3.68it/s]
60%	213/357	[00:58<00:39,	3.68it/s]
60%	214/357	[00:59<00:38,	3.68it/s]
60%	215/357	[00:59<00:38,	3.68it/s]
61%	216/357	[00:59<00:38,	3.68it/s]
61%	217/357	[00:59<00:38,	3.68it/s]
61%	218/357	[01:00<00:37,	3.68it/s]
61%	219/357	[01:00<00:37,	3.68it/s]
62%	220/357	[01:00<00:37,	3.68it/s]
62%	221/357	[01:00<00:36,	3.68it/s]
62%	222/357	[01:01<00:36,	3.68it/s]
62%	223/357	[01:01<00:36,	3.68it/s]
63%	224/357	[01:01<00:36,	3.63it/s]
63%	225/357	[01:02<00:36,	3.64it/s]
63%	226/357	[01:02<00:35,	3.65it/s]
64%	227/357	[01:02<00:38,	3.38it/s]
64%	228/357	[01:02<00:37,	3.46it/s]
64%	229/357	[01:03<00:36,	3.53it/s]
64%	230/357	[01:03<00:35,	3.57it/s]
65%	231/357	[01:03<00:34,	3.60it/s]
65%	232/357	[01:03<00:34,	3.63it/s]
65%	233/357	[01:04<00:34,	3.64it/s]
66%	234/357	[01:04<00:33,	3.64it/s]
66%	235/357	[01:05<01:03,	1.92it/s]
66%	236/357	[01:05<00:54,	2.24it/s]
66%	237/357	[01:06<00:47,	2.54it/s]
67%	238/357	[01:06<00:42,	2.79it/s]
67%	239/357	[01:06<00:39,	2.97it/s]
67%	240/357	[01:07<00:37,	3.13it/s]
68%	241/357	[01:07<00:35,	3.28it/s]
68%	242/357	[01:07<00:34,	3.38it/s]



68%		243/357	[01:07<00:32,	3.47it/s]
68%		244/357	[01:08<00:32,	3.53it/s]
69%		245/357	[01:08<00:31,	3.58it/s]
69%		246/357	[01:08<00:30,	3.61it/s]
69%		247/357	[01:08<00:30,	3.63it/s]
69%		248/357	[01:09<00:29,	3.64it/s]
70%		249/357	[01:09<00:29,	3.66it/s]
70%		250/357	[01:09<00:30,	3.56it/s]
70%		251/357	[01:10<00:29,	3.58it/s]
71%		252/357	[01:10<00:29,	3.60it/s]
71%		253/357	[01:10<00:28,	3.62it/s]
71%		254/357	[01:10<00:28,	3.62it/s]
71%		255/357	[01:11<00:28,	3.58it/s]
72%		256/357	[01:11<00:28,	3.59it/s]
72%		257/357	[01:11<00:28,	3.56it/s]
72%		258/357	[01:11<00:27,	3.60it/s]
73%		259/357	[01:12<00:27,	3.60it/s]
73%		260/357	[01:12<00:27,	3.58it/s]
73%		261/357	[01:12<00:26,	3.59it/s]
73%		262/357	[01:13<00:26,	3.62it/s]
74%		263/357	[01:13<00:26,	3.53it/s]
74%		264/357	[01:13<00:26,	3.54it/s]
74%		265/357	[01:13<00:25,	3.56it/s]
75%		266/357	[01:14<00:25,	3.56it/s]
75%		267/357	[01:14<00:25,	3.56it/s]
75%		268/357	[01:14<00:26,	3.31it/s]
75%		269/357	[01:15<00:25,	3.40it/s]
76%		270/357	[01:15<00:25,	3.46it/s]
76%		271/357	[01:15<00:24,	3.50it/s]
76%		272/357	[01:15<00:24,	3.52it/s]
76%		273/357	[01:16<00:23,	3.55it/s]
77%		274/357	[01:16<00:23,	3.56it/s]
77%		275/357	[01:16<00:23,	3.55it/s]
77%		276/357	[01:17<00:23,	3.50it/s]
78%		277/357	[01:17<00:22,	3.49it/s]
78%		278/357	[01:17<00:22,	3.48it/s]
78%		279/357	[01:17<00:22,	3.47it/s]
78%		280/357	[01:18<00:21,	3.52it/s]
79%		281/357	[01:18<00:21,	3.52it/s]
79%		282/357	[01:18<00:21,	3.54it/s]
79%		283/357	[01:19<00:21,	3.50it/s]
80%		284/357	[01:19<00:20,	3.50it/s]
80%		285/357	[01:19<00:20,	3.52it/s]
80%		286/357	[01:19<00:20,	3.52it/s]
80%		287/357	[01:20<00:19,	3.54it/s]
81%		288/357	[01:20<00:19,	3.47it/s]
81%		289/357	[01:20<00:19,	3.49it/s]
81%		290/357	[01:21<00:19,	3.51it/s]

82%		291/357	[01:21<00:18,	3.51it/s]
82%		292/357	[01:21<00:18,	3.48it/s]
82%		293/357	[01:21<00:18,	3.43it/s]
82%		294/357	[01:22<00:18,	3.40it/s]
83%		295/357	[01:22<00:18,	3.39it/s]
83%		296/357	[01:22<00:17,	3.42it/s]
83%		297/357	[01:23<00:17,	3.46it/s]
83%		298/357	[01:23<00:16,	3.48it/s]
84%		299/357	[01:23<00:16,	3.48it/s]
84%		300/357	[01:23<00:16,	3.47it/s]
84%		301/357	[01:24<00:16,	3.46it/s]
85%		302/357	[01:24<00:16,	3.43it/s]
85%		303/357	[01:24<00:15,	3.46it/s]
85%		304/357	[01:25<00:15,	3.47it/s]
85%		305/357	[01:25<00:15,	3.46it/s]
86%		306/357	[01:25<00:14,	3.45it/s]
86%		307/357	[01:26<00:14,	3.45it/s]
86%		308/357	[01:26<00:14,	3.39it/s]
87%		309/357	[01:26<00:13,	3.47it/s]
87%		310/357	[01:26<00:13,	3.49it/s]
87%		311/357	[01:27<00:13,	3.35it/s]
87%		312/357	[01:27<00:13,	3.23it/s]
88%		313/357	[01:27<00:13,	3.29it/s]
88%		314/357	[01:28<00:12,	3.35it/s]
88%		315/357	[01:28<00:12,	3.33it/s]
89%		316/357	[01:28<00:12,	3.37it/s]
89%		317/357	[01:29<00:11,	3.36it/s]
89%		318/357	[01:29<00:11,	3.38it/s]
89%		319/357	[01:29<00:11,	3.37it/s]
90%		320/357	[01:29<00:10,	3.37it/s]
90%		321/357	[01:30<00:10,	3.39it/s]
90%		322/357	[01:30<00:10,	3.37it/s]
90%		323/357	[01:30<00:10,	3.35it/s]
91%		324/357	[01:31<00:09,	3.35it/s]
91%		325/357	[01:31<00:09,	3.34it/s]
91%		326/357	[01:31<00:09,	3.37it/s]
92%		327/357	[01:31<00:08,	3.45it/s]
92%		328/357	[01:32<00:08,	3.41it/s]
92%		329/357	[01:32<00:08,	3.40it/s]
92%		330/357	[01:32<00:08,	3.37it/s]
93%		331/357	[01:33<00:07,	3.36it/s]
93%		332/357	[01:33<00:07,	3.37it/s]
93%		333/357	[01:33<00:07,	3.38it/s]
94%		334/357	[01:34<00:07,	3.26it/s]
94%		335/357	[01:34<00:06,	3.27it/s]
94%		336/357	[01:34<00:06,	3.29it/s]
94%		337/357	[01:34<00:06,	3.31it/s]
95%		338/357	[01:35<00:05,	3.31it/s]

```

95%|      | 339/357 [01:35<00:05,  3.29it/s]
95%|      | 340/357 [01:35<00:05,  3.31it/s]
96%|      | 341/357 [01:36<00:04,  3.33it/s]
96%|      | 342/357 [01:36<00:04,  3.29it/s]
96%|      | 343/357 [01:36<00:04,  3.27it/s]
96%|      | 344/357 [01:37<00:03,  3.25it/s]
97%|      | 345/357 [01:37<00:03,  3.18it/s]
97%|      | 346/357 [01:37<00:03,  3.23it/s]
97%|      | 347/357 [01:38<00:03,  3.23it/s]
97%|      | 348/357 [01:38<00:02,  3.24it/s]
98%|      | 349/357 [01:38<00:02,  3.25it/s]
98%|      | 350/357 [01:38<00:02,  3.25it/s]
98%|      | 351/357 [01:39<00:01,  3.26it/s]
99%|      | 352/357 [01:39<00:01,  3.24it/s]
99%|      | 353/357 [01:39<00:01,  3.25it/s]
99%|      | 354/357 [01:40<00:00,  3.26it/s]
99%|      | 355/357 [01:40<00:00,  3.27it/s]
100%|     | 356/357 [01:40<00:00,  3.27it/s]
100%|     | 357/357 [01:41<00:00,  3.53it/s]
90%|      | 9/10 [15:29<01:42, 102.16s/it]

```

For - 8 Loss: tensor(0.0002, device='cuda:0', grad\_fn=<MeanBackward0>)

```

0%|      | 0/357 [00:00<?, ?it/s]
0%|      | 1/357 [00:00<01:45,  3.37it/s]
1%|      | 2/357 [00:00<01:39,  3.55it/s]
1%|      | 3/357 [00:00<01:37,  3.61it/s]
1%|      | 4/357 [00:01<01:37,  3.64it/s]
1%|      | 5/357 [00:01<01:36,  3.66it/s]
2%|      | 6/357 [00:01<01:35,  3.67it/s]
2%|      | 7/357 [00:01<01:35,  3.68it/s]
2%|      | 8/357 [00:02<01:34,  3.68it/s]
3%|      | 9/357 [00:02<01:34,  3.68it/s]
3%|      | 10/357 [00:02<01:34,  3.68it/s]
3%|      | 11/357 [00:03<01:33,  3.69it/s]
3%|      | 12/357 [00:03<01:33,  3.69it/s]
4%|      | 13/357 [00:03<01:33,  3.68it/s]
4%|      | 14/357 [00:03<01:33,  3.69it/s]
4%|      | 15/357 [00:04<01:32,  3.69it/s]
4%|      | 16/357 [00:04<01:32,  3.69it/s]
5%|      | 17/357 [00:04<01:32,  3.69it/s]
5%|      | 18/357 [00:04<01:31,  3.69it/s]
5%|      | 19/357 [00:05<01:31,  3.69it/s]
6%|      | 20/357 [00:05<01:31,  3.68it/s]
6%|      | 21/357 [00:05<01:31,  3.68it/s]
6%|      | 22/357 [00:05<01:30,  3.68it/s]
6%|      | 23/357 [00:06<01:30,  3.68it/s]
7%|      | 24/357 [00:06<01:30,  3.68it/s]

```

7%	25/357	[00:06<01:30,	3.68it/s]
7%	26/357	[00:07<01:29,	3.68it/s]
8%	27/357	[00:07<01:29,	3.68it/s]
8%	28/357	[00:07<01:29,	3.69it/s]
8%	29/357	[00:07<01:28,	3.69it/s]
8%	30/357	[00:08<01:28,	3.69it/s]
9%	31/357	[00:08<01:28,	3.68it/s]
9%	32/357	[00:08<01:28,	3.68it/s]
9%	33/357	[00:08<01:27,	3.68it/s]
10%	34/357	[00:09<01:27,	3.68it/s]
10%	35/357	[00:09<01:27,	3.68it/s]
10%	36/357	[00:09<01:27,	3.68it/s]
10%	37/357	[00:10<01:26,	3.69it/s]
11%	38/357	[00:10<01:26,	3.68it/s]
11%	39/357	[00:10<01:26,	3.69it/s]
11%	40/357	[00:10<01:25,	3.69it/s]
11%	41/357	[00:11<01:25,	3.69it/s]
12%	42/357	[00:11<01:25,	3.69it/s]
12%	43/357	[00:11<01:25,	3.69it/s]
12%	44/357	[00:11<01:24,	3.69it/s]
13%	45/357	[00:12<01:24,	3.69it/s]
13%	46/357	[00:12<01:24,	3.69it/s]
13%	47/357	[00:12<01:24,	3.68it/s]
13%	48/357	[00:13<01:23,	3.69it/s]
14%	49/357	[00:13<01:23,	3.68it/s]
14%	50/357	[00:13<01:23,	3.69it/s]
14%	51/357	[00:13<01:22,	3.69it/s]
15%	52/357	[00:14<01:22,	3.69it/s]
15%	53/357	[00:14<01:22,	3.68it/s]
15%	54/357	[00:14<01:22,	3.68it/s]
15%	55/357	[00:14<01:21,	3.69it/s]
16%	56/357	[00:15<01:21,	3.68it/s]
16%	57/357	[00:15<01:21,	3.68it/s]
16%	58/357	[00:15<01:21,	3.68it/s]
17%	59/357	[00:16<01:20,	3.68it/s]
17%	60/357	[00:16<01:20,	3.68it/s]
17%	61/357	[00:16<01:20,	3.68it/s]
17%	62/357	[00:16<01:20,	3.68it/s]
18%	63/357	[00:17<01:19,	3.68it/s]
18%	64/357	[00:17<01:19,	3.68it/s]
18%	65/357	[00:17<01:19,	3.68it/s]
18%	66/357	[00:17<01:19,	3.68it/s]
19%	67/357	[00:18<01:18,	3.68it/s]
19%	68/357	[00:18<01:18,	3.68it/s]
19%	69/357	[00:18<01:18,	3.68it/s]
20%	70/357	[00:19<01:17,	3.68it/s]
20%	71/357	[00:19<01:17,	3.68it/s]
20%	72/357	[00:19<01:17,	3.68it/s]

20%	73/357	[00:19<01:17,	3.68it/s]
21%	74/357	[00:20<01:16,	3.68it/s]
21%	75/357	[00:20<01:16,	3.68it/s]
21%	76/357	[00:20<01:16,	3.68it/s]
22%	77/357	[00:20<01:15,	3.69it/s]
22%	78/357	[00:21<01:15,	3.68it/s]
22%	79/357	[00:21<01:15,	3.69it/s]
22%	80/357	[00:21<01:15,	3.68it/s]
23%	81/357	[00:22<01:14,	3.68it/s]
23%	82/357	[00:22<01:14,	3.68it/s]
23%	83/357	[00:22<01:14,	3.68it/s]
24%	84/357	[00:22<01:14,	3.68it/s]
24%	85/357	[00:23<01:13,	3.68it/s]
24%	86/357	[00:23<01:13,	3.68it/s]
24%	87/357	[00:23<01:13,	3.68it/s]
25%	88/357	[00:23<01:12,	3.69it/s]
25%	89/357	[00:24<01:12,	3.68it/s]
25%	90/357	[00:24<01:12,	3.68it/s]
25%	91/357	[00:24<01:12,	3.68it/s]
26%	92/357	[00:24<01:11,	3.69it/s]
26%	93/357	[00:25<01:11,	3.68it/s]
26%	94/357	[00:26<02:09,	2.03it/s]
27%	95/357	[00:26<01:51,	2.35it/s]
27%	96/357	[00:26<01:39,	2.64it/s]
27%	97/357	[00:27<01:30,	2.88it/s]
27%	98/357	[00:27<01:24,	3.08it/s]
28%	99/357	[00:27<01:19,	3.24it/s]
28%	100/357	[00:27<01:16,	3.36it/s]
28%	101/357	[00:28<01:14,	3.45it/s]
29%	102/357	[00:28<01:12,	3.52it/s]
29%	103/357	[00:28<01:11,	3.57it/s]
29%	104/357	[00:28<01:10,	3.60it/s]
29%	105/357	[00:29<01:09,	3.62it/s]
30%	106/357	[00:29<01:08,	3.64it/s]
30%	107/357	[00:29<01:08,	3.65it/s]
30%	108/357	[00:30<01:07,	3.66it/s]
31%	109/357	[00:30<01:07,	3.67it/s]
31%	110/357	[00:30<01:07,	3.68it/s]
31%	111/357	[00:30<01:06,	3.68it/s]
31%	112/357	[00:31<01:08,	3.60it/s]
32%	113/357	[00:31<01:07,	3.62it/s]
32%	114/357	[00:31<01:06,	3.64it/s]
32%	115/357	[00:31<01:06,	3.66it/s]
32%	116/357	[00:32<01:05,	3.66it/s]
33%	117/357	[00:32<01:05,	3.67it/s]
33%	118/357	[00:32<01:05,	3.67it/s]
33%	119/357	[00:33<01:04,	3.68it/s]
34%	120/357	[00:33<01:04,	3.68it/s]

34%	121/357	[00:33<01:04,	3.68it/s]
34%	122/357	[00:33<01:03,	3.68it/s]
34%	123/357	[00:34<01:03,	3.68it/s]
35%	124/357	[00:34<01:03,	3.68it/s]
35%	125/357	[00:34<01:03,	3.68it/s]
35%	126/357	[00:34<01:02,	3.69it/s]
36%	127/357	[00:35<01:02,	3.69it/s]
36%	128/357	[00:35<01:02,	3.69it/s]
36%	129/357	[00:35<01:01,	3.68it/s]
36%	130/357	[00:36<01:01,	3.68it/s]
37%	131/357	[00:36<01:01,	3.68it/s]
37%	132/357	[00:36<01:01,	3.68it/s]
37%	133/357	[00:36<01:00,	3.68it/s]
38%	134/357	[00:37<01:00,	3.68it/s]
38%	135/357	[00:37<01:00,	3.69it/s]
38%	136/357	[00:37<00:59,	3.69it/s]
38%	137/357	[00:37<00:59,	3.68it/s]
39%	138/357	[00:38<00:59,	3.69it/s]
39%	139/357	[00:38<00:59,	3.69it/s]
39%	140/357	[00:38<00:58,	3.68it/s]
39%	141/357	[00:39<00:58,	3.69it/s]
40%	142/357	[00:39<00:58,	3.69it/s]
40%	143/357	[00:39<00:58,	3.68it/s]
40%	144/357	[00:39<00:57,	3.69it/s]
41%	145/357	[00:40<00:57,	3.69it/s]
41%	146/357	[00:40<00:57,	3.69it/s]
41%	147/357	[00:40<00:58,	3.62it/s]
41%	148/357	[00:40<00:57,	3.63it/s]
42%	149/357	[00:41<00:56,	3.65it/s]
42%	150/357	[00:41<00:56,	3.66it/s]
42%	151/357	[00:41<00:56,	3.66it/s]
43%	152/357	[00:42<00:55,	3.67it/s]
43%	153/357	[00:42<00:55,	3.67it/s]
43%	154/357	[00:42<00:55,	3.68it/s]
43%	155/357	[00:42<00:54,	3.67it/s]
44%	156/357	[00:43<00:54,	3.68it/s]
44%	157/357	[00:43<00:54,	3.68it/s]
44%	158/357	[00:43<00:54,	3.68it/s]
45%	159/357	[00:43<00:53,	3.68it/s]
45%	160/357	[00:44<00:53,	3.68it/s]
45%	161/357	[00:44<00:53,	3.68it/s]
45%	162/357	[00:44<00:52,	3.68it/s]
46%	163/357	[00:45<00:52,	3.68it/s]
46%	164/357	[00:45<00:52,	3.68it/s]
46%	165/357	[00:45<00:52,	3.68it/s]
46%	166/357	[00:45<00:51,	3.68it/s]
47%	167/357	[00:46<00:51,	3.68it/s]
47%	168/357	[00:46<00:51,	3.68it/s]

47%	169/357	[00:46<00:51,	3.68it/s]
48%	170/357	[00:46<00:50,	3.68it/s]
48%	171/357	[00:47<00:50,	3.68it/s]
48%	172/357	[00:47<00:50,	3.68it/s]
48%	173/357	[00:47<00:50,	3.68it/s]
49%	174/357	[00:48<00:49,	3.68it/s]
49%	175/357	[00:48<00:49,	3.68it/s]
49%	176/357	[00:48<00:49,	3.68it/s]
50%	177/357	[00:48<00:49,	3.67it/s]
50%	178/357	[00:49<00:49,	3.64it/s]
50%	179/357	[00:49<00:48,	3.65it/s]
50%	180/357	[00:49<00:48,	3.66it/s]
51%	181/357	[00:49<00:48,	3.67it/s]
51%	182/357	[00:50<00:47,	3.67it/s]
51%	183/357	[00:50<00:47,	3.67it/s]
52%	184/357	[00:50<00:47,	3.61it/s]
52%	185/357	[00:51<00:47,	3.61it/s]
52%	186/357	[00:51<00:47,	3.63it/s]
52%	187/357	[00:51<00:46,	3.65it/s]
53%	188/357	[00:51<00:46,	3.66it/s]
53%	189/357	[00:52<00:45,	3.67it/s]
53%	190/357	[00:52<00:45,	3.67it/s]
54%	191/357	[00:52<00:47,	3.49it/s]
54%	192/357	[00:53<00:46,	3.54it/s]
54%	193/357	[00:53<00:45,	3.58it/s]
54%	194/357	[00:53<00:45,	3.61it/s]
55%	195/357	[00:53<00:44,	3.63it/s]
55%	196/357	[00:54<00:44,	3.64it/s]
55%	197/357	[00:54<00:43,	3.65it/s]
55%	198/357	[00:54<00:43,	3.67it/s]
56%	199/357	[00:54<00:43,	3.67it/s]
56%	200/357	[00:55<00:42,	3.67it/s]
56%	201/357	[00:55<00:42,	3.68it/s]
57%	202/357	[00:55<00:42,	3.68it/s]
57%	203/357	[00:56<00:42,	3.61it/s]
57%	204/357	[00:56<00:42,	3.63it/s]
57%	205/357	[00:56<00:41,	3.64it/s]
58%	206/357	[00:56<00:41,	3.65it/s]
58%	207/357	[00:57<00:40,	3.67it/s]
58%	208/357	[00:57<00:40,	3.67it/s]
59%	209/357	[00:57<00:40,	3.68it/s]
59%	210/357	[00:57<00:39,	3.68it/s]
59%	211/357	[00:58<00:39,	3.68it/s]
59%	212/357	[00:58<00:39,	3.68it/s]
60%	213/357	[00:58<00:39,	3.68it/s]
60%	214/357	[00:58<00:38,	3.68it/s]
60%	215/357	[00:59<00:38,	3.68it/s]
61%	216/357	[00:59<00:38,	3.69it/s]

61%	217/357	[00:59<00:38,	3.68it/s]
61%	218/357	[01:00<00:37,	3.68it/s]
61%	219/357	[01:00<00:37,	3.68it/s]
62%	220/357	[01:00<00:37,	3.68it/s]
62%	221/357	[01:00<00:36,	3.68it/s]
62%	222/357	[01:01<00:36,	3.68it/s]
62%	223/357	[01:01<00:36,	3.68it/s]
63%	224/357	[01:01<00:36,	3.68it/s]
63%	225/357	[01:01<00:35,	3.68it/s]
63%	226/357	[01:02<00:35,	3.68it/s]
64%	227/357	[01:02<00:35,	3.67it/s]
64%	228/357	[01:02<00:35,	3.67it/s]
64%	229/357	[01:03<00:34,	3.67it/s]
64%	230/357	[01:03<00:34,	3.67it/s]
65%	231/357	[01:03<00:34,	3.67it/s]
65%	232/357	[01:03<00:34,	3.67it/s]
65%	233/357	[01:04<00:33,	3.67it/s]
66%	234/357	[01:04<00:33,	3.68it/s]
66%	235/357	[01:04<00:33,	3.68it/s]
66%	236/357	[01:04<00:32,	3.68it/s]
66%	237/357	[01:05<00:32,	3.68it/s]
67%	238/357	[01:05<00:32,	3.68it/s]
67%	239/357	[01:06<00:54,	2.18it/s]
67%	240/357	[01:06<00:47,	2.47it/s]
68%	241/357	[01:06<00:42,	2.74it/s]
68%	242/357	[01:07<00:38,	2.97it/s]
68%	243/357	[01:07<00:36,	3.15it/s]
68%	244/357	[01:07<00:34,	3.24it/s]
69%	245/357	[01:08<00:33,	3.36it/s]
69%	246/357	[01:08<00:32,	3.45it/s]
69%	247/357	[01:08<00:31,	3.51it/s]
69%	248/357	[01:08<00:30,	3.56it/s]
70%	249/357	[01:09<00:30,	3.59it/s]
70%	250/357	[01:09<00:29,	3.61it/s]
70%	251/357	[01:09<00:29,	3.63it/s]
71%	252/357	[01:09<00:28,	3.65it/s]
71%	253/357	[01:10<00:28,	3.64it/s]
71%	254/357	[01:10<00:28,	3.64it/s]
71%	255/357	[01:10<00:28,	3.59it/s]
72%	256/357	[01:11<00:27,	3.62it/s]
72%	257/357	[01:11<00:27,	3.61it/s]
72%	258/357	[01:11<00:27,	3.63it/s]
73%	259/357	[01:11<00:26,	3.65it/s]
73%	260/357	[01:12<00:26,	3.62it/s]
73%	261/357	[01:12<00:26,	3.63it/s]
73%	262/357	[01:12<00:26,	3.65it/s]
74%	263/357	[01:13<00:25,	3.65it/s]
74%	264/357	[01:13<00:25,	3.64it/s]



74%	265/357	[01:13<00:25,	3.65it/s]
75%	266/357	[01:13<00:25,	3.64it/s]
75%	267/357	[01:14<00:24,	3.60it/s]
75%	268/357	[01:14<00:24,	3.62it/s]
75%	269/357	[01:14<00:24,	3.59it/s]
76%	270/357	[01:14<00:24,	3.59it/s]
76%	271/357	[01:15<00:24,	3.58it/s]
76%	272/357	[01:15<00:23,	3.55it/s]
76%	273/357	[01:15<00:23,	3.50it/s]
77%	274/357	[01:16<00:23,	3.49it/s]
77%	275/357	[01:16<00:23,	3.54it/s]
77%	276/357	[01:16<00:22,	3.55it/s]
78%	277/357	[01:16<00:22,	3.53it/s]
78%	278/357	[01:17<00:22,	3.50it/s]
78%	279/357	[01:17<00:22,	3.51it/s]
78%	280/357	[01:17<00:22,	3.43it/s]
79%	281/357	[01:18<00:21,	3.47it/s]
79%	282/357	[01:18<00:22,	3.32it/s]
79%	283/357	[01:18<00:21,	3.38it/s]
80%	284/357	[01:19<00:21,	3.42it/s]
80%	285/357	[01:19<00:20,	3.44it/s]
80%	286/357	[01:19<00:20,	3.49it/s]
80%	287/357	[01:19<00:19,	3.50it/s]
81%	288/357	[01:20<00:19,	3.51it/s]
81%	289/357	[01:20<00:19,	3.48it/s]
81%	290/357	[01:20<00:19,	3.52it/s]
82%	291/357	[01:20<00:18,	3.50it/s]
82%	292/357	[01:21<00:19,	3.38it/s]
82%	293/357	[01:21<00:18,	3.43it/s]
82%	294/357	[01:21<00:18,	3.41it/s]
83%	295/357	[01:22<00:18,	3.38it/s]
83%	296/357	[01:22<00:17,	3.40it/s]
83%	297/357	[01:22<00:17,	3.39it/s]
83%	298/357	[01:23<00:17,	3.42it/s]
84%	299/357	[01:23<00:17,	3.40it/s]
84%	300/357	[01:23<00:16,	3.42it/s]
84%	301/357	[01:23<00:16,	3.42it/s]
85%	302/357	[01:24<00:16,	3.43it/s]
85%	303/357	[01:24<00:16,	3.37it/s]
85%	304/357	[01:24<00:15,	3.41it/s]
85%	305/357	[01:25<00:16,	3.25it/s]
86%	306/357	[01:25<00:15,	3.27it/s]
86%	307/357	[01:25<00:15,	3.33it/s]
86%	308/357	[01:26<00:14,	3.35it/s]
87%	309/357	[01:26<00:14,	3.39it/s]
87%	310/357	[01:26<00:13,	3.40it/s]
87%	311/357	[01:26<00:13,	3.42it/s]
87%	312/357	[01:27<00:13,	3.40it/s]

88%		313/357	[01:27<00:12,	3.39it/s]
88%		314/357	[01:27<00:12,	3.35it/s]
88%		315/357	[01:28<00:12,	3.32it/s]
89%		316/357	[01:28<00:12,	3.34it/s]
89%		317/357	[01:28<00:12,	3.27it/s]
89%		318/357	[01:29<00:11,	3.34it/s]
89%		319/357	[01:29<00:11,	3.34it/s]
90%		320/357	[01:29<00:11,	3.36it/s]
90%		321/357	[01:29<00:10,	3.38it/s]
90%		322/357	[01:30<00:10,	3.39it/s]
90%		323/357	[01:30<00:10,	3.38it/s]
91%		324/357	[01:30<00:09,	3.36it/s]
91%		325/357	[01:31<00:09,	3.34it/s]
91%		326/357	[01:31<00:09,	3.35it/s]
92%		327/357	[01:31<00:08,	3.36it/s]
92%		328/357	[01:32<00:08,	3.34it/s]
92%		329/357	[01:32<00:08,	3.33it/s]
92%		330/357	[01:32<00:08,	3.32it/s]
93%		331/357	[01:32<00:07,	3.30it/s]
93%		332/357	[01:33<00:07,	3.23it/s]
93%		333/357	[01:33<00:07,	3.28it/s]
94%		334/357	[01:33<00:07,	3.26it/s]
94%		335/357	[01:34<00:06,	3.31it/s]
94%		336/357	[01:34<00:06,	3.34it/s]
94%		337/357	[01:34<00:05,	3.35it/s]
95%		338/357	[01:35<00:05,	3.33it/s]
95%		339/357	[01:35<00:05,	3.32it/s]
95%		340/357	[01:35<00:05,	3.32it/s]
96%		341/357	[01:35<00:04,	3.27it/s]
96%		342/357	[01:36<00:04,	3.27it/s]
96%		343/357	[01:36<00:04,	3.24it/s]
96%		344/357	[01:36<00:04,	3.21it/s]
97%		345/357	[01:37<00:03,	3.19it/s]
97%		346/357	[01:37<00:03,	3.20it/s]
97%		347/357	[01:37<00:03,	3.26it/s]
97%		348/357	[01:38<00:02,	3.26it/s]
98%		349/357	[01:38<00:02,	3.23it/s]
98%		350/357	[01:38<00:02,	3.24it/s]
98%		351/357	[01:39<00:01,	3.22it/s]
99%		352/357	[01:39<00:01,	3.21it/s]
99%		353/357	[01:39<00:01,	3.24it/s]
99%		354/357	[01:39<00:00,	3.26it/s]
99%		355/357	[01:40<00:00,	3.27it/s]
100%		356/357	[01:40<00:00,	3.28it/s]
100%		357/357	[01:40<00:00,	3.54it/s]
100%		10/10	[17:10<00:00,	103.10s/it]

For - 9 Loss: tensor(0.0002, device='cuda:0', grad\_fn=<MeanBackward0>)

### 3.4 Downstream Tuning and Evaluation

#### Metrics Descriptions in CASS Paper

- **Metric Used:** F1 Score
- **Definition:**
- F1 Score is calculated as  $F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$ , where Precision is the ratio of true positive predictions to the total positive predictions, and Recall is the ratio of true positive predictions to the total actual positives.
- **Reason for Choice:**
- Selected based on previous work or as defined by the dataset provider.

```
[10]: """
Define F1 score metric
"""
class MyF1Score(Metric):
    def __init__(self, cfg, threshold: float = 0.5, dist_sync_on_step=False):
        super().__init__(dist_sync_on_step=dist_sync_on_step)
        self.cfg = cfg
        self.threshold = threshold
        self.add_state("tp", default=torch.tensor(0), dist_reduce_fx="sum")
        self.add_state("fp", default=torch.tensor(0), dist_reduce_fx="sum")
        self.add_state("fn", default=torch.tensor(0), dist_reduce_fx="sum")

    def update(self, preds: torch.Tensor, target: torch.Tensor):
        assert preds.shape == target.shape
        preds_str_batch = self.num_to_str(torch.sigmoid(preds))
        target_str_batch = self.num_to_str(target)
        tp, fp, fn = 0, 0, 0
        for pred_str_list, target_str_list in zip(preds_str_batch,
↪target_str_batch):
            for pred_str in pred_str_list:
                if pred_str in target_str_list:
                    tp += 1
                if pred_str not in target_str_list:
                    fp += 1

            for target_str in target_str_list:
                if target_str not in pred_str_list:
                    fn += 1

        self.tp += tp
        self.fp += fp
        self.fn += fn

    def compute(self):
```

```

        f1 = 2.0 * self.tp / (2.0 * self.tp + self.fn + self.fp)
        return f1

    def num_to_str(self, ts: torch.Tensor) -> list:
        batch_bool_list = (ts > self.threshold).detach().cpu().numpy().tolist()
        batch_str_list = []
        for one_sample_bool in batch_bool_list:
            lb_str_list = [self.cfg.label_num2str[lb_idx] for lb_idx, bool_val
in enumerate(one_sample_bool) if bool_val]
            batch_str_list.append(lb_str_list)
        return batch_str_list

```

```

[12]: if cfg.CNN:

    model_vit=torch.load('./cass-vit.pt')
    model_cnn=torch.load('./cass-r50.pt')
    last_loss=math.inf
    val_loss_arr=[]
    train_loss_arr=[]
    counter=0

    model_cnn.to(device)
    model_vit.to(device)
    print('*'*10)

    #Train Correspong Supervised CNN
    print('Fine tunning Cov-T')
    writer = SummaryWriter()
    model_cnn.fc=nn.Linear(in_features=2048, out_features=4, bias=True)
    criterion = FocalLoss(cfg.fl_alpha, cfg.fl_gamma)
    metric = MyF1Score(cfg)
    val_metric=MyF1Score(cfg)
    optimizer = torch.optim.Adam(model_cnn.parameters(), lr = 3e-4)
    scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer,T_max=cfg.
t_max,eta_min=cfg.min_lr)
    model_cnn.train()
    from torch.autograd import Variable
    best=0
    best_val=0
    for epoch in tqdm(range(200)):
        total_loss = 0
        for images, label in train_loader:
            model_cnn.train()
            images = images.to(device)
            label = label.to(device)
            model_cnn.to(device)

```

```

        pred_ts=model_cnn(images)
        label_one_hot = torch.nn.functional.one_hot(label, num_classes=4).
        ↪float()

        # print(pred_ts.shape, label.shape,label)
        loss = criterion(pred_ts, label_one_hot)
        score = metric(pred_ts,label_one_hot)
        loss.backward()
        optimizer.step()
        optimizer.zero_grad()
        scheduler.step()
        total_loss += loss.detach()
        avg_loss=total_loss/ len(train_loader)
        train_score=metric.compute()
        logs = {'train_loss': avg_loss, 'train_f1': train_score, 'lr':
        ↪optimizer.param_groups[0]['lr']}
        writer.add_scalar("CNN Supervised Loss/train", loss, epoch)
        writer.add_scalar("CNN Supervised F1/train", train_score, epoch)
        print(logs)
        if best < train_score:
            best=train_score
            model_cnn.eval()
            total_loss = 0
            for images,label in val_loader:
                images = images.to(device)
                label = label.to(device)
                model_cnn.to(device)
                pred_ts=model_cnn(images)
                label_one_hot = torch.nn.functional.one_hot(label,
        ↪num_classes=4).float()
                score_val = val_metric(pred_ts, label_one_hot)
                val_loss = criterion(pred_ts, label_one_hot)
                total_loss += val_loss.detach()
            avg_loss=total_loss/ len(train_loader)
            print('Val Loss:',avg_loss)
            val_score=val_metric.compute()
            print('CNN Validation Score:',val_score)
            writer.add_scalar("CNN Supervised F1/Validation", val_score, epoch)
            if avg_loss > last_loss:
                counter+=1
            else:
                counter=0

            last_loss = avg_loss
            if counter > 5:
                print('Early Stopping!')
                break
            else:

```

```

        if val_score > best_val:
            best_val=val_score
            print('Saving')
            torch.save(model_cnn,
                        './cass-r50-tuned.pt')

if cfg.ViT:
    print('Fine tuning Cov-T')
    writer = SummaryWriter()

    model_vit=torch.load('./cass-vit.pt')
    model_cnn=torch.load('./cass-r50.pt')
    last_loss=math.inf
    val_loss_arr=[]
    train_loss_arr=[]
    counter=0

    model_cnn.to(device)
    model_vit.to(device)
    print('*'*10)

    writer.flush()
    last_loss=999999999
    val_loss_arr=[]
    train_loss_arr=[]
    counter=0
    # Training the Corresponding ViT
    model_vit.head=nn.Linear(in_features=768, out_features=4, bias=True)
    criterion = FocalLoss(cfg.fl_alpha, cfg.fl_gamma)
    metric = MyF1Score(cfg)
    optimizer = torch.optim.Adam(model_vit.parameters(), lr = 3e-4)
    scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer,T_max=cfg.
    ↪t_max,eta_min=cfg.min_lr)
    model_vit.train()
    val_metric=MyF1Score(cfg)
    writer = SummaryWriter()
    from torch.autograd import Variable
    best=0
    best_val=0
    for epoch in tqdm(range(200)):
        total_loss = 0
        for images,label in train_loader:
            model_vit.train()
            images = images.to(device)
            label = label.to(device)

```

```

        label = torch.nn.functional.one_hot(label, num_classes=4).float().
    ↪to(device)
        model_vit.to(device)
        pred_ts=model_vit(images)
        loss = criterion(pred_ts, label)
        score = metric(pred_ts,label)
        loss.backward()
        optimizer.step()
        optimizer.zero_grad()
        scheduler.step()
        total_loss += loss.detach()
    avg_loss=total_loss/ len(train_loader)
    train_score=metric.compute()
    logs = {'train_loss': loss, 'train_f1': train_score, 'lr': optimizer.
    ↪param_groups[0]['lr']}
    writer.add_scalar("ViT Supervised Loss/train", loss, epoch)
    writer.add_scalar("ViT Supervised F1/train", train_score, epoch)
    print(logs)
    if best < train_score:
        best=train_score
        model_vit.eval()
        total_loss = 0
        for images,label in val_loader:
            images = images.to(device)
            label = label.to(device)
            label = torch.nn.functional.one_hot(label, num_classes=4).
    ↪float().to(device)
            model_vit.to(device)
            pred_ts=model_vit(images)
            score_val = val_metric(pred_ts,label)
            val_loss = criterion(pred_ts, label)
            total_loss += val_loss.detach()
        avg_loss=total_loss/ len(train_loader)
        val_score=val_metric.compute()
        print('ViT Validation Score:',val_score)
        print('Val Loss:',avg_loss)
        writer.add_scalar("ViT Supervised F1/Validation", val_score, epoch)
        if avg_loss > last_loss:
            counter+=1
        else:
            counter=0

    last_loss = avg_loss
    if counter > 5:
        print('Early Stopping!')
        break
    else:

```

```

        if val_score > best_val:
            best_val=val_score
            print('Saving')
            torch.save(model_vit,
                        './cass-vit-tuned.pt')

writer.flush()
print('*'*10)

```

\*\*\*\*\*

Fine tuning Cov-T

```

0%|          | 0/200 [00:00<?, ?it/s]

{'train_loss': tensor(0.0624, device='cuda:0'), 'train_f1': tensor(0.3112),
'lr': 0.00023355774983644742}

0%|          | 1/200 [00:39<2:12:34, 39.97s/it]

Val Loss: tensor(0.0129, device='cuda:0')
CNN Validation Score: tensor(0.6094)
Saving
{'train_loss': tensor(0.0427, device='cuda:0'), 'train_f1': tensor(0.5353),
'lr': 9.328882686143623e-05}

1%|          | 2/200 [01:18<2:08:17, 38.87s/it]

Val Loss: tensor(0.0100, device='cuda:0')
CNN Validation Score: tensor(0.6573)
Saving
{'train_loss': tensor(0.0368, device='cuda:0'), 'train_f1': tensor(0.6180),
'lr': 3.872600579717116e-06}

2%|          | 3/200 [01:56<2:06:30, 38.53s/it]

Val Loss: tensor(0.0088, device='cuda:0')
CNN Validation Score: tensor(0.6889)
Saving
{'train_loss': tensor(0.0343, device='cuda:0'), 'train_f1': tensor(0.6605),
'lr': 4.478753621262174e-05}

2%|          | 4/200 [02:34<2:05:10, 38.32s/it]

Val Loss: tensor(0.0077, device='cuda:0')
CNN Validation Score: tensor(0.7110)
Saving
{'train_loss': tensor(0.0313, device='cuda:0'), 'train_f1': tensor(0.6914),
'lr': 0.00017966600314148415}

2%|          | 5/200 [03:12<2:04:43, 38.38s/it]

Val Loss: tensor(0.0070, device='cuda:0')
CNN Validation Score: tensor(0.7302)

```



```

Saving
{'train_loss': tensor(0.0301, device='cuda:0'), 'train_f1': tensor(0.7131),
'lr': 0.00028861999011041363}

3%|          | 6/200 [03:50<2:03:55, 38.33s/it]

Val Loss: tensor(0.0077, device='cuda:0')
CNN Validation Score: tensor(0.7397)
Saving
{'train_loss': tensor(0.0302, device='cuda:0'), 'train_f1': tensor(0.7274),
'lr': 0.00027480470703923767}

4%|          | 7/200 [04:29<2:03:38, 38.44s/it]

Val Loss: tensor(0.0071, device='cuda:0')
CNN Validation Score: tensor(0.7466)
Saving
{'train_loss': tensor(0.0281, device='cuda:0'), 'train_f1': tensor(0.7404),
'lr': 0.00015050000000011167}

4%|          | 8/200 [05:07<2:02:58, 38.43s/it]

Val Loss: tensor(0.0064, device='cuda:0')
CNN Validation Score: tensor(0.7569)
Saving
{'train_loss': tensor(0.0263, device='cuda:0'), 'train_f1': tensor(0.7514),
'lr': 2.6195292960795287e-05}

4%|          | 9/200 [05:46<2:02:20, 38.43s/it]

Val Loss: tensor(0.0062, device='cuda:0')
CNN Validation Score: tensor(0.7662)
Saving
{'train_loss': tensor(0.0256, device='cuda:0'), 'train_f1': tensor(0.7614),
'lr': 1.2380009889563584e-05}

5%|          | 10/200 [06:24<2:01:41, 38.43s/it]

Val Loss: tensor(0.0059, device='cuda:0')
CNN Validation Score: tensor(0.7739)
Saving
{'train_loss': tensor(0.0232, device='cuda:0'), 'train_f1': tensor(0.7708),
'lr': 0.00012133399685859641}

6%|          | 11/200 [07:03<2:00:54, 38.38s/it]

Val Loss: tensor(0.0056, device='cuda:0')
CNN Validation Score: tensor(0.7810)
Saving
{'train_loss': tensor(0.0231, device='cuda:0'), 'train_f1': tensor(0.7786),
'lr': 0.0002562124637874285}

6%|          | 12/200 [07:41<2:00:24, 38.43s/it]

```

```

Val Loss: tensor(0.0055, device='cuda:0')
CNN Validation Score: tensor(0.7875)
Saving
{'train_loss': tensor(0.0242, device='cuda:0'), 'train_f1': tensor(0.7843),
'lr': 0.000297127399420361}

6%|          | 13/200 [08:20<1:59:48, 38.44s/it]

Val Loss: tensor(0.0060, device='cuda:0')
CNN Validation Score: tensor(0.7925)
Saving
{'train_loss': tensor(0.0214, device='cuda:0'), 'train_f1': tensor(0.7910),
'lr': 0.00020771117313866285}

7%|          | 14/200 [08:58<1:59:05, 38.42s/it]

Val Loss: tensor(0.0049, device='cuda:0')
CNN Validation Score: tensor(0.7982)
Saving
{'train_loss': tensor(0.0215, device='cuda:0'), 'train_f1': tensor(0.7965),
'lr': 6.744225016353256e-05}

8%|          | 15/200 [09:37<1:58:37, 38.48s/it]

Val Loss: tensor(0.0057, device='cuda:0')
CNN Validation Score: tensor(0.8023)
Saving
{'train_loss': tensor(0.0209, device='cuda:0'), 'train_f1': tensor(0.8018),
'lr': 1e-06}

8%|          | 16/200 [10:16<1:58:41, 38.71s/it]

Val Loss: tensor(0.0067, device='cuda:0')
CNN Validation Score: tensor(0.8043)
Saving
{'train_loss': tensor(0.0206, device='cuda:0'), 'train_f1': tensor(0.8063),
'lr': 6.744225016353646e-05}

8%|          | 17/200 [10:54<1:57:37, 38.57s/it]

Val Loss: tensor(0.0047, device='cuda:0')
CNN Validation Score: tensor(0.8088)
Saving
{'train_loss': tensor(0.0192, device='cuda:0'), 'train_f1': tensor(0.8110),
'lr': 0.00020771117313852486}

9%|          | 18/200 [11:33<1:57:04, 38.59s/it]

Val Loss: tensor(0.0043, device='cuda:0')
CNN Validation Score: tensor(0.8132)
Saving
{'train_loss': tensor(0.0191, device='cuda:0'), 'train_f1': tensor(0.8151),
'lr': 0.00029712739942025447}

```

```

10%|          | 19/200 [12:12<1:56:34, 38.64s/it]
Val Loss: tensor(0.0054, device='cuda:0')
CNN Validation Score: tensor(0.8158)
Saving
{'train_loss': tensor(0.0199, device='cuda:0'), 'train_f1': tensor(0.8187),
'lr': 0.0002562124637873881}

10%|          | 20/200 [12:50<1:55:35, 38.53s/it]
Val Loss: tensor(0.0050, device='cuda:0')
CNN Validation Score: tensor(0.8178)
Saving
{'train_loss': tensor(0.0183, device='cuda:0'), 'train_f1': tensor(0.8223),
'lr': 0.00012133399685859908}

10%|          | 21/200 [13:28<1:54:57, 38.53s/it]
Val Loss: tensor(0.0046, device='cuda:0')
CNN Validation Score: tensor(0.8206)
Saving
{'train_loss': tensor(0.0178, device='cuda:0'), 'train_f1': tensor(0.8259),
'lr': 1.2380009889564631e-05}

11%|          | 22/200 [14:07<1:54:10, 38.48s/it]
Val Loss: tensor(0.0040, device='cuda:0')
CNN Validation Score: tensor(0.8242)
Saving
{'train_loss': tensor(0.0180, device='cuda:0'), 'train_f1': tensor(0.8290),
'lr': 2.619529296078206e-05}

12%|          | 23/200 [14:46<1:54:13, 38.72s/it]
Val Loss: tensor(0.0050, device='cuda:0')
CNN Validation Score: tensor(0.8267)
Saving
{'train_loss': tensor(0.0171, device='cuda:0'), 'train_f1': tensor(0.8322),
'lr': 0.00015050000000009026}

12%|          | 24/200 [15:24<1:53:21, 38.64s/it]
Val Loss: tensor(0.0039, device='cuda:0')
CNN Validation Score: tensor(0.8298)
Saving
{'train_loss': tensor(0.0168, device='cuda:0'), 'train_f1': tensor(0.8351),
'lr': 0.000274804707039427}

12%|          | 25/200 [16:03<1:52:30, 38.57s/it]
Val Loss: tensor(0.0044, device='cuda:0')
CNN Validation Score: tensor(0.8324)
Saving

```

```
{'train_loss': tensor(0.0155, device='cuda:0'), 'train_f1': tensor(0.8382),  
'lr': 0.0002886199901106825}
```

```
13%|          | 26/200 [16:41<1:51:35, 38.48s/it]
```

```
Val Loss: tensor(0.0037, device='cuda:0')
```

```
CNN Validation Score: tensor(0.8352)
```

```
Saving
```

```
{'train_loss': tensor(0.0156, device='cuda:0'), 'train_f1': tensor(0.8409),  
'lr': 0.00017966600314159198}
```

```
14%|          | 27/200 [17:20<1:50:56, 38.48s/it]
```

```
Val Loss: tensor(0.0048, device='cuda:0')
```

```
CNN Validation Score: tensor(0.8372)
```

```
Saving
```

```
{'train_loss': tensor(0.0163, device='cuda:0'), 'train_f1': tensor(0.8433),  
'lr': 4.47875362126684e-05}
```

```
14%|          | 28/200 [17:58<1:50:19, 38.49s/it]
```

```
Val Loss: tensor(0.0040, device='cuda:0')
```

```
CNN Validation Score: tensor(0.8395)
```

```
Saving
```

```
{'train_loss': tensor(0.0144, device='cuda:0'), 'train_f1': tensor(0.8458),  
'lr': 3.87260057971705e-06}
```

```
14%|          | 29/200 [18:36<1:49:27, 38.41s/it]
```

```
Val Loss: tensor(0.0035, device='cuda:0')
```

```
CNN Validation Score: tensor(0.8421)
```

```
Saving
```

```
{'train_loss': tensor(0.0153, device='cuda:0'), 'train_f1': tensor(0.8481),  
'lr': 9.328882686131896e-05}
```

```
15%|          | 30/200 [19:15<1:48:52, 38.43s/it]
```

```
Val Loss: tensor(0.0037, device='cuda:0')
```

```
CNN Validation Score: tensor(0.8441)
```

```
Saving
```

```
{'train_loss': tensor(0.0145, device='cuda:0'), 'train_f1': tensor(0.8503),  
'lr': 0.00023355774983626116}
```

```
16%|          | 31/200 [19:53<1:48:03, 38.36s/it]
```

```
Val Loss: tensor(0.0033, device='cuda:0')
```

```
CNN Validation Score: tensor(0.8466)
```

```
Saving
```

```
{'train_loss': tensor(0.0149, device='cuda:0'), 'train_f1': tensor(0.8523),  
'lr': 0.000299999999999980303}
```

```
16%|          | 32/200 [20:31<1:47:11, 38.28s/it]
```

```
Val Loss: tensor(0.0040, device='cuda:0')
```

```
CNN Validation Score: tensor(0.8481)
```

```

Saving
{'train_loss': tensor(0.0144, device='cuda:0'), 'train_f1': tensor(0.8545),
'lr': 0.0002335577498362998}

16%|          | 33/200 [21:10<1:47:11, 38.51s/it]

Val Loss: tensor(0.0040, device='cuda:0')
CNN Validation Score: tensor(0.8498)
Saving
{'train_loss': tensor(0.0139, device='cuda:0'), 'train_f1': tensor(0.8565),
'lr': 9.328882686140936e-05}

17%|          | 34/200 [21:48<1:46:24, 38.46s/it]

Val Loss: tensor(0.0050, device='cuda:0')
CNN Validation Score: tensor(0.8511)
Saving
{'train_loss': tensor(0.0137, device='cuda:0'), 'train_f1': tensor(0.8584),
'lr': 3.87260057971682e-06}

18%|          | 35/200 [22:27<1:45:44, 38.45s/it]

Val Loss: tensor(0.0033, device='cuda:0')
CNN Validation Score: tensor(0.8530)
Saving
{'train_loss': tensor(0.0138, device='cuda:0'), 'train_f1': tensor(0.8603),
'lr': 4.478753621268148e-05}

18%|          | 36/200 [23:05<1:45:00, 38.42s/it]

Val Loss: tensor(0.0028, device='cuda:0')
CNN Validation Score: tensor(0.8551)
Saving
{'train_loss': tensor(0.0124, device='cuda:0'), 'train_f1': tensor(0.8622),
'lr': 0.00017966600314182446}

18%|          | 37/200 [23:44<1:44:19, 38.40s/it]

Val Loss: tensor(0.0029, device='cuda:0')
CNN Validation Score: tensor(0.8571)
Saving
{'train_loss': tensor(0.0134, device='cuda:0'), 'train_f1': tensor(0.8638),
'lr': 0.0002886199901111289}

19%|          | 38/200 [24:22<1:43:56, 38.50s/it]

Val Loss: tensor(0.0035, device='cuda:0')
CNN Validation Score: tensor(0.8586)
Saving
{'train_loss': tensor(0.0130, device='cuda:0'), 'train_f1': tensor(0.8654),
'lr': 0.00027480470703994186}

20%|          | 39/200 [25:01<1:43:13, 38.47s/it]

```

```

Val Loss: tensor(0.0037, device='cuda:0')
CNN Validation Score: tensor(0.8601)
Saving
{'train_loss': tensor(0.0114, device='cuda:0'), 'train_f1': tensor(0.8671),
'lr': 0.00015050000000038435}

20%|          | 40/200 [25:49<1:50:39, 41.50s/it]

Val Loss: tensor(0.0037, device='cuda:0')
CNN Validation Score: tensor(0.8616)
Saving
{'train_loss': tensor(0.0132, device='cuda:0'), 'train_f1': tensor(0.8686),
'lr': 2.6195292960851588e-05}

20%|          | 41/200 [26:36<1:54:14, 43.11s/it]

Val Loss: tensor(0.0028, device='cuda:0')
CNN Validation Score: tensor(0.8632)
Saving
{'train_loss': tensor(0.0114, device='cuda:0'), 'train_f1': tensor(0.8703),
'lr': 1.2380009889574763e-05}

21%|          | 42/200 [27:20<1:54:00, 43.30s/it]

Val Loss: tensor(0.0029, device='cuda:0')
CNN Validation Score: tensor(0.8646)
Saving
{'train_loss': tensor(0.0114, device='cuda:0'), 'train_f1': tensor(0.8718),
'lr': 0.00012133399685866668}

22%|          | 43/200 [28:04<1:54:03, 43.59s/it]

Val Loss: tensor(0.0025, device='cuda:0')
CNN Validation Score: tensor(0.8663)
Saving
{'train_loss': tensor(0.0127, device='cuda:0'), 'train_f1': tensor(0.8731),
'lr': 0.0002562124637876399}

22%|          | 44/200 [28:49<1:54:19, 43.97s/it]

Val Loss: tensor(0.0023, device='cuda:0')
CNN Validation Score: tensor(0.8681)
Saving
{'train_loss': tensor(0.0114, device='cuda:0'), 'train_f1': tensor(0.8745),
'lr': 0.0002971273994206079}

22%|          | 45/200 [29:34<1:54:37, 44.37s/it]

Val Loss: tensor(0.0031, device='cuda:0')
CNN Validation Score: tensor(0.8695)
Saving
{'train_loss': tensor(0.0104, device='cuda:0'), 'train_f1': tensor(0.8760),
'lr': 0.00020771117313882662}

```

```

23%|          | 46/200 [30:19<1:54:21, 44.55s/it]
Val Loss: tensor(0.0025, device='cuda:0')
CNN Validation Score: tensor(0.8711)
Saving
{'train_loss': tensor(0.0121, device='cuda:0'), 'train_f1': tensor(0.8772),
'lr': 6.744225016368753e-05}

24%|          | 47/200 [31:03<1:53:02, 44.33s/it]
Val Loss: tensor(0.0030, device='cuda:0')
CNN Validation Score: tensor(0.8722)
Saving
{'train_loss': tensor(0.0104, device='cuda:0'), 'train_f1': tensor(0.8786),
'lr': 1e-06}

24%|          | 48/200 [31:47<1:51:44, 44.11s/it]
Val Loss: tensor(0.0032, device='cuda:0')
CNN Validation Score: tensor(0.8733)
Saving
{'train_loss': tensor(0.0113, device='cuda:0'), 'train_f1': tensor(0.8798),
'lr': 6.744225016348969e-05}

24%|          | 49/200 [32:30<1:50:40, 43.98s/it]
Val Loss: tensor(0.0026, device='cuda:0')
CNN Validation Score: tensor(0.8747)
Saving
{'train_loss': tensor(0.0093, device='cuda:0'), 'train_f1': tensor(0.8811),
'lr': 0.00020771117313845902}

25%|          | 50/200 [33:14<1:49:29, 43.80s/it]
Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.8762)
Saving
{'train_loss': tensor(0.0105, device='cuda:0'), 'train_f1': tensor(0.8823),
'lr': 0.0002971273994201495}

26%|          | 51/200 [33:58<1:48:54, 43.86s/it]
Val Loss: tensor(0.0027, device='cuda:0')
CNN Validation Score: tensor(0.8773)
Saving
{'train_loss': tensor(0.0119, device='cuda:0'), 'train_f1': tensor(0.8834),
'lr': 0.00025621246378732313}

26%|          | 52/200 [34:39<1:46:31, 43.19s/it]
Val Loss: tensor(0.0028, device='cuda:0')
CNN Validation Score: tensor(0.8784)
Saving

```

```

{'train_loss': tensor(0.0088, device='cuda:0'), 'train_f1': tensor(0.8847),
'lr': 0.00012133399685855329}

26%|          | 53/200 [35:18<1:42:31, 41.84s/it]
Val Loss: tensor(0.0023, device='cuda:0')
CNN Validation Score: tensor(0.8797)
Saving
{'train_loss': tensor(0.0088, device='cuda:0'), 'train_f1': tensor(0.8860),
'lr': 1.2380009889571853e-05}

27%|          | 54/200 [35:57<1:39:33, 40.92s/it]
Val Loss: tensor(0.0024, device='cuda:0')
CNN Validation Score: tensor(0.8810)
Saving
{'train_loss': tensor(0.0089, device='cuda:0'), 'train_f1': tensor(0.8872),
'lr': 2.6195292960834478e-05}

28%|          | 55/200 [36:35<1:37:15, 40.24s/it]
Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.8821)
Saving
{'train_loss': tensor(0.0091, device='cuda:0'), 'train_f1': tensor(0.8883),
'lr': 0.00015050000000003558}

28%|          | 56/200 [37:15<1:35:42, 39.88s/it]
Val Loss: tensor(0.0017, device='cuda:0')
CNN Validation Score: tensor(0.8835)
Saving
{'train_loss': tensor(0.0087, device='cuda:0'), 'train_f1': tensor(0.8894),
'lr': 0.00027480470703996934}

28%|          | 57/200 [37:53<1:34:20, 39.59s/it]
Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.8846)
Saving
{'train_loss': tensor(0.0082, device='cuda:0'), 'train_f1': tensor(0.8905),
'lr': 0.0002886199901112604}

29%|          | 58/200 [38:32<1:32:57, 39.28s/it]
Val Loss: tensor(0.0022, device='cuda:0')
CNN Validation Score: tensor(0.8857)
Saving
{'train_loss': tensor(0.0091, device='cuda:0'), 'train_f1': tensor(0.8916),
'lr': 0.000179666003141181}

30%|          | 59/200 [39:11<1:31:58, 39.14s/it]
Val Loss: tensor(0.0022, device='cuda:0')
CNN Validation Score: tensor(0.8867)

```



```

Saving
{'train_loss': tensor(0.0084, device='cuda:0'), 'train_f1': tensor(0.8926),
'lr': 4.4787536212573e-05}

30%|          | 60/200 [39:50<1:31:34, 39.25s/it]

Val Loss: tensor(0.0025, device='cuda:0')
CNN Validation Score: tensor(0.8877)
Saving
{'train_loss': tensor(0.0092, device='cuda:0'), 'train_f1': tensor(0.8935),
'lr': 3.87260057971705e-06}

30%|          | 61/200 [40:29<1:30:35, 39.10s/it]

Val Loss: tensor(0.0025, device='cuda:0')
CNN Validation Score: tensor(0.8887)
Saving
{'train_loss': tensor(0.0082, device='cuda:0'), 'train_f1': tensor(0.8945),
'lr': 9.328882686153678e-05}

31%|          | 62/200 [41:08<1:29:59, 39.13s/it]

Val Loss: tensor(0.0023, device='cuda:0')
CNN Validation Score: tensor(0.8896)
Saving
{'train_loss': tensor(0.0084, device='cuda:0'), 'train_f1': tensor(0.8955),
'lr': 0.00023355774983663865}

32%|          | 63/200 [41:51<1:31:49, 40.22s/it]

Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.8906)
Saving
{'train_loss': tensor(0.0071, device='cuda:0'), 'train_f1': tensor(0.8965),
'lr': 0.000300000000000003921}

32%|          | 64/200 [42:44<1:39:40, 43.97s/it]

Val Loss: tensor(0.0029, device='cuda:0')
CNN Validation Score: tensor(0.8913)
Saving
{'train_loss': tensor(0.0079, device='cuda:0'), 'train_f1': tensor(0.8974),
'lr': 0.0002335577498367809}

32%|          | 65/200 [43:27<1:38:40, 43.86s/it]

Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.8923)
Saving
{'train_loss': tensor(0.0069, device='cuda:0'), 'train_f1': tensor(0.8985),
'lr': 9.328882686155512e-05}

33%|          | 66/200 [44:11<1:37:57, 43.86s/it]

```

```

Val Loss: tensor(0.0022, device='cuda:0')
CNN Validation Score: tensor(0.8931)
Saving
{'train_loss': tensor(0.0073, device='cuda:0'), 'train_f1': tensor(0.8994),
'lr': 3.872600579714875e-06}

34%|          | 67/200 [44:53<1:35:50, 43.23s/it]

Val Loss: tensor(0.0019, device='cuda:0')
CNN Validation Score: tensor(0.8941)
Saving
{'train_loss': tensor(0.0079, device='cuda:0'), 'train_f1': tensor(0.9002),
'lr': 4.478753621274724e-05}

34%|          | 68/200 [45:43<1:39:49, 45.37s/it]

Val Loss: tensor(0.0020, device='cuda:0')
CNN Validation Score: tensor(0.8950)
Saving
{'train_loss': tensor(0.0070, device='cuda:0'), 'train_f1': tensor(0.9012),
'lr': 0.00017966600314214807}

34%|          | 69/200 [46:29<1:39:24, 45.53s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.8961)
Saving
{'train_loss': tensor(0.0070, device='cuda:0'), 'train_f1': tensor(0.9021),
'lr': 0.00028861999011171044}

35%|          | 70/200 [47:12<1:36:58, 44.76s/it]

Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.8970)
Saving
{'train_loss': tensor(0.0076, device='cuda:0'), 'train_f1': tensor(0.9029),
'lr': 0.0002748047070404597}

36%|          | 71/200 [47:54<1:34:35, 43.99s/it]

Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.8979)
Saving
{'train_loss': tensor(0.0066, device='cuda:0'), 'train_f1': tensor(0.9037),
'lr': 0.00015050000000077501}

36%|          | 72/200 [48:37<1:32:49, 43.51s/it]

Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.8990)
Saving
{'train_loss': tensor(0.0077, device='cuda:0'), 'train_f1': tensor(0.9045),
'lr': 2.6195292960917382e-05}

```

```

36%|          | 73/200 [49:19<1:31:24, 43.18s/it]
Val Loss: tensor(0.0020, device='cuda:0')
CNN Validation Score: tensor(0.8998)
Saving
{'train_loss': tensor(0.0068, device='cuda:0'), 'train_f1': tensor(0.9053),
'lr': 1.2380009889533543e-05}

37%|          | 74/200 [50:02<1:30:22, 43.04s/it]
Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9006)
Saving
{'train_loss': tensor(0.0071, device='cuda:0'), 'train_f1': tensor(0.9061),
'lr': 0.0001213339968583858}

38%|          | 75/200 [50:44<1:29:14, 42.83s/it]
Val Loss: tensor(0.0017, device='cuda:0')
CNN Validation Score: tensor(0.9014)
Saving
{'train_loss': tensor(0.0066, device='cuda:0'), 'train_f1': tensor(0.9069),
'lr': 0.00025621246378689595}

38%|          | 76/200 [51:27<1:28:28, 42.81s/it]
Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9022)
Saving
{'train_loss': tensor(0.0061, device='cuda:0'), 'train_f1': tensor(0.9076),
'lr': 0.00029712739941983}

38%|          | 77/200 [52:10<1:28:08, 43.00s/it]
Val Loss: tensor(0.0017, device='cuda:0')
CNN Validation Score: tensor(0.9030)
Saving
{'train_loss': tensor(0.0062, device='cuda:0'), 'train_f1': tensor(0.9084),
'lr': 0.0002077111731383063}

39%|          | 78/200 [52:54<1:27:49, 43.19s/it]
Val Loss: tensor(0.0023, device='cuda:0')
CNN Validation Score: tensor(0.9037)
Saving
{'train_loss': tensor(0.0059, device='cuda:0'), 'train_f1': tensor(0.9091),
'lr': 6.744225016347936e-05}

40%|          | 79/200 [53:38<1:27:17, 43.28s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9045)
Saving

```

```

{'train_loss': tensor(0.0054, device='cuda:0'), 'train_f1': tensor(0.9099),
'lr': 1e-06}

40%|          | 80/200 [54:20<1:26:18, 43.16s/it]
Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.9052)
Saving
{'train_loss': tensor(0.0058, device='cuda:0'), 'train_f1': tensor(0.9106),
'lr': 6.74422501635965e-05}

40%|          | 81/200 [55:03<1:25:18, 43.01s/it]
Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9060)
Saving
{'train_loss': tensor(0.0060, device='cuda:0'), 'train_f1': tensor(0.9113),
'lr': 0.00020771117313883955}

41%|          | 82/200 [55:45<1:24:01, 42.72s/it]
Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9068)
Saving
{'train_loss': tensor(0.0050, device='cuda:0'), 'train_f1': tensor(0.9120),
'lr': 0.0002971273994207407}

42%|          | 83/200 [56:28<1:23:04, 42.60s/it]
Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9076)
Saving
{'train_loss': tensor(0.0055, device='cuda:0'), 'train_f1': tensor(0.9128),
'lr': 0.00025621246378779844}

42%|          | 84/200 [57:10<1:22:22, 42.61s/it]
Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9083)
Saving
{'train_loss': tensor(0.0062, device='cuda:0'), 'train_f1': tensor(0.9134),
'lr': 0.00012133399685888545}

42%|          | 85/200 [57:56<1:23:32, 43.59s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9091)
Saving
{'train_loss': tensor(0.0056, device='cuda:0'), 'train_f1': tensor(0.9141),
'lr': 1.2380009889605367e-05}

43%|          | 86/200 [58:39<1:22:43, 43.54s/it]
Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9097)

```

```

Saving
{'train_loss': tensor(0.0043, device='cuda:0'), 'train_f1': tensor(0.9148),
'lr': 2.6195292960868003e-05}

44%|          | 87/200 [59:24<1:22:33, 43.84s/it]

Val Loss: tensor(0.0014, device='cuda:0')
CNN Validation Score: tensor(0.9104)
Saving
{'train_loss': tensor(0.0060, device='cuda:0'), 'train_f1': tensor(0.9154),
'lr': 0.00015050000000068933}

44%|          | 88/200 [1:00:08<1:21:48, 43.83s/it]

Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9111)
Saving
{'train_loss': tensor(0.0050, device='cuda:0'), 'train_f1': tensor(0.9161),
'lr': 0.00027480470704045496}

44%|          | 89/200 [1:00:51<1:20:43, 43.64s/it]

Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9118)
Saving
{'train_loss': tensor(0.0054, device='cuda:0'), 'train_f1': tensor(0.9167),
'lr': 0.00028861999011183827}

45%|          | 90/200 [1:01:35<1:20:03, 43.67s/it]

Val Loss: tensor(0.0017, device='cuda:0')
CNN Validation Score: tensor(0.9124)
Saving
{'train_loss': tensor(0.0055, device='cuda:0'), 'train_f1': tensor(0.9173),
'lr': 0.0001796660031423195}

46%|          | 91/200 [1:02:18<1:18:56, 43.46s/it]

Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9130)
Saving
{'train_loss': tensor(0.0052, device='cuda:0'), 'train_f1': tensor(0.9179),
'lr': 4.478753621283384e-05}

46%|          | 92/200 [1:03:01<1:18:15, 43.48s/it]

Val Loss: tensor(0.0017, device='cuda:0')
CNN Validation Score: tensor(0.9136)
Saving
{'train_loss': tensor(0.0049, device='cuda:0'), 'train_f1': tensor(0.9185),
'lr': 3.87260057971705e-06}

46%|          | 93/200 [1:03:44<1:17:14, 43.31s/it]

```

```

Val Loss: tensor(0.0019, device='cuda:0')
CNN Validation Score: tensor(0.9142)
Saving
{'train_loss': tensor(0.0052, device='cuda:0'), 'train_f1': tensor(0.9190),
'lr': 9.328882686120297e-05}

47%|          | 94/200 [1:04:28<1:16:33, 43.34s/it]

Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9148)
Saving
{'train_loss': tensor(0.0046, device='cuda:0'), 'train_f1': tensor(0.9196),
'lr': 0.0002335577498360557}

48%|          | 95/200 [1:05:11<1:15:47, 43.31s/it]

Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9154)
Saving
{'train_loss': tensor(0.0050, device='cuda:0'), 'train_f1': tensor(0.9202),
'lr': 0.000299999999999996008}

48%|          | 96/200 [1:05:54<1:15:17, 43.44s/it]

Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9160)
Saving
{'train_loss': tensor(0.0058, device='cuda:0'), 'train_f1': tensor(0.9207),
'lr': 0.00023355774983613306}

48%|          | 97/200 [1:06:38<1:14:49, 43.59s/it]

Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9166)
Saving
{'train_loss': tensor(0.0041, device='cuda:0'), 'train_f1': tensor(0.9213),
'lr': 9.328882686140056e-05}

49%|          | 98/200 [1:07:22<1:14:16, 43.69s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9172)
Saving
{'train_loss': tensor(0.0044, device='cuda:0'), 'train_f1': tensor(0.9218),
'lr': 3.872600579726197e-06}

50%|          | 99/200 [1:08:06<1:13:26, 43.62s/it]

Val Loss: tensor(0.0024, device='cuda:0')
CNN Validation Score: tensor(0.9176)
Saving
{'train_loss': tensor(0.0049, device='cuda:0'), 'train_f1': tensor(0.9223),
'lr': 4.478753621265868e-05}

```

```

50%|          | 100/200 [1:08:49<1:12:28, 43.49s/it]
Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9181)
Saving
{'train_loss': tensor(0.0053, device='cuda:0'), 'train_f1': tensor(0.9228),
'lr': 0.00017966600314171346}

50%|          | 101/200 [1:09:32<1:11:42, 43.46s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9187)
Saving
{'train_loss': tensor(0.0038, device='cuda:0'), 'train_f1': tensor(0.9234),
'lr': 0.00028861999011091236}

51%|          | 102/200 [1:10:19<1:12:36, 44.45s/it]
Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9192)
Saving
{'train_loss': tensor(0.0047, device='cuda:0'), 'train_f1': tensor(0.9239),
'lr': 0.00027480470703978926}

52%|          | 103/200 [1:11:07<1:13:33, 45.50s/it]
Val Loss: tensor(0.0023, device='cuda:0')
CNN Validation Score: tensor(0.9196)
Saving
{'train_loss': tensor(0.0046, device='cuda:0'), 'train_f1': tensor(0.9244),
'lr': 0.00015050000000034044}

52%|          | 104/200 [1:11:48<1:10:49, 44.26s/it]
Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9201)
Saving
{'train_loss': tensor(0.0040, device='cuda:0'), 'train_f1': tensor(0.9249),
'lr': 2.6195292960829026e-05}

52%|          | 105/200 [1:12:30<1:08:48, 43.46s/it]
Val Loss: tensor(0.0020, device='cuda:0')
CNN Validation Score: tensor(0.9206)
Saving
{'train_loss': tensor(0.0042, device='cuda:0'), 'train_f1': tensor(0.9254),
'lr': 1.2380009889623425e-05}

53%|          | 106/200 [1:13:12<1:07:09, 42.87s/it]
Val Loss: tensor(0.0028, device='cuda:0')
CNN Validation Score: tensor(0.9209)
Saving

```

```
{'train_loss': tensor(0.0044, device='cuda:0'), 'train_f1': tensor(0.9259),  
'lr': 0.00012133399685908332}
```

```
54%|          | 107/200 [1:13:53<1:05:59, 42.57s/it]
```

```
Val Loss: tensor(0.0010, device='cuda:0')
```

```
CNN Validation Score: tensor(0.9214)
```

```
Saving
```

```
{'train_loss': tensor(0.0031, device='cuda:0'), 'train_f1': tensor(0.9264),  
'lr': 0.0002562124637886046}
```

```
54%|          | 108/200 [1:14:35<1:04:58, 42.38s/it]
```

```
Val Loss: tensor(0.0011, device='cuda:0')
```

```
CNN Validation Score: tensor(0.9220)
```

```
Saving
```

```
{'train_loss': tensor(0.0042, device='cuda:0'), 'train_f1': tensor(0.9269),  
'lr': 0.00029712739942177295}
```

```
55%|          | 109/200 [1:15:17<1:03:56, 42.16s/it]
```

```
Val Loss: tensor(0.0014, device='cuda:0')
```

```
CNN Validation Score: tensor(0.9225)
```

```
Saving
```

```
{'train_loss': tensor(0.0051, device='cuda:0'), 'train_f1': tensor(0.9273),  
'lr': 0.00020771117313963202}
```

```
55%|          | 110/200 [1:15:58<1:02:49, 41.89s/it]
```

```
Val Loss: tensor(0.0018, device='cuda:0')
```

```
CNN Validation Score: tensor(0.9229)
```

```
Saving
```

```
{'train_loss': tensor(0.0042, device='cuda:0'), 'train_f1': tensor(0.9278),  
'lr': 6.74422501639976e-05}
```

```
56%|          | 111/200 [1:16:39<1:01:45, 41.64s/it]
```

```
Val Loss: tensor(0.0010, device='cuda:0')
```

```
CNN Validation Score: tensor(0.9235)
```

```
Saving
```

```
{'train_loss': tensor(0.0040, device='cuda:0'), 'train_f1': tensor(0.9283),  
'lr': 1e-06}
```

```
56%|          | 112/200 [1:17:28<1:04:19, 43.86s/it]
```

```
Val Loss: tensor(0.0015, device='cuda:0')
```

```
CNN Validation Score: tensor(0.9239)
```

```
Saving
```

```
{'train_loss': tensor(0.0037, device='cuda:0'), 'train_f1': tensor(0.9287),  
'lr': 6.744225016345307e-05}
```

```
56%|          | 113/200 [1:18:20<1:06:51, 46.11s/it]
```

```
Val Loss: tensor(0.0015, device='cuda:0')
```

```
CNN Validation Score: tensor(0.9244)
```



```

Saving
{'train_loss': tensor(0.0040, device='cuda:0'), 'train_f1': tensor(0.9292),
'lr': 0.00020771117313832848}

57%|          | 114/200 [1:19:10<1:08:05, 47.50s/it]

Val Loss: tensor(0.0019, device='cuda:0')
CNN Validation Score: tensor(0.9248)
Saving
{'train_loss': tensor(0.0035, device='cuda:0'), 'train_f1': tensor(0.9297),
'lr': 0.0002971273994199362}

57%|          | 115/200 [1:20:01<1:08:46, 48.55s/it]

Val Loss: tensor(0.0025, device='cuda:0')
CNN Validation Score: tensor(0.9251)
Saving
{'train_loss': tensor(0.0043, device='cuda:0'), 'train_f1': tensor(0.9301),
'lr': 0.00025621246378719167}

58%|          | 116/200 [1:20:52<1:08:36, 49.01s/it]

Val Loss: tensor(0.0019, device='cuda:0')
CNN Validation Score: tensor(0.9254)
Saving
{'train_loss': tensor(0.0033, device='cuda:0'), 'train_f1': tensor(0.9305),
'lr': 0.00012133399685950653}

58%|          | 117/200 [1:21:38<1:06:51, 48.33s/it]

Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9259)
Saving
{'train_loss': tensor(0.0039, device='cuda:0'), 'train_f1': tensor(0.9309),
'lr': 1.2380009889717456e-05}

59%|          | 118/200 [1:22:25<1:05:33, 47.97s/it]

Val Loss: tensor(0.0008, device='cuda:0')
CNN Validation Score: tensor(0.9264)
Saving
{'train_loss': tensor(0.0029, device='cuda:0'), 'train_f1': tensor(0.9314),
'lr': 2.6195292960747454e-05}

60%|          | 119/200 [1:23:13<1:04:39, 47.90s/it]

Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9268)
Saving
{'train_loss': tensor(0.0044, device='cuda:0'), 'train_f1': tensor(0.9318),
'lr': 0.000150500000000019673}

60%|          | 120/200 [1:24:03<1:04:28, 48.36s/it]

```

```

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9272)
Saving
{'train_loss': tensor(0.0032, device='cuda:0'), 'train_f1': tensor(0.9322),
'lr': 0.0002748047070397519}

60%|          | 121/200 [1:24:53<1:04:30, 48.99s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9276)
Saving
{'train_loss': tensor(0.0033, device='cuda:0'), 'train_f1': tensor(0.9326),
'lr': 0.0002886199901110604}

61%|          | 122/200 [1:25:44<1:04:26, 49.57s/it]

Val Loss: tensor(0.0014, device='cuda:0')
CNN Validation Score: tensor(0.9281)
Saving
{'train_loss': tensor(0.0043, device='cuda:0'), 'train_f1': tensor(0.9330),
'lr': 0.00017966600314180863}

62%|          | 123/200 [1:26:34<1:03:57, 49.84s/it]

Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9284)
Saving
{'train_loss': tensor(0.0034, device='cuda:0'), 'train_f1': tensor(0.9334),
'lr': 4.478753621269042e-05}

62%|          | 124/200 [1:27:24<1:03:01, 49.76s/it]

Val Loss: tensor(0.0010, device='cuda:0')
CNN Validation Score: tensor(0.9289)
Saving
{'train_loss': tensor(0.0030, device='cuda:0'), 'train_f1': tensor(0.9338),
'lr': 3.87260057971705e-06}

62%|          | 125/200 [1:28:11<1:01:19, 49.06s/it]

Val Loss: tensor(0.0019, device='cuda:0')
CNN Validation Score: tensor(0.9292)
Saving
{'train_loss': tensor(0.0039, device='cuda:0'), 'train_f1': tensor(0.9342),
'lr': 9.328882686086911e-05}

63%|          | 126/200 [1:28:59<1:00:04, 48.72s/it]

Val Loss: tensor(0.0019, device='cuda:0')
CNN Validation Score: tensor(0.9296)
Saving
{'train_loss': tensor(0.0038, device='cuda:0'), 'train_f1': tensor(0.9346),
'lr': 0.00023355774983535962}

```

```

64%|          | 127/200 [1:29:48<59:20, 48.78s/it]
Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9299)
Saving
{'train_loss': tensor(0.0033, device='cuda:0'), 'train_f1': tensor(0.9350),
'lr': 0.00029999999999988098}

64%|          | 128/200 [1:30:37<58:32, 48.78s/it]
Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9304)
Saving
{'train_loss': tensor(0.0031, device='cuda:0'), 'train_f1': tensor(0.9354),
'lr': 0.00023355774983559836}

64%|          | 129/200 [1:31:26<57:45, 48.81s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9307)
Saving
{'train_loss': tensor(0.0029, device='cuda:0'), 'train_f1': tensor(0.9357),
'lr': 9.328882686111984e-05}

65%|          | 130/200 [1:32:15<57:08, 48.98s/it]
Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9311)
Saving
{'train_loss': tensor(0.0029, device='cuda:0'), 'train_f1': tensor(0.9361),
'lr': 3.872600579710993e-06}

66%|          | 131/200 [1:33:05<56:41, 49.30s/it]
Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9315)
Saving
{'train_loss': tensor(0.0034, device='cuda:0'), 'train_f1': tensor(0.9365),
'lr': 4.4787536212974674e-05}

66%|          | 132/200 [1:33:55<55:55, 49.34s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9319)
Saving
{'train_loss': tensor(0.0030, device='cuda:0'), 'train_f1': tensor(0.9368),
'lr': 0.0001796660031429285}

66%|          | 133/200 [1:34:44<54:59, 49.25s/it]
Val Loss: tensor(0.0010, device='cuda:0')
CNN Validation Score: tensor(0.9323)
Saving

```

```

{'train_loss': tensor(0.0030, device='cuda:0'), 'train_f1': tensor(0.9372),
'lr': 0.0002886199901128738}

67%|          | 134/200 [1:35:30<53:01, 48.21s/it]
Val Loss: tensor(0.0010, device='cuda:0')
CNN Validation Score: tensor(0.9327)
Saving
{'train_loss': tensor(0.0023, device='cuda:0'), 'train_f1': tensor(0.9376),
'lr': 0.0002748047070416466}

68%|          | 135/200 [1:36:17<51:49, 47.84s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9330)
Saving
{'train_loss': tensor(0.0028, device='cuda:0'), 'train_f1': tensor(0.9379),
'lr': 0.00015050000000142217}

68%|          | 136/200 [1:37:04<51:00, 47.83s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9334)
Saving
{'train_loss': tensor(0.0027, device='cuda:0'), 'train_f1': tensor(0.9383),
'lr': 2.6195292961048892e-05}

68%|          | 137/200 [1:37:54<50:40, 48.26s/it]
Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9337)
Saving
{'train_loss': tensor(0.0026, device='cuda:0'), 'train_f1': tensor(0.9386),
'lr': 1.238000988955618e-05}

69%|          | 138/200 [1:38:43<50:05, 48.47s/it]
Val Loss: tensor(0.0009, device='cuda:0')
CNN Validation Score: tensor(0.9341)
Saving
{'train_loss': tensor(0.0025, device='cuda:0'), 'train_f1': tensor(0.9390),
'lr': 0.00012133399685880306}

70%|          | 139/200 [1:39:32<49:40, 48.86s/it]
Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9344)
Saving
{'train_loss': tensor(0.0036, device='cuda:0'), 'train_f1': tensor(0.9393),
'lr': 0.0002562124637879552}

70%|          | 140/200 [1:40:23<49:16, 49.27s/it]
Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9347)

```

```

Saving
{'train_loss': tensor(0.0034, device='cuda:0'), 'train_f1': tensor(0.9396),
'lr': 0.00029712739942098154}

70%|          | 141/200 [1:41:08<47:24, 48.22s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9350)
Saving
{'train_loss': tensor(0.0030, device='cuda:0'), 'train_f1': tensor(0.9399),
'lr': 0.0002077111731390489}

71%|          | 142/200 [1:41:59<47:10, 48.80s/it]

Val Loss: tensor(0.0020, device='cuda:0')
CNN Validation Score: tensor(0.9353)
Saving
{'train_loss': tensor(0.0031, device='cuda:0'), 'train_f1': tensor(0.9402),
'lr': 6.744225016367634e-05}

72%|          | 143/200 [1:42:45<45:42, 48.11s/it]

Val Loss: tensor(0.0022, device='cuda:0')
CNN Validation Score: tensor(0.9355)
Saving
{'train_loss': tensor(0.0036, device='cuda:0'), 'train_f1': tensor(0.9405),
'lr': 1e-06}

72%|          | 144/200 [1:43:33<44:52, 48.07s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9358)
Saving
{'train_loss': tensor(0.0031, device='cuda:0'), 'train_f1': tensor(0.9408),
'lr': 6.74422501631966e-05}

72%|          | 145/200 [1:44:21<44:03, 48.06s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9361)
Saving
{'train_loss': tensor(0.0028, device='cuda:0'), 'train_f1': tensor(0.9411),
'lr': 0.0002077111731376919}

73%|          | 146/200 [1:45:10<43:30, 48.34s/it]

Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9365)
Saving
{'train_loss': tensor(0.0027, device='cuda:0'), 'train_f1': tensor(0.9414),
'lr': 0.0002971273994191602}

74%|          | 147/200 [1:45:59<42:48, 48.46s/it]

```

```

Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9368)
Saving
{'train_loss': tensor(0.0022, device='cuda:0'), 'train_f1': tensor(0.9418),
'lr': 0.0002562124637864871}

74%|          | 148/200 [1:46:48<42:17, 48.79s/it]

Val Loss: tensor(0.0017, device='cuda:0')
CNN Validation Score: tensor(0.9371)
Saving
{'train_loss': tensor(0.0021, device='cuda:0'), 'train_f1': tensor(0.9421),
'lr': 0.00012133399685817215}

74%|          | 149/200 [1:47:39<41:53, 49.29s/it]

Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9374)
Saving
{'train_loss': tensor(0.0019, device='cuda:0'), 'train_f1': tensor(0.9424),
'lr': 1.2380009889515308e-05}

75%|          | 150/200 [1:48:28<41:00, 49.22s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9377)
Saving
{'train_loss': tensor(0.0022, device='cuda:0'), 'train_f1': tensor(0.9427),
'lr': 2.6195292961010603e-05}

76%|          | 151/200 [1:49:18<40:25, 49.49s/it]

Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9379)
Saving
{'train_loss': tensor(0.0016, device='cuda:0'), 'train_f1': tensor(0.9430),
'lr': 0.00015050000000108447}
Val Loss: tensor(0.0014, device='cuda:0')
CNN Validation Score: tensor(0.9382)
Saving

76%|          | 152/200 [1:50:05<39:00, 48.76s/it]

{'train_loss': tensor(0.0030, device='cuda:0'), 'train_f1': tensor(0.9433),
'lr': 0.00027480470704150154}

76%|          | 153/200 [1:50:51<37:34, 47.96s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9385)
Saving
{'train_loss': tensor(0.0024, device='cuda:0'), 'train_f1': tensor(0.9436),
'lr': 0.000288619990112992}

```

```

77%|          | 154/200 [1:51:36<36:09, 47.17s/it]
Val Loss: tensor(0.0010, device='cuda:0')
CNN Validation Score: tensor(0.9388)
Saving
{'train_loss': tensor(0.0030, device='cuda:0'), 'train_f1': tensor(0.9439),
'lr': 0.0001796660031430805}

78%|          | 155/200 [1:52:23<35:17, 47.06s/it]
Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.9391)
Saving
{'train_loss': tensor(0.0025, device='cuda:0'), 'train_f1': tensor(0.9442),
'lr': 4.4787536213047424e-05}

78%|          | 156/200 [1:53:12<34:51, 47.52s/it]
Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9394)
Saving
{'train_loss': tensor(0.0019, device='cuda:0'), 'train_f1': tensor(0.9444),
'lr': 3.87260057971705e-06}

78%|          | 157/200 [1:54:01<34:19, 47.91s/it]
Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9397)
Saving
{'train_loss': tensor(0.0022, device='cuda:0'), 'train_f1': tensor(0.9447),
'lr': 9.328882686163865e-05}

79%|          | 158/200 [1:54:51<34:02, 48.64s/it]
Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9399)
Saving
{'train_loss': tensor(0.0025, device='cuda:0'), 'train_f1': tensor(0.9450),
'lr': 0.00023355774983703688}

80%|          | 159/200 [1:55:41<33:30, 49.03s/it]
Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9402)
Saving
{'train_loss': tensor(0.0019, device='cuda:0'), 'train_f1': tensor(0.9453),
'lr': 0.000300000000000007792}

80%|          | 160/200 [1:56:29<32:33, 48.84s/it]
Val Loss: tensor(0.0007, device='cuda:0')
CNN Validation Score: tensor(0.9405)
Saving

```

```

{'train_loss': tensor(0.0028, device='cuda:0'), 'train_f1': tensor(0.9455),
'lr': 0.00023355774983721086}

80%|      | 161/200 [1:57:16<31:20, 48.22s/it]
Val Loss: tensor(0.0014, device='cuda:0')
CNN Validation Score: tensor(0.9407)
Saving
{'train_loss': tensor(0.0029, device='cuda:0'), 'train_f1': tensor(0.9458),
'lr': 9.32888268618173e-05}

81%|      | 162/200 [1:58:04<30:23, 47.98s/it]
Val Loss: tensor(0.0020, device='cuda:0')
CNN Validation Score: tensor(0.9409)
Saving
{'train_loss': tensor(0.0028, device='cuda:0'), 'train_f1': tensor(0.9460),
'lr': 3.872600579748834e-06}

82%|      | 163/200 [1:58:54<29:58, 48.60s/it]
Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9412)
Saving
{'train_loss': tensor(0.0017, device='cuda:0'), 'train_f1': tensor(0.9463),
'lr': 4.478753621238592e-05}

82%|      | 164/200 [1:59:43<29:16, 48.81s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9414)
Saving
{'train_loss': tensor(0.0016, device='cuda:0'), 'train_f1': tensor(0.9466),
'lr': 0.00017966600314070995}

82%|      | 165/200 [2:00:32<28:27, 48.80s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9417)
Saving
{'train_loss': tensor(0.0014, device='cuda:0'), 'train_f1': tensor(0.9469),
'lr': 0.00028861999010941985}

83%|      | 166/200 [2:01:22<27:55, 49.27s/it]
Val Loss: tensor(0.0007, device='cuda:0')
CNN Validation Score: tensor(0.9420)
Saving
{'train_loss': tensor(0.0034, device='cuda:0'), 'train_f1': tensor(0.9471),
'lr': 0.00027480470703829664}

84%|      | 167/200 [2:02:12<27:11, 49.44s/it]
Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9422)

```



```

Saving
{'train_loss': tensor(0.0018, device='cuda:0'), 'train_f1': tensor(0.9474),
'lr': 0.000150499999999947113}

84%|      | 168/200 [2:03:02<26:29, 49.66s/it]

Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9425)
Saving
{'train_loss': tensor(0.0021, device='cuda:0'), 'train_f1': tensor(0.9476),
'lr': 2.619529296065222e-05}

84%|      | 169/200 [2:03:49<25:16, 48.92s/it]

Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9427)
Saving
{'train_loss': tensor(0.0022, device='cuda:0'), 'train_f1': tensor(0.9478),
'lr': 1.2380009889594027e-05}

85%|      | 170/200 [2:05:16<30:12, 60.40s/it]

Val Loss: tensor(0.0028, device='cuda:0')
CNN Validation Score: tensor(0.9429)
Saving
{'train_loss': tensor(0.0025, device='cuda:0'), 'train_f1': tensor(0.9481),
'lr': 0.00012133399685950057}

86%|      | 171/200 [2:05:55<26:02, 53.89s/it]

Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9431)
Saving
{'train_loss': tensor(0.0018, device='cuda:0'), 'train_f1': tensor(0.9483),
'lr': 0.0002562124637895693}

86%|      | 172/200 [2:06:33<22:57, 49.20s/it]

Val Loss: tensor(0.0008, device='cuda:0')
CNN Validation Score: tensor(0.9434)
Saving
{'train_loss': tensor(0.0018, device='cuda:0'), 'train_f1': tensor(0.9486),
'lr': 0.00029712739942295115}

86%|      | 173/200 [2:07:12<20:40, 45.95s/it]

Val Loss: tensor(0.0021, device='cuda:0')
CNN Validation Score: tensor(0.9436)
Saving
{'train_loss': tensor(0.0033, device='cuda:0'), 'train_f1': tensor(0.9488),
'lr': 0.00020771117314050008}

87%|      | 174/200 [2:07:50<18:54, 43.62s/it]

```

```

Val Loss: tensor(0.0022, device='cuda:0')
CNN Validation Score: tensor(0.9437)
Saving
{'train_loss': tensor(0.0020, device='cuda:0'), 'train_f1': tensor(0.9490),
'lr': 6.744225016419463e-05}

88%|      | 175/200 [2:08:28<17:30, 42.02s/it]

Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9440)
Saving
{'train_loss': tensor(0.0017, device='cuda:0'), 'train_f1': tensor(0.9493),
'lr': 1e-06}

88%|      | 176/200 [2:09:06<16:21, 40.89s/it]

Val Loss: tensor(0.0019, device='cuda:0')
CNN Validation Score: tensor(0.9442)
Saving
{'train_loss': tensor(0.0028, device='cuda:0'), 'train_f1': tensor(0.9495),
'lr': 6.744225016377972e-05}

88%|      | 177/200 [2:09:45<15:24, 40.20s/it]

Val Loss: tensor(0.0017, device='cuda:0')
CNN Validation Score: tensor(0.9443)
Saving
{'train_loss': tensor(0.0020, device='cuda:0'), 'train_f1': tensor(0.9497),
'lr': 0.00020771117313896385}

89%|      | 178/200 [2:10:24<14:34, 39.73s/it]

Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9446)
Saving
{'train_loss': tensor(0.0017, device='cuda:0'), 'train_f1': tensor(0.9500),
'lr': 0.0002971273994210919}

90%|      | 179/200 [2:11:02<13:43, 39.23s/it]

Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9448)
Saving
{'train_loss': tensor(0.0021, device='cuda:0'), 'train_f1': tensor(0.9502),
'lr': 0.00025621246378823825}

90%|      | 180/200 [2:11:40<13:00, 39.05s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9450)
Saving
{'train_loss': tensor(0.0018, device='cuda:0'), 'train_f1': tensor(0.9504),
'lr': 0.00012133399685905992}

```

```

90%|      | 181/200 [2:12:19<12:17, 38.84s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9453)
Saving
{'train_loss': tensor(0.0019, device='cuda:0'), 'train_f1': tensor(0.9506),
'lr': 1.2380009889627376e-05}

91%|      | 182/200 [2:12:57<11:37, 38.77s/it]
Val Loss: tensor(0.0010, device='cuda:0')
CNN Validation Score: tensor(0.9455)
Saving
{'train_loss': tensor(0.0014, device='cuda:0'), 'train_f1': tensor(0.9509),
'lr': 2.6195292960657258e-05}

92%|      | 183/200 [2:13:35<10:55, 38.55s/it]
Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9457)
Saving
{'train_loss': tensor(0.0015, device='cuda:0'), 'train_f1': tensor(0.9511),
'lr': 0.00015049999999948346}

92%|      | 184/200 [2:14:14<10:15, 38.46s/it]
Val Loss: tensor(0.0023, device='cuda:0')
CNN Validation Score: tensor(0.9459)
Saving
{'train_loss': tensor(0.0021, device='cuda:0'), 'train_f1': tensor(0.9513),
'lr': 0.00027480470703834613}

92%|      | 185/200 [2:14:52<09:38, 38.56s/it]
Val Loss: tensor(0.0012, device='cuda:0')
CNN Validation Score: tensor(0.9461)
Saving
{'train_loss': tensor(0.0027, device='cuda:0'), 'train_f1': tensor(0.9515),
'lr': 0.00028861999010950827}

93%|      | 186/200 [2:15:31<08:59, 38.50s/it]
Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9463)
Saving
{'train_loss': tensor(0.0018, device='cuda:0'), 'train_f1': tensor(0.9517),
'lr': 0.0001796660031410522}

94%|      | 187/200 [2:16:09<08:19, 38.43s/it]
Val Loss: tensor(0.0009, device='cuda:0')
CNN Validation Score: tensor(0.9465)
Saving

```

```

{'train_loss': tensor(0.0021, device='cuda:0'), 'train_f1': tensor(0.9519),
'lr': 4.478753621259585e-05}

94%|      | 188/200 [2:16:47<07:40, 38.41s/it]

Val Loss: tensor(0.0010, device='cuda:0')
CNN Validation Score: tensor(0.9467)
Saving
{'train_loss': tensor(0.0014, device='cuda:0'), 'train_f1': tensor(0.9521),
'lr': 3.87260057971705e-06}

94%|      | 189/200 [2:17:26<07:01, 38.31s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9470)
Saving
{'train_loss': tensor(0.0026, device='cuda:0'), 'train_f1': tensor(0.9523),
'lr': 9.328882686215698e-05}

95%|      | 190/200 [2:18:04<06:23, 38.32s/it]

Val Loss: tensor(0.0015, device='cuda:0')
CNN Validation Score: tensor(0.9471)
Saving
{'train_loss': tensor(0.0015, device='cuda:0'), 'train_f1': tensor(0.9525),
'lr': 0.00023355774983848816}

96%|      | 191/200 [2:18:42<05:45, 38.33s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9473)
Saving
{'train_loss': tensor(0.0016, device='cuda:0'), 'train_f1': tensor(0.9527),
'lr': 0.000300000000000027487}

96%|      | 192/200 [2:19:21<05:06, 38.34s/it]

Val Loss: tensor(0.0011, device='cuda:0')
CNN Validation Score: tensor(0.9475)
Saving
{'train_loss': tensor(0.0010, device='cuda:0'), 'train_f1': tensor(0.9529),
'lr': 0.000233557749838597}

96%|      | 193/200 [2:19:59<04:27, 38.27s/it]

Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9477)
Saving
{'train_loss': tensor(0.0020, device='cuda:0'), 'train_f1': tensor(0.9531),
'lr': 9.328882686226352e-05}

97%|      | 194/200 [2:20:37<03:49, 38.25s/it]

Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9479)

```

```

Saving
{'train_loss': tensor(0.0016, device='cuda:0'), 'train_f1': tensor(0.9533),
'lr': 3.87260057973363e-06}

98%|      | 195/200 [2:21:16<03:12, 38.52s/it]

Val Loss: tensor(0.0014, device='cuda:0')
CNN Validation Score: tensor(0.9481)
Saving
{'train_loss': tensor(0.0018, device='cuda:0'), 'train_f1': tensor(0.9535),
'lr': 4.478753621260576e-05}

98%|      | 196/200 [2:21:54<02:33, 38.42s/it]

Val Loss: tensor(0.0016, device='cuda:0')
CNN Validation Score: tensor(0.9482)
Saving
{'train_loss': tensor(0.0022, device='cuda:0'), 'train_f1': tensor(0.9537),
'lr': 0.00017966600314179158}

98%|      | 197/200 [2:22:33<01:55, 38.51s/it]

Val Loss: tensor(0.0008, device='cuda:0')
CNN Validation Score: tensor(0.9485)
Saving
{'train_loss': tensor(0.0020, device='cuda:0'), 'train_f1': tensor(0.9539),
'lr': 0.00028861999011127725}

99%|      | 198/200 [2:23:11<01:16, 38.48s/it]

Val Loss: tensor(0.0014, device='cuda:0')
CNN Validation Score: tensor(0.9486)
Saving
{'train_loss': tensor(0.0016, device='cuda:0'), 'train_f1': tensor(0.9541),
'lr': 0.000274804707040154}

100%|     | 199/200 [2:23:49<00:38, 38.36s/it]

Val Loss: tensor(0.0018, device='cuda:0')
CNN Validation Score: tensor(0.9488)
Saving
{'train_loss': tensor(0.0011, device='cuda:0'), 'train_f1': tensor(0.9543),
'lr': 0.000150500000000055294}

100%|     | 200/200 [2:24:28<00:00, 43.34s/it]

Val Loss: tensor(0.0013, device='cuda:0')
CNN Validation Score: tensor(0.9490)
Saving

```

## 3.5 Results

### 3.5.1 Performance Summary

#### Autoimmune Diseases Biopsy Slides Dataset

- CASS Resnet-50
- F1 Score: 0.8621
- CASS ViT B/16 g
- F1 Score: 0.8781

#### Dermofit Dataset

- CASS Resnet-50
- F1 Score: 0.7112
- CASS ViT B/16 g
- F1 Score: 0.6675

#### Brain Tumor MRI Dataset

- CASS Resnet-50
- F1 Score: 0.9490
- CASS ViT B/16 g
- F1 Score: 0.9211

### 3.5.2 Analysis

- **Replication:** The replication of the study yielded results that are comparable to those presented in the original paper, affirming the robustness and reliability of the CASS model.
- **Performance Across Datasets:** The model’s effectiveness is highlighted across various datasets: in autoimmune disease biopsy analysis, it demonstrated improved F1 scores with 100% labeled data. In the Dermofit dataset, CASS outshone both supervised and other self-supervised methods, showcasing its proficiency in handling diverse skin lesion types. For the Brain MRI Classification dataset, the model showed a notable improvement in bringing the performance of CNNs and Transformers closer. In the challenging ISIC-2019 dataset, known for class imbalances and inconsistent images, CASS again proved superior, especially in scenarios with limited labeled data.
- **Model’s Strengths:** These results across different medical imaging datasets emphasize CASS’s adaptability to varied image characteristics and its robustness in scenarios with limited labeled data.

### 3.5.3 Conclusion

- The CASS model stands out as a significant advancement in the field of medical image analysis. Its ability to maintain high accuracy and efficiency across diverse and challenging datasets positions it as a powerful tool for AI-driven medical diagnostics, especially in conditions with sparse data availability.

#### 3.5.4 Future Plan: Ablation Studies

- **Objective:** These studies aim to identify the optimal training configurations for the CASS model, ensuring it is well-tuned for diverse medical imaging tasks.
- **Focus Areas:**
- **Training Epochs:** Assessing the impact of varying the number of epochs on the model's accuracy and learning efficiency.
- **Optimizers:** Examining how different optimizer choices influence the model's performance.
- **Batch Size:** Understanding the effects of batch size variations on the model's training dynamics.
- **Encoder Size:** Exploring the influence of different encoder sizes on the model's capability to process and learn from medical images.

[ ]: