

# Chem 277B - Fall 2024 - Homework 6

## CNN - Image Processing Using an Encoder-Decoder Structure

*Submit this notebook to bCourses to receive a credit for this assignment.*

due: **Nov 18th 2024**

**Please upload both, the .ipynb file and the corresponding .pdf**

## 120 Points Total

### Problem

Your task is building an anomaly detection model to identify parasitized cells among images of infected and uninfected cells using a deep learning autoencoder structure (see **tutorial 7** at *bCourses*). The goal is to develop a system that detects anomalies based on reconstruction errors and density estimations of encoded features.

The learning goal of this homework assignment is to work on a real-life scenario with a realistic workflow. This task also serves a preparation for the Capstone Project concerning complexity and difficulty.

## Note: Optimize for Computational Efficiency

Autoencoders require significant computational resources, so consider the following strategies to reduce computation time while maintaining accuracy:

**Use Smaller Image Sizes:** Start with 64, 32, or even 16 pixels. Smaller images can drastically cut computational cost. Reduce Filters and Layers: Experiment with fewer filters in each layer and avoid building a very dense network. Aim for a trade-off between model complexity and compute efficiency.

**Skip the Final Conv2D Layer:** Instead of reconstructing at full resolution, consider using a smaller output size, which reduces the computation load without sacrificing much accuracy.

**Consider a Shallow Network:** A simpler encoder-decoder structure with fewer layers may still work effectively for anomaly detection.

**Preprocess Images:** Apply preprocessing techniques such as thresholding to simplify images before passing them through the autoencoder. Smarter preprocessing may reduce the need for a dense network. You may also make this a single channel image with smart pre processing.

These adjustments are intended to balance computational efficiency with detection accuracy. Use these hints to guide your network design and experiment with different configurations for the best trade-off.

Please avoid building a network that takes in **128x128** images without preprocessing. Training such a model will be infeasible with limited compute resources. Although smaller image sizes or simpler networks may not yield the highest possible accuracy, experimenting with these configurations is essential.

Grading will be based on the workflow and application of concepts—not purely on accuracy. Focus on understanding the process and applying techniques for effective anomaly detection with practical constraints. (Hint:- Opencv & KNN)

## Dataset

The dataset contains images of cell samples, both infected (parasitized) and uninfected (healthy), which can be downloaded from the National Library of Medicine: [Malaria Dataset](#)

## Guided Workflow and Hints:

```
In [141... import tensorflow as tf
import numpy as np
import random
import matplotlib.pyplot as plt
import os
from sklearn.model_selection import train_test_split
from PIL import Image
```

```
In [141... from tensorflow.keras.layers import Input, Dense, Conv2D, Conv2DTranspose, M

from tensorflow.keras.losses import mse
from tensorflow.keras import layers, models, backend as K
```

```
In [141... from tensorflow.keras import layers, models
```

## 1) Preprocessing and Data Loading

**Objective:** Load and preprocess images of healthy and infected cells.

**Hint:** Apply preprocessing techniques such as thresholding to simplify images before passing them through the autoencoder. Smarter preprocessing may reduce the need for a dense network.

You may also make this a single channel image with smart pre processing.

```
In [154... import os
import random
import numpy as np
from PIL import Image
from sklearn.model_selection import train_test_split

base_dir = "/Users/stemesghen/Downloads/HW6/cell_images"

def load_balanced_cell_images(base_dir, sample_size=None):
    data = []
    labels = []
    for folder, label in zip(["Parasitized", "Uninfected"], [1, 0]): # Para
        folder_path = os.path.join(base_dir, folder)
        files = [
            file_name
            for file_name in os.listdir(folder_path)
            if file_name.endswith(('.jpg', '.jpeg', '.png'))
        ]
        if sample_size is not None:
            files = random.sample(files, min(sample_size, len(files)))

        for file_name in files:
            img = Image.open(os.path.join(folder_path, file_name)).convert("
            data.append(np.array(img) / 255.0) # Normalize to [0, 1]
            labels.append(label)

    return np.array(data), np.array(labels)

sample_size = 200
labels = np.array(labels)

data, labels = load_balanced_cell_images(base_dir, sample_size)

# Filter only healthy cells for training
uninfected_samples = data[labels == 0] # Label 0 corresponds to healthy cel
parasitized_samples = data[labels == 1] # Parasitized (label = 1)

# Split healthy samples into training and test sets
x_train, x_test = train_test_split(uninfected_samples, test_size=0.2, random

# Reshape for the VAE
x_train = x_train.reshape(len(x_train), 28, 28, 1).astype("float32")
x_test = x_test.reshape(len(x_test), 28, 28, 1).astype("float32")

#reshaping
parasitized_samples = parasitized_samples.reshape(-1, 28, 28, 1).astype("flo
uninfected_samples = uninfected_samples.reshape(-1, 28, 28, 1).astype("float
```

```
In [141... #VAE - Autoencoder
```

```
latent_dim = 2 # Size of the latent space
```

## 2) Autoencoder Model Setup

**Objective:** Build an autoencoder with a bottleneck layer (see **tutorial 7**) to capture compressed representations of cell images.

**Hint:** Use *Conv2D*, *MaxPooling2D*, and *UpSampling2D* layers to build an encoder-decoder structure. Keep the bottleneck layer relatively small to capture key features in the latent space.

**Question:** Why is it important to use a small bottleneck layer in anomaly detection with autoencoders?

Answer: If the bottleneck layer is larger, then we can create overfitting and may classify more images as anomalous. It would be more difficult for the model to categorize between parasitized and uninfected images. Reconstruction errors would be lower in comparison to a smaller bottleneck.

```
In [141... # Encoder
inputs = layers.Input(shape=(28, 28, 1))
x = layers.Conv2D(32, (3, 3), activation="relu", padding="same")(inputs)
x = layers.MaxPooling2D((2, 2), padding="same")(x) # Downsampling
x = layers.Conv2D(64, (3, 3), activation="relu", padding="same")(x)
x = layers.MaxPooling2D((2, 2), padding="same")(x) # Downsampling
x = layers.Flatten()(x)
encoded = layers.Dense(latent_dim, activation="relu")(x)
```

```
In [141... # Sampling function
def sampling(args):
    z_mean, z_log_var = args
    epsilon = K.random_normal(shape=(K.shape(z_mean)[0], latent_dim))
    return z_mean + K.exp(0.5 * z_log_var) * epsilon

z_mean = layers.Dense(latent_dim, name="z_mean")(x)
z_log_var = layers.Dense(latent_dim, name="z_log_var")(x)

z = layers.Lambda(sampling, output_shape=(latent_dim,), name="z")([z_mean, z_log_var])
```

```
In [141... # Decoder
decoder_input = layers.Input(shape=(latent_dim,))
x = layers.Dense(7 * 7 * 64, activation="relu")(decoder_input)
x = layers.Reshape((7, 7, 64))(x)
x = layers.UpSampling2D((2, 2))(x) # Upsampling
x = layers.Conv2D(64, (3, 3), activation="relu", padding="same")(x)
x = layers.UpSampling2D((2, 2))(x) # Upsampling
x = layers.Conv2D(32, (3, 3), activation="relu", padding="same")(x)
outputs = layers.Conv2D(1, (3, 3), activation="sigmoid", padding="same")(x)
```

```
In [142... # Define encoder and decoder models
encoder = models.Model(inputs, [z_mean, z_log_var, z], name="encoder")
encoder.summary()
```

**Model: "encoder"**

Layer (type)	Output Shape	Param #	Connected to
input_layer_78 (InputLayer)	(None, 28, 28, 1)	0	-
conv2d_128 (Conv2D)	(None, 28, 28, 32)	320	input_layer_7
max_pooling2d_50 (MaxPooling2D)	(None, 14, 14, 32)	0	conv2d_128[0]
conv2d_129 (Conv2D)	(None, 14, 14, 64)	18,496	max_pooling2d
max_pooling2d_51 (MaxPooling2D)	(None, 7, 7, 64)	0	conv2d_129[0]
flatten_26 (Flatten)	(None, 3136)	0	max_pooling2d
z_mean (Dense)	(None, 2)	6,274	flatten_26[0]
z_log_var (Dense)	(None, 2)	6,274	flatten_26[0]
z (Lambda)	(None, 2)	0	z_mean[0][0], z_log_var[0][

**Total params:** 31,364 (122.52 KB)

**Trainable params:** 31,364 (122.52 KB)

**Non-trainable params:** 0 (0.00 B)

```
In [142... decoder = models.Model(decoder_input, outputs, name="decoder")
decoder.summary()
```

**Model: "decoder"**

Layer (type)	Output Shape	Par
input_layer_79 (InputLayer)	(None, 2)	
dense_52 (Dense)	(None, 3136)	9
reshape_28 (Reshape)	(None, 7, 7, 64)	
up_sampling2d_53 (UpSampling2D)	(None, 14, 14, 64)	
conv2d_130 (Conv2D)	(None, 14, 14, 64)	36
up_sampling2d_54 (UpSampling2D)	(None, 28, 28, 64)	
conv2d_131 (Conv2D)	(None, 28, 28, 32)	18
conv2d_132 (Conv2D)	(None, 28, 28, 1)	

**Total params:** 65,089 (254.25 KB)

**Trainable params:** 65,089 (254.25 KB)

**Non-trainable params:** 0 (0.00 B)

```
In [142... # VAE model - AUTOENCODER
outputs = decoder(encoder(inputs)[2])
vae = models.Model(inputs, outputs, name="vae")

vae.summary()
```

**Model: "vae"**

Layer (type)	Output Shape	Par
input_layer_78 (InputLayer)	(None, 28, 28, 1)	
encoder (Functional)	[(None, 2), (None, 2), (None, 2)]	31
decoder (Functional)	(None, 28, 28, 1)	65

**Total params:** 96,453 (376.77 KB)

**Trainable params:** 96,453 (376.77 KB)

**Non-trainable params:** 0 (0.00 B)

#### 4) Evaluation Using Reconstruction Error

**Objective:** Evaluate the reconstruction error on both healthy and parasitized cells.

**Hint:** Calculate the reconstruction error for each type of image. High errors for parasitized cells should indicate anomalies.

**Question:** Why might parasitized cells have a higher reconstruction error than healthy cells?





























Answer: Parasitized cells have a higher reconstruction error because they are harder to reconstruct through the VAE. The model was trained exclusively on healthy cells, and parasitized cells, having a more "de-normalized" shape compared to normal cells, do not effectively pass through the bottleneck layer. As a result, the VAE struggles to reconstruct these anomalous patterns. Parasitized cells can be considered anomalous because they reside in low-probability regions of the latent space, making them inherently harder for the VAE to encode and reconstruct accurately.

```
In [148... # VAE model
outputs = decoder(encoder(inputs)[2])
vae = models.Model(inputs, outputs, name="vae")

# Custom VAE loss function
def vae_loss(inputs, outputs):
    z_mean, z_log_var = encoder(inputs)[:2]
    reconstruction_loss = tf.reduce_sum(tf.keras.losses.binary_crossentropy(
    kl_loss = -0.5 * tf.reduce_sum(1 + z_log_var - tf.square(z_mean) - tf.ex
    total_loss = tf.reduce_mean(reconstruction_loss + kl_loss)
    return total_loss























# Compile the model with custom loss
vae.compile(optimizer="adam", loss=vae_loss)

In [150... # VAE fitting/train
vae.fit(
    x_train, x_train, # using x_train as both inputs and outputs since this
    epochs=50, # ----- try increase epochs per Dr.Hohle
    batch_size=128,
    validation_data=(x_test, x_test))
```

Epoch 1/50					
2/2		2s	87ms/step	- loss: 412.2075	- val_loss: 441.8917
Epoch 2/50					
2/2		0s	52ms/step	- loss: 411.4007	- val_loss: 447.7387
Epoch 3/50					
2/2		0s	50ms/step	- loss: 411.1259	- val_loss: 446.6953
Epoch 4/50					
2/2		0s	50ms/step	- loss: 410.1857	- val_loss: 445.1451
Epoch 5/50					
2/2		0s	50ms/step	- loss: 410.5730	- val_loss: 447.7159
Epoch 6/50					
2/2		0s	52ms/step	- loss: 409.6284	- val_loss: 451.9763
Epoch 7/50					
2/2		0s	50ms/step	- loss: 411.0636	- val_loss: 447.3581
Epoch 8/50					
2/2		0s	51ms/step	- loss: 410.4064	- val_loss: 447.3599
Epoch 9/50					
2/2		0s	50ms/step	- loss: 409.7982	- val_loss: 450.3166
Epoch 10/50					
2/2		0s	52ms/step	- loss: 410.3199	- val_loss: 450.4366
Epoch 11/50					
2/2		0s	50ms/step	- loss: 410.1905	- val_loss: 449.5333
Epoch 12/50					
2/2		0s	52ms/step	- loss: 410.1762	- val_loss: 448.6089
Epoch 13/50					
2/2		0s	49ms/step	- loss: 409.8383	- val_loss: 445.3858
Epoch 14/50					
2/2		0s	50ms/step	- loss: 410.0227	- val_loss: 447.3242
Epoch 15/50					
2/2		0s	50ms/step	- loss: 409.7674	- val_loss: 445.6519
Epoch 16/50					
2/2		0s	52ms/step	- loss: 409.6660	- val_loss: 448.7770
Epoch 17/50					
2/2		0s	50ms/step	- loss: 409.8935	- val_loss: 451.3096
Epoch 18/50					
2/2		0s	50ms/step	- loss: 410.8784	- val_loss: 452.4666
Epoch 19/50					
2/2		0s	50ms/step	- loss: 410.0218	- val_loss: 450.3943
Epoch 20/50					
2/2		0s	52ms/step	- loss: 410.1609	- val_loss: 452.9582
Epoch 21/50					
2/2		0s	51ms/step	- loss: 409.8729	- val_loss: 448.4160
Epoch 22/50					
2/2		0s	53ms/step	- loss: 410.9162	- val_loss: 446.6230
Epoch 23/50					
2/2		0s	50ms/step	- loss: 410.1003	- val_loss: 442.2557
Epoch 24/50					
2/2		0s	51ms/step	- loss: 410.2444	- val_loss: 448.7794
Epoch 25/50					
2/2		0s	50ms/step	- loss: 409.8417	- val_loss: 448.8539
Epoch 26/50					
2/2		0s	51ms/step	- loss: 409.8215	- val_loss: 447.9888
Epoch 27/50					
2/2		0s	72ms/step	- loss: 410.4017	- val_loss: 453.5186
Epoch 28/50					
2/2		0s	51ms/step	- loss: 410.2200	- val_loss: 451.9731



```

Epoch 29/50
2/2  0s 51ms/step - loss: 409.6290 - val_loss: 446.3042
Epoch 30/50
2/2  0s 52ms/step - loss: 410.8588 - val_loss: 447.4992
Epoch 31/50
2/2  0s 53ms/step - loss: 410.9018 - val_loss: 452.4814
Epoch 32/50
2/2  0s 51ms/step - loss: 409.9342 - val_loss: 447.0909
Epoch 33/50
2/2  0s 50ms/step - loss: 410.3923 - val_loss: 450.7503
Epoch 34/50
2/2  0s 51ms/step - loss: 410.8555 - val_loss: 448.8303
Epoch 35/50
2/2  0s 51ms/step - loss: 410.9175 - val_loss: 446.3905
Epoch 36/50
2/2  0s 57ms/step - loss: 409.9967 - val_loss: 449.5060
Epoch 37/50
2/2  0s 51ms/step - loss: 410.4698 - val_loss: 447.1417
Epoch 38/50
2/2  0s 51ms/step - loss: 410.0083 - val_loss: 451.1116
Epoch 39/50
2/2  0s 51ms/step - loss: 409.6778 - val_loss: 446.9775
Epoch 40/50
2/2  0s 52ms/step - loss: 409.7659 - val_loss: 449.6711
Epoch 41/50
2/2  0s 49ms/step - loss: 410.1628 - val_loss: 449.8413
Epoch 42/50
2/2  0s 72ms/step - loss: 410.1683 - val_loss: 449.6611
Epoch 43/50
2/2  0s 51ms/step - loss: 410.6559 - val_loss: 450.1601
Epoch 44/50
2/2  0s 51ms/step - loss: 409.9575 - val_loss: 450.7484
Epoch 45/50
2/2  0s 51ms/step - loss: 409.0106 - val_loss: 446.5891
Epoch 46/50
2/2  0s 54ms/step - loss: 409.5655 - val_loss: 445.7636
Epoch 47/50
2/2  0s 63ms/step - loss: 409.1613 - val_loss: 448.6289
Epoch 48/50
2/2  0s 66ms/step - loss: 409.8012 - val_loss: 449.9738
Epoch 49/50
2/2  0s 62ms/step - loss: 409.6128 - val_loss: 449.9853
Epoch 50/50
2/2  0s 59ms/step - loss: 410.0572 - val_loss: 452.9315

```

Out[150... <keras.src.callbacks.history.History at 0x34c378190>

```
In [150... print(inputs.shape)
```

(None, 28, 28, 1)

```
In [150... print("x_train shape:", x_train.shape)
print("x_test shape:", x_test.shape)
print("VAE input shape:", vae.input.shape)
print("VAE output shape:", vae.output.shape)
```

```
x_train shape: (160, 28, 28, 1)
x_test shape: (40, 28, 28, 1)
VAE input shape: (None, 28, 28, 1)
VAE output shape: (None, 28, 28, 1)
```

```
In [150... print(parasitized_samples.shape)
           print(uninfected_samples.shape)
```

```
(200, 28, 28, 1)
```

```
(200, 28, 28, 1)
```

```
In [150... #run data samples through VAE
```

```
uninfected_reconstruct_vae = vae.predict(uninfected_samples)# run uninfected
parasitized_reconstruct_vae= vae.predict(parasitized_samples)# run parasitized
```

```
print(uninfected_reconstruct_vae.shape)
print(parasitized_reconstruct_vae.shape)
```

```
7/7 ————— 0s 12ms/step
```

```
7/7 ————— 0s 9ms/step
```

```
(200, 28, 28, 1)
```

```
(200, 28, 28, 1)
```

```
In [150... uninfected_reconstruct_vae = vae.predict(uninfected_samples)
           print("Reconstructed uninfected samples shape:", uninfected_reconstruct_vae.
```

```
7/7 ————— 0s 9ms/step
```

```
Reconstructed uninfected samples shape: (200, 28, 28, 1)
```

```
In [151... # calculate reconstruction error (margin of error from original images (before and after VAE))
def calculate_reconstruction_error(before_vae, after_vae):
    errors = np.mean(np.square(before_vae - after_vae), axis=(1, 2, 3))
    return errors
```

```
In [151... # calculate reconstruction errors applied to data
uninfected_reconstruction_errors = calculate_reconstruction_error(uninfected_reconstruct_vae, uninfected_samples)
average_uninfected_reconstruction_errors = np.mean(uninfected_reconstruction_errors)
parasitized_reconstruction_errors = calculate_reconstruction_error(parasitized_reconstruct_vae, parasitized_samples)
average_parasitized_reconstruction_errors = np.mean(parasitized_reconstruction_errors)

print (f'''Mean Calculated Reconstruction Error
        Parasitized: {average_parasitized_reconstruction_errors}
        Uninfected:  {average_uninfected_reconstruction_errors}
        ''')
```

```
Mean Calculated Reconstruction Error
```

```
Parasitized: 0.03310062736272812
```

```
Uninfected:  0.011748772114515305
```

```
In [151... std_uninfected_reconstruction_errors = np.std(uninfected_reconstruction_errors)
std_parasitized_reconstruction_errors = np.std(parasitized_reconstruction_errors)
```

```
print (f'''Standard Deviation Calculated Reconstruction Error
        Parasitized: {std_parasitized_reconstruction_errors}
        Uninfected:   {std_uninfected_reconstruction_errors}
    ''')
```

```
Standard Deviation Calculated Reconstruction Error
Parasitized: 0.024932581931352615
Uninfected:  0.012165304273366928
```

## 5) Latent Space Representation and Density Calculation

**Objective:** Extract the compressed (latent) representations of the healthy cells from the bottleneck layer and calculate a density threshold for anomaly detection.

**Hint:** Use *KernelDensity* from *sklearn.neighbors* to fit a density model on the compressed representations of healthy cells, then set a threshold for density values.

**Question:** What role does density estimation play in enhancing anomaly detection beyond reconstruction error alone?

Having a second form of estimation, such as density estimation, can improve the accuracy of classifying images as parasitized vs. uninfected. Areas of low density in the latent space can be considered anomalous, as parasitized cells are more likely to fall outside the high-density regions where uninfected cells are concentrated.

```
In [151]: from sklearn.neighbors import KernelDensity
          np.random.seed(42)
          latent_representations = encoder.predict(uninfected_samples)[2]
```

7/7 ————— 0s 3ms/step

```
In [151]: kde = KernelDensity(kernel="gaussian", bandwidth=0.1)
          kde.fit(latent_representations)
```

```
Out[151]: KernelDensity
          KernelDensity(bandwidth=0.1)
```

```
In [151]: #Density
          log_density_scores = kde.score_samples(latent_representations)
          density_scores = np.exp(log_density_scores)

          #Threshold
          threshold = np.quantile(density_scores, 0.02) #Precentile = 2
```

```
for density in density_scores:
    if density >= threshold:
        print(f"Density Score {density:.4f} is >= threshold ({threshold:.4f})
    else:
        print(f"Density Score {density:.4f} is <= threshold ({threshold:.4f})
```

[illegible]

[illegible]

[illegible]

```

Density Score 0.1293 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0866 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.2110 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.4887 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1788 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0796 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1966 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.3377 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1001 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0856 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1888 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0796 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.3157 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.3958 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1419 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0841 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.5298 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0931 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1001 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.4860 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.3354 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0896 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1047 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1667 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.4365 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0867 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.0897 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1379 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.4454 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.3680 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.2465 is >= threshold (0.0796): Healthy (Uninfected)
Density Score 0.1192 is >= threshold (0.0796): Healthy (Uninfected)

```

```

In [151]: print("Density scores (uninfected cells):", density_scores[:5])
          print("Density threshold:", threshold)

```

```

Density scores (uninfected cells): [0.30237699 0.29957937 0.11229144 0.30780
57 0.24338546]
Density threshold: 0.07957747154595692

```

```

In [151]: #Anomaly

          anomalies = density_scores < threshold
          print("Anomaly in uninfected cells:", np.sum(anomalies))

          idx = np.where(density_scores <= threshold)
          values = uninfected_samples[idx]

```

```
Anomaly in uninfected cells: 4
```

## 6) Threshold Setting and Anomaly Detection

**Objective:** Define thresholds for density and reconstruction error to classify new images as healthy or parasitized.

**Hint:** Test different threshold values based on mean and standard deviation of errors in both healthy and parasitized cells.



**Question:** How would changing the density threshold or reconstruction error threshold affect the model's performance?

High: When I increase the thresholds for either density or reconstruction, more cells were considered uninfected.

Low: When I changed the density or reconstruction threshold by making it lower, there were more anomaly cells.

Higher the threshold then the more cells will be classified as "uninfected"

Lower the threshold then the more cells will be classified as "anomalous"

So it is important to try to find the balance between the thresholds to maximize accuracy of image classification.

```
In [151]... mean_density = np.mean(density_scores)
          std_density = np.std(density_scores)
```

```
In [151]... #K reconstruction = 1.5 and k density set at 1.5
          threshold_reconstruction_error = average_uninfected_reconstruction_errors +
          threshold_density = mean_density - 1.5 * std_density

          print(f"Threshold Density: {threshold_density:.4f}")
          print(f"Threshold Reconstruction Error: {threshold_reconstruction_error:.4f}")
```

Threshold Density: 0.0073

Threshold Reconstruction Error: 0.0300

```
In [154]... for error in parasitized_reconstruction_errors:
          if error <= threshold_reconstruction_error:
              print(f"Reconstruction Error {error:.4f} is <= to the threshold ({th
          else:
              print(f"Reconstruction Error {error:.4f} > than the threshold ({thre
```

Reconstruction Error 0.0160 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0236 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0776 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0376 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0402 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0338 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0114 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0210 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0162 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0199 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0180 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.1395 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0284 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0264 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0270 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0113 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0187 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0471 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0697 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0442 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0149 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0250 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0865 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0703 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0638 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0491 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0239 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0108 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0285 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0220 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0145 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0119 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0065 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0390 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0211 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0099 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0125 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0149 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0143 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0247 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0257 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0203 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0849 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0881 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0233 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0268 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0385 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0102 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.1120 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0730 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0552 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0220 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0109 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0161 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0188 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0389 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0091 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0192 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0255 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0137 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0047 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0061 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0632 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0246 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0663 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0074 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0207 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0281 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0114 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0341 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0314 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0400 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0296 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0182 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0318 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0324 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0356 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0184 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0275 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0914 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0223 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0122 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0161 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0055 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.1055 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0521 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0885 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0103 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0540 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0340 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0185 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0174 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0809 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0108 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0480 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0861 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0143 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0210 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0258 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0235 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0537 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0289 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0417 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0458 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0192 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.1005 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0319 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0155 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0197 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0213 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0560 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0144 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0176 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0456 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0049 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.1088 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0079 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0136 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0081 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0102 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0090 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0833 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0138 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0220 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0408 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0165 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0588 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0710 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0083 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0158 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0283 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0238 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0228 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0136 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0811 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0388 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0292 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0500 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0561 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0065 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0323 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0345 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0111 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0355 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0152 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0124 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0548 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0201 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0748 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0274 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0317 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0206 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0439 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0145 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0188 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0205 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0292 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0099 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0214 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0672 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0367 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0510 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0208 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0763 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0151 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0223 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0381 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0397 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0379 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0196 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0264 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0468 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0198 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0454 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0393 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0527 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0160 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0124 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0432 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.1160 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0193 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0154 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0443 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0125 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0237 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0186 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0137 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0141 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0186 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0324 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0327 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0307 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0138 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0101 is <= to the threshold (0.0300): Healthy (Uninfected)

Reconstruction Error 0.0588 > than the threshold (0.0300): Anomalous (Parasitized)

Reconstruction Error 0.0761 > than the threshold (0.0300): Anomalous (Parasitized)



Reconstruction Error 0.0109 is <= to the threshold (0.0300): Healthy (Uninfected)  
Reconstruction Error 0.0156 is <= to the threshold (0.0300): Healthy (Uninfected)  
Reconstruction Error 0.0154 is <= to the threshold (0.0300): Healthy (Uninfected)  
Reconstruction Error 0.0245 is <= to the threshold (0.0300): Healthy (Uninfected)

```
In [154... for error in uninfected_reconstruction_errors:
            if error <= threshold_reconstruction_error:
                print(f"Reconstruction Error for Healthy (Uninfected): {error:.4f}")
            else:
                print(f"Reconstruction Error for Anomalous (Parasitized): {error:.4f}")
```

[illegible]

Reconstruction Error for Healthy (Uninfected): 0.0056  
Reconstruction Error for Healthy (Uninfected): 0.0177  
Reconstruction Error for Healthy (Uninfected): 0.0202  
Reconstruction Error for Healthy (Uninfected): 0.0034  
Reconstruction Error for Healthy (Uninfected): 0.0100  
Reconstruction Error for Healthy (Uninfected): 0.0067  
Reconstruction Error for Healthy (Uninfected): 0.0053  
Reconstruction Error for Healthy (Uninfected): 0.0070  
Reconstruction Error for Healthy (Uninfected): 0.0099  
Reconstruction Error for Healthy (Uninfected): 0.0099  
Reconstruction Error for Healthy (Uninfected): 0.0093  
Reconstruction Error for Anomalous (Parasitized): 0.0538  
Reconstruction Error for Healthy (Uninfected): 0.0059  
Reconstruction Error for Healthy (Uninfected): 0.0137  
Reconstruction Error for Healthy (Uninfected): 0.0117  
Reconstruction Error for Healthy (Uninfected): 0.0215  
Reconstruction Error for Healthy (Uninfected): 0.0154  
Reconstruction Error for Healthy (Uninfected): 0.0177  
Reconstruction Error for Healthy (Uninfected): 0.0090  
Reconstruction Error for Anomalous (Parasitized): 0.1026  
Reconstruction Error for Healthy (Uninfected): 0.0030  
Reconstruction Error for Healthy (Uninfected): 0.0095  
Reconstruction Error for Healthy (Uninfected): 0.0168  
Reconstruction Error for Healthy (Uninfected): 0.0063  
Reconstruction Error for Healthy (Uninfected): 0.0059  
Reconstruction Error for Healthy (Uninfected): 0.0108  
Reconstruction Error for Healthy (Uninfected): 0.0290  
Reconstruction Error for Healthy (Uninfected): 0.0103  
Reconstruction Error for Healthy (Uninfected): 0.0172  
Reconstruction Error for Healthy (Uninfected): 0.0020  
Reconstruction Error for Healthy (Uninfected): 0.0059  
Reconstruction Error for Healthy (Uninfected): 0.0062  
Reconstruction Error for Healthy (Uninfected): 0.0081  
Reconstruction Error for Healthy (Uninfected): 0.0198  
Reconstruction Error for Healthy (Uninfected): 0.0073  
Reconstruction Error for Healthy (Uninfected): 0.0109  
Reconstruction Error for Healthy (Uninfected): 0.0058  
Reconstruction Error for Anomalous (Parasitized): 0.0598  
Reconstruction Error for Healthy (Uninfected): 0.0061  
Reconstruction Error for Anomalous (Parasitized): 0.0561  
Reconstruction Error for Healthy (Uninfected): 0.0038  
Reconstruction Error for Healthy (Uninfected): 0.0060  
Reconstruction Error for Healthy (Uninfected): 0.0071  
Reconstruction Error for Healthy (Uninfected): 0.0116  
Reconstruction Error for Healthy (Uninfected): 0.0089  
Reconstruction Error for Healthy (Uninfected): 0.0092  
Reconstruction Error for Healthy (Uninfected): 0.0084  
Reconstruction Error for Healthy (Uninfected): 0.0033  
Reconstruction Error for Healthy (Uninfected): 0.0123  
Reconstruction Error for Healthy (Uninfected): 0.0062  
Reconstruction Error for Healthy (Uninfected): 0.0135  
Reconstruction Error for Healthy (Uninfected): 0.0095  
Reconstruction Error for Healthy (Uninfected): 0.0112  
Reconstruction Error for Healthy (Uninfected): 0.0033  
Reconstruction Error for Healthy (Uninfected): 0.0124  
Reconstruction Error for Healthy (Uninfected): 0.0212

Reconstruction Error for Healthy (Uninfected): 0.0091  
Reconstruction Error for Healthy (Uninfected): 0.0110  
Reconstruction Error for Healthy (Uninfected): 0.0037  
Reconstruction Error for Healthy (Uninfected): 0.0155  
Reconstruction Error for Healthy (Uninfected): 0.0254  
Reconstruction Error for Healthy (Uninfected): 0.0084  
Reconstruction Error for Healthy (Uninfected): 0.0025  
Reconstruction Error for Healthy (Uninfected): 0.0043  
Reconstruction Error for Healthy (Uninfected): 0.0097  
Reconstruction Error for Healthy (Uninfected): 0.0128  
Reconstruction Error for Healthy (Uninfected): 0.0092  
Reconstruction Error for Healthy (Uninfected): 0.0124  
Reconstruction Error for Anomalous (Parasitized): 0.0731  
Reconstruction Error for Healthy (Uninfected): 0.0065  
Reconstruction Error for Healthy (Uninfected): 0.0029  
Reconstruction Error for Healthy (Uninfected): 0.0068  
Reconstruction Error for Healthy (Uninfected): 0.0090  
Reconstruction Error for Healthy (Uninfected): 0.0067  
Reconstruction Error for Healthy (Uninfected): 0.0031  
Reconstruction Error for Healthy (Uninfected): 0.0033  
Reconstruction Error for Healthy (Uninfected): 0.0261  
Reconstruction Error for Healthy (Uninfected): 0.0079  
Reconstruction Error for Healthy (Uninfected): 0.0052  
Reconstruction Error for Healthy (Uninfected): 0.0139  
Reconstruction Error for Healthy (Uninfected): 0.0058  
Reconstruction Error for Healthy (Uninfected): 0.0095  
Reconstruction Error for Healthy (Uninfected): 0.0076  
Reconstruction Error for Healthy (Uninfected): 0.0086  
Reconstruction Error for Healthy (Uninfected): 0.0045  
Reconstruction Error for Healthy (Uninfected): 0.0154  
Reconstruction Error for Healthy (Uninfected): 0.0082  
Reconstruction Error for Healthy (Uninfected): 0.0047  
Reconstruction Error for Healthy (Uninfected): 0.0049  
Reconstruction Error for Healthy (Uninfected): 0.0056  
Reconstruction Error for Healthy (Uninfected): 0.0084  
Reconstruction Error for Healthy (Uninfected): 0.0104  
Reconstruction Error for Healthy (Uninfected): 0.0099  
Reconstruction Error for Healthy (Uninfected): 0.0110  
Reconstruction Error for Healthy (Uninfected): 0.0157  
Reconstruction Error for Healthy (Uninfected): 0.0169  
Reconstruction Error for Anomalous (Parasitized): 0.0707  
Reconstruction Error for Healthy (Uninfected): 0.0122  
Reconstruction Error for Healthy (Uninfected): 0.0061  
Reconstruction Error for Healthy (Uninfected): 0.0038  
Reconstruction Error for Healthy (Uninfected): 0.0029  
Reconstruction Error for Healthy (Uninfected): 0.0123  
Reconstruction Error for Anomalous (Parasitized): 0.0332  
Reconstruction Error for Healthy (Uninfected): 0.0024  
Reconstruction Error for Healthy (Uninfected): 0.0100  
Reconstruction Error for Healthy (Uninfected): 0.0055  
Reconstruction Error for Healthy (Uninfected): 0.0101  
Reconstruction Error for Healthy (Uninfected): 0.0031  
Reconstruction Error for Healthy (Uninfected): 0.0256  
Reconstruction Error for Anomalous (Parasitized): 0.0364  
Reconstruction Error for Healthy (Uninfected): 0.0215  
Reconstruction Error for Healthy (Uninfected): 0.0054

```

Reconstruction Error for Healthy (Uninfected): 0.0070
Reconstruction Error for Healthy (Uninfected): 0.0126
Reconstruction Error for Healthy (Uninfected): 0.0071
Reconstruction Error for Healthy (Uninfected): 0.0057
Reconstruction Error for Healthy (Uninfected): 0.0150
Reconstruction Error for Healthy (Uninfected): 0.0103
Reconstruction Error for Healthy (Uninfected): 0.0084
Reconstruction Error for Healthy (Uninfected): 0.0059
Reconstruction Error for Healthy (Uninfected): 0.0047
Reconstruction Error for Anomalous (Parasitized): 0.0309
Reconstruction Error for Healthy (Uninfected): 0.0036
Reconstruction Error for Healthy (Uninfected): 0.0050
Reconstruction Error for Healthy (Uninfected): 0.0120
Reconstruction Error for Healthy (Uninfected): 0.0076
Reconstruction Error for Healthy (Uninfected): 0.0031
Reconstruction Error for Healthy (Uninfected): 0.0034
Reconstruction Error for Healthy (Uninfected): 0.0044
Reconstruction Error for Healthy (Uninfected): 0.0101
Reconstruction Error for Healthy (Uninfected): 0.0050
Reconstruction Error for Healthy (Uninfected): 0.0046
Reconstruction Error for Healthy (Uninfected): 0.0177
Reconstruction Error for Healthy (Uninfected): 0.0086
Reconstruction Error for Healthy (Uninfected): 0.0122
Reconstruction Error for Healthy (Uninfected): 0.0104
Reconstruction Error for Healthy (Uninfected): 0.0171
Reconstruction Error for Healthy (Uninfected): 0.0038
Reconstruction Error for Healthy (Uninfected): 0.0083
Reconstruction Error for Healthy (Uninfected): 0.0067
Reconstruction Error for Healthy (Uninfected): 0.0102
Reconstruction Error for Healthy (Uninfected): 0.0159
Reconstruction Error for Healthy (Uninfected): 0.0090
Reconstruction Error for Healthy (Uninfected): 0.0038

```

```
In [152... reconstruction_errors = np.concatenate([uninfected_reconstruction_errors, pa
```

```
In [152... # Combined Density with Errors for Uninfected and Parasitized Errors
for density, error in zip(density_scores, reconstruction_errors):
    if density >= threshold_density and error <= threshold_reconstruction_er
        print(f"Density Score = {density:.4f}, Reconstruction Error = {error
    else:
        print(f"Density Score = {density:.4f}, Reconstruction Error = {error
```

Density Score = 0.3024, Reconstruction Error = 0.0107: Healthy  
Density Score = 0.2996, Reconstruction Error = 0.0052: Healthy  
Density Score = 0.1123, Reconstruction Error = 0.0124: Healthy  
Density Score = 0.3078, Reconstruction Error = 0.0087: Healthy  
Density Score = 0.2434, Reconstruction Error = 0.0058: Healthy  
Density Score = 0.0798, Reconstruction Error = 0.0037: Healthy  
Density Score = 0.1294, Reconstruction Error = 0.0051: Healthy  
Density Score = 0.0851, Reconstruction Error = 0.0027: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0049: Healthy  
Density Score = 0.2209, Reconstruction Error = 0.0061: Healthy  
Density Score = 0.6008, Reconstruction Error = 0.0065: Healthy  
Density Score = 0.1960, Reconstruction Error = 0.0131: Healthy  
Density Score = 0.1665, Reconstruction Error = 0.0095: Healthy  
Density Score = 0.4148, Reconstruction Error = 0.0065: Healthy  
Density Score = 0.2441, Reconstruction Error = 0.0069: Healthy  
Density Score = 0.3601, Reconstruction Error = 0.0179: Healthy  
Density Score = 0.1708, Reconstruction Error = 0.0265: Healthy  
Density Score = 0.2488, Reconstruction Error = 0.0063: Healthy  
Density Score = 0.2249, Reconstruction Error = 0.0057: Healthy  
Density Score = 0.2741, Reconstruction Error = 0.0079: Healthy  
Density Score = 0.0805, Reconstruction Error = 0.0125: Healthy  
Density Score = 0.5170, Reconstruction Error = 0.0083: Healthy  
Density Score = 0.5641, Reconstruction Error = 0.0103: Healthy  
Density Score = 0.3050, Reconstruction Error = 0.0045: Healthy  
Density Score = 0.0828, Reconstruction Error = 0.0237: Healthy  
Density Score = 0.0851, Reconstruction Error = 0.0035: Healthy  
Density Score = 0.4374, Reconstruction Error = 0.0052: Healthy  
Density Score = 0.4576, Reconstruction Error = 0.0089: Healthy  
Density Score = 0.1220, Reconstruction Error = 0.0100: Healthy  
Density Score = 0.1978, Reconstruction Error = 0.0063: Healthy  
Density Score = 0.1728, Reconstruction Error = 0.0171: Healthy  
Density Score = 0.4586, Reconstruction Error = 0.0177: Healthy  
Density Score = 0.2227, Reconstruction Error = 0.0059: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0109: Healthy  
Density Score = 0.2208, Reconstruction Error = 0.0107: Healthy  
Density Score = 0.2731, Reconstruction Error = 0.0095: Healthy  
Density Score = 0.1175, Reconstruction Error = 0.0118: Healthy  
Density Score = 0.1992, Reconstruction Error = 0.0167: Healthy  
Density Score = 0.1866, Reconstruction Error = 0.0172: Healthy  
Density Score = 0.2610, Reconstruction Error = 0.0103: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0047: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0021: Healthy  
Density Score = 0.4399, Reconstruction Error = 0.0100: Healthy  
Density Score = 0.1706, Reconstruction Error = 0.0074: Healthy  
Density Score = 0.0992, Reconstruction Error = 0.0079: Healthy  
Density Score = 0.2054, Reconstruction Error = 0.0122: Healthy  
Density Score = 0.4601, Reconstruction Error = 0.0062: Healthy  
Density Score = 0.1396, Reconstruction Error = 0.0073: Healthy  
Density Score = 0.1146, Reconstruction Error = 0.0093: Healthy  
Density Score = 0.2651, Reconstruction Error = 0.0080: Healthy  
Density Score = 0.2295, Reconstruction Error = 0.0093: Healthy  
Density Score = 0.1124, Reconstruction Error = 0.0106: Healthy  
Density Score = 0.1355, Reconstruction Error = 0.0187: Healthy  
Density Score = 0.1541, Reconstruction Error = 0.0079: Healthy  
Density Score = 0.4584, Reconstruction Error = 0.0214: Healthy  
Density Score = 0.0884, Reconstruction Error = 0.0261: Healthy

Density Score = 0.3547, Reconstruction Error = 0.0056: Healthy  
Density Score = 0.2308, Reconstruction Error = 0.0177: Healthy  
Density Score = 0.1115, Reconstruction Error = 0.0202: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0034: Healthy  
Density Score = 0.4048, Reconstruction Error = 0.0100: Healthy  
Density Score = 0.0800, Reconstruction Error = 0.0067: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0053: Healthy  
Density Score = 0.5403, Reconstruction Error = 0.0070: Healthy  
Density Score = 0.2007, Reconstruction Error = 0.0099: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0099: Healthy  
Density Score = 0.1873, Reconstruction Error = 0.0093: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0538: Parasitized  
Density Score = 0.3490, Reconstruction Error = 0.0059: Healthy  
Density Score = 0.1560, Reconstruction Error = 0.0137: Healthy  
Density Score = 0.1399, Reconstruction Error = 0.0117: Healthy  
Density Score = 0.1572, Reconstruction Error = 0.0215: Healthy  
Density Score = 0.5107, Reconstruction Error = 0.0154: Healthy  
Density Score = 0.0884, Reconstruction Error = 0.0177: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0090: Healthy  
Density Score = 0.1373, Reconstruction Error = 0.1026: Parasitized  
Density Score = 0.0959, Reconstruction Error = 0.0030: Healthy  
Density Score = 0.2508, Reconstruction Error = 0.0095: Healthy  
Density Score = 0.3002, Reconstruction Error = 0.0168: Healthy  
Density Score = 0.1334, Reconstruction Error = 0.0063: Healthy  
Density Score = 0.1344, Reconstruction Error = 0.0059: Healthy  
Density Score = 0.6046, Reconstruction Error = 0.0108: Healthy  
Density Score = 0.2345, Reconstruction Error = 0.0290: Healthy  
Density Score = 0.0957, Reconstruction Error = 0.0103: Healthy  
Density Score = 0.0969, Reconstruction Error = 0.0172: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0020: Healthy  
Density Score = 0.2798, Reconstruction Error = 0.0059: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0062: Healthy  
Density Score = 0.0854, Reconstruction Error = 0.0081: Healthy  
Density Score = 0.1004, Reconstruction Error = 0.0198: Healthy  
Density Score = 0.4079, Reconstruction Error = 0.0073: Healthy  
Density Score = 0.2328, Reconstruction Error = 0.0109: Healthy  
Density Score = 0.0992, Reconstruction Error = 0.0058: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0598: Parasitized  
Density Score = 0.1373, Reconstruction Error = 0.0061: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0561: Parasitized  
Density Score = 0.1430, Reconstruction Error = 0.0038: Healthy  
Density Score = 0.0823, Reconstruction Error = 0.0060: Healthy  
Density Score = 0.3008, Reconstruction Error = 0.0071: Healthy  
Density Score = 0.2282, Reconstruction Error = 0.0116: Healthy  
Density Score = 0.5647, Reconstruction Error = 0.0089: Healthy  
Density Score = 0.1758, Reconstruction Error = 0.0092: Healthy  
Density Score = 0.1237, Reconstruction Error = 0.0084: Healthy  
Density Score = 0.3092, Reconstruction Error = 0.0033: Healthy  
Density Score = 0.1964, Reconstruction Error = 0.0123: Healthy  
Density Score = 0.1765, Reconstruction Error = 0.0062: Healthy  
Density Score = 0.3657, Reconstruction Error = 0.0135: Healthy  
Density Score = 0.1669, Reconstruction Error = 0.0095: Healthy  
Density Score = 0.3695, Reconstruction Error = 0.0112: Healthy  
Density Score = 0.0880, Reconstruction Error = 0.0033: Healthy  
Density Score = 0.1025, Reconstruction Error = 0.0124: Healthy  
Density Score = 0.1740, Reconstruction Error = 0.0212: Healthy

Density Score = 0.2259, Reconstruction Error = 0.0091: Healthy  
Density Score = 0.3390, Reconstruction Error = 0.0110: Healthy  
Density Score = 0.5410, Reconstruction Error = 0.0037: Healthy  
Density Score = 0.0800, Reconstruction Error = 0.0155: Healthy  
Density Score = 0.2265, Reconstruction Error = 0.0254: Healthy  
Density Score = 0.2250, Reconstruction Error = 0.0084: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0025: Healthy  
Density Score = 0.2970, Reconstruction Error = 0.0043: Healthy  
Density Score = 0.2400, Reconstruction Error = 0.0097: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0128: Healthy  
Density Score = 0.0837, Reconstruction Error = 0.0092: Healthy  
Density Score = 0.1205, Reconstruction Error = 0.0124: Healthy  
Density Score = 0.0867, Reconstruction Error = 0.0731: Parasitized  
Density Score = 0.3559, Reconstruction Error = 0.0065: Healthy  
Density Score = 0.3751, Reconstruction Error = 0.0029: Healthy  
Density Score = 0.1816, Reconstruction Error = 0.0068: Healthy  
Density Score = 0.1058, Reconstruction Error = 0.0090: Healthy  
Density Score = 0.3855, Reconstruction Error = 0.0067: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0031: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0033: Healthy  
Density Score = 0.1462, Reconstruction Error = 0.0261: Healthy  
Density Score = 0.1052, Reconstruction Error = 0.0079: Healthy  
Density Score = 0.0822, Reconstruction Error = 0.0052: Healthy  
Density Score = 0.1871, Reconstruction Error = 0.0139: Healthy  
Density Score = 0.2240, Reconstruction Error = 0.0058: Healthy  
Density Score = 0.1107, Reconstruction Error = 0.0095: Healthy  
Density Score = 0.4503, Reconstruction Error = 0.0076: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0086: Healthy  
Density Score = 0.1741, Reconstruction Error = 0.0045: Healthy  
Density Score = 0.4555, Reconstruction Error = 0.0154: Healthy  
Density Score = 0.0941, Reconstruction Error = 0.0082: Healthy  
Density Score = 0.2324, Reconstruction Error = 0.0047: Healthy  
Density Score = 0.5952, Reconstruction Error = 0.0049: Healthy  
Density Score = 0.0937, Reconstruction Error = 0.0056: Healthy  
Density Score = 0.2819, Reconstruction Error = 0.0084: Healthy  
Density Score = 0.2842, Reconstruction Error = 0.0104: Healthy  
Density Score = 0.1091, Reconstruction Error = 0.0099: Healthy  
Density Score = 0.1411, Reconstruction Error = 0.0110: Healthy  
Density Score = 0.2086, Reconstruction Error = 0.0157: Healthy  
Density Score = 0.4149, Reconstruction Error = 0.0169: Healthy  
Density Score = 0.1053, Reconstruction Error = 0.0707: Parasitized  
Density Score = 0.0926, Reconstruction Error = 0.0122: Healthy  
Density Score = 0.2990, Reconstruction Error = 0.0061: Healthy  
Density Score = 0.2992, Reconstruction Error = 0.0038: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0029: Healthy  
Density Score = 0.2543, Reconstruction Error = 0.0123: Healthy  
Density Score = 0.0798, Reconstruction Error = 0.0332: Parasitized  
Density Score = 0.0796, Reconstruction Error = 0.0024: Healthy  
Density Score = 0.0998, Reconstruction Error = 0.0100: Healthy  
Density Score = 0.1025, Reconstruction Error = 0.0055: Healthy  
Density Score = 0.1922, Reconstruction Error = 0.0101: Healthy  
Density Score = 0.0796, Reconstruction Error = 0.0031: Healthy  
Density Score = 0.1327, Reconstruction Error = 0.0256: Healthy  
Density Score = 0.1624, Reconstruction Error = 0.0364: Parasitized  
Density Score = 0.3710, Reconstruction Error = 0.0215: Healthy  
Density Score = 0.3057, Reconstruction Error = 0.0054: Healthy



```

Density Score = 0.1293, Reconstruction Error = 0.0070: Healthy
Density Score = 0.0866, Reconstruction Error = 0.0126: Healthy
Density Score = 0.2110, Reconstruction Error = 0.0071: Healthy
Density Score = 0.4887, Reconstruction Error = 0.0057: Healthy
Density Score = 0.1788, Reconstruction Error = 0.0150: Healthy
Density Score = 0.0796, Reconstruction Error = 0.0103: Healthy
Density Score = 0.1966, Reconstruction Error = 0.0084: Healthy
Density Score = 0.3377, Reconstruction Error = 0.0059: Healthy
Density Score = 0.1001, Reconstruction Error = 0.0047: Healthy
Density Score = 0.0856, Reconstruction Error = 0.0309: Parasitized
Density Score = 0.1888, Reconstruction Error = 0.0036: Healthy
Density Score = 0.0796, Reconstruction Error = 0.0050: Healthy
Density Score = 0.3157, Reconstruction Error = 0.0120: Healthy
Density Score = 0.3958, Reconstruction Error = 0.0076: Healthy
Density Score = 0.1419, Reconstruction Error = 0.0031: Healthy
Density Score = 0.0841, Reconstruction Error = 0.0034: Healthy
Density Score = 0.5298, Reconstruction Error = 0.0044: Healthy
Density Score = 0.0931, Reconstruction Error = 0.0101: Healthy
Density Score = 0.1001, Reconstruction Error = 0.0050: Healthy
Density Score = 0.4860, Reconstruction Error = 0.0046: Healthy
Density Score = 0.3354, Reconstruction Error = 0.0177: Healthy
Density Score = 0.0896, Reconstruction Error = 0.0086: Healthy
Density Score = 0.1047, Reconstruction Error = 0.0122: Healthy
Density Score = 0.1667, Reconstruction Error = 0.0104: Healthy
Density Score = 0.4365, Reconstruction Error = 0.0171: Healthy
Density Score = 0.0867, Reconstruction Error = 0.0038: Healthy
Density Score = 0.0897, Reconstruction Error = 0.0083: Healthy
Density Score = 0.1379, Reconstruction Error = 0.0067: Healthy
Density Score = 0.4454, Reconstruction Error = 0.0102: Healthy
Density Score = 0.3680, Reconstruction Error = 0.0159: Healthy
Density Score = 0.2465, Reconstruction Error = 0.0090: Healthy
Density Score = 0.1192, Reconstruction Error = 0.0038: Healthy

```

```

In [152... import matplotlib.pyplot as plt
import numpy as np

# Function to plot original and reconstructed images
def plot_reconstructions(original, reconstructed, num_images=5, title="Recon

    plt.figure(figsize=(10, 4))
    for i in range(num_images):
        # Plot original image
        plt.subplot(2, num_images, i + 1)
        plt.imshow(original[i].squeeze(), cmap='gray')
        plt.title("Original")
        plt.axis('off')

        # Plot reconstructed image
        plt.subplot(2, num_images, i + 1 + num_images)
        plt.imshow(reconstructed[i].squeeze(), cmap='gray')
        plt.title("Reconstructed")
        plt.axis('off')

    plt.suptitle(title, fontsize=16)
    plt.tight_layout()
    plt.show()

```

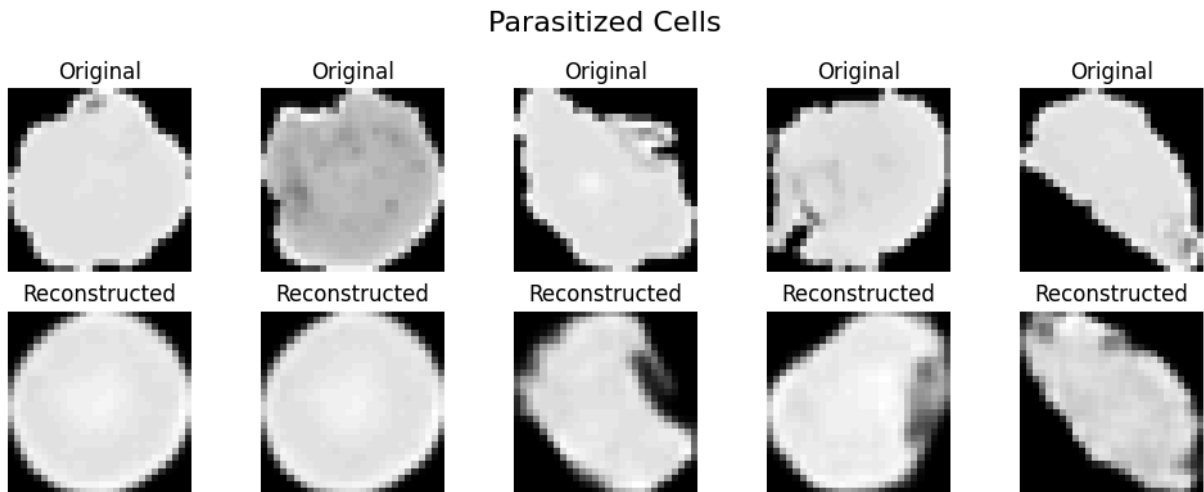
```

# Example usage with parasitized and uninfected samples
# Assuming `parasitized_samples` and `parasitized_reconstruct_vae` exist
print("Parasitized Cell Reconstructions")
plot_reconstructions(parasitized_samples, parasitized_reconstruct_vae, num_images)

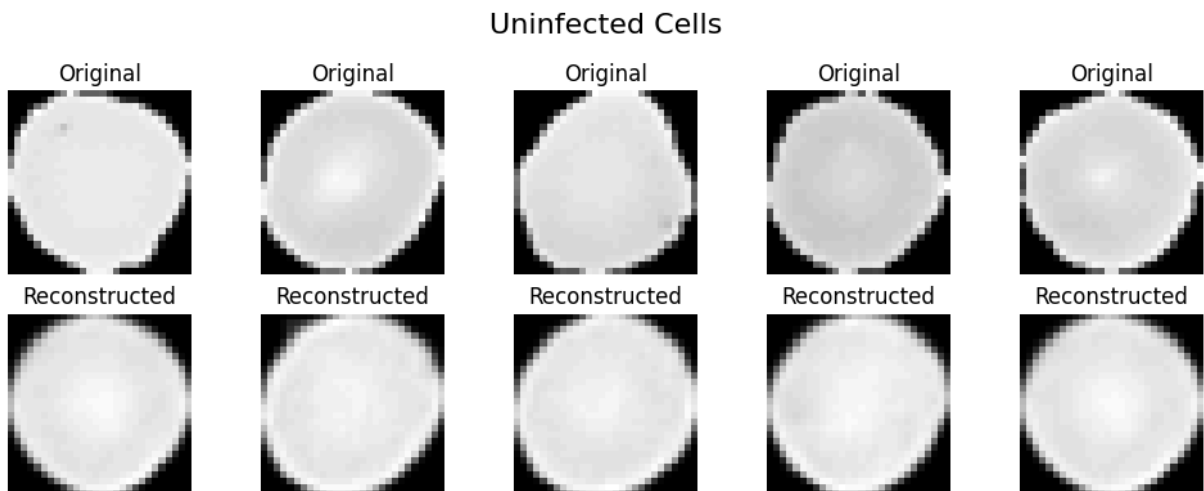
# Assuming `uninfected_samples` and `uninfected_reconstruct_vae` exist
print("Uninfected Cell Reconstructions")
plot_reconstructions(uninfected_samples, uninfected_reconstruct_vae, num_images)

```

#### Parasitized Cell Reconstructions



#### Uninfected Cell Reconstructions



```

In [155... def plot_reconstruction_errors(x_test, reconstructed_images, n=50):
    """
    Plots a histogram of reconstruction errors.
    """
    # Compute reconstruction errors
    errors = np.mean((x_test[:n] - reconstructed_images[:n])**2, axis=(1, 2))

    # Plot histogram of reconstruction errors
    plt.figure(figsize=(10, 5))
    plt.hist(errors, bins=15, color='green', alpha=0.7)
    plt.title("Distribution of Reconstruction Errors")
    plt.xlabel("Reconstruction Error")
    plt.ylabel("Frequency")

```

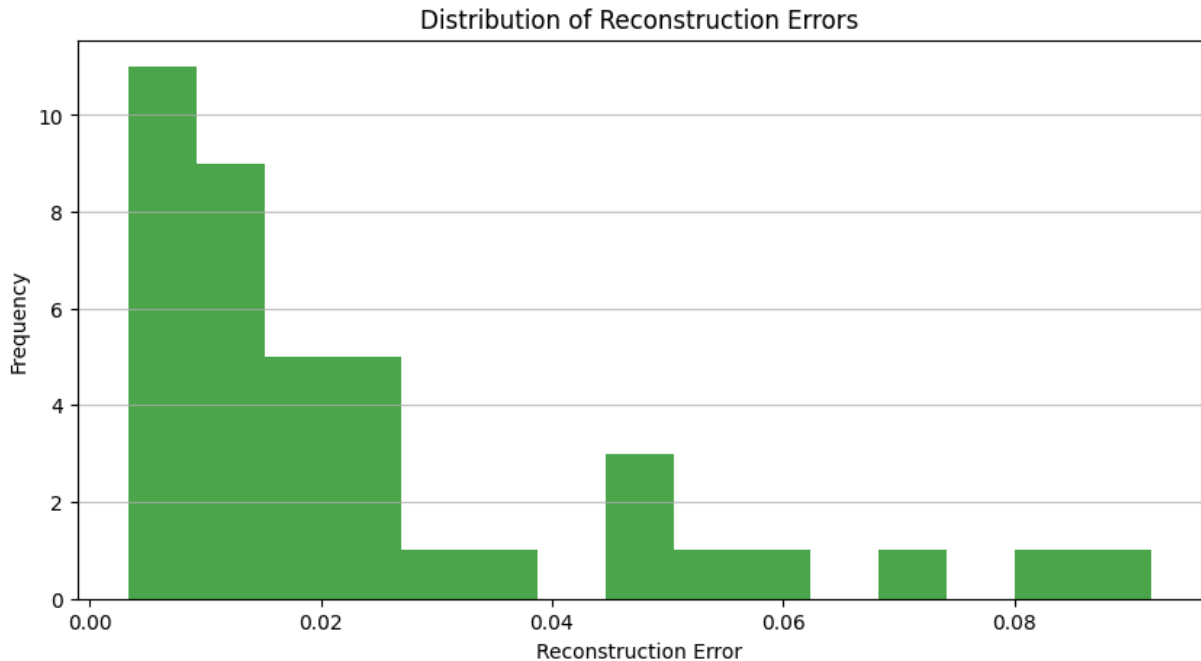
```
plt.grid(axis='y', alpha=0.75)
plt.show()

# Usage Example:
# Get reconstructed images
z_mean, _, _ = encoder.predict(x_test)
reconstructed_images = decoder.predict(z_mean)

# Plot the reconstruction errors
plot_reconstruction_errors(x_test, reconstructed_images, n=100)
```

2/2 ————— 0s 3ms/step

2/2 ————— 0s 4ms/step



## 7) Testing on New Images

**Objective:** Test the model's performance on a mix of new healthy and parasitized images.

**Hint:** Use the *check\_anomaly* function to predict whether new images are anomalies. Adjust threshold values if necessary.

**Question:** What can you conclude if a healthy cell image is incorrectly classified as parasitized?

As previously stated, if a healthy cell image is classified as parasitized, it could come from my threshold being set too low causing overfitting. It can also be from the image size, batch, and/or epoch size. Latent dimension can be too small. Image size may be too small - missing key features. Batch size - too large, can affecting performance epoch size - lower epoch can cause undertraining and lead to overfitting as well.

NOTES:

Anomaly: have high reconstruction error because the VAE could not construct it well. If there is a low density score then that means it doesn't belong to the normal region of the latent space

parasitized cells have anomalies (high reconstruction errors because they are harder to reconstruct. Thus, they are low probability in latent space).

Latent space has a Healthy cells (normal) vs. Parasitized cells (anomalies)

- Healthy are clustered around the center and represent high-probability regions
- Anomalies/Parasitized - spread far from the center and low probability regions.

```
In [152... ##### NEW CODE TO IMPLEMENT##
```

```
In [152... new_base_dir = "/Users/stemesghen/Downloads/HW6/cell_images"
new_data, new_labels = load_balanced_cell_images(new_base_dir, sample_size)
```

Loaded 200 images from 'Parasitized' folder.

Loaded 200 images from 'Uninfected' folder.

```
In [155... def check_anomaly_new_images(density_score, reconstruction_error, threshold_
    """
    Classify an image as healthy or parasitized based on density score and r
    """
    if density_score < threshold_density or reconstruction_error > threshold_
        return "Parasitized"
    else:
        return "Healthy"
```

```
In [155... # Test new images

new_data = new_data.reshape(-1, 28, 28, 1).astype("float32")

# NEW Latent Representations and Reconstruction ERRORS
latent_representations_new = encoder.predict(new_data)[2]
reconstructed_new_images = vae.predict(new_data)
new_reconstruction_errors = calculate_reconstruction_error(new_data, reconst

# NEW density scores
log_density_scores_new = kde.score_samples(latent_representations_new)
density_scores_new = np.exp(log_density_scores_new)

# Classify new images
predicted_labels = [
    0 if density_score >= threshold_density and reconstruction_error <= thre
    for density_score, reconstruction_error in zip(density_scores_new, new_r
]
```

```

# TRAIN NEW data
train_data, test_data, train_labels, test_labels = train_test_split(
    new_data, new_labels, test_size=0.2, random_state=42
)

# Evaluate model
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score

accuracy = accuracy_score(new_labels, predicted_labels)
precision = precision_score(new_labels, predicted_labels)
recall = recall_score(new_labels, predicted_labels)
f1 = f1_score(new_labels, predicted_labels)

print(f"Accuracy: {accuracy:.2f}")
print(f"Precision: {precision:.2f}")
print(f"Recall: {recall:.2f}")
print(f"F1 Score: {f1:.2f}")

```

13/13 ————— 0s 2ms/step

13/13 ————— 0s 8ms/step

Accuracy: 0.61

Precision: 0.65

Recall: 0.47

F1 Score: 0.54

## VISUALIZATION

In [155... *# Function to plot the latent space with labels*

```

def plot_latent_space(encoder, uninfected_samples, parasitized_samples):
    # Get latent space representations for both classes
    z_uninfected = encoder.predict(uninfected_samples)[2]
    z_parasitized = encoder.predict(parasitized_samples)[2]

    # Create labels for the two classes
    labels_uninfected = np.zeros(len(z_uninfected)) # Label 0 for uninfected
    labels_parasitized = np.ones(len(z_parasitized)) # Label 1 for parasitized

    # Combine data and labels
    z_combined = np.concatenate([z_uninfected, z_parasitized])
    labels_combined = np.concatenate([labels_uninfected, labels_parasitized])

    # Plot the latent space
    plt.figure(figsize=(10, 10))
    plt.scatter(z_combined[:, 0], z_combined[:, 1], c=labels_combined, cmap=
    plt.colorbar(label="Label (0: Uninfected, 1: Parasitized)")
    plt.xlabel("z[0]")
    plt.ylabel("z[1]")
    plt.title("2D Latent Space of VAE")
    plt.show()

# Plot the latent space
plot_latent_space(encoder, uninfected_samples, parasitized_samples)

# Function to generate and visualize images by tweaking the latent vector

```

```

def visualize_latent_images(decoder, n=10, range_min=-3, range_max=3):
    img_width, img_height, num_channels = 28, 28, 1
    figure = np.zeros((img_width * n, img_height * n))

    # Define grid of values in latent space
    grid_x = np.linspace(range_min, range_max, n)
    grid_y = np.linspace(range_min, range_max, n)[::-1]

    for i, yi in enumerate(grid_y):
        for j, xi in enumerate(grid_x):
            z_sample = np.array([[xi, yi]])
            x_decoded = decoder.predict(z_sample)
            image = x_decoded[0].reshape(img_width, img_height)
            figure[i * img_width: (i + 1) * img_width,
                  j * img_height: (j + 1) * img_height] = image

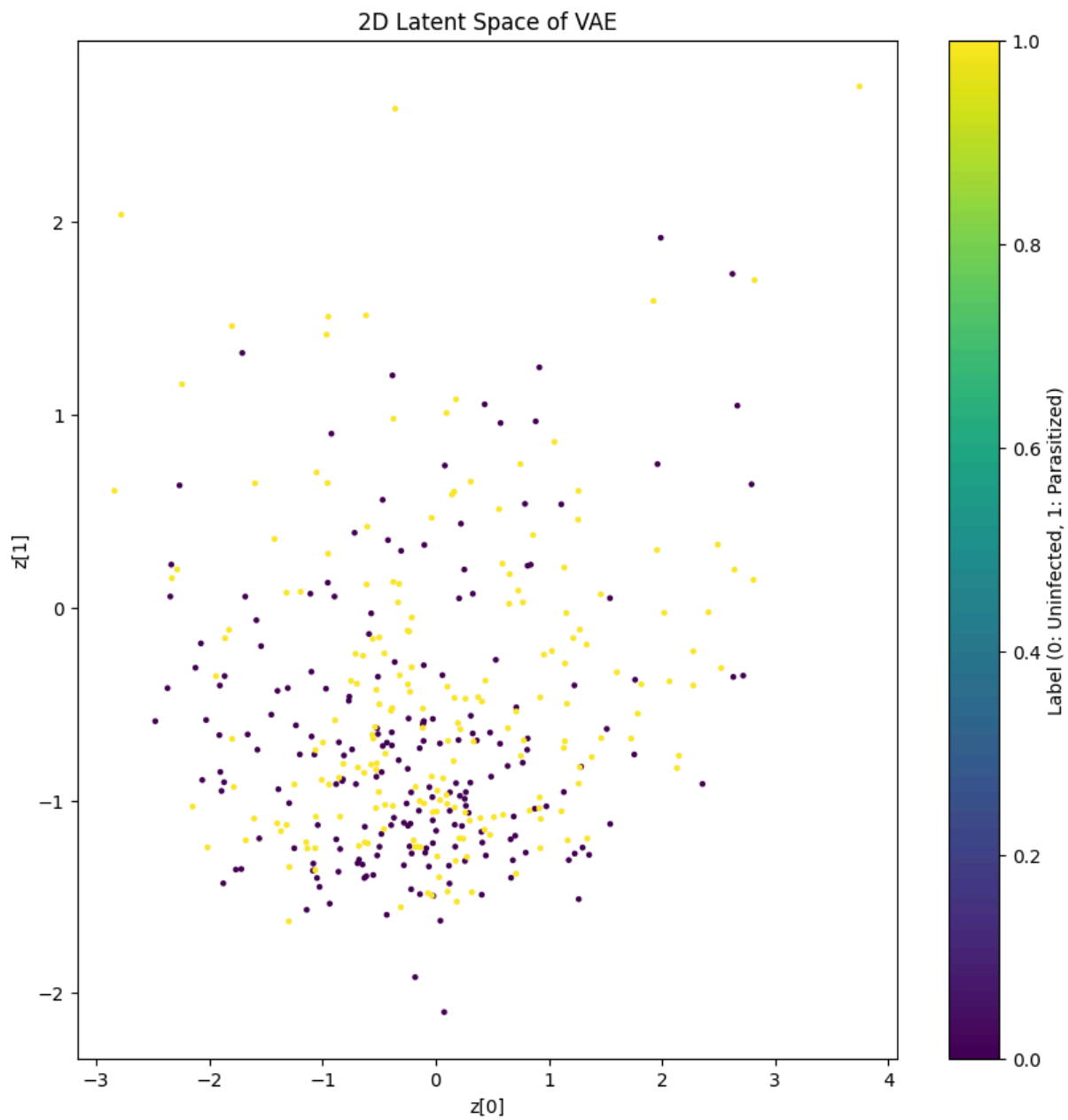
    plt.figure(figsize=(10, 10))
    plt.imshow(figure, cmap='gray')
    plt.axis('off')
    plt.title("Generated Images from Latent Space")
    plt.show()

# Visualize images generated from the latent space
visualize_latent_images(decoder)

```

7/7  0s 2ms/step

7/7  0s 2ms/step

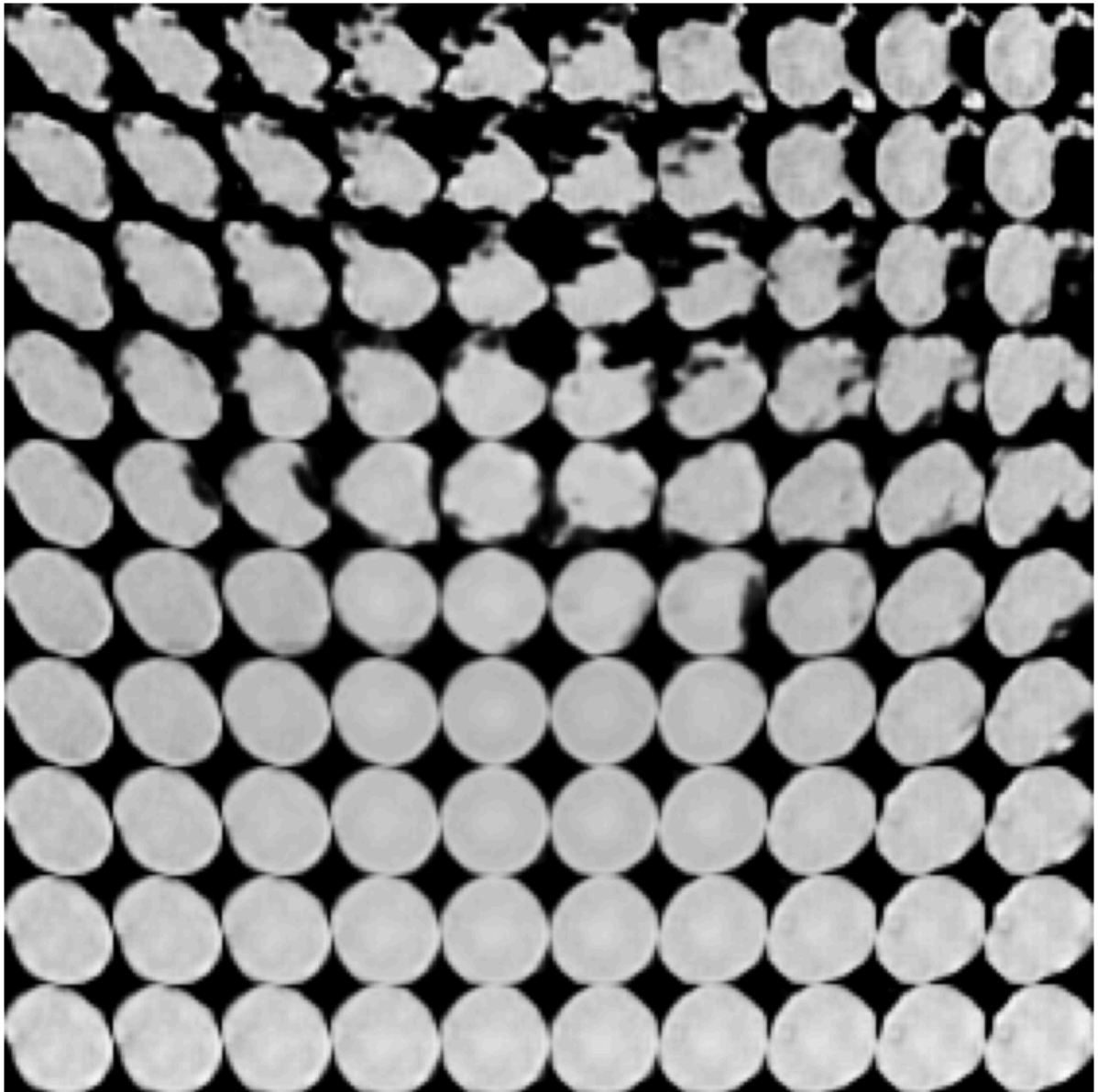


1/1	0s	11ms/step
1/1	0s	9ms/step
1/1	0s	8ms/step
1/1	0s	8ms/step
1/1	0s	10ms/step
1/1	0s	9ms/step
1/1	0s	8ms/step
1/1	0s	10ms/step
1/1	0s	12ms/step
1/1	0s	11ms/step
1/1	0s	8ms/step
1/1	0s	9ms/step
1/1	0s	11ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	10ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	8ms/step
1/1	0s	10ms/step
1/1	0s	10ms/step
1/1	0s	10ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	8ms/step
1/1	0s	10ms/step
1/1	0s	9ms/step
1/1	0s	10ms/step
1/1	0s	9ms/step
1/1	0s	10ms/step
1/1	0s	11ms/step
1/1	0s	9ms/step
1/1	0s	11ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	11ms/step
1/1	0s	9ms/step
1/1	0s	10ms/step
1/1	0s	11ms/step
1/1	0s	13ms/step
1/1	0s	9ms/step
1/1	0s	8ms/step
1/1	0s	9ms/step
1/1	0s	10ms/step
1/1	0s	10ms/step
1/1	0s	9ms/step
1/1	0s	11ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step
1/1	0s	11ms/step
1/1	0s	9ms/step
1/1	0s	9ms/step



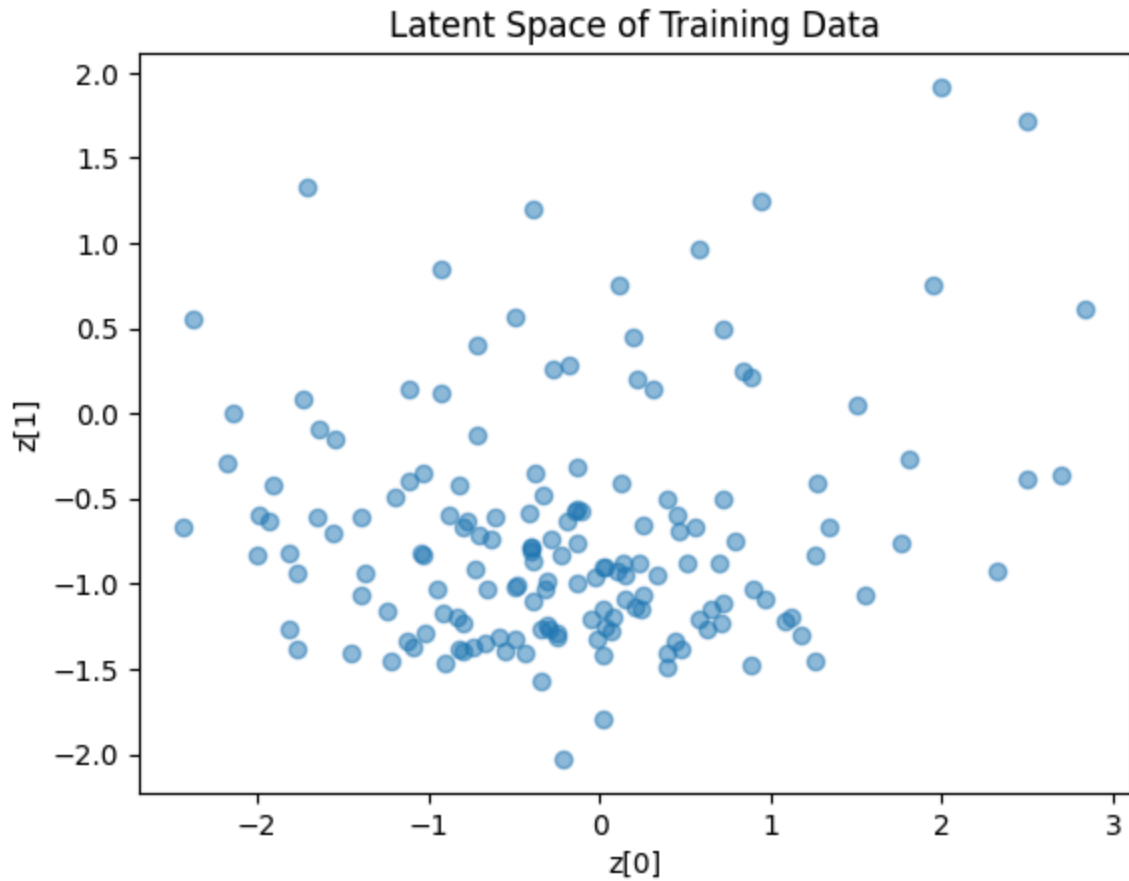
1/1	_____	0s	8ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	8ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	11ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	25ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	11ms/step
1/1	_____	0s	8ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	10ms/step
1/1	_____	0s	8ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	11ms/step
1/1	_____	0s	9ms/step
1/1	_____	0s	8ms/step
1/1	_____	0s	11ms/step
1/1	_____	0s	9ms/step

Generated Images from Latent Space



```
In [155... z_mean, _, _ = encoder.predict(x_train)
plt.scatter(z_mean[:, 0], z_mean[:, 1], alpha=0.5)
plt.xlabel("z[0]")
plt.ylabel("z[1]")
plt.title("Latent Space of Training Data")
plt.show()
```

5/5 ————— 0s 3ms/step



A	B	C
Section	Points	Criteria
1) Preprocessing and Data Loading	30	- Effective preprocessing of images using thresholding or other methods to simplify inputs. - Use of OpenCV and KNN if needed. - Single-channel optimization.
2) Autoencoder Model Setup	30	- Building an encoder-decoder structure with appropriate layers and filters. - Use of a small bottleneck layer to capture key features.
3) Model Training	5	- Training only on healthy images. - Use of mean_squared_error loss and ensuring model does not train on parasitized cells.
4) Evaluation Using Reconstruction Error	15	- Calculating reconstruction errors effectively for both healthy and parasitized cells. - Identification of high error for anomalies.
5) Latent Space Representation and Density	15	- Extracting and using latent representations from the bottleneck layer. - Effective use of KernelDensity to set density thresholds.
6) Threshold Setting and Anomaly Detection	10	- Thoughtful selection and testing of density and reconstruction error thresholds. - Balancing sensitivity and specificity in threshold setting.
7) Testing on New Images	5	- Testing model on new healthy and parasitized images. - Correct use of check_anomaly function to classify new images.
8) Documentation and Workflow	10	- Clear documentation of experiments, parameter choices, and results. - Insights on balancing accuracy with computational efficiency.