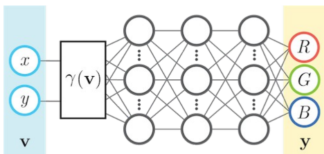
Assignment 2

In this assignment you will create a coordinate-based multilayer perceptron in numpy from scratch. For each input image coordinate (x, y), the model predicts the associated color (r, g, b).



You will then compare the following input feature mappings $\gamma(v)$.

- No mapping: $\gamma(v) = v$.
- Basic mapping: $\gamma(v) = [\cos(2\pi v), \sin(2\pi v)]^T$.
- Gaussian Fourier feature mapping: $\gamma(v) = [\cos(2\pi B v), \sin(2\pi B v)]^T$, where each entry in $B \in R^{m \times d}$ is sampled from $N(0, \sigma^2)$.

Some notes to help you with that:

- You will implement the mappings in the helper functions `get_B_dict` and `input_mapping`.
- The basic mapping can be considered a case where $B \in R^{2 \times 2}$ is the indentity matrix.
- For this assignment, d is 2 because the input coordinates in two dimensions.
- You can experiment with m and σ values e.g. $m = 256$ and $\sigma \in \{1, 10, 100\}$.

Source: https://bmild.github.io/fourfeat/ This assignment is inspired by and built off of the authors' demo.

Setup

(Optional) Colab Setup

If you aren't using Colab, you can delete the following code cell. Replace the path below with the path in your Google Drive to the uploaded assignment folder. Mounting to Google Drive will allow you access the other .py files in the assignment folder and save outputs to this folder

```python
# you will be prompted with a window asking to grant permissions
# click connect to google drive, choose your account, and click allow
from google.colab import drive
drive.mount("/content/drive")
```

```
Mounted at /content/drive
```

```
# TODO: fill in the path in your Google Drive in the string below
# Note: do not escape slashes or spaces in the path string
import os
datadir = "/content/assignment2"
if not os.path.exists(datadir):
  !ln -s "/content/drive/My Drive/CS444_Neha/assignment2/" $datadir
os.chdir(datadir)
!pwd
```

```
/content/drive/My Drive/CS444_Neha/assignment2
```

Imports

```
import matplotlib.pyplot as plt
from tqdm.notebook import tqdm
import os, imageio
import cv2
import numpy as np

# imports /content/assignment2/models/neural_net.py if you mounted
correctly
from models.neural_net import NeuralNetwork

# makes sure your NeuralNetwork updates as you make changes to the .py
file
%load_ext autoreload
%autoreload 2

# sets default size of plots
%matplotlib inline
plt.rcParams['figure.figsize'] = (10.0, 8.0)
```

Helper Functions

Experiment Runner (Fill in TODOs)

https://github.com/lionelmessi6410/Neural-Networks-from-Scratch/blob/main/NN-from-Scratch.ipynb

https://campuswire.com/c/G333B6F49/feed/206

```
def NN_experiment(X_train, y_train, X_test, y_test, input_size,
num_layers,\
                  hidden_size, hidden_sizes, output_size, epochs,\
                  learning_rate, opt):

  # Initialize a new neural network model
  net = NeuralNetwork(input_size, hidden_sizes, output_size,
num_layers, opt)

  # Variables to store performance for each epoch
  train_loss = np.zeros(epochs)
```

```
  train_psnr = np.zeros(epochs)
  test_psnr = np.zeros(epochs)
  predicted_images = np.zeros((epochs, y_test.shape[0],
y_test.shape[1]))

  # For each epoch...
  for epoch in tqdm(range(epochs)):

      # Shuffle the dataset
      # TODO implement this

      indices = np.random.permutation(X_train.shape[0])
      X_train = X_train[indices,:]
      y_train = y_train[indices,:]

      # Training
      # Run the forward pass of the model to get a prediction and
record the psnr
      # TODO implement this
      fwd_output = net.forward(X_train)
      train_psnr[epoch] = psnr(y_train,fwd_output)

      # Run the backward pass of the model to compute the loss, record
the loss, and update the weights
      # TODO implement this
      loss = net.backward(y_train)
      net.update(lr = learning_rate)
      train_loss[epoch] = loss

      # Testing
      # No need to run the backward pass here, just run the forward
pass to compute and record the psnr
      # TODO implement this
      fwd_output_test = net.forward(X_test)
      test_psnr[epoch] = psnr(y_test,fwd_output_test)
      predicted_images[epoch] = fwd_output_test

  return net, train_psnr, test_psnr, train_loss, predicted_images
```

Image Data and Feature Mappings (Fill in TODOs)

```
# Data loader - already done for you
def get_image(size=512, \

image_url='https://bmild.github.io/fourfeat/img/lion_orig.png'):

  # Download image, take a square crop from the center
  img = imageio.imread(image_url)[..., :3] / 255.
  c = [img.shape[0]//2, img.shape[1]//2]
  r = 256
  img = img[c[0]-r:c[0]+r, c[1]-r:c[1]+r]
```

```python
  if size != 512:
    img = cv2.resize(img, (size, size))

  plt.imshow(img)
  plt.show()

  # Create input pixel coordinates in the unit square
  coords = np.linspace(0, 1, img.shape[0], endpoint=False)
  x_test = np.stack(np.meshgrid(coords, coords), -1)
  test_data = [x_test, img]
  train_data = [x_test[::2, ::2], img[::2, ::2]]

  return train_data, test_data

# Create the mappings dictionary of matrix B -  you will implement
this
def get_B_dict():
  B_dict = {}
  B_dict['none'] = None

  # add B matrix for basic, gauss_1.0, gauss_10.0, gauss_100.0
  # TODO implement this

  # Basic mapping as identity matrix
  B_dict['basic'] = np.eye(2)

  for scale in [1., 10., 100.]:
    B_dict[f'gauss_{scale}'] = np.random.normal(loc=0.0, scale=scale,
size=(256,2))

  return B_dict

# Given tensor x of input coordinates, map it using B - you will
implement
def input_mapping(x, B):
  if B is None:
    # "none" mapping - just returns the original input coordinates
    return x
  else:
    # "basic" mapping and "gauss_X" mappings project input features
using B
    # TODO implement this
    x_proj = (2.*np.pi*x) @ B.T
    return np.hstack([np.sin(x_proj), np.cos(x_proj)])

# Apply the input feature mapping to the train and test data - already
done for you
def get_input_features(B_dict, mapping):
  # mapping is the key to the B_dict, which has the value of B
```

```
  # B is then used with the function `input_mapping` to map x
  y_train = train_data[1].reshape(-1, output_size)
  y_test = test_data[1].reshape(-1, output_size)
  X_train = input_mapping(train_data[0].reshape(-1, 2),
B_dict[mapping])
  X_test = input_mapping(test_data[0].reshape(-1, 2), B_dict[mapping])
  print(y_train.shape)
  print(X_train.shape)

  return X_train, y_train, X_test, y_test
```

MSE Loss and PSNR Error (Fill in TODOs)

```
def mse(y, p):
  # TODO implement this
  # make sure it is consistent with your implementation in
neural_net.py
  return np.mean((y - p)**2)

def psnr(y, p):
  # TODO implement this
  psnr_err = (20*np.log10(np.max(y))) - (10 * np.log10(np.mean((y-
p)**2)))
  return psnr_err
```

Plotting

```
def plot_training_curves(train_loss, train_psnr, test_psnr):
  # plot the training loss
  plt.subplot(2, 1, 1)
  plt.plot(train_loss)
  plt.title('MSE history')
  plt.xlabel('Iteration')
  plt.ylabel('MSE Loss')

  # plot the training and testing psnr
  plt.subplot(2, 1, 2)
  plt.plot(train_psnr, label='train')
  plt.plot(test_psnr, label='test')
  plt.title('PSNR history')
  plt.xlabel('Iteration')
  plt.ylabel('PSNR')
  plt.legend()

  plt.tight_layout()
  plt.show()

def plot_reconstruction(p, y):
  p_im = p.reshape(size,size,3)
  y_im = y.reshape(size,size,3)

  plt.figure(figsize=(12,6))
```

```python
  # plot the reconstruction of the image
  plt.subplot(1,2,1), plt.imshow(p_im), plt.title("reconstruction")

  # plot the ground truth image
  plt.subplot(1,2,2), plt.imshow(y_im), plt.title("ground truth")

  print("Final Test MSE", mse(y, p))
  print("Final Test psnr",psnr(y, p))

def plot_reconstruction_progress(predicted_images, y, N=8):
  total = len(predicted_images)
  step = total // N
  plt.figure(figsize=(24, 4))

  # plot the progress of reconstructions
  for i, j in enumerate(range(0,total, step)):
      plt.subplot(1, N, i+1)
      plt.imshow(predicted_images[j].reshape(size,size,3))
      plt.axis("off")
      plt.title(f"iter {j}")

  # plot ground truth image
  plt.subplot(1, N+1, N+1)
  plt.imshow(y.reshape(size,size,3))
  plt.title('GT')
  plt.axis("off")
  plt.show()

def plot_feature_mapping_comparison(outputs, gt):
  # plot reconstruction images for each mapping
  plt.figure(figsize=(24, 4))
  N = len(outputs)
  for i, k in enumerate(outputs):
      plt.subplot(1, N+1, i+1)
      plt.imshow(outputs[k]['pred_imgs'][-1].reshape(size, size, -1))
      plt.title(k)
  plt.subplot(1, N+1, N+1)
  plt.imshow(gt)
  plt.title('GT')
  plt.show()


  # plot train/test error curves for each mapping
  iters = len(outputs[k]['train_psnrs'])
  plt.figure(figsize=(16, 6))
  plt.subplot(121)
  for i, k in enumerate(outputs):
      plt.plot(range(iters), outputs[k]['train_psnrs'], label=k)
  plt.title('Train error')
```

```
  plt.ylabel('PSNR')
  plt.xlabel('Training iter')
  plt.legend()
  plt.subplot(122)
  for i, k in enumerate(outputs):
      plt.plot(range(iters), outputs[k]['test_psnrs'], label=k)
  plt.title('Test error')
  plt.ylabel('PSNR')
  plt.xlabel('Training iter')
  plt.legend()
  plt.show()
```

Low Resolution Reconstruction

```
size = 32
train_data, test_data = get_image(size)
```

Some suggested hyperparameter choices to help you start

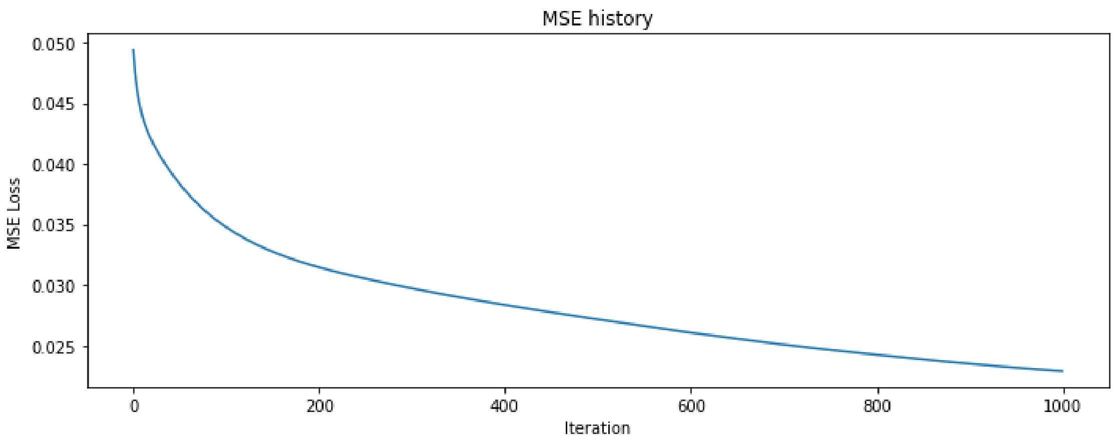
- hidden layer count: 4
- hidden layer size: 256
- number of epochs: 1000
- learning reate: 1e-4

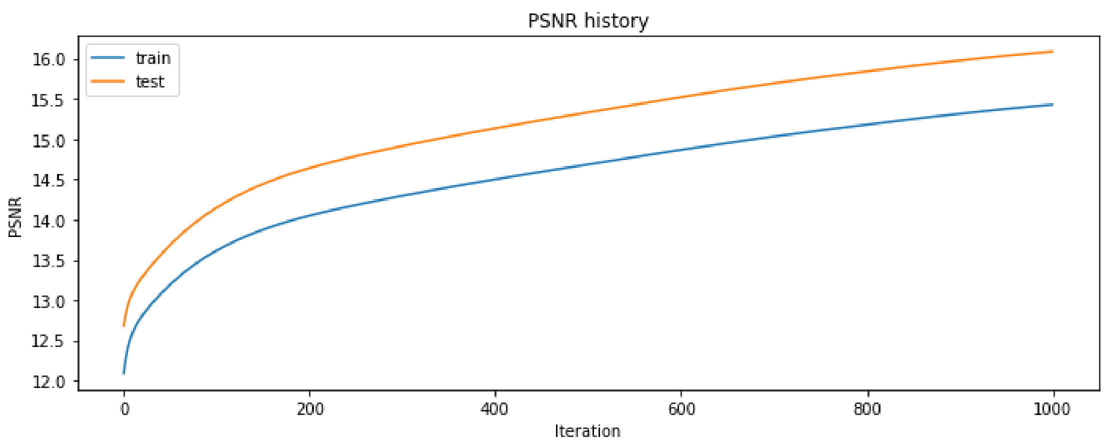
Low Resolution Reconstruction - SGD - None Mapping

```
# get input features
# TODO implement this by using the get_B_dict() and
get_input_features() helper functions
output_size = 3
num_layers = 5
hidden_size = 256
epochs = 1000
lr = 1e-1
opt = "SGD"

X_train, y_train, X_test, y_test = get_input_features(get_B_dict(),
'none')

# run NN experiment on input features
# TODO implement by using the NN_experiment() helper function
input_size = input_mapping(train_data[0].reshape(-1, 2), get_B_dict()
['none']).shape[-1]

net, train_psnr, test_psnr, train_loss, predicted_images =
NN_experiment(X_train, y_train, X_test, y_test,

input_size = input_size,

num_layers = num_layers,

hidden_sizes = [hidden_size] * (num_layers - 1),

hidden_size = hidden_size,

output_size = output_size,

epochs = epochs,

learning_rate = lr,

opt = "SGD")
# plot results of experiment
plot_training_curves(train_loss, train_psnr, test_psnr)
plot_reconstruction(net.forward(X_test), y_test)
plot_reconstruction_progress(predicted_images, y_test)
```

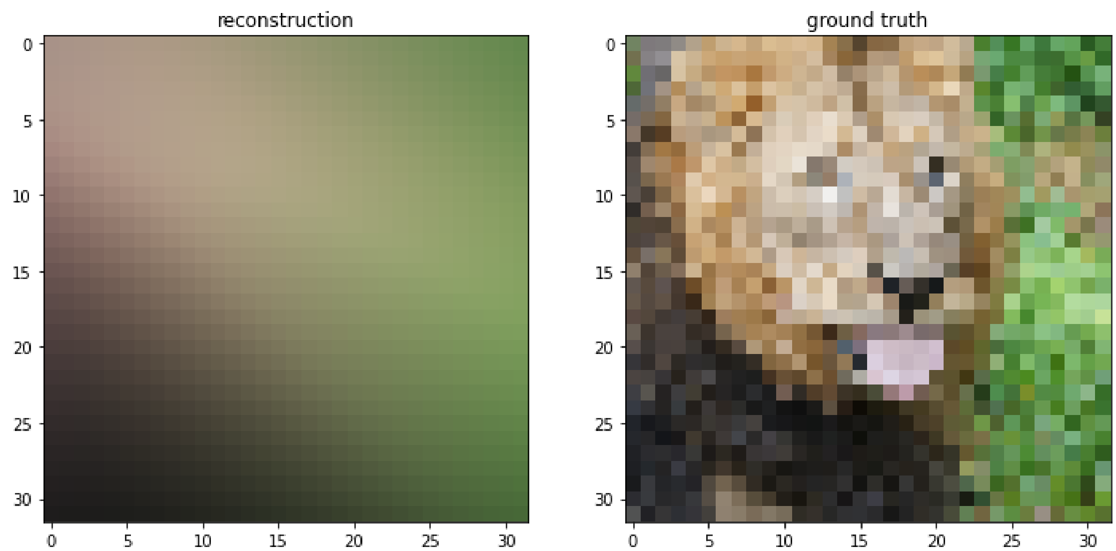
```
(256, 3)
(256, 2)
```

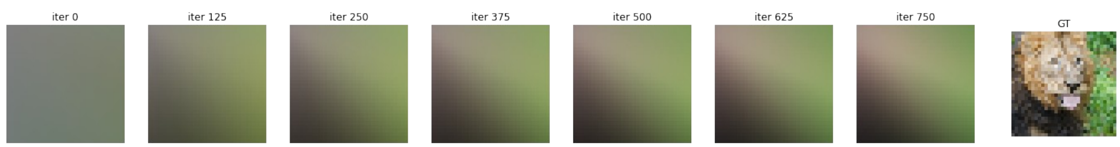
{"model_id":"171d04ef05d944d8b54db23945e77c69","version_major":2,"version_minor":0}






```
Final Test MSE 0.02198921222932762
Final Test psnr 16.087440533832602
```





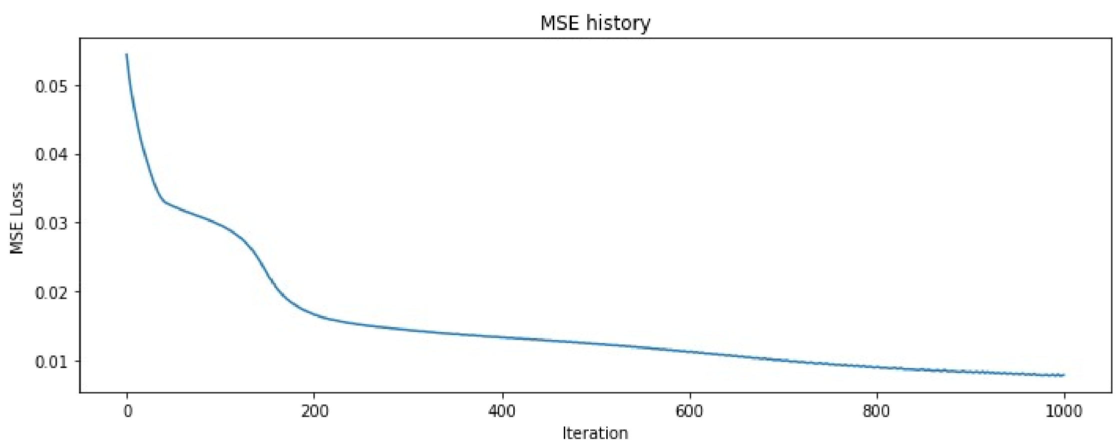
Low Resolution Reconstruction - Adam - None Mapping

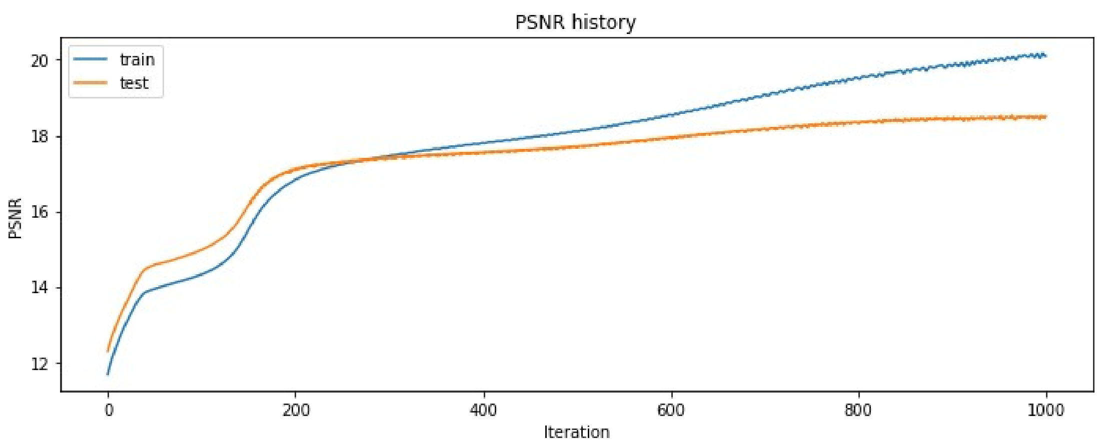
```
# get input features
# TODO implement this by using the get_B_dict() and
get_input_features() helper functions
output_size = 3
num_layers = 7
hidden_size = 256
epochs = 1000
lr = 1e-4

X_train, y_train, X_test, y_test = get_input_features(get_B_dict(),
'none')

input_size = input_mapping(train_data[0].reshape(-1, 2), get_B_dict()
['none']).shape[-1]

# run NN experiment on input features
# TODO implement by using the NN_experiment() helper function

net, train_psnr, test_psnr, train_loss, predicted_images =
NN_experiment(X_train, y_train, X_test, y_test,

input_size = input_size,

num_layers = num_layers,

hidden_sizes = [hidden_size] * (num_layers - 1),
```

```
hidden_size = hidden_size,

output_size = output_size,

epochs = epochs,

learning_rate = lr,

opt = "adam")

# plot results of experiment
plot_training_curves(train_loss, train_psnr, test_psnr)
plot_reconstruction(net.forward(X_test), y_test)
plot_reconstruction_progress(predicted_images, y_test)
```

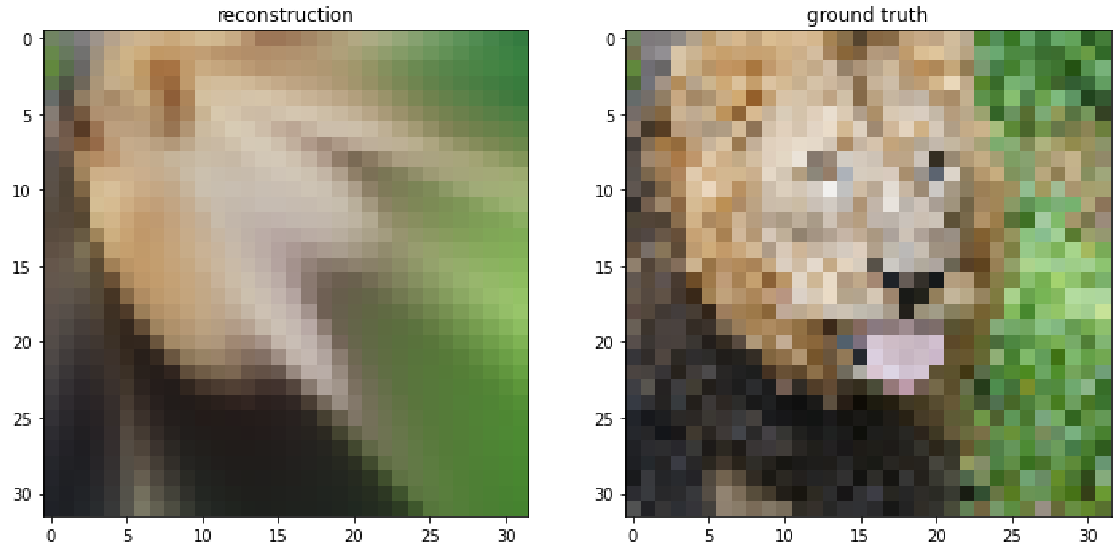
```
(256, 3)
(256, 2)
```

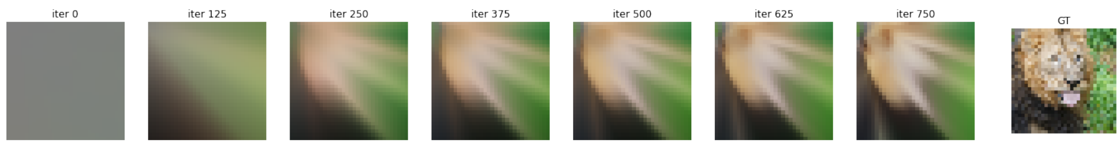
{"model_id":"39c0767589e64f37aa2033a353d1ae3a","version_major":2,"version_minor":0}






```
Final Test MSE 0.012607398818157233
Final Test psnr 18.50328232907115
```





Low Resolution Reconstruction - ADAM - Various Input Mapping Stategies

```
def train_wrapper(mapping, size, opt):
  # TODO implement
  # makes it easy to run all your mapping experiments in a for loop
  # this will similar to what you did previously in the last two
sections

    X_train, y_train, X_test, y_test =
get_input_features(get_B_dict(),mapping)

    input_size = input_mapping(train_data[0].reshape(-1, 2),
get_B_dict()[mapping]).shape[-1]

    net, train_psnr, test_psnr, train_loss, predicted_images =
NN_experiment(X_train, y_train, X_test, y_test,

input_size = input_size,

num_layers = num_layers,

hidden_sizes = [hidden_size] * (num_layers - 1),

hidden_size = hidden_size,

output_size = output_size,

epochs = epochs,
```

```
learning_rate = lr,

opt = opt)

    return {
        'net': net,
        'train_psnrs': train_psnr,
        'test_psnrs': test_psnr,
        'train_loss': train_loss,
        'pred_imgs': predicted_images
    }
```

```
outputs = {}
output_size = 3
hidden_size = 256
epochs = 1000
lr = 1e-4
opt = "adam"
B_dict = get_B_dict()
for k in tqdm(B_dict):
  print("training", k)
  if k == "none":
    num_layers = 7
  if k == "basic":
    num_layers = 5
  if k == "gauss_1.0":
    num_layers = 4
  else:
    num_layers = 5
  outputs[k] = train_wrapper(k, size, opt)
```

{"model_id":"54a2746869f64c58a24195a42e69bbaf","version_major":2,"version_minor":0}

```
training none
(256, 3)
(256, 2)
```

{"model_id":"0b754d5b7433402a921c1ec4dc682498","version_major":2,"version_minor":0}

```
training basic
(256, 3)
(256, 4)
```

{"model_id":"f0107bd4688349308c1498ddcb3fcbe4","version_major":2,"version_minor":0}

```
training gauss_1.0
(256, 3)
(256, 512)
```

{"model_id":"efd9c6ac5c4c4b7fb8a6afd252a6351a","version_major":2,"version_minor":0}

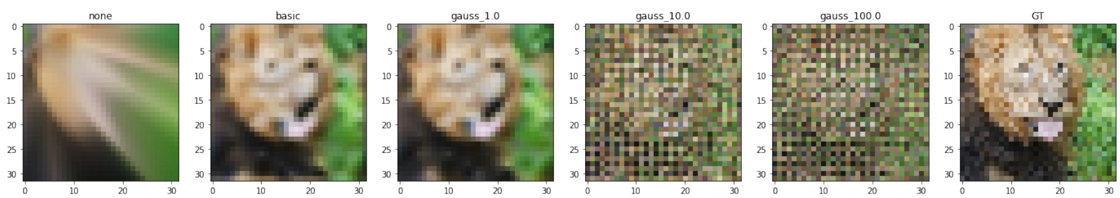
```
training gauss_10.0
(256, 3)
(256, 512)
```

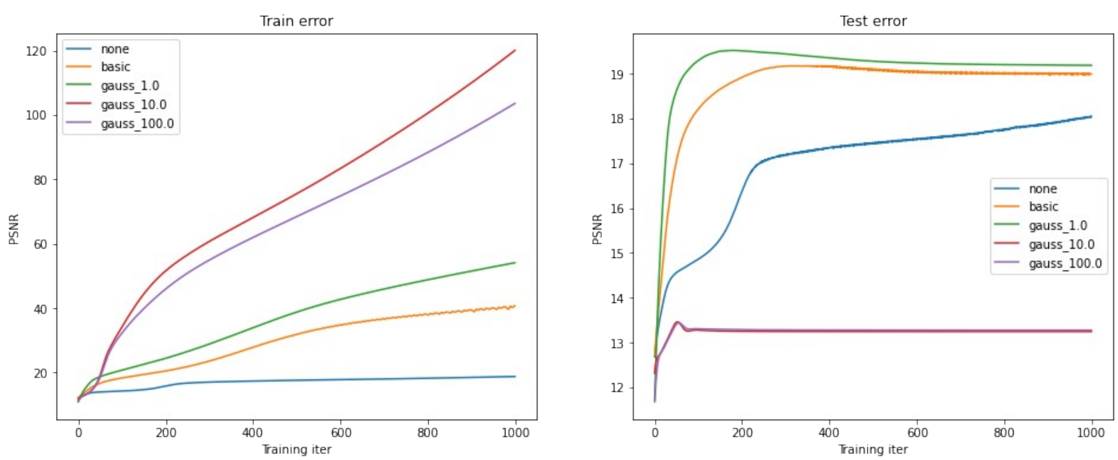
{"model_id":"b9fb1e0e96494d66a11ff18bc1572917","version_major":2,"version_minor":0}

```
training gauss_100.0
(256, 3)
(256, 512)
```

{"model_id":"c5b501106b804a74b71ff2537125ed9b","version_major":2,"version_minor":0}

```
# if you did everything correctly so far, this should output a nice figure you can use in your report
plot_feature_mapping_comparison(outputs, y_test.reshape(size,size,3))
```





Low Resolution Reconstruction - SGD - Various Input Mapping Stategies

```
outputs = {}
output_size = 3
num_layers = 5
hidden_size = 256
epochs = 1000
lr = 1e-1
opt = "SGD"
B_dict = get_B_dict()
for k in tqdm(B_dict):
```

```
  print("training", k)
  outputs[k] = train_wrapper(k, size, opt)
```

{"model_id":"85438c5a94fc410ba5faa3640c4ae299","version_major":2,"version_minor":0}

```
training none
(256, 3)
(256, 2)
```

{"model_id":"5f6532deadb34835bd357d091f73ad25","version_major":2,"version_minor":0}

```
training basic
(256, 3)
(256, 4)
```

{"model_id":"f51a94e2ad5c4b23875d260c684f7153","version_major":2,"version_minor":0}

```
training gauss_1.0
(256, 3)
(256, 512)
```

{"model_id":"21c7383dcdf74e7dafdae05796e95366","version_major":2,"version_minor":0}

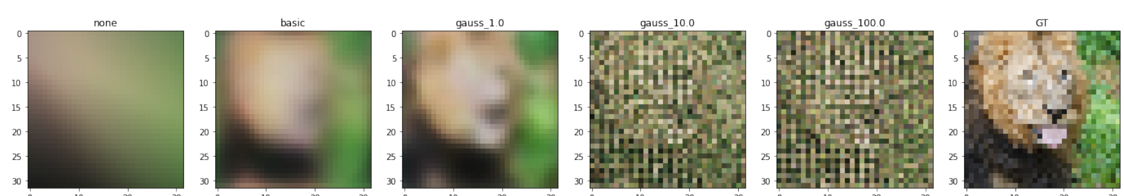
```
training gauss_10.0
(256, 3)
(256, 512)
```

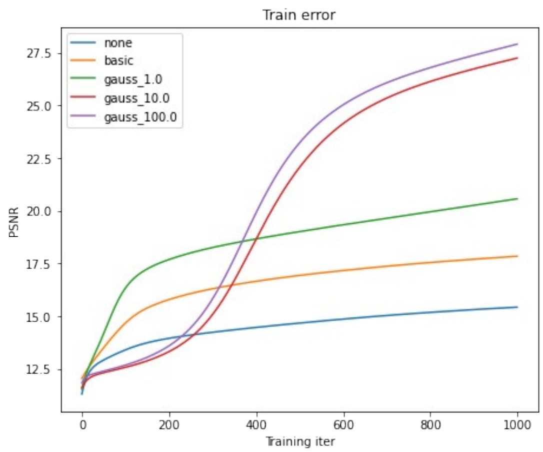
{"model_id":"f26e04991b4f4e6cb68ddb4d23464d21","version_major":2,"version_minor":0}

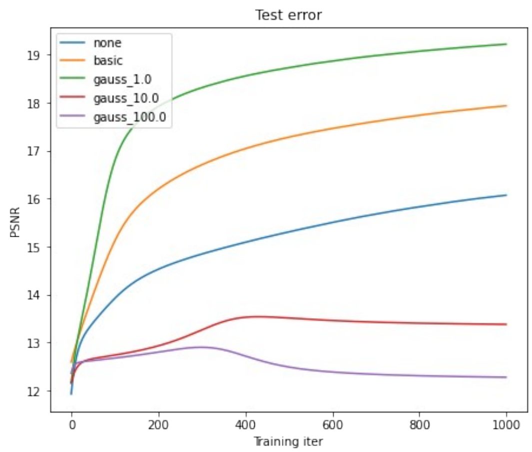
```
training gauss_100.0
(256, 3)
(256, 512)
```

{"model_id":"c35ce4985b954cbfae05507493ab973a","version_major":2,"version_minor":0}

```
# if you did everything correctly so far, this should output a nice figure you can use in your report
plot_feature_mapping_comparison(outputs, y_test.reshape(size,size,3))
```





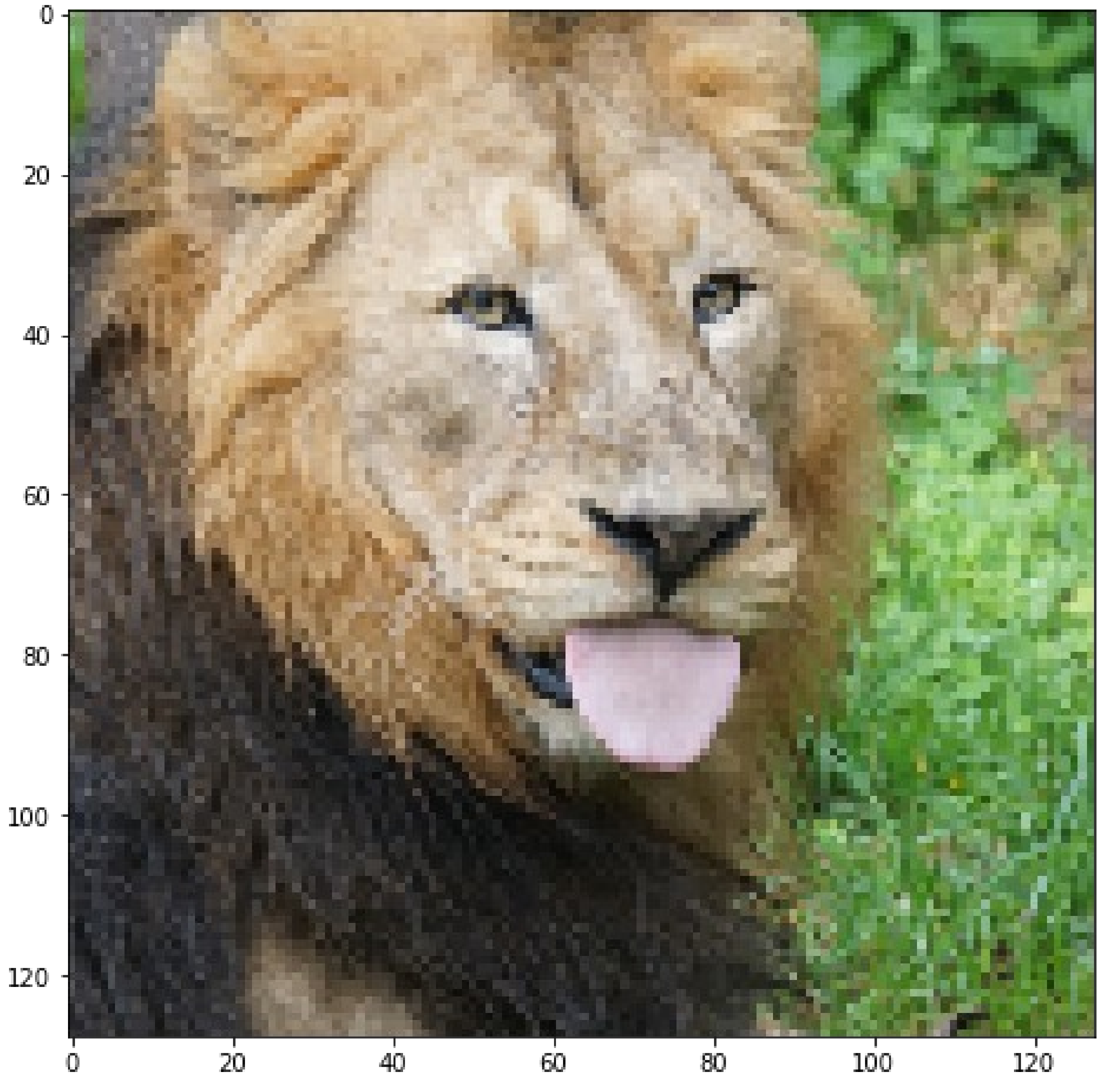


High Resolution Reconstruction

High Resolution Reconstruction - ADAM Optimizer - Various Input Mapping Stategies

Repeat the previous experiment, but at the higher resolution. The reason why we have you first experiment with the lower resolution since it is faster to train and debug. Additionally, you will see how the mapping strategies perform better or worse at the two different input resolutions.

```
size = 128
train_data, test_data = get_image(size)
```




```
outputs = {}
output_size = 3
num_layers = 5
hidden_size = 256
epochs = 1000
lr = 1e-4
opt = "adam"
B_dict = get_B_dict()
for k in tqdm(B_dict):
  print("training", k)
  outputs[k] = train_wrapper(k, size,opt)
```

{"model_id":"2d7b60361898418d82f5b586f931c497","version_major":2,"version_minor":0}

```
training none
(4096, 3)
(4096, 2)
```

{"model_id":"6aed2aaf66834eddbaf38acb843411cf","version_major":2,"version_minor":0}

```
training basic
(4096, 3)
(4096, 4)
```

{"model_id":"9eb36ca7ed194550a4530ac40a84bc31","version_major":2,"version_minor":0}

```
training gauss_1.0
(4096, 3)
(4096, 512)
```

{"model_id":"216e6975010c40b6a46eb32f975605eb","version_major":2,"version_minor":0}

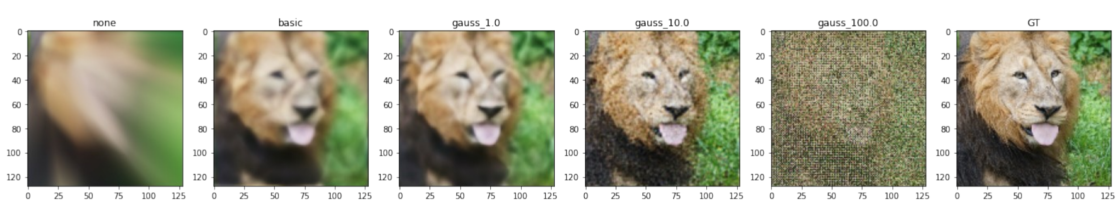
```
training gauss_10.0
(4096, 3)
(4096, 512)
```

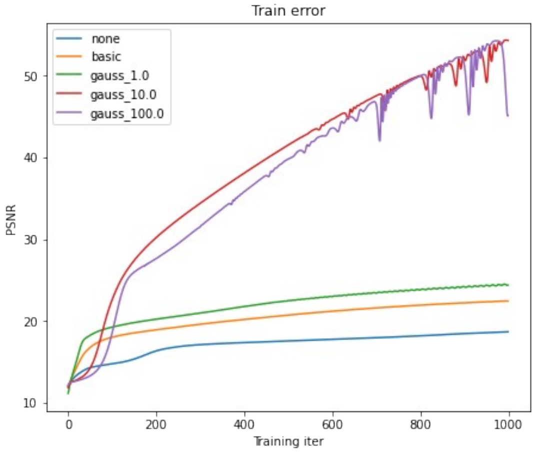
{"model_id":"63b9e5f352914b9f87866497e028c500","version_major":2,"version_minor":0}

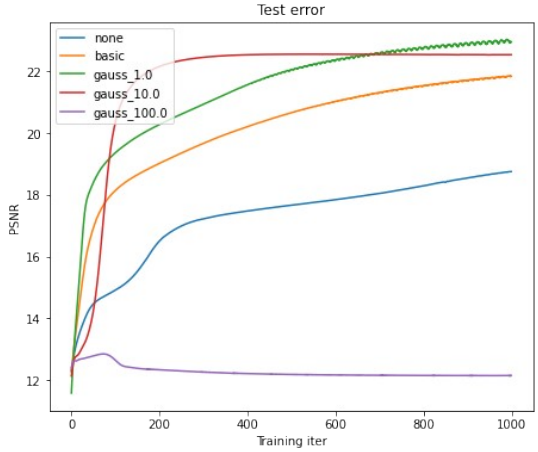
```
training gauss_100.0
(4096, 3)
(4096, 512)
```

{"model_id":"dd6124253d68425ab4cbc69649243691","version_major":2,"version_minor":0}

```python
# if you did everything correctly so far, this should output a nice
figure you can use in your report
plot_feature_mapping_comparison(outputs, y_test.reshape(size,size,3))
```







High Resolution Reconstruction - Image of your Choice

When choosing an image select one that you think will give you interesting results or a better insight into the performance of different feature mappings and explain why in your report template.

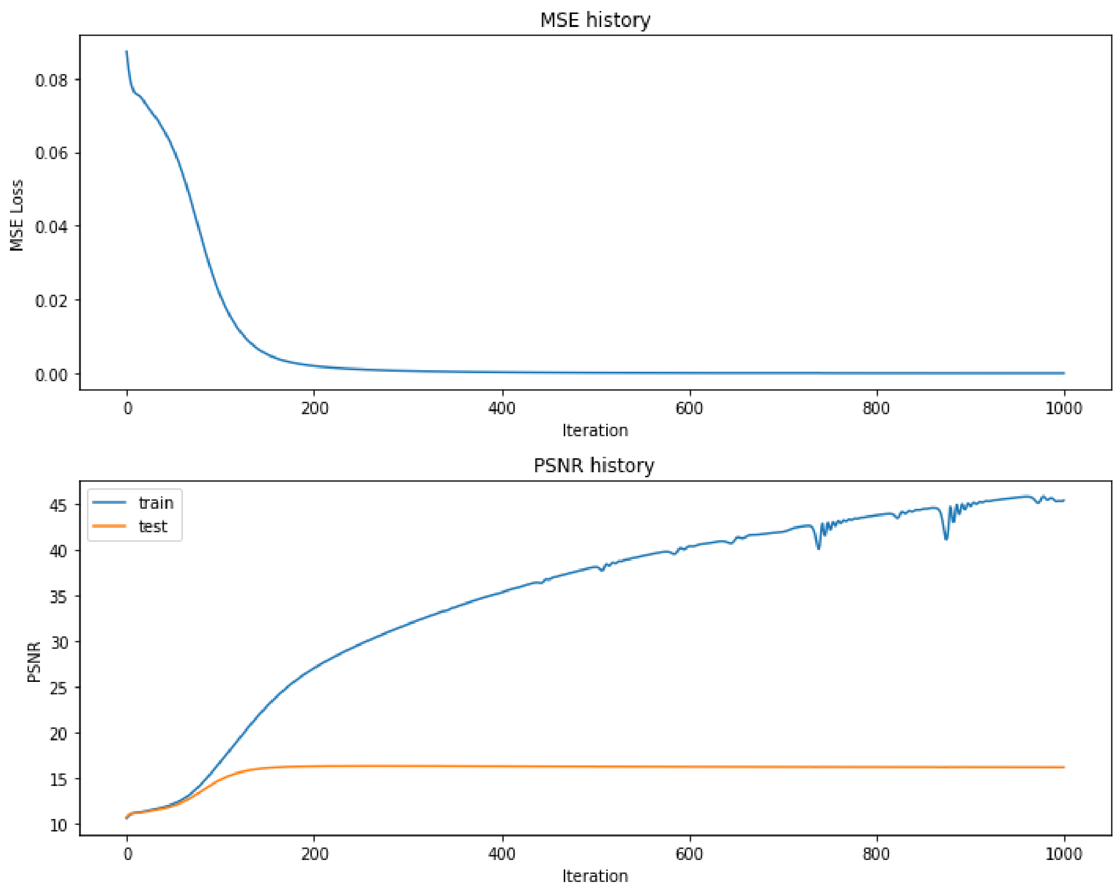
```
size = 128
# TODO pick an image and replace the url string
train_data, test_data = get_image(size, image_url="Burano-Island-is-one-of-the-nicest-places-to-see-in-Venice-Italy.jpeg")
```


```
# get input features
# TODO implement this by using the get_B_dict() and
get_input_features() helper functions

# run NN experiment on input features
# TODO implement by using the NN_experiment() helper function
output_size = 3
num_layers = 5
hidden_size = 256
epochs = 1000
lr = 1e-4

X_train, y_train, X_test, y_test = get_input_features(get_B_dict(),
'gauss_10.0')

# run NN experiment on input features
# TODO implement by using the NN_experiment() helper function
```

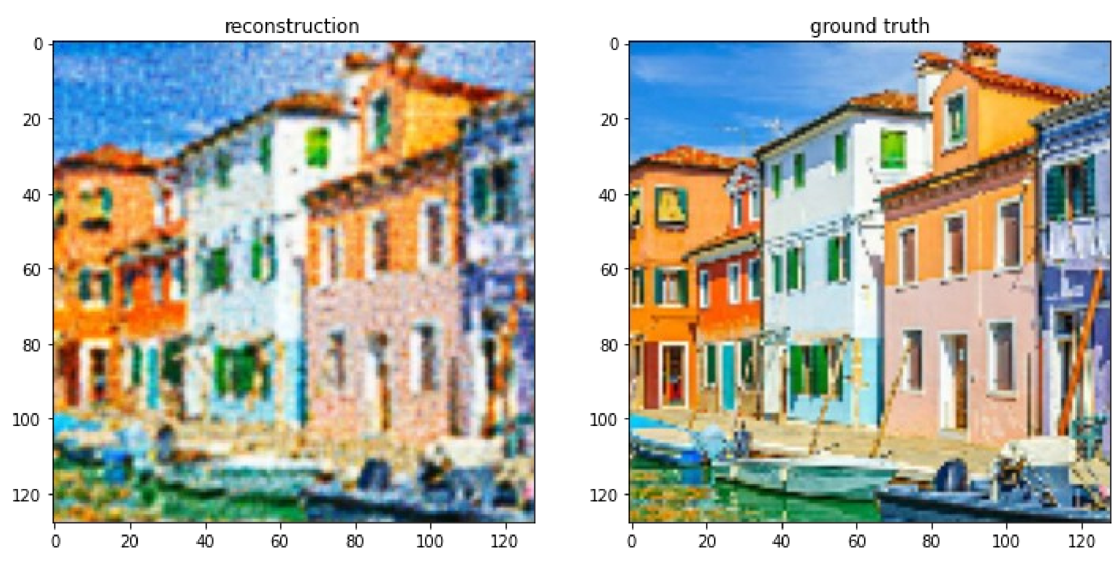
```
input_size = input_mapping(train_data[0].reshape(-1, 2), get_B_dict()
['gauss_10.0']).shape[-1]

net, train_psnr, test_psnr, train_loss, predicted_images =
NN_experiment(X_train, y_train, X_test, y_test,

input_size = input_size ,

num_layers = num_layers,

hidden_sizes = [hidden_size] * (num_layers - 1),

hidden_size = hidden_size,

output_size = output_size,

epochs = epochs,

learning_rate = lr,

opt = "adam")

plot_training_curves(train_loss, train_psnr, test_psnr)
plot_reconstruction(net.forward(X_test), y_test)
plot_reconstruction_progress(predicted_images, y_test)
```

```
(4096, 3)
(4096, 512)
```

{"model_id":"a720768a905540bb9540537f684477bc","version_major":2,"version_minor":0}



Final Test MSE 0.024155154211335l1
Final Test psnr 16.16990185545787



Reconstruction Process Video (Optional)

(For Fun!) Visualize the progress of training in a video

```
# requires installing this additional dependency
!pip install imageio-ffmpeg
```

```
Looking in indexes: https://pypi.org/simple, https://us-
python.pkg.dev/colab-wheels/public/simple/
Collecting imageio-ffmpeg
  Downloading imageio_ffmpeg-0.4.8-py3-none-manylinux2010_x86_64.whl
(26.9 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 26.9/26.9 MB 18.7 MB/s eta
0:00:00
ageio-ffmpeg
Successfully installed imageio-ffmpeg-0.4.8
```

```
# Save out video
def create_and_visualize_video(outputs, size=size, epochs=epochs,
filename='training_convergence.mp4'):
  all_preds = np.concatenate([outputs[n]
['pred_imgs'].reshape(epochs,size,size,3)[::25] for n in outputs],
axis=-2)
  data8 = (255*np.clip(all_preds, 0, 1)).astype(np.uint8)
  f = os.path.join(filename)
  imageio.mimwrite(f, data8, fps=20)

  # Display video inline
  from IPython.display import HTML
  from base64 import b64encode
  mp4 = open(f, 'rb').read()
  data_url = "data:video/mp4;base64," + b64encode(mp4).decode()

  N = len(outputs)
  if N == 1:
    return HTML(f'''
    <video width=256 controls autoplay loop>
          <source src="{data_url}" type="video/mp4">
    </video>
    ''')
  else:
    return HTML(f'''
    <video width=1000 controls autoplay loop>
          <source src="{data_url}" type="video/mp4">
```

```
    </video>
    <table width="1000" cellspacing="0" cellpadding="0">
      <tr>{''.join(N*[f'<td
width="{1000//len(outputs)}"></td>'])}</tr>
      <tr>{''.join(N*['<td style="text-align:center">{}</td>'])}</tr>
    </table>
    '''.format(*list(outputs.keys())))
```

```
# single video example
create_and_visualize_video({"gauss": {"pred_imgs": predicted_images}},
filename="training_high_res_gauss.mp4")
```

```
<IPython.core.display.HTML object>
```

```
# multi video example
create_and_visualize_video(outputs, epochs=1000, size=32)
```