

Image_Captioning

April 18, 2020

Table of Contents

Image Captioning with Visual Attention

Introduction

Step1: Downloading and preparing the MS-COCO dataset

Step 2: Image and caption preparation

Step2a: Defining an Image feature extractor (InceptionV3)

Step 2b: Initialize and loading InceptionV3 with the pretrained Imagenet weights

Step2c: Caching the features extracted from InceptionV3

Step 2d: Caption tokenization and preprocessing

Step 3: Dataset preparation

Step 3a: Splitting data in train and validation set

Step 3b: Creating dataset pipeline

Step 4: Model Definition

Step 4a: Bahdanau's Attention

Step 4b: CNN Encoder

Step 4c: Long short-term memory Decoder

Step 4d: Defining optimizer and checkpoints

Step 5: Evaluation, attention plot and caption generation

Saving Weights

Step 6: Training

Loading weights back

Step 7: Prediction

Testing on other images

*Team Members * Saurabh Mishra - 64 * Vaibhav Singh - 67 * Virag Dosani - 77*

1 Image Captioning with Visual Attention

Given an image like the example below, our goal is to generate a caption such as “**Dog playing with a ball**”.

2 Introduction

Automatically generating captions of an image is a task very close to the heart of scene understanding — one of the primary goals of computer vision. Not only must caption generation models be powerful enough to solve the computer vision challenges of determining which objects are in an image, but they must also be capable of capturing and expressing their relationships in a natural language.

```
[1]: !pip install -q tqdm  
!pip install -q wandb  
!wandb login 984196dfc7bc6ae6093fed3667fd5da413300d29
```



```
|           | 1.4MB 3.4MB/s  
|           | 112kB 40.5MB/s  
|           | 102kB 8.3MB/s  
|           | 102kB 10.4MB/s  
|           | 460kB 35.6MB/s  
|           | 71kB 7.0MB/s  
|           | 71kB 6.7MB/s  
Building wheel for gql (setup.py) ... done  
Building wheel for subprocess32 (setup.py) ... done  
Building wheel for watchdog (setup.py) ... done  
Building wheel for graphql-core (setup.py) ... done  
Building wheel for pathtools (setup.py) ... done  
wandb: Appending key for api.wandb.ai to your netrc file:  
/root/.netrc  
Successfully logged in to Weights & Biases!
```

```
[2]: !pip install -q kaggle  
!mkdir ~/.kaggle  
!cp kaggle.json ~/.kaggle/  
!chmod 600 ~/.kaggle/kaggle.json  
!kaggle datasets download -d thanakomsn/glove6b300dtxt  
!unzip /content/glove6b300dtxt.zip  
!rm -rf /content/glove6b300dtxt.zip
```

```
Downloading glove6b300dtxt.zip to /content  
99% 382M/386M [00:05<00:00, 73.2MB/s]  
100% 386M/386M [00:05<00:00, 72.3MB/s]  
Archive: /content/glove6b300dtxt.zip  
inflating: glove.6B.300d.txt
```

```
[3]: import wandb
wandb.init(project="image-captioning-test")

<IPython.core.display.HTML object>

[3]: W&B Run: https://app.wandb.ai/veb-101/image-captioning-test/runs/14dmbbj2

[0]: import tensorflow as tf

import matplotlib.pyplot as plt

from sklearn.model_selection import train_test_split
from sklearn.utils import shuffle

import re
import numpy as np
import os
import time
import json
from PIL import Image
import pickle
from tqdm import tqdm
import matplotlib.image as mpimg
```

3 Step1: Downloading and preparing the MS-COCO dataset

```
[5]: # downloading caption annotation files
annotation_folder = '/annotations/'
if not os.path.exists(os.path.abspath('.') + annotation_folder):
    annotation_zip = tf.keras.utils.get_file('captions.zip',
                                              cache_subdir=os.path.abspath('.'),
                                              origin = 'http://images.
→cocodataset.org/annotations/annotations_trainval2014.zip',
                                              extract = True)
    annotation_file = os.path.dirname(annotation_zip)+'/annotations/
→captions_train2014.json'
    os.remove(annotation_zip)
else:
    annotation_file = r"/content/annotations/captions_train2014.json"

# downloading image files
image_folder = '/train2014/'
if not os.path.exists(os.path.abspath('.') + image_folder):
    image_zip = tf.keras.utils.get_file('train2014.zip',
```

```

        cache_subdir=os.path.abspath('.'),
        origin = 'http://images.cocodataset.org/
zips/train2014.zip',
        extract = True)
PATH = os.path.dirname(image_zip) + image_folder
os.remove(image_zip)
else:
    PATH = os.path.abspath('..') + image_folder

```

Downloading data from
http://images.cocodataset.org/annotations/annotations_trainval2014.zip
252878848/252872794 [=====] - 3s 0us/step
Downloading data from <http://images.cocodataset.org/zips/train2014.zip>
13510574080/13510573713 [=====] - 333s 0us/step

```
[0]: with open(annotation_file, 'r') as f:
    annotations = json.load(f)

    # storing each captions and image name
    all_captions = []
    all_img_name_vector = []

    for annot in annotations['annotations']:
        caption = '<start> ' + annot['caption'] + ' <end>'
        image_id = annot['image_id']
        full_coco_image_path = PATH + 'COCO_train2014_' + '%012d.jpg' % (image_id)

        all_img_name_vector.append(full_coco_image_path)
        all_captions.append(caption)

    train_captions, img_name_vector = shuffle(all_captions,
                                                all_img_name_vector,
                                                random_state=42)
```

```
[0]: # Select the first 50000 captions from the shuffled set
num_examples = 45000
train_captions = train_captions[:num_examples]
img_name_vector = img_name_vector[:num_examples]
```

```
[8]: len(train_captions), len(img_name_vector), len(all_captions)
```

```
[8]: (45000, 45000, 414113)
```

```
[9]: # Demo Images
```

```
plt.figure(figsize=(18, 15))
for i in range(4):
```

```

plt.subplot(3, 2, i+1)
caption_id = np.random.choice(range(len(train_captions)))
image = mpimg.imread(img_name_vector[caption_id])
caption = train_captions[caption_id].replace("<start>", '')
replace("<end>", '')
caption.strip()
plt.title(caption)
plt.imshow(image)
plt.axis("off")

```

Two giraffes facing opposite directions of each other.



a close up of a plate of food with broccoli



Two children skiing out in the snow together.



A clock on a green lamp post on a sidewalk



4 Step 2: Image and caption preparation

4.1 Step2a: Defining an Image feature extractor (InceptionV3)

```
[0]: def load_image(image_path):
    img = tf.io.read_file(image_path)
    img = tf.image.decode_jpeg(img, channels=3)
    img = tf.image.resize(img, (299, 299))
    img = tf.keras.applications.inception_v3.preprocess_input(img)
    return img, image_path
```

4.2 Step 2b: Initialize and loading InceptionV3 with the pretrained Imagenet weights

```
[11]: image_model = tf.keras.applications.InceptionV3(include_top=False, ↪weights='imagenet')
new_input = image_model.input
hidden_layer = image_model.layers[-1].output

image_features_extract_model = tf.keras.Model(new_input, hidden_layer)
```

Downloading data from https://storage.googleapis.com/tensorflow/keras-applications/inception_v3/inception_v3_weights_tf_dim_ordering_tf_kernels_notop.h5
87916544/87910968 [=====] - 1s 0us/step

4.3 Step2c: Caching the features extracted from InceptionV3

```
[12]: encode_train = sorted(set(img_name_vector))
print(len(encode_train))

# extracting features from images
image_dataset = tf.data.Dataset.from_tensor_slices(encode_train)
image_dataset = image_dataset.map(
    load_image, num_parallel_calls=tf.data.experimental.AUTOTUNE).batch(16)

# caching
for img, path in tqdm(image_dataset):
    batch_features = image_features_extract_model(img)
    # temp = batch_features.shape
    batch_features = tf.reshape(batch_features,
                                (batch_features.shape[0], -1, batch_features.
↪shape[3]))

    for bf, p in zip(batch_features, path):
        path_of_feature = p.numpy().decode("utf-8")
        np.save(path_of_feature, bf.numpy())
```

36168

2261it [10:19, 3.65it/s]

```
[0]: # temp, batch_features.shape
```

4.4 Step 2d: Caption tokenization and preprocessing

```
[0]: # train_captions[:2], img_name_vector[:2]
```

```
[0]: # Find the maximum length of any caption in our dataset
def calc_max_length(tensor):
    return max(len(t) for t in tensor)
```

```
[0]: # Choosing all words in the embedding
top_k = 7500
tokenizer = tf.keras.preprocessing.text.Tokenizer(num_words=top_k,
                                                 oov_token=<unk>,
                                                 filters='!"#$%&()*+.,-/:=?
                                                 →@[\]^_`{|}~ ')
tokenizer.fit_on_texts(train_captions)
```

```
[16]: tokenizer.word_index['<pad>'] = 0
tokenizer.index_word[0] = '<pad>'

word_index = {}
k = 0
for i, j in tokenizer.word_index.items():
    if k < top_k:
        word_index[i] = j
    k += 1

print(len(tokenizer.word_index))
print(len(word_index))
```

9741
7500

```
[0]: # Create the tokenized vectors
train_seqs = tokenizer.texts_to_sequences(train_captions)
# train_seqs[:2]

# Pad each vector to the max_length of the captions
cap_vector = tf.keras.preprocessing.sequence.pad_sequences(train_seqs, ↴
    padding='post')
# cap_vector[:2]

# Calculates the max_length, which is used to store the attention weights
max_length = calc_max_length(train_seqs)
# max_length
```

```
[0]: import csv
def create_csv(file, dict):
    with open(file, 'w') as csvfile:
        writer = csv.writer(csvfile)
        for key in dict.keys():
            writer.writerow([key, dict[key]])
```

```
create_csv('idx2word.csv', tokenizer.index_word)
```

5 Step 3: Dataset preparation

5.1 Step 3a: Splitting data in train and validation set

```
[19]: # Create training and validation sets using an 80-20 split
img_name_train, img_name_val, cap_train, cap_val =_
    train_test_split(img_name_vector,
                     cap_vector,
                     test_size=0.
    ↪2,
    ↪
    ↪shuffle=True,
    ↪
    ↪random_state=47)

len(img_name_train), len(cap_train), len(img_name_val), len(cap_val)
```

```
[19]: (36000, 36000, 9000, 9000)
```

```
[0]: # word embedding initialization

embeddings_index = {};
with open('/content/glove.6B.300d.txt') as f:
    for line in f:
        values = line.split();
        word = values[0];
        coefs = np.asarray(values[1:], dtype='float32');
        embeddings_index[word] = coefs;
```

5.2 Step 3b: Creating dataset pipeline

```
[0]: # imp constants used later
EPOCHS = 20
BATCH_SIZE = 64
BUFFER_SIZE = 1000
embedding_dim = 300
units = 512
vocab_size = len(word_index) + 1
# num_steps = len(img_name_train)
```

```
[0]: # Shape of the vector extracted from InceptionV3 is (64, 2048)
# These two variables represent that vector shape
features_shape = 2048
attention_features_shape = 64
```

```
[0]: # !rm -rf /content/glove.6B.300d.txt  
embeddings_matrix = np.zeros((vocab_size, embedding_dim));  
for word, i in word_index.items():  
    embedding_vector = embeddings_index.get(word);  
    if embedding_vector is not None:  
        embeddings_matrix[i] = embedding_vector;
```

```
[25]: len(embeddings_matrix), len(embeddings_matrix[0])
```

```
[25]: (7501, 300)
```

```
[0]: config = wandb.config  
config.CNN = "InceptionV3"  
config.CNN_out_shape = "8x8x2048"  
config.optimizer = "Adam-default"  
config.captions = num_examples  
config.vocab_size = vocab_size  
config.BATCH_SIZE = BATCH_SIZE  
config.embedding_dim = embedding_dim  
config.Decoder = "LSTM"  
config.EPOCHS = EPOCHS  
config.img_size = "299x299"
```

```
[0]: def map_func(img_name, cap):  
    img_tensor = np.load(img_name.decode('utf-8') + '.npy')  
    return img_tensor, cap
```

```
[0]: def createDataset(img_name, cap):  
    dataset = tf.data.Dataset.from_tensor_slices((img_name, cap))  
    # Use map to load the numpy files in parallel  
    dataset = dataset.map(lambda item1, item2: tf.numpy_function(  
        map_func, [item1, item2], [tf.float32, tf.int32]),  
        num_parallel_calls=tf.data.experimental.AUTOTUNE)  
  
    # Shuffle and batch  
    dataset = dataset.shuffle(BUFFER_SIZE).batch(BATCH_SIZE)  
    dataset = dataset.prefetch(buffer_size=tf.data.experimental.AUTOTUNE)  
    return dataset
```

```
[0]: train_dataset = createDataset(img_name_train, cap_train)  
valid_dataset = createDataset(img_name_val, cap_val)
```

6 Step 4: Model Definition

6.1 Step 4a: Bahdanau's Attention

$$score(s_t, h_i) = v_a^T \tanh(W_a[s_t : h_i])$$

$$\alpha_{t,i} = \frac{\exp(score(s_{t-1}, h_i))}{\sum_{i'=1}^n \exp(score(s_{t-1}, h_{i'}))}$$

$$ContextVector_t = \sum_{i=1}^n \alpha_{t,i} h_i$$

α = attention weights

s_t = previous state output

h_i = image feature 'i'

```
[0]: class SoftAttention(tf.keras.Model):
```

```
    def __init__(self, units):
        super().__init__()
        self.W1 = tf.keras.layers.Dense(units)
        self.W2 = tf.keras.layers.Dense(units)
        self.V = tf.keras.layers.Dense(1)

    def call(self, features, hidden):
        hidden_with_time_axis = tf.expand_dims(hidden, 1)
        score = tf.nn.tanh(self.W1(features) + self.W2(hidden_with_time_axis))
        attention_weights = tf.nn.softmax(self.V(score), axis=1)
        context_vector = attention_weights * features
        context_vector = tf.reduce_sum(context_vector, axis=1)
        return context_vector, attention_weights
```

6.2 Step 4b: CNN Encoder

This encoder maps the input feature of size $(64, 2048)$ to dimesion $(64, 300)$ using a fully connected layer and applies a **relu** activation on the output.

```
[0]: class CNN_Encoder(tf.keras.Model):
```

```
    def __init__(self, embedding_dim):
        super().__init__()
        self.fc = tf.keras.layers.Dense(embedding_dim)

    def call(self, x):
        x = self.fc(x)
        x = tf.nn.relu(x)
        return x
```

6.3 Step 4c: Long short-term memory Decoder

LSTM as used in paper by [Recurrent Neural Network Regularization](#)

h_{t-1} = previous state output

Ey_{t-1} = Word Embedding of previous state output (used as this state's input)

\hat{z}_t = Context Vector t from Attention unit

```
[0]: class RNN_Decoder(tf.keras.Model):
    def __init__(self, attention, embedding_dim, units, vocab_size):
        super().__init__()
        self.units = units

        self.embedding = tf.keras.layers.Embedding(vocab_size, embedding_dim,
                                                input_length=max_length,
                                                weights=[embeddings_matrix],
                                                trainable=False)
        self.lstm = tf.keras.layers.LSTM(self.units,
                                         return_sequences=True,
                                         return_state=True,
                                         recurrent_initializer='glorot_uniform')
        self.gru = tf.keras.layers.GRU(self.units,
                                       return_sequences=True,
                                       return_state=True,
                                       recurrent_initializer='glorot_uniform')

        self.fc1 = tf.keras.layers.Dense(self.units)
        self.fc2 = tf.keras.layers.Dense(vocab_size)

        # need to change here
        self.attention = attention

    def call(self, x, features, hidden, dec="lstm"):
        context_vector, attention_weights = self.attention(features, hidden)
        x = self.embedding(x)
        x = tf.concat([tf.expand_dims(context_vector, 1), x], axis=-1)
        if dec == "lstm":
            lstm1, hidden_state_op, state_cell = self.lstm(x)
            x = self.fc1(hidden_state_op)
            x = self.fc2(x)
            return x, hidden_state_op, attention_weights
        elif dec == "gru":
            hidden_state_op, state = self.gru(x)
            x = self.fc1(hidden_state_op)
            x = tf.reshape(x, (-1, x.shape[2]))
            x = self.fc2(x)
            return x, state, attention_weights

    def reset_state(self, batch_size):
        return tf.zeros((batch_size, self.units))
```

6.4 Step 4d: Defining optimizer and checkpoints

```
[0]: optimizer = tf.keras.optimizers.Adam()
loss_object = tf.keras.losses.SparseCategoricalCrossentropy(
    from_logits=True, reduction='none')

def loss_function(real, pred):
    mask = tf.math.logical_not(tf.math.equal(real, 0))
    loss_ = loss_object(real, pred)
    mask = tf.cast(mask, dtype=loss_.dtype)
    loss_ *= mask
    return tf.reduce_mean(loss_)

[0]: import shutil
try:
    shutil.rmtree("checkpoints")
except:
    pass

[0]: start_epoch = 0

[0]: # checkpoint
checkpoint_path = "./checkpoints/train"
ckpt = tf.train.Checkpoint(encoder=encoder,
                            attention=attention,
                            decoder=decoder,
                            optimizer = optimizer)
ckpt_manager = tf.train.CheckpointManager(ckpt, checkpoint_path, max_to_keep=2)

[0]: # # for restoring latest checkpoint
if ckpt_manager.latest_checkpoint:
    start_epoch = int(ckpt_manager.latest_checkpoint.split('-')[-1])
    ckpt.restore(ckpt_manager.latest_checkpoint)
```

7 Step 5: Evaluation, attention plot and caption generation

```
[0]: def evaluate(image, encoder=None, decoder=None):
    attention_plot = np.zeros((max_length, attention_features_shape))

    hidden = decoder.reset_state(batch_size=1)

    temp_input = tf.expand_dims(load_image(image)[0], 0)
    img_tensor_val = image_features_extract_model(temp_input)
    img_tensor_val = tf.reshape(img_tensor_val, (img_tensor_val.shape[0], -1, img_tensor_val.shape[3]))
```

```

features = encoder(img_tensor_val)

dec_input = tf.expand_dims([tokenizer.word_index['<start>']], 0)
result = []

for i in range(max_length):
    predictions, hidden, attention_weights = decoder(dec_input, features, hidden)
    attention_plot[i] = tf.reshape(attention_weights, (-1, )).numpy()

    predicted_id = tf.random.categorical(predictions, 1)[0][0].numpy()
    result.append(tokenizer.index_word[predicted_id])

    if tokenizer.index_word[predicted_id] == '<end>':
        return result, attention_plot

    dec_input = tf.expand_dims([predicted_id], 0)

attention_plot = attention_plot[:len(result), :]
return result, attention_plot

```

```

[0]: def plot_attention(image, result, attention_plot):
    temp_image = np.array(Image.open(image))

    fig = plt.figure(figsize=(15, 12))

    len_result = len(result)
    for l in range(len_result):
        temp_att = np.resize(attention_plot[l], (8, 8))
        ax = fig.add_subplot(len_result//2, len_result//2, l+1)
        ax.set_title(result[l])
        img = ax.imshow(temp_image)
        ax.axis("off")
        ax.imshow(temp_att, cmap='gray', alpha=0.6, extent=img.get_extent())

    plt.tight_layout()
    plt.show()

```

```

[0]: def predict(id=None, name=None, encoder=encoder, decoder=decoder):

    # captions on the validation set
    if id == None and name==None:
        rid = np.random.randint(0, len(img_name_val))
    else:
        rid = id

    if rid:

```

```

    real_caption = ' '.join([tokenizer.index_word[i] for i in cap_val[rid]
    ↵if i not in [0]])
    print ('Real Caption:', real_caption)

if name == None:
    image = img_name_val[rid]
else:
    image = name
    rid=None

result, attention_plot = evaluate(image, encoder=encoder, decoder=decoder)
print ('Prediction Caption:', ' '.join(result))
print(rid, image)
plot_attention(image, result, attention_plot)
plt.figure(figsize=(12, 10))
img = mpimg.imread(image)
plt.imshow(img)
plt.axis("off")

```

8 Saving Weights

```
[0]: import shutil

try:
    shutil.rmtree("/content/Attention")
    shutil.rmtree("/content/Decoder")
    shutil.rmtree("/content/Encoder")
except:
    pass

[0]: # SAVING
## ENCODER
from google.colab import files

def save():
    encoder.save_weights("./Encoder/encoder_weights", save_format='tf')

    ## Attention
    attention.save_weights("./Attention/attention_weights", save_format='tf')

    ## Decoder
    decoder.save_weights("./Decoder/decoder_weights", save_format='tf')

# DOWNLOADING WEIGHTS
```

```

!zip -rq Attention.zip ./Attention
!zip -rq Encoder.zip ./Encoder
!zip -rq Decoder.zip ./Decoder

!zip -rq all.zip Attention.zip Encoder.zip Decoder.zip
!rm -rf Attention.zip Decoder.zip Encoder.zip

shutil.rmtree("/content/Attention")
shutil.rmtree("/content/Decoder")
shutil.rmtree("/content/Encoder")

# files.download('all.zip')
# !rm -rf all.zip

```

9 Step 6: Training

- You extract the features stored in the respective .npy files and then pass those features through the encoder.
- The encoder output, hidden state(initialized to 0) and the decoder input (which is the start token) is passed to the decoder.
- The decoder returns the predictions and the decoder hidden state.
- The decoder hidden state is then passed back into the model and the predictions are used to calculate the loss.
- Use teacher forcing to decide the next input to the decoder.
- Teacher forcing is the technique where the target word is passed as the next input to the decoder.
- The final step is to calculate the gradients and apply it to the optimizer and backpropagate.

```
[0]: train_loss_plot = []
valid_loss_plot = []
```

```
[0]: # @tf.function
def train_step(img_tensor, target, encoder=None, decoder=None):
    loss = 0

    # initializing the hidden state for each batch
    # because the captions are not related from image to image
    hidden = decoder.reset_state(batch_size=target.shape[0])

    dec_input = tf.expand_dims([tokenizer.word_index['<start>']] * target.
    ↪shape[0], 1)

    with tf.GradientTape() as tape:
        features = encoder(img_tensor)
```

```

        for i in range(1, target.shape[1]):
            predictions, hidden, _ = decoder(dec_input, features, hidden, "lstm")
            loss += loss_function(target[:, i], predictions)
            dec_input = tf.expand_dims(target[:, i], 1)

        total_loss = (loss / int(target.shape[1]))
        trainable_variables = encoder.trainable_variables + decoder.trainable_variables
        gradients = tape.gradient(loss, trainable_variables)
        optimizer.apply_gradients(zip(gradients, trainable_variables))
    return loss, total_loss

# @tf.function
def valid_step(img_tensor, target, encoder=None, decoder=None):
    loss = 0
    hidden = decoder.reset_state(batch_size=target.shape[0])
    dec_input = tf.expand_dims([tokenizer.word_index['<start>']] * target.shape[0], 1)

    features = encoder(img_tensor)
    for i in range(1, target.shape[1]):
        predictions, hidden, _ = decoder(dec_input, features, hidden, "lstm")
        loss += loss_function(target[:, i], predictions)
        dec_input = tf.expand_dims(target[:, i], 1)

    total_loss = (loss / int(target.shape[1]))

    return loss, total_loss

```

```

[0]: def startTraining(start_epoch, EPOCHS, training_data=None, validation_data=None, encoder=None, decoder=None):
    for epoch in range(start_epoch, EPOCHS):
        start = time.time()
        train_total_loss = 0
        valid_total_loss = 0

        train_num_steps = 0
        valid_num_steps = 0

        for (batch, (img_tensor, target)) in enumerate(training_data):
            batch_loss, t_loss = train_step(img_tensor, target, encoder, decoder)
            train_total_loss += t_loss
            train_num_steps += 1

```

```

        if batch % 100 == 0:
            print (f'Epoch {epoch + 1} Batch {batch} Loss {batch_loss.
→numpy()/int(target.shape[1]):.4f}')

    for (b, (tensor, tar)) in enumerate(validation_data):
        # print(tensor, tar)
        _, t_loss = valid_step(tensor, tar, encoder, decoder)
        valid_total_loss += t_loss
        valid_num_steps += 1

    train_loss_steps = train_total_loss / train_num_steps
    valid_loss_steps = valid_total_loss / valid_num_steps

    train_loss_plot.append(train_loss_steps)
    valid_loss_plot.append(valid_loss_steps)

# if epoch % 3 == 0:
#     ckpt_manager.save()

    print (f'Epoch: {epoch + 1}; Train Loss: {train_loss_steps:.6f};\n
→Validation loss: {valid_loss_steps:.6f}')
    print (f'Time taken for {epoch + 1}; Epoch {time.time() - start} sec\n')

    wandb.log({"train_loss": train_loss_steps, "valid_loss":\n
→valid_loss_steps, "epoch": epoch + 1})

    if input("Save weights: ") == "y":
        save()
        print("Weights Saved")

    if input("Terminate: ") == "y":
        print("finished Training")
        return None

```

[0]:

```
encoder = CNN_Encoder(embedding_dim)
attention = SoftAttention(units)
decoder = RNN_Decoder(attention, embedding_dim, units, vocab_size)
```

[61]:

```
startTraining(start_epoch, EPOCHS, training_data=train_dataset,\n
→validation_data=valid_dataset, encoder=encoder, decoder=decoder)
```

Epoch 1 Batch 0 Loss 2.1394
Epoch 1 Batch 100 Loss 1.3223
Epoch 1 Batch 200 Loss 1.2835

```
Epoch 1 Batch 300 Loss 1.2329
Epoch 1 Batch 400 Loss 1.1657
Epoch 1 Batch 500 Loss 1.0535
Epoch: 1; Train Loss: 1.228358; Validation loss: 1.005958
Time taken for 1; Epoch 850.2502801418304 sec
```

```
Save weights: y
Weights Saved
Terminate: n
Epoch 2 Batch 0 Loss 0.9921
Epoch 2 Batch 100 Loss 0.9718
Epoch 2 Batch 200 Loss 0.9550
Epoch 2 Batch 300 Loss 0.9418
Epoch 2 Batch 400 Loss 0.8109
Epoch 2 Batch 500 Loss 0.8075
Epoch: 2; Train Loss: 0.878002; Validation loss: 0.826471
Time taken for 2; Epoch 864.7819554805756 sec
```

```
Save weights: y
Weights Saved
Terminate: n
Epoch 3 Batch 0 Loss 0.7966
Epoch 3 Batch 100 Loss 0.8063
Epoch 3 Batch 200 Loss 0.8142
Epoch 3 Batch 300 Loss 0.7205
Epoch 3 Batch 400 Loss 0.6949
Epoch 3 Batch 500 Loss 0.7867
Epoch: 3; Train Loss: 0.766892; Validation loss: 0.792328
Time taken for 3; Epoch 866.5353519916534 sec
```

```
Save weights: y
Weights Saved
Terminate: n
Epoch 4 Batch 0 Loss 0.7862
Epoch 4 Batch 100 Loss 0.6735
Epoch 4 Batch 200 Loss 0.7137
Epoch 4 Batch 300 Loss 0.7040
Epoch 4 Batch 400 Loss 0.6861
Epoch 4 Batch 500 Loss 0.7380
Epoch: 4; Train Loss: 0.714045; Validation loss: 0.779327
Time taken for 4; Epoch 873.8853802680969 sec
```

```
Save weights: y
Weights Saved
Terminate: n
Epoch 5 Batch 0 Loss 0.7145
Epoch 5 Batch 100 Loss 0.6868
Epoch 5 Batch 200 Loss 0.7124
```

```
Epoch 5 Batch 300 Loss 0.6534
Epoch 5 Batch 400 Loss 0.6597
Epoch 5 Batch 500 Loss 0.6590
Epoch: 5; Train Loss: 0.676530; Validation loss: 0.774025
Time taken for 5; Epoch 872.0828177928925 sec
```

```
Save weights: y
Weights Saved
Terminate: n
Epoch 6 Batch 0 Loss 0.7419
Epoch 6 Batch 100 Loss 0.6085
Epoch 6 Batch 200 Loss 0.6504
Epoch 6 Batch 300 Loss 0.6121
Epoch 6 Batch 400 Loss 0.6141
Epoch 6 Batch 500 Loss 0.7018
Epoch: 6; Train Loss: 0.646019; Validation loss: 0.781322
Time taken for 6; Epoch 877.4589960575104 sec
```

```
Save weights: n
Terminate: n
Epoch 7 Batch 0 Loss 0.6581
Epoch 7 Batch 100 Loss 0.6662
Epoch 7 Batch 200 Loss 0.6422
Epoch 7 Batch 300 Loss 0.6173
Epoch 7 Batch 400 Loss 0.5924
Epoch 7 Batch 500 Loss 0.5870
Epoch: 7; Train Loss: 0.618556; Validation loss: 0.787416
Time taken for 7; Epoch 872.9390561580658 sec
```

```
Save weights: n
Terminate: n
Epoch 8 Batch 0 Loss 0.6305
Epoch 8 Batch 100 Loss 0.5757
Epoch 8 Batch 200 Loss 0.5977
Epoch 8 Batch 300 Loss 0.6197
Epoch 8 Batch 400 Loss 0.5765
Epoch 8 Batch 500 Loss 0.5455
Epoch: 8; Train Loss: 0.592001; Validation loss: 0.796695
Time taken for 8; Epoch 865.4335935115814 sec
```

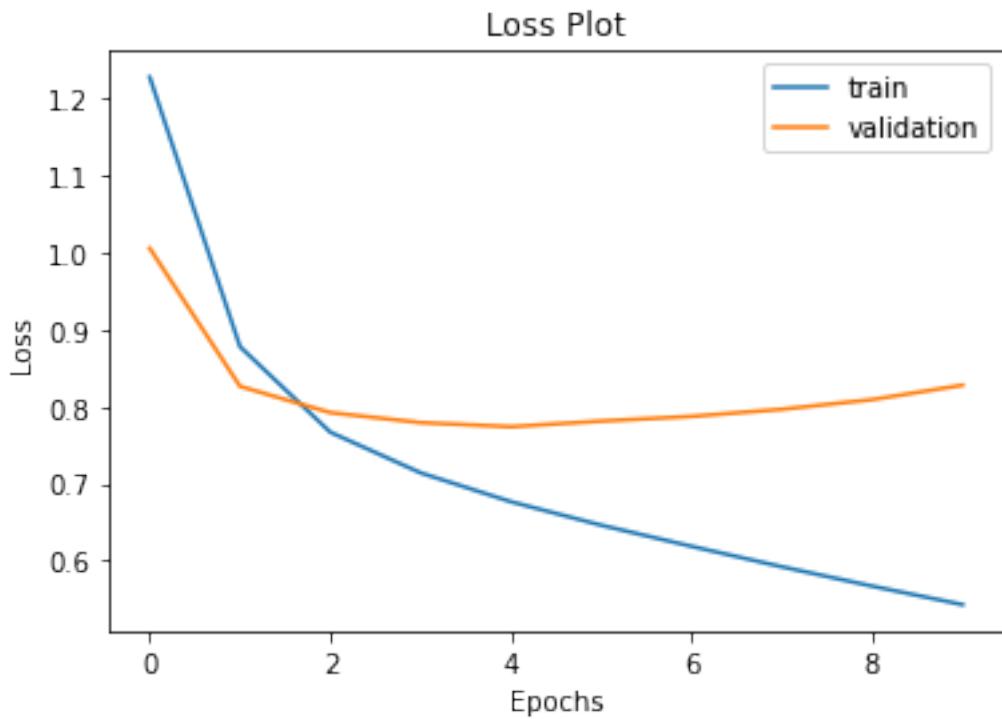
```
Save weights: n
Terminate: n
Epoch 9 Batch 0 Loss 0.6293
Epoch 9 Batch 100 Loss 0.6162
Epoch 9 Batch 200 Loss 0.5778
Epoch 9 Batch 300 Loss 0.5799
Epoch 9 Batch 400 Loss 0.5657
Epoch 9 Batch 500 Loss 0.5653
```

```
Epoch: 9; Train Loss: 0.566862; Validation loss: 0.809250
Time taken for 9; Epoch 853.5799715518951 sec
```

```
Save weights: n
Terminate: n
Epoch 10 Batch 0 Loss 0.5499
Epoch 10 Batch 100 Loss 0.5461
Epoch 10 Batch 200 Loss 0.5043
Epoch 10 Batch 300 Loss 0.5540
Epoch 10 Batch 400 Loss 0.5671
Epoch 10 Batch 500 Loss 0.5759
Epoch: 10; Train Loss: 0.542814; Validation loss: 0.828003
Time taken for 10; Epoch 862.9952504634857 sec
```

```
Save weights: n
Terminate: y
finished Training
```

```
[62]: plt.plot(train_loss_plot, label="train")
plt.plot(valid_loss_plot, label="validation")
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.title('Loss Plot')
plt.legend()
plt.show()
```



```
[0]: # wandb.save('./checkpoints/train/*ckpt*')
```

10 Loading weights back

```
[0]: !unzip -qq all.zip  
!unzip -qq Attention.zip  
!unzip -qq Encoder.zip  
!unzip -qq Decoder.zip  
  
!rm -rf Attention.zip Encoder.zip Decoder.zip
```

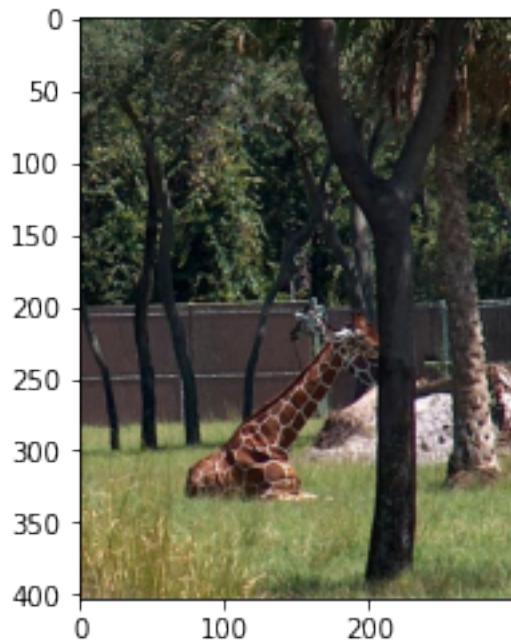
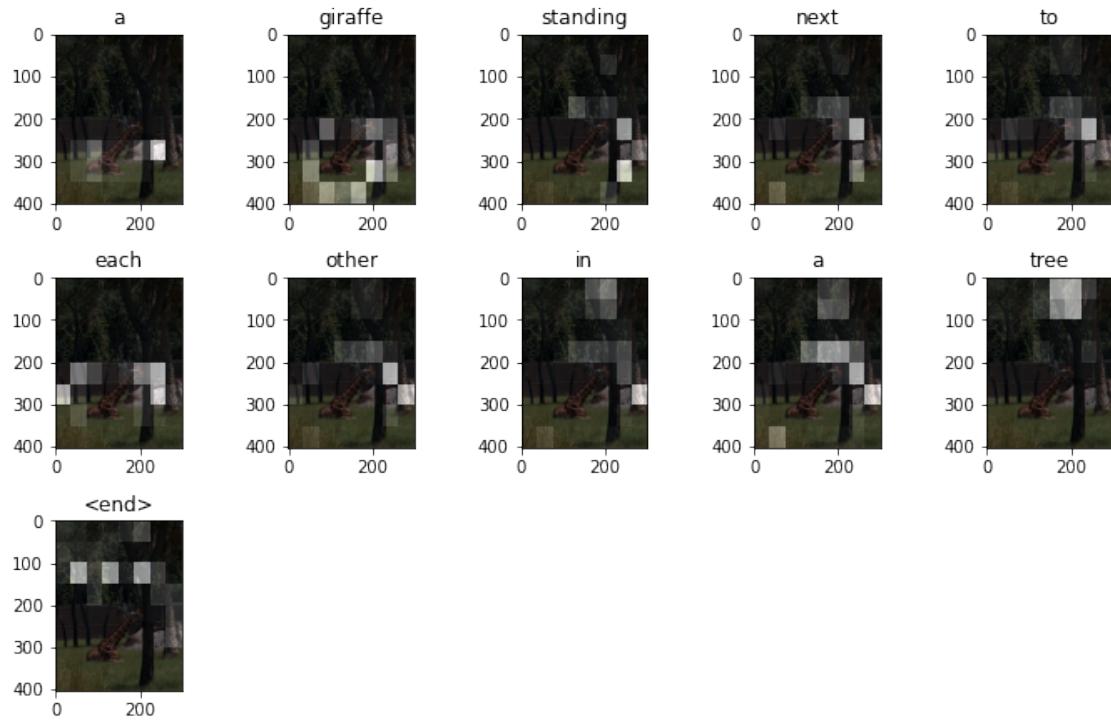
```
[65]: encoder2 = CNN_Encoder(embedding_dim)  
encoder2.load_weights("./Encoder/encoder_weights")  
  
attention2 = SoftAttention(units)  
attention2.load_weights("./Attention/attention_weights")  
  
decoder2 = RNN_Decoder(attention, embedding_dim, units, vocab_size)  
decoder2.load_weights("./Decoder/decoder_weights")
```

```
[65]: <tensorflow.python.training.util.CheckpointLoadStatus at  
0x7f8cf7e25ba8>
```

11 Step 7: Prediction

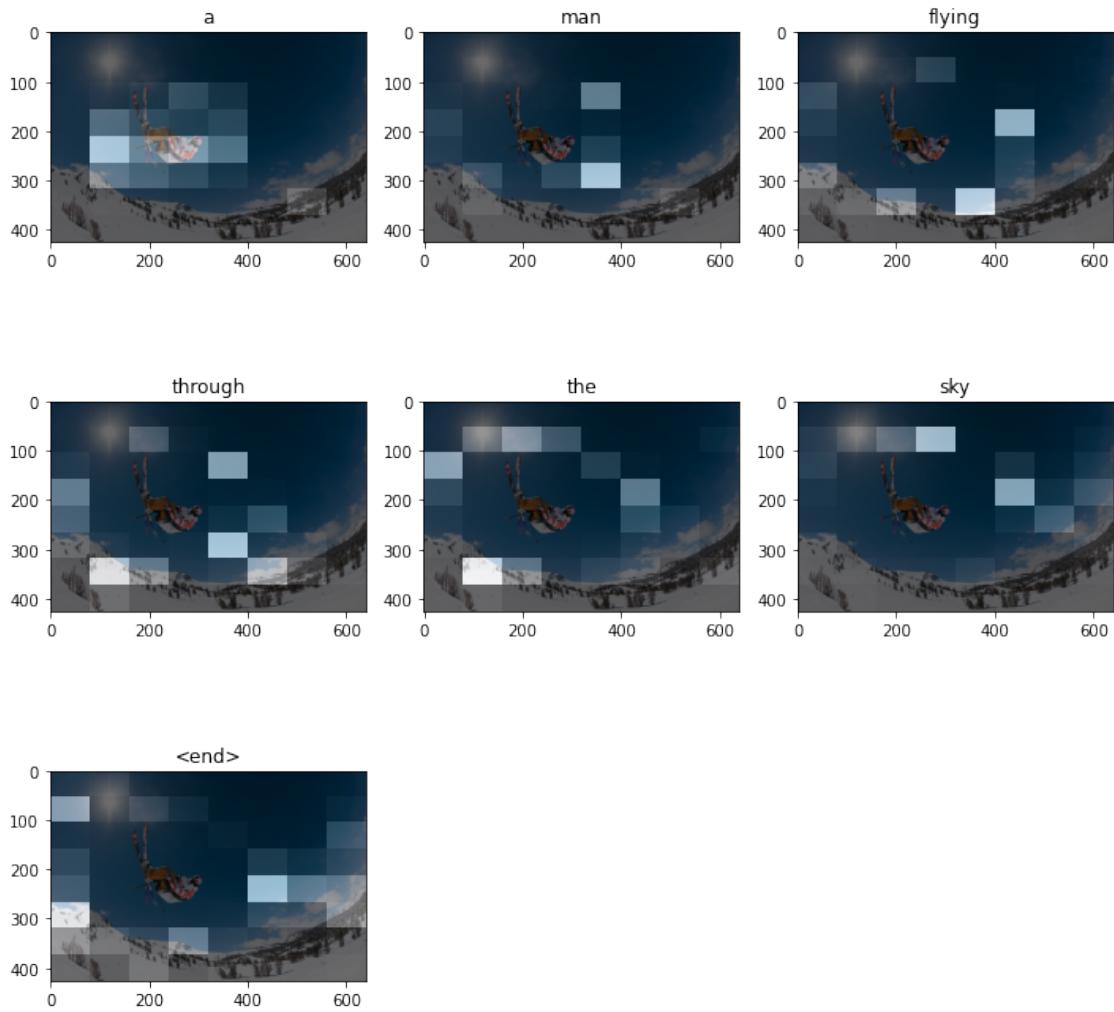
```
[0]: # testing on validation set
```

WARNING:tensorflow:10 out of the last 10 calls to <function load_image at 0x7fe7a0ebb840> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings is likely due to passing python objects instead of tensors. Also, tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. Please refer to https://www.tensorflow.org/tutorials/customization/performance#python_or_tensor_args and https://www.tensorflow.org/api_docs/python/tf/function for more details.
Real Caption: <start> a giraffe laying on lush green grass next to trees <end>
Prediction Caption: a giraffe standing next to each other in a tree <end>
2963 /content/train2014/COCO_train2014_000000424978.jpg



[0] : predict()

WARNING:tensorflow:11 out of the last 11 calls to <function load_image at 0x7fe7a0ebb840> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings is likely due to passing python objects instead of tensors. Also, tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. Please refer to https://www.tensorflow.org/tutorials/customization/performance#python_or_tensor_args and https://www.tensorflow.org/api_docs/python/tf/function for more details.
Real Caption: <start> a man on snow skis who is performing a jump <end>
Prediction Caption: a man flying through the sky <end>
8032 /content/train2014/COCO_train2014_000000247906.jpg





```
[71]: predict(2936, encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> bird on edge of sill looking to the left <end>

Prediction Caption: a bird is eating treats surrounded by flowers <end>

2936 /content/train2014/COCO_train2014_000000164997.jpg





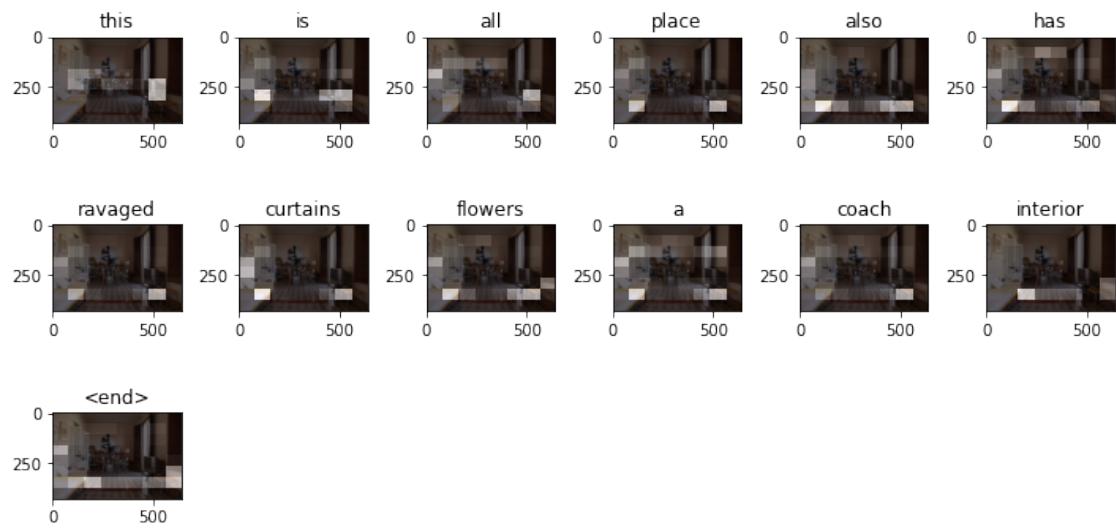
[0] : predict()

WARNING:tensorflow:11 out of the last 11 calls to <function load_image at 0x7fe7a0ebb840> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings is likely due to passing python objects instead of tensors. Also, tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. Please refer to https://www.tensorflow.org/tutorials/customization/performance#python_or_tensor_args and https://www.tensorflow.org/api_docs/python/tf/function for more details.

Real Caption: <start> a empty kitchen that has table with two chairs <end>

Prediction Caption: this is all place also has ravaged curtains flowers a coach interior <end>

6286 /content/train2014/COCO_train2014_000000365003.jpg



```
[76]: predict(encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> a baby is lying next to a large teddy bear <end>
 Prediction Caption: a teddy bear is decorated with two teddy bear <end>
 5402 /content/train2014/COCO_train2014_000000080016.jpg





[78]: predict(id=4046, encoder=encoder2, decoder=decoder2)

Real Caption: <start> a couple of elephants that are by the pond <end>

Prediction Caption: a group of elephants relax along water in a body of water

<end>

4046 /content/train2014/COCO_train2014_000000080105.jpg



```
[0]: predict(id=2187, encoder=encoder2, decoder=decoder2)
```

WARNING:tensorflow:11 out of the last 11 calls to <function load_image at 0x7fe7a0ebb840> triggered tf.function retracing. Tracing is expensive and the excessive number of tracings is likely due to passing python objects instead of tensors. Also, tf.function has experimental_relax_shapes=True option that relaxes argument shapes that can avoid unnecessary retracing. Please refer to https://www.tensorflow.org/tutorials/customization/performance#python_or_tensor_args and https://www.tensorflow.org/api_docs/python/tf/function for more details.
Real Caption: <start> a room with rows of televisions and a vhs tapes <end>
Prediction Caption: desk filled with television on top of it <end>
2187 /content/train2014/COCO_train2014_000000442774.jpg



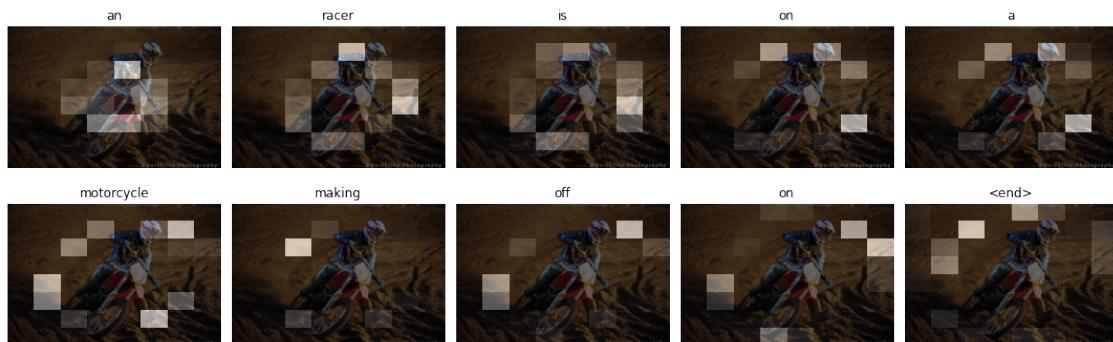


```
[82]: predict(id=10, encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> a man riding a dirt bike on top of a dirt road <end>

Prediction Caption: an racer is on a motorcycle making off on <end>

10 /content/train2014/COCO_train2014_000000439268.jpg

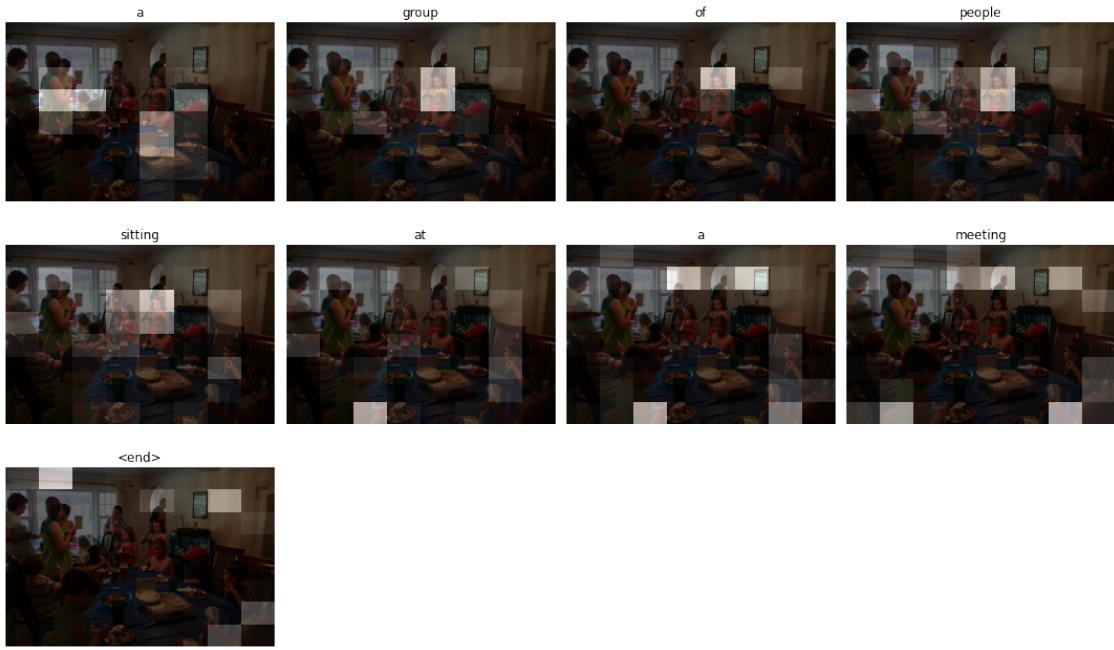




```
[83]: predict(882, encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> a group of people sitting and standing around a table
<end>

Prediction Caption: a group of people sitting at a meeting <end>
882 /content/train2014/COCO_train2014_000000468125.jpg

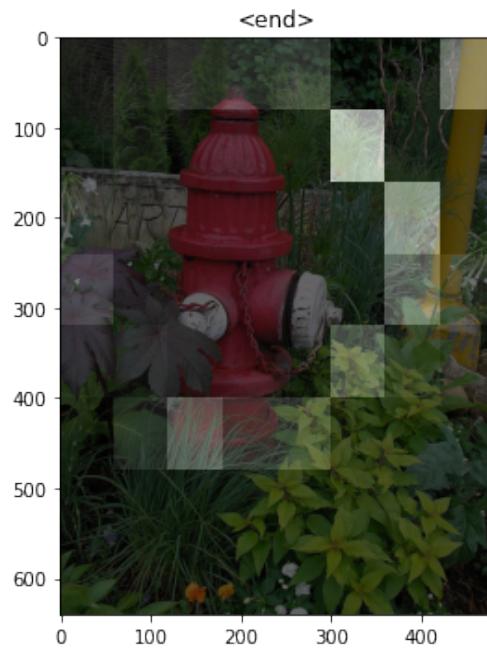
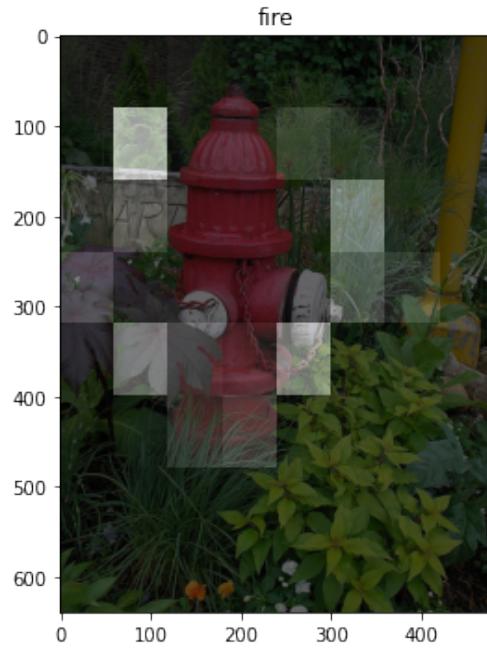
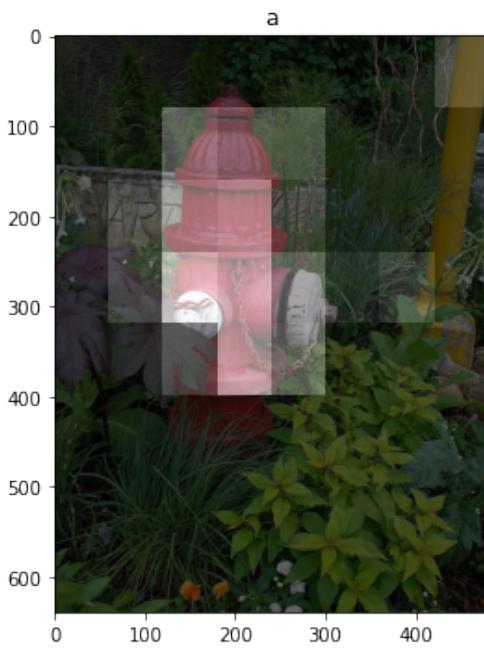


```
[0]: predict(341, encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> a red fire hydrant surrounded by a lush garden <end>

Prediction Caption: a fire hydrant <end>

341 /content/train2014/COCO_train2014_000000143339.jpg





[91]: predict(8800, encoder=encoder2, decoder=decoder2)

Real Caption: <start> a man holding a piece of luggage with a kid inside of it

<end>

Prediction Caption: a man and boy sitting at a video game with his computer

<end>

8800 /content/train2014/COCO_train2014_000000283548.jpg





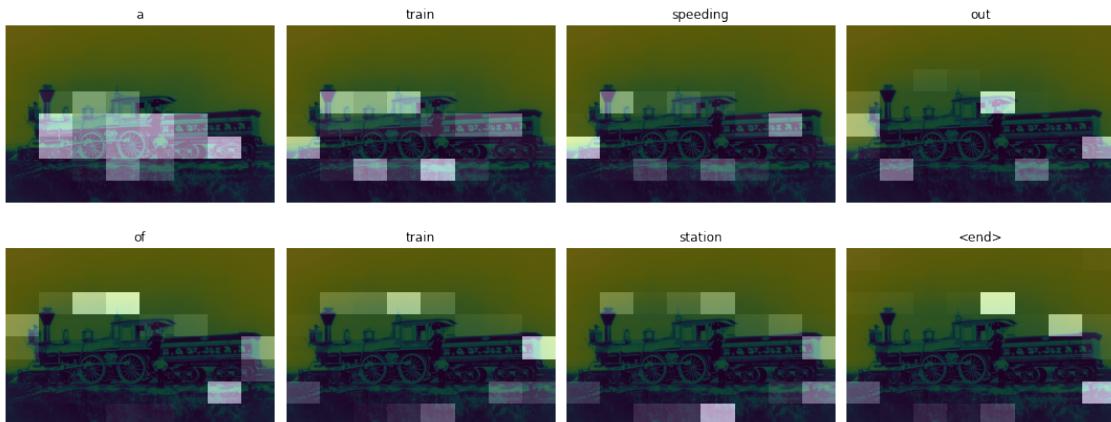
[92]: predict(6607, encoder=encoder2, decoder=decoder2)

Real Caption: <start> a train with two people in front of it during a sunny day

<end>

Prediction Caption: a train speeding out of train station <end>

6607 /content/train2014/COCO_train2014_000000480482.jpg



```
[93]: predict(326, encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> several dishes of various types of fruits vegetables <unk>
and soup <end>

Prediction Caption: a restaurant table topped with plates are layed along by a

table near plastic bottles <end>

326 /content/train2014/COCO_train2014_000000419666.jpg



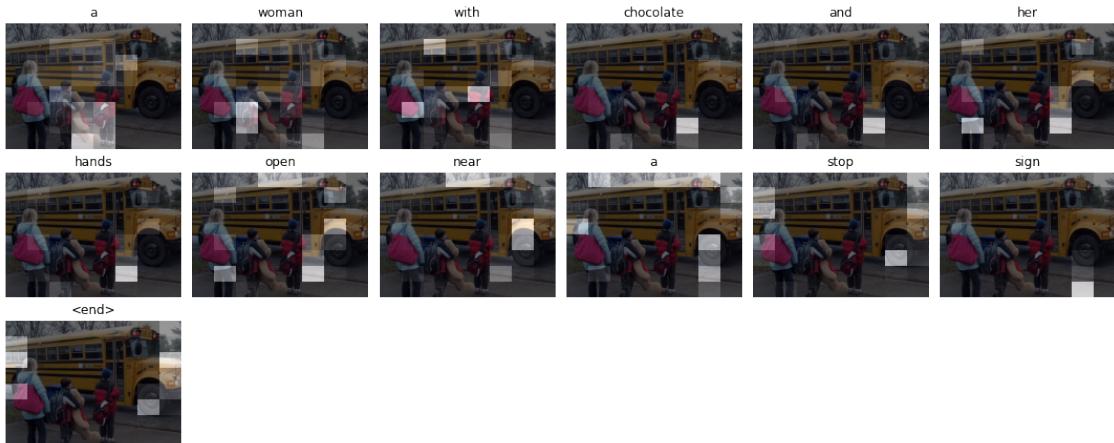


[98]: predict(8834, encoder=encoder2, decoder=decoder2)

Real Caption: <start> there are children that are boarding the school bus <end>

Prediction Caption: a woman with chocolate and her hands open near a stop sign
<end>

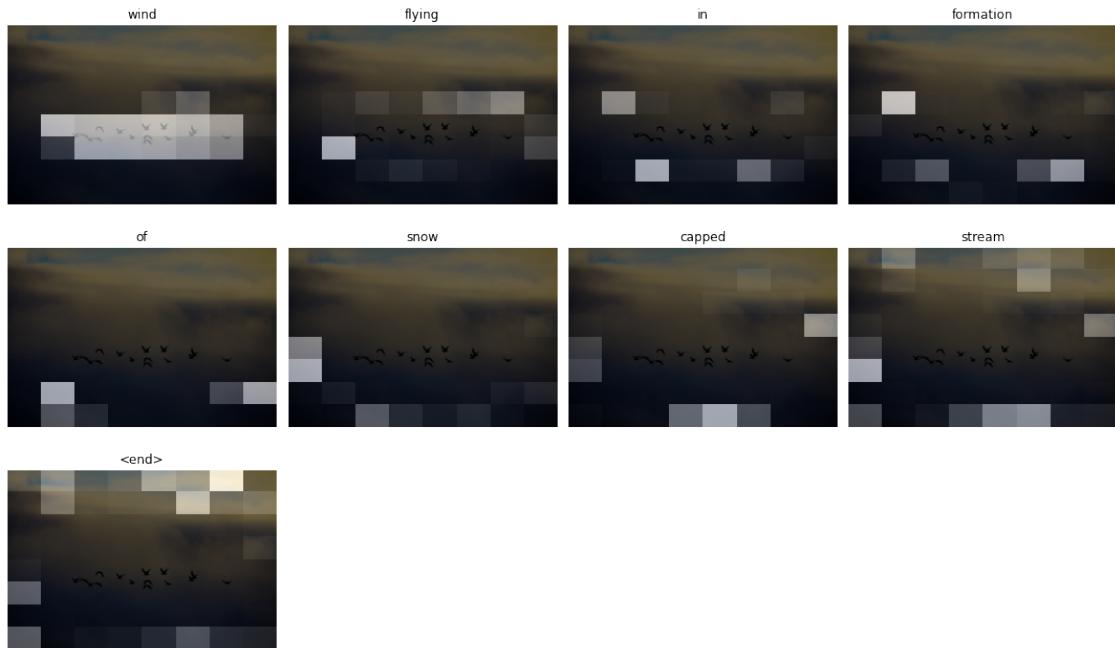
8834 /content/train2014/COCO_train2014_000000205283.jpg



[103]: predict(5863, encoder=encoder2, decoder=decoder2)

Real Caption: <start> black birds are flying against a cloudy sky <end>

Prediction Caption: wind flying in formation of snow capped stream <end>
5863 /content/train2014/COCO_train2014_000000514356.jpg

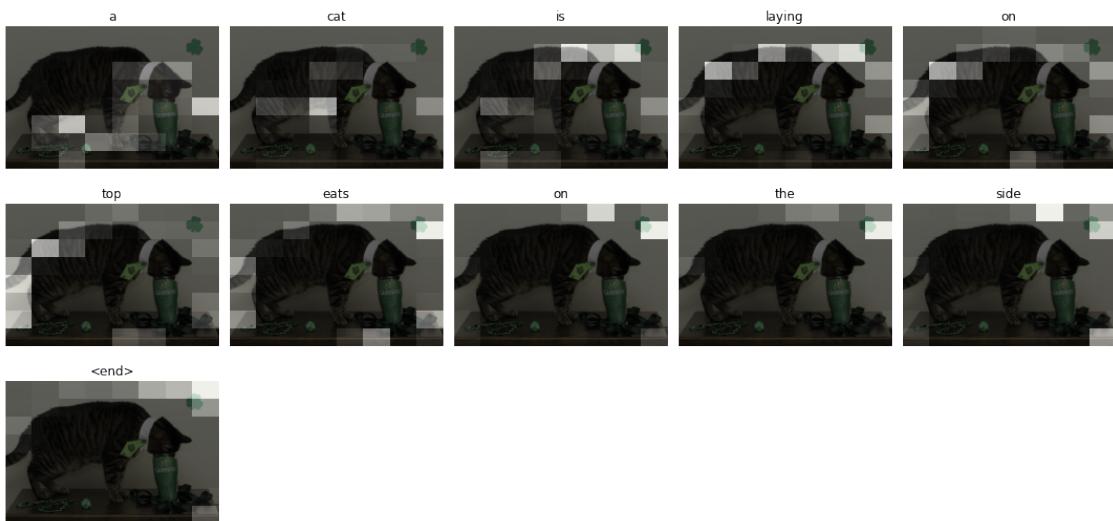


```
[109]: predict(2722, encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> a brown and black cat puts his face in a cup <end>

Prediction Caption: a cat is laying on top eats on the side <end>

2722 /content/train2014/COCO_train2014_000000276514.jpg

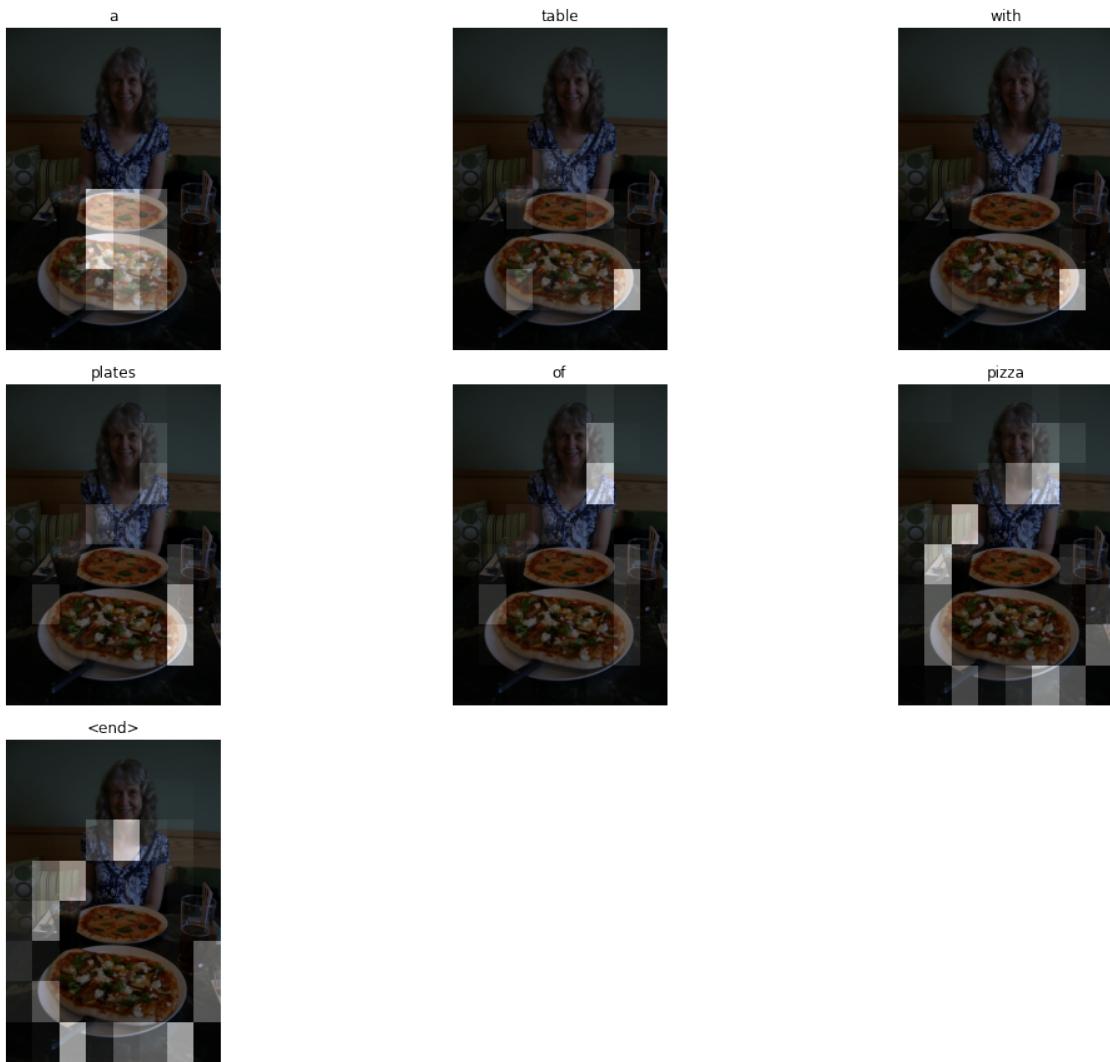


```
[114]: predict(6739, encoder=encoder2, decoder=decoder2)
```

Real Caption: <start> a woman sits at a table with pizza and drinks <end>

Prediction Caption: a table with plates of pizza <end>

6739 /content/train2014/COCO_train2014_000000081201.jpg

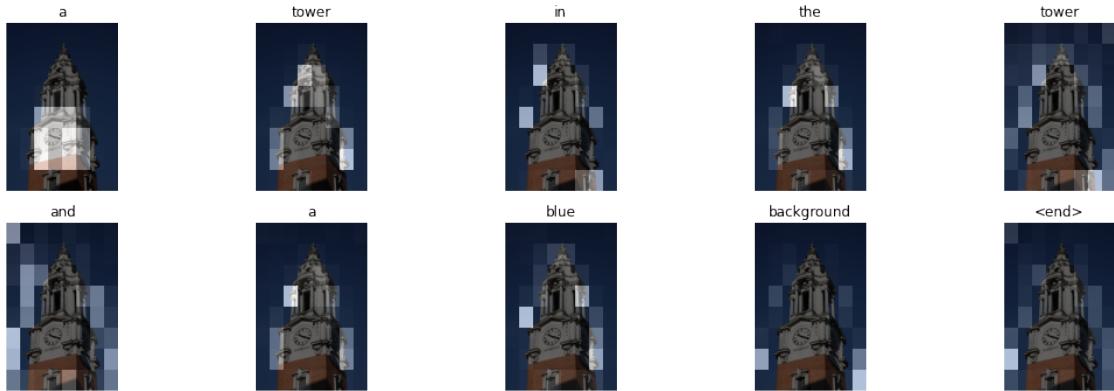




[0]: predict(id=884, encoder=encoder2, decoder=decoder2)

Real Caption: <start> a large tower has a clock towards the top <end>

Prediction Caption: a tower in the tower and a blue background <end>
884 /content/train2014/COCO_train2014_000000576608.jpg

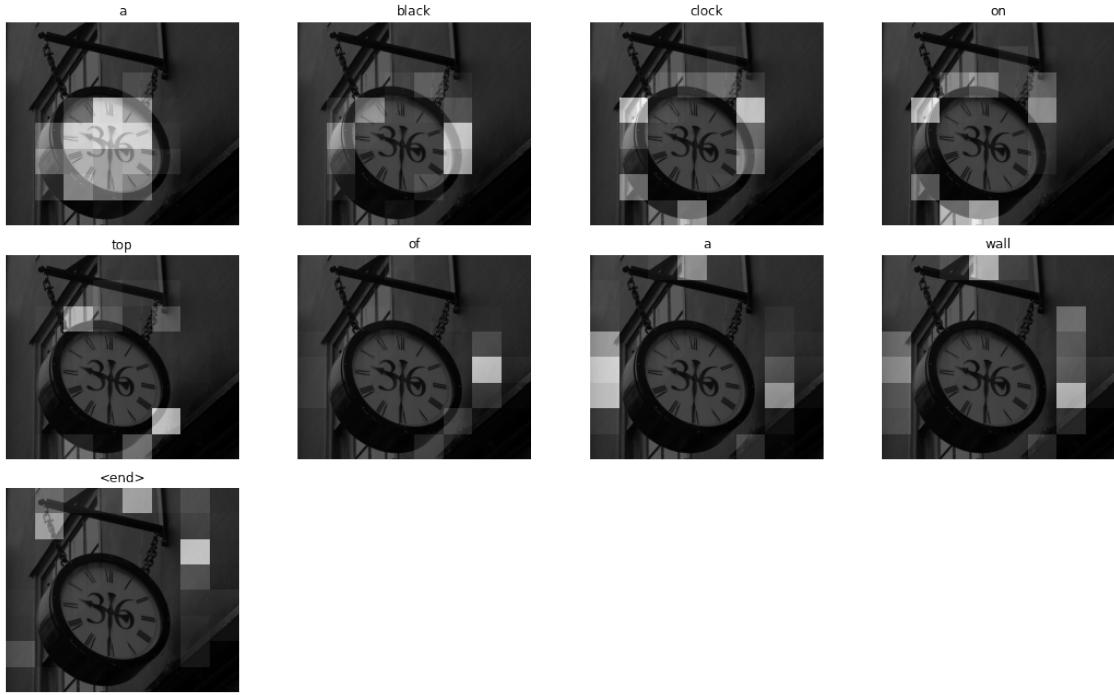




[147]: predict(id=5555, encoder=encoder2, decoder=decoder2)

Real Caption: <start> a round clock with the numbers 3 and 6 in the center is

hanging from two black metal chains with a time of 9 31 <end>
Prediction Caption: a black clock on top of a wall <end>
5555 /content/train2014/COCO_train2014_000000160589.jpg





12 Testing on other images

```
[163]: image_url = "https://www.tensorflow.org/tutorials/text/image_captioning_files/
         →output_9Psd1quzaAWg_2.png" #@param {type:"string"}
image_extension = image_url[-4:]
image_path = tf.keras.utils.get_file('image'+image_extension,
                                      cache_subdir="/content/",
                                      origin=image_url)
predict(name=image_path, encoder=encoder2, decoder=decoder2)
os.remove(image_path)
```

```
Downloading data from https://www.tensorflow.org/tutorials/text/image_captioning_
_files/output_9Psd1quzaAWg_2.png
458752/458338 [=====] - 0s 0us/step
Prediction Caption: a man in suit and a military smiling <end>
None /content/image.png
```





```
[167]: image_url = "https://raw.githubusercontent.com/veb-101/Image-Captioning/master/  
    ↪test_images/laptop_working.jpg" #@param {type:"string"}  
image_extension = image_url[-4:]  
image_path = tf.keras.utils.get_file('image'+image_extension,  
                                    cache_subdir="/content/",  
                                    origin=image_url)  
predict(name=image_path, encoder=encoder2, decoder=decoder2)
```

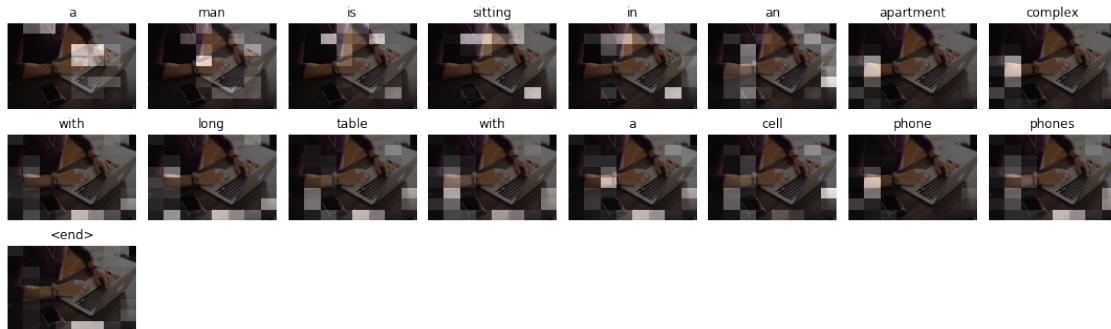
```
os.remove(image_path)
```

Downloading data from https://raw.githubusercontent.com/veb-101/Image-Captioning/master/test_images/laptop_working.jpg

8192/7003 [=====] - 0s 0us/step

Prediction Caption: a man is sitting in an apartment complex with long table with a cell phone phones <end>

None /content/image.jpg



[0] :


```

image_extension = image_url[-4:]
image_path = tf.keras.utils.get_file('image'+image_extension,
                                      cache_subdir="/content/",
                                      origin=image_url)

predict(name=image_path)
os.remove(image_path)

```

Downloading data from data:image/jpeg;base64,/9j/4AAQSkZJRgABAQAAAQABAAD/2wCEAAk
GBxMTEhUTEExIWFhUWFx0XGBgYGB0gIBgYHRgYHR4aHR0dICggGBolHxgXITEiJSkrLi4uGB8zODMtNyg
tLisBCgoKDg00GxAQGy8lHyYtLS0tMC0tLy8tLS0vLS0tLy0tLS8vLS0tLS0tLS0tLS0tLS0tLS0tLS0
tLS0tLS0tLf/AABEIALcBEwMBIgACEQEDEQH/xAAcAACAgMBAQAAAAAAAAAFBgMEAAIHAQj/xAB
FEAACAQIEAgcEBwQJBQADAAABAhEAAwQSITEFQQYTlFhcYEykaHRFEJSU5KwSNi4fAHFTNDcoKi0vE
WJGOTw1SDsv/EABgBAAMBAQAAAAAAAAAAECAwAE/8QALBEAagIBAwQBAwQCAwAAAAAAAECERI
DITETQVFhIgRwgZGhwdEy8FKC4f/aAAwDAQACEQMRAD8AU+k2PuPcdkwt6JUjNbbtDKoKka6SoPjQnA3
M7t1mEurmJKi0rgCeU04AAiRrzPhXYmdogsTG0x8tq0PWRMA+Gf4+xFnk+wcfIudGuJFFW0cKyZULMUD
sWYt2cxY5RCg+z09RdLL16+qC1h78gtJNvYFTEQd8wWmHE4t0Us1swP31+VLuH6c2XP701eaI2UHQ89D
oPGire5tkLnD8JjDdVruHvtChRnzAAQBJVgFHdTNxBpwtPaVRbJuC40kAP2co70jTkJgcvSr+B6Wp
d7CWnZ94IUGBvBLQazi+IxF6y6WsJczkAiWtgaMD9ueXIVr13BS0e423c6wPcWAwb7QaH5EgwJ1nbmD3
Vd4ZhXuEIiFmMgZA0mSIkAaga++qy9G+JzPOR2SMoVyrJEzEZo3189aZuF804ha/aCxbRh1MWZS4wXQK
WYwQNNJA08qt1YknBi71ck5SZE/uuBHMCR39+1WrWIvWVYK95D9Vhcf skaToSG17xHdFNt7huJG62b6x
sRkcDuBiKD4wpqtxbtiQVh11TJB9rWdVHdUbZSqIbWLw10gvYs9YTJdusLM25Y1d7U6yw3q9Y4Ph7rHK
mEZuc0+YGBus77d1U8TwbrA3V9UzGSDJBEsDt+Ia9/hQjFcPdIFxSDE9oH1hh6bGtaMMGJwzDHKeq61
9rSlhtuZJMISJn0pebpTf65gt91XP2IYABZ74mor5dhHWMQAQA5LABjqA26gx3cqpcoWLJeVc2UtoMyK
4M90ghiTbKnUmkCkx14fj7mLDG+bmItiALJIMGWoaSBMDTXv2ogiZRFrCXAOahrYH/wDWvuoRwrD3sNc
YG/YEzmOQiHBiAitljVu606ieL4uwQkXrTEDzbZJpkCr610rd/2PdIpCQ4gyume2UydooYJZWkTIaAA
RMc4G1VOGYIYm/cuNeuG3uACOToInuA13mhff8fdZwzG3GTUsBMSOLBnLsDB763w0MuFOrRUZdSVVRzM
xCrIBJ3nnVIzUVQstKUt9w3wjjuMtB1s2GIzSub0xcZhqCWhNNdN8vjWcVxV+yi4i/h+rDF11yQSVJkC
CwOp5/VPPYxhtkqtloSYmRIMHQkRJ00mrG04pbxFrqbmCZGQqwZihVlaU6vNrGr5ioPKg5q01yDCrT4
AOLsWMNasE2uvuXpZ811wFntQMhH2gIrfh17B4i4lq5g0XrGW2biXbspngBocsDl0me6psXjs03t3VXK
fYKo7DcEAFTzjbuqDh/SHqLiHqrPVE5utawhYAlgCAI2I2j8qpKqFSd1jF9CbYYgXGQgkEEA6jxBX8qH
Hom2cjr1MakAnMBY7JH61txvpVibZRhcBa6WuPKjYsIAB21zVxt9Lr8yCjTvmtrt7gaVQr/I0pJ8FD6X
YRj+3xBglYFtI00v1wTt3VNea6dbKvcQri2dQfECd/OrvE7aLaw+Mt2LYNzrLdxEtnKHR5zwrDJK0k6x
PnVbjQYphrYKB87XGUTGpChSOYhTz+dDYZJ1d71xVJuWbijmSjR5yQQKiTFW2kdmT3qPOIq3xDEWcLey
CyGbIuabjkAzCSuVR21gga9xqLDmxebJ1WSCAWViw15w+/wAKDiu4VvVdzRbQymEU6icrMNg8Ec68Ns
Ror7jZgw2Pdy1onf6HsNbzb4EMh/VfjWtzh14C0fuLKc+g8RoPMnnS0uxrfcrYkFy01sSNV8dvhW2Mw
padV9tt/8K+X0o7qYiSRbYqNo198SK2uumUEvDg6hkI1MSTlnupGmuDWVMZg2BEKTbOfD1qriLbA6q4
OHLwohcx4DBQynSznQeBmINeJi1YSYMGJBNOXgNoGOI79h9XwFZrtCIHzNZsdT0Y6gw8KsYVY7WSQDr
I08jQgcPx5XP11iNtQRJ8BzqNhxECoSw5HdD71WvY11DptaUiesyEoOhnZvdQRpI0VtdzHia53bFpZ/Z
1YGp6wx75g+1N/Sbo5xDF1CzYcZJiC43jwPdQzDdD0J2xCX7QEZGc/qlMpVsb8fqH0jPB81oXjcVFCSV
bGfxMBEAsNSzRI3Bp14rfNuxdfaW2Sv+KNPjS/wTo/is0iANaLFv2jEt0pP1e8+Y9av8c4MHuG+XfW2
LbqumA+bQjtAzGxFJFuT3Karil8RNt/OocTH98hgwJtJoJ8quYf+1XiUgE2D52vkwojiui+AGUBMjEE
sXvNE5ojtXB4660Zwh9H/AA4ssY1CCwELeg+0RpLhtR19/hr0SjjycZJk3Qnp/icRiFt3rViDPaVXHg
IlyPhzo1i8KrZ1YSJIg+BI/SrN/8Ao6sohGEvXFuzAd3zC3njMOCCWhRGo2FQpYNsC2zZmXslvtEbtrr
qZNRk0ysUxRx/A7aNE1E7x9TxHMelCcTxC9h3a3di6oMa/ofnTdxth1oB0z4P1SrcLgloGXmD1E89RI+
NNB3tIE1W6FXi69sXbRyi4JC6S0Rkbby0+r0HW7ZE3VU0D1tuYIEGTvoIgS04VWwgK3UKDOZnLHOTp+s
0yvhcTiCFGHCwG12upCBkKk8o0un8kaSx2Qsd9y0/SxV0t5YIB1tmTI710UwZG24NRYzpJYya0+XFQxRh
HutvBjXWP0rfDuJphrFpWRg02odBmVypkuDJJBzR5gj1Qi70uRXzpb5y00P1tSpeh79iJe4yXua5ifZ1

YiOUaEDnFW+051KILsxo2Vz0YxI0sxsBPcaesZ02tFT+wBldJMxI/w99c74VY6+6WYEIAbj5SJgaQOUkwNRzp4xaZnNNbk6W3tg0/Wr4szjX3609vgkxaAXcd1DIg6tWHthRm7UwdQDG4pdRMC2jWsQ0d98/8Uw8HfBvbuW7Vlme1bNwLeKkBvZQ3a3GjSSe6m1B8oWOpGmmgPxPorhb0e59LLXEBfKTa1KiQIJB5Da17BcTz4m3mMKWUbbAa7ESPQiny1za8kDhtgyYFwXFGWJ8wRPhQnHdF73a6vBAP9V0uWjB7xA1pKQ6m1wGcC/YZrvV9XmlCwGimN58fLehbcdRLuQWLTKxzAAmRM5TI0nloZofTzcTsoetw56tVMtKDKo3Y6mYHdG1VMeMYG6wsdo6fSCFPDXU93gPfValqkTw8mNeP6QrcW1ZuJaUKWkm1cDZC0AyMgnYaabUms+Iw91utEP1KyCDJU6kQZAgu3o7YxhEQ/DNt1YD0man4pwdcaRcuNNwdmbJDDZR05CkCNZ1ApHHuMpKq7iraxQuGXuWyR9plXSdtSs/x3orggrgM3BAW3tBjBYaDXMQRMc6st/RwF9nEH3j5UPx2MGG/YF1JVxJnmBPuOceo8KWMkGScnbYUXpRfYmLUMT9Ro0u++1Wb1vEXbZuqlxCGAZYcTmBgrPiGB/y99V0Fv1v7ZsotgGCD7TLyAiTqIMVZsW7gQG5iT24KnNc56jSas0qtfwR+V71IYDEMZYuT+8CfzE1RuYcnf4H9D86a0Eccsocj41XYvIJLagAddrXcGhNjAvCujaSRqwGxI5nvFRjQw34R0z0VGMWuXygLd4dP2W8Nj8f0Jqlf4WdgCD9k6fA60xHCToJXU5R2hqZiBrqajfUAvgw7m1Hp03pTKUZ80yUoS/j/kq+4u9ReGmvvNZR0/5vRzHxBryjQp0w8ZsffJ76wcWw/3yfiFRNwXC/dr7yPyNQtwDC/Y/1n51L4+yvyLg4rh/vrf4h86w8Vw/39r8a/0qH/TmGP1T+M/0tk6J4SZIJ8M9b4+zXIurxSwf7+1/7F+dBum/SQWMMhw+IQ0bymUZSRCsZ5mBFxsR0Ywrakt4duAB4DYUNxfQzBXBBBnWf7St8VujW2K1vpfauW1bEZ718MSXgagsTM66+1G+D90cKjqSLNjyICDBMkgE5dBJnerS/wBHmB3Y3P8A2Vf4b0TwmHYtzBiIkvnOrTPVYvTQe4d/SLNk3rpjKdQoHagM0LrziJnnWXLwVZJnT36UEv4XDpdW+dbiAqpnkdwQNG9aJG8AvW4lgidZ0befP038qlux1secNw5uN19wfs0PZ/eb15gfn60r9LeKJfuZerkrnJIA8BG4H/ABWvSDpjcu910xbGgA7vShWFtE5Hui4bbsUVbQD0zD1G6iSBMHXYVWcd2Tby2RDhLRVutDtbtIQhvZCQJ5BRq093qYpkt4brequ4q1oYaxhpzG4SVVrz9kMlnVCZ35QPa9w/B+pdHvK13FwVtoqxbw4Vc6m91JD300D1H2p10IsYnh18ubq4lusbfMAddJAMAhdNqSU7YyjSDd/FHsZyJ0UbKPJVmAaw0dDcdaB0wPoKGXsDjAT/3CHO+a1BheG4644UX18T3D8NDH2NfocMFh16izKrsx2GxaRQTpDhszBdw2yIkKqajcb860u4DiU9nEJAEDyH+S1fjvQviN+4bj37RmN3YagRySKdVF3YLfg0Laug60YGoUAdhvCkd+10XQjBXHuFuvAULBQqGzSNdTBWu0W+geMn+3sgd/Wv8qP8L60YnDgXLWPVL6nT9t+z3iCpEns z6009XJVQkIYux14/w029+4y3LqjNEI8ARoYEGNQagudG1ZATi7wI0y5oY7DeNe+hbcJtMM13Hk3Dq5V1gtzI8JmocXwvLZJsYxLnahka6A+WNSkSWbYAd9LG/P7Bf2Jcf0QR1dPpd+SsQzgjUGMwgSPCud8d4ZiWYDq2ItrlkK0TMkrI8hp3eNWMT0jxCXGFm6cohSQndBqdd9Sa3sdLMXJK3WIInPlttVYryxWwJ9DvACUbu2189dY8dqNdFnBm1o5YuLBFxsoJkASRJGs7Ci2Pxd+7gwwYFyHJKqCQFcAhtzqrEjyEurdGrQONsoSDmdQZBGh108GdaD1Gth1Cn8jomH4Jw917V60rBZIN4wTzCkn3TvQ7FdG+Gtr11v8A9p/3U1X0jWF+5H4m+dZ/0hhsgc2gVLZfbaQfHWpZL2NT9HMcVi1sXhs2srWUYEEMSDmWY58830rX9etcVbZtLBjKQD08aSx30g9KFYhLbAkmdIK5e45sw90zQ5bu8HYfEUkpNukzph8IW0nfklxtzNfN7q8q55ImIIg7Tp5UfvcWR7VpLU+3m0usBwSpqizB/0gN42+qMxn8jPtVrwa2ouj0wAKHU8pUgT6ke6tGVqhJrGaa/wBsZ8Ldt0XuLZLBFz5c79mCJYazpI7/AJX8Jh7N8MuRrTEKAA+6tAmGEg6mDJ1FKGBxbKGgnVCu/JhRFMYzBZcgqISDq00NoPhU61HhnQtTTm3FxXev0CKYNiBABAeEHcefOsoXicMxZiqKwk6kgE9+mU86yr9Y4+idnyk1PZCASQWbu2X1I1Pwrwmoy9IY3Z/TwFR9YeVbC131IABRARC0TvUgUCo7uIA50LxHECxy2wWY7ACsYI3sUq7mh/0p7py2lnvPIeZ5VSxLWrWuJu9r7pDLevJff60KxvSW7d/ZYZ0qSDopgxzLOYCjvOnmaZQbFckg3jcdZwYkkXr/AHckP6fn5UuXcNjMaGvsItKCTcs21A5LPteSyaL8A4BDowspjXm1u3FmyREZmykXDrMe4HeiuIvYaxipYfSMS8vpHWsqn2BpPsxrqI+rzbJR4FpvkDYDo61q3axQvLaUS1y7i7Qyxp16q0TqZ+sxE6R3VK2Jti/ikwCBMSYNzF3NWYvqwtjXqxE8hruNJqLiCHEm79LN11crlgN2Y6zYAHKII/WmKyc4mIB5kH8qlKTbHSop4XhVtS2RT23a4xJ0rNudSd4FS3baJvJPLU6T67VNevqui+tUArXGyqMxP86nkKWhiJ1LkBQSToACa04bBC0uQElj7Zk/hFa2MMLQIqg3ToXIOXvA+dUbyYyTlezHiGn9adKzN13E4G2+jAkeZ+dC8TwLD/AHf+pvnUN/6cP7yz6fxWqLYTiDGetX0gf/NMk/IrfoJW+j0H+wfxFN862PrnDfZb8RoS2A4h94fR6C9JFxqWij3G/a9gS5iNCwnaYIEdzHwpv+37gSviI0Do3gnjLcUkiYF2SPQTWu06M2VtRhxb68Gc73DEa6EBZ7hvXKrnR/EAf2Y9CP4Vs0H4jQmyQRzGUfEGarg+9kslynuGeM9GMTaT6RcSy6Bu0tu5tyBggN103E0Cu4y2w0sIp5N1oMa9x0vdXUBYxGIyDfYqrWwLz04ALR1g8yRBafEUk9J+gF7CFTdVQrt7VtpULJzLqAc8ZSBtE710Sx2XBW0q5u5chzopePOUGQBnkgMSSFJ+q0zHjM6DxjzEYjCtdtXSt3NzfOsLbwTp7WXUjSg/C7/UXLSWndrBabhZFDKAZ0I38d0VP3/U+G51vw1GEa5RfW1c634VFJ+mafdt8PnVZumiT/Zt8P8AdRJ+kuE8fwVAvSzCJnyhSxtsnatn6w8t5inaX/EirvkR+mtsLiQFRUHVkcqgAalidB4mgVjAu5K20ZzBMAchvTz016NX715r4a0EyLu0EALJnTTWaE2+CXkRLoZQGBAAch4IIJiNAR+YqEU70yepFwruKV86D31qVIY

jy/Q1exuGdXCB3IgST+XxraxaZnysrA+05Hu+01VUWkQ1KMpWRAkAEgwdJjSQNp7610K1T110/qDREm
 4Ee209U8aEgwRsw+yw7xvtQvD4Niygjcj4kVpLYGi6nZ0tej+H75mTy5mayjGJ6E4t3ZssZiTcqiHl2q
 ythE3UkHgh09SrA2q05eC7mhuI4pJy2wWJ2jU1hAldxAx0LxHE50VAWJ2AqLEYXKM+Kui0vJZ1j5Afp
 NCr/SoLnVb2cp0mdhmc+nL1nyp1FvgVySCmIw4ResxV3q10yDV28AP+fS13HdInMpYAs2+cHtETEu521
 8Y1qZuCYi4ys6m/fuf3QcSi5SSQ959rduQooyDB01EU8W8NasYcpiVs2xGZr0HntFdYZ2Idtu4bak0cox
 9gpsRMD0dNy2bgZcRckRh7DqWIJE13mEAGv6ij+j4TzW+DniFm2FSLhw+H0rsNusuE5mInaQuu5qPhnS
 PrMIfoeHTC2i7IoWAdFXtHKvtHMe86b1DwnhLpYVHYXDmZmYySzM2g1mTtJpJTb5GUUj3inELuJwdtbK
 rh7bqCLayqhNezIAkbeBqzcwys4dVJIUrPP3kzEN5anuq+0H5diuw1BA5bb6+cVCmHuEDL1gjSGGvrzo
 Yy8B2MS0qe0ZPMtt76zEY6dtBWlgLhEkgDlzLeQUEkeMRVf6RYtN/3DHNythG+Jjaji/BrXJPgsE13Wc
 qDdj+neaK28qLltCBzbm1L0L4s91oTFgLyQWSI+BJqvcGM3F5iPjh/wdNNNgLmNwq07cikjEYjFj+9b8R
 FDrgvse1cJ87o/VqPT9mz9HQMpY/wqTqn5N8K5u2Hud4/Gp/Wq9xLn8mm6XsXqeh+xHE1XNOLtgIe2ZU
 5JMdr10156VMMRwJj1tzE2WuECWNzUmANh5RtXPeL80b6PcTDgu8r18HUXqFUDWBLCNzrp0KTjgryme
 qdTvsQR76ySHknGt963R3zBcL4LfJFvFqWMmFxWvoG02vdWvShoRhUw966b93q1AILsrgsYACnTSSJJP
 5VxFMSSsXGPqg0m3nT1/R1hVuWsRYZ+qs38uS6y9jrEdYnw3EzAJH1W1NpA0tNyl8dnyjz6U9sFXdrnV
 rktrMZIBAWIEEEzHht3/SRcLWLk1Sxk0RrGqkTpz0IhrQTpX/AEfvg1t4psQLttHQXBkK1QTGf2iX1IE
 aHWe+ouE8VsKzvcxt261wgnrQTEtosKIGu21Jsk1FbejbybcnuAUw3/hH+r5142GH2CPKf1p4XjuE++X
 1B+Va/wBd4XX9snupylqPwDFeRDuYRFH4f0tbvAz1PXA9mSqjmWj14fKnhuKYSNbLo+cfrW68WwzBVN
 611UyFLLAnuHKs9RvsFRp3ZypeK0gKi5prIMhMqdx4Gvf69ufeL6haqYPAdb1jZo7ZM+E98jv01XsLOX
 a/0RgAIMsT6DU8+/1X0opyxR2vVkoZv8An+q/crx0JuTJuCfCB+RqH6e07ie/v9Zq/wD9I4hfa7X+Ekj
 0qN+h2LaSqAgci0H4xPpTrTmuzJS14NXaIfprcw8yf8AdUmGvksgAUQw28/Pasx/Rm7a3Xk0egbXMcy
 1ggiRpI0pqTo7gl0Mt2zMFZ02+Uz+tCndMfqLHJ17BZ6R8VdQ6paKsAQZTYjxNZXnDOD4nql6pzkiF0t
 bAkC1mvatcDjqXop4nBhFNzFXsoGuRdSfDz/maDDpBeuHqsDYyTzUznPiSdFHnMd9XsVOUZVZ2ZsXfbW
 LctAMsJzuT2BE6z79qL8Is4ixZjEPH7A01uyAp3+2Q01HcpPjS3Few02JeFt2BixZxLXL98vle3ak5Wk
 e24ksNZOXaDrTfxvg+G+j58Wy4LDje3aK5nhgcPIXUkqNAC2+1DsT0qyY21hcPYtKlsG5c3ZpzTMbtC7
 sW3r3inD/AKUv7V3uCQdwJI8QJA15UspN8hUaLHEcaPofU80RrCXEUowMMQwBkknNmynckt4iq2H4b+z
 VWB0EEZzPr3zqTrrNXrFpbauqyooqqJIMKqgDU7aAb1Hcx06IsfvH9B/I86QYmw+ESo0Ac0o02g1MT6mP
 E1z3ph0svm4Ldi5kVe1KbtzUk8hHajxHkGhpJiCmHuNm7RGWe7MQDHdoTtXNb5VwYsWYyTC7n3R7qaN
 KRSMG40uQnh+k0JDyMS7jScwB7pMe+r7cfu6ZcS4BEZeqQgCI2J/W12zhSDMuB/hkH3VJnI3PwPzqu12
 Tx1KpHTug3SEuty3iCD1SmDESoloI56EkevdvHPBxpB8XAkKjAFZjsHx5NIbXb2aV0j3E0tXTm0ZLgCt
 CtIGu098ketMtnF23tZEsrz7PtEAXm1p193Kk6nKZZ/SynjJbbU/uv7VEuDwti+zDDYgMygNBB2Pcw0M
 bGnq3fhGKX2WA/wALkfKhSm97FoXAAWCTpMw0z4LUbYe99Zbh8w361aLclyck4qLo0W8PjxrJP8AmQ/
 nVn6bi1HatzH+H8taVXttzU+6vcGqnM7yUTcDcnko/U91NgLkX8e9242Zkb8J0+FU7FtusSVPtCdDt0v
 wotb4thwQvONZ2iJpxM1P/W0FE58IR3EZh3yNH15fGjTRtvIWuLzBhv3qIYJwbMTBIAJB5CT+poZjk4Z
 iyOuVxlBiGuAK0Z5qPPwpU6X9G1whzYe5dZIBILCVzLmHsxpE7jXKe41nqu01A6N72dGXhSk5usUnxKj
 9BTL0e6MpfV1xKB7Q9hA/M7t2DthImvnHD46+0C3cvkkwArNqe7Rq6R0T6f47hdsjGWxv2SwALXAWQwd
 AZ0YGOe0VKeqpLEbTjKEsk90dA/pcxk4J0E6XkTzCgEn0JA8xxFwK6S1t8eRfvuDauWkZLKz+zZpdpaE
 0czkbDRV7q1udDs0dsw8m+YpYaiWwzg2c3Y1Fcauh30h1j7x/evyqpe6G2fVw/wBNP1YgwYgm25UsAco
 3PduWc99dHbozaN1bILCDLMCO0fKNOXuFvx6F2Y9p/h8q3UibBnNOG3czdWikux023qxj0J0p6q1cZUT
 s9k+0frN05208AKY+NcEt4NTcRml1I0g2JA07j2t/Ck651caZge6B86EZIZw25/Bds469EnEXPfV6xx0
 +I/7i571/UUIDWiQSvubSsKd/HwrazfoggNI85/jTuTXcmo3s0+Kxj3rdosx1DWmIAMT21fTn0v+Xxpc4
 Uzdb+0c9YAQV06kd48vzohwviA6u4gYScsamdDrGmhIkpona4GpudeDJZQ0XcPjpUm0313KuEofG9ud
 glw/j+It21RFBUTBKsdyTyaNzX1MXB8ZatWURg8rMwikaknfP415T5exafgg/r11BfrzMGEQBQ04QABz
 3AHjrrS10aulMz4pszs+diz5mYquVRoZ3u0d/q0D1jvFYTHypembM6Pg+j2ANLq9+uUen6VBielf1DAd
 Se8CQPUD4CueFyfCtay0kbNjzc6QWGMtckjbsnTyEfxqRekGH+8/0t8qRr0HzvAd/yqU3Qmianvpukgd
 Rh3prjUfDDLcGtxQVgzEN2tREAxp40hZ/5BFw+LmzIZJJ0oMIjWZqc4K9i2nrNkg3Z41fUAkRA20X+Zq
 LHcRvXQof6sxCvx5UIS6RPiI/L5VqCaDbfLMpxTtINcIVetXrVYoDjywCR4TpN0+BZkJeItskDb2gwHe
 YMEug4KzdS2bwjK0R30oG3LenDBXPpWEtrLD28wU/vaCe6ADp30mKss/qXiooeuA3U61GJAJBB8QLlyJ
 99E/pKfaFI/DcJ1VsIgMCrDXHHI09157Y3vibcEFhBEGe46frSZxL6KXyo6ogJhQHOZiInMAdRJ3Gunl

VfG4o5chJXrDE6nbWq6jSGcNEZcugAAAAid9JnxpXNw2RaGnpnjcufH8/+bFu+110KtcUSQQxZhliZiU
 A1/Sor2CtMIS+g/8A2Wz+bA1pZuWgSHF5h0hDgnLA3k+1M1mLawQera8DPZDhCPHNofhVM5JXa0bpQus
 WTXeA3nACLKhcy1TAJgkkE0c6bdHrj3FuW2XJ1SI0nZ1mBH1+Z76TrTujzaDEsTmydlog8xGkxptpTj
 hbuKWwFxQy whole string continues

1tvi1+dLrVaTj2IW02D5ijjnF8oDi+zGvpDOTsFQbIIFs1CJ1M5rgJPMwY9wpW/qFe9vhRrgfHrl5msuB
 2xoQfrAiIHfpFaBbxclVUr3HY98QDrM7kVCaV7FEmrBacCXuP8+tWE4Gvcw/zH50cUMP7sVtmP038TS
 4hsE4LhK2zKgyfGrvVHuqx1n7p95rRrx8adIVsitWCzKo+sQo9TFdKvkIhJAhVn0A/hXObWMdGDqYZTI
 0+vrVvFdKMQ6sjMpVgVPYGx3qidCSVgbP0prK8zfzFe01D2b5yeZ99bCmA8CT7bfCo34D3XPev8aKkgY
 sBzXhNGDwFuTqfQ1G/A3+0vx+VHNAXYZ6ie5RS5wa53r7z8qp4jhdwgcT4Gg5IOLIcFhmulUgd5J0A
 /Xypqs27WFtkzA+sx3Y/zsKUm4Tc+xPqPnWh4Vd+waeM0hJQbJ8f0rbrJKwhnKOa7anvP8+djCdKVpts
 hHf1K/maC4zg9w7o3uNDrnB7o+o3uNT10V20oqqHq3x+y311Pr/GpPpl1vvqa52eGXfu2/Ca3t8Hvna0
 /qi/Oh1pI3TR0A3cNubQP11/20/8AQDpFZMWx1g/xBYUR4HT3Vwy3wfELuD+OI+NwcIbluQbQadybon+
 FJLvc1Q0YJHe+n3GLLImFLAm7mZ1/8aoZOn7zLpzg0n9Ks/Eby0zulq0Zt21IgfvNoZYjTw2HMnnHRm3
 dt387toVKsSZMGNPhT/Y4sgEAn3U1NjJovDB6RWNhBVb+uV7/AIVGeLp9r4GjTNaLTyVaq3c0KgfjNv7
 Q+Pyqu3FkP1h7xW3NsR4y1EaTu2u0CAZHPRIYqfphxA2rV3DqkNcZMxE/U11AERBzTvsB3Cqd7iVr0oa
 GUo4ILQN5keMetBuI9KbgzqIcEKGLCSQhEQQZAhRPfSynKKpLn9vZ0fT6cJScp0qVr2/AoXF7R868C0
 Tv3Guu1wBQNyI9JE6nx1qXYCFrmgt2zB1JkT4aA6VeLvhEJxp22CQK3Ao+vR+9GYWFkzrDjw01iPPxof
 jME1vRrJUkmJfujTTn6011yiaSfDI+HX11GG4YU7Yu4AYWQwZiwHdMr7wSR4Uj8PydZ+1JVB3AmTyG8
 +PpTVwzH11W3nZkBzAnY0UganWN9KjKayT/B16ek5abXjf8dy4mIbvqT6QfCpAR4V51HdVD1Ije8B76j
 dweXxqco06o2tjuHxrALHDuBviFZkiAY1aNYnTTxFVeK8Haxl6z60xBB2j510DotgwmGTs6tLe86fCKV
 +nbTfVY91PiST+QFO0qFTdizkXx91ZXotnur21GHDNWTWV1THNXuVBcu17WUDBXg/Ry5iENwOEUmFJ1n
 v0GOVX430bbDKGa4jZjA9qSfdG3jWV1GgXuCbdupmAUSa9rKCCyjieIqpjkSapvxQ8kHvrKyqqKJ5MiP
 En7gPT+NQXMe/fHoP1WV1bFGt105cLbknzra1brKyjRrLK6Vt1tZWuAatdqC5fPfWV1ZmKt3EHvqs98
 11ZQCVnYltNZGWPmg/pVK7dBJmsrKjPYtpkCMVMg1NaxlxTKuQfA/zNe11LFjNfp004kCBea07Sob3FL
 r+22aNpjT3V1ZT7+Sey4RWQlmA7zHKmfBqiwEUBiRr4TtP87V1ZSpb1Yv4thUXDWc7XtZVzmM67xrU4j
 xrKytQQjZ6VY1FCrdOUAAFRoBsNqoYzib3XLvByxJiNhH5CsrKwCIX/CvKysouE//2Q==
 16384/12316 [=====] - 0s 0us/step

Prediction Caption: various vehicles on both sides <end>
 None /content/image2Q==





```
[0]: image_url = "https://images.unsplash.com/photo-1565560665589-37da0a389949?  
↪ixlib=rb-1.2.1&ixid=eyJhcHBfaWQiOjEyMDd9&auto=format&fit=crop&w=500&q=60"  
↪#@param {type:"string"}  
image_extension = image_url[-4:]  
image_path = tf.keras.utils.get_file('image'+image_extension,  
                                      cache_subdir="/content/",  
                                      origin=image_url)  
predict(name=image_path, encoder=encoder2, decoder=decoder2)  
os.remove(image_path)
```

```
Downloading data from https://images.unsplash.com/photo-1565560665589-37da0a3899  
49?ixlib=rb-1.2.1&ixid=eyJhcHBfaWQiOjEyMDd9&auto=format&fit=crop&w=500&q=60  
98304/95861 [=====] - 0s 0us/step  
Prediction Caption: men with bags on street <end>  
None /content/imageq=60
```



