In [1]:
```
!jupyter nbextension enable --py widgetsnbextension
```

```
Enabling notebook extension jupyter-js-widgets/extension...
      - Validating: OK
```

In [2]:
```
!pip install tensorflow numpy tqdm pillow setuptools
```

```
Requirement already satisfied: tensorflow in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (
2.16.2)
Requirement already satisfied: numpy in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (1.26.
4)
Requirement already satisfied: tqdm in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (4.64.
0)
Requirement already satisfied: pillow in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (9.0.
1)
Requirement already satisfied: setuptools in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (
61.2.0)
Requirement already satisfied: absl-py>=1.0.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packag
es (from tensorflow) (2.1.0)
Requirement already satisfied: packaging in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (f
rom tensorflow) (21.3)
Requirement already satisfied: libclang>=13.0.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pack
ages (from tensorflow) (18.1.1)
Requirement already satisfied: flatbuffers>=23.5.26 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-
packages (from tensorflow) (24.3.25)
Requirement already satisfied: astunparse>=1.6.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pac
kages (from tensorflow) (1.6.3)
Requirement already satisfied: google-pasta>=0.1.1 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-p
ackages (from tensorflow) (0.2.0)
Requirement already satisfied: keras>=3.0.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages
(from tensorflow) (3.5.0)
Requirement already satisfied: six>=1.12.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages
(from tensorflow) (1.16.0)
Requirement already satisfied: wrap>=1.11.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-package
s (from tensorflow) (1.12.1)
Requirement already satisfied: requests<3,>=2.21.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-p
```

```
ackages (from tensorflow) (2.27.1)
Requirement already satisfied: protobuf!=4.21.0,!=4.21.1,!=4.21.2,!=4.21.3,!=4.21.4,!=4.21.5,<5.0.0dev,>=3.20.3
in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (from tensorflow) (4.25.5)
Requirement already satisfied: termcolor>=1.1.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pack
ages (from tensorflow) (2.4.0)
Requirement already satisfied: opt-einsum>=2.3.2 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pac
kages (from tensorflow) (3.3.0)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /Users/adityagadagandla/opt/anaconda3/li
b/python3.9/site-packages (from tensorflow) (0.37.1)
Requirement already satisfied: typing-extensions>=3.6.6 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/s
ite-packages (from tensorflow) (4.12.2)
Requirement already satisfied: tensorboard<2.17,>=2.16 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/si
te-packages (from tensorflow) (2.16.2)
Requirement already satisfied: gast!=0.5.0,!=0.5.1,!=0.5.2,>=0.2.1 in /Users/adityagadagandla/opt/anaconda3/lib/
python3.9/site-packages (from tensorflow) (0.6.0)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-p
ackages (from tensorflow) (1.66.1)
Requirement already satisfied: h5py>=3.10.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages
(from tensorflow) (3.11.0)
Requirement already satisfied: ml-dtypes~=0.3.1 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pack
ages (from tensorflow) (0.3.2)
Requirement already satisfied: wheel<1.0,>=0.23.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pa
ckages (from astunparse>=1.6.0->tensorflow) (0.37.1)
Requirement already satisfied: namex in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (from
keras>=3.0.0->tensorflow) (0.0.8)
Requirement already satisfied: optree in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (from
keras>=3.0.0->tensorflow) (0.12.1)
Requirement already satisfied: rich in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (from k
eras>=3.0.0->tensorflow) (13.8.1)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site
-packages (from requests<3,>=2.21.0->tensorflow) (1.26.9)
Requirement already satisfied: certifi>=2017.4.17 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pa
ckages (from requests<3,>=2.21.0->tensorflow) (2021.10.8)
Requirement already satisfied: charset-normalizer~=2.0.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/
site-packages (from requests<3,>=2.21.0->tensorflow) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages
(from requests<3,>=2.21.0->tensorflow) (3.3)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in /Users/adityagadagandla/opt/anaconda3/li
```

```
b/python3.9/site-packages (from tensorboard<2.17,>=2.16->tensorflow) (0.7.2)
Requirement already satisfied: werkzeug>=1.0.1 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packa
ges (from tensorboard<2.17,>=2.16->tensorflow) (2.0.3)
Requirement already satisfied: markdown>=2.6.8 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packa
ges (from tensorboard<2.17,>=2.16->tensorflow) (3.3.4)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/s
ite-packages (from packaging->tensorflow) (3.0.4)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/si
te-packages (from rich->keras>=3.0.0->tensorflow) (2.18.0)
Requirement already satisfied: markdown-it-py>=2.2.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site
-packages (from rich->keras>=3.0.0->tensorflow) (3.0.0)
Requirement already satisfied: mdurl~=0.1 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (
from markdown-it-py>=2.2.0->rich->keras>=3.0.0->tensorflow) (0.1.2)
```

In [3]:
```python
import ssl
ssl._create_default_https_context = ssl._create_unverified_context
```

In [4]:
```python
pip install --upgrade certifi
```

```
Requirement already satisfied: certifi in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site
-packages (2024.8.30)
Note: you may need to restart the kernel to use updated packages.
```

In [5]:
```python
pip install --upgrade pip
```

```
Requirement already satisfied: pip in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-pac
kages (24.3.1)
Note: you may need to restart the kernel to use updated packages.
```

```
In [3]:  import os
         import pickle
         import numpy as np
         from tqdm.notebook import tqdm
         from tensorflow.keras.applications.vgg16 import VGG16, preprocess_input
         from tensorflow.keras.preprocessing.image import load_img, img_to_array
         from tensorflow.keras.preprocessing.text import Tokenizer
         from tensorflow.keras.preprocessing.sequence import pad_sequences
         from tensorflow.keras.models import Model
         from tensorflow.keras.utils import to_categorical, plot_model
         from tensorflow.keras.layers import Input, Dense, LSTM, Embedding, Dropout, add
```

```
In [7]:  BASE_DIR = '/Users/adityagadagandla/Desktop/project/flickr8k'
         WORKING_DIR = '/Users/adityagadagandla/Desktop/project'
```

Extract Image Features

```
In [8]:  # load vgg16 model
         model = VGG16()
         # restructure the model
         model = Model(inputs=model.inputs, outputs=model.layers[-2].output)
         # summarize
         print(model.summary())
```

Model: "functional"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1,792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36,928

block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73,856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147,584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295,168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590,080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590,080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1,180,160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2,359,808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2,359,808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2,359,808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2,359,808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2,359,808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102,764,544
fc2 (Dense)	(None, 4096)	16,781,312

Total params: 134,260,544 (512.16 MB)

Trainable params: 134,260,544 (512.16 MB)

Non-trainable params: 0 (0.00 B)

None

In [9]:
```
pip install ipywidgets
```

Requirement already satisfied: ipywidgets in /Library/Frameworks/Python.framework/Versions/3.12/site-packages (8.1.5)
Requirement already satisfied: comm>=0.1.3 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipywidgets) (0.2.1)
Requirement already satisfied: ipython>=6.1.0 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipywidgets) (8.21.0)
Requirement already satisfied: traitlets>=4.3.1 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipywidgets) (5.14.1)
Requirement already satisfied: widgetsnbextension~=4.0.12 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from ipywidgets) (4.0.13)
Requirement already satisfied: jupyterlab-widgets~=3.0.12 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages (from ipywidgets) (3.0.13)
Requirement already satisfied: decorator in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipython>=6.1.0->ipywidgets) (5.1.1)
Requirement already satisfied: jedi>=0.16 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipython>=6.1.0->ipywidgets) (0.19.1)
Requirement already satisfied: matplotlib-inline in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipython>=6.1.0->ipywidgets) (0.1.6)
Requirement already satisfied: prompt-toolkit<3.1.0,>=3.0.41 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipython>=6.1.0->ipywidgets) (3.0.43)
Requirement already satisfied: pygments>=2.4.0 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipython>=6.1.0->ipywidgets) (2.17.2)
Requirement already satisfied: stack-data in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipython>=6.1.0->ipywidgets) (0.6.3)
Requirement already satisfied: pexpect>4.3 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from ipython>=6.1.0->ipywidgets) (4.9.0)
Requirement already satisfied: parso<0.9.0,>=0.8.3 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (from jedi>=0.16->ipython>=6.1.0->ipywidgets) (0.8.3)
Requirement already satisfied: ptyprocess>=0.5 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-pa

```
ckages (from pexpect>4.3->ipython>=6.1.0->ipywidgets) (0.7.0)
Requirement already satisfied: wcwidth in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages (
from prompt-toolkit<3.1.0,>=3.0.41->ipython>=6.1.0->ipywidgets) (0.2.13)
Requirement already satisfied: executing>=1.2.0 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-p
ackages (from stack-data->ipython>=6.1.0->ipywidgets) (2.0.1)
Requirement already satisfied: asttokens>=2.1.0 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-p
ackages (from stack-data->ipython>=6.1.0->ipywidgets) (2.4.1)
Requirement already satisfied: pure-eval in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages
(from stack-data->ipython>=6.1.0->ipywidgets) (0.2.2)
Requirement already satisfied: six>=1.12.0 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packag
es (from asttokens>=2.1.0->stack-data->ipython>=6.1.0->ipywidgets) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

In [10]:
```python
# extract features from image
features = {}
directory = os.path.join(BASE_DIR, 'Images')

for img_name in tqdm(os.listdir(directory)):
    # load the image from file
    img_path = directory + '/' + img_name
    image = load_img(img_path, target_size=(224, 224))
    # convert image pixels to numpy array
    image = img_to_array(image)
    # reshape data for model
    image = image.reshape((1, image.shape[0], image.shape[1], image.shape[2]))
    # preprocess image for vgg
    image = preprocess_input(image)
    # extract features
    feature = model.predict(image, verbose=0)
    # get image ID
    image_id = img_name.split('.')[0]
    # store feature
    features[image_id] = feature
```

```
  0%|          | 0/8091 [00:00<?, ?it/s]
```

```
In [11]:   # store features in pickle
           pickle.dump(features, open(os.path.join(WORKING_DIR, 'features.pkl'), 'wb'))
```

```
In [9]:    # load features from pickle
           with open(os.path.join(WORKING_DIR, 'features.pkl'), 'rb') as f:
               features = pickle.load(f)
```

Load captions data

```
In [10]:   with open(os.path.join(BASE_DIR, 'captions.txt'), 'r') as f:
               next(f)
               captions_doc = f.read()
```

```
In [12]:   # create mapping of image to captions
           mapping = {}
           # process lines
           for line in tqdm(captions_doc.split('\n')):
               # split the line by comma(,)
               tokens = line.split(',')
               if len(line) < 2:
                   continue
               image_id, caption = tokens[0], tokens[1:]
               # remove extension from image ID
               image_id = image_id.split('.')[0]
               # convert caption list to string
               caption = " ".join(caption)
               # create list if needed
               if image_id not in mapping:
                   mapping[image_id] = []
               # store the caption
               mapping[image_id].append(caption)
```

```
0%|          | 0/40456 [00:00<?, ?it/s]
```

In [13]:
```python
len(mapping)
```

Out[13]: 8091

Preprocess Text Data

In [14]:
```python
import re   # Import the regex module

def clean(mapping):
    for key, captions in mapping.items():
        for i in range(len(captions)):
            # take one caption at a time
            caption = captions[i]
            # preprocessing steps
            # convert to lowercase
            caption = caption.lower()
            # delete digits, special chars, etc., using regex
            caption = re.sub(r'[^A-Za-z\s]', '', caption)
            # delete additional spaces using regex
            caption = re.sub(r'\s+', ' ', caption).strip()
            # add start and end tags to the caption
            caption = 'startseq ' + " ".join([word for word in caption.split() if len(word) > 1]) + ' endseq'
            captions[i] = caption
```

In [15]:
```python
# before preprocess of text
mapping['1000268201_693b08cb0e']
```

Out[15]:
```
['A child in a pink dress is climbing up a set of stairs in an entry way .',
 'A girl going into a wooden building .',
 'A little girl climbing into a wooden playhouse .',
 'A little girl climbing the stairs to her playhouse .',
 'A little girl in a pink dress going into a wooden cabin .']
```

In [16]:
```python
# preprocess the text
clean(mapping)
```

In [17]:
```python
# after preprocess of text
mapping['1000268201_693b08cb0e']
```

Out[17]:
```
['startseq child in pink dress is climbing up set of stairs in an entry way endseq',
 'startseq girl going into wooden building endseq',
 'startseq little girl climbing into wooden playhouse endseq',
 'startseq little girl climbing the stairs to her playhouse endseq',
 'startseq little girl in pink dress going into wooden cabin endseq']
```

In [18]:
```python
all_captions = []
for key in mapping:
    for caption in mapping[key]:
        all_captions.append(caption)
```

In [19]:
```python
len(all_captions)
```

Out[19]:
```
40455
```

In [20]:
```python
all_captions[:10]
```

Out[20]:
```
['startseq child in pink dress is climbing up set of stairs in an entry way endseq',
 'startseq girl going into wooden building endseq',
 'startseq little girl climbing into wooden playhouse endseq',
 'startseq little girl climbing the stairs to her playhouse endseq',
 'startseq little girl in pink dress going into wooden cabin endseq',
 'startseq black dog and spotted dog are fighting endseq',
 'startseq black dog and tricolored dog playing with each other on the road endseq',
 'startseq black dog and white dog with brown spots are staring at each other in the street endseq',
 'startseq two dogs of different breeds looking at each other on the road endseq',
 'startseq two dogs on pavement moving toward each other endseq']
```

In [21]:
```python
# tokenize the text
tokenizer = Tokenizer()
tokenizer.fit_on_texts(all_captions)
vocab_size = len(tokenizer.word_index) + 1
```

In [22]:
```python
vocab_size
```

Out[22]: 8768

In [23]:
```python
# get maximum length of the caption available
max_length = max(len(caption.split()) for caption in all_captions)
max_length
```

Out[23]: 34

In [27]:
```python
pip install nbconvert
```

```
Collecting nbconvert
  Downloading nbconvert-7.16.4-py3-none-any.whl.metadata (8.5 kB)
Collecting beautifulsoup4 (from nbconvert)
  Downloading beautifulsoup4-4.12.3-py3-none-any.whl.metadata (3.8 kB)
Collecting bleach!=5.0.0 (from nbconvert)
  Downloading bleach-6.2.0-py3-none-any.whl.metadata (30 kB)
Collecting defusedxml (from nbconvert)
  Downloading defusedxml-0.7.1-py2.py3-none-any.whl.metadata (32 kB)
Collecting jinja2>=3.0 (from nbconvert)
  Downloading jinja2-3.1.4-py3-none-any.whl.metadata (2.6 kB)
Requirement already satisfied: jupyter-core>=4.7 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-
packages (from nbconvert) (5.7.1)
Collecting jupyterlab-pygments (from nbconvert)
  Downloading jupyterlab_pygments-0.3.0-py3-none-any.whl.metadata (4.4 kB)
Requirement already satisfied: markupsafe>=2.0 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python
3.12/site-packages (from nbconvert) (2.1.5)
Collecting mistune<4,>=2.0.3 (from nbconvert)
  Downloading mistune-3.0.2-py3-none-any.whl.metadata (1.7 kB)
Collecting nbclient>=0.5.0 (from nbconvert)
```

```
        Downloading nbclient-0.10.0-py3-none-any.whl.metadata (7.8 kB)
Collecting nbformat>=5.7 (from nbconvert)
        Downloading nbformat-5.10.4-py3-none-any.whl.metadata (3.6 kB)
Requirement already satisfied: packaging in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages
(from nbconvert) (23.2)
Collecting pandocfilters>=1.4.1 (from nbconvert)
        Downloading pandocfilters-1.5.1-py2.py3-none-any.whl.metadata (9.0 kB)
Requirement already satisfied: pygments>=2.4.1 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-pa
ckages (from nbconvert) (2.17.2)
Collecting tinycss2 (from nbconvert)
        Downloading tinycss2-1.4.0-py3-none-any.whl.metadata (3.0 kB)
Requirement already satisfied: traitlets>=5.1 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-pac
kages (from nbconvert) (5.14.1)
Collecting webencodings (from bleach!=5.0.0->nbconvert)
        Downloading webencodings-0.5.1-py2.py3-none-any.whl.metadata (2.1 kB)
Requirement already satisfied: platformdirs>=2.5 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-
packages (from jupyter-core>=4.7->nbconvert) (4.2.0)
Requirement already satisfied: jupyter-client>=6.1.12 in /Users/adityagadagandla/Library/Python/3.12/lib/python/
site-packages (from nbclient>=0.5.0->nbconvert) (8.6.0)
Collecting fastjsonschema>=2.15 (from nbformat>=5.7->nbconvert)
        Downloading fastjsonschema-2.20.0-py3-none-any.whl.metadata (2.1 kB)
Collecting jsonschema>=2.6 (from nbformat>=5.7->nbconvert)
        Downloading jsonschema-4.23.0-py3-none-any.whl.metadata (7.9 kB)
Collecting soupsieve>1.2 (from beautifulsoup4->nbconvert)
        Downloading soupsieve-2.6-py3-none-any.whl.metadata (4.6 kB)
Collecting attrs>=22.2.0 (from jsonschema>=2.6->nbformat>=5.7->nbconvert)
        Downloading attrs-24.2.0-py3-none-any.whl.metadata (11 kB)
Collecting jsonschema-specifications>=2023.03.6 (from jsonschema>=2.6->nbformat>=5.7->nbconvert)
        Downloading jsonschema_specifications-2024.10.1-py3-none-any.whl.metadata (3.0 kB)
Collecting referencing>=0.28.4 (from jsonschema>=2.6->nbformat>=5.7->nbconvert)
        Downloading referencing-0.35.1-py3-none-any.whl.metadata (2.8 kB)
Collecting rpds-py>=0.7.1 (from jsonschema>=2.6->nbformat>=5.7->nbconvert)
        Downloading rpds_py-0.20.1-cp312-cp312-macosx_11_0_arm64.whl.metadata (4.2 kB)
Requirement already satisfied: python-dateutil>=2.8.2 in /Users/adityagadagandla/Library/Python/3.12/lib/python/
site-packages (from jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (2.8.2)
Requirement already satisfied: pyzmq>=23.0 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packag
es (from jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (25.1.2)
Requirement already satisfied: tornado>=6.2 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packa
```

```
ges (from jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (6.4)
Requirement already satisfied: six>=1.5 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages
(from python-dateutil>=2.8.2->jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (1.16.0)
Downloading nbconvert-7.16.4-py3-none-any.whl (257 kB)
Downloading bleach-6.2.0-py3-none-any.whl (163 kB)
Downloading jinja2-3.1.4-py3-none-any.whl (133 kB)
Downloading mistune-3.0.2-py3-none-any.whl (47 kB)
Downloading nbclient-0.10.0-py3-none-any.whl (25 kB)
Downloading nbformat-5.10.4-py3-none-any.whl (78 kB)
Downloading pandocfilters-1.5.1-py2.py3-none-any.whl (8.7 kB)
Downloading beautifulsoup4-4.12.3-py3-none-any.whl (147 kB)
Downloading defusedxml-0.7.1-py2.py3-none-any.whl (25 kB)
Downloading jupyterlab_pygments-0.3.0-py3-none-any.whl (15 kB)
Downloading tinycss2-1.4.0-py3-none-any.whl (26 kB)
Downloading fastjsonschema-2.20.0-py3-none-any.whl (23 kB)
Downloading jsonschema-4.23.0-py3-none-any.whl (88 kB)
Downloading soupsieve-2.6-py3-none-any.whl (36 kB)
Downloading webencodings-0.5.1-py2.py3-none-any.whl (11 kB)
Downloading attrs-24.2.0-py3-none-any.whl (63 kB)
Downloading jsonschema_specifications-2024.10.1-py3-none-any.whl (18 kB)
Downloading referencing-0.35.1-py3-none-any.whl (26 kB)
Downloading rpds_py-0.20.1-cp312-cp312-macosx_11_0_arm64.whl (321 kB)
Installing collected packages: webencodings, fastjsonschema, tinycss2, soupsieve, rpds-py, pandocfilters, mistun
e, jupyterlab-pygments, jinja2, defusedxml, bleach, attrs, referencing, beautifulsoup4, jsonschema-specification
s, jsonschema, nbformat, nbclient, nbconvert
Successfully installed attrs-24.2.0 beautifulsoup4-4.12.3 bleach-6.2.0 defusedxml-0.7.1 fastjsonschema-2.20.0 ji
nja2-3.1.4 jsonschema-4.23.0 jsonschema-specifications-2024.10.1 jupyterlab-pygments-0.3.0 mistune-3.0.2 nbclien
t-0.10.0 nbconvert-7.16.4 nbformat-5.10.4 pandocfilters-1.5.1 referencing-0.35.1 rpds-py-0.20.1 soupsieve-2.6 ti
nycss2-1.4.0 webencodings-0.5.1
Note: you may need to restart the kernel to use updated packages.
```

In [28]:
```python
from nbconvert.nbconvertapp import main
```

Test Train split

In [24]:
```python
image_ids = list(mapping.keys())
split = int(len(image_ids) * 0.90)
train = image_ids[:split]
test = image_ids[split:]
```

In [68]:
```python
# startseq girl going into wooden building endseq
#        X                      y
# startseq                     girl
# startseq girl                going
# startseq girl going          into
# ...........
# startseq girl going into wooden building     endseq
```

```python
In [25]:  # create data generator to get data in batch (avoids session crash)
          def data_generator(data_keys, mapping, features, tokenizer, max_length, vocab_size, batch_size):
              # loop over images
              X1, X2, y = list(), list(), list()
              n = 0
              while 1:
                  for key in data_keys:
                      n += 1
                      captions = mapping[key]
                      # process each caption
                      for caption in captions:
                          # encode the sequence
                          seq = tokenizer.texts_to_sequences([caption])[0]
                          # split the sequence into X, y pairs
                          for i in range(1, len(seq)):
                              # split into input and output pairs
                              in_seq, out_seq = seq[:i], seq[i]
                              # pad input sequence
                              in_seq = pad_sequences([in_seq], maxlen=max_length)[0]
                              # encode output sequence
                              out_seq = to_categorical([out_seq], num_classes=vocab_size)[0]

                              # store the sequences
                              X1.append(features[key][0])
                              X2.append(in_seq)
                              y.append(out_seq)
                  if n == batch_size:
                      X1, X2, y = np.array(X1), np.array(X2), np.array(y)
                      yield {"image": X1, "text": X2}, y
                      X1, X2, y = list(), list(), list()
                      n = 0
```

Model

In [43]:
```python
# encoder model
# image feature layers
inputs1 = Input(shape=(4096,), name="image")
fe1 = Dropout(0.4)(inputs1)
fe2 = Dense(256, activation='relu')(fe1)
# sequence feature layers
inputs2 = Input(shape=(max_length,), name="text")
se1 = Embedding(vocab_size, 256, mask_zero=True)(inputs2)
se2 = Dropout(0.4)(se1)
se3 = LSTM(256)(se2)

# decoder model
decoder1 = add([fe2, se3])
decoder2 = Dense(256, activation='relu')(decoder1)
outputs = Dense(vocab_size, activation='softmax')(decoder2)

model = Model(inputs=[inputs1, inputs2], outputs=outputs)
model.compile(loss='categorical_crossentropy', optimizer='adam')
```

In [44]:
```python
# train the model
epochs = 15
batch_size = 64
steps = len(train) // batch_size

for i in range(epochs):
    # create data generator
    generator = data_generator(train, mapping, features, tokenizer, max_length, vocab_size, batch_size)
    # fit for one epoch
    model.fit(generator, epochs=1, steps_per_epoch=steps, verbose=1)
```

```
113/113 ━━━━━━━━━━━━━━━━ 331s 3s/step - loss: 6.1861
113/113 ━━━━━━━━━━━━━━━━ 345s 3s/step - loss: 4.5077
113/113 ━━━━━━━━━━━━━━━━ 345s 3s/step - loss: 3.7963
113/113 ━━━━━━━━━━━━━━━━ 328s 3s/step - loss: 3.4692
113/113 ━━━━━━━━━━━━━━━━ 315s 3s/step - loss: 3.2514
113/113 ━━━━━━━━━━━━━━━━ 318s 3s/step - loss: 3.0935
113/113 ━━━━━━━━━━━━━━━━ 564s 5s/step - loss: 2.9610
113/113 ━━━━━━━━━━━━━━━━ 662s 6s/step - loss: 2.8610
113/113 ━━━━━━━━━━━━━━━━ 409s 4s/step - loss: 2.7730
113/113 ━━━━━━━━━━━━━━━━ 469s 4s/step - loss: 2.6940
113/113 ━━━━━━━━━━━━━━━━ 482s 4s/step - loss: 2.6227
113/113 ━━━━━━━━━━━━━━━━ 468s 4s/step - loss: 2.5573
113/113 ━━━━━━━━━━━━━━━━ 489s 4s/step - loss: 2.5009
113/113 ━━━━━━━━━━━━━━━━ 478s 4s/step - loss: 2.4514
113/113 ━━━━━━━━━━━━━━━━ 491s 4s/step - loss: 2.4094
```

In [45]:
```python
# save the model
model.save(WORKING_DIR+'/new_model.h5')
```

```
WARNING:absl:You are saving your model as an HDF5 file via `model.save()` or `keras.saving.save_model(model)`. This file format is considered legacy. We recommend using instead the native Keras format, e.g. `model.save('my_model.keras')` or `keras.saving.save_model(model, 'my_model.keras')`.
```

Generate captions for image

In [46]:
```python
def idx_to_word(integer, tokenizer):
    for word, index in tokenizer.word_index.items():
        if index == integer:
            return word
    return None
```

In [47]:
```python
# generate caption for an image
def predict_caption(model, image, tokenizer, max_length):
    # add start tag for generation process
    in_text = 'startseq'
    # iterate over the max length of sequence
    for i in range(max_length):
        # encode input sequence
        sequence = tokenizer.texts_to_sequences([in_text])[0]
        # pad the sequence
        sequence = pad_sequences([sequence], max_length)
        # predict next word
        yhat = model.predict([image, sequence], verbose=0)
        # get index with high probability
        yhat = np.argmax(yhat)
        # convert index to word
        word = idx_to_word(yhat, tokenizer)
        # stop if word not found
        if word is None:
            break
        # append word as input for generating next word
        in_text += " " + word
        # stop if we reach end tag
        if word == 'endseq':
            break

    return in_text
```

In [45]:
```python
!pip install nltk
```

```
Requirement already satisfied: nltk in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (3.7)
Requirement already satisfied: tqdm in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (from n
ltk) (4.64.0)
Requirement already satisfied: joblib in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (from
nltk) (1.1.0)
Requirement already satisfied: click in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (from
nltk) (8.0.4)
Requirement already satisfied: regex>=2021.8.3 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packa
ges (from nltk) (2022.3.15)
```

In [46]:
```python
import nltk
```

```
---------------------------------------------------------------------
ModuleNotFoundError                              Traceback (most recent call last)
Cell In[46], line 1
----> 1 import nltk

ModuleNotFoundError: No module named 'nltk'
```

In [47]:
```python
import sys
print(sys.executable)
```

```
/usr/local/bin/python3
```

In [48]:
```python
!{sys.executable} -m pip install nltk
```

```
Collecting nltk
  Downloading nltk-3.9.1-py3-none-any.whl.metadata (2.9 kB)
Collecting click (from nltk)
  Downloading click-8.1.7-py3-none-any.whl.metadata (3.0 kB)
Collecting joblib (from nltk)
  Downloading joblib-1.4.2-py3-none-any.whl.metadata (5.4 kB)
Collecting regex>=2021.8.3 (from nltk)
  Downloading regex-2024.9.11-cp312-cp312-macosx_11_0_arm64.whl.metadata (40 kB)
Requirement already satisfied: tqdm in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-pa
ckages (from nltk) (4.66.5)
Downloading nltk-3.9.1-py3-none-any.whl (1.5 MB)
  ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 1.5/1.5 MB 7.9 MB/s eta 0:00:00ta 0:00:01
Downloading regex-2024.9.11-cp312-cp312-macosx_11_0_arm64.whl (284 kB)
Downloading click-8.1.7-py3-none-any.whl (97 kB)
Downloading joblib-1.4.2-py3-none-any.whl (301 kB)
Installing collected packages: regex, joblib, click, nltk
Successfully installed click-8.1.7 joblib-1.4.2 nltk-3.9.1 regex-2024.9.11
```

In [58]:
```python
import nltk
```

In [48]:
```python
from nltk.translate.bleu_score import corpus_bleu
# validate with test data
actual, predicted = list(), list()

for key in tqdm(test):
    # get actual caption
    captions = mapping[key]
    # predict the caption for image
    y_pred = predict_caption(model, features[key], tokenizer, max_length)
    # split into words
    actual_captions = [caption.split() for caption in captions]
    y_pred = y_pred.split()
    # append to the list
    actual.append(actual_captions)
    predicted.append(y_pred)

# calcuate BLEU score
print("BLEU-1: %f" % corpus_bleu(actual, predicted, weights=(1.0, 0, 0, 0)))
print("BLEU-2: %f" % corpus_bleu(actual, predicted, weights=(0.5, 0.5, 0, 0)))
```

```
  0%|          | 0/810 [00:00<?, ?it/s]
BLEU-1: 0.557839
BLEU-2: 0.334501
```

Visualize the result

In [52]:
```python
!pip install matplotlib
```

```
Requirement already satisfied: matplotlib in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (
3.5.1)
Requirement already satisfied: fonttools>=4.22.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pac
kages (from matplotlib) (4.25.0)
Requirement already satisfied: python-dateutil>=2.7 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-
packages (from matplotlib) (2.8.2)
Requirement already satisfied: pyparsing>=2.2.1 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pack
ages (from matplotlib) (3.0.4)
Requirement already satisfied: kiwisolver>=1.0.1 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-pac
kages (from matplotlib) (1.3.2)
Requirement already satisfied: numpy>=1.17 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages
(from matplotlib) (1.26.4)
Requirement already satisfied: pillow>=6.2.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-package
s (from matplotlib) (9.0.1)
Requirement already satisfied: cycler>=0.10 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages
(from matplotlib) (0.11.0)
Requirement already satisfied: packaging>=20.0 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packa
ges (from matplotlib) (21.3)
Requirement already satisfied: six>=1.5 in /Users/adityagadagandla/opt/anaconda3/lib/python3.9/site-packages (fr
om python-dateutil>=2.7->matplotlib) (1.16.0)
```

In [54]:
```
import sys
!{sys.executable} -m pip install matplotlib
```

```
Collecting matplotlib
  Downloading matplotlib-3.9.2-cp312-cp312-macosx_11_0_arm64.whl.metadata (11 kB)
Collecting contourpy>=1.0.1 (from matplotlib)
  Downloading contourpy-1.3.0-cp312-cp312-macosx_11_0_arm64.whl.metadata (5.4 kB)
Collecting cycler>=0.10 (from matplotlib)
  Downloading cycler-0.12.1-py3-none-any.whl.metadata (3.8 kB)
Collecting fonttools>=4.22.0 (from matplotlib)
  Downloading fonttools-4.54.1-cp312-cp312-macosx_11_0_arm64.whl.metadata (163 kB)
Collecting kiwisolver>=1.3.1 (from matplotlib)
  Downloading kiwisolver-1.4.7-cp312-cp312-macosx_11_0_arm64.whl.metadata (6.3 kB)
Requirement already satisfied: numpy>=1.23 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/
site-packages (from matplotlib) (1.26.4)
Requirement already satisfied: packaging>=20.0 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-pa
ckages (from matplotlib) (23.2)
Requirement already satisfied: pillow>=8 in /Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/si
te-packages (from matplotlib) (10.4.0)
Collecting pyparsing>=2.3.1 (from matplotlib)
  Downloading pyparsing-3.1.4-py3-none-any.whl.metadata (5.1 kB)
Requirement already satisfied: python-dateutil>=2.7 in /Users/adityagadagandla/Library/Python/3.12/lib/python/si
te-packages (from matplotlib) (2.8.2)
Requirement already satisfied: six>=1.5 in /Users/adityagadagandla/Library/Python/3.12/lib/python/site-packages
(from python-dateutil>=2.7->matplotlib) (1.16.0)
Downloading matplotlib-3.9.2-cp312-cp312-macosx_11_0_arm64.whl (7.8 MB)
  ──────────────────────────────────────── 7.8/7.8 MB 9.0 MB/s eta 0:00:00a 0:00:01m
Downloading contourpy-1.3.0-cp312-cp312-macosx_11_0_arm64.whl (251 kB)
Downloading cycler-0.12.1-py3-none-any.whl (8.3 kB)
Downloading fonttools-4.54.1-cp312-cp312-macosx_11_0_arm64.whl (2.3 MB)
  ──────────────────────────────────────── 2.3/2.3 MB 5.3 MB/s eta 0:00:00ta 0:00:01
Downloading kiwisolver-1.4.7-cp312-cp312-macosx_11_0_arm64.whl (63 kB)
Downloading pyparsing-3.1.4-py3-none-any.whl (104 kB)
Installing collected packages: pyparsing, kiwisolver, fonttools, cycler, contourpy, matplotlib
Successfully installed contourpy-1.3.0 cycler-0.12.1 fonttools-4.54.1 kiwisolver-1.4.7 matplotlib-3.9.2 pyparsin
g-3.1.4
```

In [49]:
```python
from PIL import Image
import matplotlib.pyplot as plt
def generate_caption(image_name):
    # load the image
    # image_name = "1001773457_577c3a7d70.jpg"
    image_id = image_name.split('.')[0]
    img_path = os.path.join(BASE_DIR, "Images", image_name)
    image = Image.open(img_path)
    captions = mapping[image_id]
    print('--------------------Actual--------------------')
    for caption in captions:
        print(caption)
    # predict the caption
    y_pred = predict_caption(model, features[image_id], tokenizer, max_length)
    print('--------------------Predicted--------------------')
    print(y_pred)
    plt.imshow(image)
```

In [50]:
```python
generate_caption("1001773457_577c3a7d70.jpg")
```

```
--------------------Actual--------------------
startseq black dog and spotted dog are fighting endseq
startseq black dog and tricolored dog playing with each other on the road endseq
startseq black dog and white dog with brown spots are staring at each other in the street endseq
startseq two dogs of different breeds looking at each other on the road endseq
startseq two dogs on pavement moving toward each other endseq
--------------------Predicted--------------------
startseq two dogs play with ball in the grass endseq
```
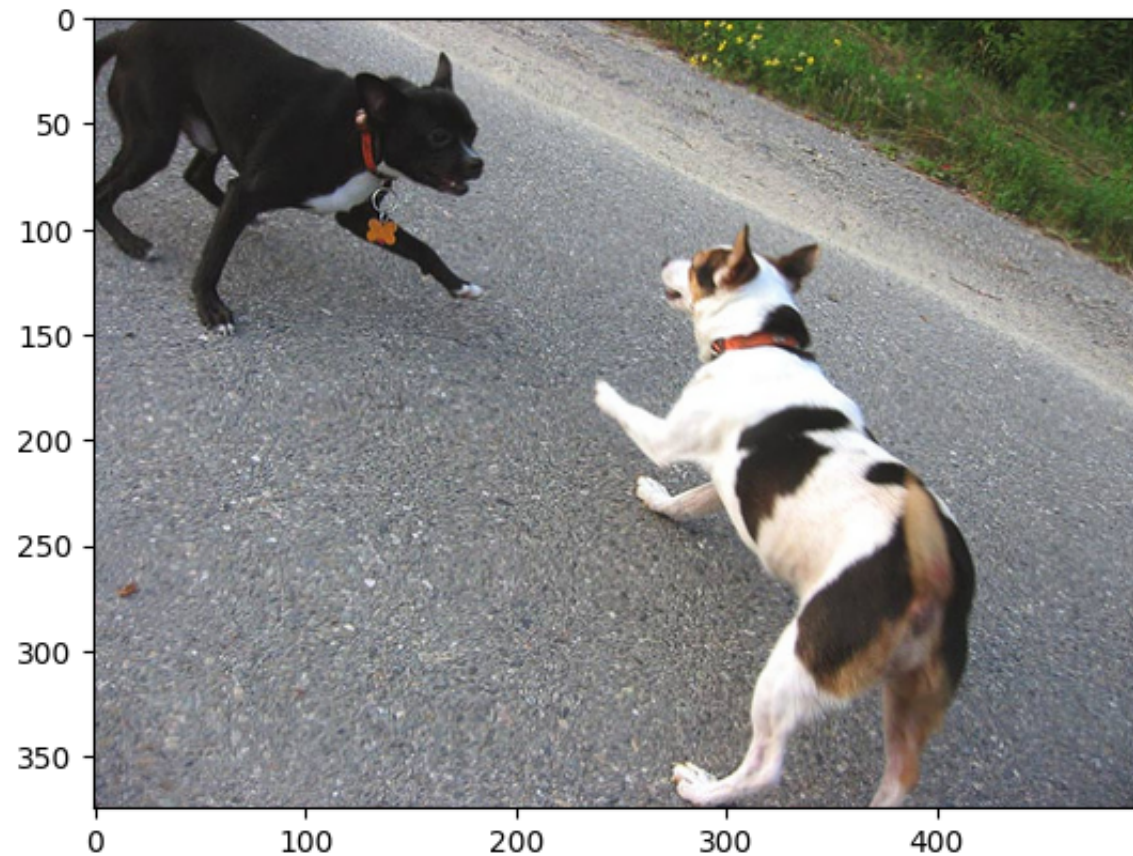
```
In [51]:   generate_caption("1002674143_1b742ab4b8.jpg")
```

```
--------------------Actual--------------------
startseq little girl covered in paint sits in front of painted rainbow with her hands in bowl endseq
startseq little girl is sitting in front of large painted rainbow endseq
startseq small girl in the grass plays with fingerpaints in front of white canvas with rainbow on it endseq
startseq there is girl with pigtails sitting in front of rainbow painting endseq
startseq young girl with pigtails painting outside in the grass endseq
--------------------Predicted--------------------
startseq girl in pigtails is sitting on the dirt with fingerpaints endseq
```

```
In [56]:  vgg_model = VGG16()
          # restructure the model
          vgg_model = Model(inputs=vgg_model.inputs, outputs=vgg_model.layers[-2].output)
```

In [60]:
```python
image_path = '/Users/adityagadagandla/Desktop/project/flickr8k/Images/101654506_8eb26cfb60.jpg'
# load image
image = load_img(image_path, target_size=(224, 224))
# convert image pixels to numpy array
image = img_to_array(image)
# reshape data for model
image = image.reshape((1, image.shape[0], image.shape[1], image.shape[2]))
# preprocess image from vgg
image = preprocess_input(image)
# extract features
feature = vgg_model.predict(image, verbose=0)
# predict from the trained model
predict_caption(model, feature, tokenizer, max_length)
```

Out[60]: `'startseq dog is running through snow endseq'`

In [63]:
```python
image_path = '/Users/adityagadagandla/Desktop/project/flickr8k/Images/132489044_3be606baf7.jpg'
# load image
image = load_img(image_path, target_size=(224, 224))
# convert image pixels to numpy array
image = img_to_array(image)
# reshape data for model
image = image.reshape((1, image.shape[0], image.shape[1], image.shape[2]))
# preprocess image from vgg
image = preprocess_input(image)
# extract features
feature = vgg_model.predict(image, verbose=0)
# predict from the trained model
predict_caption(model, feature, tokenizer, max_length)
```

Out[63]: `'startseq man in suit is sitting on chair in front of an empty transit endseq'`