Text Recognition System

ARTIFICIAL NEURAL NETWORKS PROJECT
VAMSHEETH VADLAMUDI | U101115FCS274 | B. TECH CSE 15-19 | SECTION 4

Optical Character Recognition

- OCR stands for Optical Character Recognition. It is a widespread technology to recognise text inside images, such as scanned documents and photos.

- OCR technology is used to convert virtually any kind of images containing written text (typed, handwritten or printed) into machine-readable text data.

Why OCR?

- Many businesses, educational institutes and other organizations make use of OCR programs and software to hasten their data entry process.

- Furthermore, using OCR program reduces human error, which can occur while the data is being input.

- However, the conversion while the program is running has to be monitored carefully, so that unrecognized characters, words and sentences can be identified.

▪ **Names of group members**

■ Vamsheeth Vadlamudi

▪ **A paragraph which describes the approach you have used to solve the problem**

A DIA consists of three main modules. The first major step is to segment the textual information from a given scanned page. The layout analysis module extracts text-lines, words or characters from a given page and the OCR module does the actual text recognition. Errors induced by these two modules are then corrected at a post-processing step using language models or dictionary correction.

The system is developed using the Python programming language and uses the Keras deep learning library. The Keras library is compatible with both Tensorflow and Theano. However this system is developed and tested with Tensorflow

Searching Algorithm:

- Word beam search decoding is a Connectionist Temporal Classification (CTC) decoding algorithm. It is used for sequence recognition tasks like Handwritten Text Recognition (HTR) or Automatic Speech Recognition (ASR).

- Word beam search decoding is placed right after the RNN layers to decode the output, see the red dashed rectangle in the illustration.

The four main properties of word beam search are:

- Words constrained by dictionary

- Allows arbitrary number of non-word characters between words (numbers, punctuation marks)

- Optional word-level Language Model (LM)

- Faster than token passing

▪ Data set information.

The IAM Handwriting Database contains forms of handwritten English text which can be used to train and test handwritten text recognizers and to perform writer identification and verification experiments.

The IAM Handwriting Database 3.0 is structured as follows:

- 657 writers contributed samples of their handwriting

- 1'539 pages of scanned text

- 5'685 isolated and labeled sentences

- 13'353 isolated and labeled text lines

- 115'320 isolated and labeled words

▪ Software used

- Python
- TensorFlow
- Tesseract
- OpenCV

▪ Inputs and Outputs (Intermediate results as well if applicable)


```
add 1 (Add)                     (None, 128, 512)      0        gru1[0][0]
                                                               gru1_b[0][0]

gru2 (GRU)                      (None, 128, 512)      1574400  add_1[0][0]

dense2 (Dense)                  (None, 128, 52)       26676    gru2[0][0]

softmax (Activation)            (None, 128, 52)       0        dense2[0][0]
=====================================================================
Total params: 3,286,020
Trainable params: 3,286,020
Non-trainable params: 0

Epoch 1/25
2018-12-06 03:23:45.754656: I tensorflow/core/platform/cpu_feature_guard.cc:141] Your CPU suppo
12800/12800 [==============================] - 52206s 4s/step - loss: 21.1394 - val_loss: 127.5

Out of 256 samples:  Mean edit distance: 11.621 Mean normalized edit distance: 0.872
/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/matplotlib/font_manager.py:1241: Us
  (prop.get_family(), self.defaultFamily[fontext]))
Epoch 2/25
   443/12800 [>................................] - ETA: 12:50:39 - loss: 14.0141
```

not-yeshwanth-reddy@tech-book: ~/OCR

```
                                          gru1_b[0][0]

gru2 (GRU)                    (None, 128, 512)    1574400    add_1[0][0]

dense2 (Dense)                (None, 128, 52)     26676      gru2[0][0]

softmax (Activation)          (None, 128, 52)     0          dense2[0][0]
================================================================================
Total params: 3,286,020
Trainable params: 3,286,020
Non-trainable params: 0

Epoch 1/25
2018-12-06 03:23:45.754656: I tensorflow/core/platform/cpu_feature_guard.cc:141] Your CPU supports instructions that this TensorFlow binary was not compiled to use: AVX2 FMA
12800/12800 [==============================] - 52206s 4s/step - loss: 21.1394 - val_loss: 127.5945

Out of 256 samples:  Mean edit distance: 11.621 Mean normalized edit distance: 0.872
/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/matplotlib/font_manager.py:1241: UserWarning: findfont: Font family ['Thaana Unicode Akeh'] not found. Falling back to DejaVu Sans.
  (prop.get_family(), self.defaultFamily[fontext]))
Epoch 2/25
  769/12800 [>.............................] - ETA: 12:30:57 - loss: 13.9113^[[3~

12800/12800 [==============================] - 55372s 4s/step - loss: 12.6385 - val_loss: 128.9450

Out of 256 samples:  Mean edit distance: 11.781 Mean normalized edit distance: 0.906
Epoch 3/25
  423/12800 [.............................] - ETA: 13:44:19 - loss: 13.7929^CTraceback (most recent call last):
  File "trainthaana.py", line 4, in <module>
    ocr.train("test_run",0,25,512)
  File "/home/not-yeshwanth-reddy/OCR/thaanaocr.py", line 609, in train
    callbacks=[viz_cb, self.img_gen], initial_epoch=start_epoch, verbose=1)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/keras/legacy/interfaces.py", line 91, in wrapper
    return func(*args, **kwargs)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/keras/engine/training.py", line 1418, in fit_generator
    initial_epoch=initial_epoch)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/keras/engine/training_generator.py", line 217, in fit_generator
    class_weight=class_weight)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/keras/engine/training.py", line 1217, in train_on_batch
    outputs = self.train_function(ins)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/keras/backend/tensorflow_backend.py", line 2715, in __call__
    return self._call(inputs)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/keras/backend/tensorflow_backend.py", line 2675, in _call
    fetched = self._callable_fn(*array_vals)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/tensorflow/python/client/session.py", line 1439, in __call__
    run_metadata_ptr)
KeyboardInterrupt
not-yeshwanth-reddy@tech-book:~/OCR$ 
```

not-yeshwanth-reddy@tech-book: ~/OCR

```
Layer (type)                  Output Shape        Param #    Connected to
================================================================================
the_input (InputLayer)        (None, 512, 64, 1)  0

conv1 (Conv2D)                (None, 512, 64, 16) 160        the_input[0][0]

max1 (MaxPooling2D)           (None, 256, 32, 16) 0          conv1[0][0]

conv2 (Conv2D)                (None, 256, 32, 16) 2320       max1[0][0]

max2 (MaxPooling2D)           (None, 128, 16, 16) 0          conv2[0][0]

reshape (Reshape)             (None, 128, 256)    0          max2[0][0]

dense1 (Dense)                (None, 128, 32)     8224       reshape[0][0]

gru1 (GRU)                    (None, 128, 512)    837120     dense1[0][0]

gru1_b (GRU)                  (None, 128, 512)    837120     dense1[0][0]

add_1 (Add)                   (None, 128, 512)    0          gru1[0][0]
                                                             gru1_b[0][0]

gru2 (GRU)                    (None, 128, 512)    1574400    add_1[0][0]

dense2 (Dense)                (None, 128, 52)     26676      gru2[0][0]

softmax (Activation)          (None, 128, 52)     0          dense2[0][0]
================================================================================
Total params: 3,286,020
Trainable params: 3,286,020
Non-trainable params: 0
_____
Traceback (most recent call last):
  File "ocrtest.py", line 32, in <module>
    ocr.loadweights('weights24.h5') # replace with your trained weight file
  File "/home/not-yeshwanth-reddy/OCR/thaanaocr.py", line 552, in loadweights
    self.model.load_weights(weightsfile)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/keras/engine/network.py", line 1157, in load_weights
    with h5py.File(filepath, mode='r') as f:
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/h5py/_hl/files.py", line 312, in __init__
    fid = make_fid(name, mode, userblock_size, fapl, swmr=swmr)
  File "/home/not-yeshwanth-reddy/.local/lib/python3.6/site-packages/h5py/_hl/files.py", line 142, in make_fid
    fid = h5f.open(name, flags, fapl=fapl)
  File "h5py/_objects.pyx", line 54, in h5py._objects.with_phil.wrapper
  File "h5py/_objects.pyx", line 55, in h5py._objects.with_phil.wrapper
  File "h5py/h5f.pyx", line 78, in h5py.h5f.open
OSError: Unable to open file (unable to open file: name = 'weights24.h5', errno = 2, error message = 'No such file or directory', flags = 0, o_flags = 0)
not-yeshwanth-reddy@tech-book:~/OCR$ 
```

▪ Final Results:
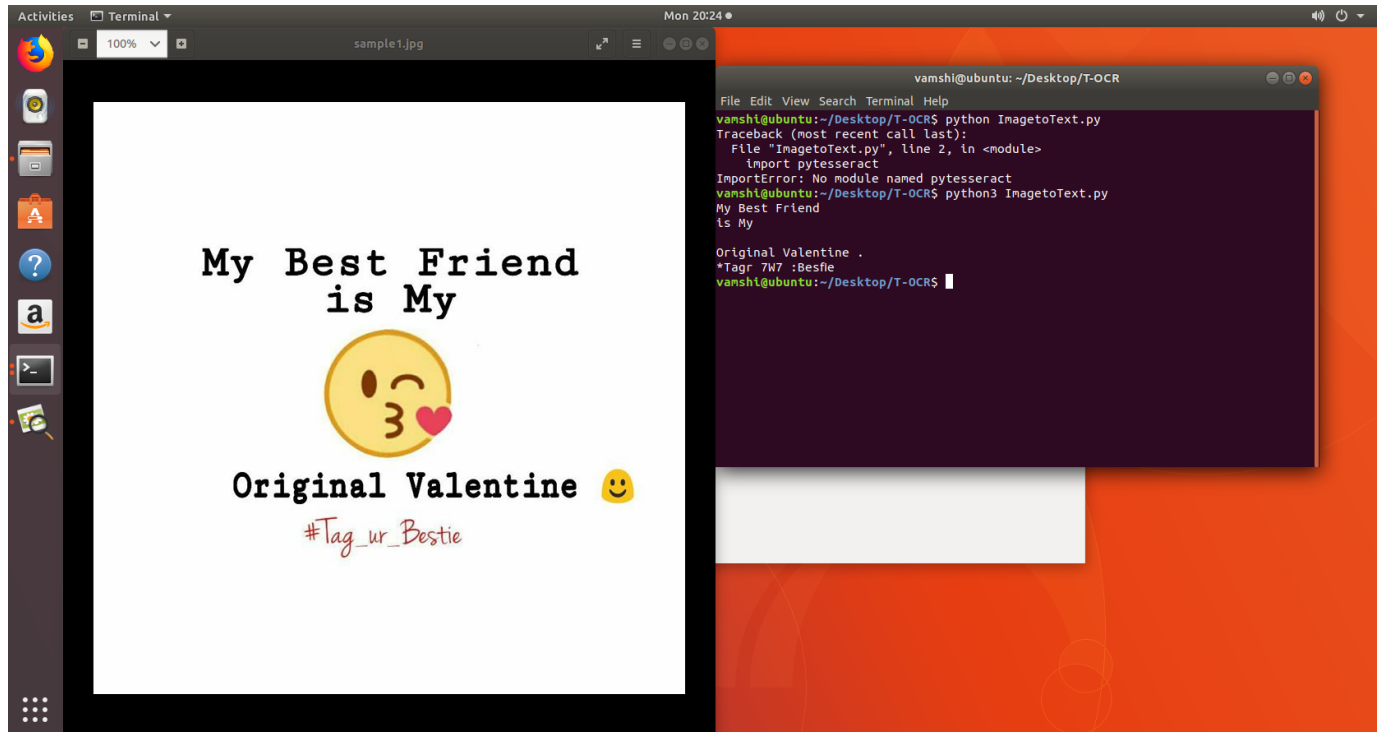
▪ Accuracy

Engine	Total char errors	Word Recall Errors	Word Precision Errors	Walltime	CPUtime*
Tess 3.04	13.9	30	31.2	3.0	2.8
Cube	15.1	29.5	30.7	3.4	3.1
Tess+Cube	11.0	24.2	25.4	5.7	5.3
LSTM	7.6	20.9	20.8	1.5	2.5

▪ Reference papers

1. http://paper.ijcsns.org/07_book/200812/20081218.pdf

2. http://ijarcet.org/wp-content/uploads/IJARCET-VOL-1-ISSUE-4-131-133.pdf

3. http://www.pnrsolution.org/Datacenter/Vol3/Issue1/8.pdf

4. https://books.google.co.in/books?hl=en&lr=&id=wovsCgAAQBAJ&oi=fnd&pg=PA288&dq= handwritten+text+recognition&ots=y4nPxm6qMv&sig=NkawbYGjWskMjnlVN8cAGKqq g Sc#v=onepage&q=handwritten%20text%20recognition&f=false

5. https://dihana.cps.unizar.es/~eduardo/trabhci/doc/2012/2012-03text_trabhci.pdf

6. http://yann.lecun.com/exdb/mnist/

7. http://www.fki.inf.unibe.ch/databases/iam-handwriting-database/download-the-iamhandwriting-database (Requires registration to download)

8.https://hackernoon.com/top-10-libraries-in-python-to-implement-machine-learning12602cf5dc61