Deep Learning Basics - Project

Phase 2 by Carmen Azorín

First try

Hyperparameters:

- Epochs = 20
- Batch size = 12
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 162 samples

Validation data:

- My own
- 55 samples

Epoch	Time	Loss	Val loss
1	22s	14,57	32,57
2	14s	13,84	31,49
3	14s	13,58	30,37
4	14s	13,29	29,81
5	14s	13,08	29,50
6	14s	12,8313	31,88
7	14s	12,6683	36,30
8	14s	12,4545	29,90
9	15s	12,3471	32,82
10	16s	12,4697	32,68
11	15s	12,0729	35,13
12	14s	11,807	36,83
13	14s	11,6737	38,05
14	17s	11,4001	38,35
15	18s	11,1775	40,49
16	15s	10,8469	39,12
17	14s	10,6936	41,09
18	18s	10,4846	37,79
19	16s	10,3153	38,18
20	18s	10,0244	37,22

Second try

Hyperparameters:

- Epochs = 20
- Batch size = 24
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 416 samples

Validation data:

- My own
- 55 samples

Epoch	Time	Loss	Val loss
1	46s	11,61	32,77
2	36s	11,09	32,72
3	36s	10,63	37,76
4	35s	10,28	88,50
5	35s	10,02	34,72
6	32s	9,65	39,32
7	33s	9,38	67,62
8	32s	8,96	38,72
9	33s	8,57	37,18
10	32s	8,20	38,55
11	32s	7,63	40,43
12	33s	7,28	42,00
13	32s	6,88	50,97
14	33s	6,37	42,77
15	32s	5,99	46,48
16	33s	5,50	43,40
17	32s	5,04	45,94
18	33s	4,75	45,23
19	32s	4,53	88,25
20	33s	4,25	47,21

Conclusions first and second try

After first try:

- Results are really bad
- Both loss and validation loss are too high, so we should increase the training data
- We are also increasing the batch size

After second try:

- Result don’t seem to improve
- Although loss decreases, validation loss increases
- Since the training dataset have many images, we are going to increase the number of samples notably

Third try

Hyperparameters:

- Epochs = 20
- Batch size = 24
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 2171 samples

Validation data:

- My own
- 55 samples

Epoch	Time	Loss	Val loss
1	170s	19,56	32,74
2	135s	14,56	28,49
3	135s	13,89	24,98
4	141s	13,52	24,62
5	136s	13,17	24,75
6	137s	12,80	24,54
7	149s	12,42	26,14
8	141s	11,91	28,77
9	137s	11,46	28,78
10	136s	11,14	31,20
11	136s	10,78	31,17
12	136s	10,83	33,15
13	141s	10,20	35,54
14	136s	9,89	34,84
15	136s	9,57	41,83
16	138s	9,16	38,84
17	136s	8,80	41,17
18	136s	8,46	38,99
19	137s	8,03	42,82
20	136s	7,64	44,04

Fourth try

Hyperparameters:

- Epochs = 20
- Batch size = 24
- Learning rate = 0.01
  ADADELTA

Training data:

- IAM Words
- 2171 samples

Validation data:

- My own
- 55 samples

Epoch	Time	Loss	Val loss
1	250	222,4785	223,6628
2	220	136,0815	161,5947
3	223	34,5549	81,4436
4	224	22,4053	50,9397
5	220	21,0182	40,6348
6	220	20,4079	37,157
7	221	19,8395	35,8214
8	219	19,2604	34,754
9	219	18,6889	33,9607
10	219	18,145	33,0269
11	219	17,6655	31,957
12	221	17,2614	31,3386
13	220	16,9145	30,5305
14	219	16,6204	29,7556
15	219	16,3818	29,33
16	222	16,184	28,7509
17	218	16,0187	28,3894
18	219	15,8867	28,0613
19	219	15,7672	27,7689
20	220	15,6667	27,5042

Fifth try

Hyperparameters:

- Epochs = 20
- Batch size = 24
- Learning rate = 0.01 Adam

Training data:

- IAM Words
- 2171 samples

Validation data:

- My own
- 55 samples

Epoch	Time	Loss	Val loss
1	231	18,9698	24,8037
2	216	14,3608	25,6011
3	220	13,7526	24,5007
4	218	13,2926	24,4112
5	218	12,9363	25,147
6	218	12,4555	24,9447
7	217	12,0371	29,4383
8	219	11,8724	27,7382
9	218	11,2983	30,07

Conclusions third, fourth and fifth try

After third try:

- Results are worse
- Loss has increased and validation loss is the same
- Also, the executing time is too high. For next try, we change optimizer to Adadelta

After fourth try:

- Results are worse
- Loss has increased and time is higher. We return to Adam's optimizer but less learning rate.

After fifth try:

- Results are pretty much the same
- For next attempts, we return to first optimizer and learning rate
- Let's try different hyperparameters

Sixth try

Hyperparameters:

- Epochs = 60
- Batch size = 24
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 2171 samples

Validation data:

- My own
- 55 samples

Epoch	Time	Loss	Val loss
3	137s	13,90	24,75
6	175s	12,83	24,86
9	136s	11,83	26,67
12	135s	10,63	27,21
15	136s	9,70	30,42
18	136s	8,55	30,72
21	139s	7,32	33,45
24	139s	5,94	38,30
27	141s	4,55	36,86
30	135s	3,28	40,68
33	135s	2,27	43,48
36	136s	1,53	47,64
39	136s	0,99	53,91
42	135s	0,75	54,27
45	135s	0,64	55,01
48	136s	0,46	53,69
51	135s	0,38	55,49
54	136s	0,28	49,71
57	136s	0,47	57,20
60	136s	0,75	64,17

Seventh try

Hyperparameters:

- Epochs = 60
- Batch size = 64
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 4521 samples (with data augmentation)

Validation data:

- My own
- 55 samples

Epoch	Time	Loss	Val loss
1	232	20,17	28,12
2	225	14,73	29,57
3	222	14,10	24,31
4	223	13,58	24,32
5	223	13,23	24,19
6	223	12,84	24,35
7	223	12,41	25,70
8	224	11,99	25,93
9	223	12,69	25,19
10	223	12,85	26,22

Conclusions sixth and seventh try

After sixth try:

- Results are much better
- Loss has decreased considerably
- Unfortunately, validation loss doesn't improve
- Let's try data augmentation: some images are darkening, other eroding and other dilating
- Also we are adding early stopping depending on validation loss

After seventh try:

- Execution stopped by early stopping
- Maybe my validation data is wrong
- I am going to split my training data so that some images are for validation
- I won't use my own validation data for next try

Eighth try

Hyperparameters:

- Epochs = 60
- Batch size = 24
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 4521 samples

Validation data:

- IAM Words
- 146 samples

Epoch	Time	Loss	Val loss
2	488	13,7769	13,2209
4	493	12,7119	12,1189
6	485	11,2379	10,6627
8	487	10,0252	10,0342
10	484	8,6019	10,8986
12	487	6,8645	8,9077
14	488	5,1497	9,1991
16	491	3,522	9,8438
18	493	2,5569	21,4657
20	482	1,5091	10,9311
22	481	1,0771	11,8585
24	484	0,7319	30,2285
26	501	0,6179	12,5426
28	481	0,4675	13,6833
30	492	0,4782	49,7148

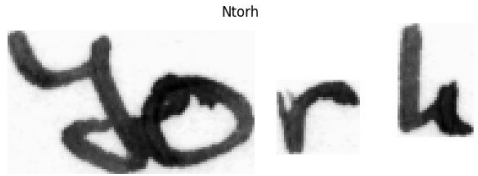
Conclusions and sanity check eighth try

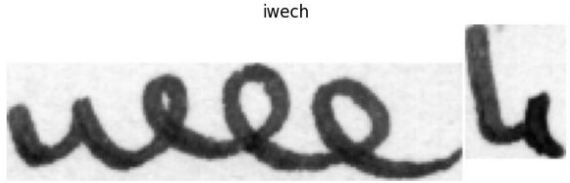
After eighth try:

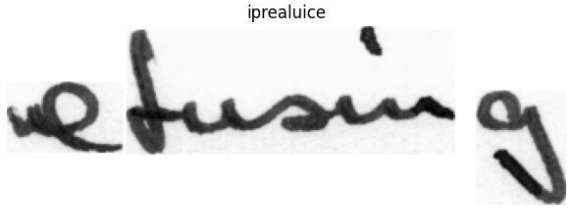
- Results are the best so far
- Stopped by early stopping because validation loss was increasing
- Therefore, validation loss doesn't improve as expected

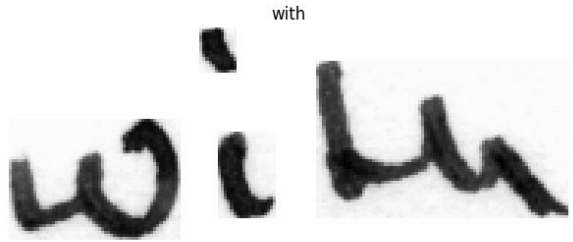
Performance metrics:

```
Correct characters predicted : 23.50%

Correct words predicted      : 19.18%

Character Error Rate         : 80.82%
```









Extra attempts

These attempts were in the middle of other executions in order to see how the model worked with different hyperparameters

Ninth try

Hyperparameters:

- Epochs = 60
- Batch size = 128
- Learning rate = 0.001 Adam

Training data:

- IAM Words
- 4512 samples (with data augmentation)

Validation data:

- My own
- 76 samples

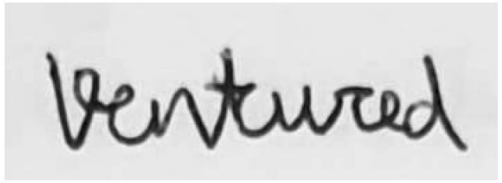
Epoch	Time	Loss	Val loss
1	364	15,2251	36,3803
2	354	13,8858	40,4283
3	350	13,1042	28,4066
4	422	12,5858	22,6433
5	436	12,0315	22,5669
6	378	11,3552	22,7396
7	361	10,8798	23,0696
8	357	10,4314	23,2438
9	357	9,8138	30,7117
10	349	9,3465	25,0862

- Best results for my own validation data
- Stopped by Early Stopping

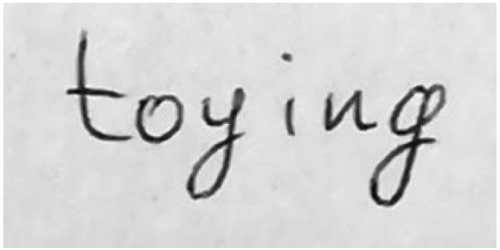
Sanity check ninth try

```
Correct characters predicted : 2.79%
Correct words predicted      : 0.00%
Mean Character Error Rate    : 100.00%
```

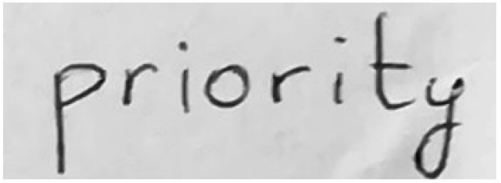
bfd



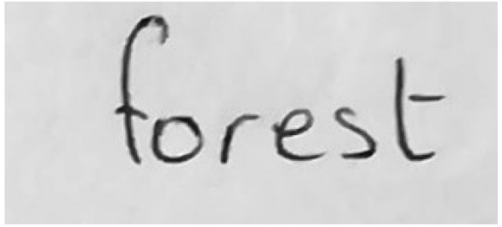
be



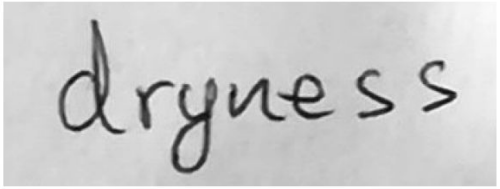
bfs



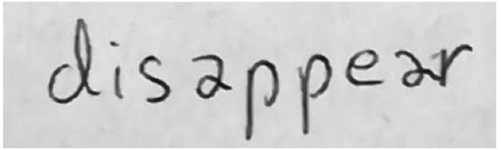
wy



bfn



bfe



Tenth try

Hyperparameters:

- Epochs = 60
- Batch size = 128
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 4512 samples (with data augmentation)

Validation data:

- IAM Words
- 146 samples

Epoch	Time	Loss	Val loss
4	308	14,2239	18,0789
8	309	13,0234	13,1023
12	311	11,9992	11,802
16	313	10,931	10,6946
20	311	9,5858	9,8535
24	312	8,2183	9,4115
28	313	6,5323	16,5898
32	314	4,8621	9,729
36	310	3,3369	9,9708

- Same case as eighth try (the best one) but batch size is higher
- Results are similar, but loss is worse
- Was stopped by Early Stopping

Eleventh try

Hyperparameters:

- Epochs = 40
- Batch size = 64
- Learning rate = 0.0001
  Adam

Training data:

- IAM Words
- 4512 samples (with data augmentation)

Validation data:

- IAM Words
- 146 samples

Epoch	Time	Loss	Val loss
4	377	13,6609	13,5483
8	372	12,2475	11,5913
12	369	10,6793	11,0072
16	371	8,6777	9,7663
20	371	6,8872	9,2536
24	372	4,2045	9,2652
28	372	2,4133	10,9367
32	374	1,3026	11,2752
36	372	0,7543	12,7587
40	371	0,4307	24,7469

- Same case as eighth try, but the number of epochs is lower
- Loss and validation loss are the same as eighth attempt
- The executing time is better

Twelfth try

Hyperparameters:

- Epochs = 30
- Batch size = 24
- Learning rate = 0.0001 Adam

Training data:

- IAM Words + My own
- 4580 samples (With data augmentation)

Validation data:

- My own + IAM Words
- 27 samples

Epoch	Time	Loss	Val loss
1	487	16,7174	23,8079
2	474	14,0061	23,3109
3	470	13,3308	23,3861
4	468	12,6621	23,8522
5	461	12,0048	23,6122
6	469	11,5257	24,1409
7	466	10,7972	24,3355

- Stopped by Early Stopping with 5 epochs of patience
- For this attempt, I mixed IAM Words dataset and my own dataset
- Results aren’t better

Final conclusions

- The program doesn’t work as expected. For next phase I will try to improve it.
- I will try more optimizers and maybe execute the program in other laptop to see if the problem is the hardware.
- In order to decrease validation loss, I will collect more words from new different people.
- If none of the above works, I will try more different hyperparameters.