exps	AE as feature extractor	AE regulizer	epochs	standard	batch size	CV	early stopping	lr decay on plateau	Noise	f1score
1	FALSO	FALSO	300	VERO	128	5	10	FALSO	False	**0.93**
2	FALSO	FALSO	300	FALSO	128	5	10	FALSO	False	**0.94**
3	FALSO	FALSO	300	FALSO	128	5	10	0.1, patience 10	False	**0.94**
4	FALSO	FALSO	300	FALSO	128	5	10	FALSO	Gaussian	**0.94**
5	FALSO	FALSO	300	FALSO	128	5	10	FALSO	saltpepper	**0.93**
6	VERO	FALSO	300	FALSO	128	FALSO	10	FALSO	FALSE	**0.91**
7	VERO	FALSO	300	FALSO	128	FALSO	10	FALSO	saltpepper	**0.93**
8	VERO	FALSO	300	FALSO	128	FALSO	10	FALSO	gaussian	**0.94**
9	VERO	KLD	300	FALSO	128	FALSO	10	FALSO	FALSO	**0.92**
10	VERO	KLD	300	FALSO	256	FALSO	10	FALSO	Gaussian	**0.94**
11	**AE ONLY**									

Experiments - Plots and models

Exp1 3

Exp2 4

Exp3 5

Exp4 6

Exp5 7

Exp6 8

Exp7 10

Exp8 11

Exp9 13

Exp10 14

Exp1

Model

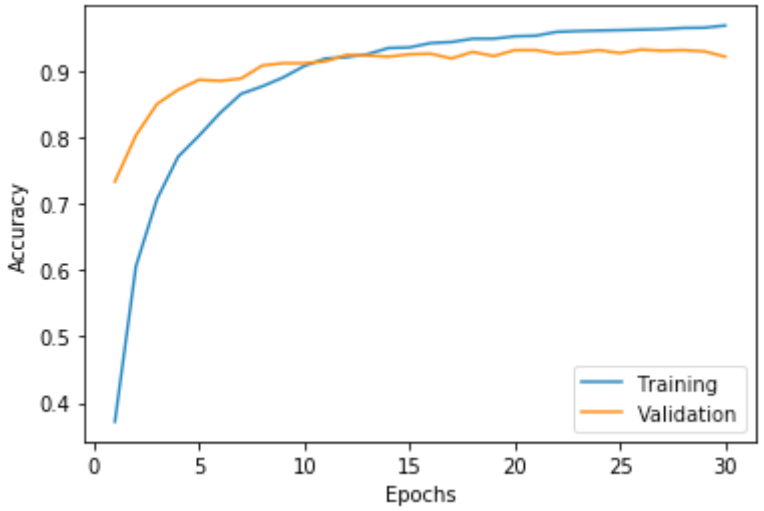
Layer (type)	Output Shape	Param #
input_20 (InputLayer)	(None, 784)	0
dense_77 (Dense)	(None, 512)	401920
batch_normalization_77 (Batc	(None, 512)	2048
dropout_77 (Dropout)	(None, 512)	0
dense_78 (Dense)	(None, 170)	87210
batch_normalization_78 (Batc	(None, 170)	680
dropout_78 (Dropout)	(None, 170)	0
dense_79 (Dense)	(None, 56)	9576
batch_normalization_79 (Batc	(None, 56)	224
dropout_79 (Dropout)	(None, 56)	0
dense_80 (Dense)	(None, 18)	1026
batch_normalization_80 (Batc	(None, 18)	72
dropout_80 (Dropout)	(None, 18)	0
classifier (Dense)	(None, 11)	209

Total params: 502,965
Trainable params: 501,453

Results

	precision	recall	f1-score	support
0	0.94	0.95	0.95	258
1	0.92	0.90	0.91	242
2	0.93	0.92	0.93	264
3	0.99	0.96	0.98	280
4	0.94	0.95	0.95	263
5	0.89	0.92	0.91	267
6	0.87	0.90	0.88	259
7	0.97	0.92	0.94	257
8	0.93	0.92	0.93	264
9	0.88	0.91	0.90	278
10	0.96	0.95	0.96	168
accuracy			0.93	2800
macro avg	0.93	0.93	0.93	2800
weighted avg	0.93	0.93	0.93	2800

Plots



Exp2

Model:

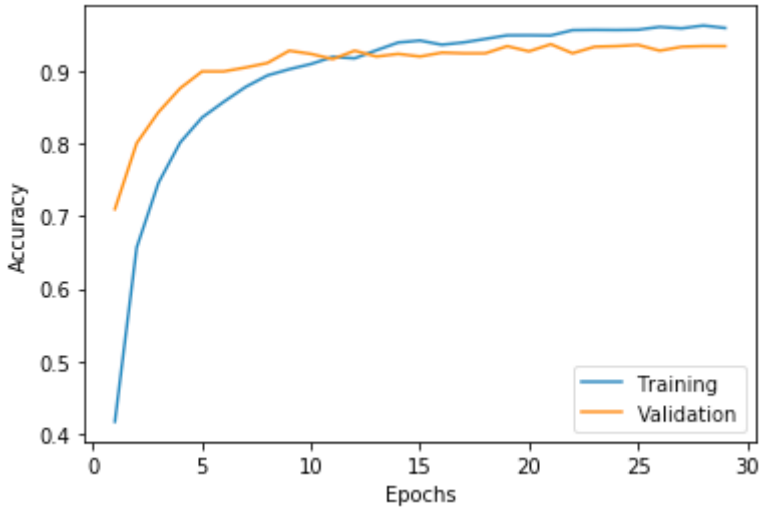
Layer (type)	Output Shape	Param #
input_20 (InputLayer)	(None, 784)	0
dense_77 (Dense)	(None, 512)	401920
batch_normalization_77 (Batc	(None, 512)	2048
dropout_77 (Dropout)	(None, 512)	0
dense_78 (Dense)	(None, 170)	87210
batch_normalization_78 (Batc	(None, 170)	680
dropout_78 (Dropout)	(None, 170)	0
dense_79 (Dense)	(None, 56)	9576
batch_normalization_79 (Batc	(None, 56)	224
dropout_79 (Dropout)	(None, 56)	0
dense_80 (Dense)	(None, 18)	1026
batch_normalization_80 (Batc	(None, 18)	72
dropout_80 (Dropout)	(None, 18)	0
classifier (Dense)	(None, 11)	209

Total params: 502,965
Trainable params: 501,453

Results

	precision	recall	f1-score	support
0	0.94	0.94	0.94	259
1	0.91	0.91	0.91	258
2	0.91	0.93	0.92	265
3	0.97	0.98	0.97	257
4	0.97	0.94	0.95	283
5	0.91	0.95	0.93	244
6	0.89	0.90	0.89	269
7	0.97	0.98	0.97	270
8	0.97	0.93	0.95	250
9	0.94	0.92	0.93	273
10	0.96	0.95	0.96	172
accuracy			0.94	2800
macro avg	0.94	0.94	0.94	2800
weighted avg	0.94	0.94	0.94	2800

Plots



Exp3

Model

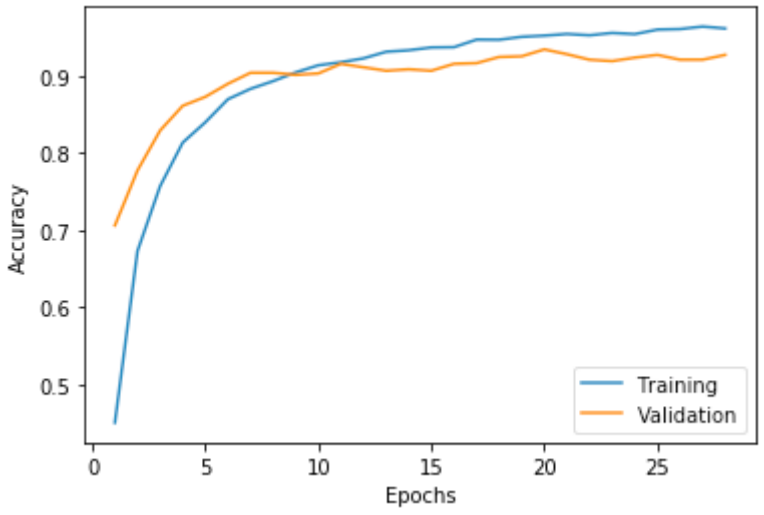
Layer (type)	Output Shape	Param #
input_20 (InputLayer)	(None, 784)	0
dense_77 (Dense)	(None, 512)	401920
batch_normalization_77 (Batc	(None, 512)	2048
dropout_77 (Dropout)	(None, 512)	0
dense_78 (Dense)	(None, 170)	87210
batch_normalization_78 (Batc	(None, 170)	680
dropout_78 (Dropout)	(None, 170)	0
dense_79 (Dense)	(None, 56)	9576
batch_normalization_79 (Batc	(None, 56)	224
dropout_79 (Dropout)	(None, 56)	0
dense_80 (Dense)	(None, 18)	1026
batch_normalization_80 (Batc	(None, 18)	72
dropout_80 (Dropout)	(None, 18)	0
classifier (Dense)	(None, 11)	209

Total params: 502,965
Trainable params: 501,453

Results

	precision	recall	f1-score	support
0	0.94	0.94	0.94	259
1	0.91	0.91	0.91	258
2	0.91	0.93	0.92	265
3	0.97	0.98	0.97	257
4	0.97	0.94	0.95	283
5	0.91	0.95	0.93	244
6	0.89	0.90	0.89	269
7	0.97	0.98	0.97	270
8	0.97	0.93	0.95	250
9	0.94	0.92	0.93	273
10	0.96	0.95	0.96	172
accuracy			0.94	2800
macro avg	0.94	0.94	0.94	2800
weighted avg	0.94	0.94	0.94	2800

Plots



Exp4

Model

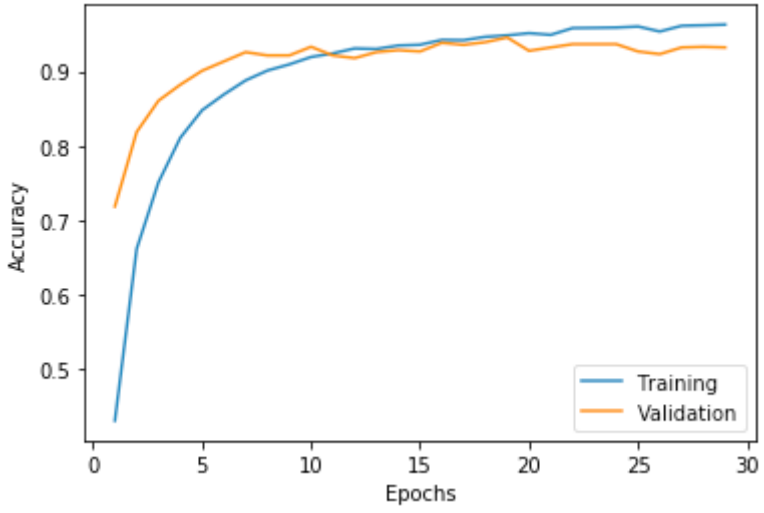
Layer (type)	Output Shape	Param #
input_20 (InputLayer)	(None, 784)	0
dense_77 (Dense)	(None, 512)	401920
batch_normalization_77 (Batc	(None, 512)	2048
dropout_77 (Dropout)	(None, 512)	0
dense_78 (Dense)	(None, 170)	87210
batch_normalization_78 (Batc	(None, 170)	680
dropout_78 (Dropout)	(None, 170)	0
dense_79 (Dense)	(None, 56)	9576
batch_normalization_79 (Batc	(None, 56)	224
dropout_79 (Dropout)	(None, 56)	0
dense_80 (Dense)	(None, 18)	1026
batch_normalization_80 (Batc	(None, 18)	72
dropout_80 (Dropout)	(None, 18)	0
classifier (Dense)	(None, 11)	209

Total params: 502,965
Trainable params: 501,453

Results

	precision	recall	f1-score	support
0	0.97	0.95	0.96	254
1	0.91	0.93	0.92	229
2	0.91	0.95	0.93	243
3	0.99	0.96	0.97	296
4	0.93	0.94	0.93	251
5	0.92	0.92	0.92	272
6	0.89	0.91	0.90	268
7	0.96	0.97	0.96	280
8	0.92	0.94	0.93	257
9	0.94	0.89	0.91	266
10	0.95	0.95	0.95	184
accuracy			0.94	2800
macro avg	0.93	0.94	0.93	2800
weighted avg	0.94	0.94	0.94	2800

Plots



Exp5

Model

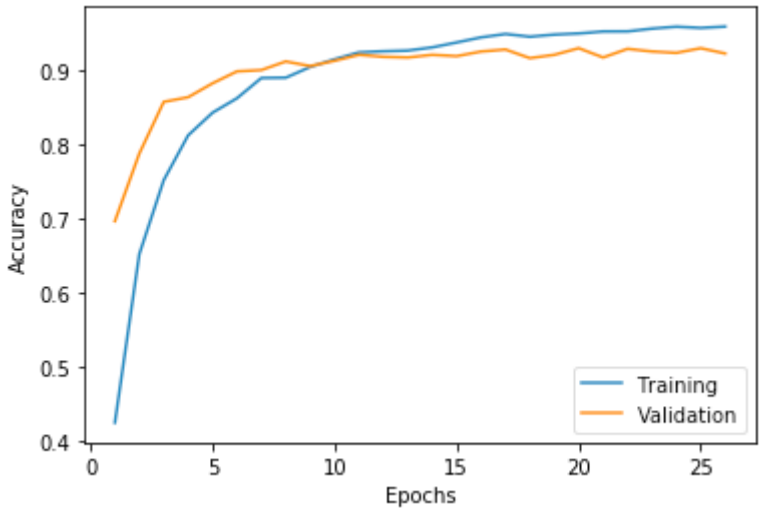
Layer (type)	Output Shape	Param #
input_20 (InputLayer)	(None, 784)	0
dense_77 (Dense)	(None, 512)	401920
batch_normalization_77 (Batc	(None, 512)	2048
dropout_77 (Dropout)	(None, 512)	0
dense_78 (Dense)	(None, 170)	87210
batch_normalization_78 (Batc	(None, 170)	680
dropout_78 (Dropout)	(None, 170)	0
dense_79 (Dense)	(None, 56)	9576
batch_normalization_79 (Batc	(None, 56)	224
dropout_79 (Dropout)	(None, 56)	0
dense_80 (Dense)	(None, 18)	1026
batch_normalization_80 (Batc	(None, 18)	72
dropout_80 (Dropout)	(None, 18)	0
classifier (Dense)	(None, 11)	209

Total params: 502,965
Trainable params: 501,453

Results

	precision	recall	f1-score	support
0	0.94	0.96	0.95	264
1	0.92	0.92	0.92	247
2	0.92	0.92	0.92	266
3	0.96	0.97	0.97	247
4	0.95	0.95	0.95	275
5	0.90	0.97	0.94	282
6	0.89	0.90	0.89	251
7	0.98	0.95	0.96	244
8	0.92	0.94	0.93	266
9	0.95	0.85	0.90	294
10	0.96	0.96	0.96	164
accuracy			0.93	2800
macro avg	0.94	0.94	0.93	2800
weighted avg	0.93	0.93	0.93	2800

Plots



Exp6

Model

Layer (type)	Output Shape	Param #	Connected to
input_4 (InputLayer)	(None, 784)	0	
dense_14 (Dense)	(None, 512)	401920	input_4[0][0]
dense_15 (Dense)	(None, 256)	131328	dense_14[0][0]
dense_16 (Dense)	(None, 128)	32896	dense_15[0][0]
encoder (Dense)	(None, 64)	8256	dense_16[0][0]
dense_17 (Dense)	(None, 128)	8320	encoder[0][0]
dense_18 (Dense)	(None, 256)	33024	dense_17[0][0]
dense_19 (Dense)	(None, 512)	131584	dense_18[0][0]
classifier (Dense)	(None, 11)	715	encoder[0][0]
decoder (Dense)	(None, 784)	402192	dense_19[0][0]

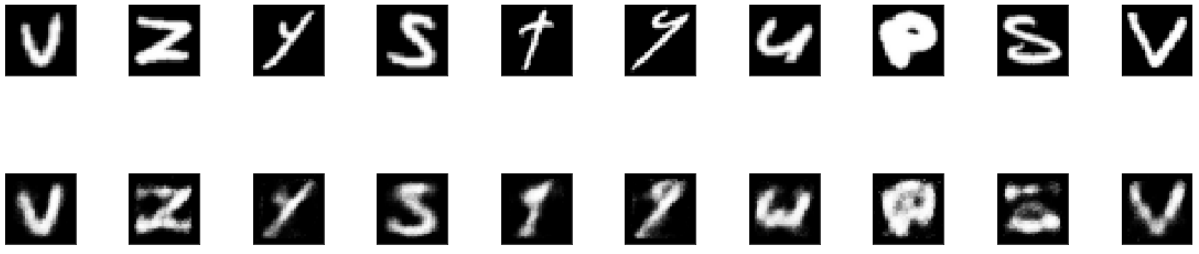
Total params: 1,150,235
Trainable params: 1,150,235
Non-trainable params: 0

Results

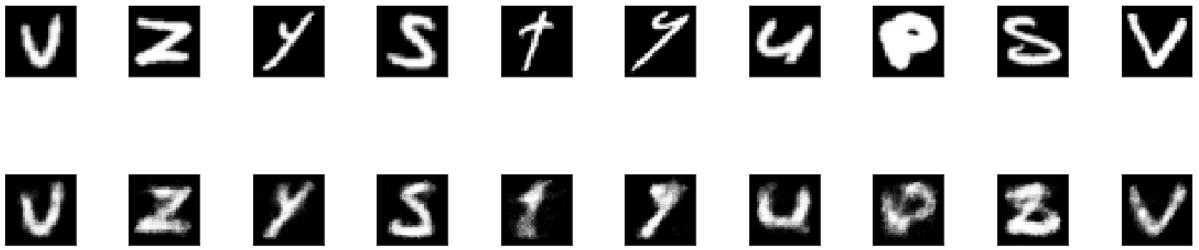
	precision	recall	f1-score	support
0	0.93	0.94	0.93	265
1	0.87	0.87	0.87	247
2	0.88	0.94	0.91	286
3	0.95	0.96	0.95	260
4	0.95	0.89	0.92	281
5	0.94	0.90	0.92	249
6	0.90	0.91	0.90	268
7	0.95	0.94	0.95	269
8	0.91	0.91	0.91	255
9	0.92	0.85	0.88	255
10	0.85	0.98	0.91	165
accuracy			0.91	2800
macro avg	0.91	0.92	0.91	2800
weighted avg	0.92	0.91	0.91	2800

Plots

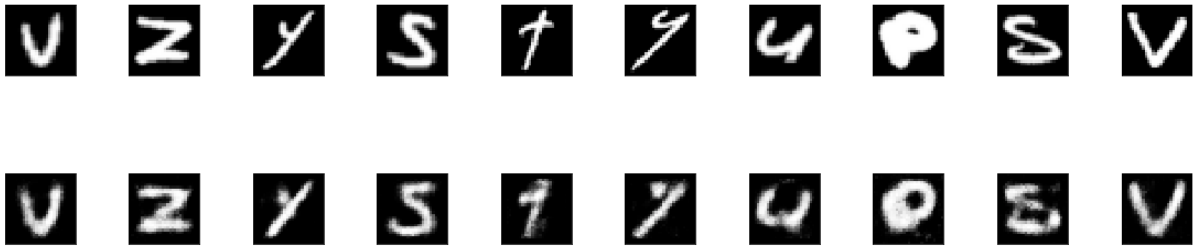
- batch 11200



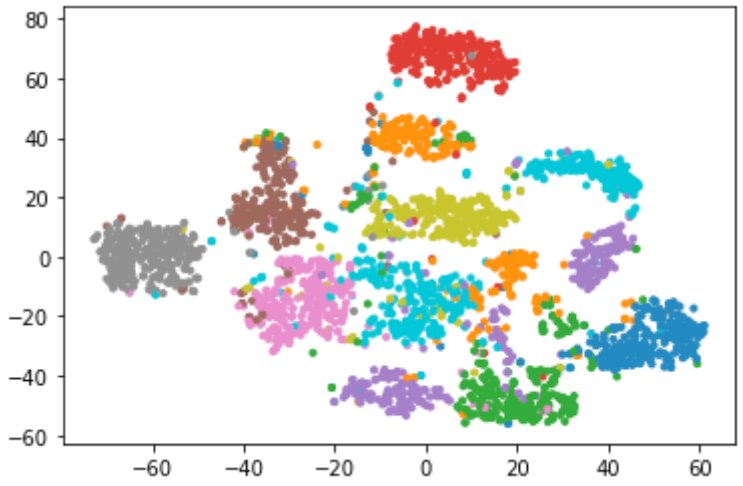
- batch 16



- batch 128



manifold (each class is a color)



Exp7

Model

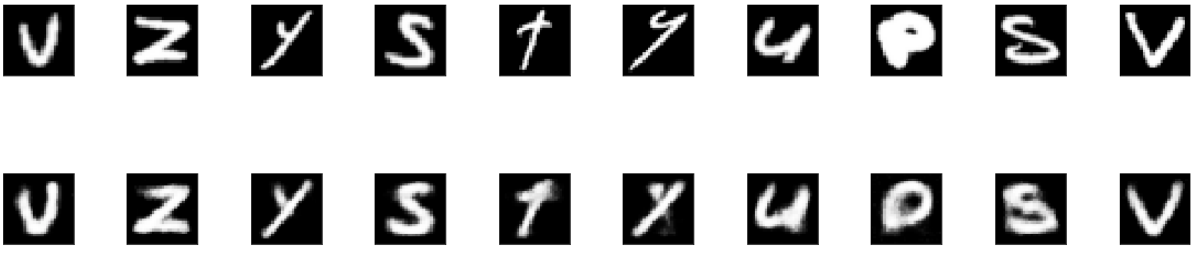
Layer (type)	Output Shape	Param #	Connected to
input_4 (InputLayer)	(None, 784)	0	
dense_14 (Dense)	(None, 512)	401920	input_4[0][0]
dense_15 (Dense)	(None, 256)	131328	dense_14[0][0]
dense_16 (Dense)	(None, 128)	32896	dense_15[0][0]
encoder (Dense)	(None, 64)	8256	dense_16[0][0]
dense_17 (Dense)	(None, 128)	8320	encoder[0][0]
dense_18 (Dense)	(None, 256)	33024	dense_17[0][0]
dense_19 (Dense)	(None, 512)	131584	dense_18[0][0]
classifier (Dense)	(None, 11)	715	encoder[0][0]
decoder (Dense)	(None, 784)	402192	dense_19[0][0]

Total params: 1,150,235
Trainable params: 1,150,235
Non-trainable params: 0

Results

	precision	recall	f1-score	support
0	0.92	0.96	0.94	265
1	0.94	0.86	0.90	247
2	0.91	0.94	0.92	286
3	0.98	0.99	0.98	260
4	0.96	0.91	0.93	281
5	0.92	0.96	0.94	249
6	0.89	0.92	0.90	268
7	0.97	0.95	0.96	269
8	0.90	0.95	0.92	255
9	0.92	0.84	0.87	255
10	0.95	0.98	0.96	165
accuracy			0.93	2800
macro avg	0.93	0.93	0.93	2800
weighted avg	0.93	0.93	0.93	2800

Plots



Exp8

Model

Layer (type)	Output Shape	Param #	Connected to
input_4 (InputLayer)	(None, 784)	0	
dense_14 (Dense)	(None, 512)	401920	input_4[0][0]
dense_15 (Dense)	(None, 256)	131328	dense_14[0][0]
dense_16 (Dense)	(None, 128)	32896	dense_15[0][0]
encoder (Dense)	(None, 64)	8256	dense_16[0][0]
dense_17 (Dense)	(None, 128)	8320	encoder[0][0]
dense_18 (Dense)	(None, 256)	33024	dense_17[0][0]
dense_19 (Dense)	(None, 512)	131584	dense_18[0][0]
classifier (Dense)	(None, 11)	715	encoder[0][0]
decoder (Dense)	(None, 784)	402192	dense_19[0][0]

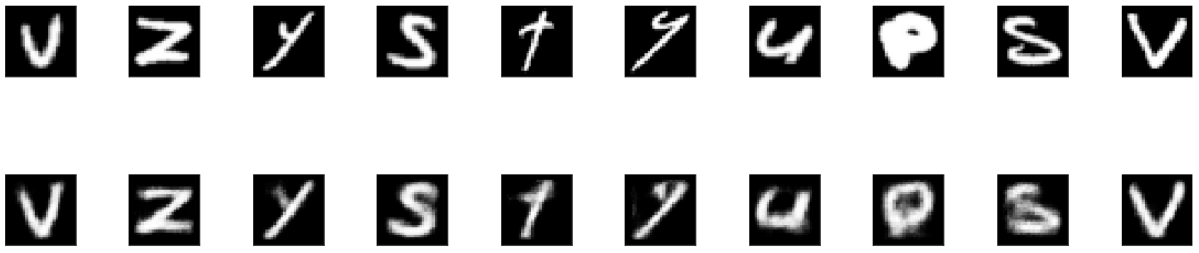
Total params: 1,150,235
Trainable params: 1,150,235
Non-trainable params: 0

Results

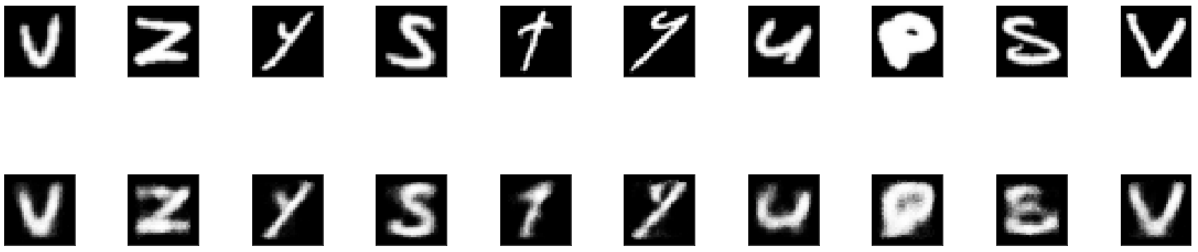
	precision	recall	f1-score	support
0	0.94	0.95	0.94	265
1	0.94	0.89	0.91	247
2	0.92	0.95	0.93	286
3	0.98	0.98	0.98	260
4	0.98	0.91	0.94	281
5	0.95	0.93	0.94	249
6	0.87	0.95	0.91	268
7	0.99	0.96	0.97	269
8	0.94	0.93	0.93	255
9	0.88	0.92	0.90	255
10	0.96	0.97	0.97	165
accuracy			0.94	2800
macro avg	0.94	0.94	0.94	2800
weighted avg	0.94	0.94	0.94	2800

Plots

- batch 128



- batch 11200



Exp9

Model

Layer (type)	Output Shape	Param #	Connected to
input_4 (InputLayer)	(None, 784)	0	
dense_14 (Dense)	(None, 512)	401920	input_4[0][0]
dense_15 (Dense)	(None, 256)	131328	dense_14[0][0]
dense_16 (Dense)	(None, 128)	32896	dense_15[0][0]
encoder (Dense)	(None, 64)	8256	dense_16[0][0]
dense_17 (Dense)	(None, 128)	8320	encoder[0][0]
dense_18 (Dense)	(None, 256)	33024	dense_17[0][0]
dense_19 (Dense)	(None, 512)	131584	dense_18[0][0]
classifier (Dense)	(None, 11)	715	encoder[0][0]
decoder (Dense)	(None, 784)	402192	dense_19[0][0]

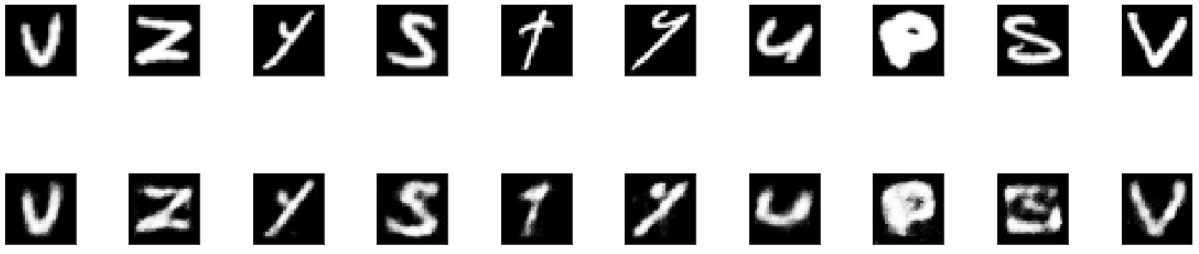
Total params: 1,150,235
Trainable params: 1,150,235
Non-trainable params: 0

Results

	precision	recall	f1-score	support
0	0.91	0.95	0.93	265
1	0.90	0.89	0.89	247
2	0.86	0.92	0.89	286
3	0.96	0.95	0.96	260
4	0.93	0.91	0.92	281
5	0.92	0.92	0.92	249
6	0.93	0.90	0.91	268
7	0.98	0.94	0.96	269
8	0.95	0.89	0.92	255
9	0.85	0.88	0.86	255
10	0.87	0.94	0.90	165
accuracy			0.92	2800
macro avg	0.92	0.92	0.92	2800
weighted avg	0.92	0.92	0.92	2800

Plots

- batch 128



Exp10

Model

Layer (type)	Output Shape	Param #	Connected to
input_4 (InputLayer)	(None, 784)	0	
dense_14 (Dense)	(None, 512)	401920	input_4[0][0]
dense_15 (Dense)	(None, 256)	131328	dense_14[0][0]
dense_16 (Dense)	(None, 128)	32896	dense_15[0][0]
encoder (Dense)	(None, 64)	8256	dense_16[0][0]
dense_17 (Dense)	(None, 128)	8320	encoder[0][0]
dense_18 (Dense)	(None, 256)	33024	dense_17[0][0]
dense_19 (Dense)	(None, 512)	131584	dense_18[0][0]
classifier (Dense)	(None, 11)	715	encoder[0][0]
decoder (Dense)	(None, 784)	402192	dense_19[0][0]

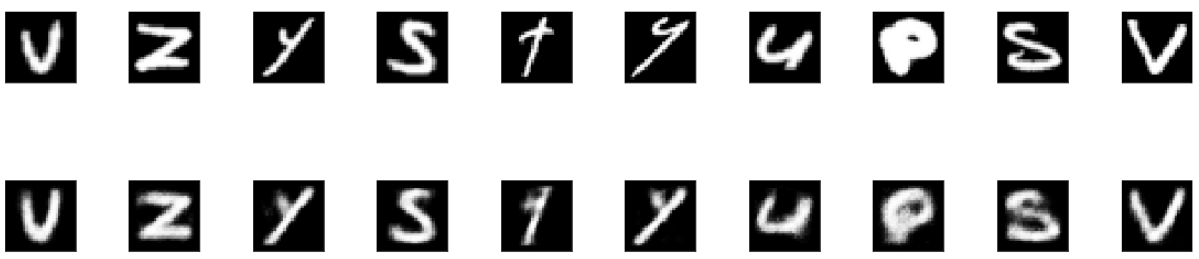
Total params: 1,150,235
Trainable params: 1,150,235
Non-trainable params: 0

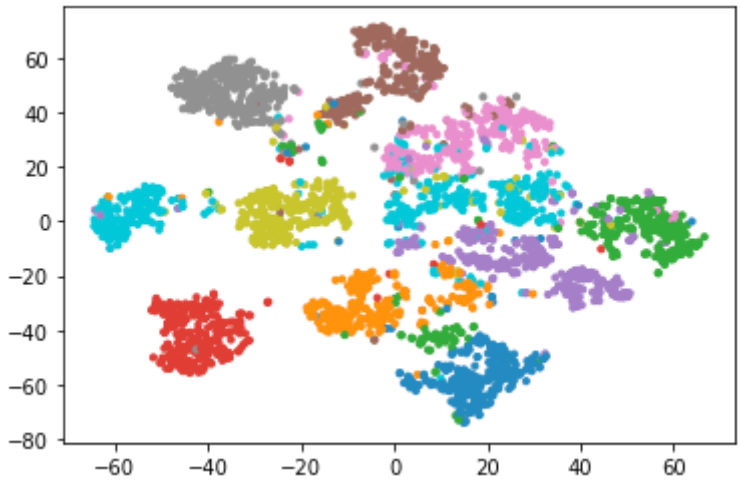
Results

	precision	recall	f1-score	support
0	0.95	0.94	0.95	265
1	0.92	0.91	0.92	247
2	0.91	0.93	0.92	286
3	0.99	0.98	0.98	260
4	0.94	0.93	0.94	281
5	0.91	0.94	0.93	249
6	0.93	0.91	0.92	268
7	0.96	0.97	0.97	269
8	0.95	0.93	0.94	255
9	0.89	0.89	0.89	255
10	0.96	0.98	0.97	165
accuracy			0.94	2800
macro avg	0.94	0.94	0.94	2800
weighted avg	0.94	0.94	0.94	2800

Plots

- batch 128




60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60

EXP11

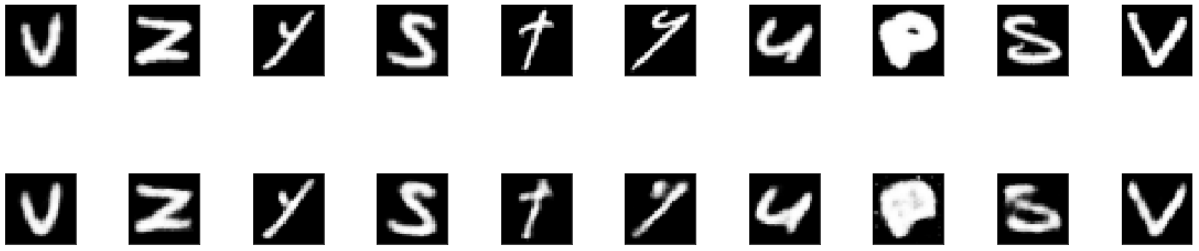
Tying AE

```
Model: "sequential_20"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
sequential_18 (Sequential)   (None, 32)                244192
_________________________________________________________________
sequential_19 (Sequential)   (None, 784)               245424
=================================================================
Total params: 245,424
Trainable params: 245,424
Non-trainable params: 0
_________________________________________________________________
```

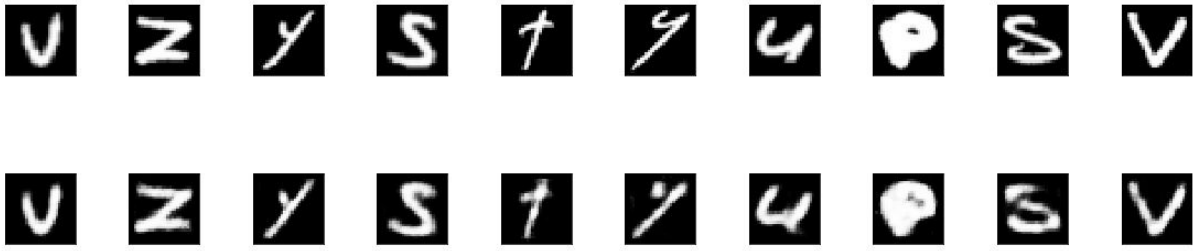
Note: encoder and decoder share their weights -> AE must be symmetric

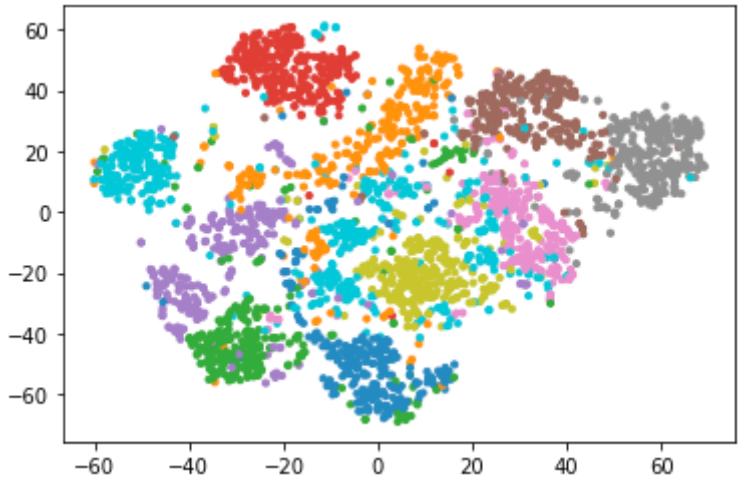
Results

batch 512

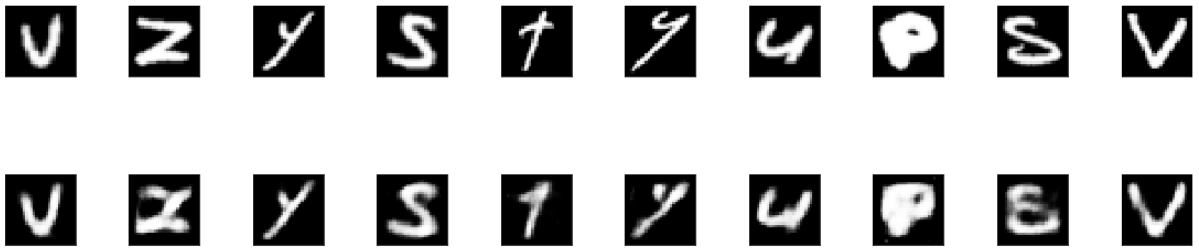
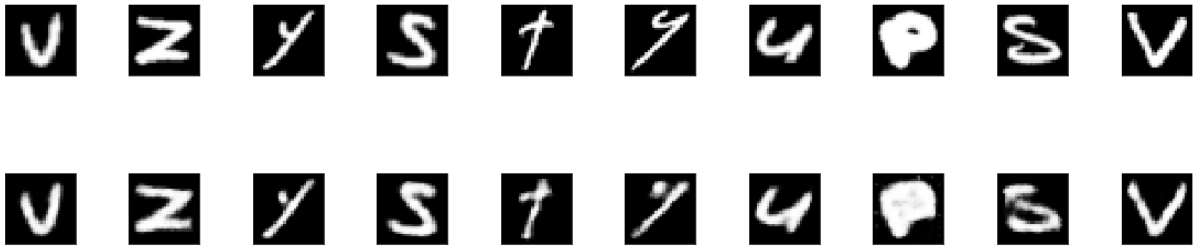


batch 128





batch 64



batch 32