7th International Big Data & Analytics Education Conference
June 3-4, 2019

Case study to develop deep learning image recognition & classification models for fashion items

by

Bharatendra Rai

Charlton College of Business

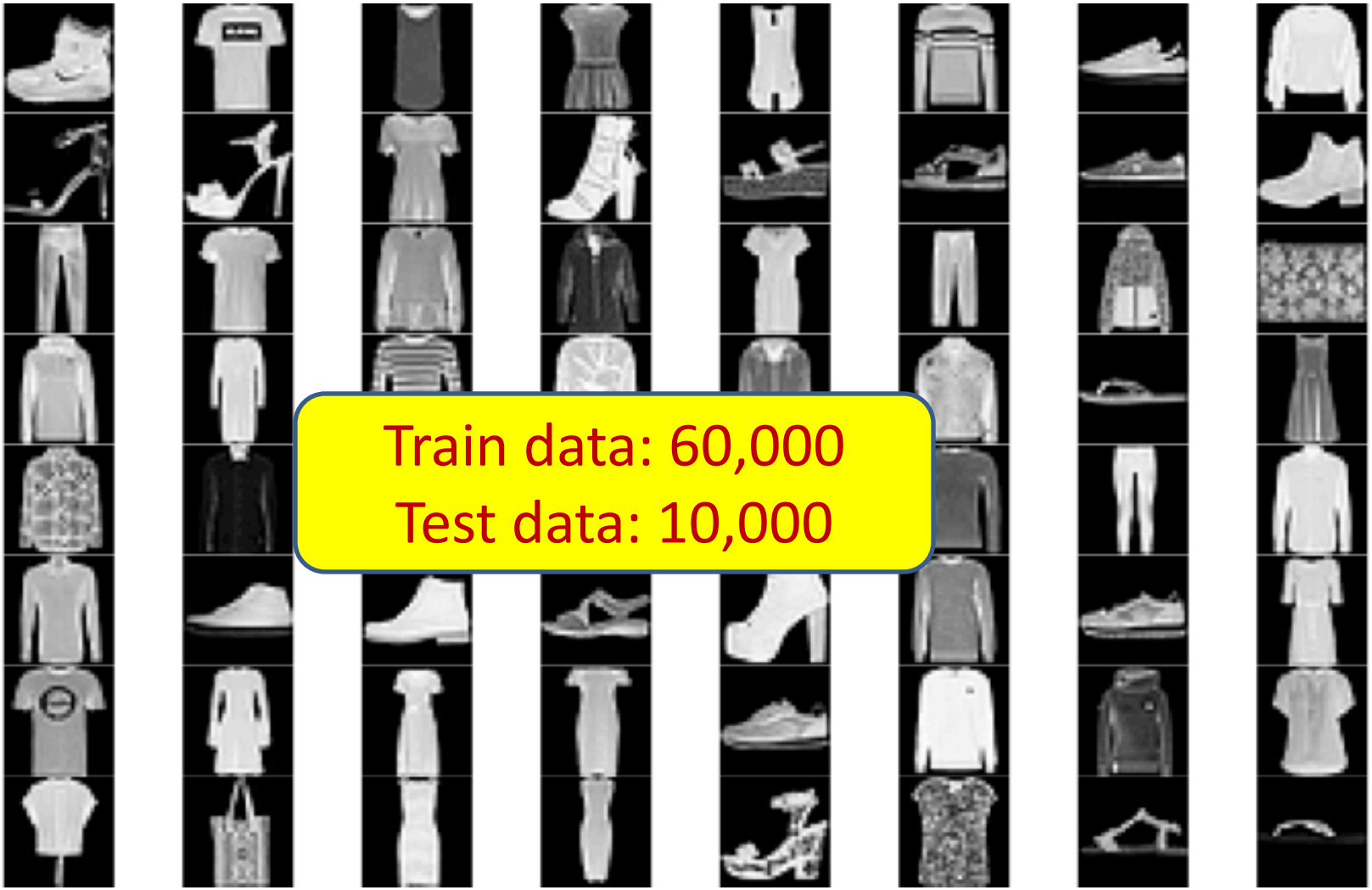
UMass-Dartmouth

Overview

- Data and objective
- MNIST fashion data
- CNN architecture
- Accuracy & loss
- Confusion matrix
- Generalization
- Conclusions

Data & Objective

MNIST Fashion Data

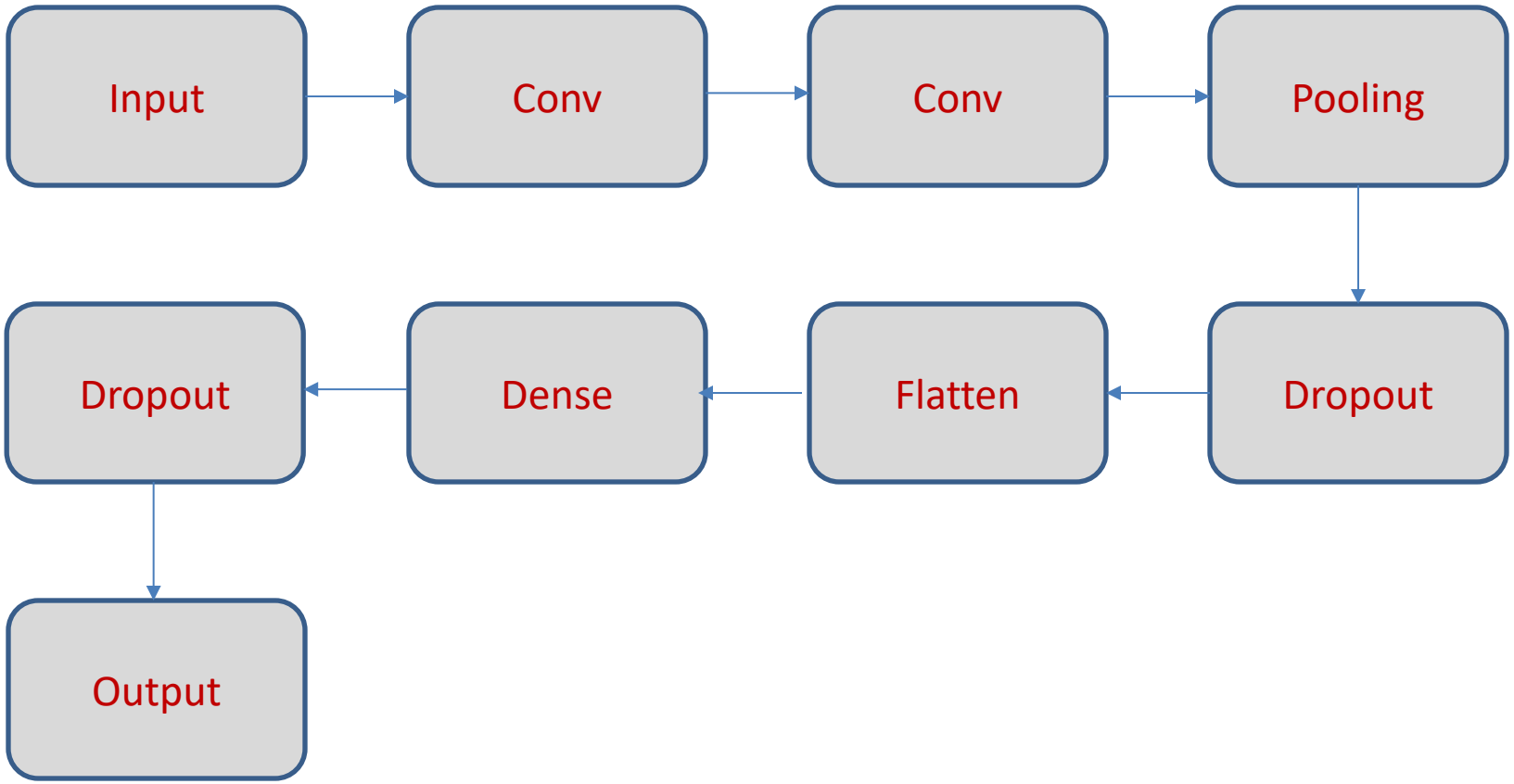


Train data: 60,000
Test data: 10,000

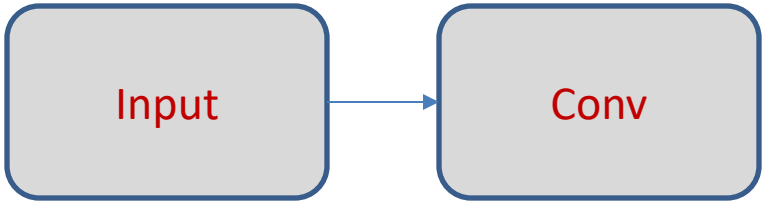
Labels

Label	Description
0	T-shirt/Top
1	Trouser
2	Pullover
3	Dress
4	Coat
5	Sandal
6	Shirt
7	Sneaker
8	Bag
9	Ankle Boot

CNN Architecture



CNN Architecture



CNN parameters:

3 x 3 x 1 x 32 + 32 = **320**

Where,

- 3 x 3 is the kernel size,
- 1 is the number of channels for the image,
- 32 is the number of output filters,
- 32 bias

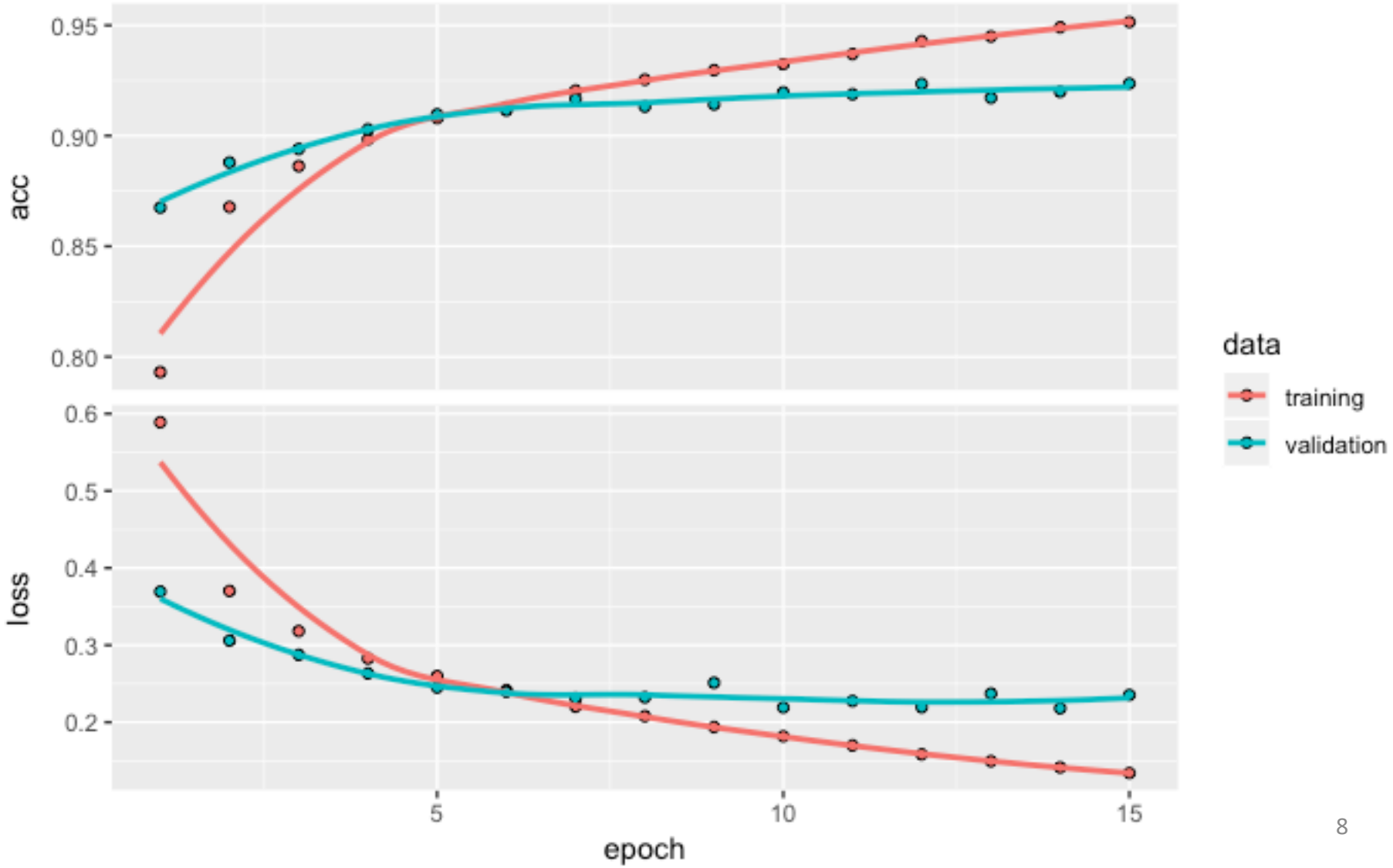
Fully connected network parameters:

Input: 28x28x1 = 784 neurons

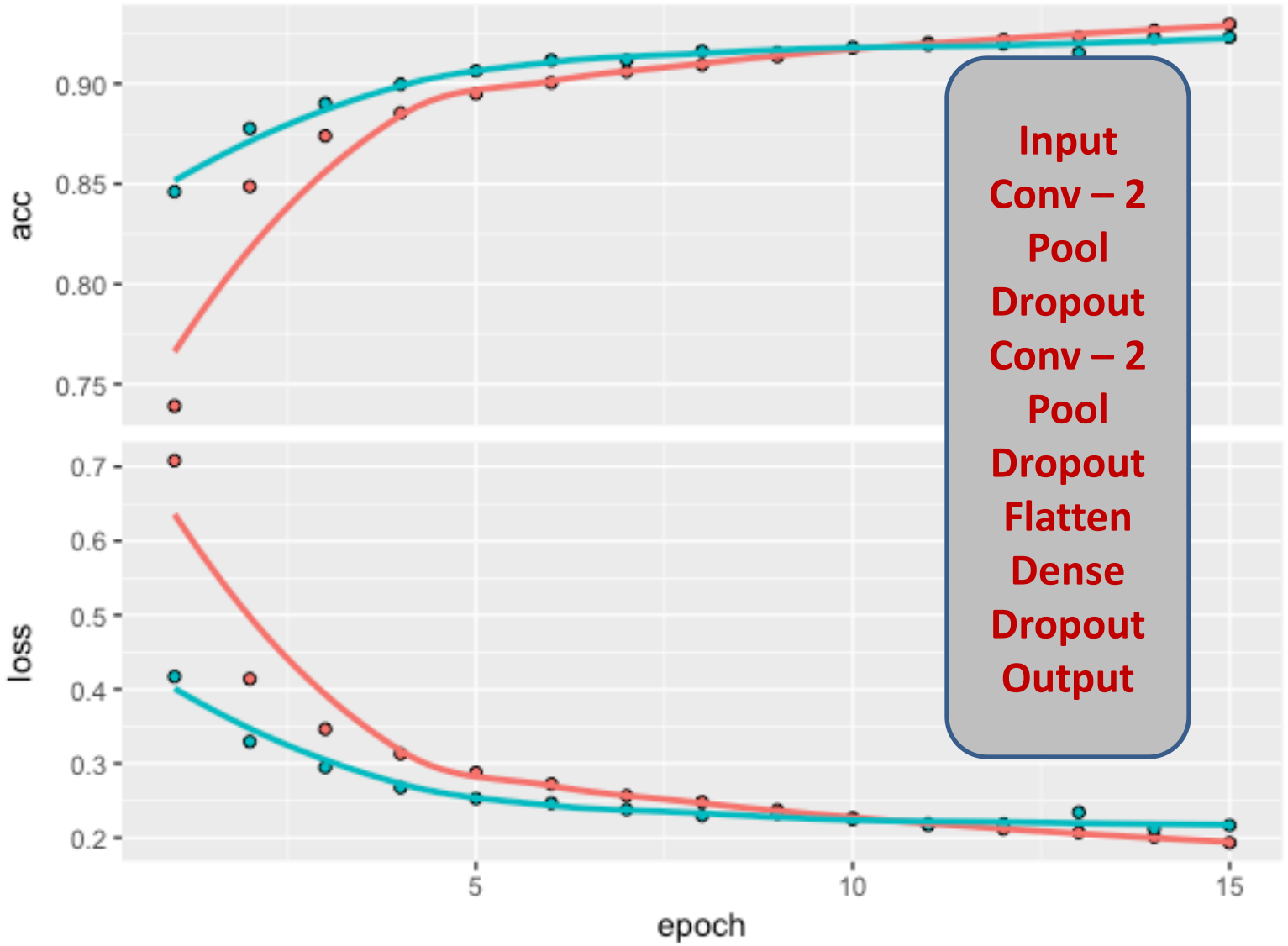
1st layer: 26x26x32 = 21,632 neurons

Total 784x21632 + 21632 = **16,981,120**

Accuracy & Loss - 1



Accuracy & Loss - 2



Confusion Matrix - Train

Predicted \ Actual	0	1	2	3	4	5	6	7	8	9
0	5499	0	58	63	3	0	456	0	4	0
1	2	5936	1	5	3	0	4	0	1	0
2	83	0	5669	13	258	0	438	0	7	0
3	69	52	48	5798	197	0	103	0	6	0
4	3	3	136	49	5348	0	265	0	5	0
5	0	0	0	0	0	5879	0	3	0	4
6	309	6	73	67	181	0	4700	0	2	0
7	0	0	0	0	0	75	0	5943	1	169
8	35	3	15	5	10	3	34	0	5974	2
9	0	0	0	0	0	43	0	54	0	5825

Label Description

0	T-shirt/Top
1	Trouser
2	Pullover
3	Dress
4	Coat
5	Sandal
6	Shirt
7	Sneaker
8	Bag
9	Ankle Boot

Accuracy: 94.3%

Confusion Matrix - Test

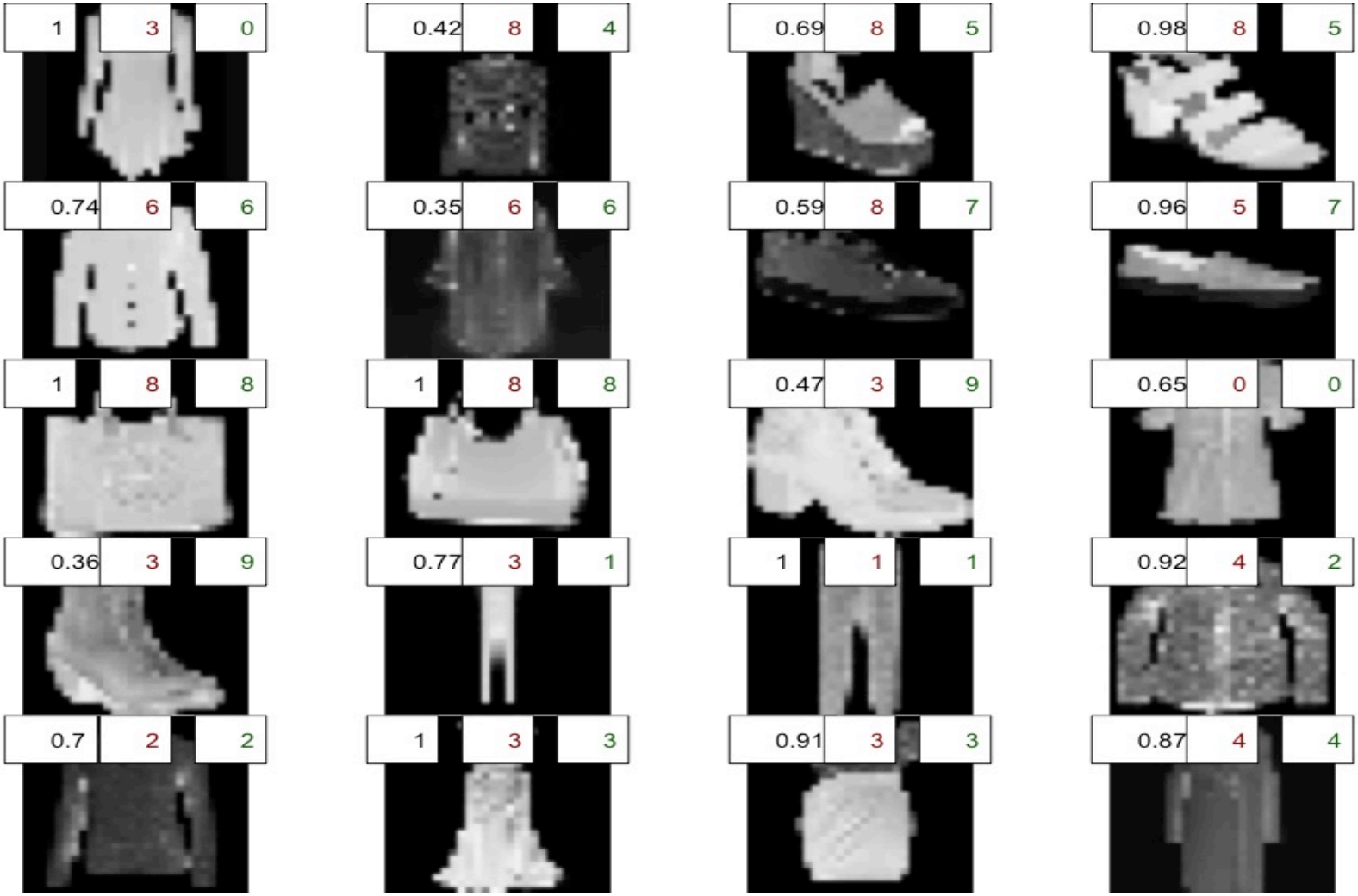
	Actual									
Predicted	0	1	2	3	4	5	6	7	8	9
0	875	1	18	8	0	0	104	0	3	0
1	0	979	0	2	0	0	0	0	0	0
2	19	0	926	9	50	0	78	0	1	0
3	10	14	9	936	35	0	19	0	3	0
4	2	0	30	12	869	0	66	0	0	0
5	0	0	0	0	0	971	0	2	1	2
6	78	3	16	29	45	0	720	0	1	0
7	0	0	0	0	0	18	0	988	1	39
8	16	3	1	4	1	0	13	0	989	1
9	0	0	0	0	0	11	0	10	1	958

Label Description

0	T-shirt/Top
1	Trouser
2	Pullover
3	Dress
4	Coat
5	Sandal
6	Shirt
7	Sneaker
8	Bag
9	Ankle Boot

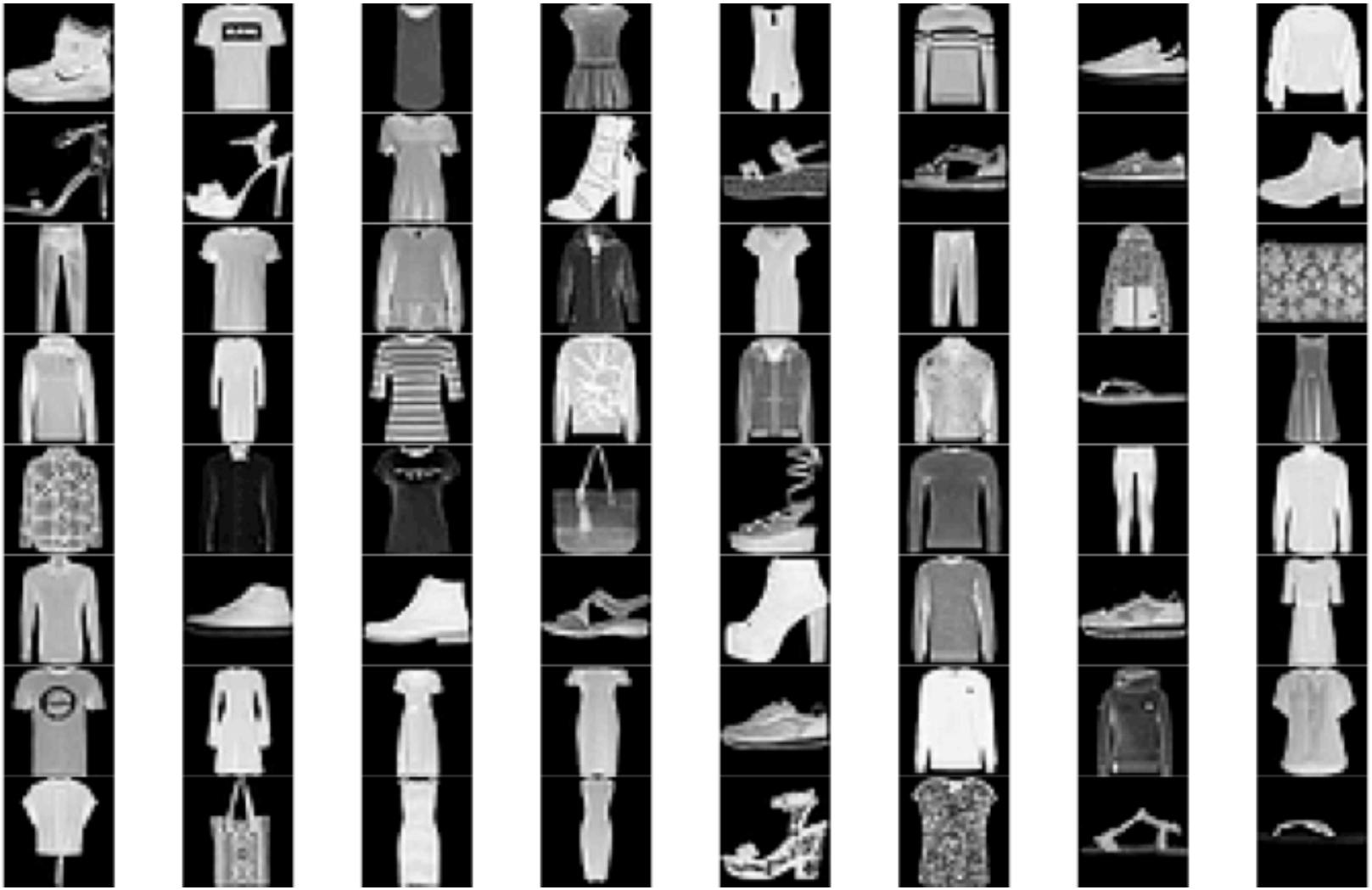
Accuracy: 92.1%

Generalization…

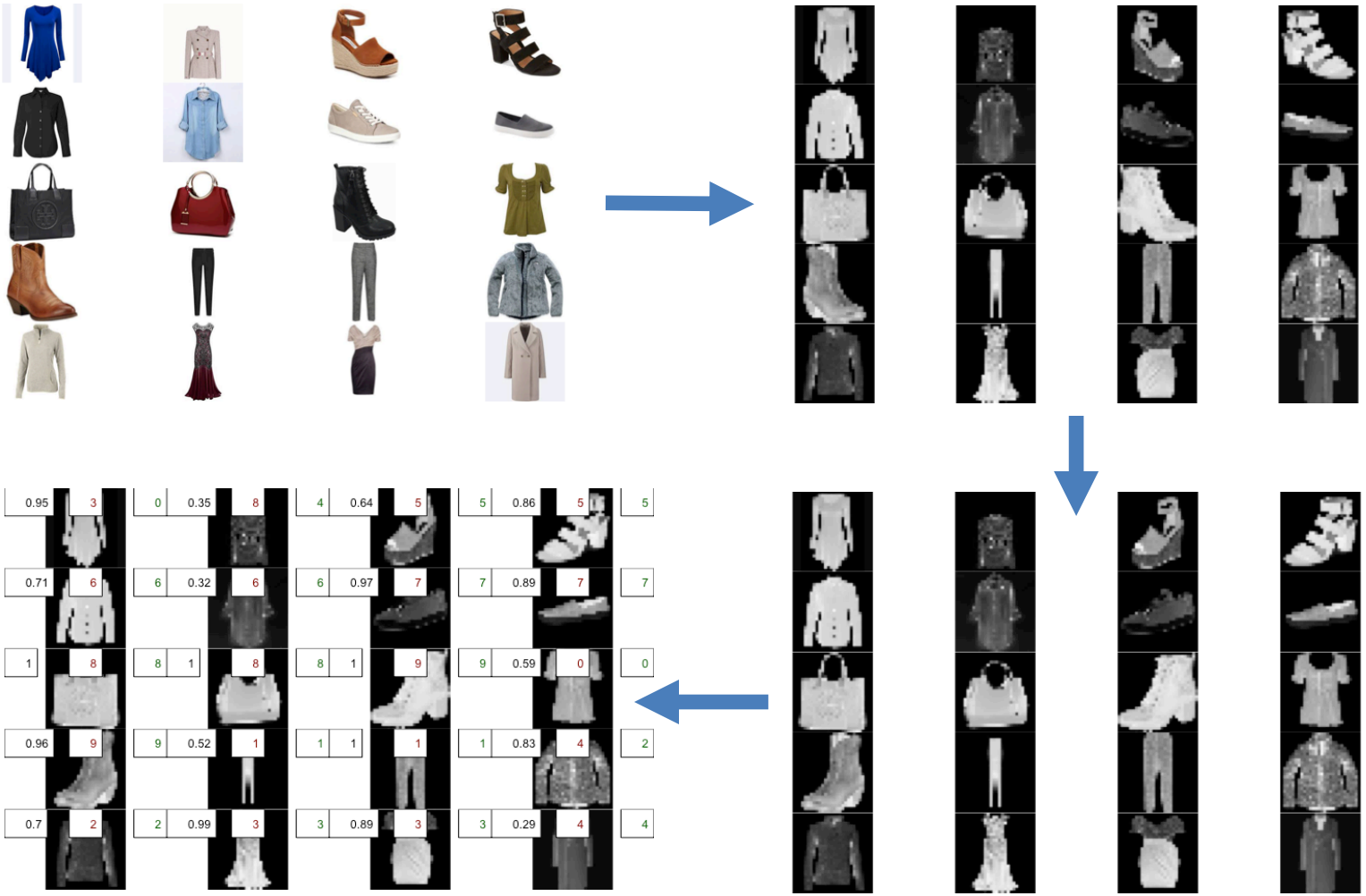


Accuracy: 50%

MNIST Fashion Data



Generalization



Accuracy: 85%

Conclusions

CNN network helps to reduce number of parameters.

Dropout layers can help reduce overfitting.

Validation split of x% chooses last x% of train data.

Generalization to new data is challenging.

Thank You