Project Machine Learning

By Leonardo pajer

- Machine Learning field is fluid
- New algorithms are complex and time consuming
- Old and sound algorithms have been discarted
- Dataset dependency of algorithms

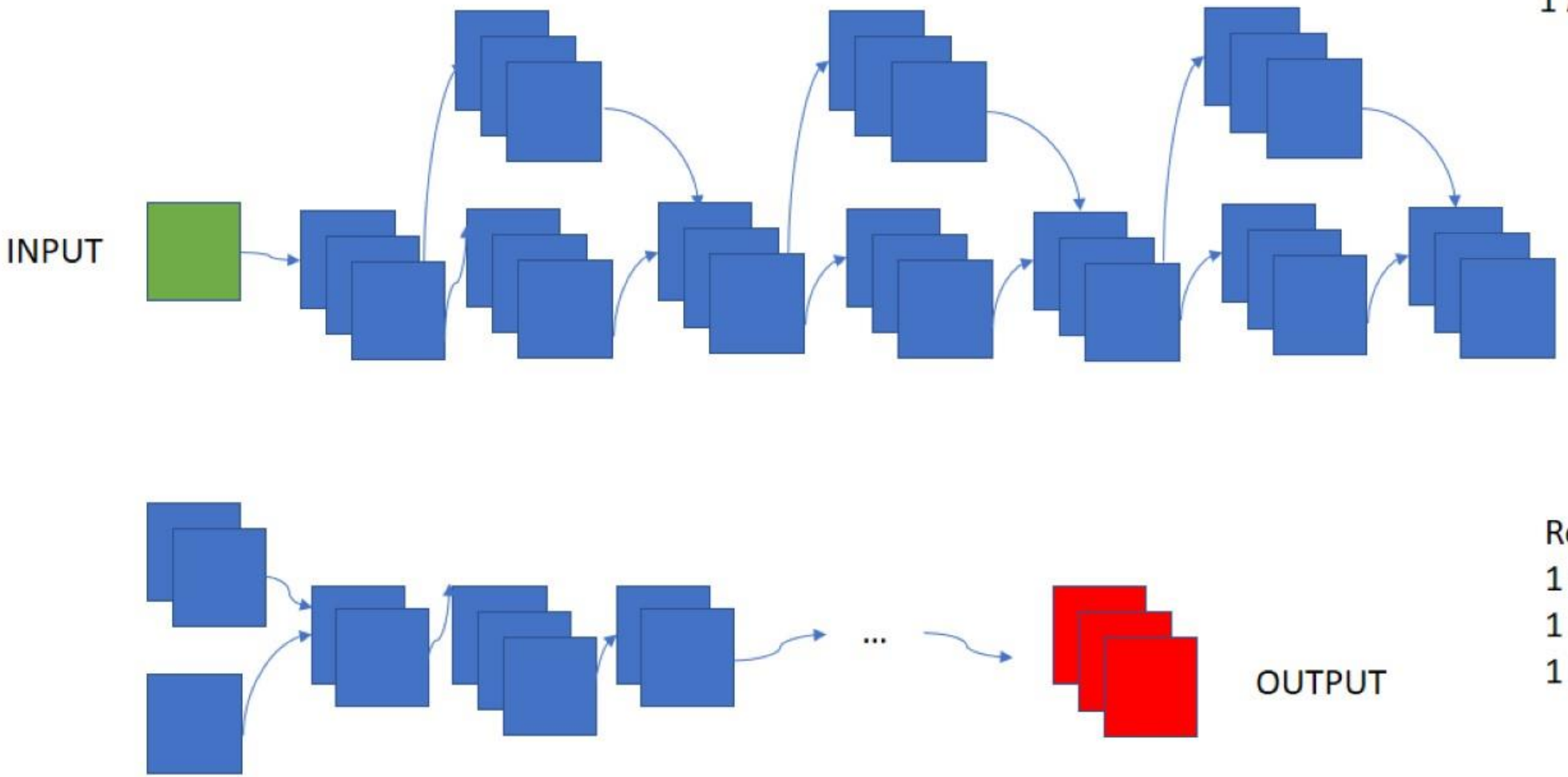
Introduction to the analyzed problem

Dataset description

- Intel Image classification dataset
- Circa 14k images
- Shape size 150x150 RGB

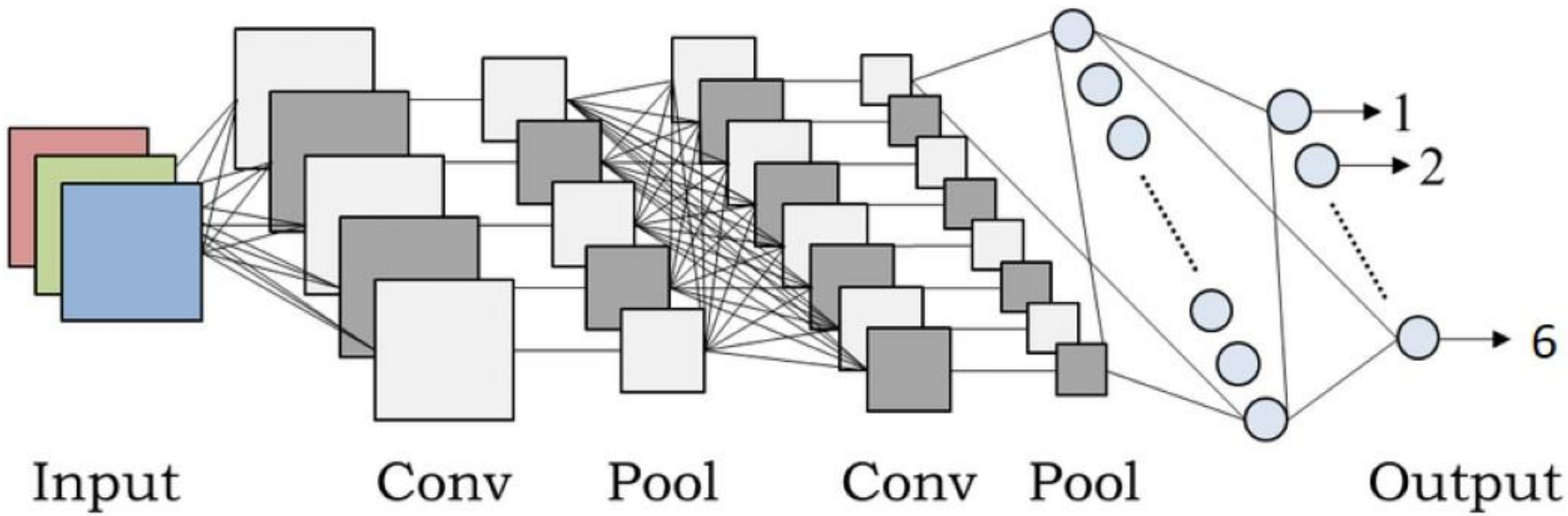
First half and output architecture ResnNet20

Blue blocks:
1 Conv2D layer
1 Batch Norm
1 Activation



Red block:
1 average Pooling
1 Flatten
1 Dense

CNN architecture



- Neurons in the input layer
- Kernel initializer
- Dropout ratio
- Activation function

List of parameters

	Training accuracy (%)	Test accuracy (%)
CNN	85	81
ResNet	96	73

Results

- ResNet20 almost perfect in training. Overfitting?
- CNN 10% better in test set
- CNN twice faster
- ResNet20 not optimized

Conclusions

References

Dataset source: https://www.kaggle.com/puneet6060/intel-image-classification