Classification of agricultural land use by ensemble of convolutional neural networks

Nathaniel Newlands and Etienne Lord

nathaniel.newlands@canada.ca, researcher

etienne.lord@canada.ca, researcher

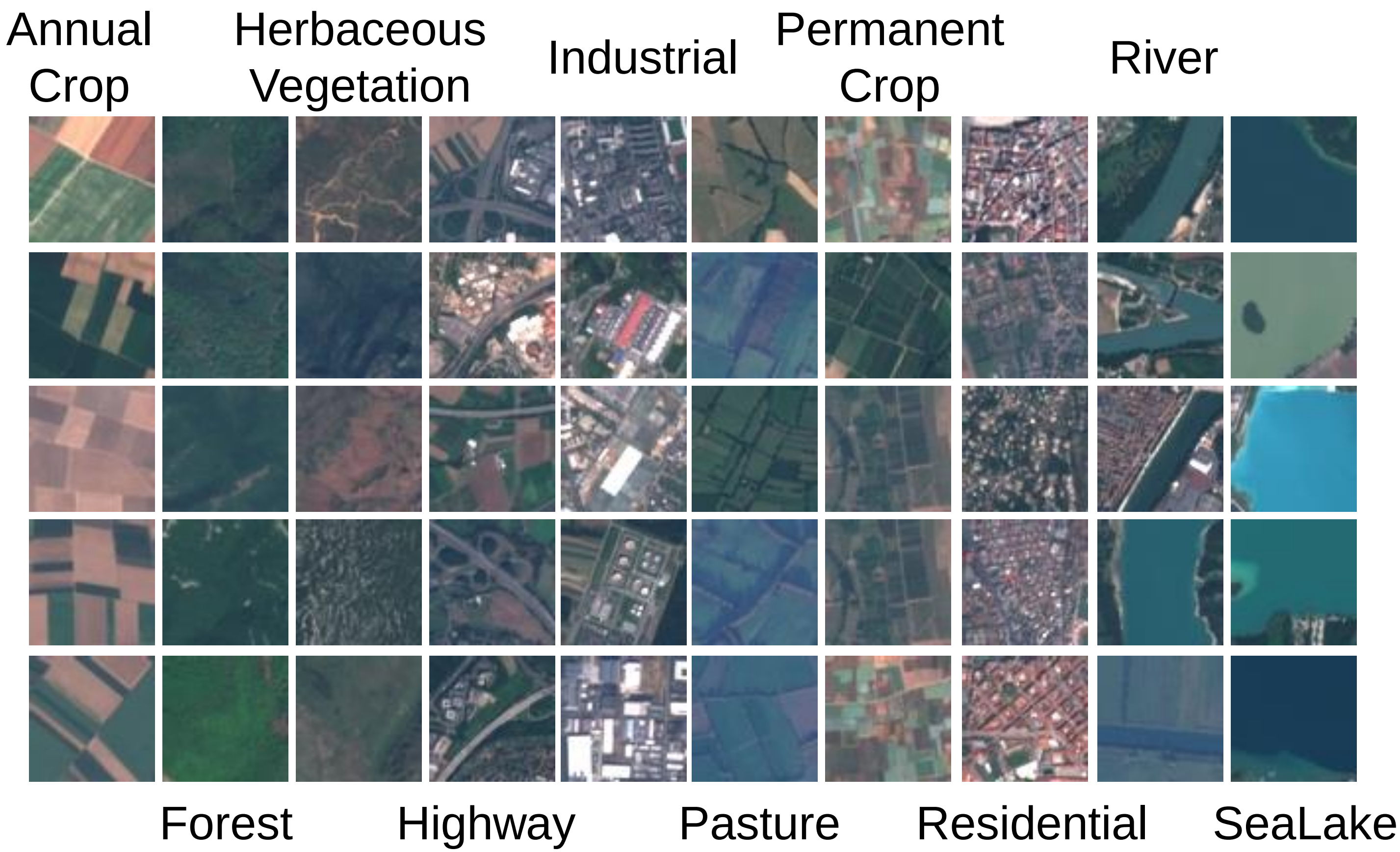
Introduction

What is the correct temporal and spatial resolution for precision agriculture?

Google image

EuroSat (sentinel-2)

10 land use and land cover classes including 27,000 images



Helber P, Bischke B, Dengel A, Borth D. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. arXiv preprint arXiv:1709.00029. 2017 Aug 31.

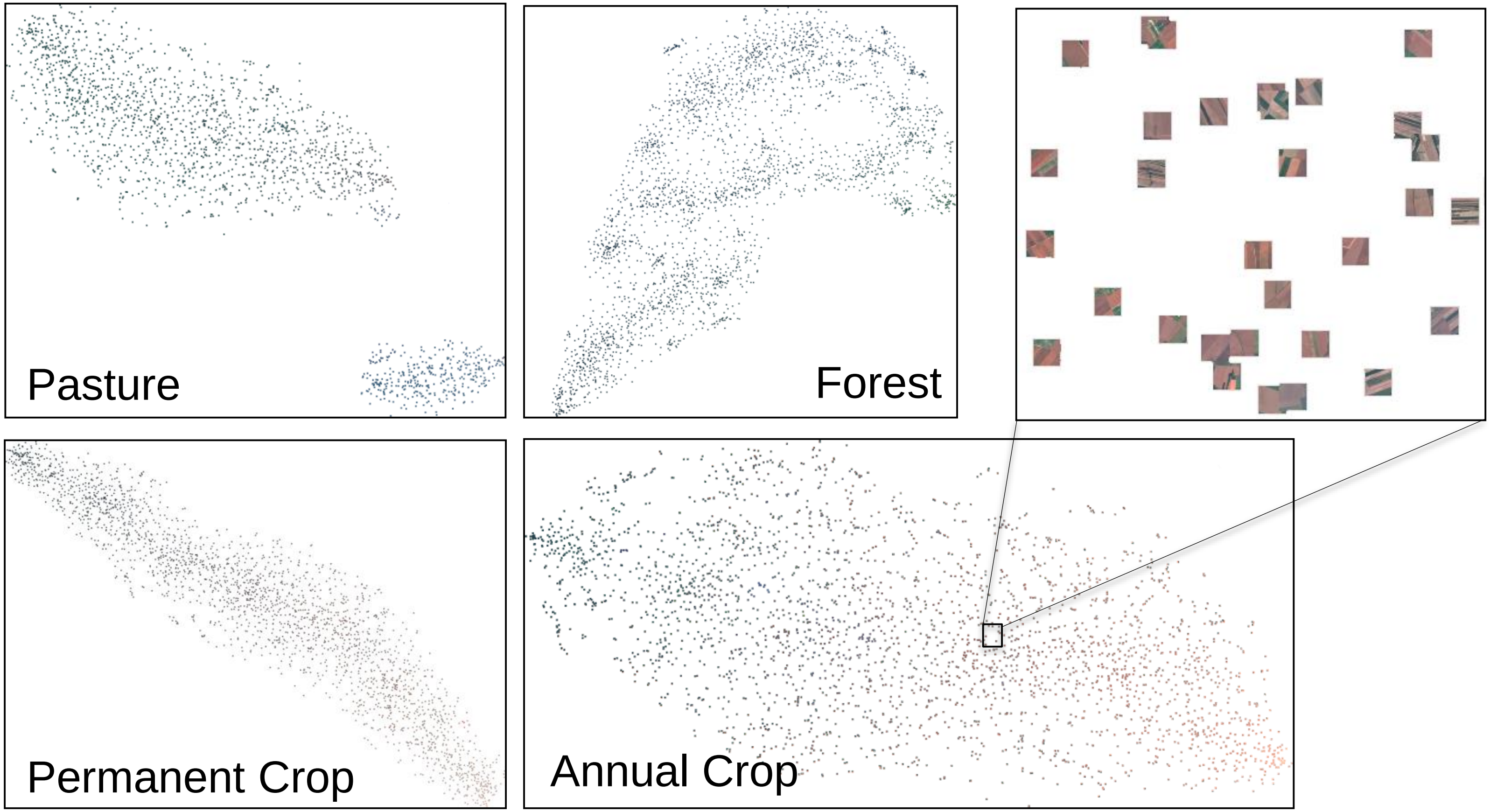
Goals is evolution of land use

However, they only verified single band or RGB classification accuracy

Helber et al. (2017) used a ResNet-50-based convolutional neural network.

t-SNE clustering

t-Distributed Stochastic Neighbor Embedding is a dimensionality reduction method*.
The dataset images show diverse distributions over the RGB bands

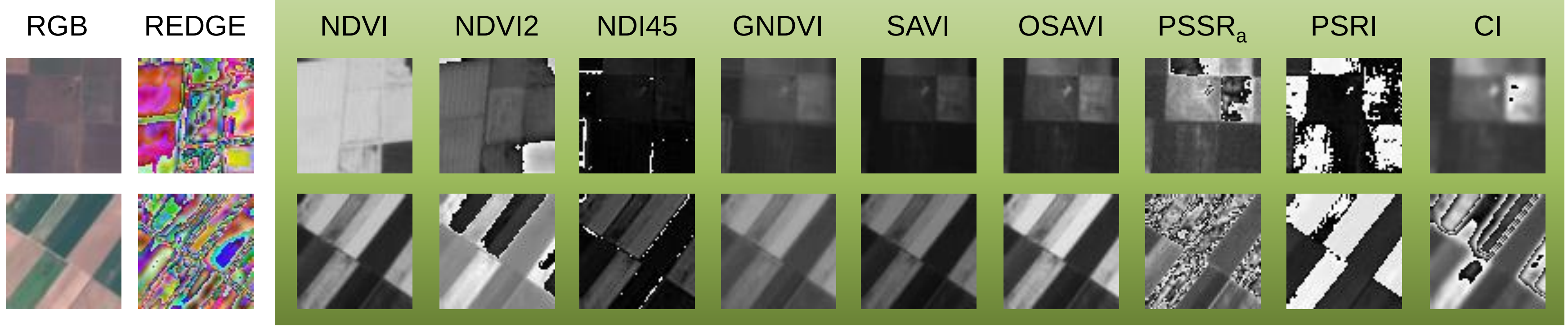


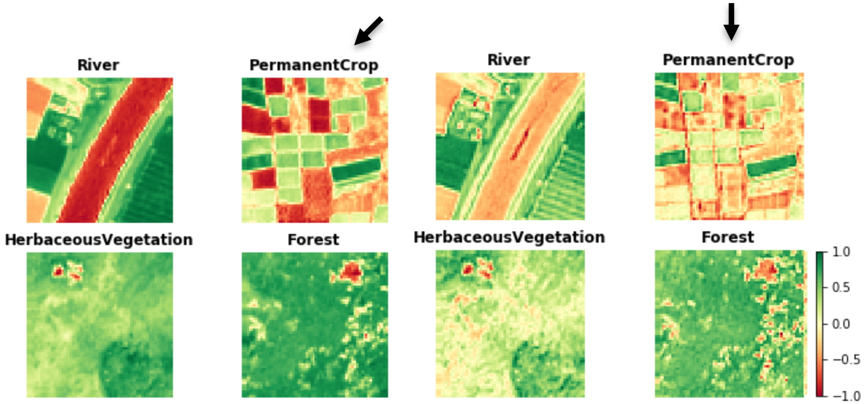
***As opposed to PCA, there is no reduction in the signal in the pairwise similarity search.**

Chan DM, Rao R, Huang F, Canny JF. t-SNE-CUDA: GPU-Accelerated t-SNE and its Applications to Modern Data. In2018 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD) 2018 Sep 24 (pp. 330-338). IEEE.

Vegetation Indices

Xue and Su (2017) reviewed 108 common vegetation indices, we selected 9



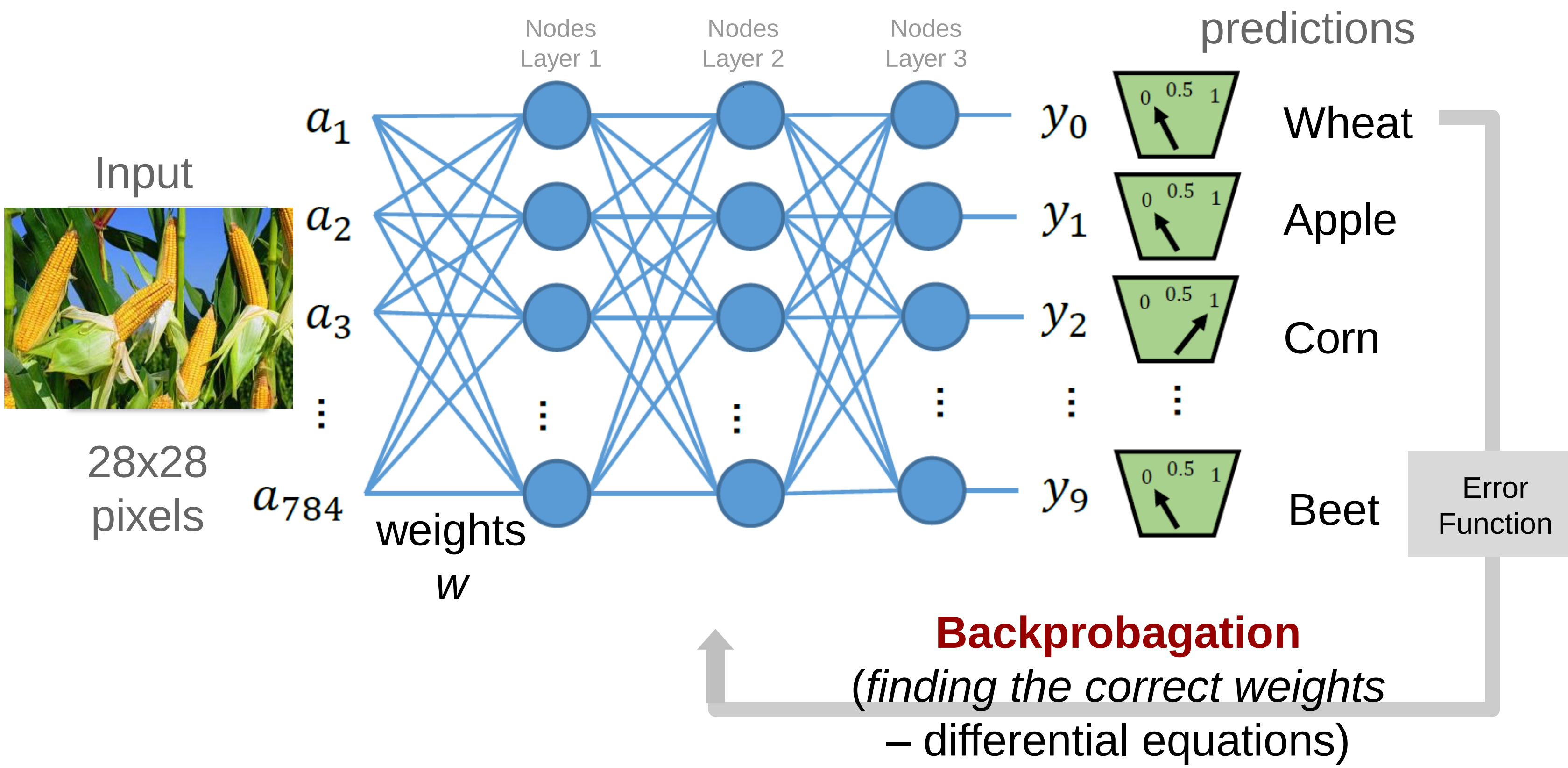


	Sentinel-2 bands formulas	References
NDVI	(B7-B4)/(B7+B4)	Rouse et al. 1973
NDVI2	(B8A-B4)/(B8A+B4)	Datt, 1999
NDI45	(B5-B4)/(B5+B4)	Delegido et al. 2011
GNDVI	(B7-B3)/(B7+B3)	Gitelson et al. 1996
SAVI	(B8-B4)/(B8+B4+0.5)*1.5	Huete, 1998
OSAVI	1.16*(B8-B4)/(B8+B4+0.61)	Rondeaux et al. 1996
PSSRa	B7/B4	Blackburn 1998
PSRI	(B4-B2)/B6	Merzlyak et al. 1999
CI	(B7/B5)-1	Gitelson et al. 2003

Xue J, Su B. Significant remote sensing vegetation indices: a review of developments and applications. J of Sensors 2017.

Deep Learning

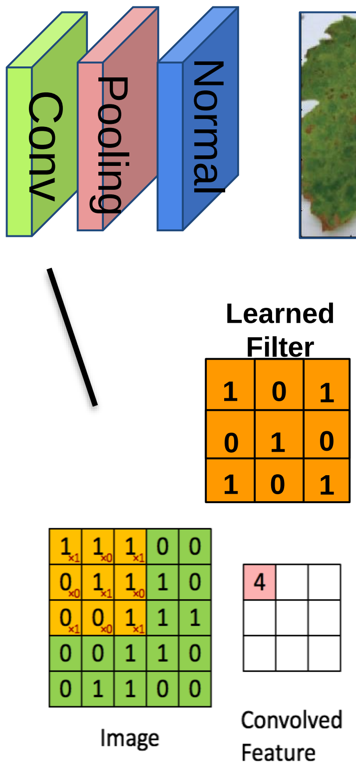
Simple model of a deep neural network



Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". Nature. 323 (6088): 533–536.

Convulational neural networks

Each layer “learn” a different representation of the data



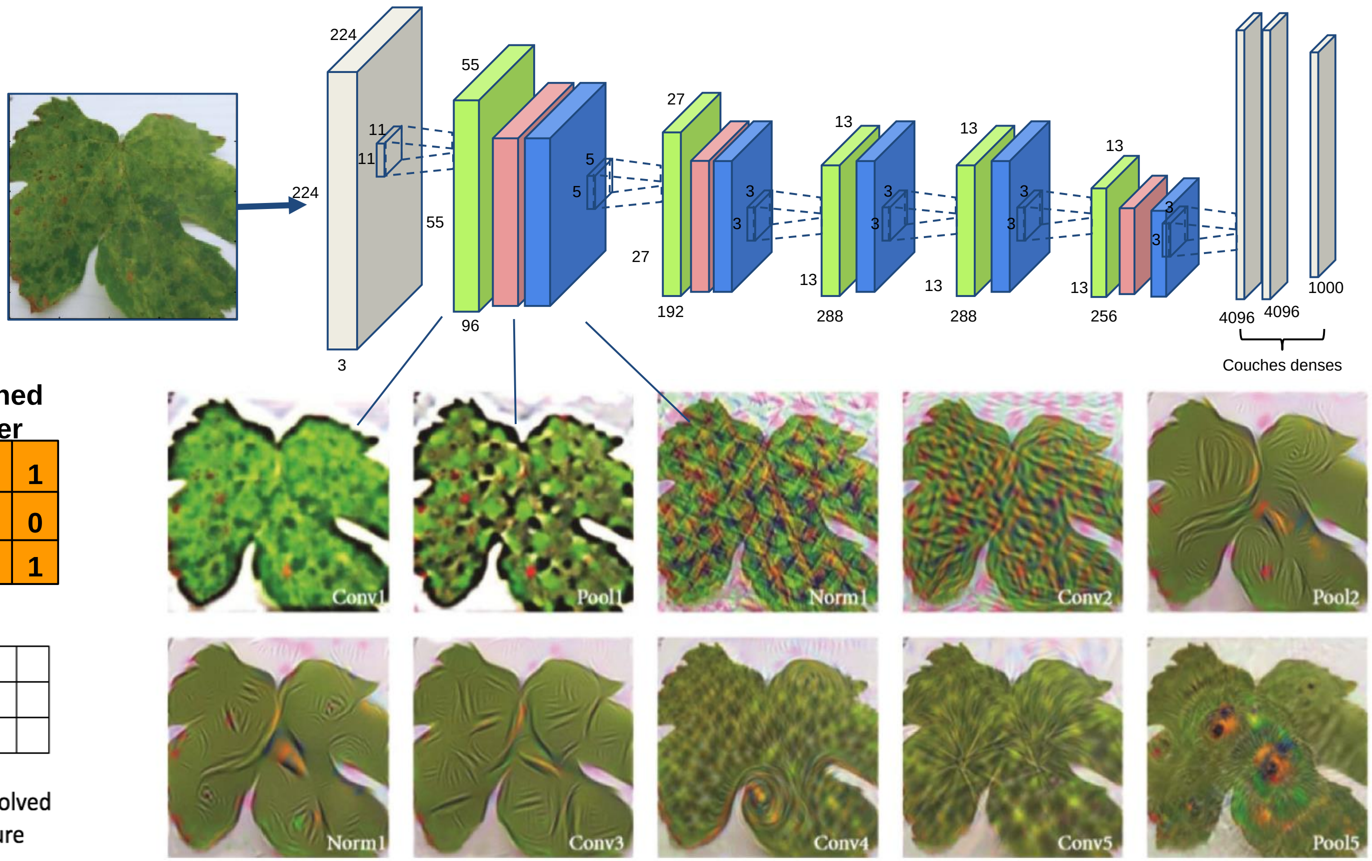
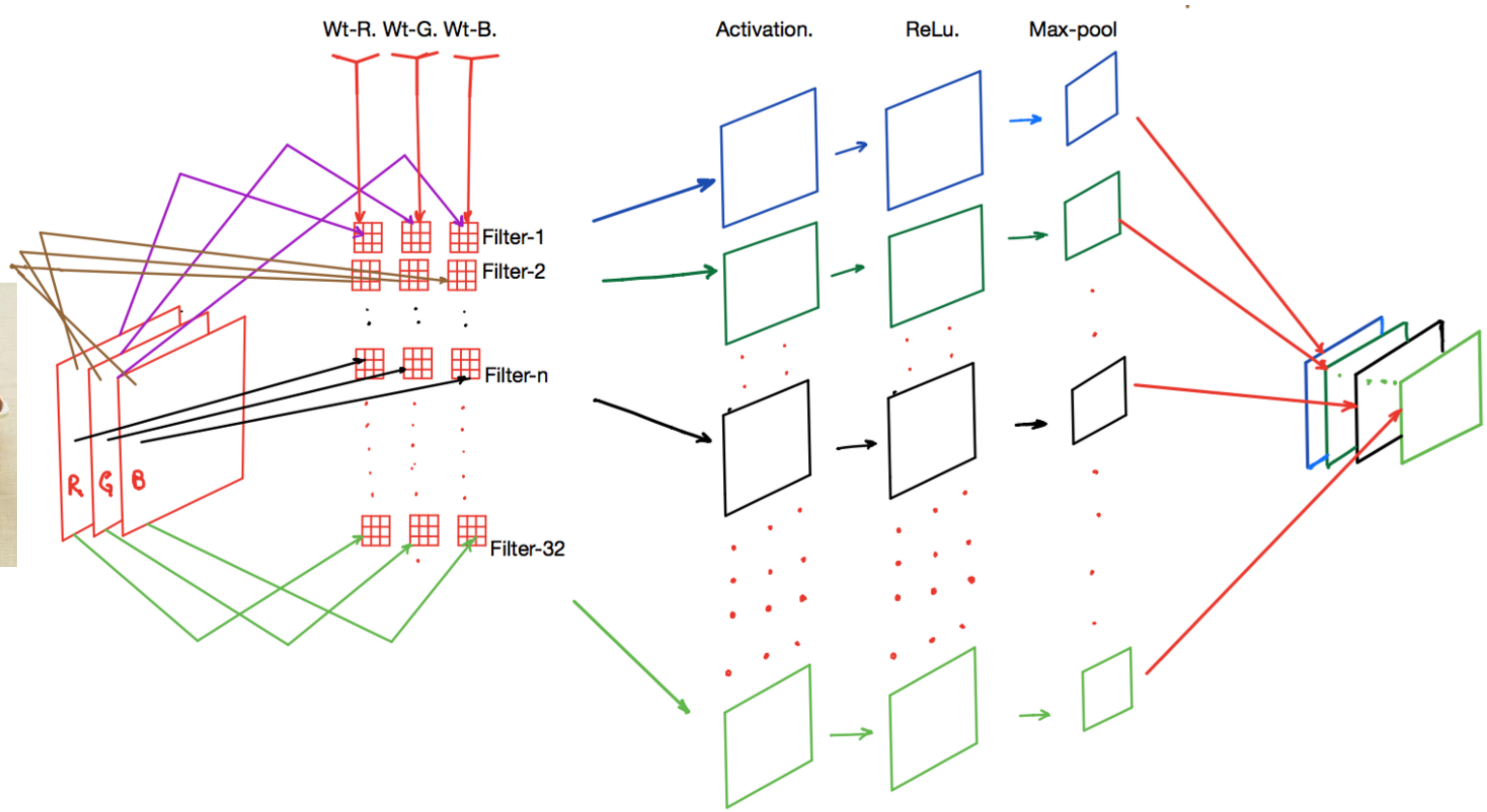


Fig. 2. Visualization of the output layers images after each processing step of the CaffeNet CNN (i.e. convolution, pooling, normalization) at a plant disease identification problem based on leaf images.
Source: Sladojevic et al. (2016).

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*, 70-90

Convolution layers

A layer where **filters are learned** and applied to each input data. The inverse operation is the **deconvolution**. In this layer, we want to "**learn"** some characteristics of the input data (shape, pattern, etc.).



https://medium.com/@apiltamang/a-gentle-dive-into-the-anatomy-of-a-convolution-layer-6f1024339aca

Convolution layers

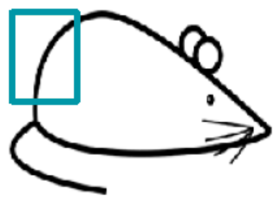
Filters are basically "multiplicative" masks applied to the different part of an image.

Filter

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

Visualization of a curve detector filter



Visualization of the receptive field

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

Pixel representation of the receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

Multiplication and Summation = (50*30)+(50*30)+(50*30)+(20*30)+(50*30) = 6600 (A large number!)

Visualization of the filter on the image

0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of receptive field

*

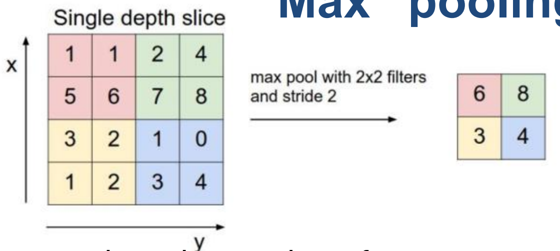
0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

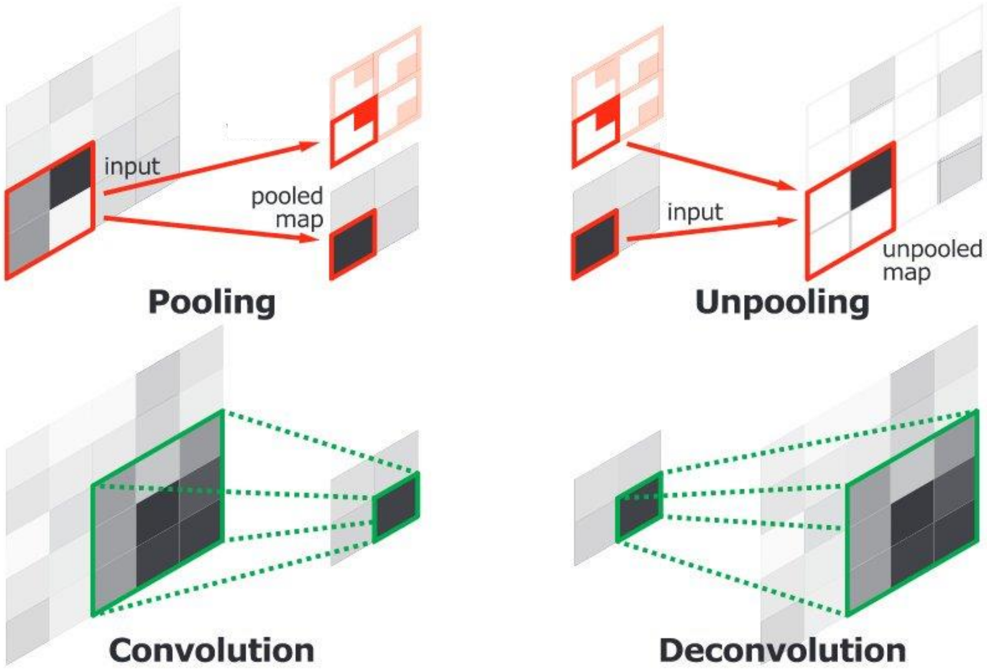
Multiplication and Summation = 0

Other layer types

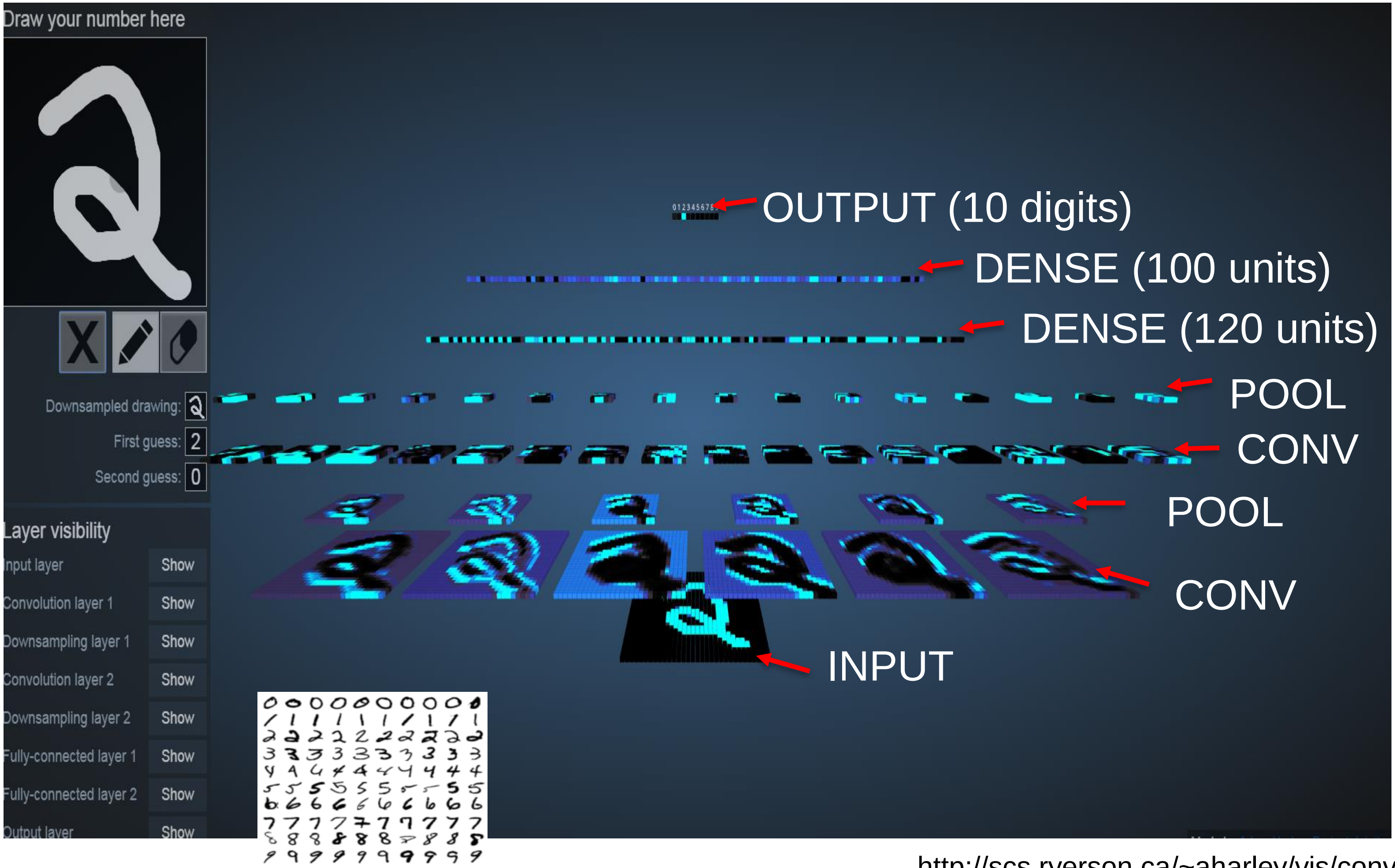
Max "pooling"



- Reduce the number of parameters
- Add resistance to local noise



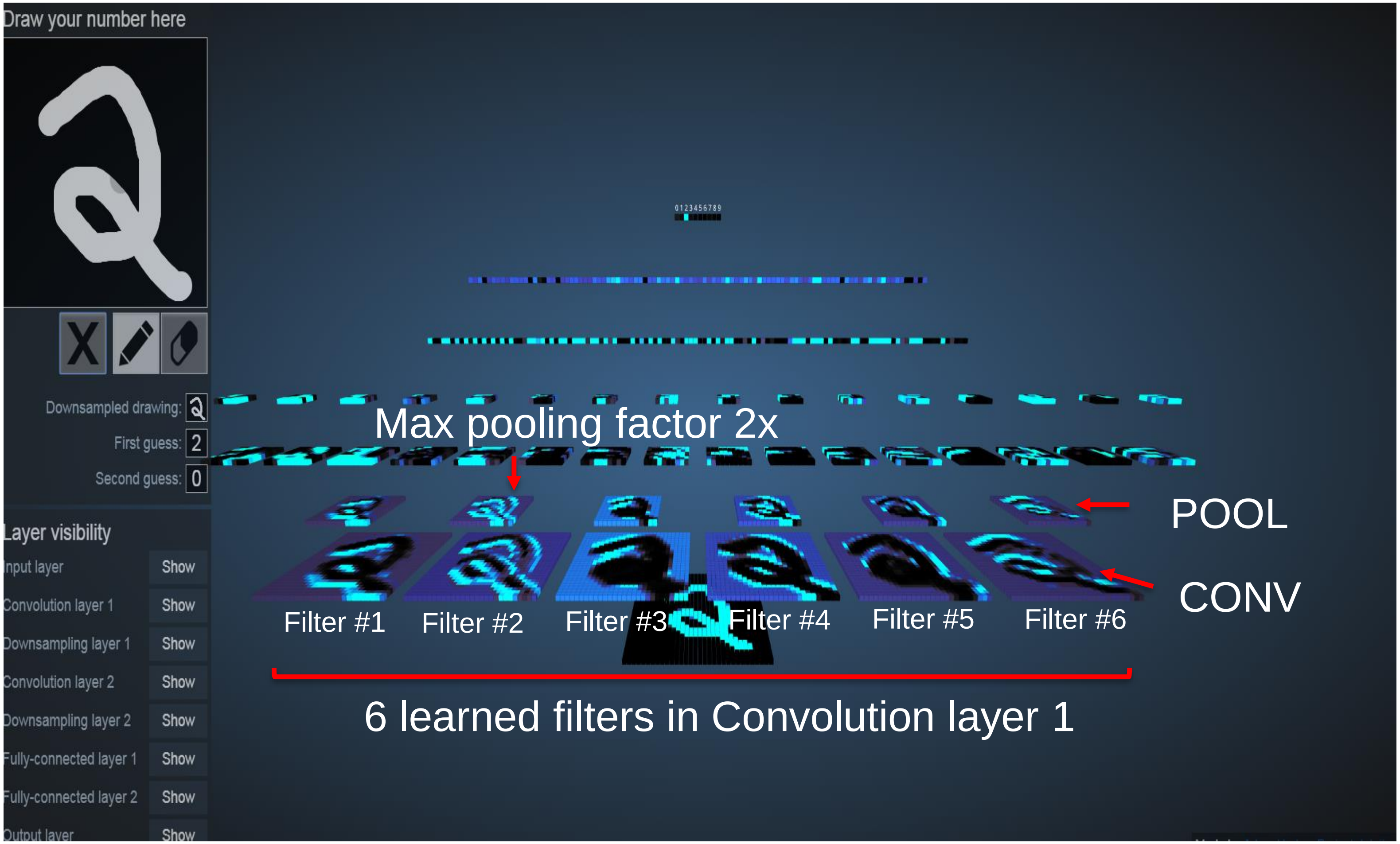
Convolution MNIST example (1/5)



yann.lecun.com/exdb/mnist

http://scs.ryerson.ca/~aharley/vis/conv/

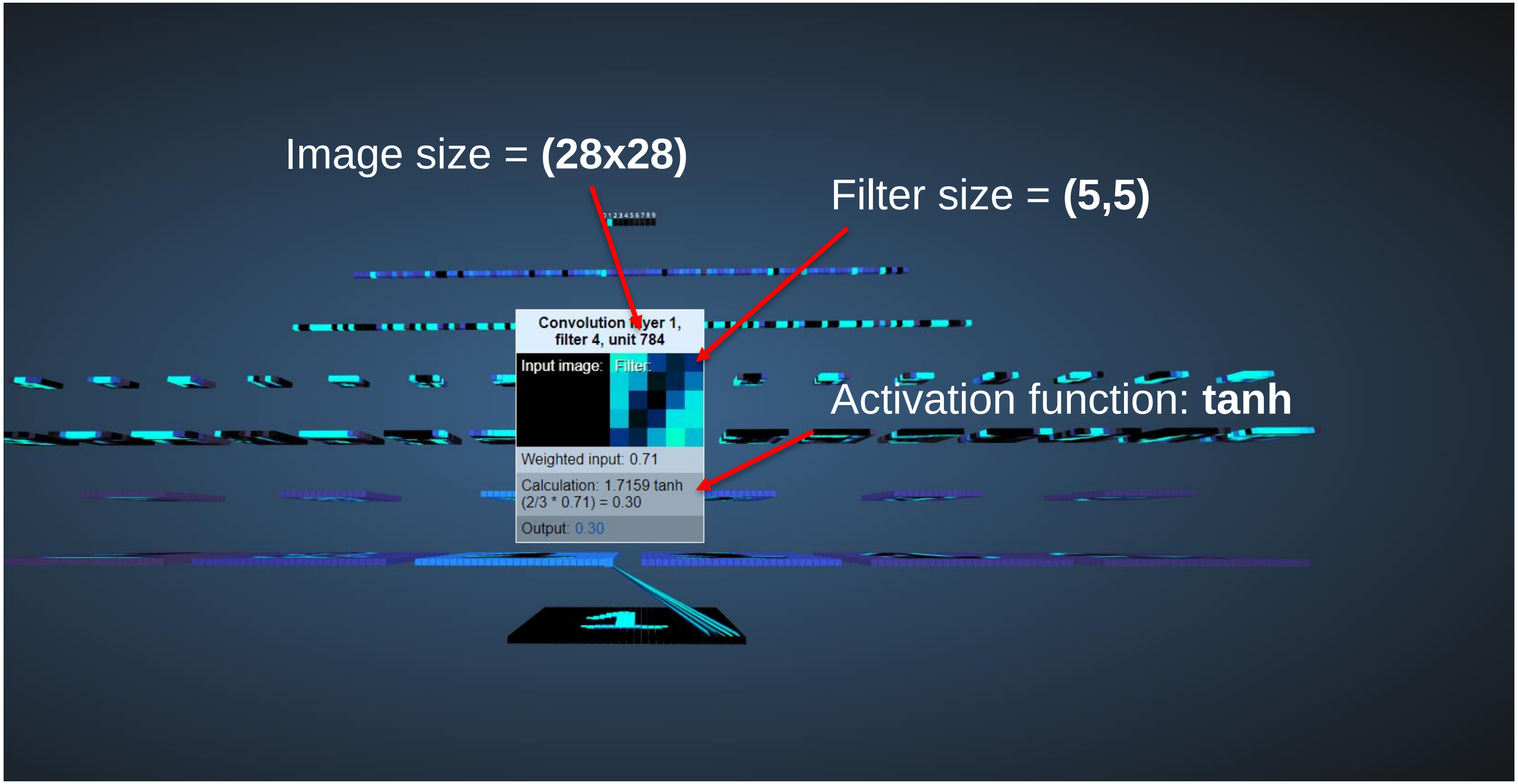
Convolution MNIST example (2/5)



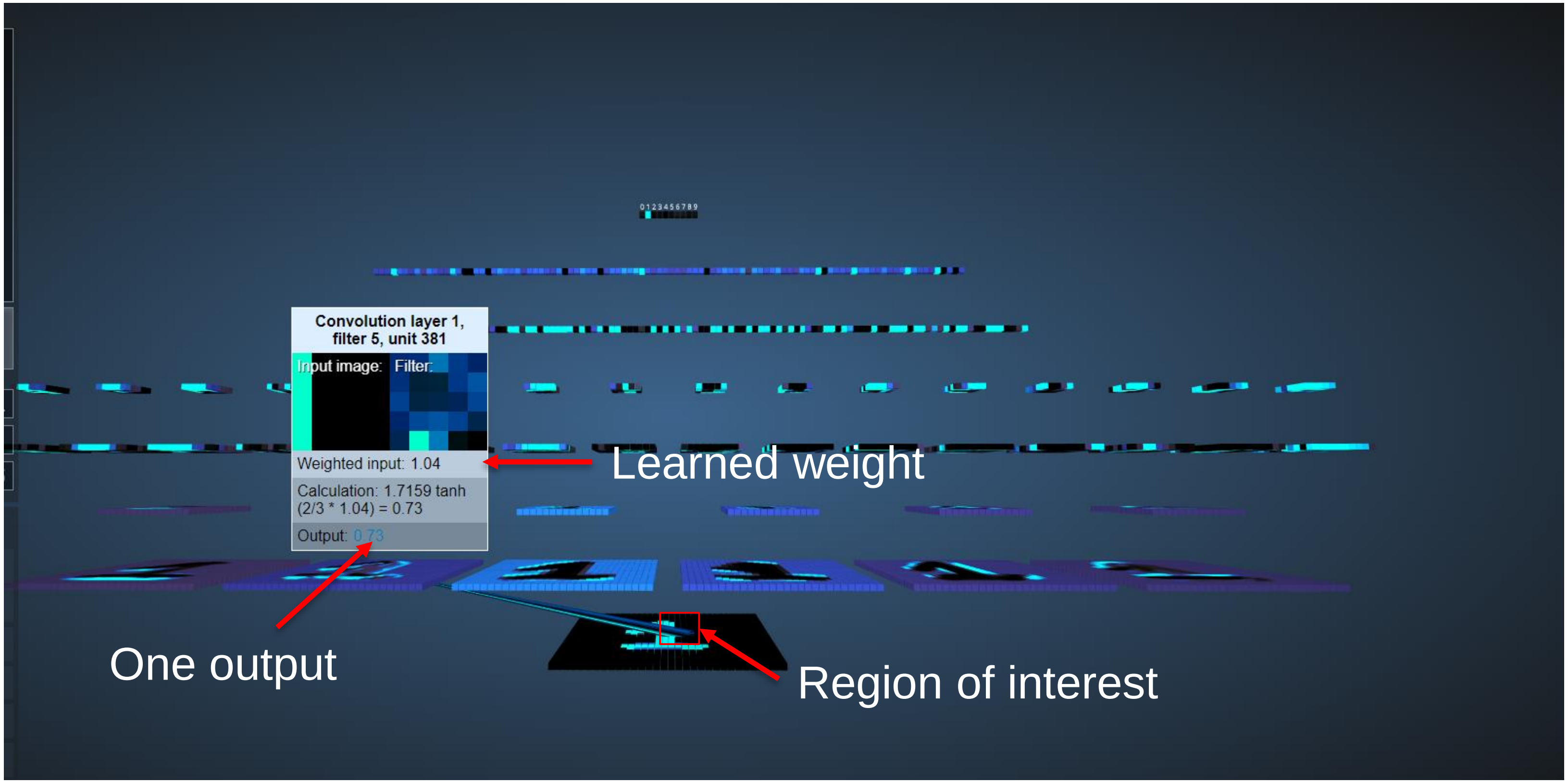
http://scs.ryerson.ca/~aharley/vis/conv/

yann.lecun.com/exdb/mnist

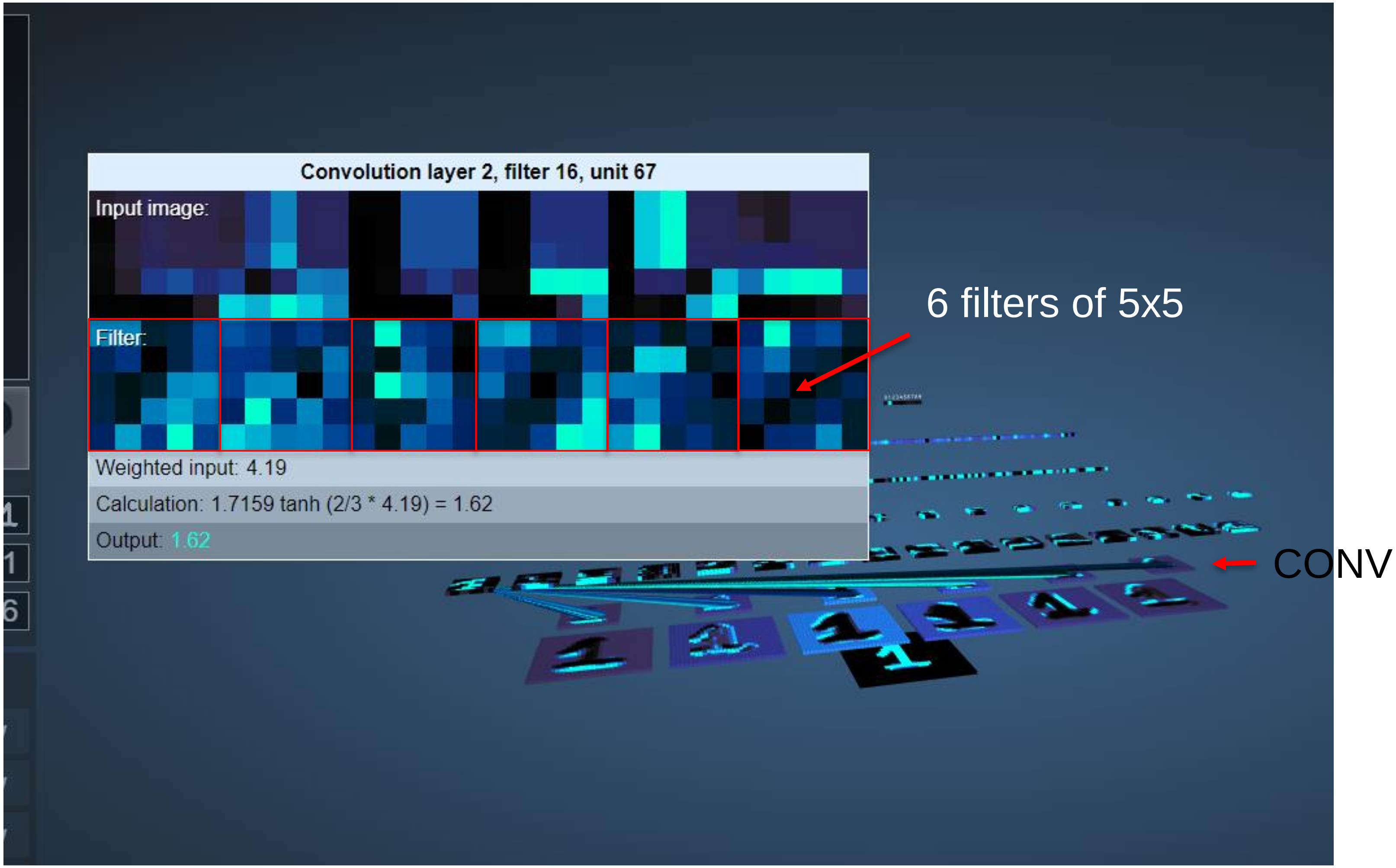
Convolution MNIST example (3/5)



Convolution MNIST example (4/5)



Convolution MNIST example (5/5)



Replicating with Keras

Keras

Number of filters

Filter size

Image size

Activation function

Pool (2x)

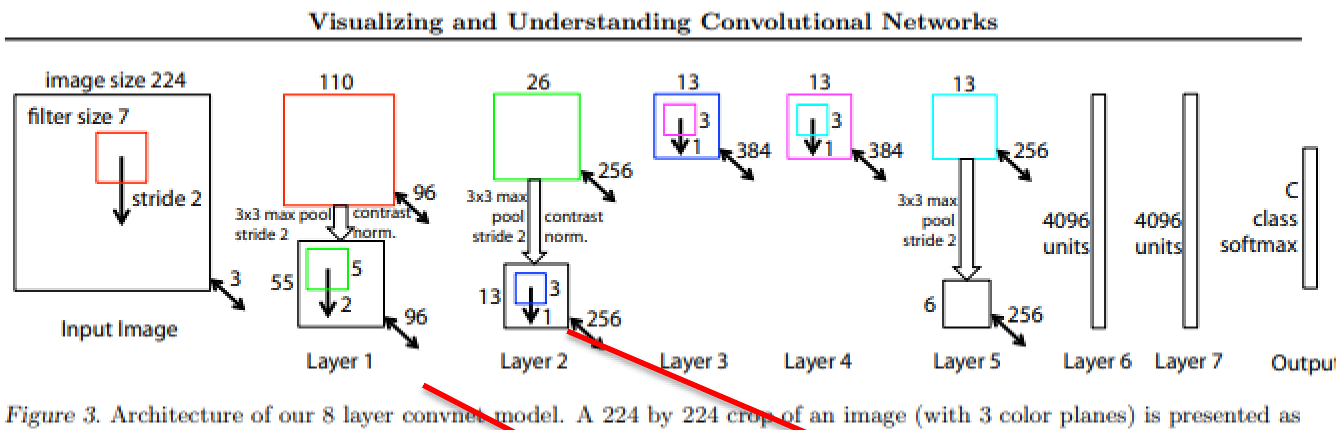
LAYER 1

LAYER 2

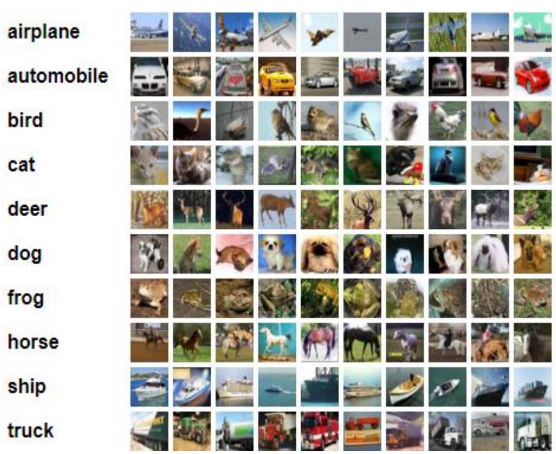
LAST LAYERS

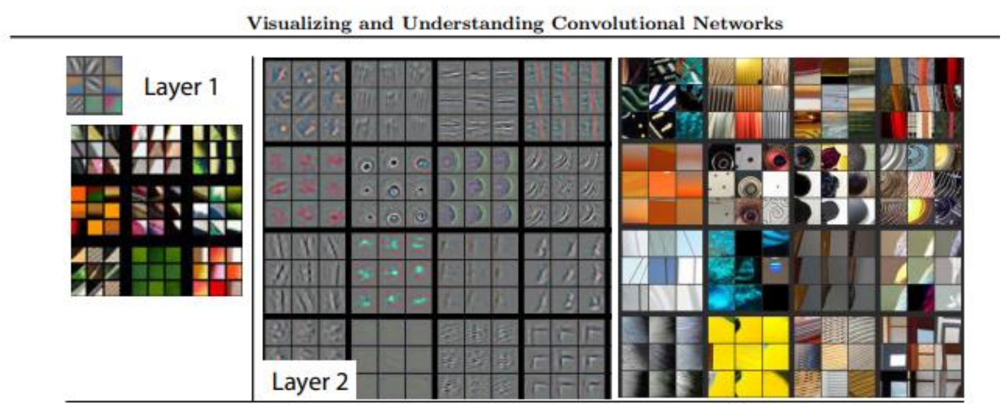
```
model = Sequential()

model.add(Conv2D(6, (5,5), input_shape=(28,28)))
model.add(Activation('tanh'))
model.add(MaxPooling2D(pool_size=(2, 2)))

model.add(Conv2D(16, (5,5)))
model.add(Activation('tanh'))
model.add(MaxPooling2D(pool_size=(2, 2)))

# Last layers
model.add(Flatten()) # We flatten to convert 2D to 1D
model.add(Dense(120))
model.add(Activation('tanh'))
model.add(Dense(100))
model.add(Activation('tanh'))
model.add(Dense(10)) # One last cell, the digits 0 to 9
model.add(Activation('tanh'))
```

Learning characteristics (1/2)



CIFAR

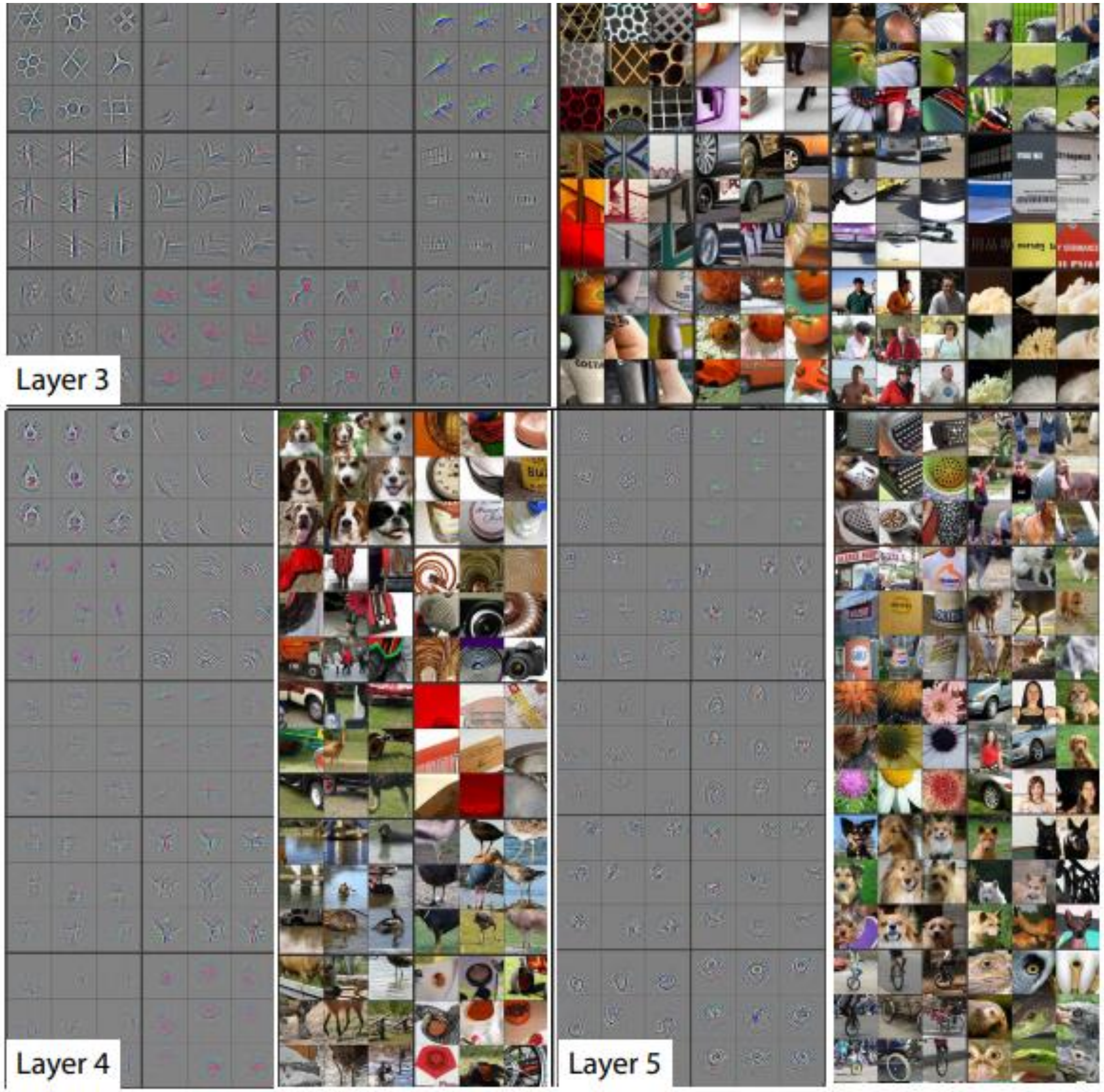




www.cs.utoronto.ca/~kriz/cifar.html

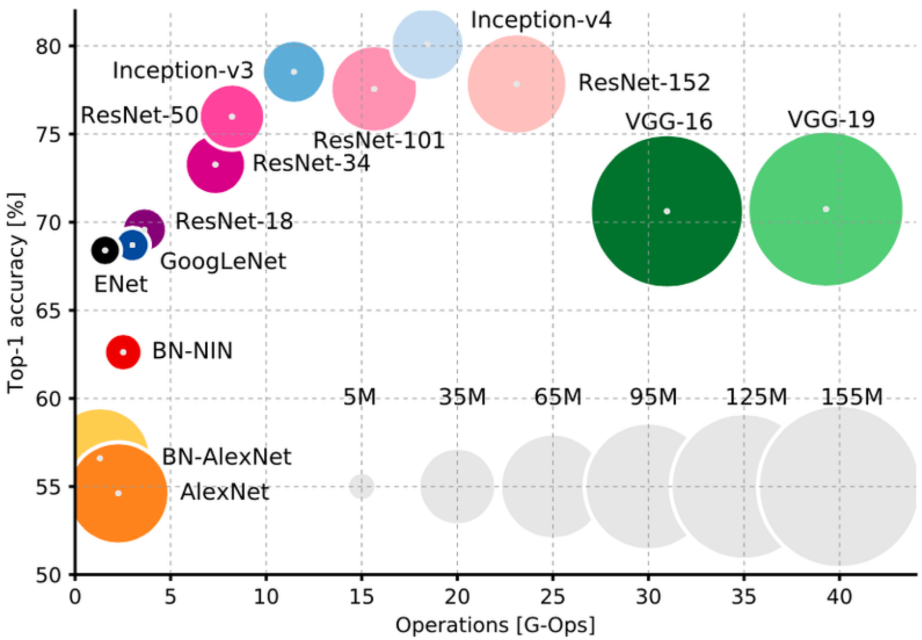
Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

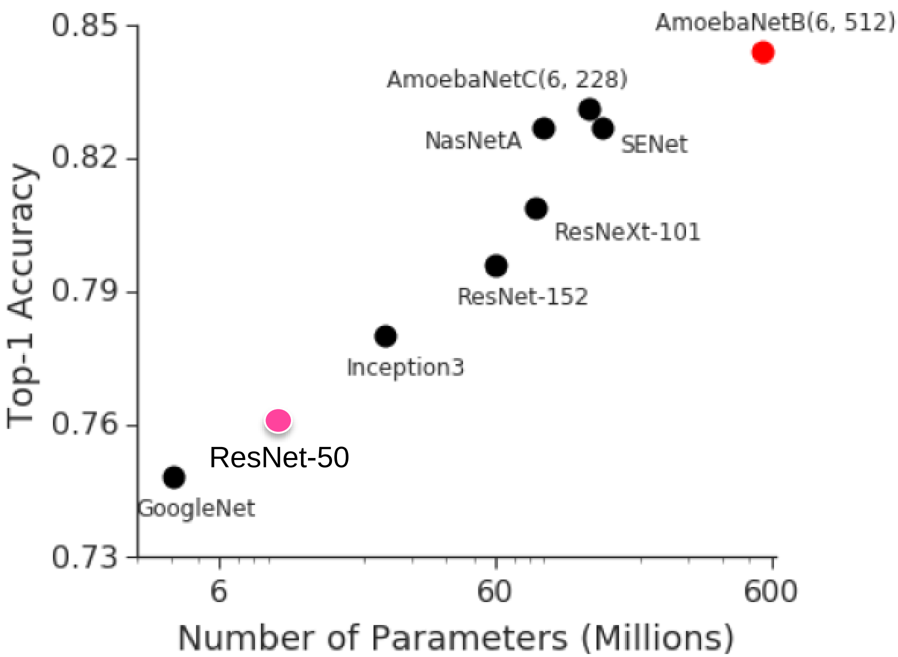
Learning characteristics (2/2)



Transfer learning with ResNet-50

Trade-off between accuracy and number of parameters



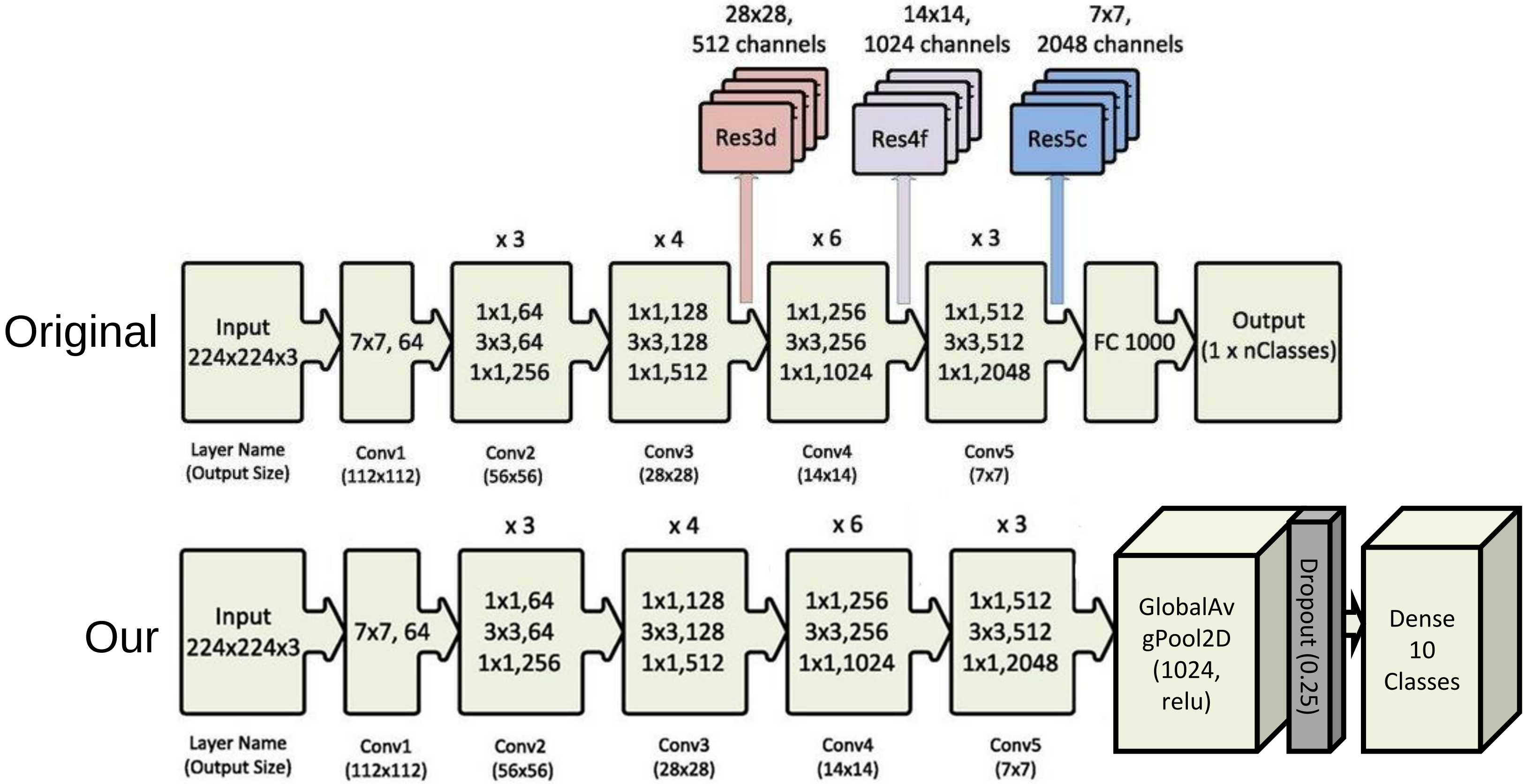


Canziani A, Paszke A, Culurciello E. An analysis of deep neural network models for practical applications. arXiv preprint arXiv:1605.07678. 2016 May 24.

Kornblith S, Shlens J, Le QV. Do better imagenet models transfer better?. arXiv preprint arXiv:1805.08974. 2018 May 23.

ResNet-50 convolutional network

We use transfer learning and added some final layers



Keras 2.2.4 implementation

Singh A, Kisku DR. Detection of Rare Genetic Diseases using Facial 2D Images with Transfer Learning. In2018 8th International Symposium on Embedded Computing and System Design (ISED) 2019 May 2 (pp. 26-30). IEEE.

Experimental conditions

	Training	Validation
Original article (Helber et al. 2017)	80%	20%

	Training	Validation	Test
Our	70%	15%	15%
	18900, 10 classes	4050	4050 images

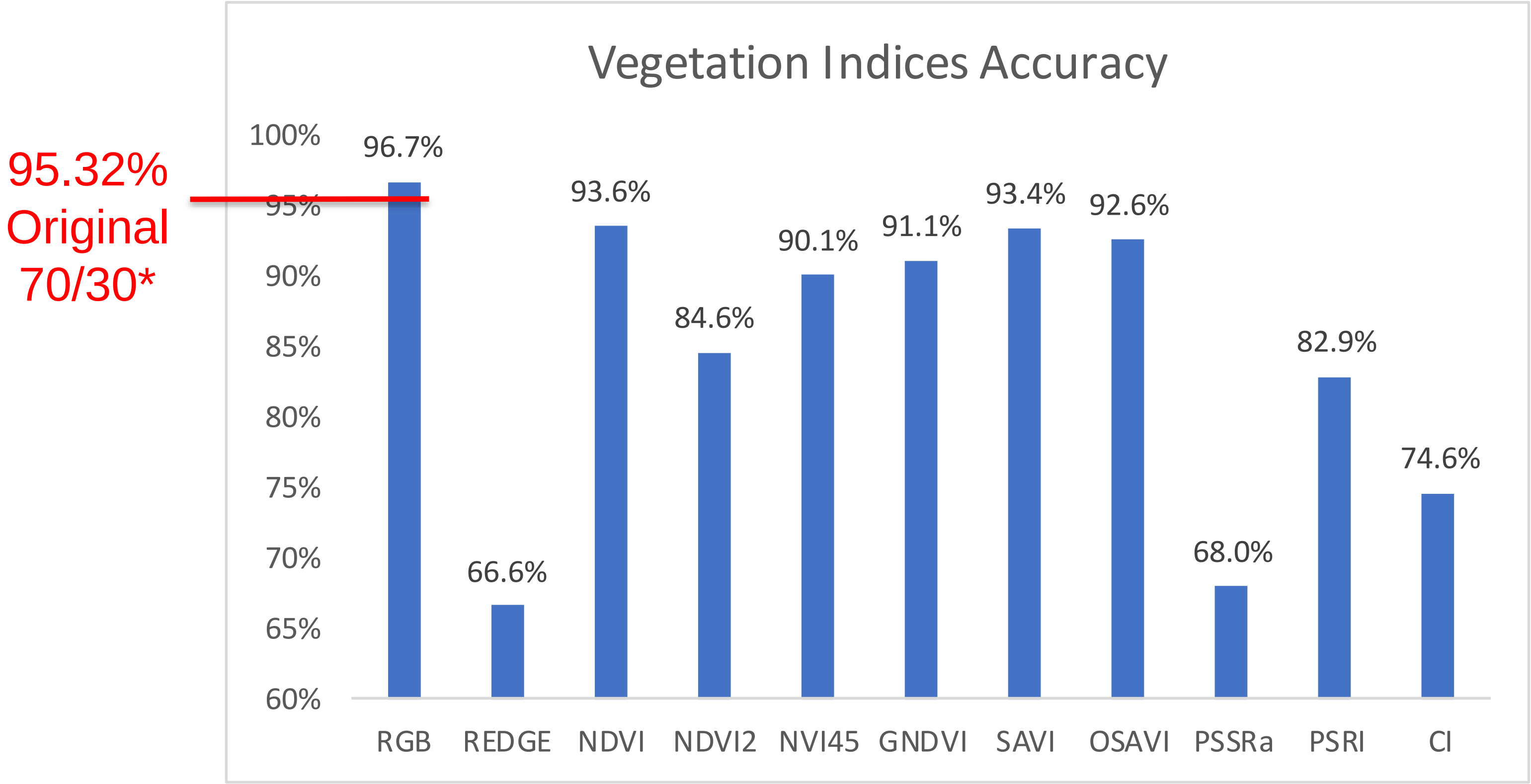
~25 min for each indices with GTX 1050 Ti.

Epochs 1-10: Learning rate: 0.01
Epochs 11-30: Learning rate: 0.0001
Total params: 25,696,138
Total Trainable params: **25,643,018**
Total Non-trainable params: 53,120
API: Keras v2.2.4 with Tensorflow
ResNet-50 on ImageNet with RMSProp
Some image transformations and dropout

Helber P, Bischke B, Dengel A, Borth D. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. arXiv preprint arXiv:1709.00029. 2017 Aug 31.

Results for convolutional networks

We achieve higher classification accuracy than original article for RGB bands



*98.57% for 80/20 training/test with ResNet-50.

Results for some Vegetation Indices

Some classes should have multi-labels

RGB

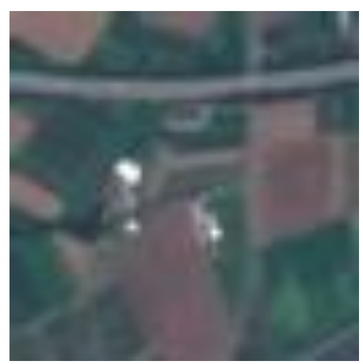
True label \ Predicted label	AnnualCrop	Forest	Herbaceous	Highway	Industrial	Pasture	PermanentCrop	Residential	River	SeaLake
AnnualCrop	429	0	0	0	0	4	2	0	0	4
Forest	2	440	1	0	0	3	0	0	1	0
Herbaceous	2	2	426	0	0	3	2	0	1	0
Highway	1	3	3	348	8	1	2	1	7	0
Industrial	2	0	0	5	371	0	1	1	0	0
Pasture	5	0	3	1	0	253	2	0	3	1
PermanentCrop	**16**	0	**32**	4	1	1	328	0	0	0
Residential	0	0	1	0	10	0	0	419	0	0
River	3	1	0	**17**	2	0	0	0	397	0
SeaLake	0	1	0	0	0	0	0	0	1	472

GNDVI

True label \ Predicted label	AnnualCrop	Forest	Herbaceous	Highway	Industrial	Pasture	PermanentCrop	Residential	River	SeaLake
AnnualCrop	415	0	1	2	0	13	6	0	1	1
Forest	2	429	0	0	0	16	0	0	0	0
Herbaceous	4	13	383	1	1	21	9	3	1	0
Highway	6	1	3	339	7	3	2	9	4	0
Industrial	2	0	0	0	362	0	1	15	0	0
Pasture	6	4	1	0	0	255	1	0	0	1
PermanentCrop	15	1	20	0	2	10	329	5	0	0
Residential	0	0	1	0	13	0	0	416	0	0
River	4	1	0	18	3	4	0	0	390	0
SeaLake	1	2	0	0	0	1	0	0	0	470

Highway

Highway

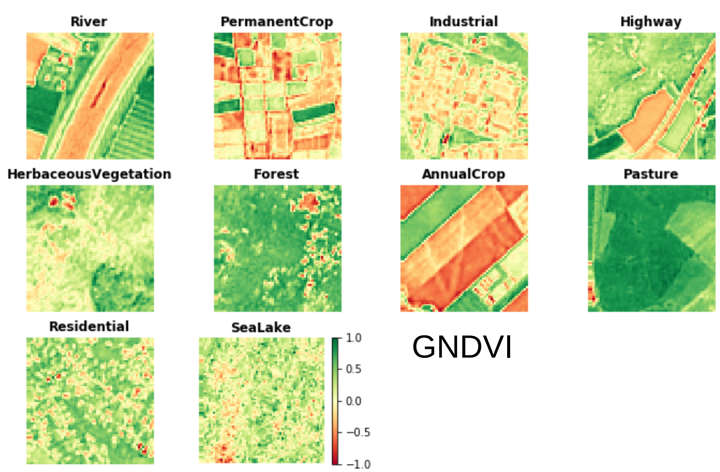


River

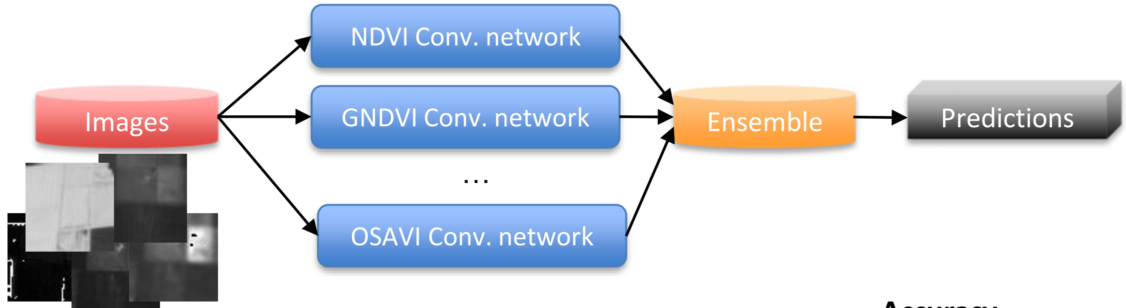
Results for every classes

Some vegetation indices have higher accuracy for some classes

	RGB	REDGE	NDVI	NDVI2	NDI45	GNDVI	SAVI	OSAVI	PSSRa	PSRI	CI
AnnualCrop	0.98	0.79	0.96	0.92	0.90	0.95	0.91	0.94	0.73	0.89	0.81
Forest	0.98	0.75	0.95	0.91	0.99	0.96	0.96	0.97	0.73	0.92	0.70
Herbaceous	0.98	0.62	0.95	0.86	0.80	0.88	0.82	0.96	0.59	0.81	0.70
Highway	0.93	0.42	0.92	0.80	0.84	0.91	0.96	0.95	0.50	0.62	0.70
Industrial	0.98	0.64	0.94	0.89	0.95	0.95	0.94	0.95	0.92	0.77	0.78
Pasture	0.94	0.32	0.80	0.83	0.81	0.95	0.77	0.85	0.31	0.70	0.54
PermanentCrop	0.86	0.42	0.90	0.72	0.85	0.86	0.83	0.77	0.57	0.74	0.42
Residential	0.97	0.71	0.98	0.88	0.97	0.97	0.97	0.98	0.58	0.88	0.87
River	0.95	0.79	0.93	0.80	0.87	0.93	0.94	0.91	0.80	0.95	0.88
SeaLake	1.00	0.99	0.98	0.99	0.99	0.99	0.99	0.93	0.94	0.90	0.97



Ensemble of convolutional networks



True label \ Predicted label	AnnualCrop	Forest	Herbaceous	Highway	Industrial	Pasture	PermanentCrop	Residential	River	SeaLake	Accuracy
AnnualCrop	**2866**	22	10	20	3	30	27	0	4	17	98.17%
Forest	0	**2899**	14	2	0	13	3	4	0	65	99.03%
Herbaceous	10	261	**2601**	12	9	34	40	25	3	5	98.03%
Highway	7	1	2	**2467**	7	0	1	1	14	0	98.44%
Industrial	2	0	1	20	**2450**	0	0	22	5	0	98.92%
Pasture	15	102	25	32	5	**1627**	138	3	5	48	91.40%
PermanentCrop	44	5	90	19	11	7	**2310**	14	0	0	90.48%
Residential	0	1	1	5	23	0	2	**2968**	0	0	99.50%
River	2	0	1	54	3	1	0	0	**2432**	7	96.96%
SeaLake	3	1	0	1	1	0	0	1	17	**2976**	96.17%

*RGB and Rededge convolutional network not included, over the whole dataset of 27,000 images.

Selection of 3 vegetation indices

Using 3 Vegetation Indices improved the overall classification accuracy by **10.7%** (OSAVI, NDVI, GNDVI) and **18.3%** (NDVI, NDVI2, NDI45) over RGB classification

	OSAVI NDVI GNDVI	NDVI NDVI2 NDI45	All	RGB
AnnualCrop	0.982	**0.986**	0.956	0.977
Forest	**0.990**	0.982	0.966	0.984
HerbaceousVegetation	**0.980**	0.978	0.867	0.977
Highway	**0.984**	0.973	0.987	0.930
Industrial	**0.989**	0.980	0.980	0.976
Pasture	0.914	0.928	0.814	**0.944**
PermanentCrop	0.905	**0.962**	0.924	0.859
Residential	**0.995**	**0.997**	0.989	0.974
River	0.970	**0.975**	0.973	0.945
SeaLake	0.962	0.986	0.992	**0.996**

Conclusions

- The EuroSat sentinel-2 dataset is a good starting point for transfer learning in agriculture/geomatic.
- Other interesting datasets are:
  - UC Merced Land Use Dataset (http://weegee.vision.ucmerced.edu/datasets/landuse.html)
  - Crop/Weed Field Image Dataset (https://github.com/cwfid/dataset)

- Using an ensemble of convolutional neural networks improve the overall accuracy of classification for single class data.