Deep Convolutional Neural Networks for Map Type Classification

Xiran Zhou, Wenwen Li, Arizona State University

Samantha T Arundel, Center of Excellence in Geographic Information Science, U.S. Geological Survey

Jun Liu, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

5/23/2018

Madison, WI

Background

Maps

Maps are an important medium that enable people to comprehensively understand the configuration of cultural activities and natural elements over different times and places.

Internet and spin-off techniques (e.g. web-based service technologies, cyberinfrastructure, and volunteered geographic information) have significantly changed the nature of map generation and the use of maps (Hurst and Clough, 2013).

Currently, maps are used by:

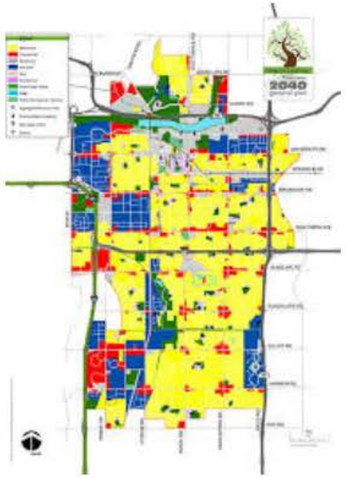
- geographical domain experts to conduct geospatial computing and analysis
- the public to better facilitate daily activities such as ridesharing, delivery, and transportation network analysis, to name just a few

Challenge 1

- A majority of maps available from the Internet lack map elements like map title, direction indicators, legends, metadata, and so on.

Challenge 2

- The creation of immense map repositories containing maps with diverse themes, configurations and designs.



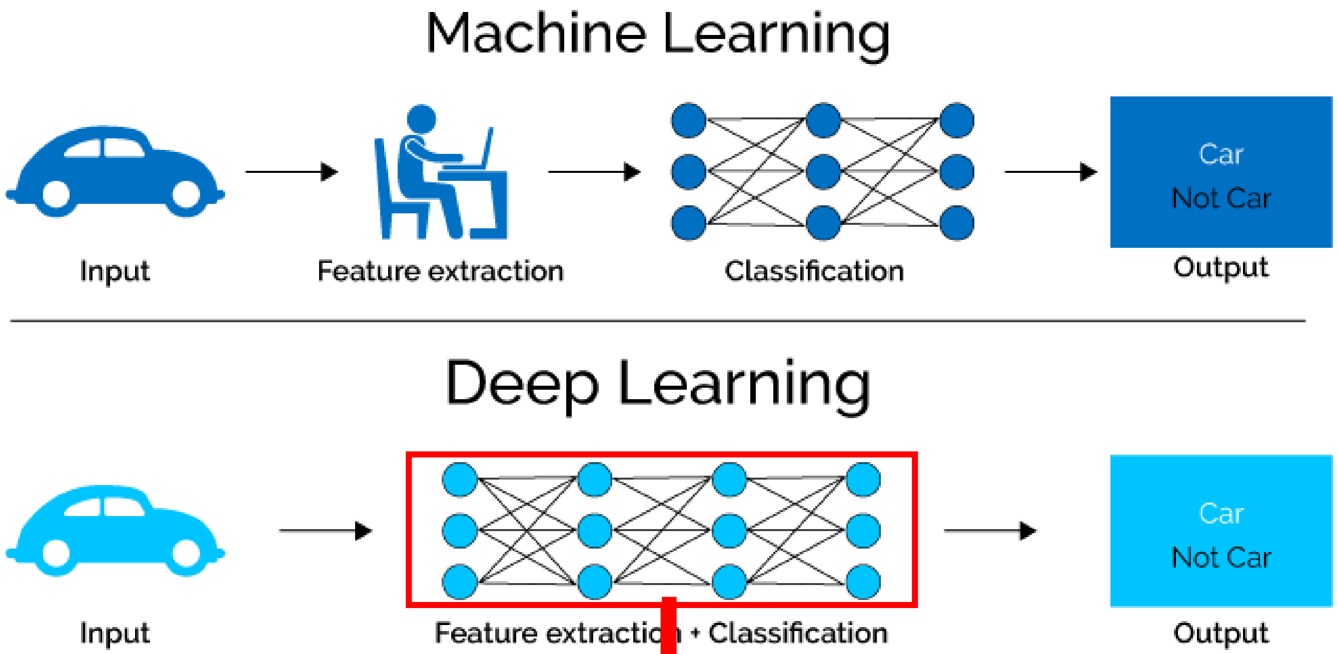
Objective

An automated approach ensures the availability of needed maps and is essential to facilitate the role of maps in geographical analysis and public activities.

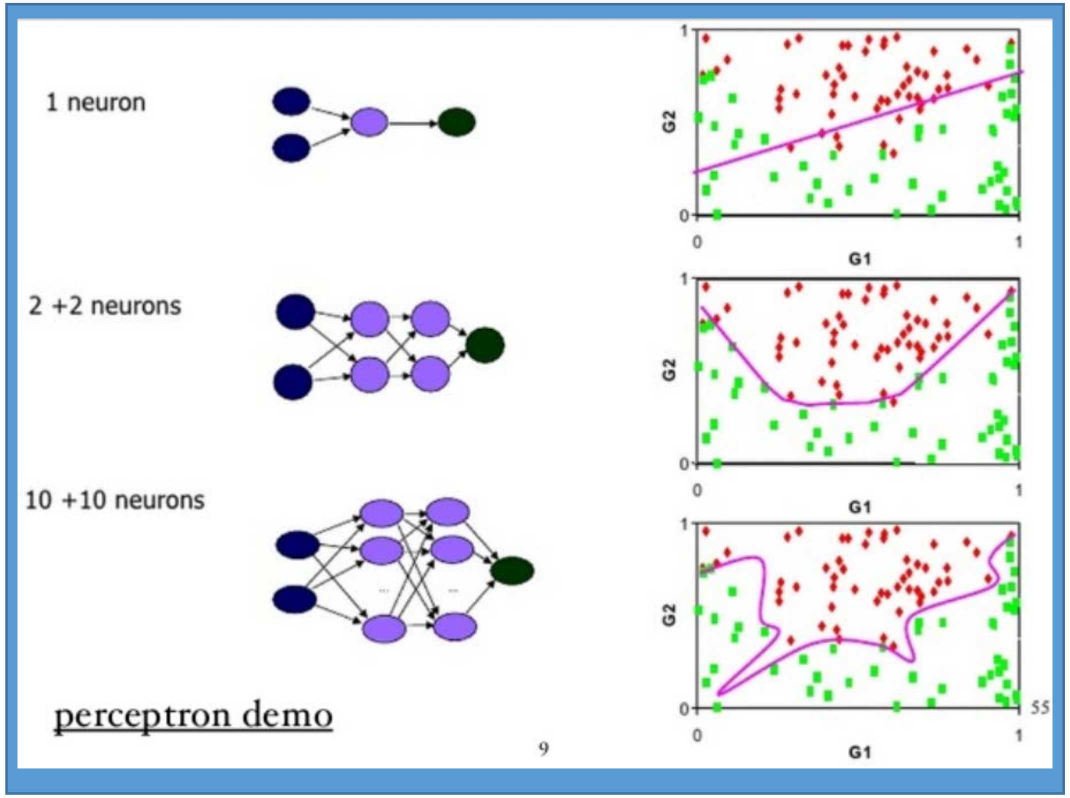
Deep learning/representation learning

Deep learning approaches enable computers to automatically discover the high-level features/representations of data based on a multi-layers processing framework (Bengio, Courville and Vincent, 2012; LeCun, Bengio and Hinton, 2015).

Machine learning vs deep learning



https://codeutsava.in/blog/40



(Pieters, 2015)

Benchmark dataset—*deepMap*

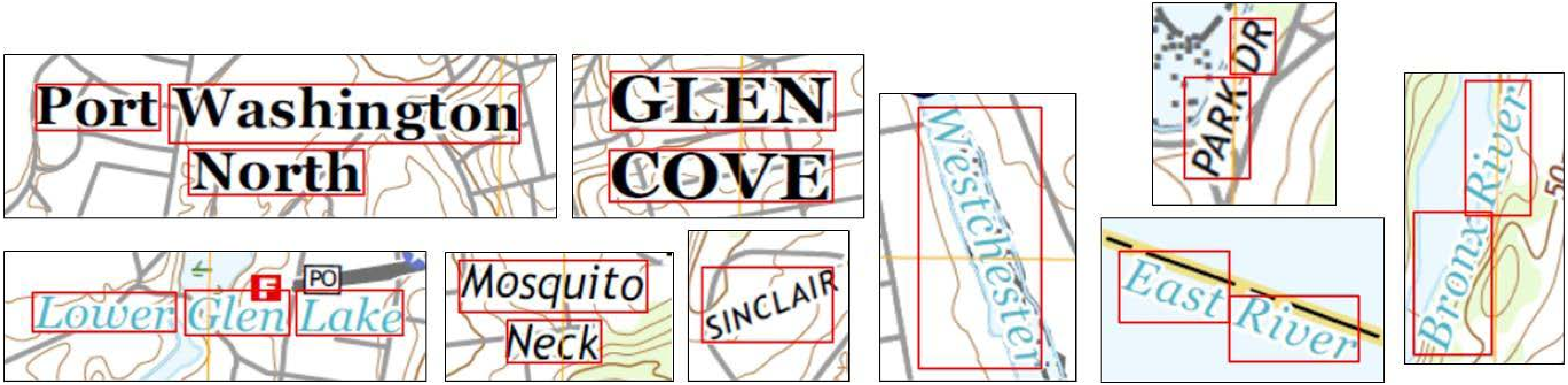
It is critical to prepare large-scale well-labeled data to feed a neural network for enhancing its capability of distinguishing different classes (Bengio, Courville and Vincent, 2012).

Map text dataset

- Data source: the National Map (the U.S. Geological Survey)
- 44 letter classes (C=c, O=o, P=p, S=s, U=u, V=v, W=w, Z=z)
- ≈ 5,000 text samples

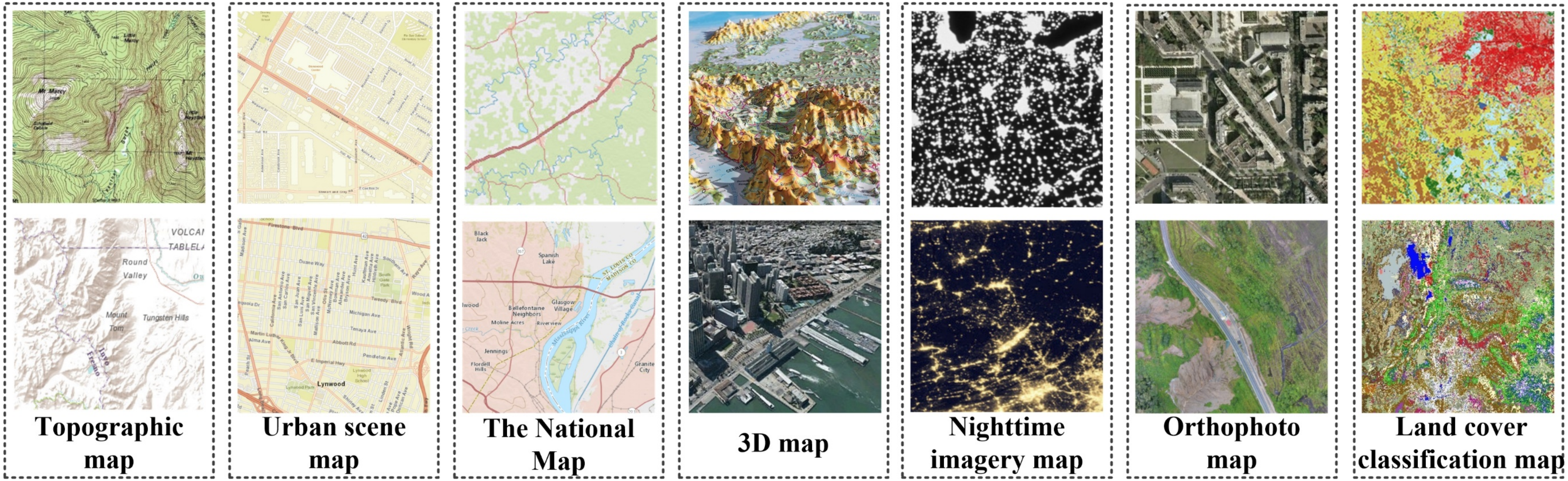
Text-labeled map dataset

- Data sources: the National Map, Google Maps, Bing Maps, OpenStreetMap
- ≈ 4,000 text-labeled samples



Map dataset

- 7 types map
- 250-350 samples for each map type
- ≈ 2,200 samples in total



Map type classification

- *AlexNet* (Krizhevsky, Sutskever and Hinton, 2012)
- *VGG Net* (Simonyan and Zisserman, 2014)
- *GoogleNet or Inception* (Szegedy, et al., 2015)
- *ResNet* (He, et al., 2016)
- *Inception-ResNet* (Szegedy, et al., 2017)

CNNs & DCNNs	*Experimental results*	
	Group 1: 60% data used for training & 40% data used for testing	*Group 2: 80%* data used for training & 20% data used for testing
AlexNet	*71% ~ 78%*	*77% ~ 83%*
VGG Net-19	*73% ~ 80%*	*78% ~ 84%*
Inception V4	*82% ~ 87%*	*88% ~ 94%*
ResNet V2-152	*82% ~ 86%*	*89% ~ 93%*
ResNet-Inception V2	*88% ~ 92%*	*95% ~ 99%*

Map text recognition (optical character recognition for cartography)

Optical character recognition (OCR): “is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text.” (*Wikipedia*)

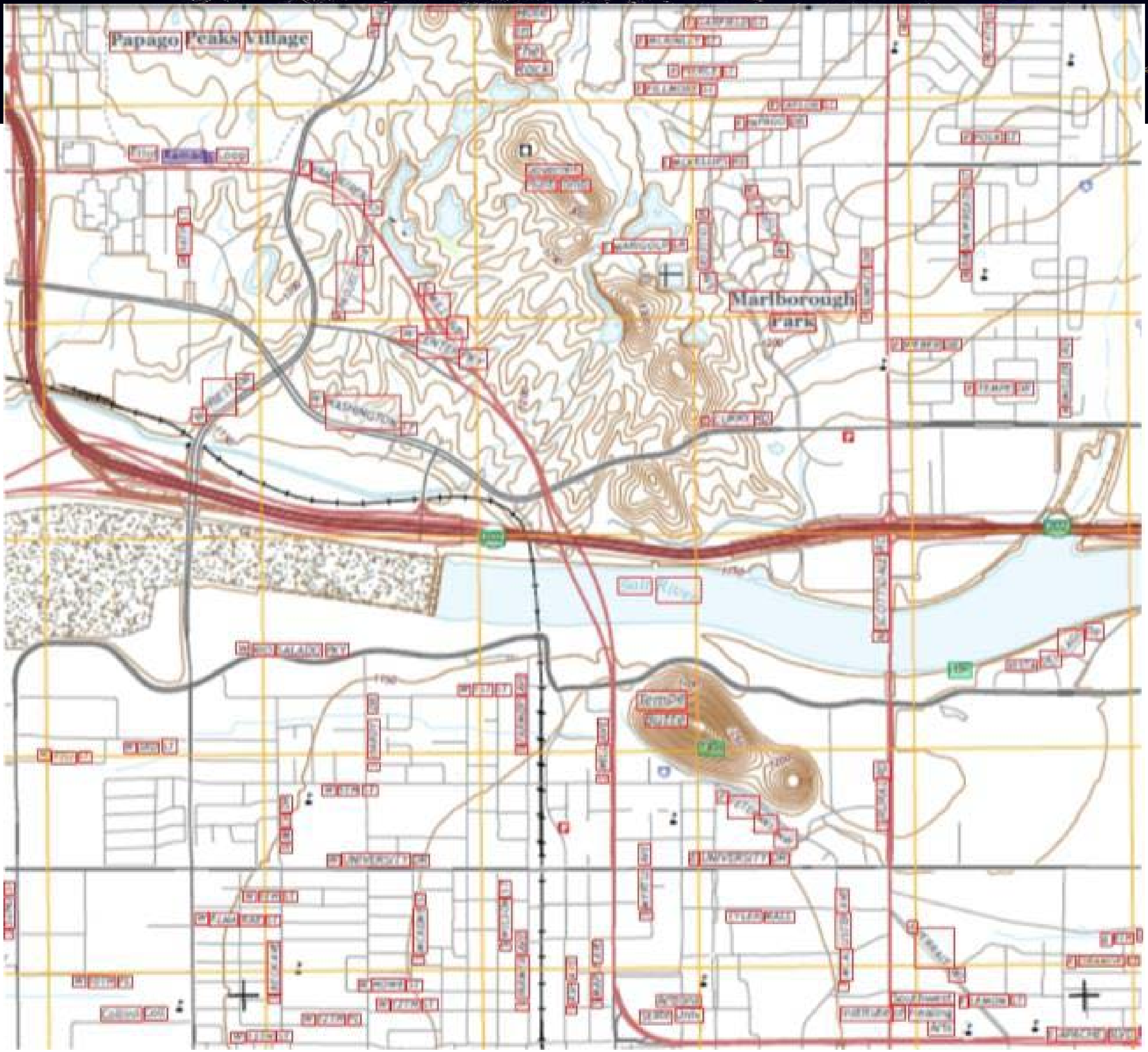
Optical Character Recognition (OCR) version 1.4 by Diego Barragán

Map text recognition

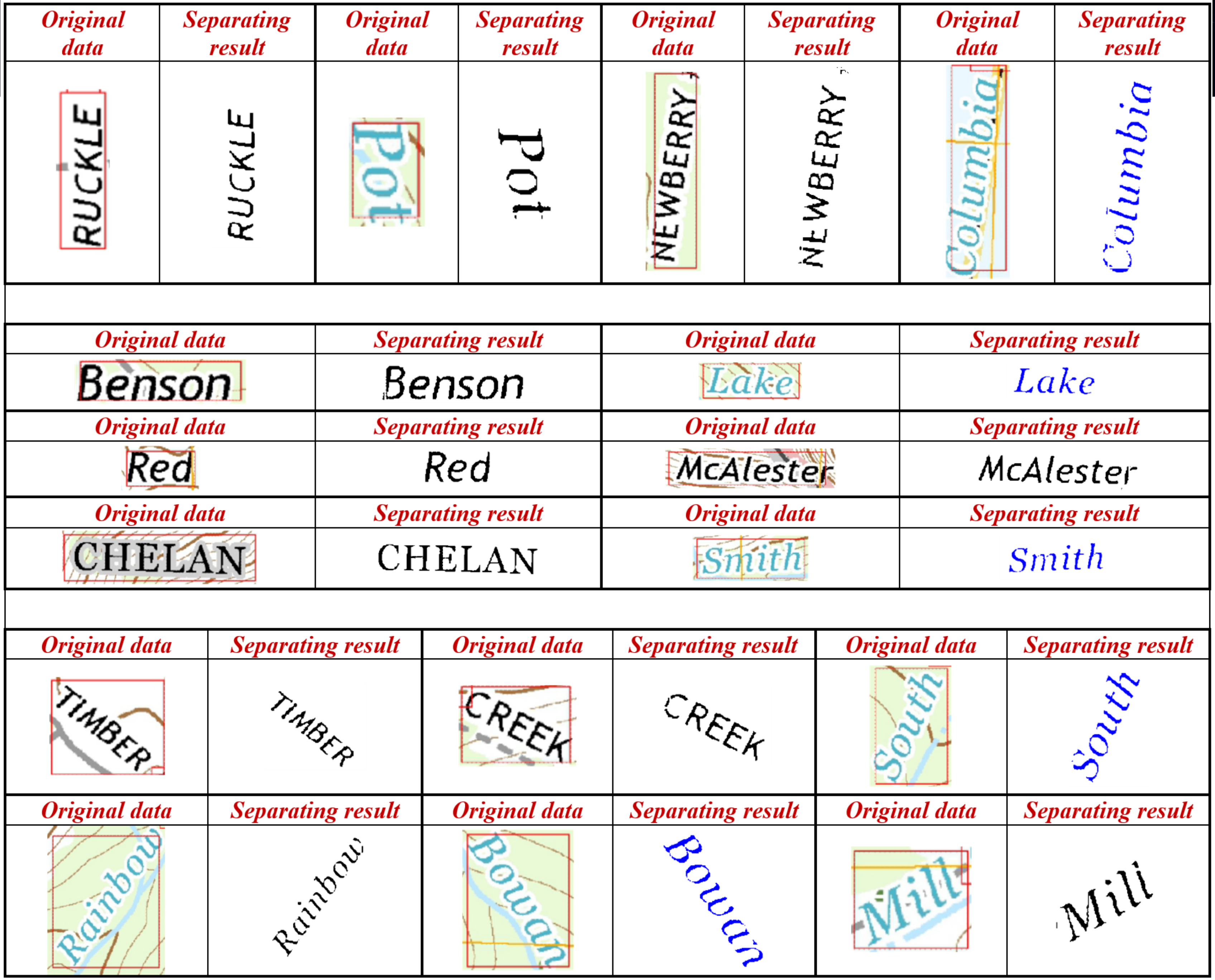
- Map text detection
- Map text separation
- Map text straightening
- Map text recognition

Intelligent map reader

Map text detection



Map text separation



Map text straightening

Original data	Straightening result	Original data	Straightening result	Original data	Straightening result
RUCKLE	RUCKLE	Pot	Pot	Mili	Mili
Original data	**Straightening result**	**Original data**	**Straightening result**	**Original data**	**Straightening result**
River	River	SKYLINE	SKYLINE	Bowan	Bowan

Limit: curved map text

Map text recognition: Google Tesseract OCR

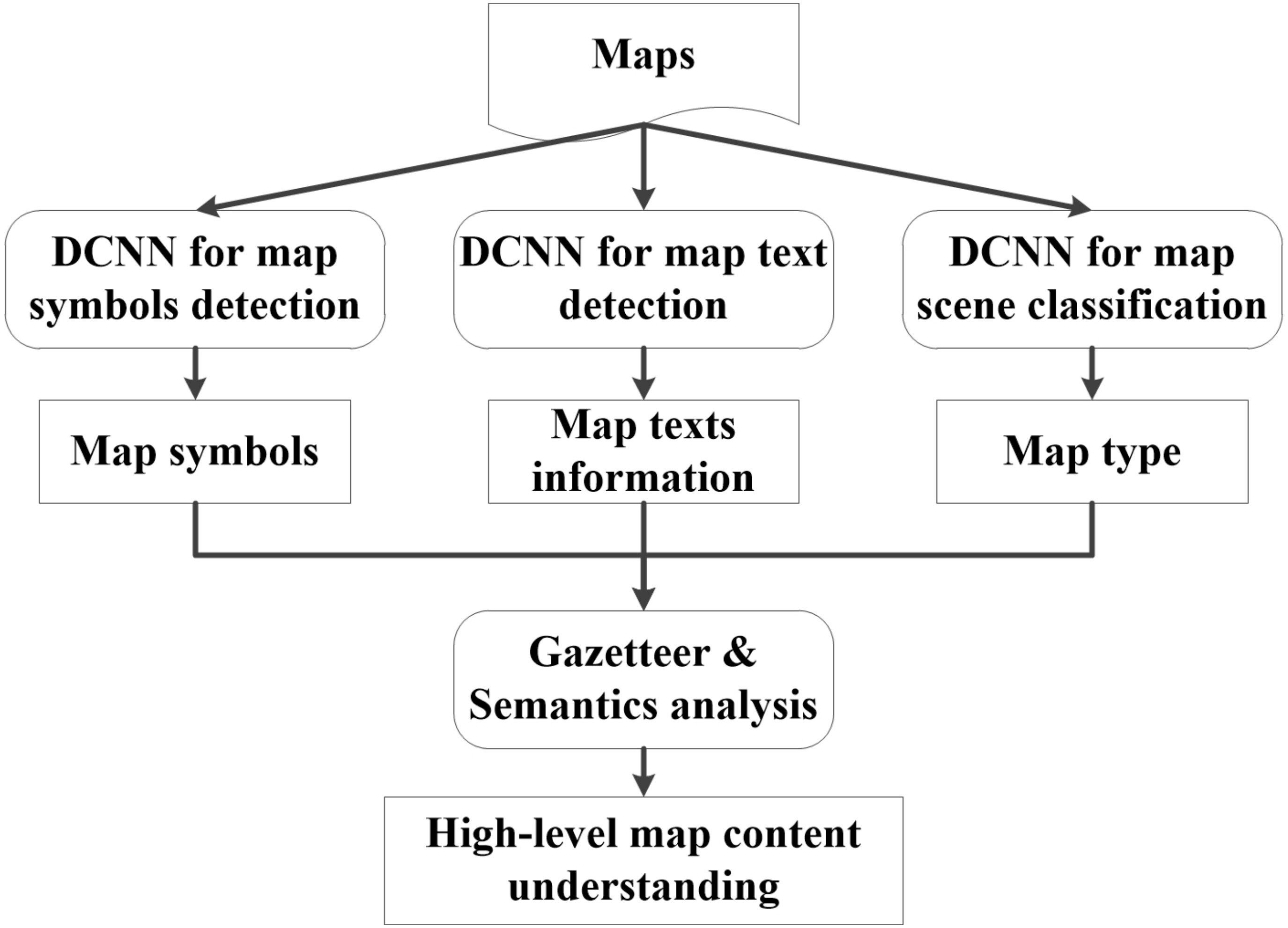
- “Tesseract is an OCR engine with support for unicode and the ability to recognize more than 100 languages out of the box.”
- https://github.com/tesseract-ocr/tesseract

Original data	*Machine-readable text*	*Original data*	*Machine-readable text*	*Original data*	*Machine-readable text*
RUCKLE	RUCKLE	Pot	Pot	Mill	Mill
Original data	*Machine-readable text*	*Original data*	*Machine-readable text*	*Original data*	*Machine-readable text*
River	River	SKYLINE	SKYLINE	Bowan	Bowan

Topographic map	AP of map text recognition (%)
USGS US Topo 7.5-minute map for Onego, WV 2016	95.57
USGS US Topo 7.5-minute map for Eagle Nest, ID 2017	95.13
USGS US Topo 7.5-minute map for Terrell Creek, MT 2017	95.35
USGS US Topo 7.5-minute map for Rocky Reach Dam, WA 2017	94.84
USGS US Topo 7.5-minute map for Skutumpah Creek, UT 2017	94.70
USGS US Topo 7.5-minute map for Billerica, MA 2015	92.97
USGS US Topo 7.5-minute map for Boston North, MA 2015	91.75
USGS US Topo 7.5-minute map for Brooklin, NY 2016	92.52
USGS US Topo 7.5-minute map for Jamaica, NY 2016	92.55
USGS US Topo 7.5-minute map for Flagstaff West, AZ 2014	94.27
USGS US Topo 7.5-minute map for Sunland, CA 2015	93.40
USGS US Topo 7.5-minute map for South Lake Tahoe, CA-NV 2015	94.73
USGS US Topo 7.5-minute map for Sled Springs, OR 2017	93.67
USGS US Topo 7.5-minute map for Evergreen, CO 2016	94.89
USGS US Topo 7.5-minute map for Marsh-Miller Lake, WI 2015	96.51
USGS US Topo 7.5-minute map for Hamilton, NY 2016	92.31
USGS US Topo 7.5-minute map for San Jose East, CA 2015	94.65
USGS US Topo 7.5-minute map for El Monte, CA 2015	94.86
USGS US Topo 7.5-minute map for Wappapello, MO 2017	94.71
USGS US Topo 7.5-minute map for Hilo, HI 2017	94.93
All 20 maps	94.78

On-going work

Intelligent map reader V2



References

Bengio, Y., Courville, A., & Vincent, P. (2012). Representation Learning: A Review and New Perspectives. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, *35*(8), 1798-1828.

Hurst, P. and Clough, P. (2013). Will we be lost without paper maps in the digital age?. *Journal of Information Science*, *39*(1), 48-60.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436.

Li, W., Yang, C. and Yang, C. (2010). An active crawler for discovering geospatial web services and their distribution pattern–A case study of OGC Web Map Service. *International Journal of Geographical Information Science*, *24*(8), 1127-1147.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Szegedy, C., et al (2015, June). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-5).

Szegedy, C., et al (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI* (Vol. 4, p. 12).

H. Li, J. Liu, X. Zhou*. Intelligent map reader: topographic map understanding with deep learning techniques and gazetteer. *IEEE Access*. (Accepted)

Thanks!

Comments and question?