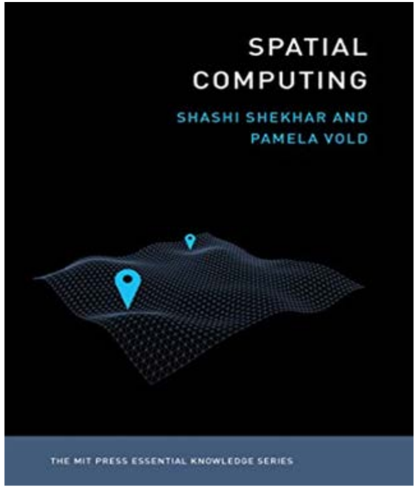
Panel on Community Voices

GEO.AI – Redefining Geospatial Conference, World Geospatial Council, Dec. 7th- 9th, 2020



What's Special About GeoAI?

Shashi Shekhar

McKnight Distinguished University Professor, Univ. of Minnesota

www.cs.umn.edu/~shekhar

UNIVERSITY OF MINNESOTA
Driven to Discover®

Acks.: NSF, USDOD-NGA, USDOE-ARPA-E, USDA

Acknowledgements

- P.I., Connecting the Smart-City Paradigm with a Sustainable Urban Infrastructure Systems Frame- work to Advance Equity in Communities, National Science Foundation (Award 1737633), $2.5 M, 9/1/2017 - 8/31/2021.

- P.I., Spatio-temporal Informatics for Transportation Science, National Science Foundation (Award 1901099), $1.2M, 8/1/2019-7/31/2023.

- P.I., EAGER: Spatiotemporal Big Data Analysis to Understand COVID-19 Effects, 100K, National Science Foundation (Grant 2040459), 9/1/2020-8/31/2021.

- P.I., Identifying Aberration Patterns in Multi-attribute Trajectory Data with Gaps, 200K, USDOD-NGA (HM0476-20-1-0009), 6/15/2020- 6/14/2025.

- Co-P.I., Increasing Low-Input Turfgrass Adoption Through Breeding, Innovation, and Public Education, $5.4 M, USDA/NIFA/SCRI (contract 2017-51181-27222), 9/2017 - 8/2021. (with E. Watkins).

- Co-P.I., Planning Grant: Engineering Research Center for Intelligent Infrastructure for Safe, Efficient and Resilient Mobility, National Science Foundation,(Award 1840432), 97K, 9/18-8/21, (PI: Anil Misra, University of Kansas).

Spatial Revolution

- GPS & Location traces
  - 2 billion GPS receivers today (7 billion by 2022)
  - Reference clock for telecom, banks, …
  - Help understand Spatio-temporal patterns of life
- (Nano-)Satellite Imagery, …

The World Economy Runs on GPS. It Needs a Backup Plan

Bloomberg Businessweek

July 25, 2018, 4:00 AM CDT

McKinsey Global Institute

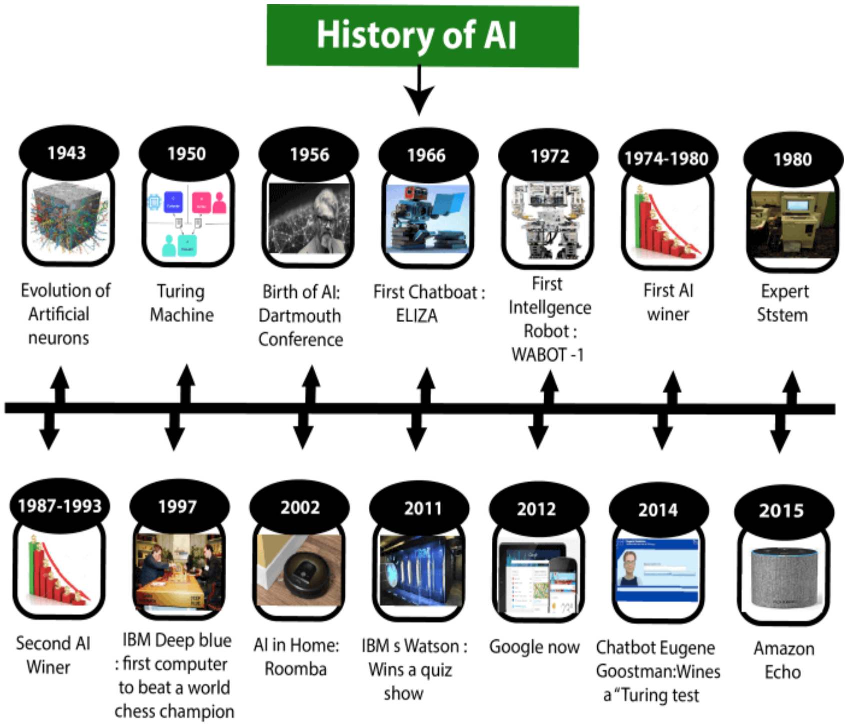
The study estimates that the use of personal location data could save consumers worldwide more than $600 billion annually by 2020. Computers determine users' whereabouts by tracking their mobile devices, like cellphones.

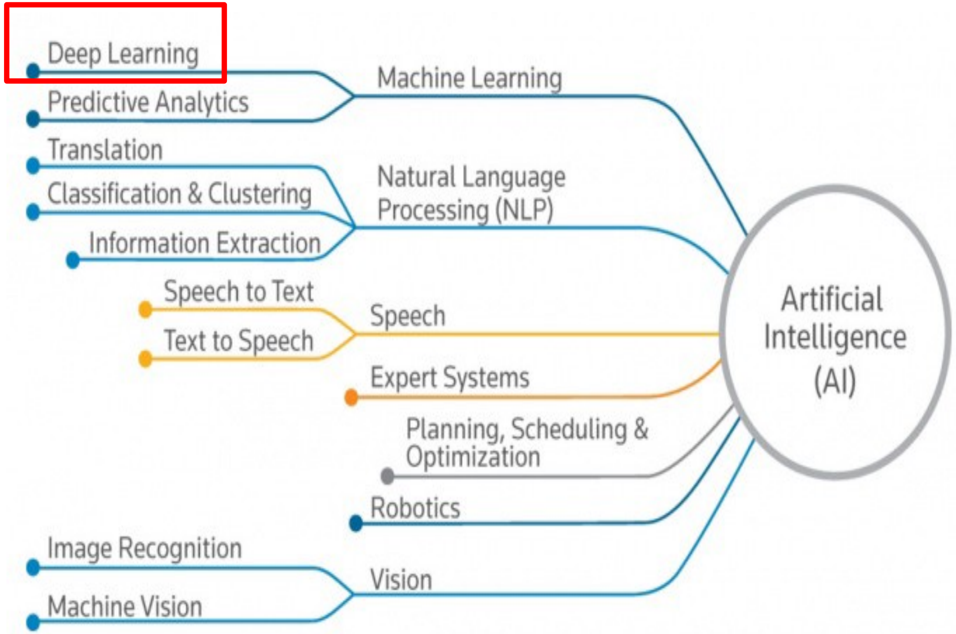
The New York Times

Published: May 13, 2011

Source: Y. Xie et al., Transforming Smart Cities With Spatial Computing, Proc. IEEE Intl. Conf. on Smart Cities, 2018.

AI: Topics, History, Recent Breakthroughs

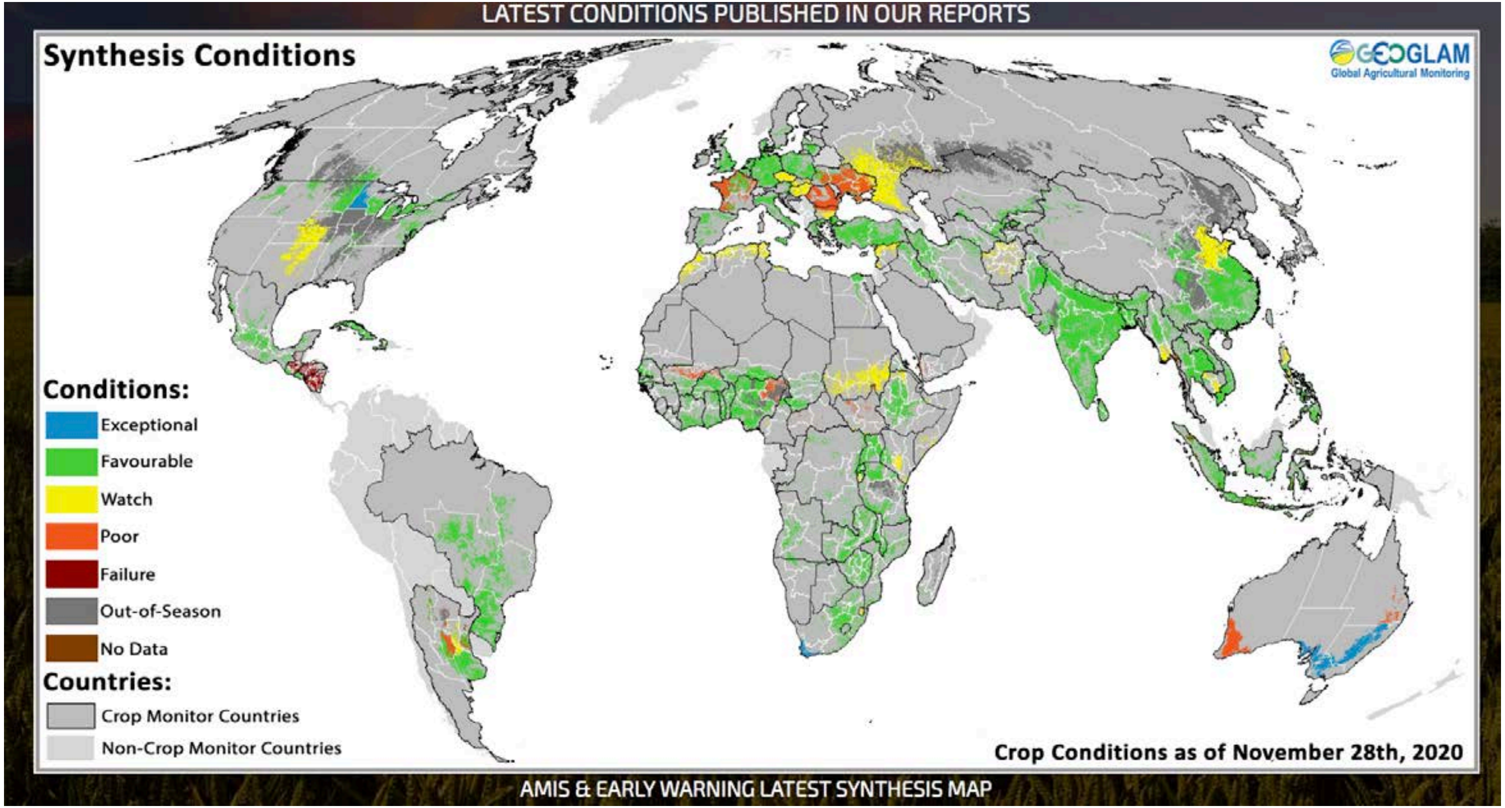




AI promise for Spatial Problems

- Cheaper, faster, and bigger Maps
  - Ex. US Natl. Wetland Inventory - $400 M over 40 years (last century)
- Inverse Geo-Problems
- Geo-Content based Querying
- But many hurdles

Cheaper Faster Bigger Maps (e.g., Crop Health)



AMIS Agricultural Market Information System

GEOGLAM Global Agricultural Monitoring

Traditional AI vs. Deep Learning

- From Satellite Imagery: Classify Land-cover, Map buildings
- Ex. 2009 Haiti Earthquake: Map building damage [1]

eCognition

Traditional AI

1. Pre-processing
2. Feature engineering & Selection
3. Prediction (class or object)
4. Post-processing (rules, manual)

Deep Learning

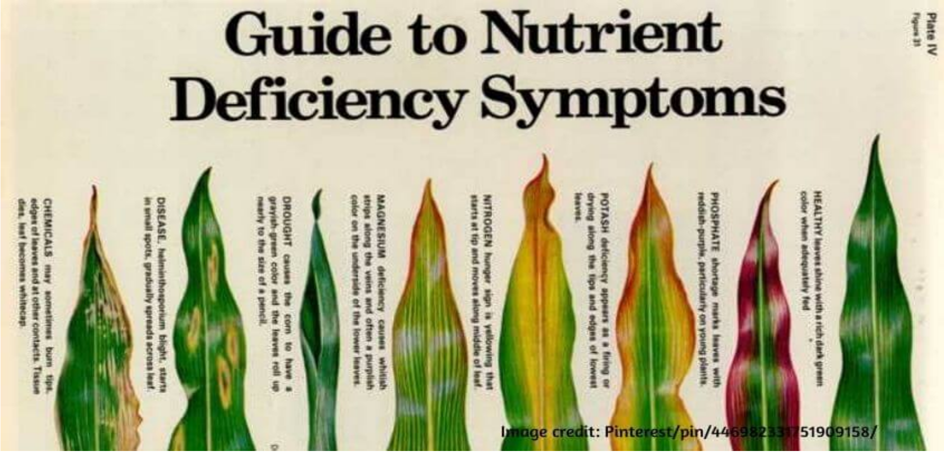
1. Pre-processing

Deep (neural network) Learning

4. Post-processing (guardrail for AI)

[1] J. Aardt et al., Geospatial Disaster Response during the Haiti Earthquake: A Case Study Spanning Airborne Deployment, Data Collection, Transfer, Processing, and Dissemination, Photo. Eng. & Remote Sens., 77(9):943-952, Sept. 2011.

https://agpartners.net/2018/08/17/corn-silage-in-2018/



Open Debates

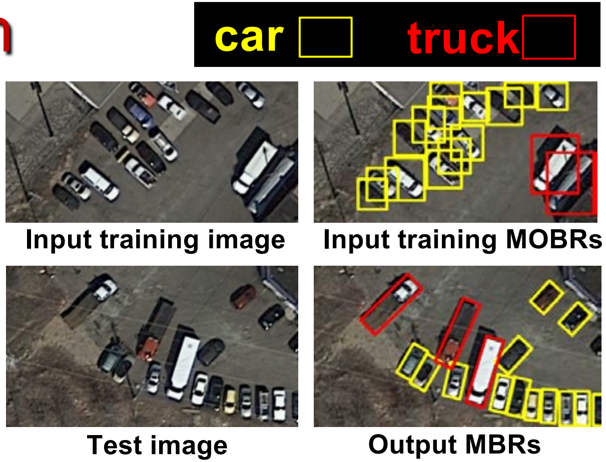
- Why Machine Constructed Features?
  - NDVI measures overall plant health
  - No diagnosis, e.g., disease, nutrients, …

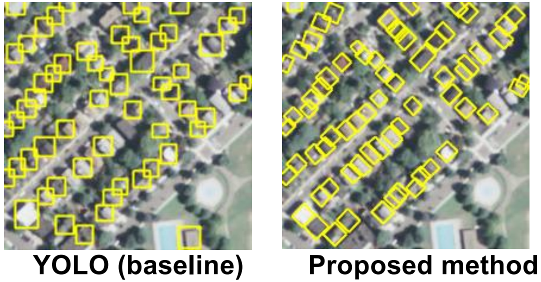
- What's value of hand-constructed features for deep learning?
  - Improve explanation & Prediction
  - Reduce computation cost, number of learning samples needed

Details: *NDVI Versus CNN Features in Deep Learning for Land Cover Classification of Aerial Images*, *IEEE Intl. Geoscience and Remote Sensing (IGARSS) Symposium* 2019, pp. 6483-6486. doi: 10.1109/IGARSS.2019.8900165 (A. Ramanath, S. Muthusrinivasan, Y. Xie, S. Shekhar and B. Ramachandra,)

Deep Learning for Geo-Object Detection

- **Q:?** How many vehicles in a parking lot? City?
- **Ex.:** Estimate truck supply in a city (CH Robinson).
- **Old Computer Vision workflow**
  - Many steps, each adds error
- **New Deep Learning Workflow – fewer steps**
  - Aerial imagery (3 inch pixels, Twincites, MN, USA)
  - NAIP Imagery (1 meter pixels, 2017)
    - MA Buildings Data (https://www.cs.toronto.edu/~vmnih/data/)
- Detected Geo-objects
  - Cars, trucks, Houses, …
  - Method: Convolutional Neural Networks (YOLO)

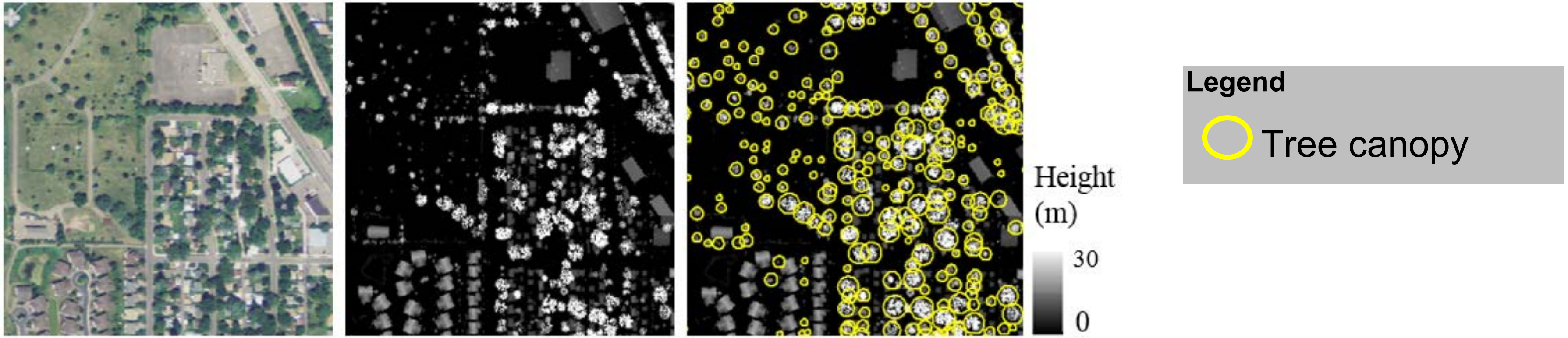




Details: Y. Xie et al., An Unsupervised Augmentation Framework for Deep Learning based Geospatial Object Detection: A Summary of Results, Proc. 26th ACM SIGSPATIAL Intl. Conf. Adv. in GIS, 2019.

Mapping Trees from Remote Sensing Imagery

- Why: Protect Powerlines, EAB Mgmt, Green infrastructure Equity
- Input: LiDAR + Remotely Sensed Imagery + (NAIP Ground Truth)
- Approach: Tree Inference by Minimizing Bound-and-band Errors (TIMBER)
  - Optimization to find tree locations and sizes
  - Deep learning constructs features separating trees & non-trees (e.g., light pole)



Details: **(a)** Revolutionizing Tree Management via Intelligent Spatial Techniques,
Proc. 27th ACM SIGSPATIAL Intl. Conf. on Adv. in GIS, Nov. 2019 Pages 71–74 (Best Vision Paper).
(b) TIMBER: A Framework for Mining Inventories of Individual Trees in Urban Environments using Remote Sensing Datasets,
IEEE Intl. Conf. on Data Mining (ICDM), 2018..

AI promise for Spatial Problems

- Cheaper, faster, and bigger Maps
- Inverse Geo-Problems, e.g., Find geo-locate of a picture
- Geo-Content based Querying
- But many hurdles

Geo-Fingerprinting: Plan B for GPS Jamming/Spoofing

Map of Mississippi River

http://worldsmap.world/mississippi-river-on-world-map/

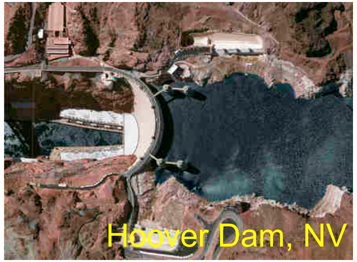
Q? Which location does this image show?

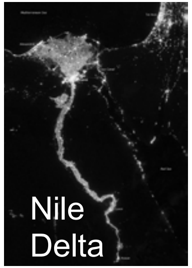
http://maps.google.com

Spatial Computing Research Group

Inverse Geo-Problem

- Ex. Find location given a geo-image or geo-video





Sources: Google Earth

AI promise for Spatial Problems

- Cheaper, faster, and bigger Maps
- Inverse Geo-Problems
- Geo-Content based Querying, e.g., factcheck geo-imagery
- But many hurdles

Inverse Geo-Problem: Factcheck a Geo-propaganda

- Ex. Google Earth shows that supposed oil storage facility in Al-Thawrah, Syria
  - was a grain storage facility located 150 Km away [1].

Russian Ministry of Defense video

Google Maps

Source: Distract, Deceive, Destroy, M. Czuperski et al., Atlantic Council, 2016. ISBN 978-1-61977-510-7.

AI promise for Spatial Problems

- Cheaper, faster, and bigger Maps
- Inverse Geo-Problems
- Geo-Content based Querying
- But many hurdles (Machine is still learning!)

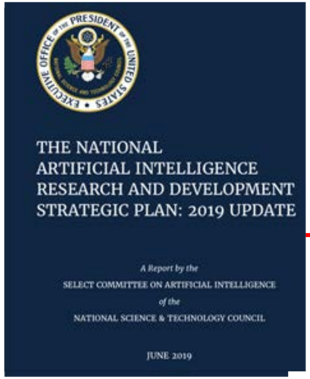
Acknowledgement: Thank you Mr. Vatyam Krishan for the phrase: "Machine is still learning"

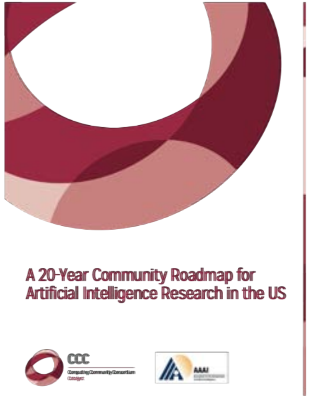
Many Hurdles: Machine is still learning

- Open Problems
- Content based Querying
  - 2017-20: DARPA Geospatial Cloud Analytics
    - food shortages, fracking, illegal fishing vessels
  - 2020-onwards: IARPA SMART (Space-based Machine Automated Recognition Technique)
    - Spatiotemporal: Construction & classify Stage
    - Map Underground: (Subterranean Challenge)

.

Sources: DARPA(https://www.darpa.mil/attachments/DARPA-2019-framework.pdf)





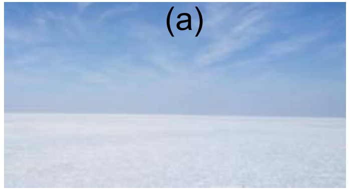
Advancing data-focused methodologies for knowledge discovery

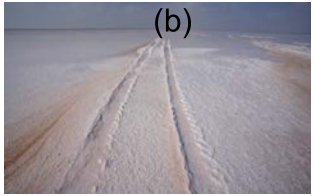
As discussed in the 2016 *Federal Big Data Research and Development Strategic Plan*,[34] many fundamental new tools and technologies are needed to achieve intelligent data understanding and knowledge discovery. Further progress is needed in the development of more advanced machine learning algorithms that can identify all the useful information hidden in big data. Many open research questions revolve around the creation and use of data, including its veracity and appropriateness for AI system training. The veracity of data is particularly challenging when dealing with vast amounts of data, making it difficult for humans to assess and extract knowledge from it. While much research has dealt with veracity through data quality assurance methods to perform data cleaning and knowledge discovery, further study is needed to improve the efficiency of data cleaning techniques, to create methods for discovering inconsistencies and anomalies in the data, and to develop approaches for incorporating human feedback. Researchers need to explore new methods to enable data and associated metadata to be mined simultaneously.

Many AI applications are interdisciplinary in nature and make use of heterogeneous data. Further investigation of multimodality machine learning is needed to enable knowledge discovery from a wide variety of different types of data (e.g., discrete, continuous, text, spatial, temporal, spatio-temporal, graphs). AI investigators must determine the amount of data needed for training and to properly address large-scale versus long-tail data needs. They must also determine how to identify and process rare events beyond purely statistical approaches; to work with knowledge sources (i.e., any type of information that explains the world, such as knowledge of the law of gravity or of social norms) as well as data sources, integrating models and ontologies in the learning process; and to obtain effective learning performance with little data when big data sources may not be available.

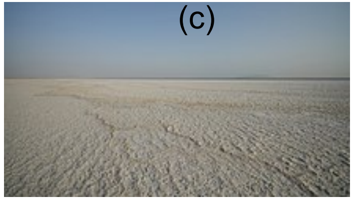
Quiz Time

Q. Which images show snow ?

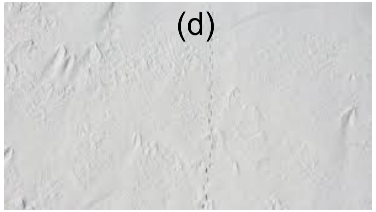




Runn of Kutch, Gujarat, India



Lake Karum, Ethiopia



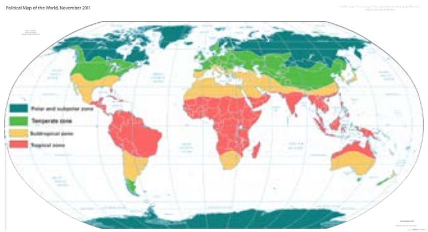
Snow

Snow

Q2. Which geo-challenges are addressed by Convolutional Neural Network (CNN) ?

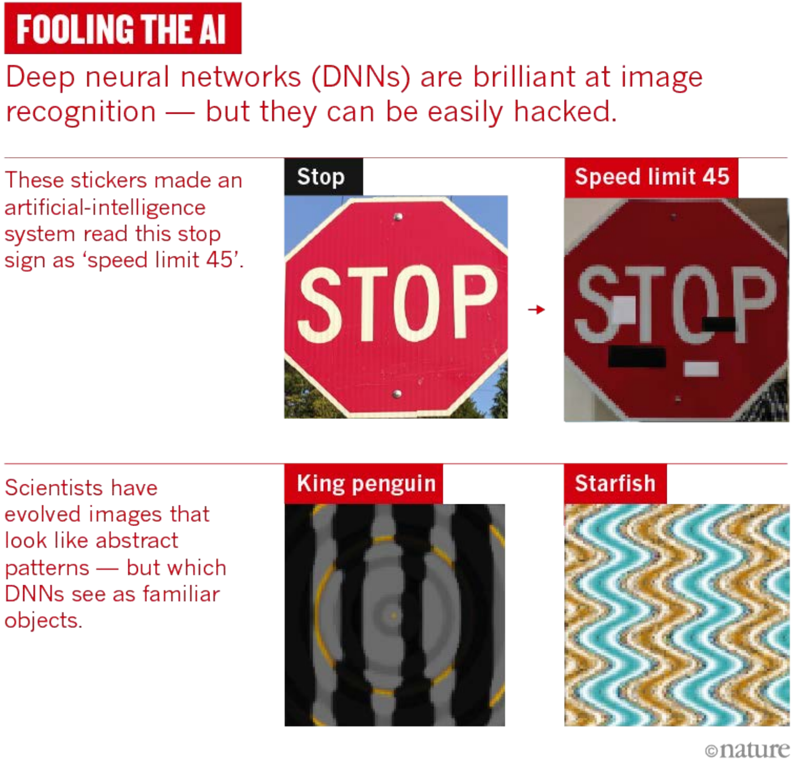
(a) High Cost of spurious and missed patterns (b) Spatial Auto-correlation

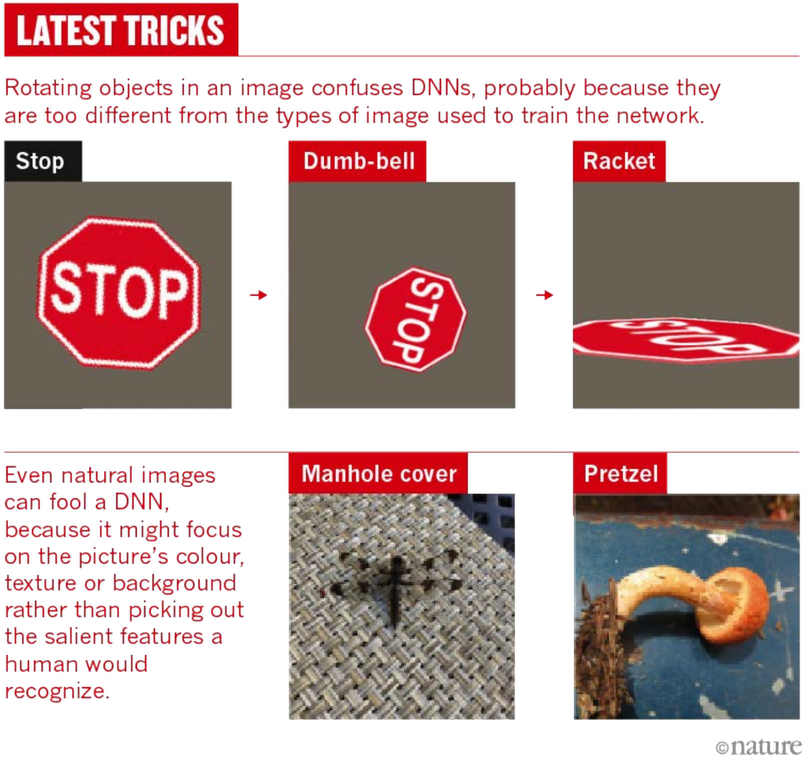
(c) Spatial Heterogeneity (d) Teleconnections



Details: Towards Spatial Variability Aware Deep Neural Networks (SVANN): A Summary of Results, J. Gupta, Y. Xie, and S. Shekhar, DeepSpatial2020 (1st ACM SIGKDD Workshop on Deep Learning for Spatiotemporal Data, Applications, and Systems). Best paper award.

Should AI (e.g., deep learning) reduce false patterns?

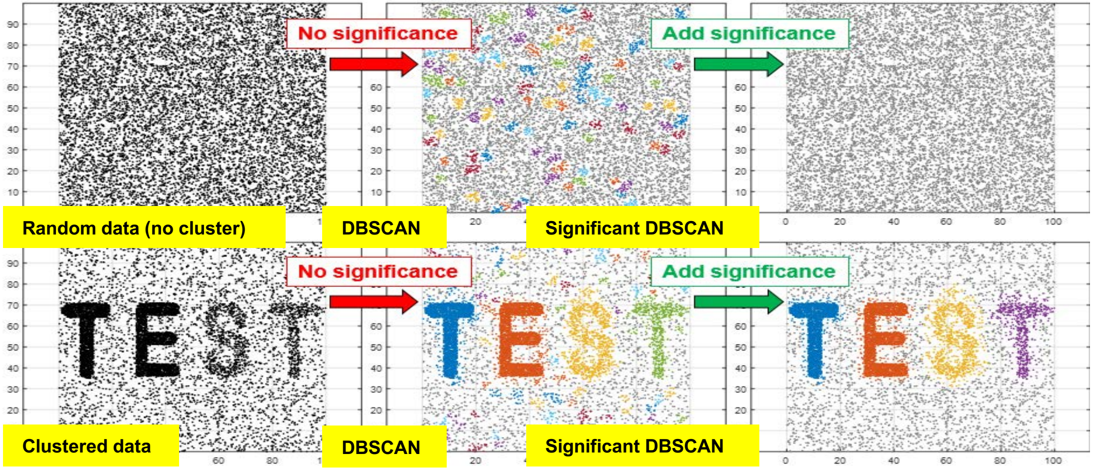


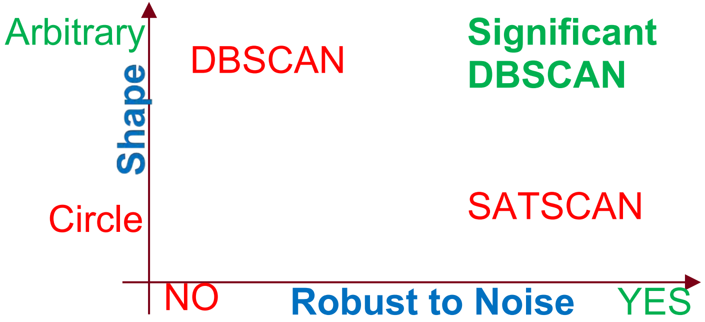


Source: [Why deep-learning AIs are so easy to fool](), D. Heaven, Nature, 09 Oct. 2019.

Guard Rails for AI: Ex. Statistical Significance

- Q? How to detect Statistically Significant Arbitrary Shape Hotspots?
- **Significant DBSCAN** [SSTD 2019]
  - Significance modeling in DBSCAN + A fast dual-convergence algorithm





Details: Significant DBSCAN towards Statistically Robust Clustering, Y. Xie & S. Shekhar, Proc. 16th Intl. Symp. on Spatial and Temporal Databases (SSTD '19), 2019, ACM **(Best Paper).**

Help AI (Vision) break out of "RGB+Lidar" box to try richer sensors?

A downpour in Las Vegas at the CES technology show a year and a half ago may prove to have been a watershed moment in the race to develop autonomous cars. While other companies promoting their experimental self-driving vehicles had to keep them parked in the rain, one company, AdaSky, demonstrated how its sensors could see people hundreds of feet ahead even in a downpour, and even when they were wearing white and standing against a white background.

"Thermal imaging is the best sensor at detecting people, day or night," Chris Posch of FLIR Systems said.

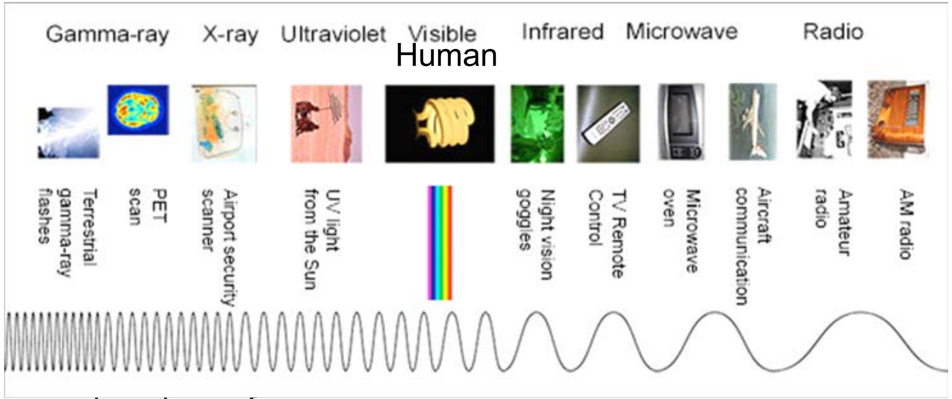
These High-Tech Sensors May Be the Key to Autonomous Cars

The New York Times

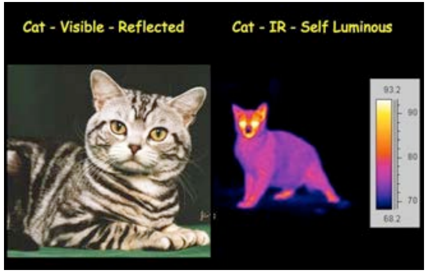
Sept. 26, 2019

Human vision → Superhuman Vision

- Sensors: Electromagnetic, sonar, …
- Ex. Electromagnetic radiation Emitted by objects above absolute zero (0° Kelvin(K), −273°C)



Source: imagine.gsfc.nasa.gov

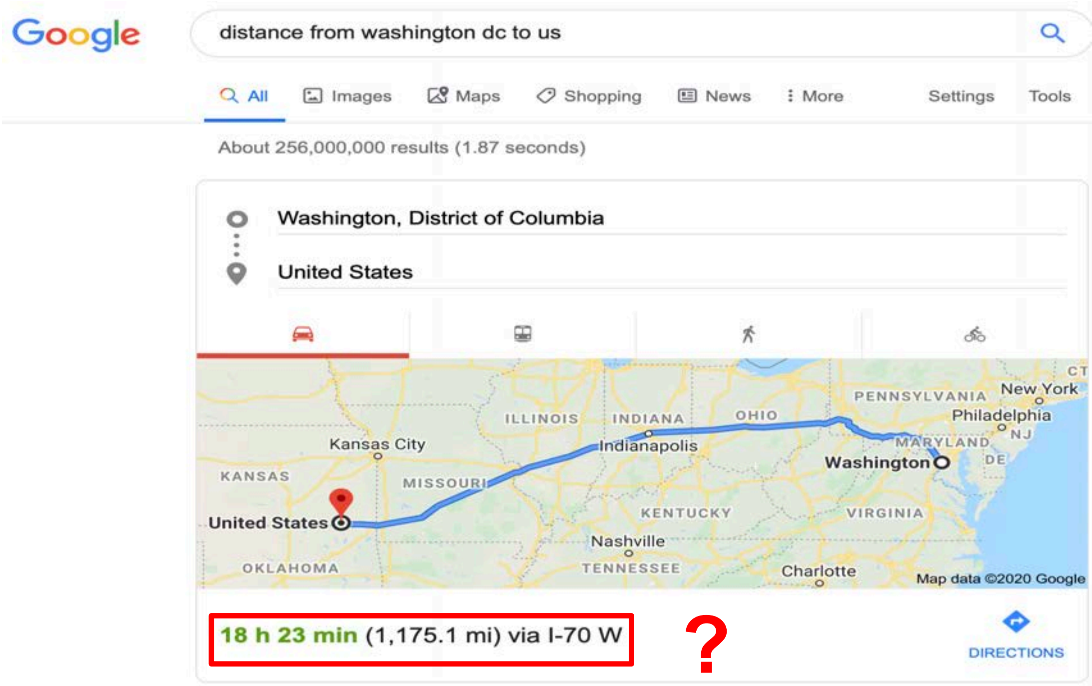


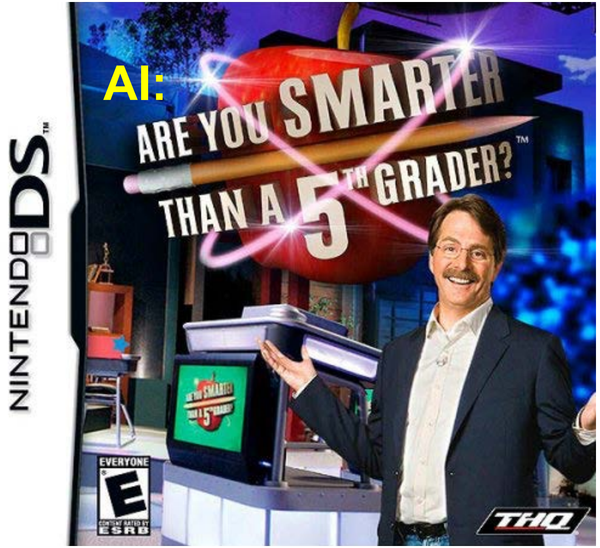
Source: directthermography.co.uk

Should AI learn Extended Spatial Data Types?

Q? What is distance between Washington D.C. and U.S.A.?

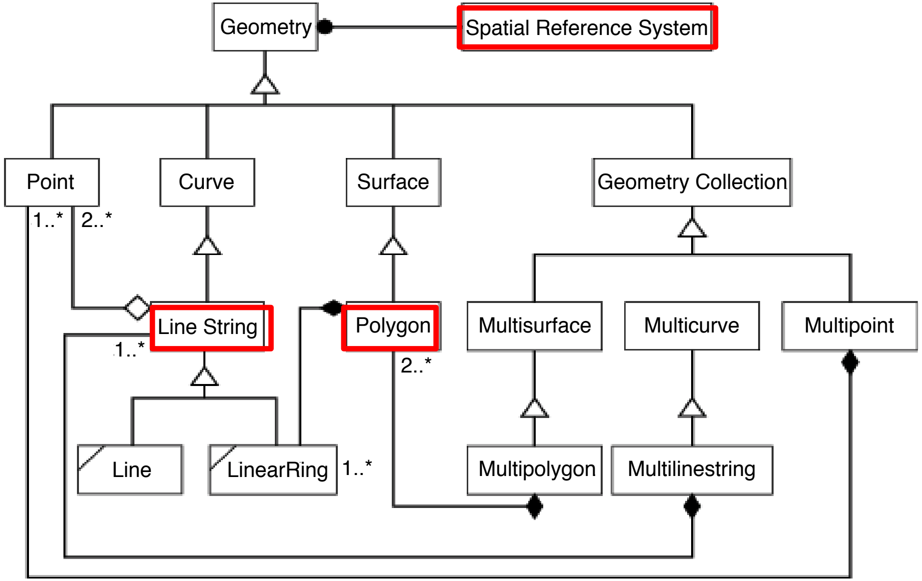
- Zero (Washington D.C. is inside U.S.A.)
- NSF OKN funded 2 grants on geo-knowledge networks!





Spatial Data Types: OGC Simple Features Standard

- Data types: Point, LineString, Polygon, Collections
- Relationships: Topological, Metric, …
- Helps feature selection for machine learning
    - Ex. Distance to key geo-features, Neighbor relationship

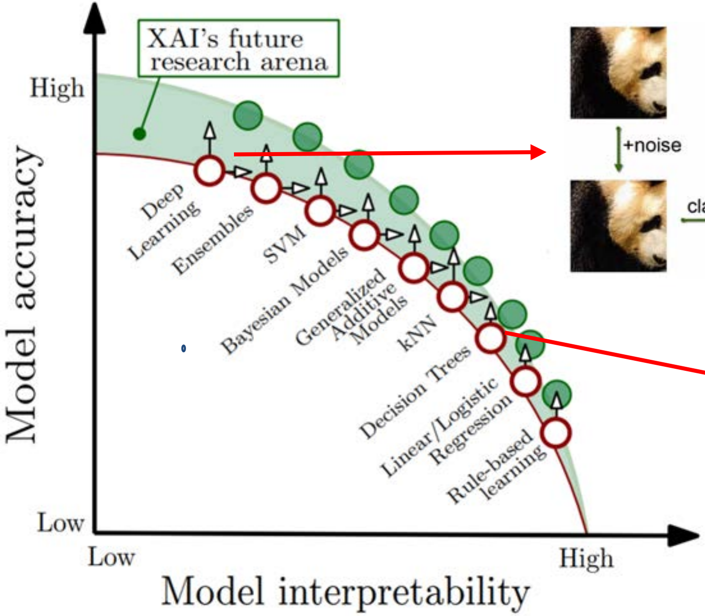


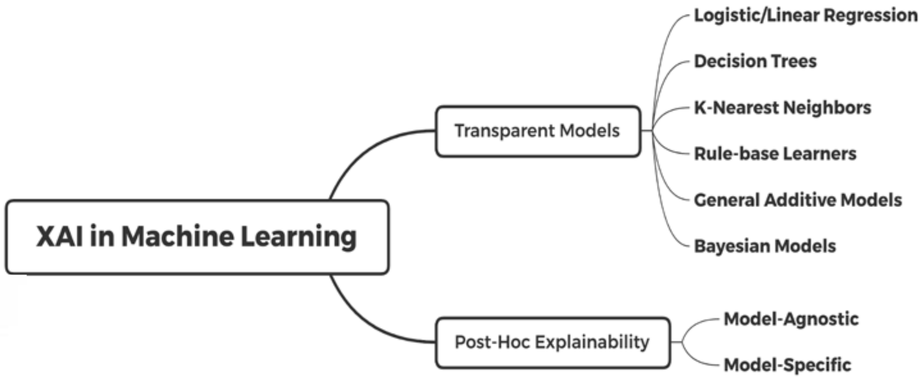
Basic Functions	SpatialReference ()
	Envelop ()
	Export ()
	IsEmpty ()
	IsSimple ()
	Boundary ()
Topological / Set Operators	Equal
	Disjoint
	Intersect
	Touch
	Cross
	Within
	Contains
	Overlap
Spatial Analysis	Distance
	Buffer
	ConvexHull
	Intersection
	Union
	Difference
	DymmDiff

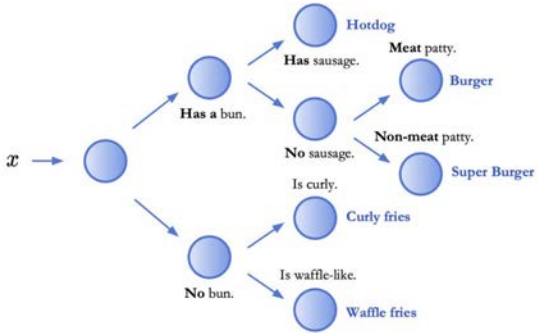
Details: [Spatial Databases: Accomplishments and Research Needs](), S. Shekhar et al., IEEE Trans. on Knowledge and Data Eng., 11(1), Jan.-Feb. 1999.

Explainable AI

- How may AI decision be understood by humans
- Accuracy Interpretability Trade-off







[Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI]().
A. Arrierta et al., Information Fusion, Volume 58, Pages 82-115, Elsevier, June 2020.

Summary

The World Economy
Runs on GPS.

- Spatial Data has already transformed our society
  - It is only a beginning!
  - It promises astonishing opportunities in coming decade

- AI has promise but faces major challenges
  - Rich Data Types, e.g., lineStrings, polygons, …
  - High cost of errors, Spatial Heterogeneity, …

- **Ask**
  - Sponsors: Nurture approaches to overcome challenges (Geo-AI
  - Academics: Include Spatial topics in courses and curricula

Resources

- Spatial Computing, MIT Press (Essential Knowledge Series), 2020. (with P. Vold)
- Spatial Computing (html , short video , tweet), Communications of the ACM, 59(1):72-81, January 2016 (With S. Feiner, and W. Aref).
- Transdisciplinary Foundations of Geospatial Data Science (html , pdf) ISPRS Intl. Jr. of Geo-Informatics, 6(12), 2017. doi:10.3390/ijgi6120395. (with Y. Xie, E. Eftelioglu, R. Ali, X. Tang, Y. Li, and R. Doshi)
- Transforming Smart Cities With Spatial Computing, Proc. IEEE Intl. Smart Cities Conference , 2018 (with Y. Xie, J. Gupta, Y. Li).
- Encyclopedia of GIS (2nd Ed.), Springer, 2017,isbn978-3-319-17884-4.(Co-Ed.w/ H. Xiong, and X. Zhou). (1st Ed. in 2008, isbn 978-0-387-30858-6).
- A Tour of Spatial Databases, Prentice Hall, 2003, isbn013-017480-7.(w/ S.Chawla).
- A scalable parallel formulation of the backpropagation algorithm for hypercubes and related architectures, IEEE Transactions on Parallel and Distributed Systems, 5(10):1073-1090, Oct. 1994 (w/ V. Kumar, M. B. Amin).
- Software Development Support for AI Programs, ,IEEE Computer, 20,(1):, 30-40, Jan. 1987, (w/ C.V. Ramamoorthy, and V. Garg,) .