Acquisition of proxy data based on satellites

Norwegian Computing Center's role in AIRQUIP

Arnt-Børre Salberg

AIRQUIP KO meeting

MET Norway, Oslo, April 19, 2017

Norsk Regnesentral (NR)

NR is a research institute that performs contract research in data analysis for the

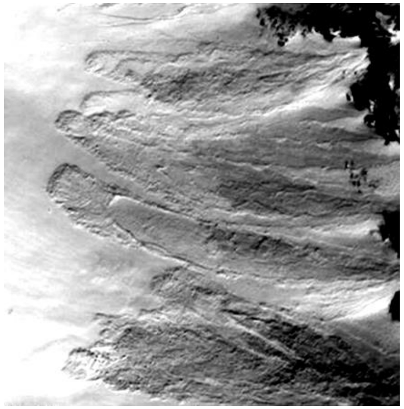
- oil and gas industry
- finance, insurance and commodity markets
- climate, environment, marine resources and health
- technology, industry and administration
- Earth observation

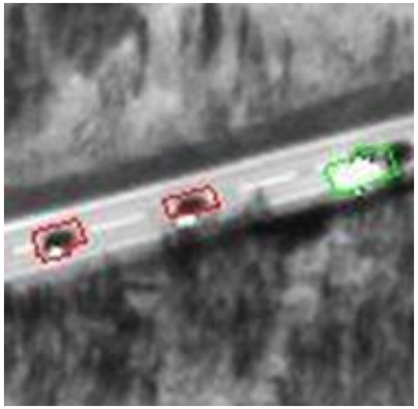
Section for Earth Observation

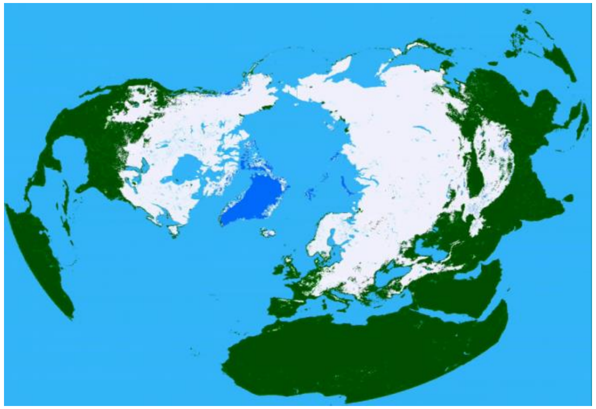
NR has been a leading research and development institute in Earth Observation since Norway started to focus on satellite remote sensing in the beginning of the 1980's.

Main focus

R&D on algorithms and methods for automatic and semi-automatic analysis of remote sensing data







Objectives

Main objective

- Acquire, process and quality control existing datasets, and produce new traffic density data using satellite imagery.

NR's focus

- Develop methodology for detecting and counting vehicles in VHR satellite images. Classify the vehicle types.
- Analyze spectral differences between Sentinel-2 images along roads in order to construct a proxy for the traffic density

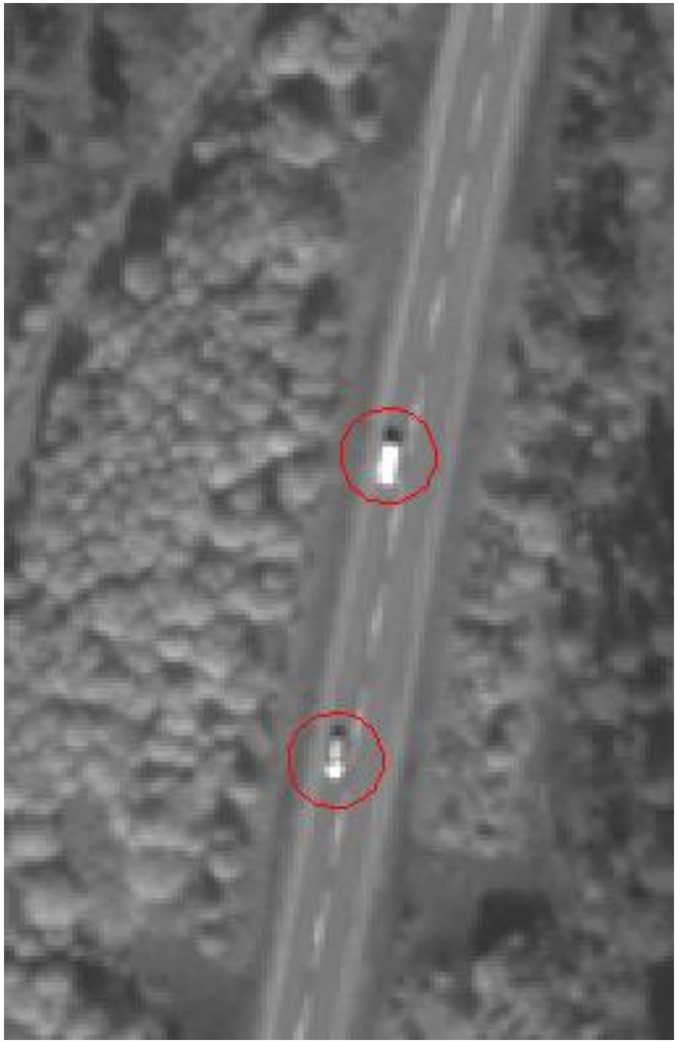
Main objective

Acquire, process and quality control existing datasets, and produce new traffic density data using satellite imagery.

Proxy dataset	Representing	Norway	Europe	China
Road network data	Traffic emissions	NVDB (www.vegdata.no)	Open street map or other routing services	Data from local governments and from the Digital Road-network Map (CDRM)
Traffic volume data	Traffic emissions	NVDB, satellite imagery for ground truthing and gap filling	Individual national databases, satellite imagery, TREMOVE scenarios for major roads.	Estimates from emission inventories, satellite imagery
Satellite imagery for traffic volume data	Traffic emissions	Google Earth	Google Earth	-
Shipping position and emission data	Shipping emissions	AIS data, available from kystverket, www.havbase.no	AIS data and national databases	-
Industrial site positions and emissions	Industrial emissions	Available from the Norw. Environment Agency (www.norskeutslipp.no)	Available from EPRTR (prtr.ec.europa.eu)	For Beijing and Guangzhou (local environm. agencies)
Population and dwelling density	Domestic heating emissions	Available from SSB at 250 m resolution for all of Norway (www.ssb.no/natur-og-miljo/geodata) or home addresses from kartverket	Available from EEA (www.eea.europa.eu/data-and-maps/data/population-density-disaggregated-with-corine-land-cover-2000-2)	1x1 km resolution from LANDSCAN
Topographical data	For pseudo dispersion modelling	Available for all of Norway at 10 - 50m (data.kartverket.no)	Available from EEA at 30 m, (www.eea.europa.eu/data-and-maps/data/eu-dem)	-
Building positions	For pseudo dispersion modelling	Available for all of Norway at municipality level on shp format (shapefiles at 'norge digitalt', www.geonorge.no/)	Unknown	To be decided during the project.
Improved temporal emission data	For EMEP calculations	Not applicable	Not applicable	Data from Tsinghua Research group

Traffic statistics

- AADT = annual average daily traffic
  - measures vehicle density
  - applied in models for noise and air pollution?
- Ground based instruments
  - current method for traffic counts
  - point-wise measurements
  - areas with high traffic density



SatTraffic project

Goal: Estimate traffic statistics automatically from very high resolution satellite images.

Funding: Project with Norwegian Road Administration.

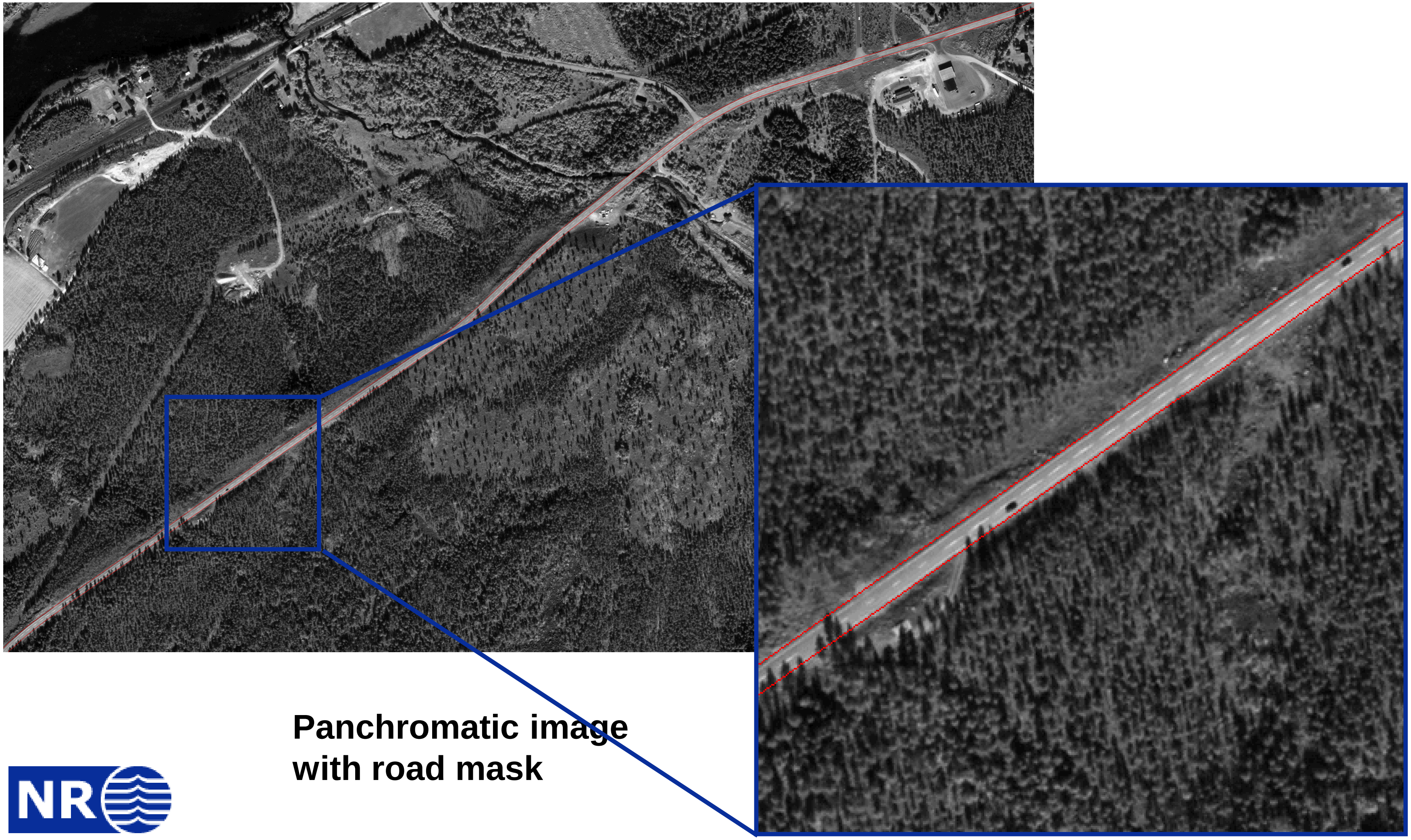
NRs role: Developed an image and pattern analysis system that

- automatically counted the number of vehicles in the satellite images, and
- calibrated “road traffic snapshots” to AADT.

Data

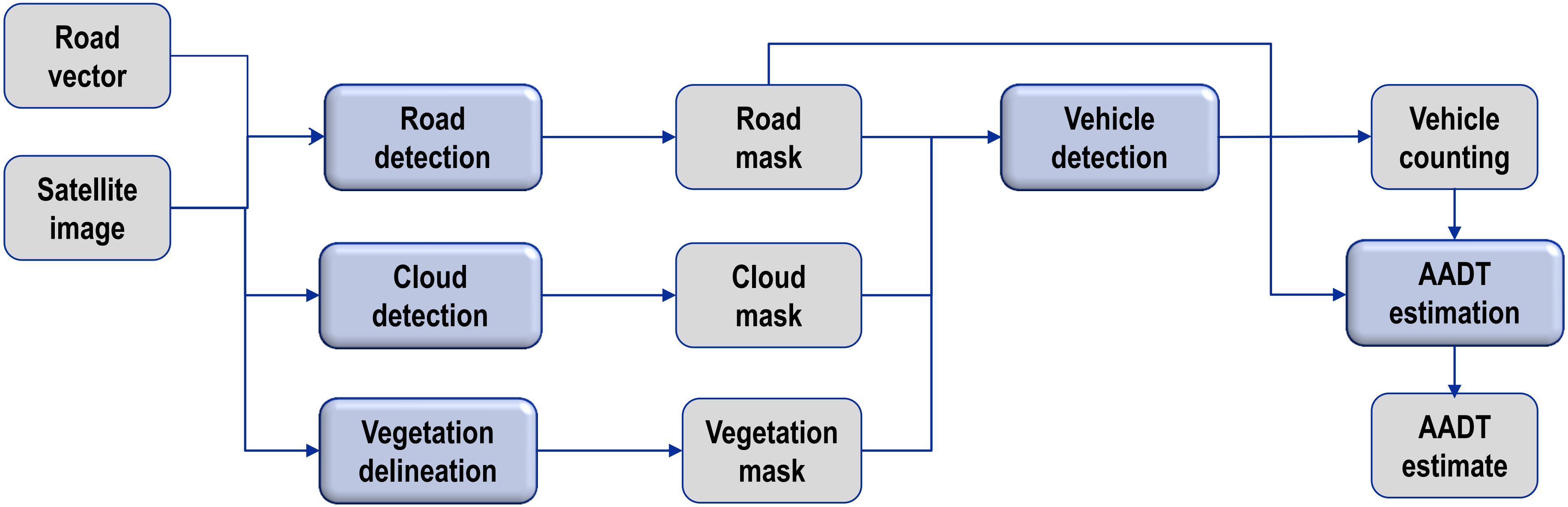
- VHR satellites (WorldView, QuickBird, etc.):
  - Approx. 0.5m ground resolution in the panchromatic band
- Road vectors
- Focus on Norwegian conditions
  - narrow roads
  - low traffic density
  - low elevation of the sun
  - trees and shadows extensive problem

Data

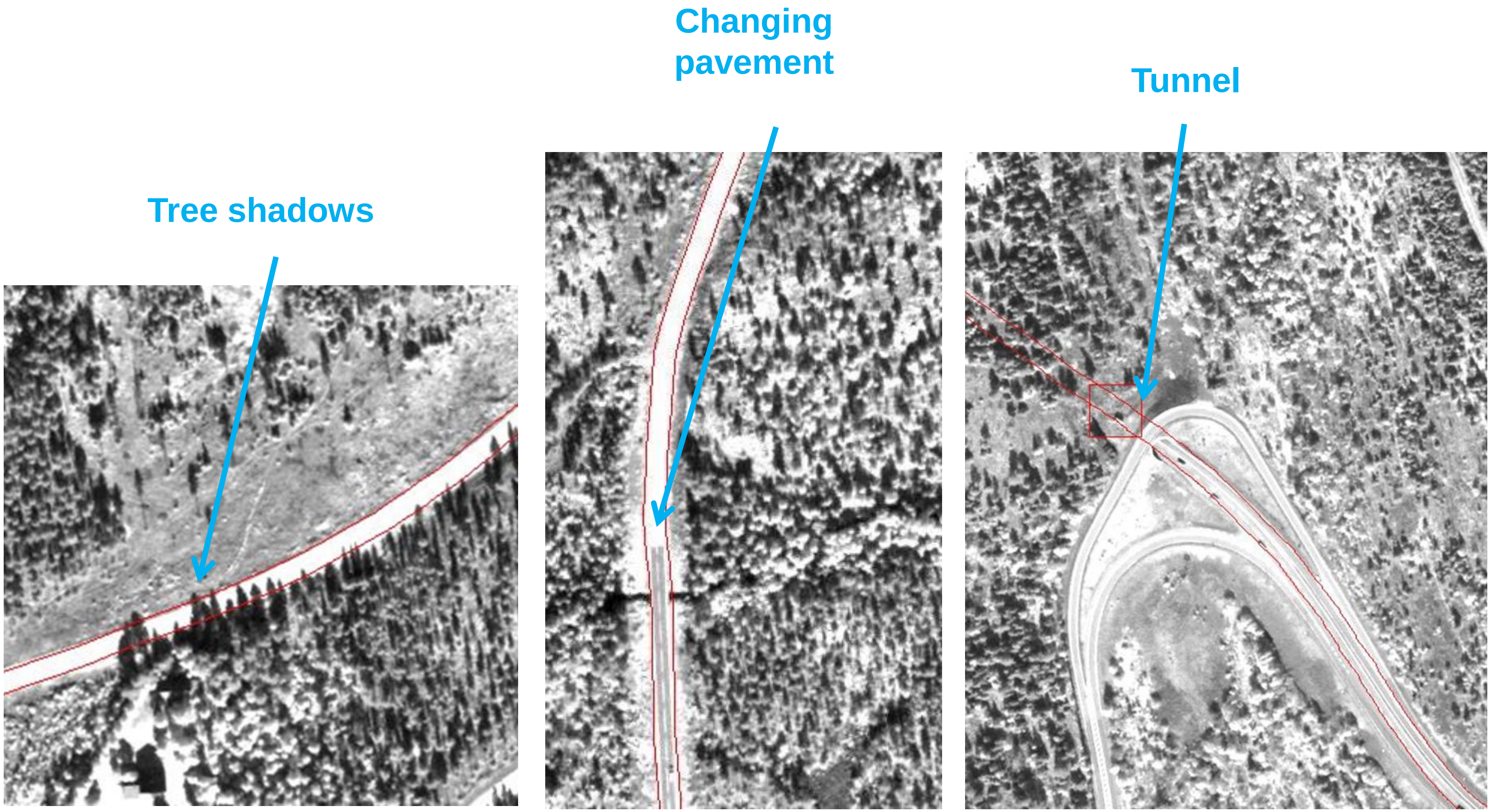


Panchromatic image with road mask

Processing chain



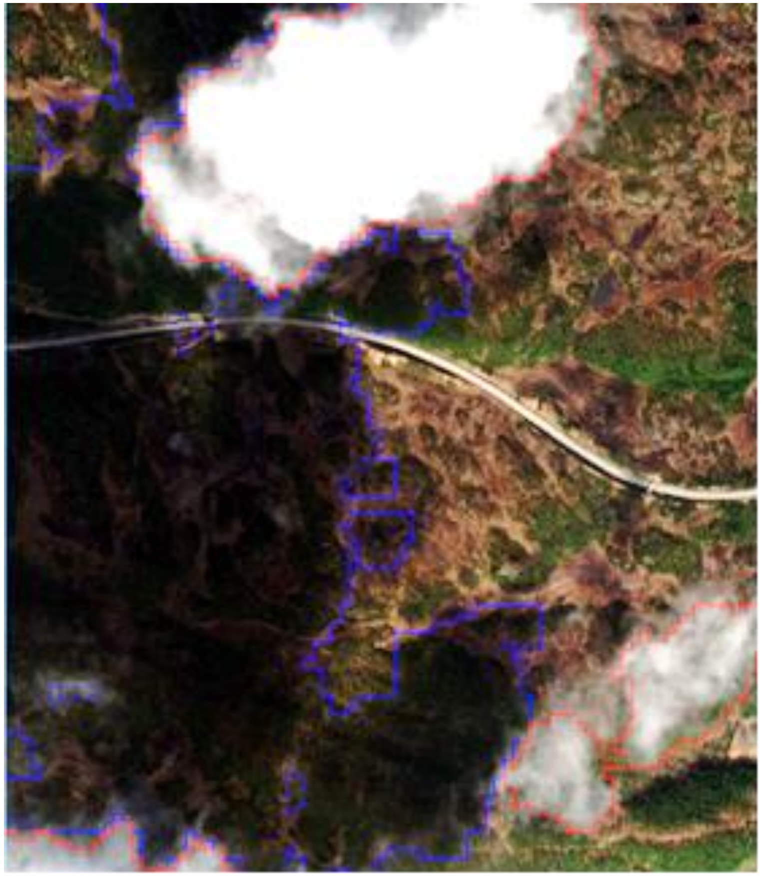
Road detection

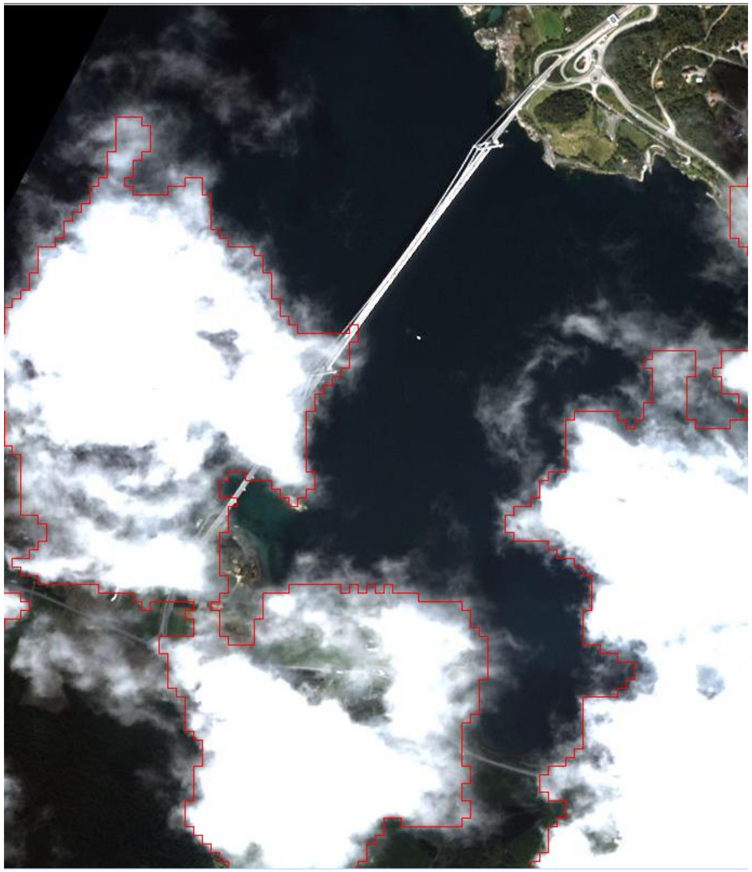


Cloud detection

Blue lines: Cloud shadows

Red lines: Clouds

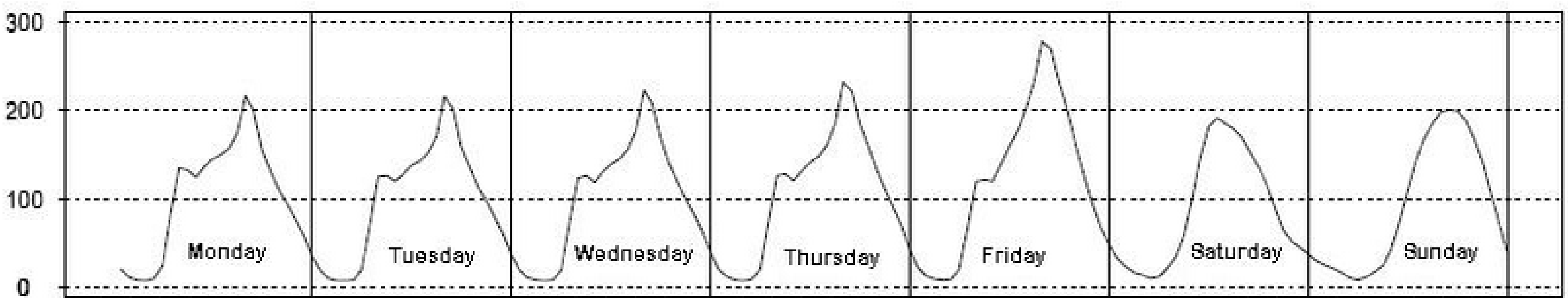




Estimation of AADT

Basis curve method:

- Enables estimation of AADT from short-time counts.
- The methods takes into account holidays and special days.



AADT estimate is computed using information about the number of vehicles, length of observed road segment and speed limit.

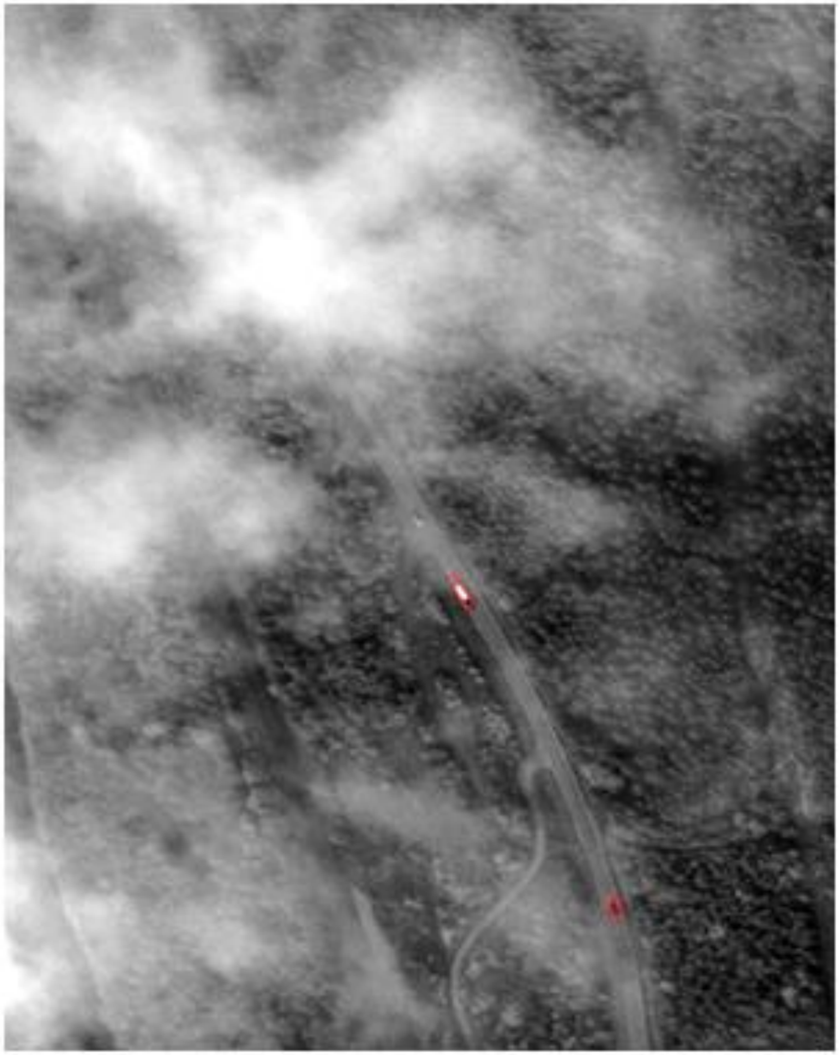
Validation

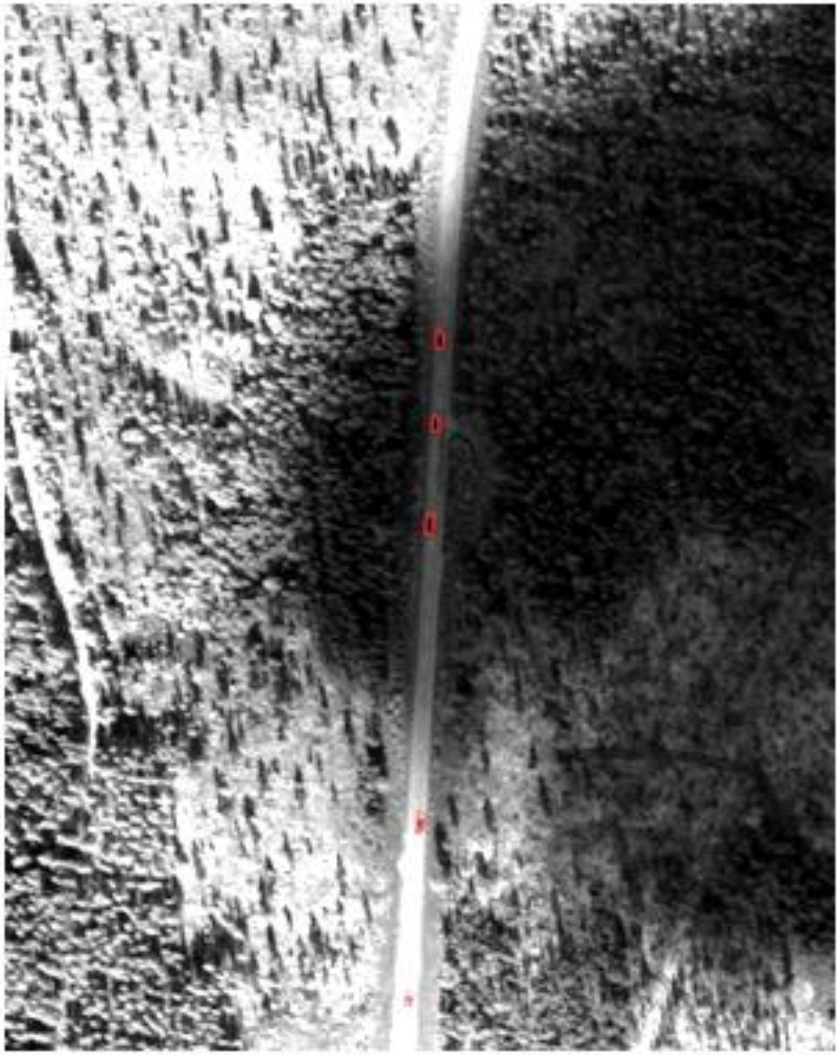
All images	Vehicle type			
	car	truck	mc	total
True	239	68	7	314
Detected	205	57	6	268
Wrong type	4	20	6	30
Double counts	2	2	0	4
False detection	24	3	0	27
Segmentation ok	228	63	7	298

Detection rate: 85.4%
False detection rate: 8.6%

Validation

Works for challenging cases.





Pros and cons

Pros

- Automatic
  - AADT estimation possible from the office.
- Costs
  - 50% cheaper than current methods.
  - Price is falling?
- Little manpower needed.
  - No personnel needed to conduct short time counting.
- Geographically independent.
  - AADT may be estimated from all road segments.
- No privacy concerns

Cons

- Precision
  - Lower than manual counting.
  - Dependent of length and speed limit of road segment.
- Quality
  - System sensitive to the precision of the segmented road.
  - Need more work on distinguishing vehicle types.
- Engine type
  - Cannot determine the engine type.
- Data are normally not free of charge.
- Method developed for rural areas.

Conclusion

- Automatic traffic counts from satellite
  - based on image analysis and pattern recognition
  - limited time resolution
  - statistical model for AADT estimation
  - no maintenance costs
  - suitable for time-limited monitoring

- Prototype ready for pre-operational testing.

- Possible to improve the detection performance
  - Development of new object recognition methods during the last years have made me with good hope ☺

Deep learning – a revolution in computer vision

The Mobile Internet Is Over. Baidu Goes All In on AI

The Chinese company has more than 1,300 people working deep learning.

CADE METZ BUSINESS 12.25.16 7:00 AM

2016: THE YEAR THAT DEEP LEARNING TOOK OVER THE INTERNET

The amazing artificial intelligence we were promised is coming, finally

Forbes / Tech

DEC 29, 2014 @ 5:37 PM 80,006 VIEWS

Tech 2015: Deep Learning And Machine Intelligence Will Eat The World

Forbes / Tech

DEC 29, 2013 @ 6:56 PM 73,405 VIEWS

Why Is Machine Learning (CS 229) The Most Popular Course At Stanford?

theguardian

UK world politics sport football opinion culture business lifestyle fashion environment tech travel

home › tech

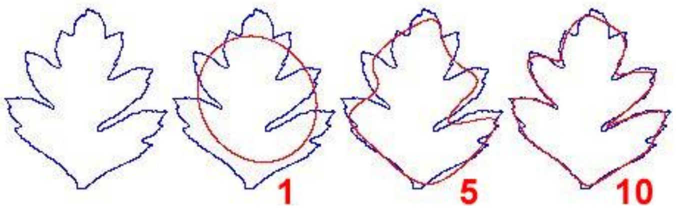
Artificial intelligence (AI)

Google buys UK artificial intelligence startup Deepmind for £400m

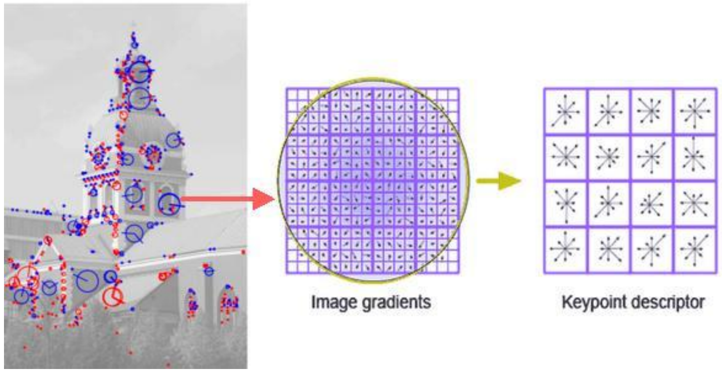
NR Microsoft Google facebook amazon web services Baidu 百度

Good **features** are essential for successful analysis in many applications!

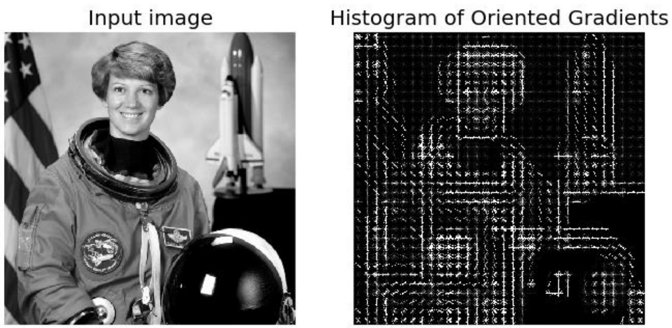
- Features = data representation
  = data description



- Good features capture posterior belief about explanatory causes and underlying factors of variation.



- Multitude of hand-designed and fixed features currently in use



Designing good features is difficult!

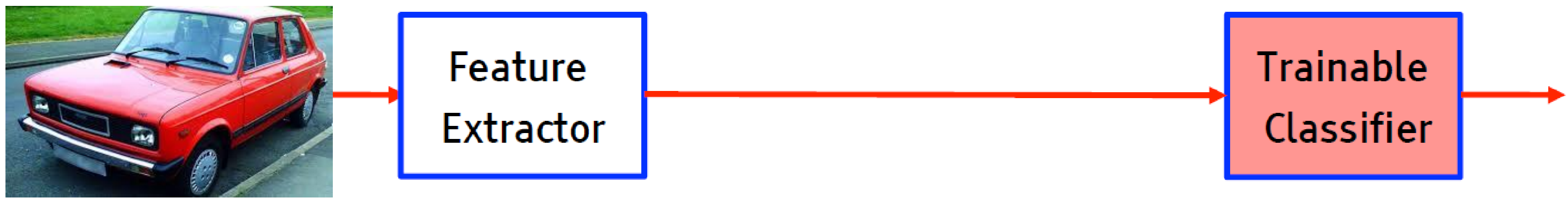
Shape and color features works fine here

But which features would you propose to describe the content in these images?

Images borrowed from LeCun et al. (2015)

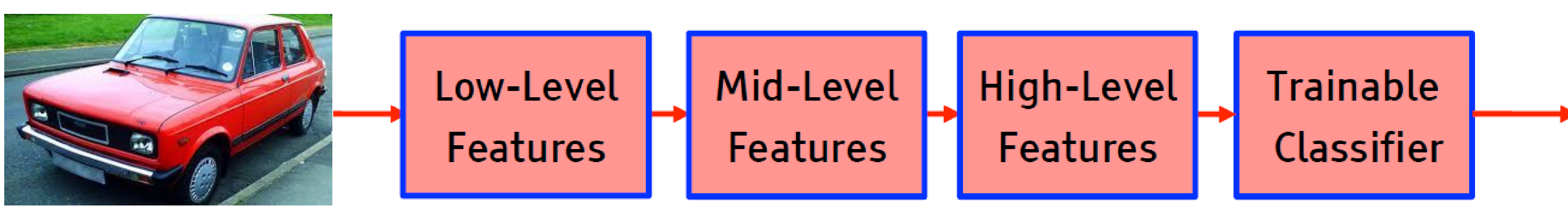
Architectures: Image recognition systems

Traditional way: Handcrafted features + supervised classifier



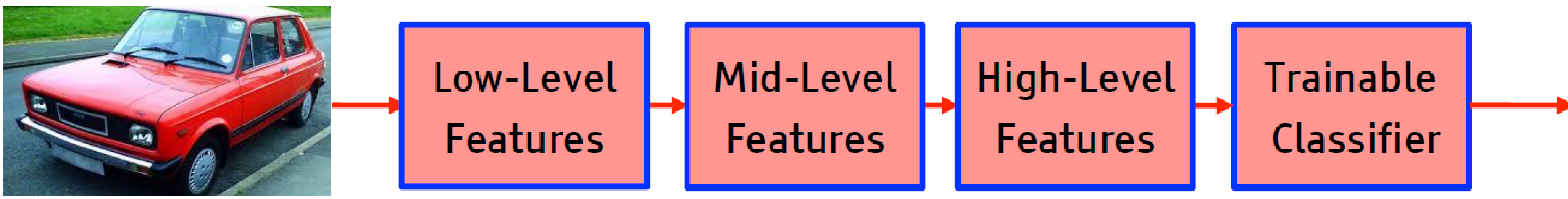
Mainstream approach (until recently) in image (and speech) recognition

Deep learning: Multiple stages/layers trained end-to-end



Deep learning: Multiple stages/layers trained end-to-end

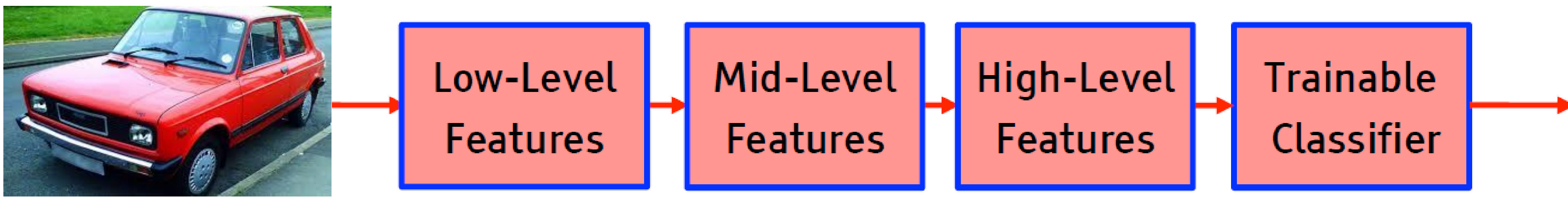
- A hierarchy of trainable feature transforms
- Each module transforms its input representation into a higher-level one.



How can we make all the modules trainable and get them to learn appropriate features (representations)?

Deep learning: Multiple stages/layers trained end-to-end

- A hierarchy of trainable feature transforms
- Each module transforms its input representation into a higher-level one.

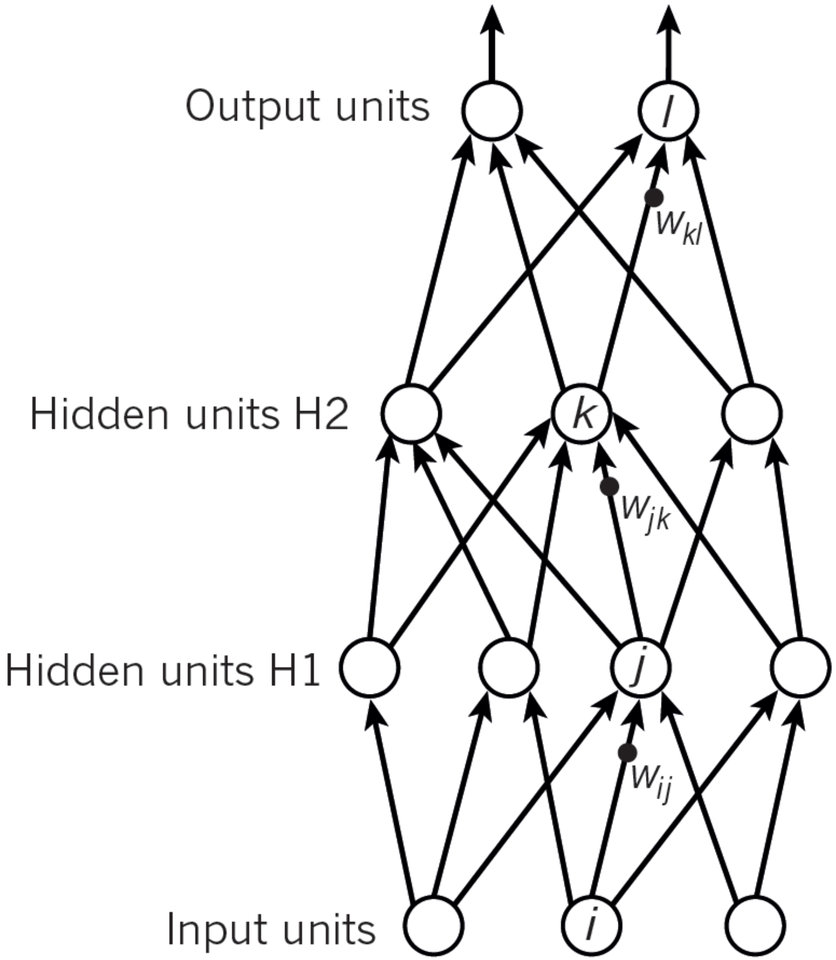


How can we make all the modules trainable and get them to learn appropriate features (representations)?

Deep neural networks!

The secret of Deep Learning

Neural networks



$$y_l = f(z_l)$$

$$z_l = \sum_{k \,\varepsilon\, \text{H2}} w_{kl} \, y_k$$

$$y_k = f(z_k)$$

$$z_k = \sum_{j \,\varepsilon\, \text{H1}} w_{jk} \, y_j$$

$$y_j = f(z_j)$$

$$z_j = \sum_{i \,\varepsilon\, \text{Input}} w_{ij} \, x_i$$

Learning:
Estimate the weights $\{w_{kl}\}$ from the data.

f(·) non-linear function

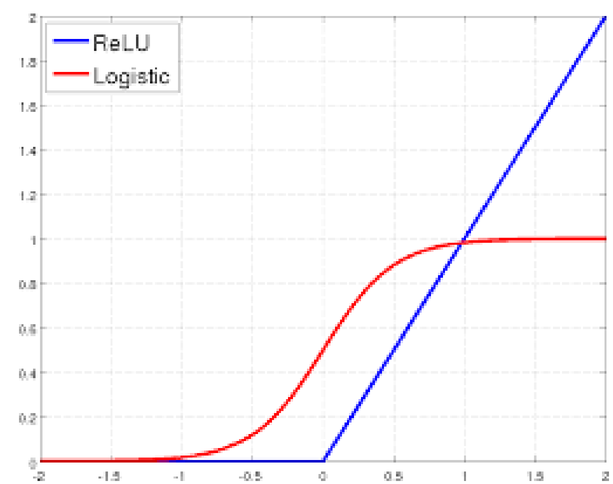
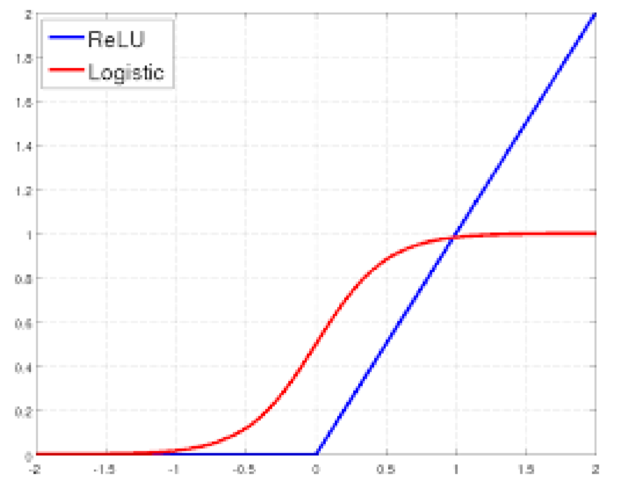


Figure borrowed from LeCun et al. (2015)

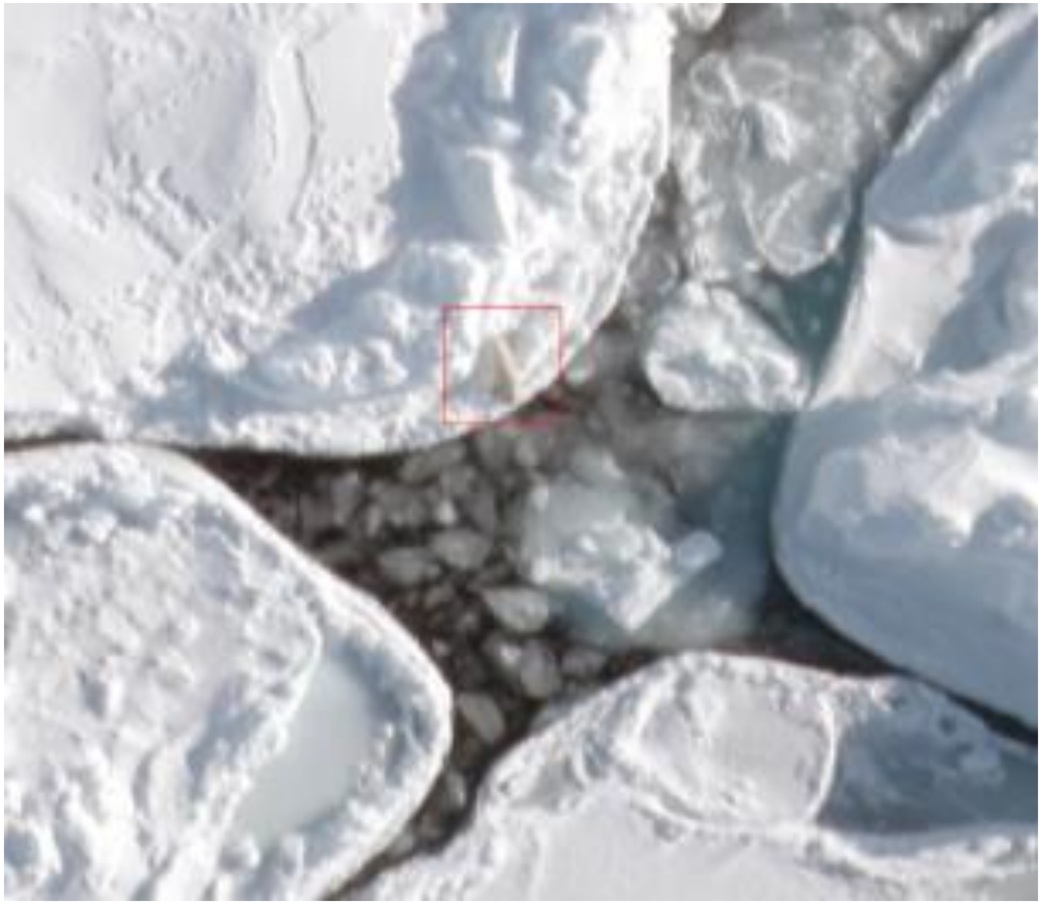
What has happened the last 10 years

- Labeled datasets got much bigger.
- Computers got much faster and use GPUs efficiently.
- Good parameter initialization (pre-training, Xavier), ReLUs instead of sigmoids, model averaging (regularization) like dropout, and batch normalization.



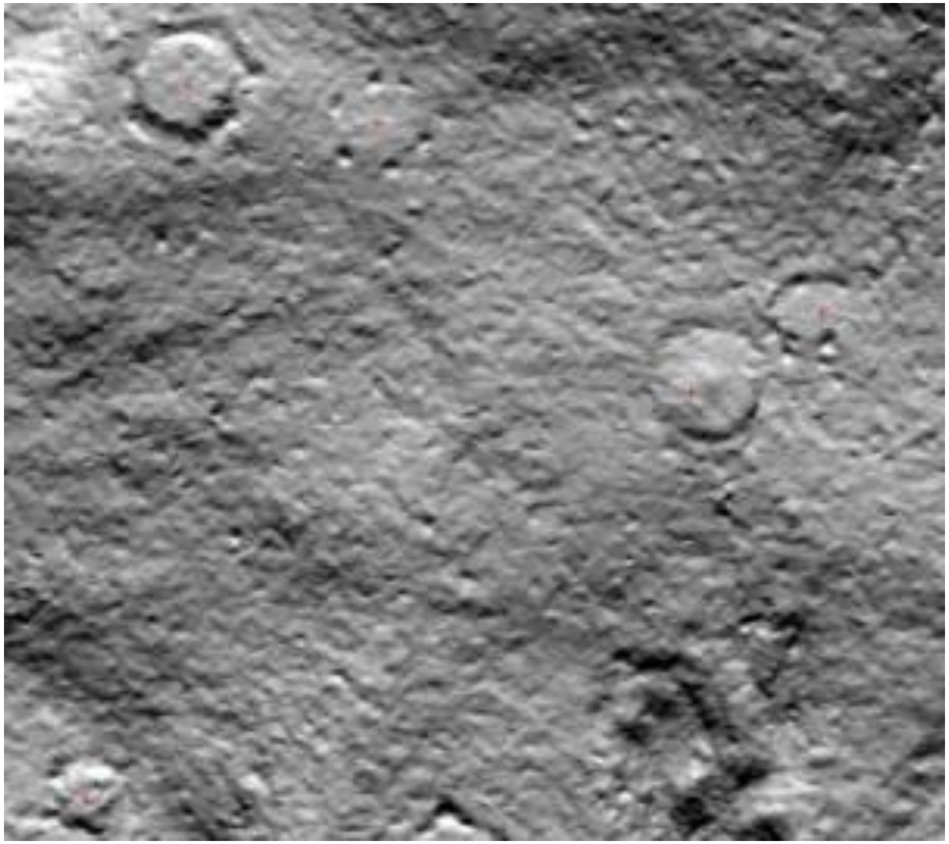
Deep learning remote sensing projects at NR

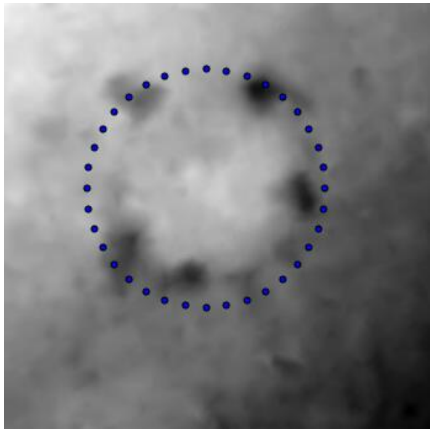
UAVSeal (Institute of Marine Research)
Detection and counting of seals from aerial photos



Deep learning remote sensing projects at NR

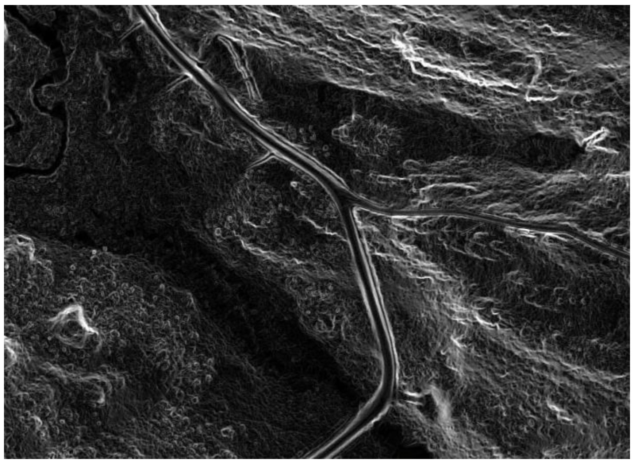
CultSearcher (Directorate for Cultural Heritage)
Detection of kilns from laser data

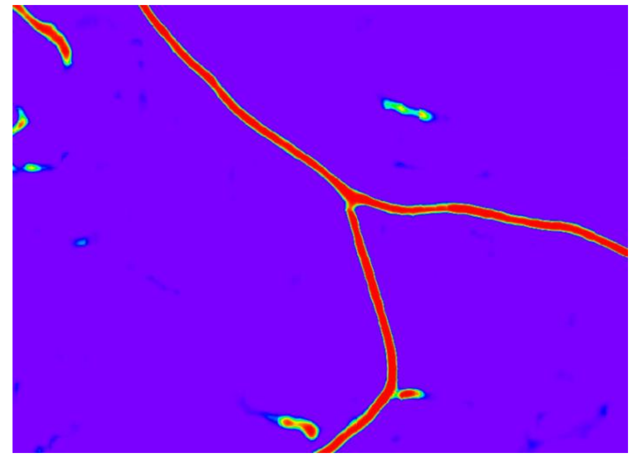


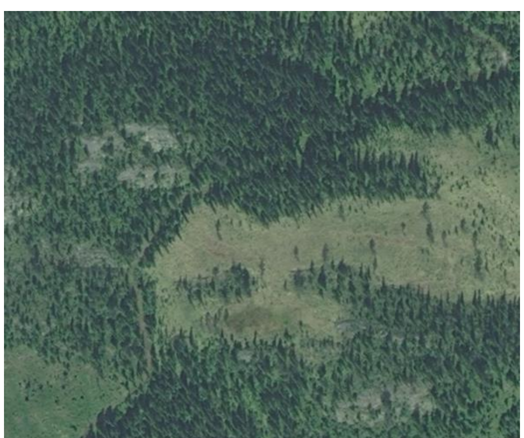


Deep learning remote sensing projects at NR

Lastrak (The Norwegian Mapping Authority)
Mapping of forest roads from laser data

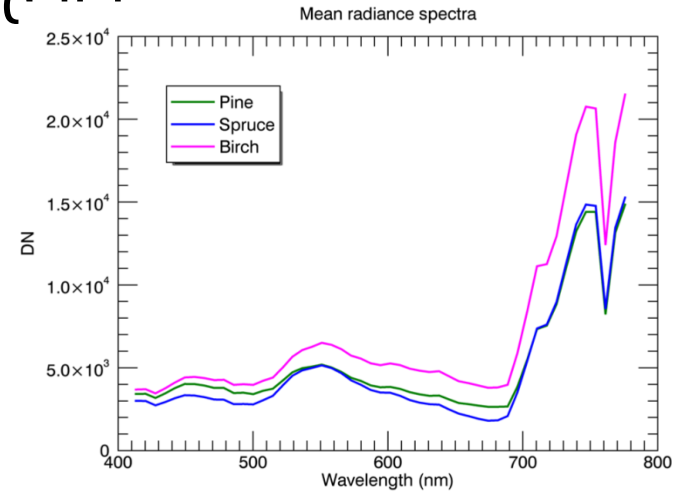


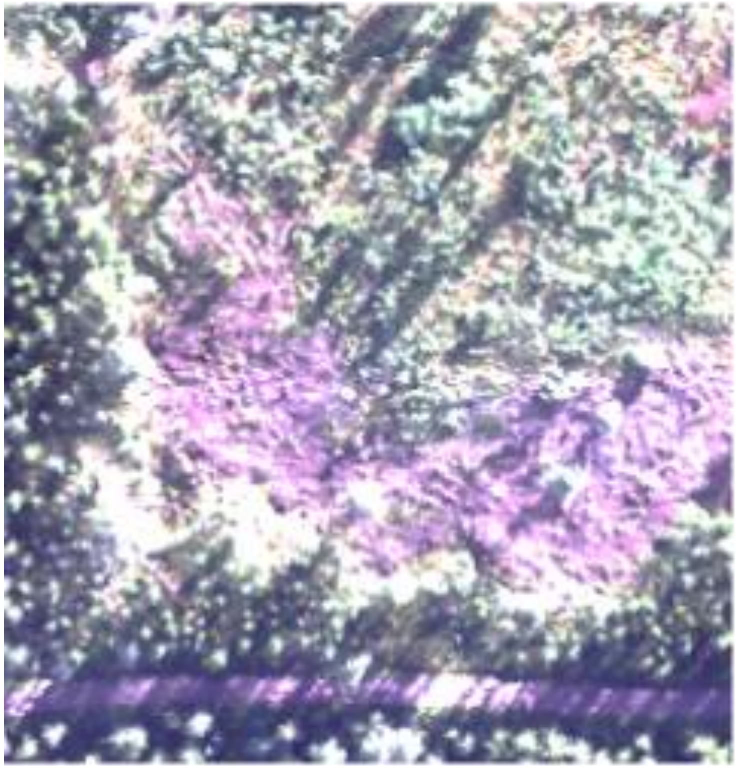




Deep learning remote sensing projects at NR

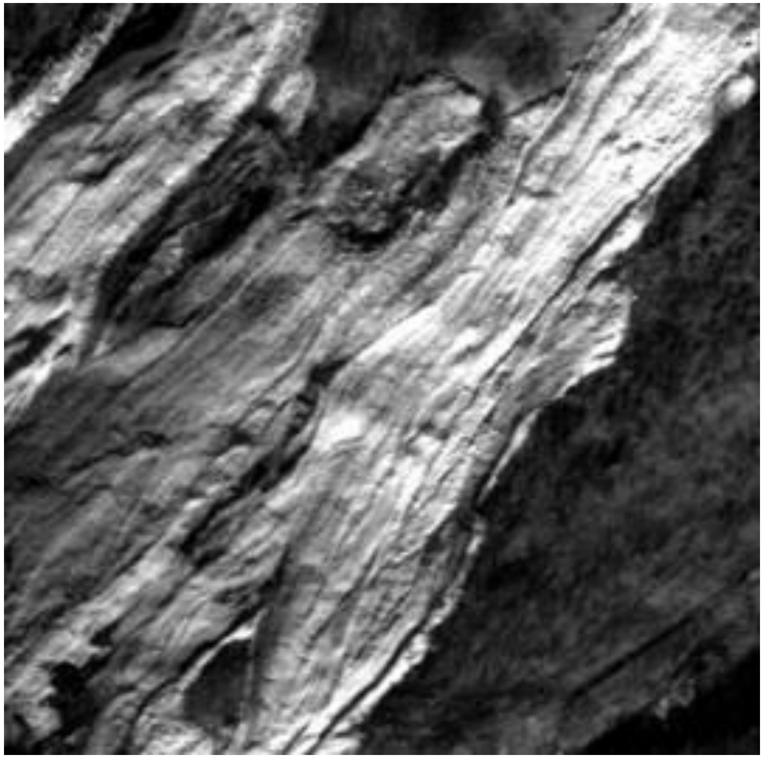
Hyperbio (TerraTec)
Mapping of forest from hyperspectral data
data

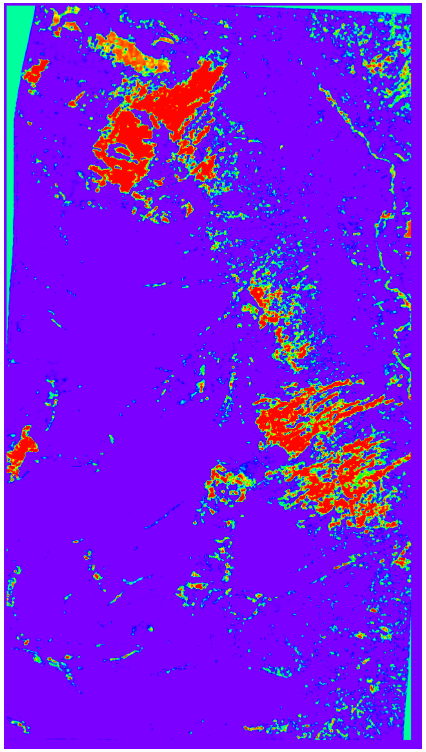




Deep learning remote sensing projects at NR

SnowBall (EEA)
Detection of snow avalanches from satellite images

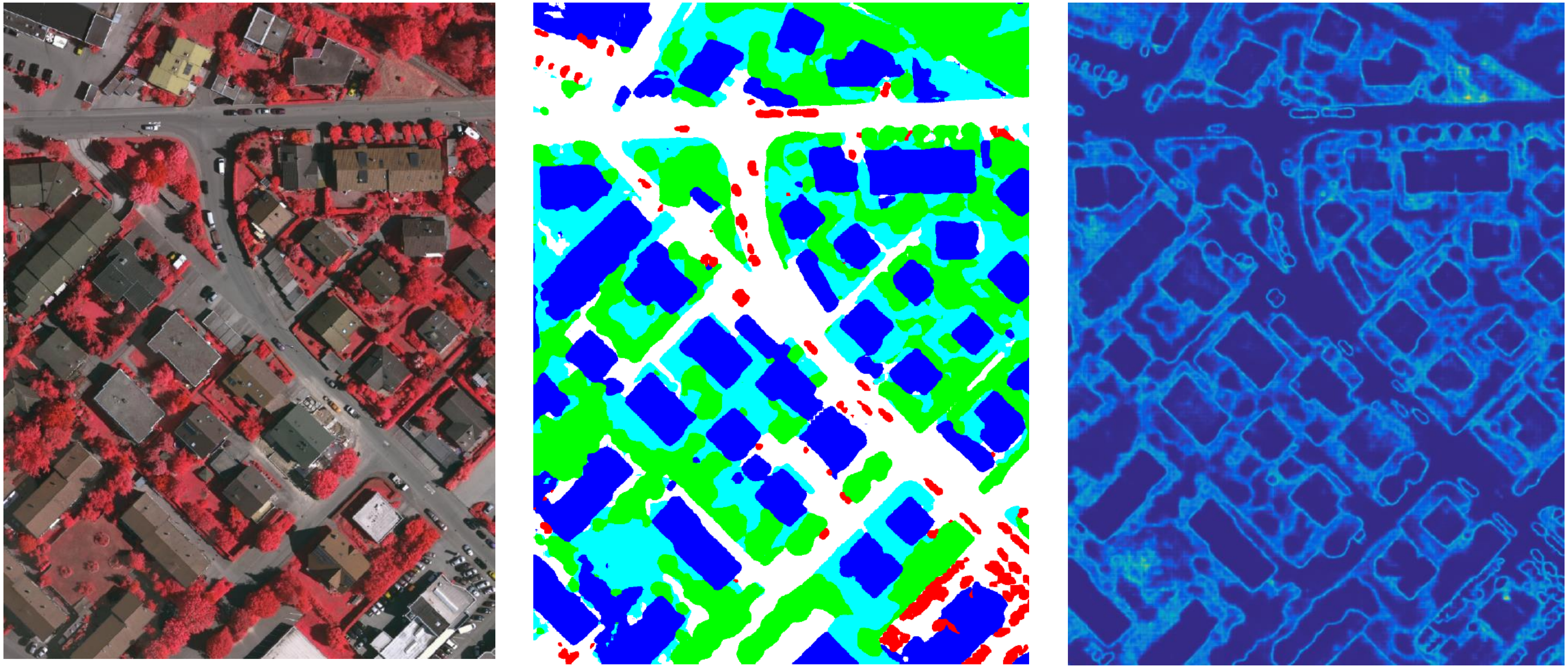




Deep learning remote sensing projects at NR

Univ. Tromsø/NR-project

Mapping of urban areas from aerial images



Potential satellite data source for China

GaoJing / SuperView Earth Observation Constellation

Panchromatic: 0.5 m
Multi spectral: 2 m

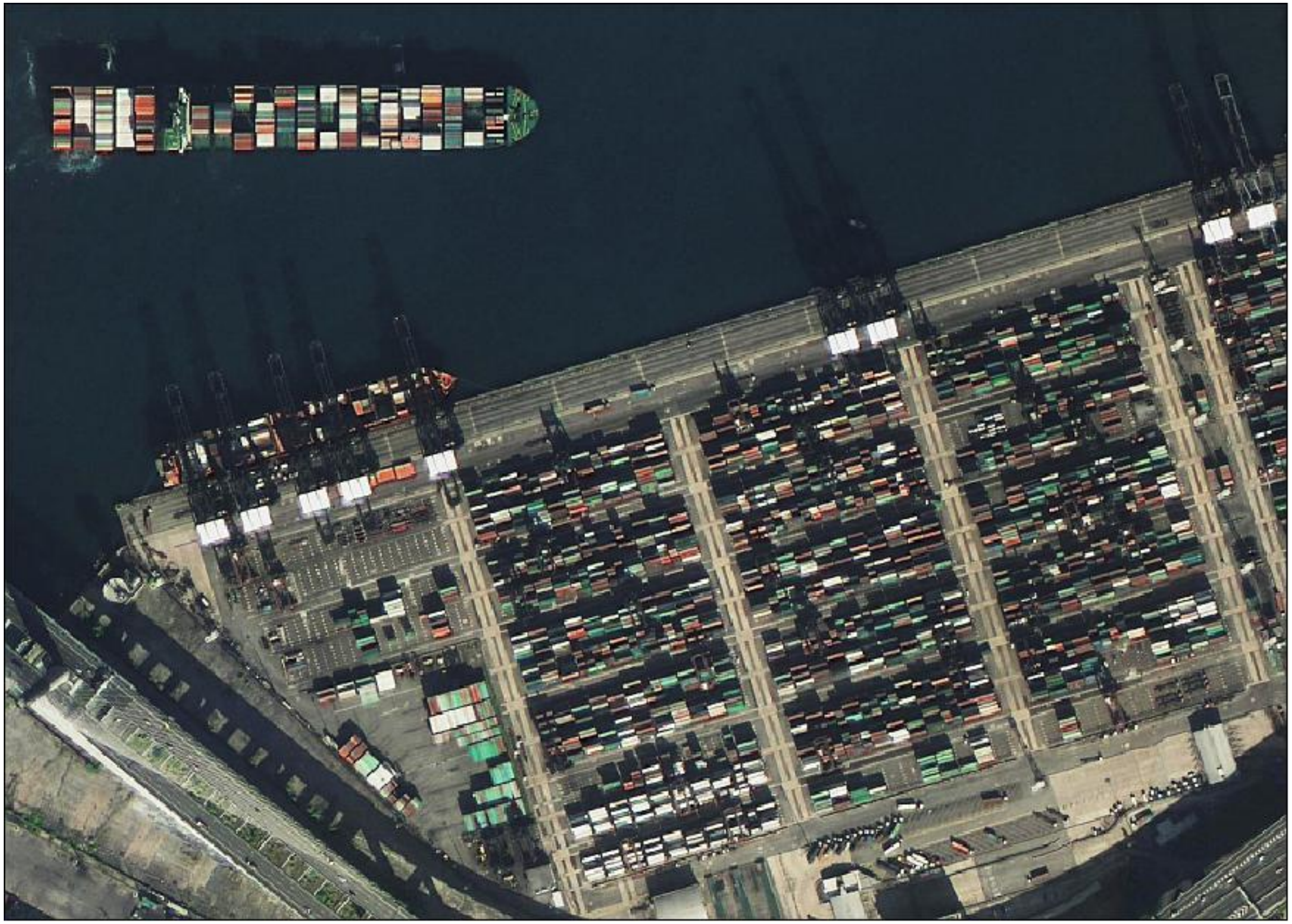
PAN: 0.45-0.89 µm
B1/blue: 0.45-0.52 µm
B2/green: 0.52-0.59 µm
B3/red: 0.63-0.69 µm
B4/NIR: 0.77-0.89 µm

Potential satellite data source for China

GaoJing / SuperView Earth Observation Constellation

Potential satellite data source for China

GaoJing / SuperView Earth Observation Constellation



Potential satellite data source for China

Digital Globe Imagery Grant

Applicants will be granted a limited amount of archive imagery at no cost, provided the following criteria are met:

- The applicant must be enrolled as a student or a faculty member at an accredited university-level educational institution.
- Applicants must agree to provide feedback in the form of an article, thesis or white paper.
- The data may not be used commercially or shared with anyone who might use it commercially.
- Applicants must agree to provide feedback in the form of an article, thesis or white paper, video testimonial, or social media outlets.

Potential sources for ship detection

- AIS
- Sentinel-1.
  SAR satellite
  - Highly suitable for ship detection.
  - Very good coverage in Europe.

New directions

Change detection in Sentinel-2 images

- Can we detect traffic flow?
- The benefit is free and frequent images.
- Possible to get a complete picture of traffic, road status, and vegetation.

Nano satellites?