Pilot Project: Geospatial Construction Starts

June 19, 2019

Nicolas Martinez
Elton Cryderman

Delivering insight through data, for a better Canada

Geospatial Survey of Construction Starts

- Purpose :
  - To develop a leading edge survey of housing starts using satellite imagery and machine learning/artificial intelligence (ML/AI)

- Justification:
  - Improve the accuracy and coverage of housing starts survey. Fill data gaps for smaller remote communities and indigenous communities which are generally excluded from the current survey programs
  - Reduce costs. CMHC currently employs a significant number of people at their head office, in addition field agents across the country, physically confirming residential starts and completions. Using remote sensing could significantly reduce this cost once developed
  - Data from this pilot project, if successful, could feed into numerous other programs

Schedule

- Publish request for proposal – May 2019
- Award contract and begin work on “proof of concept” – July 2019
- Begin work on “proof of concept”, including Machine Learning/Artificial Intelligence testing – August/September 2019
- Evaluate pilot project (measures of success) and decide on whether to move project into second phase, or full production – Winter 2019
- Continue with phased approach or expand data collection (January 2020 and beyond)

Deliverables

- Images from three regions will be gathered to meet various criteria of the pilot project:
  - A small to mid-sized CMA (Kitchener-Cambridge-Waterloo)
  - A CA (Red Deer)
  - A remote community (Iqaluit)

- Images will be preprocessed so that they can be used in a ML/AI data processing model
- A training set of images for processing data will be developed
- Construction of a “Convolutionary Neural Network” (CNN) that will be able to identify buildings and the start of construction

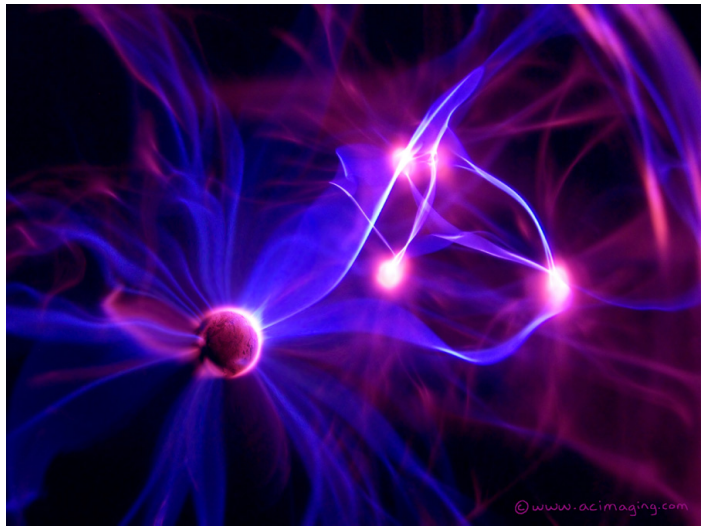
- End of project:
  - Using geospatial data, assess cost, efficiency, and accuracy of estimating building starts in various regions using ML/AI, and provide a written report recommending next steps

Construction of CNN

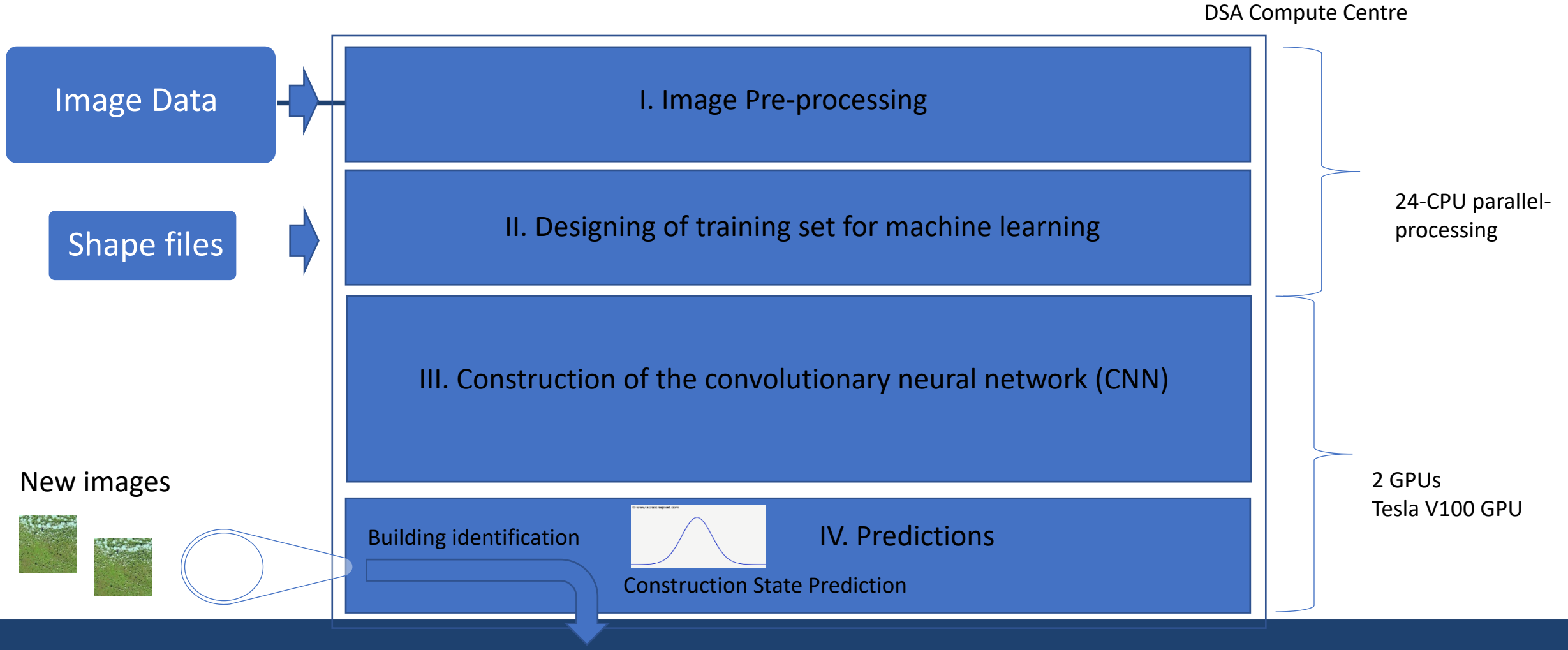
- Evaluation of different NN architectures: pre-trained vs self-trained models
  - Use pre-trained approach

- Use open-source software Tensorflow and Keras to build NNs
- Object tagging combined with construction state identification

- Hyper-parameter tuning
  - Learning rate schedule
  - Loss function
  - Optimizer



Overview: AI Setup for Construction Identification



Foundation

Characteristics

- Foundation walls are generally very bright and angular – stand out from surroundings
- Several stages
  - Footprints
  - Wall castings
  - Walls and floors complete
- Construction materials commonly seen in and around footprint

Superstructure (floors and walls)

Characteristics

- Wooden construction
- Walls and subfloors visible
- Pre-built segments in or around building

Roof under construction

Characteristics

- Rafters and roof structures visible
- Roof substrate visible
- Pre-built segments sometimes visible around site

Roof completed, no landscaping

Characteristics

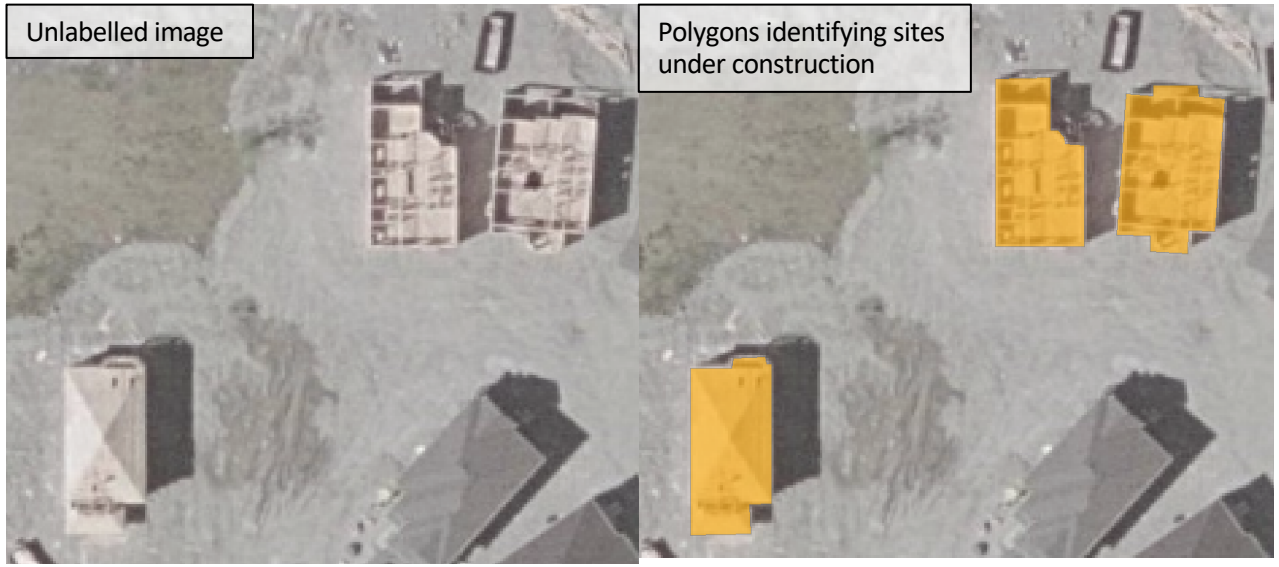
- Change of roof colour from light (substrate) to dark (singles)
- Sites often have construction material around
- Angular roof features and shadows from peaks

Completed building and Landscaping

Characteristics

- Change of site from bare to vegetated
- No construction materials near/around site
- Signs of occupation (cars in driveway, patio furniture, etc.)

Example GIS polygons and attributes



Attributes

- Building Type
- Phase
- Phase Detail / Sub-phase
- Date (Imagery/Observed)
- Notes
- Area

More attribute entries likely to be added later

	OBJECTID *	SHAPE *	Bldg_Type	Phase	SubPhase	ImageDate	ImageSource	Notes	SHAPE_Area
	3	Polygon	Single	Superstructure	Roof - Substrate Complete	2005	City of Ottawa		164.676279
	4	Polygon	Single	Superstructure	Walls - In Progress	2005	City of Ottawa		159.576575
	5	Polygon	Single	Superstructure	Walls - In Progress	2005	City of Ottawa		182.571779

Table 1 – Sample attribute table for building polygons above (Numbered left to right)

Measures of Success

- Cost
  - Demonstrate through the pilot project that building starts could be estimated at a cost equal to or less than existing methods
- Timeliness
  - Demonstrate through the pilot project that monthly building starts can be estimated using geospatial images and ML/AI as per existing program timelines
- Quality
  - Successful generation of a building shape file for all buildings within the selected test areas
  - Successful development of training images for machine learning/AI process
  - Successful change detection for the start of construction with an acceptable level of error (at a “human level”):
    - False positives: <5%
    - False Negatives: <5%
    - Valid Interpretation of Polygons: >95%

STATISTICS CANADA THANKS YOU!

Elton Cryderman – (613) 878-0323

elton.cryderman@canada.ca

Nicolas Martinez – (613) 302-9291

nicolas.martinez@canada.ca