INL/CON-20-60136-Revision-0

Critical Infrastructure Classification via CNN-based Modeling and Image Analysis

October 2020

Ashley Jeanette Brockman Shields, Shiloh Elliott

Changing the World's Energy Future

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

INL/CON-20-60136-Revision-0

Critical Infrastructure Classification via CNN-based Modeling and Image Analysis

Ashley Jeanette Brockman Shields, Shiloh Elliott

October 2020

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

October 2020

Ashley Shields & Shiloh Elliott

Modeling and Simulation Scientist

Infrastructure Assurance & Analysis

National & Homeland Security

Critical Infrastructure Classification via CNN-based Modeling and Image Analysis

Idaho National Laboratory

Critical Infrastructure

- *"Critical infrastructure describes the physical and cyber systems and assets that are so vital to the United States that their incapacity or destruction would have a debilitating impact on our physical or economic security or public health or safety."* https://www.dhs.gov/topic/critical-infrastructure-security

- A few examples…
  - Nuclear Power Plants
  - Water Treatment Facilities
  - Hydroelectric Plants
  - Airports

https://www.energy.gov/ne/nuclear-reactor-technologies

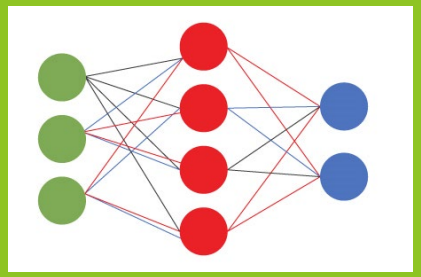
Project Overview

- Implement computer vision techniques to analyze imagery and return a list of included features

Input: Image

https://www.portlandoregon.gov/bes/article/40645

Assessment: Run the Model (CNN)



https://acadgild.com/blog/convolutional-neural-network-cnn

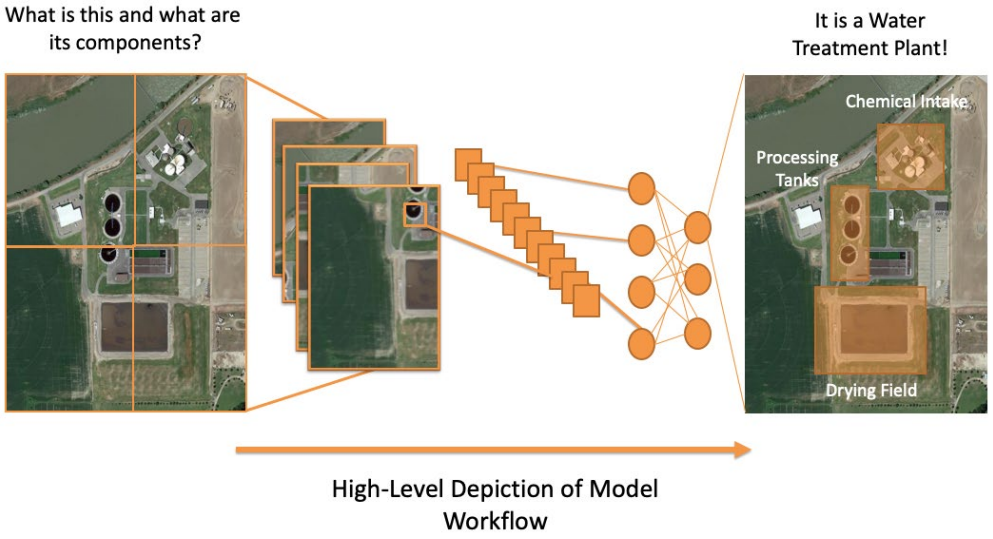
Output: Feature Labels

QAQC: Explainability

What's our current status

- Densenet-161 and ResNeXt-101 Prototypes
  - Can differentiate between airports and water treatment plants
    - Expand to include additional labels
- Now we need to…
  - Make some choices
    - Architecture
  - Expand the model
    - Additional facility labels
    - Inclusion of sublabels
  - Evaluate model performance
    - Error catching
    - Explainability
    - Verification and Validation
  - Adapt the model for other datasets
    - Transfer learning

https://www.usgs.gov/media/images/central-wastewater-treatment-plant-nashville



High-Level Depiction of Model Workflow

Choices
Architecture

Model Expansion
Additional labels and sublabels

Quality
Error catching, verification and validation, explainability

Adaptability
Transfer Learning

Choices: Architecture

- ResNeXt
  - Behaves like ResNet, but with improved accuracy
  - Performs well with imagery
  - Skip connections: allow for the efficient use of deep networks
- DenseNet
  - Developed for visual object recognition
  - Requires fewer parameters
  - *Functional Map of the World*

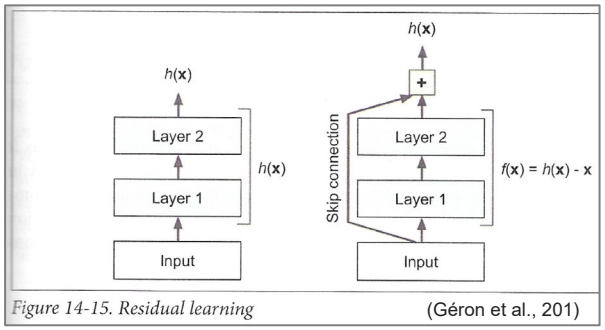
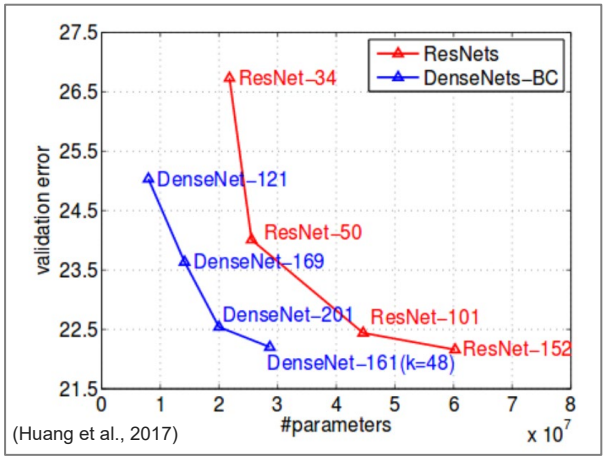
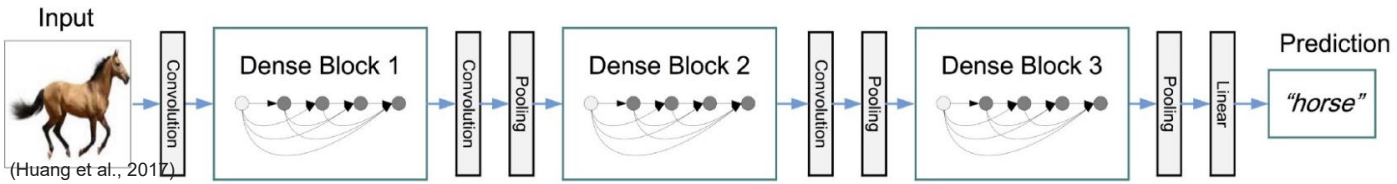


Figure 14-15. Residual learning (Géron et al., 201)

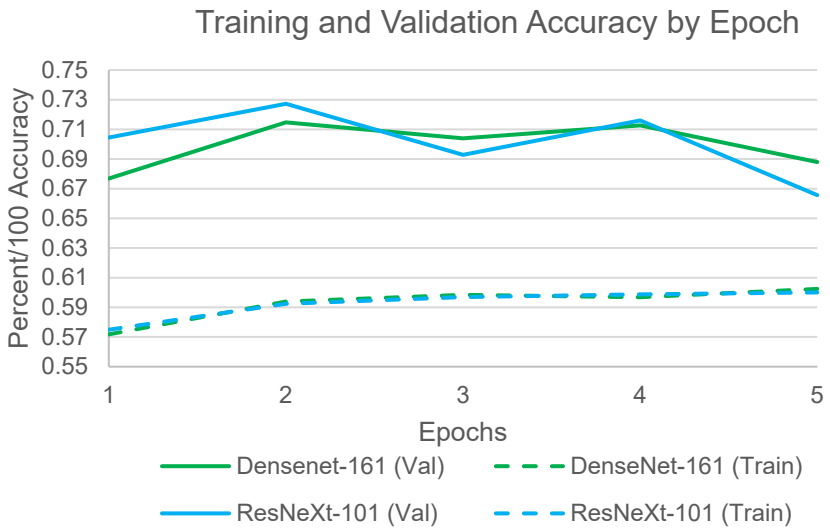


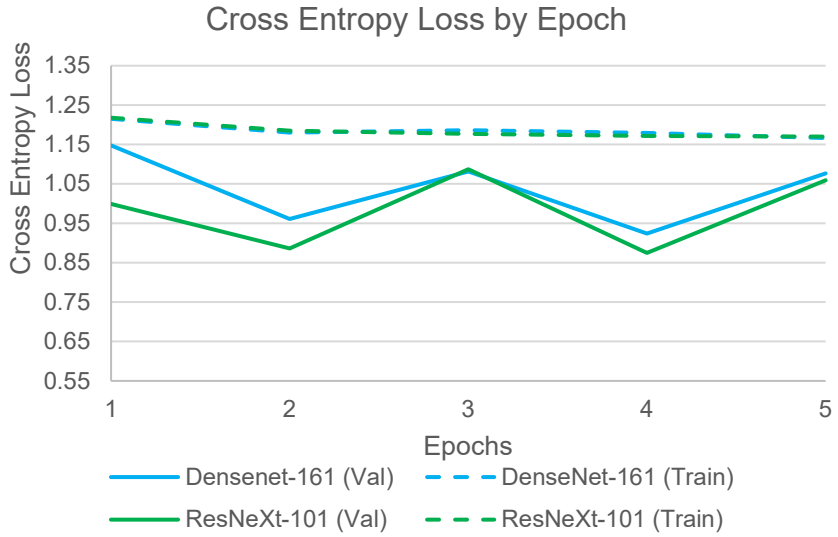
(Huang et al., 2017)



(Huang et al., 2017)

Choices:
Preliminary Architecture Comparison





DenseNet-161 (Percent)	Airport	Dam	Factory or PowerPlant	Hospital	Solar Farm	Water Treatment Facility
Airport	0.998	0.002	0.000	0.000	0.000	0.000
Dam	0.033	0.818	0.007	0.022	0.024	0.095
Factory or PowerPlant	0.036	0.142	0.239	0.333	0.095	0.156
Hospital	0.036	0.014	0.026	0.787	0.062	0.075
Solar Farm	0.006	0.013	0.020	0.064	0.817	0.080
Water Treatment Facility	0.013	0.082	0.031	0.052	0.088	0.733

ResNeXt-101 (Percent)	Airport	Dam	Factory or PowerPlant	Hospital	Solar Farm	Water Treatment Facility
Airport	0.987	0.004	0.000	0.004	0.000	0.004
Dam	0.006	0.753	0.021	0.024	0.027	0.169
Factory or PowerPlant	0.010	0.116	0.301	0.345	0.085	0.144
Hospital	0.002	0.009	0.016	0.053	0.665	0.067
Solar Farm	0.002	0.011	0.020	0.066	0.818	0.083
Water Treatment Facility	0.002	0.066	0.039	0.059	0.064	0.770

(Rows = True Labels, Columns = Predictions, Units = Percent/100)

Model Expansion: Multilabel Semantic Segmentation

- Classifying data into categories and subcategories
  - Category = Water Treatment Plant
  - Subcategory = Water Tank
  - How should we assess sub-features?
    - Defining characteristics?
    - Dataset development method?
    - Prodigy as a tool for segmentation?
  - Which classification structures ‘play nicely’ with the model?
  - At what stage are sub-features assessed within the model?
    - Assessed in the primary CNN?
    - Nested classification structure?

It is a Water Treatment Plant!

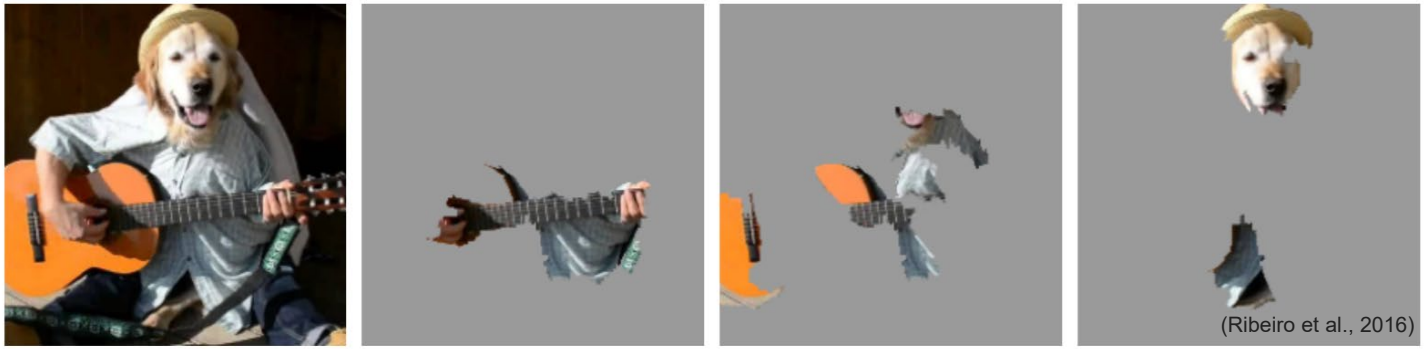
Quality Control: Error Catching, Verification, and Validation

- What if there is no Critical Infrastructure (CI) in an image?
- What if there are multiple CI types in an image?
- How does the model handle feature variability?
  - E.g. Small vs. large hydroelectric plants
- Why did the model make the choice that it did?
  - Explainability!
- How accurate do we want to be? Is there a floor?

https://landslides.photoshelter.com/image/I0000IrFCOLPEkf4

Quality Control: Explainability

- Explainability: additional scripts to explain why a model came to a decision
  - provides a ‘gut check’ step to ensure that the model is focusing on the correct features
- Can the model be trusted?
  - And why? Why not?
- Why did the model make the choices that it did?
- Where are the problem areas?

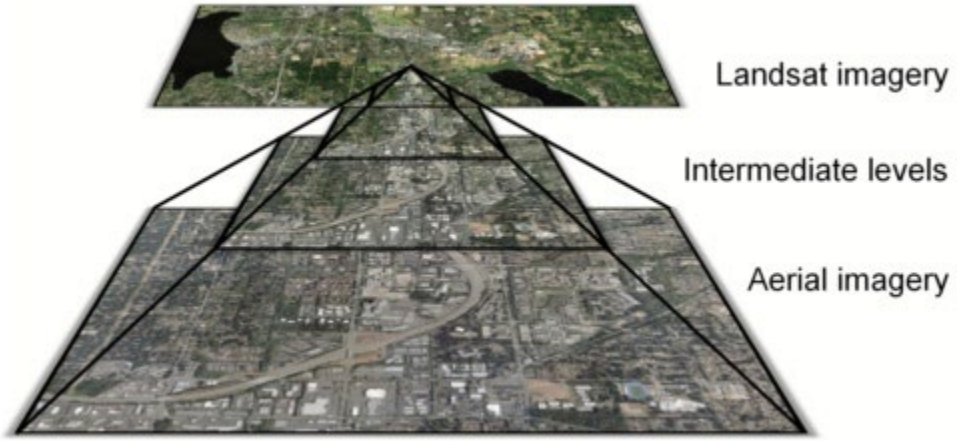


(a) Original Image (b) Explaining *Electric guitar* (c) Explaining *Acoustic guitar* (d) Explaining *Labrador*

Adaptability: Transfer Learning

- How can we take what a model has learned and apply it to a dataset with different characteristics?

- Extrapolate information from the current model so that it can be applied to alternative datasets
  - Bypass development//implementation of a new model



http://hhoppe.com/proj/pyramid/

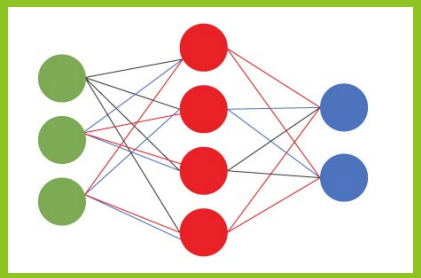
References

- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).

Input: Image

https://www.portlandoregon.gov/bes/article/40645

Assessment: Run the Model (CNN)



https://acadgild.com/blog/convolutional-neural-network-cnn

Output: Feature Labels

QAQC: Explainability