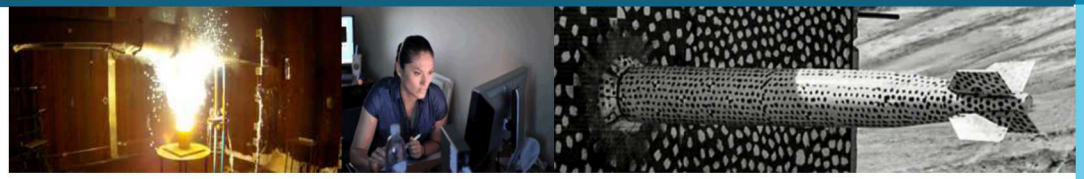
SAND2019-9268PE

Spatio-Temporal Anomaly Detection in Video



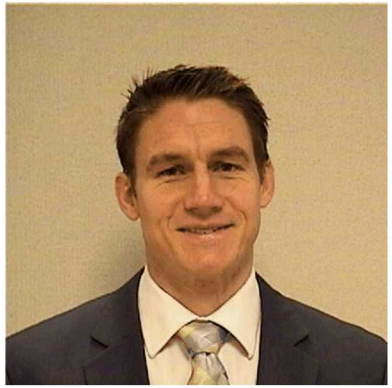
Presented by: Michael R. Smith (5852)

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

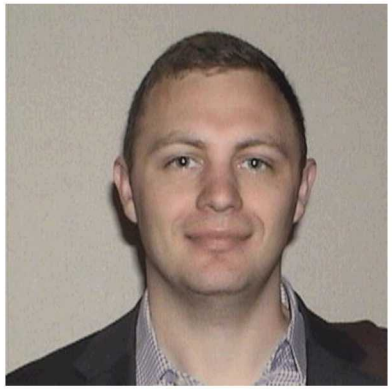
SAND2019-XXXX

Team Members

Moe:
Josh Rutkowski
6832



Larry:
Mike Smith
5852



Curly:
Mike Hamel
6831

The NPT:
Treaty on the Non-Proliferation of Nuclear Weapons

"The NPT aims to prevent the spread of nuclear weapons and weapons technology, to foster the peaceful uses of nuclear energy, and to further the goal of disarmament. The Treaty **establishes a safeguards system under the responsibility of the IAEA**, which also plays a central role under the Treaty in areas of technology transfer for peaceful purposes."*

Date of entry into force: 5 March 1970

Number of States Parties: 191**

* https://www.iaea.org/publications/documents/treaties/npt
** http://disarmament.un.org/treaties/t/npt

IAEA Department of Safeguards

The Department of Safeguards independently **verifies States legal obligations of the peaceful use of nuclear materials and technology**.

The verification process includes:

- Onsite inspections
- State declaration verification
- Open source information analysis
- **Containment and surveillance**
  - **including camera surveillance**

IAEA
International Atomic Energy Agency
Atoms for Peace and Development

https://www.iaea.org/about/business

Next Generation Surveillance System Cameras

The NGSS is a camera -based surveillance system with specific measures implemented to ensure technical reliability and security *

https://www.iaea.org/newscenter/news/surveying-safeguarded-material-24/7

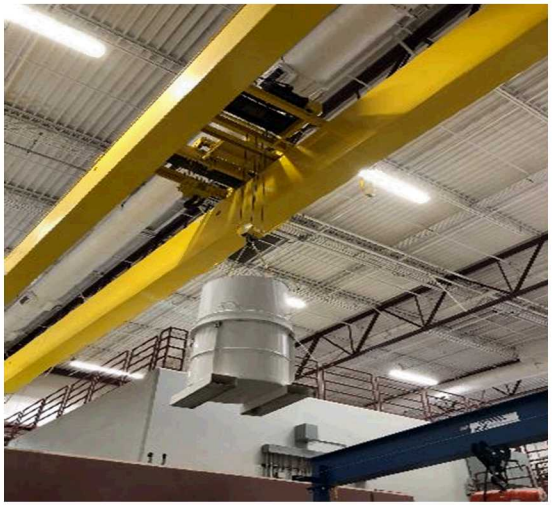
https://www.iaea.org/newscenter/multimedia/photoessays/faces-of-safeguards

Over a million pieces of encrypted safeguards data are collected by over **1400 surveillance cameras**, and 400 radiation and other sensors around the world.**

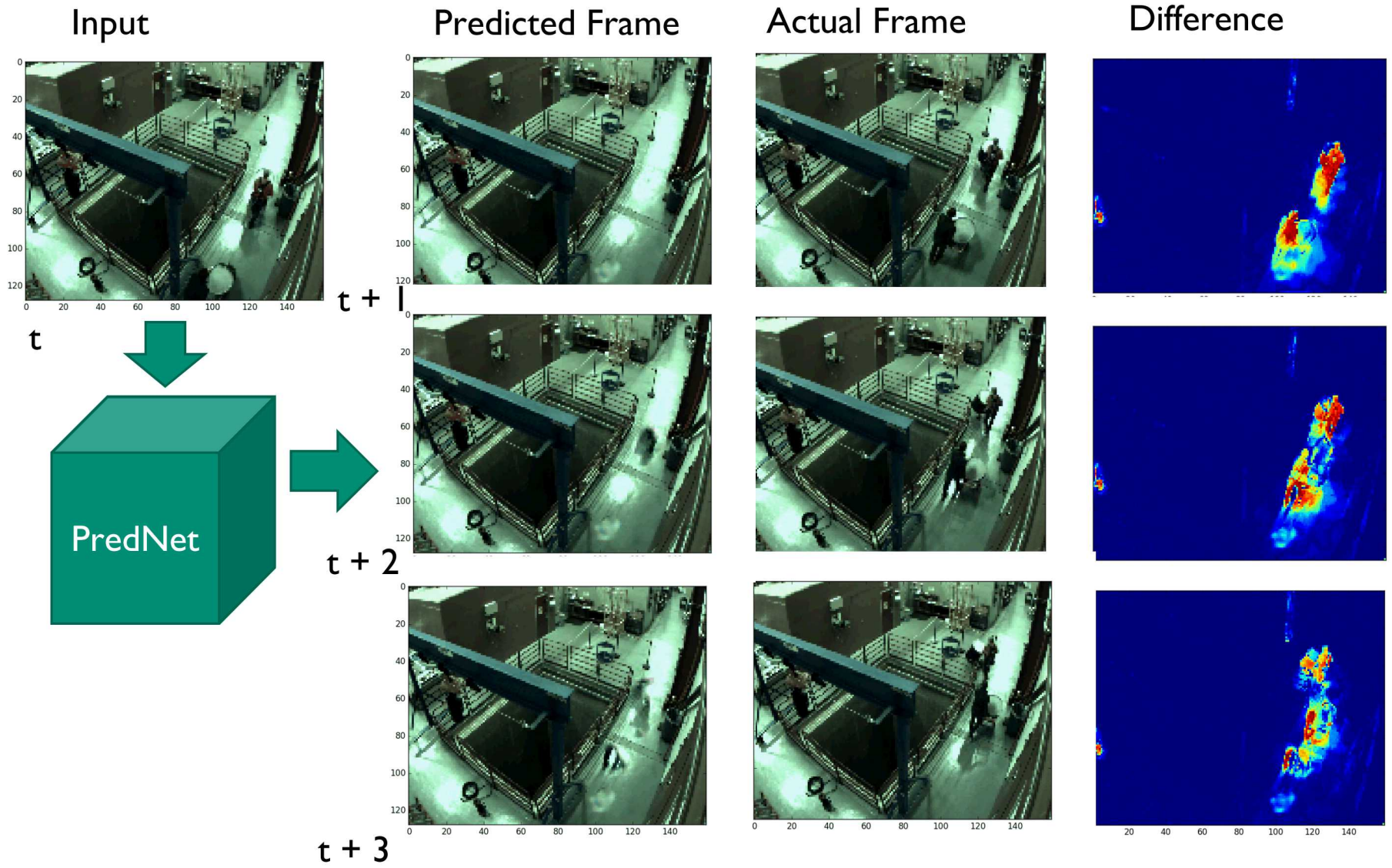
* https://www.osti.gov/servlets/purl/1116295
** https://www.iaea.org/newscenter/news/surveying-safeguarded-material-24/7

Problem Space and Use Case

- Review of NGSS surveillance data by IAEA inspectors is **mundane and tedious**
  - Look for anomalous activity (unknown unknowns)
- Common situation is transfer of spent fuel to storage and transportation casks
- Sandia developed a proxy use-case to transfer a large (approx. 5ft. tall by 3 ft. wide) container into and out of a floor vault
- Assumptions:
  - No labelled training data (cannot enumerate all anomalies)
  - Data cannot leave facility
  - Non ML expert users
  - Environments and processes change significantly across facilities

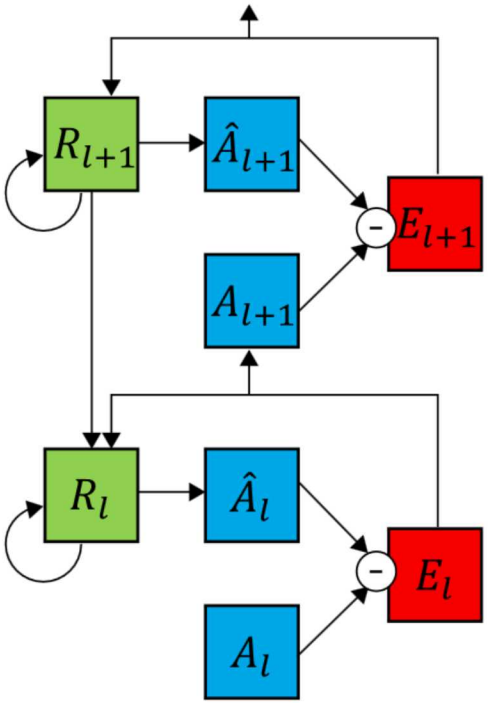


Solution: Deep Predictive Coding Networks for Video Prediction and Unsupervised Learning (PredNet)

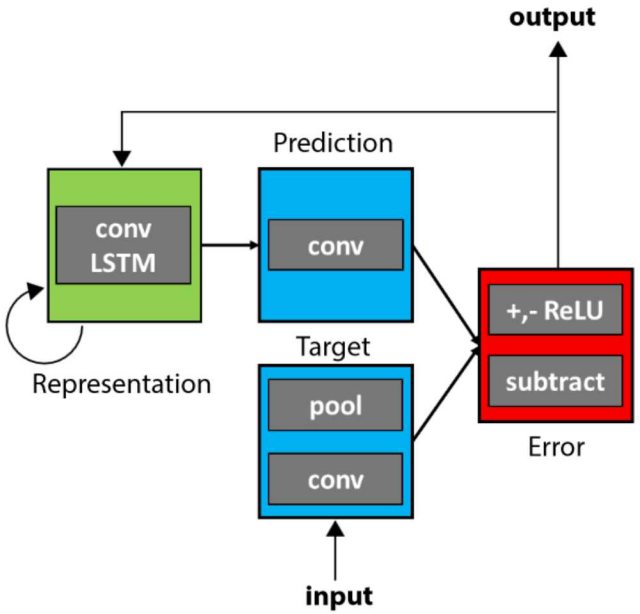


PredNet Architecture

- Each layer in PredNet consists of:
  - R_l: representation neurons
  - $\hat{A}_l$: layer-specific predictions at each time step
  - A_l: layer-specific target
  - E_l: layer-specific error term



- Information flow within 2 layers

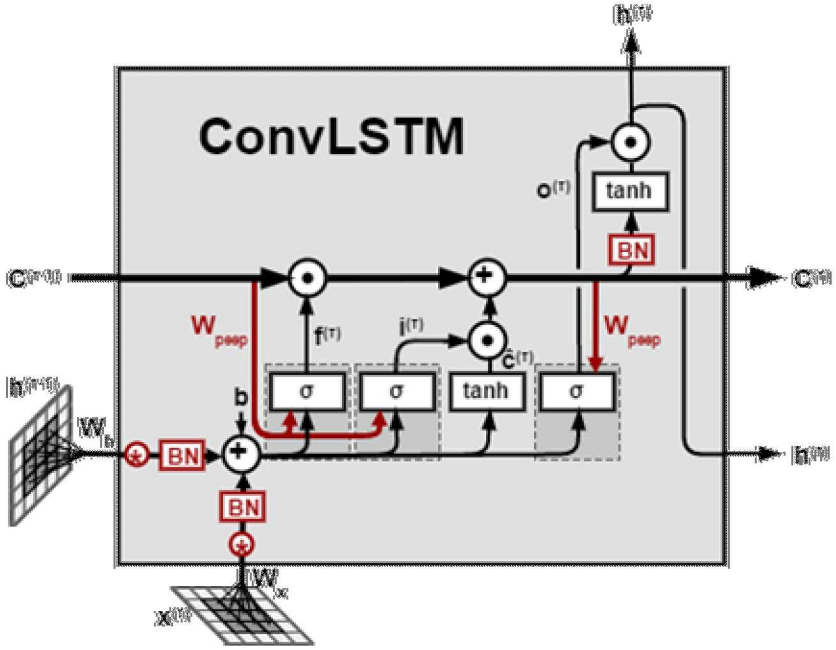


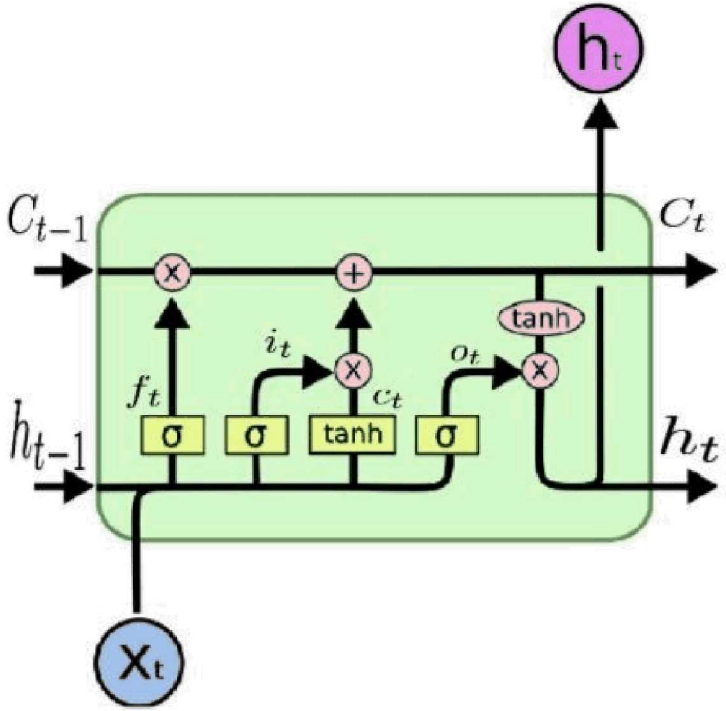
- Module operations

ConvLSTM -- Pictures

Models spatio-temporal relationships in the data

- Integration of CNN and LSTM
- Recurrent layer (like and LSTM)
- Internal standard matrix multiplications exchanged with convolution operations
- Retains multiple-dimension data (LSTM is one dimensional)





ConvLSTM -- MATH

LSTM

$$
\begin{aligned}
i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \\
f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \\
c_t &= f_t \circ c_{t-1} + i_t \circ \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\
o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \circ c_t + b_o) \\
h_t &= o_t \circ \tanh(c_t)
\end{aligned}
$$

ConvLSTM

$$
\begin{aligned}
i_t &= \sigma(W_{xi} * \mathcal{X}_t + W_{hi} * \mathcal{H}_{t-1} + W_{ci} \circ \mathcal{C}_{t-1} + b_i) \\
f_t &= \sigma(W_{xf} * \mathcal{X}_t + W_{hf} * \mathcal{H}_{t-1} + W_{cf} \circ \mathcal{C}_{t-1} + b_f) \\
\mathcal{C}_t &= f_t \circ \mathcal{C}_{t-1} + i_t \circ \tanh(W_{xc} * \mathcal{X}_t + W_{hc} * \mathcal{H}_{t-1} + b_c) \\
o_t &= \sigma(W_{xo} * \mathcal{X}_t + W_{ho} * \mathcal{H}_{t-1} + W_{co} \circ \mathcal{C}_t + b_o) \\
\mathcal{H}_t &= o_t \circ \tanh(\mathcal{C}_t)
\end{aligned}
$$

$*$ represents the convolution operator

Variables are capitalized in ClnvLSTM because they are 3D tensors

Calculating image differences

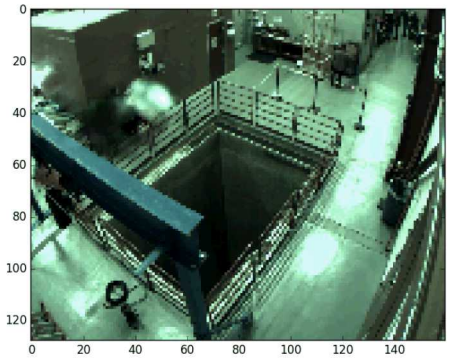
Compare **Predicted Image** to **Actual Image**

1. Convert both images to grayscale
2. Calculate Squared Error, E, for each pixel i

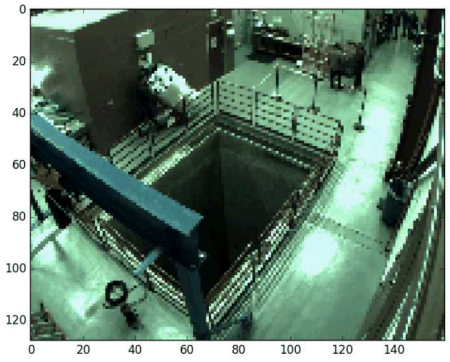
$P = Predicted\ Image$ $A = Actual\ Image$

$$E_i = (P_i - A_i)^2$$

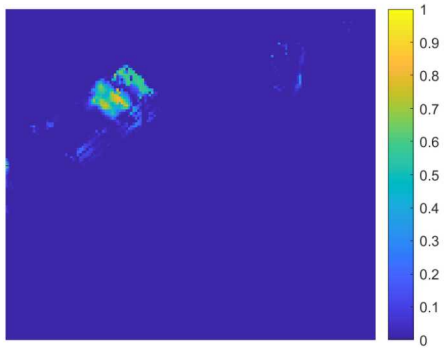
Predicted Image



Actual Image

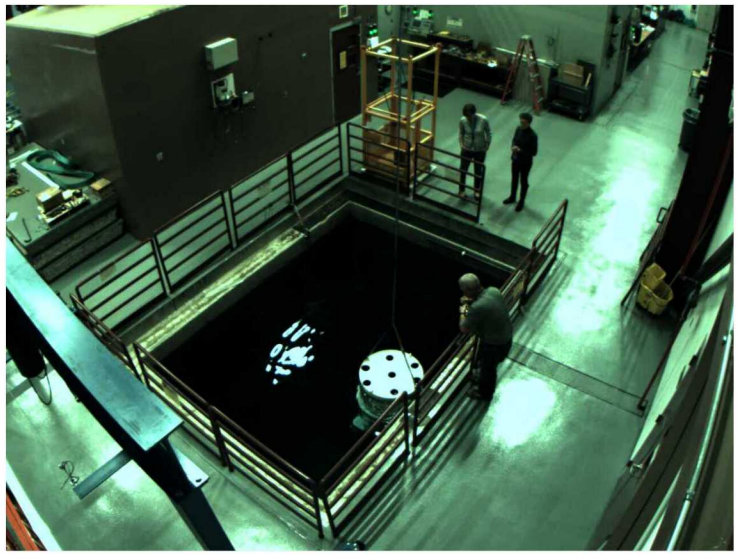


Squared Error Image



Data

- Sandia deployed two NGSS cameras in the Gamma Irradiation Facility (GIF)
- Collected down-time data and active scripted container movements over multiple days
- Collections include both full (water) and empty floor vault scenarios



Scenarios for Data Analysis Plan

Evaluate what the PredNet algorithm determines as "anomalous" and its relevance to safeguards

Test four categories of potentially anomalous scenarios:

1. **Unintentional Anomalies** – examine anomalies that are identified in "normal" operational scenarios
2. **Intentional Anomalies** – intentionally insert anomalous frames to determine algorithm response
3. **Operational Anomalies** – change operational activities within a facility, including types of containers present, appearance of containers, areas in which container are located
4. **Safeguards scenarios** – experiment with scenarios that are determined to be of high safeguards interest, e.g. greyscale images, longer time lapse, and play-back loops

Results

Experiment trained only on containers leaving the facility

Significantly larger irregulatity scores for containers entering the facility

Calculate Mean Squared Error for images in a series

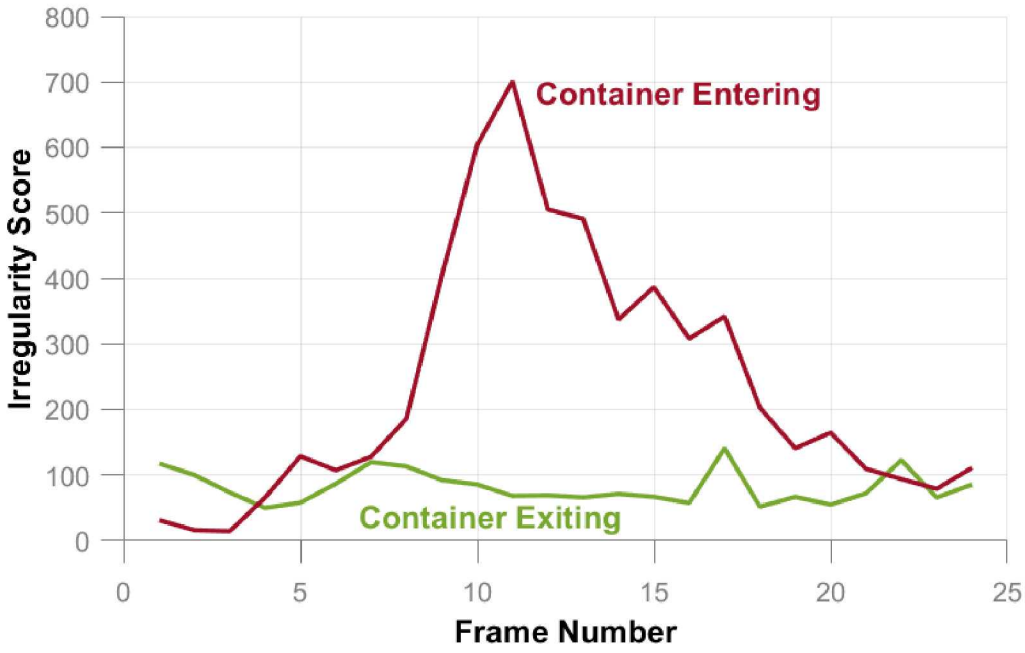
1. Convert both images to grayscale
2. Calculate Squared Error, E, for each pixel i

P = Pixel values from predicted image

A = Pixel values from actual image

N = Number of pixels

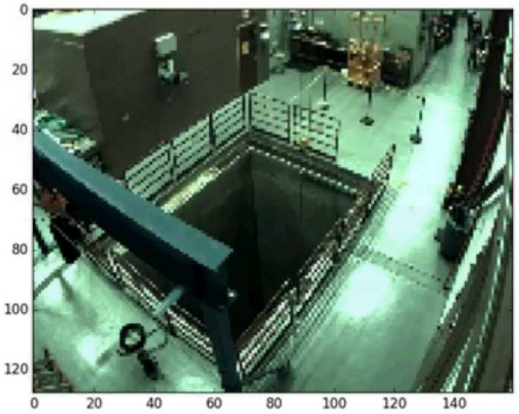
$$\frac{1}{N}\sum_{i=1}^{N}(P_i - A_i)^2$$



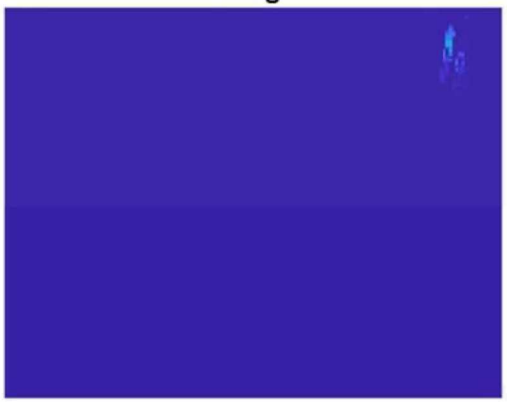
Video showing the sequence of containers entering and exiting the facility

Container Entering

Actual Image



Difference Between Predicted and Actual Images



Frame Number: 1

Conclusions and Future Work

PredNet is a viable solution for detecting spatio-temporal anomalies

- Does not require labelled data (which can be time consuming and labor intensive)
- Does not require (potentially sensitive) data to leave given facilities
- Demonstration of detection of normal objects and people doing anomalous activities

Cons:

- Time consuming (in computational time) to train (but alleviates human burden)

Future work

- Examine PredNet on more extensive analyses
  - What does PredNet detect in day to day activities
  - Does PredNet overly detect anomalies?
- Extend to work with supervised approaches
  - Anomalous activities near objects of interest
  - Can the supervised and unsupervised share weights?