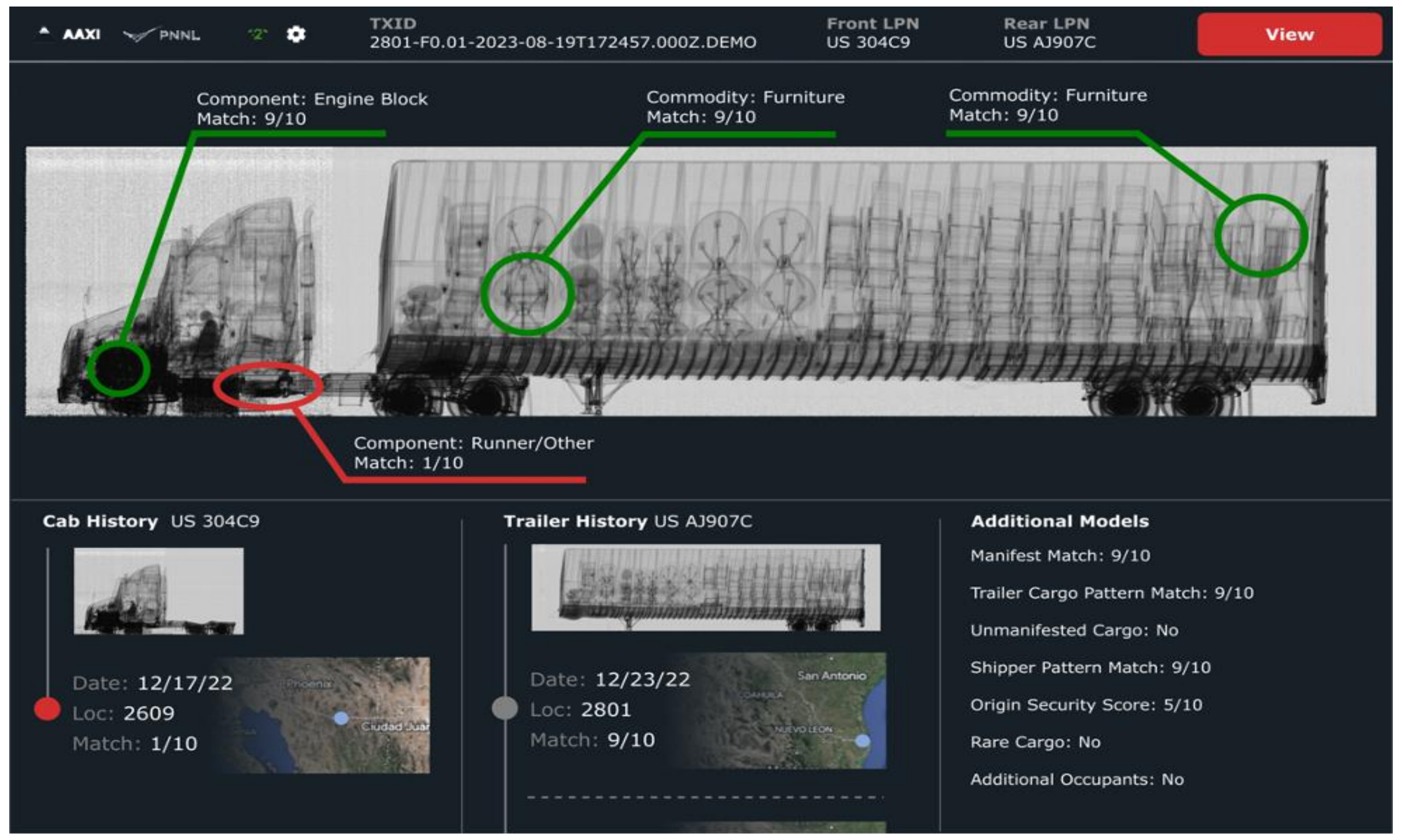
Advanced Analytics in X-ray Images

Benny Bellmar, Giovanni De Geronimo, Vinay Gupta, Thomas Lu, Brice Rider

PURDUE UNIVERSITY | The Data Mine | PNNL

INTRODUCTION



Project Motivation:

- Secure the borders to protect public health and safety.

Project Goal:

- Detect contraband before it enters the USA using X-ray analytics and machine learning.

Abstract:

- In this project, we were given x-rays of variable size and tasked with creating a tool that would find contraband both hidden or in plain sight. The backbone of this project is an autoencoder which enables unsupervised learning of image features. The comparison are made by reducing the autoencoder's latent representations using a vision transformer and using hierarchal navigable small words to find vectors similar to the reduced latent representation.

WORKFLOW

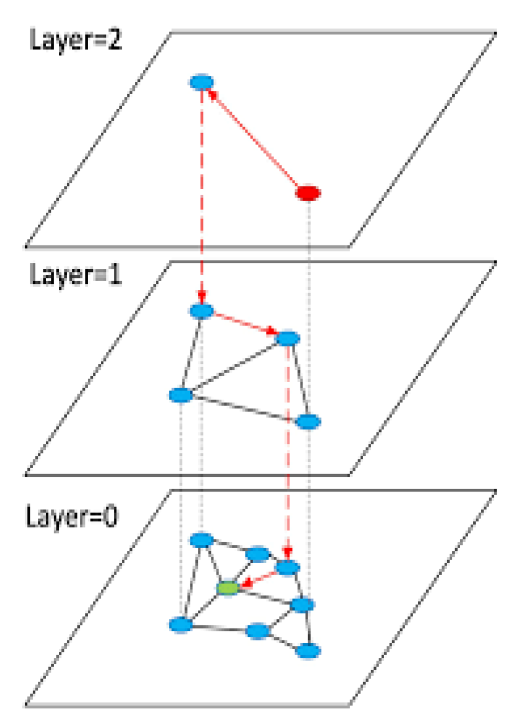
Research into autoencoders and convolutional neural networks. → Created a flexible autoencoder with easily swappable reduction operations. → Tested various reduction operations that reduce differently sized vectors to the same size. → Researched how to add invariance and equivariances into the model pipeline. → Explored distance metrics and decided to use HNSW for nearest neighbors operations. → Experimented with a vision transformer as the reduction operation. Determining decision logic for program.

Fall 2023 Semester | Spring 2024 Semester

AUTOENCODER

- Original Image — Takes in an image of any size.
- Convolution Encoder — Extract information about the image's features and splits it into latent representations.
- Latent Representation — Each latent representation is portion of the original image focusing on different features.
- Reduction Operation — Reduces data in latent representations to a fixed amount regardless of starting image.
- Reduced Latent Representation — Data that can be used to conduct object comparison for the sub-image.
- Convolutional Decoder — Model attempts to reassemble original images from latent representations.
- Reconstructed Image — Reconstructed image is compared with original.

NEAREST NEIGHBORS AND DISTANCE METRICS



- To compare the reduced representation vectors that have high dimensionality, we use a measure of similarity, cosine distance.
- We utilize a method called Hierarchal Navigable Small Worlds (HNSW) to organize our database of vectors in layers.
- When finding the nearest neighbors of a new vector, HNSW starts in the top layer and descends following the path that decreases the distance most.
- This lets us efficiently add a new vector to the existing database and find its nearest neighbors.

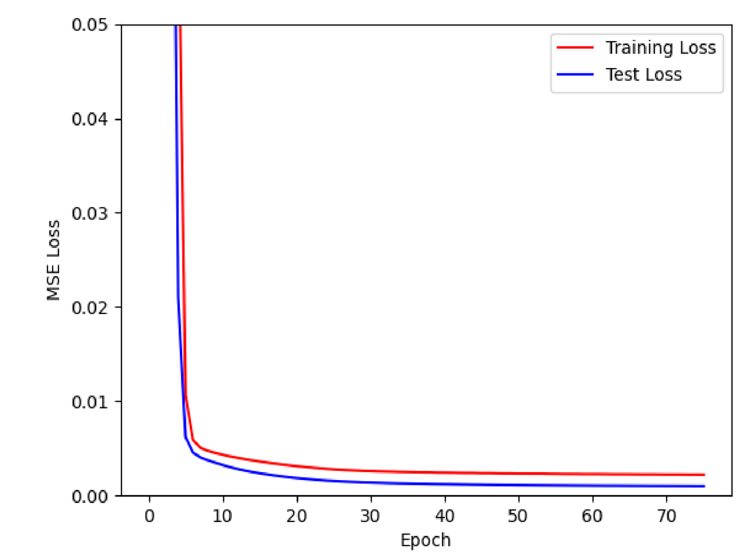
DECISION LOGIC

- The purpose of the decision logic is to determine whether a car has contraband.
- We use information from a vehicle's prior border crossing along with HNSW to check whether the current shape matches its last recorded shape.
- From there we use decision logic such as the diagram to the right.
- If the model flags a car, when it is inspected, the model receives feedback whether the car had contraband and can learn from it.
- There is still an issue of false negatives that do not flag vehicles, so they are not physically checked.

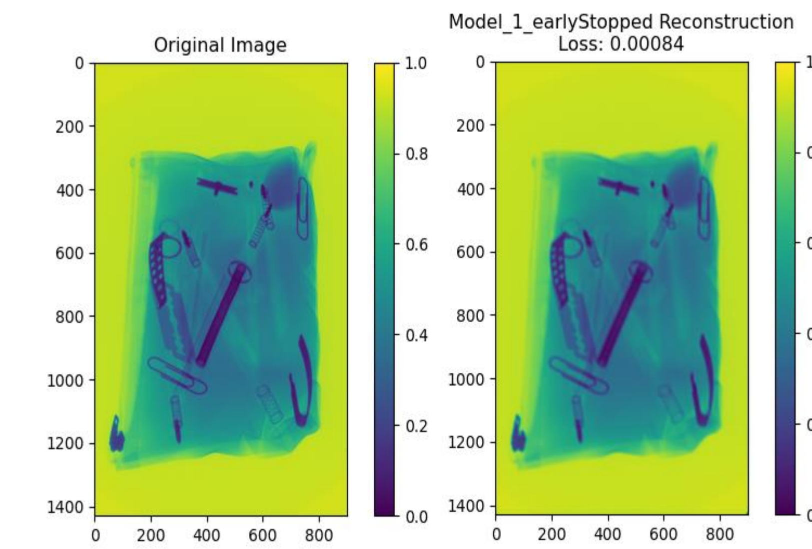
The vehicles nearest neighbor is itself in a previous crossing
- The past crossing was marked as suspicious → Flag the vehicle, inspect it, and add it to the database
- The past crossing was not marked as suspicious → Flag the vehicle, inspect it, and add it to the database

RESULTS

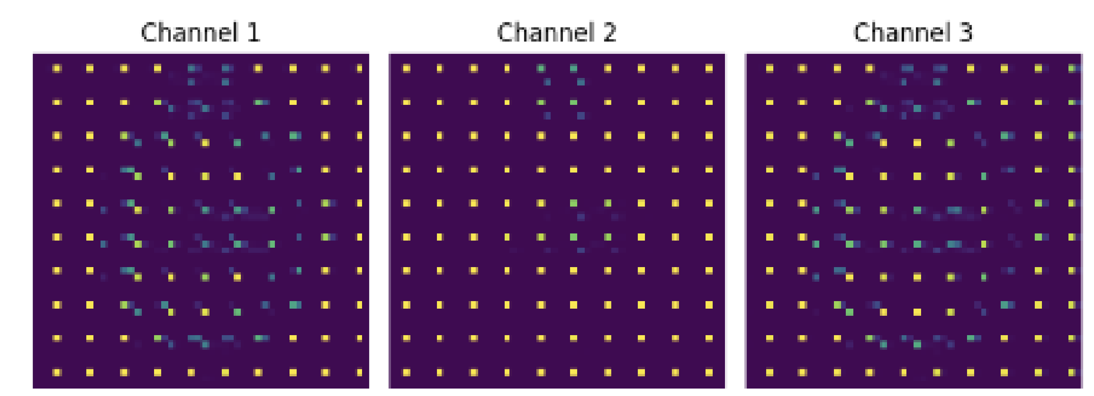
Autoencoder Loss Curve



Original vs. reconstruction



Patched histogram reduction operation



CONCLUSION

- The autoencoder was successful in that its latent vectors accurately captured the important details of the input image
- Reduction operations need some work as patched histogram was inefficient; looking into AdaptiveAvgPool2d
- The HNSW using cosine distance metric was effective in preliminary tests to find approximate nearest neighbors, but more advanced experimenting is necessary
- Decision logic to compare sets of images may be more effective than current idea
- Vision transformer may improve upon autoencoder

FUTURE GOALS AND ACKNOWLEDGEMENTS

Future Goals:

- Explore different reduction operations like patched histograms or vision transformers
- Try a variable stride autoencoder that doesn't need an additional reduction operation
- Explore the possibility of comparing sets of vehicles in the decision logic

Acknowledgements:

- Alexander Hagen
- Rachel Pfeifer
- Lauren Dalder
- Kali Lacy

Dataset:

https://domingomery.ing.puc.cl/material/gdxray/

The Data Mine Corporate Partners Symposium 2024