**Slide 1: Title Slide**

**Anomaly Detection On Low Resolution Photography On The Martian Surface Using High Dimensional Tensor Decomposition**

**Practicum Final Report**

**Anthony Chan**

**July, 2024**

*Script:* Welcome everyone, and thank you for joining me today. I am Anthony Chan, and this presentation covers my practicum final report on "Anomaly Detection on Low Resolution Photography on the Martian Surface Using High Dimensional Tensor Decomposition." This study is created to fulfill the requirement from Sandia National Laboratories, the project sponsor, to find a methodology for using Machine Learning or Computer Vision to Signal Analysis to automatically predict signal degradation

Let's dive in.

**Slide 2: Executive Summary**

* **High-dimensional tensor decomposition**: Tucker & CP (PARAFAC)
* **Goal**: Reduce dimension of image sequences into signals for anomaly detection
* **Best Performance**: Tucker decomposition with autoencoder (58% accuracy)

*Script:* To begin, I’ll provide an executive summary of the study. The primary focus was on using high-dimensional tensor decomposition techniques, specifically Tucker and CP (PARAFAC) decompositions, to reduce the dimensionality of image sequences into signals for training anomaly detection models. The study found that Tucker decomposition combined with autoencoder neural networks achieved the highest performance, identifying anomalies with 58% accuracy.

**Slide 5: Introduction - Background**

* **Importance**: Anomaly detection in space exploration, medical imaging, industrial inspection, etc.
* **Dataset**: Images from NASA's Mars Science Laboratory (MSL) Curiosity rover
* **Subset Used**: 200 images for training, 50 images for testing

*Script:* Moving on to the background, detecting novel or anomalous objects in images is crucial in various domains such as space exploration, medical imaging, and industrial inspection. This study used a dataset of images captured by NASA's Mars Science Laboratory (MSL) on the Curiosity rover. Due to computational constraints, a subset of 200 images for training and 50 images for testing was used.

**Slide 6: Introduction - Objectives**

1. **Evaluate Decomposition Methods**: CP vs. Tucker
2. **Assess Anomaly Detection Models**: OC-SVM, autoencoders, random forest, combined approaches
3. **Determine Optimal Dimensions**: Find ranks that maximize performance
4. **Performance Metrics Analysis**: Accuracy and other metrics

*Script:* The main objectives of the study were to evaluate the effectiveness of CP and Tucker decomposition techniques, assess the performance of different anomaly detection models including OC-SVM, autoencoders, random forest, and an ensemble model of OC-SVM and autoencoder to determine the optimal decomposition ranks for maximizing performance, and analyze performance metrics such as prediction accuracy.

**Slide 7: Methodology - Data Collection**

* **Dataset Source**: Kerner et al.'s research
* **Subset**: 200 training images, 50 test images
* **Image Sets**: 6 images per set, 64x64 resolution
* **Representative Sample**: Ensures efficient processing on a mid-performance laptop

*Script:* For data collection, the image dataset was sourced from Kerner and company’s research. The subset used consisted of 200 training images and 50 test images, with each set containing 6 images at a 64x64 resolution. This representative sample allowed for efficient processing on a mid-performance laptop.

**Slide 8: Methodology - Image Examples**

* **Non-anomalous Image Sets**: Examples
* **Anomalous Image Sets**: Examples

*Script:* Here are some examples of the image sets used in the study. On the left, we have non-anomalous image sets, and on the right, we have anomalous image sets. These examples illustrate the types of images analyzed for anomaly detection.

**Slide 9: Methodology - Signal Transformation**

* **Image Decomposition**: Transforms images into low-dimensional 2D signals
* **Example**: Non-anomalous image sets decomposed into signals

*Script:* The image decomposition process transforms the images into low-dimensional 2D signals. The example shown here depicts how non-anomalous image sets are decomposed into signals, which can then be analyzed by the anomaly detection models.

**Slide 10: Experiments**

* **Two Categories**:
  1. CP Decomposition
  2. Tucker Decomposition
* **Flatten Data**: Into 2D arrays for model training
* **Models Used**: OC-SVM, autoencoder, random forest, combined autoencoder + OC-SVM

*Script:* The experiments were divided into two main categories: CP decomposition and Tucker decomposition. After decomposing the images, the data was flattened into 2D arrays for model training. The models used included OC-SVM, autoencoder, random forest, and an ensemble autoencoder + OC-SVM approach.

**Slide 11: CP Decomposition Results**

* **OC-SVM**: 44% accuracy, rank 80
* **Autoencoder**: 46% accuracy, rank 85
* **Random Forest**: 38% accuracy, rank 10
* **Autoencoder + OC-SVM**: 56% accuracy, rank 35

*Script:* Here are the results for CP decomposition. The OC-SVM model achieved 44% accuracy with a decomposition rank of 80. The autoencoder reached 46% accuracy with a rank of 85. The random forest model had 38% accuracy with a rank of 10, and the combined autoencoder + OC-SVM model achieved 56% accuracy with a rank of 35. The optimal rank for CP decomposition is found after testing each rank from 5 to 100, with increments of 5 for the highest accuracy.

**Slide 12: Tucker Decomposition Results**

* **OC-SVM**: 56% accuracy, rank (65, 35, 65)
* **Autoencoder**: 58% accuracy, rank (95, 65, 65)
* **Random Forest**: 46% accuracy, rank (5, 65, 5)
* **Autoencoder + OC-SVM**: 48% accuracy, rank (65, 35, 5)

*Script:* For Tucker decomposition, the OC-SVM model achieved 56% accuracy with a decomposition rank of (65, 35, 65). The autoencoder achieved the highest accuracy of 58% with a rank of (95, 65, 65). The random forest model had 46% accuracy with a rank of (5, 65, 5), and the combined autoencoder + OC-SVM model achieved 48% accuracy with a rank of (65, 35, 5).

The optimal rank for Tucker’s decomposition is found after testing each combination rank of 3-tuple values from 5 to 95 with increments of 5 for the highest accuracy.

**Slide 13: Results Summary**

* **CP Decomposition**: Best accuracy - 56% (Autoencoder + OC-SVM)
* **Tucker Decomposition**: Best accuracy - 58% (Autoencoder)
* **Performance Comparison**: Tucker outperformed CP in most cases
* **Best Overall Model**: Tucker decomposition with autoencoder

*Script:* Summarizing the results, the CP decomposition methods showed varying degrees of success, with the highest accuracy of 56% achieved by the combined autoencoder + OC-SVM approach. Tucker decomposition methods generally outperformed CP decomposition, with the best accuracy of 58% achieved by the autoencoder model. Overall, Tucker decomposition combined with autoencoder was the best-performing model.

**Slide 14: Comparative Analysis**

* **Variability**: Consistent performance across multiple runs despite hardware limitations
* **Overall Performance**: Tucker decomposition superior to CP
* **Best Model**: Tucker with autoencoder (58% accuracy)
* **Optimal Ranks**: (95, 65, 65) for Tucker, 35 for CP with autoencoder + OC-SVM

*Script:* In the comparative analysis, it was observed that the performance was consistent across multiple runs, despite hardware limitations. Tucker decomposition methods demonstrated superior performance compared to CP decomposition methods. The best model overall was Tucker with autoencoder, achieving 58% accuracy. The optimal ranks identified were (95, 65, 65) for Tucker decomposition and 35 for CP decomposition with autoencoder + OC-SVM.

**Slide 15: Conclusion**

* **High-dimensional tensor decomposition**: Effective for anomaly detection in image set and HD signals.
* **Best Technique**: Tucker decomposition with autoencoder
* **Future Work**: Explore more ranks, neural network architectures, and hybrid methods
* **Applicability**: Techniques generalizable to various domains

*Script:* In conclusion, this study demonstrates the effectiveness of high-dimensional tensor decomposition methods, particularly Tucker decomposition, in improving anomaly detection in image datasets and more generally high dimensional signals. The best-performing technique was Tucker decomposition with autoencoder. Future work should explore additional decomposition ranks, neural network architectures, and hybrid methods to further enhance performance. The methodologies developed in this study are generalizable to various domains beyond