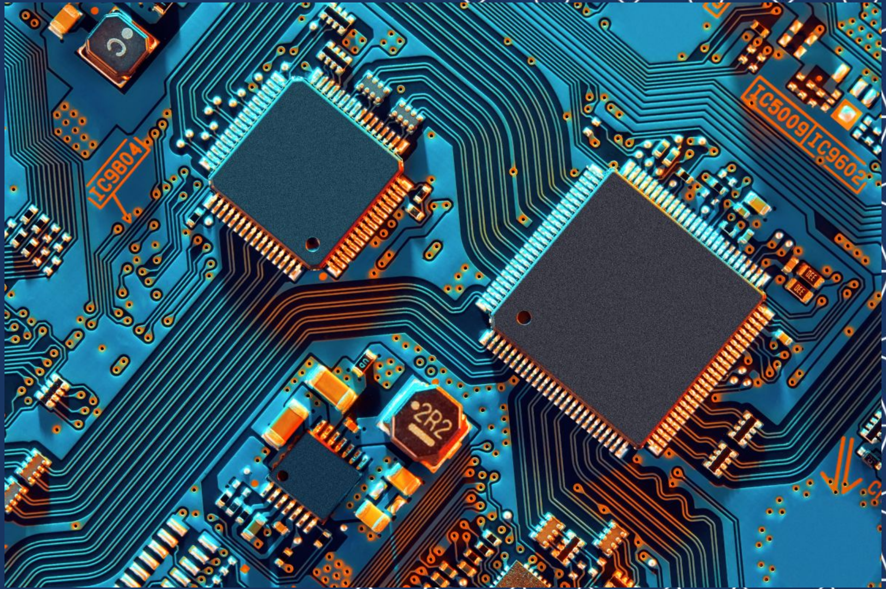
Automatic Wafer Defect Detection



Oluwafemi Adelegan

Jorge Bonilla

Berkeley
UNIVERSITY OF CALIFORNIA

Problem Statement:

Variations in the wafer fabrication process can lead to defects that may render an individual product useless or even necessitate the disposal of an entire batch. So, high-volume manufacturing processes require rapid and precise identification of products with surface defects. Since humans are subject to error, reducing their involvement in defect identification in manufacturing is crucial.

Our objective is to identify defective production items using images taken in a controlled industrial setting. This process of identification involves the use of 10 distinct machine learning algorithms on both defective and non-defective images.

Wafer Dataset

The dataset consists of :

- 246 images with visible defects
- 2085 images without any defect
- image sizes of approximately 230 x 630 pixels
- training dataset with 190 positive and 1558 negative images prior to augmentation
- test dataset with 56 positive and 527 negative images prior to augmentation
- several different types of defects (scratches, surface imperfections, etc.)

Image without defect

Images with defect

Block Diagram

Problem Statement

Using an image dataset of labeled images of defected items, apply machine learning algorithms to train models to accurately detect which production items have surface defects

Data cleaning and pre-processing

- Load labeled images
- Principal Component Analysis for image compression
- Perform data analysis with augmentation and without augmentation to compare the results
- Data augmentation includes:
  - Normalize images by diving the number of pixels by 255
  - Flattening images
  - Resizing the images to remove any unnecessary pixels such as white color around the image.
  - Apply the SMOTE procedure to augment the data with synthetic images to bring balance to the dataset.
- Split the data into training, testing, and validation sets

Modeling Approaches

1. Logistic Regression
2. Decision Trees
3. Random Forest Classifier
4. K-Nearest Neighbors
5. Naive Bayes Classifier
6. Support Vector Machines
7. Gradient Boosting Classifier
8. Convolutional Neural Networks
9. ResNet Transfer Learning
10. XCeption Transfer Learning

Evaluation Metrics

1. Train and Test Accuracy
2. Precision
3. Recall
4. AUC
5. F1 Scores

Model Selection

Compare models and choose appropriate model based on industry standards in manufacturing.

Ethical Concerns

Manage expectations and be transparent with the business on how this automation can impact current workflows

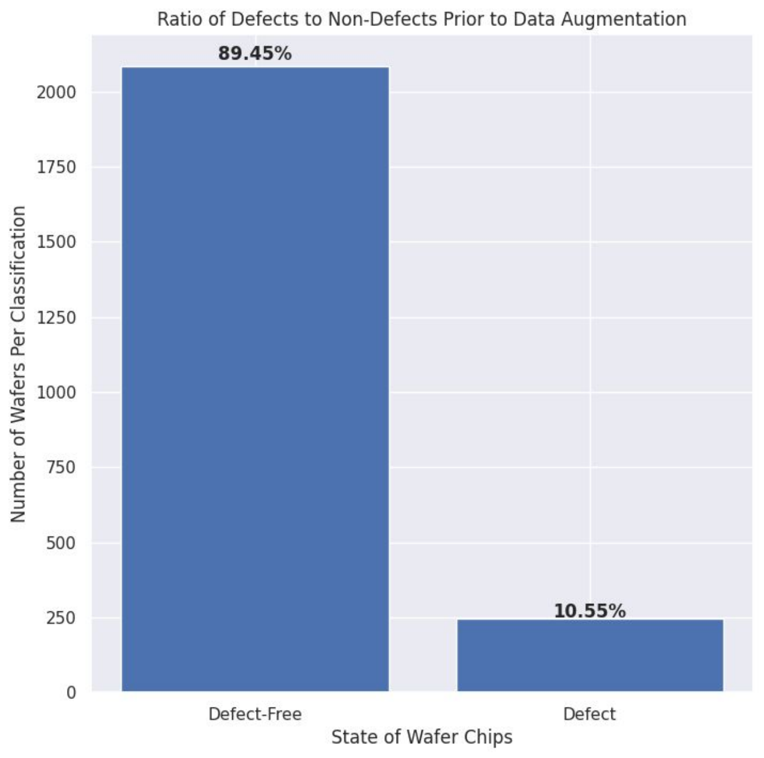
Business Impact

Discuss the business impact in terms of cost savings and process improvements

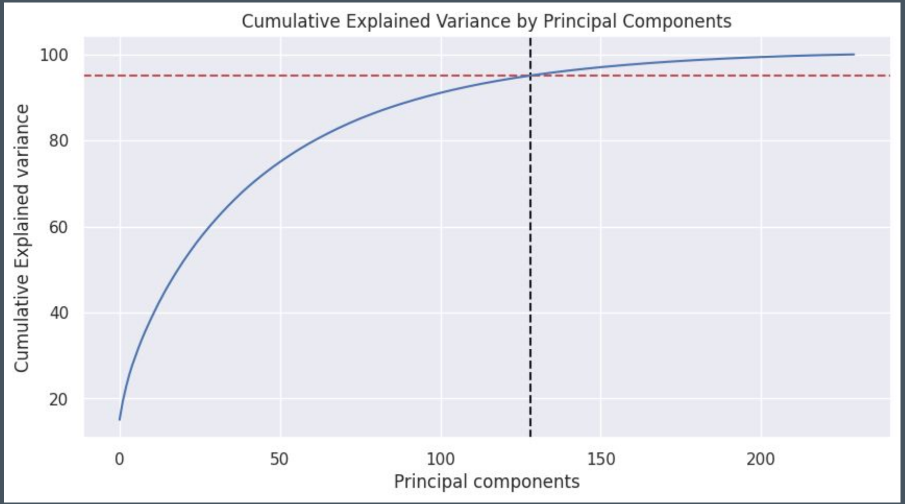
Berkeley
UNIVERSITY OF CALIFORNIA

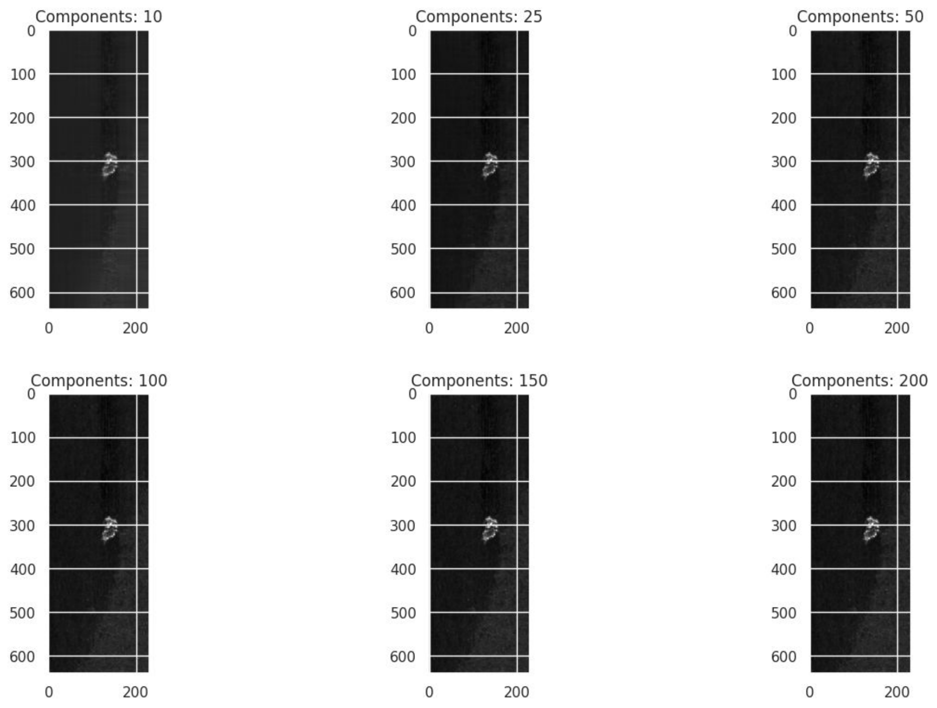
Methods for Imbalanced Dataset

- Adjust brightness and contrast
- Random Image Flips
- SMOTE Augmentation



Principal Component Analysis





Approach/Methodology

Logistic Regression (Base Model)	Decision Trees
K-Nearest Neighbors	Support Vector Machines
Naive Bayes	Convolutional Neural Networks
Gradient Boosting	Random Forest
ResNet50 Transfer Learning	Xception Transfer Learning

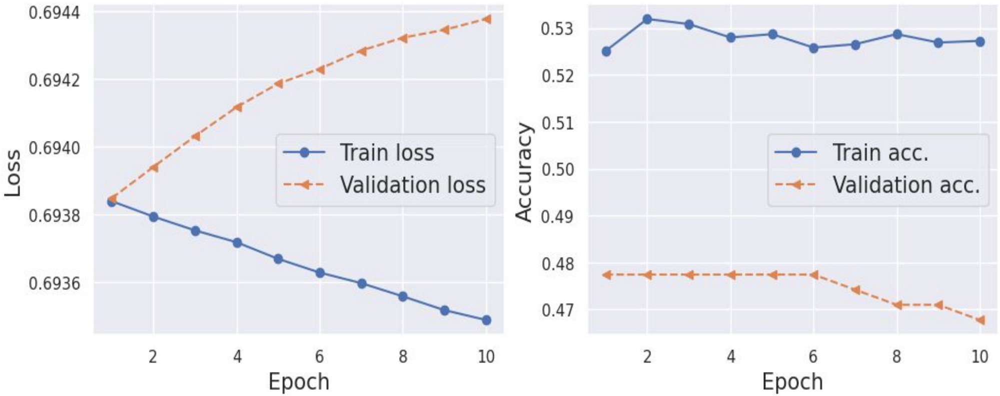
Logistic Regression

No Data Augmentation

- Train Accuracy: 89.44%
- Test Accuracy: 9.3%

Data Augmentation

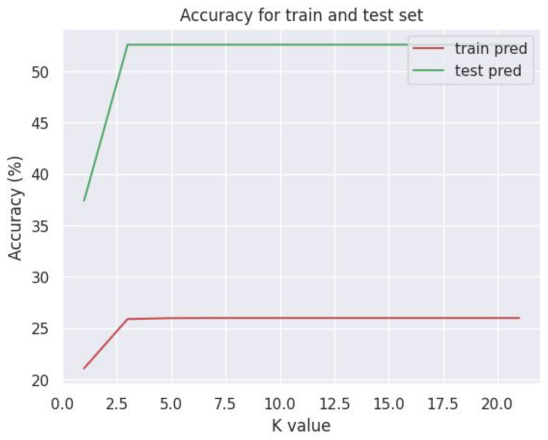
- Train Accuracy: 52.72%
- Test Accuracy: 7.9%



K-Nearest Neighbors

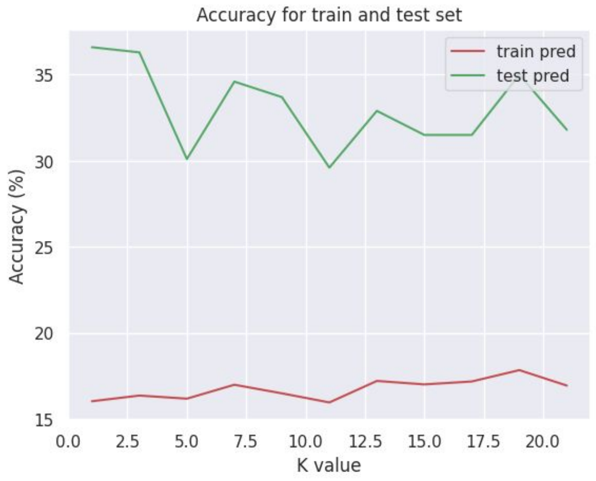
No Data Augmentation

- K-Value at 5
- Train Accuracy: 25.44%
- Test Accuracy: 51.0%



Data Augmentation

- K-Value at 5
- Train Accuracy: 15.0%
- Test Accuracy: 30.0%



Support Vector Machine (SVM)

No Data Augmentation

- Training Accuracy: 92%
- Test Accuracy: 89%
- Precision: 90%
- Recall: 1.0
- F1 Score: 0.95

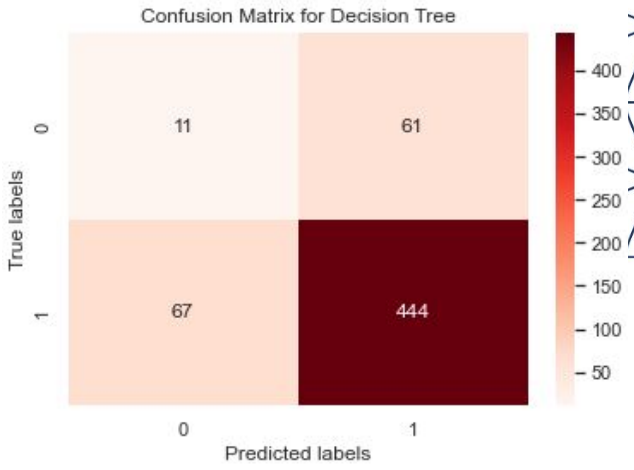
Data Augmentation

- Training Accuracy: 94
- Test Accuracy: 90%
- Precision: 90%
- Recall: 1.0
- F1 Score: 0.94

Decision Tree

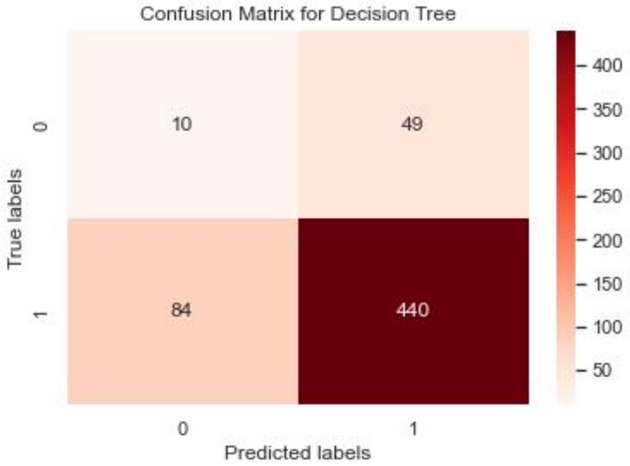
No Data Augmentation

- Train Accuracy: 98%
- Test Accuracy: 80%
- Precision: 90%
- Recall: 0.89
- F1 Score: 0.89



Data Augmentation

- Train Accuracy: 98%
- Test Accuracy: 78%
- Precision: 90%
- Recall: 0.84
- F1 Score: 0.86



Random Forest

No Data Augmentation

- Train Accuracy: 99%
- Test Accuracy: 90%
- Precision: 92%
- Recall: 0.99
- F1 Score: 0.95

Data Augmentation

- Train Accuracy: 99%
- Test Accuracy: 89%
- Precision: 92%
- Recall: 0.99
- F1 Score: 0.94

Naive Bayes Classification

No Data Augmentation

- Train Accuracy: 80.66%
- Test Accuracy: 77%
- Precision: 90.50%
- Recall: 0.82
- F1 Score: 0.86

Data Augmentation

- Train Accuracy: 89.5%
- Test Accuracy: 90.7%
- Precision: 92.25%
- Recall: 1.0
- F1 Score: 0.96

Gradient Boosting Ensemble Method

No Data Augmentation

- Test Accuracy: 88.33%
- Precision: 89.86%
- Recall: 0.98
- F1 Score: 0.98

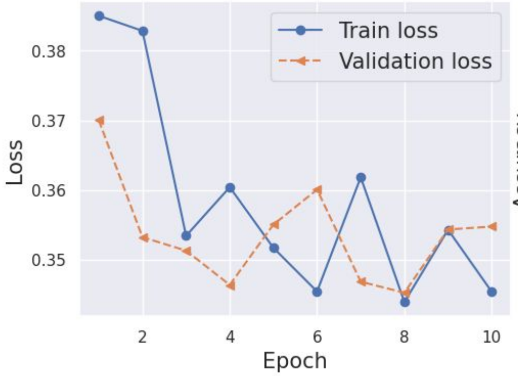
Data Augmentation

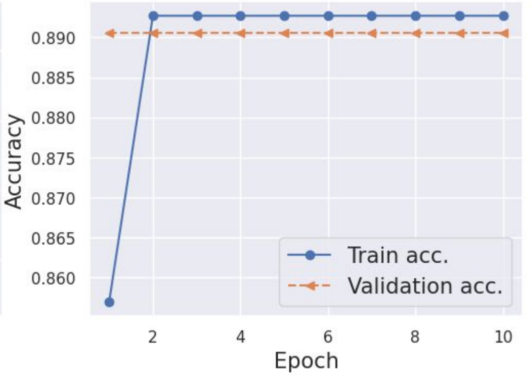
- Test Accuracy: 90.39%
- Precision: 90.34%
- Recall: 1.0
- F1 Score: 1.0

Convolutional Neural Networks

No Data Augmentation

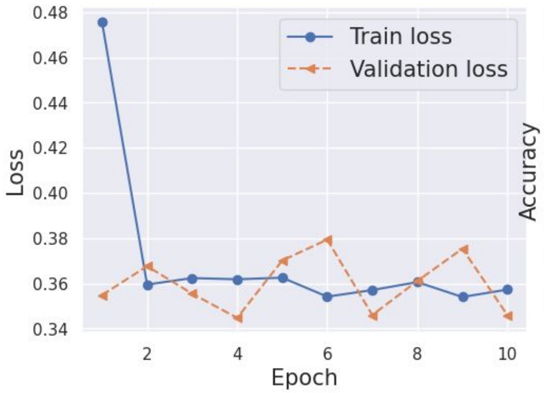
- Train Accuracy: 89.27%
- Precision: 89.27%
- Recall: 1.0
- Test Accuracy: 90.36%

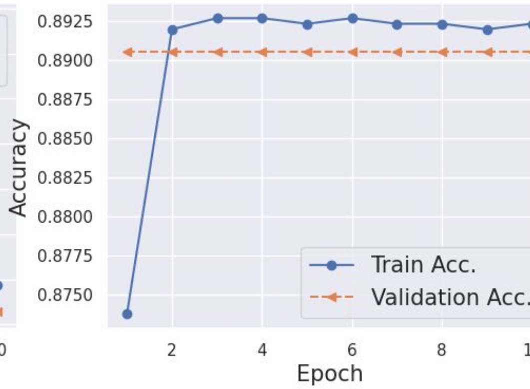




Data Augmentation

- Train Accuracy: 89.23%
- Precision: 89.26%
- Recall: 0.99
- Test Accuracy: 90.36%

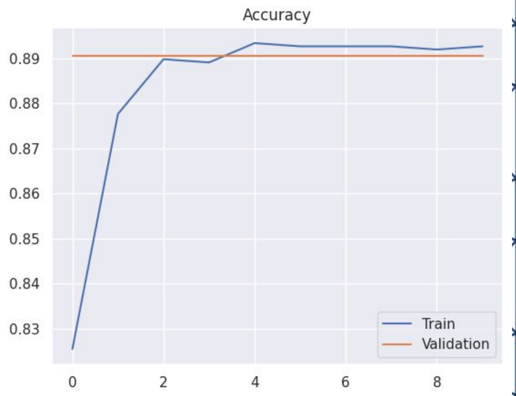




ResNet50 Transfer Learning

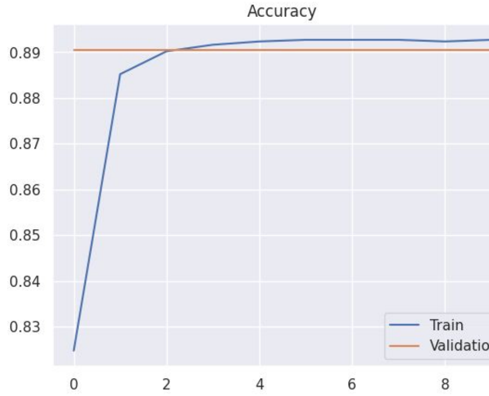
No Data Augmentation

- Train Accuracy: 89.27%
- Test Accuracy: 90.36%



Data Augmentation

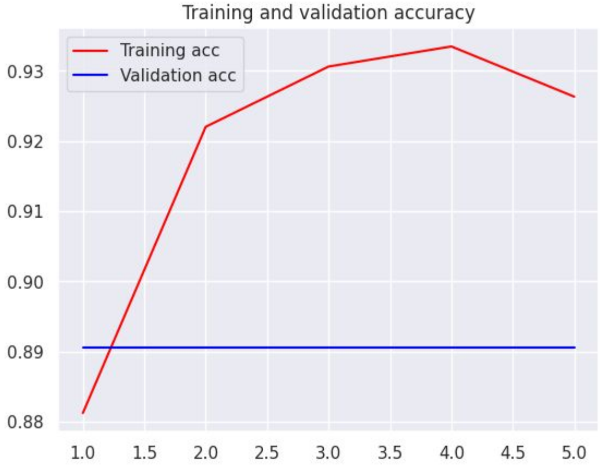
- Train Accuracy: 89.27%
- Test Accuracy: 90.36%

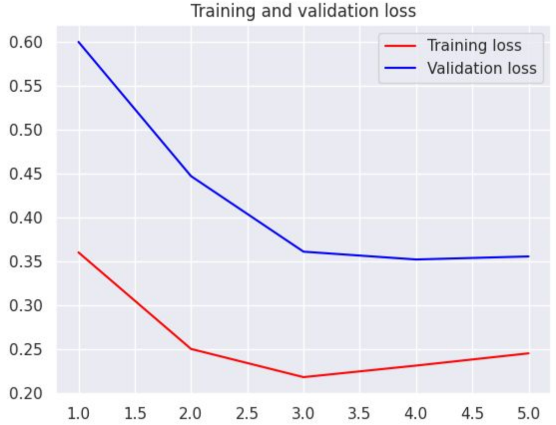


Xception Transfer Learning

No Data Augmentation

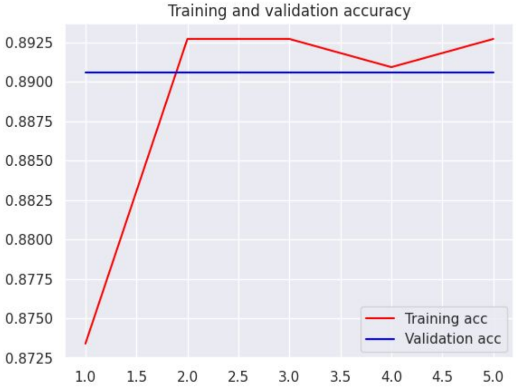
- Train Accuracy: 92.63%
- Test Accuracy: 90.36%

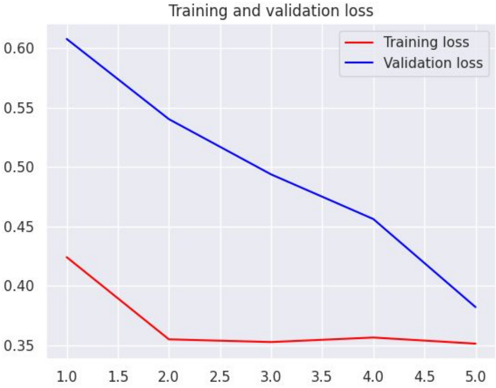




Data Augmentation

- Train Accuracy: 89.27%
- Test Accuracy: 90.36%

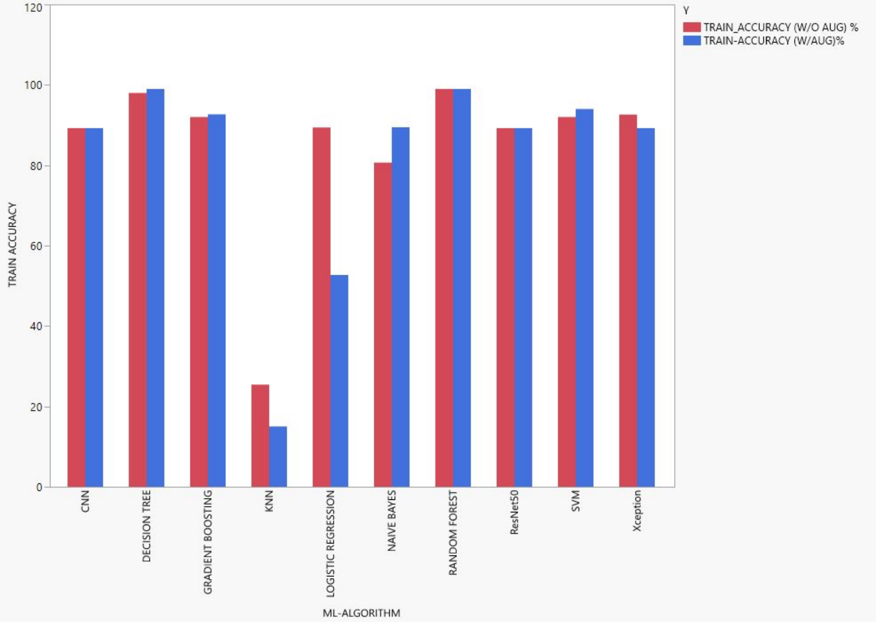


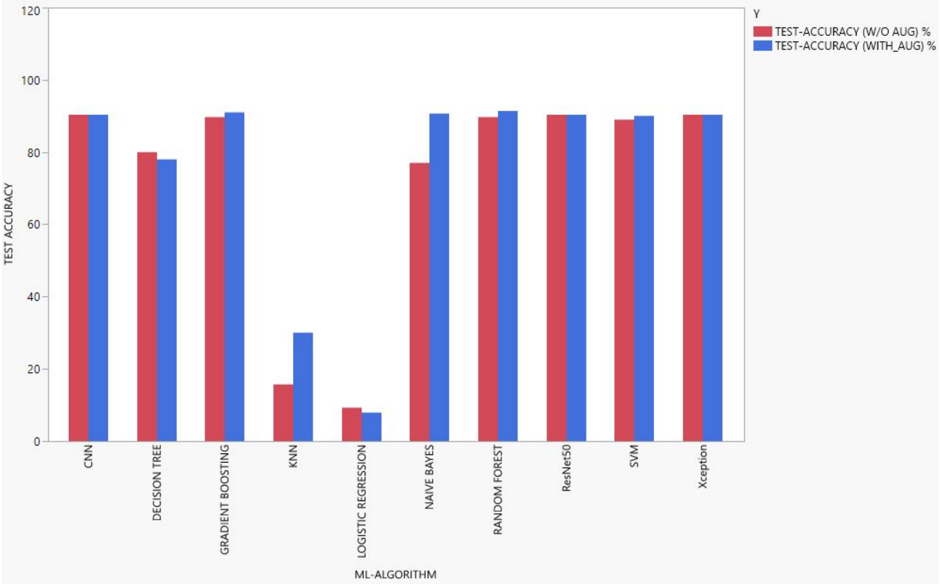


ML Algorithm		Train Acc(%)	Test Acc(%)	Precision (%)	Recall	F1- Score
Logistic Regression	W/O Aug	89.44	9.3	88.33	1.0	0.93
	W/ Aug	52.72	7.9	89.19	1.0	0.94
K-Nearest Neighbors	W/O Aug	25.44	51.0	0.88	0.06	0.11
	W/ Aug	15	30	0.88	1.0	0.93
SVM	W/O Aug	92	89	90	1.0	0.95
	W/ Aug	94	90	90	1.0	0.94
Decision Tree	W/O Aug	98	80	90	0.89	0.89
	W/ Aug	99	78	90	0.84	0.86
Random Forest	W/O Aug	99	90	92	0.99	0.95
	W/ Aug	99	89	92	0.99	0.94
Naive Bayes Classification	W/O Aug	80.66	77	90.5	0.82	0.86
	W/ Aug	89.5	90.7	92.25	1.0	0.96
Gradient Boosting Ensemble	W/O Aug	92	**88.33**	89.86	0.98	0.94
	W/ Aug	92.7	**90.39**	90.34	1.0	0.95
CNN	W/O Aug	89.27	**90.36**	89.27	1.0	0.94
	W/ Aug	89.23	**90.36**	89.26	0.99	0.99
ResNet50	W/O Aug	89.27	**90.36**			
	W/ Aug	89.27	**90.36**			
Xception	W/O Aug	92.63	**90.36**			
	W/ Aug	89.27	**90.36**			

Comparisons and Summary

Comparisons and Summary -





Bar charts showing both the train and test accuracies of the ML algorithms - With and without data augmentation

Limitations

- No way to tell the morphology of the defects
- Imbalanced dataset

Ethical Concerns

- Data privacy and security form a critical ethical consideration
- The potential displacement of human roles due to automation
- Issues of algorithmic bias and transparency
- Transparency and explainability of machine learning models

Future Work

- Wafers with multiple classes of defects
- Higher resolution images to better identify defects
- Test trained model on other types of industrial settings such as steel manufacturing
- Ensemble methods of transfer learning models.

References

[1]Božič, J., Tabernik, D., & Skočaj, D. (2021). Mixed supervision for surface-defect detection: From weakly to fully supervised learning. Computers in Industry, 129, 103459.

[2]Tao, Xian, et al. "Deep learning for unsupervised anomaly localization in industrial images: A survey." IEEE Transactions on Instrumentation and Measurement (2022).