From Pixels to Precision: Computer Vision for Defect Identification

AI-DetectifyLabs

Sharper Insights, Smarter Decisions

Charting Our Path Forward

Mission:

Our mission is to revolutionize manufacturing quality control through cutting-edge machine learning tools.

Goals:

- Positive impact on Industry Standards
- Cost-effectiveness
- Real-time Detection
- Accurate Anomaly Detection

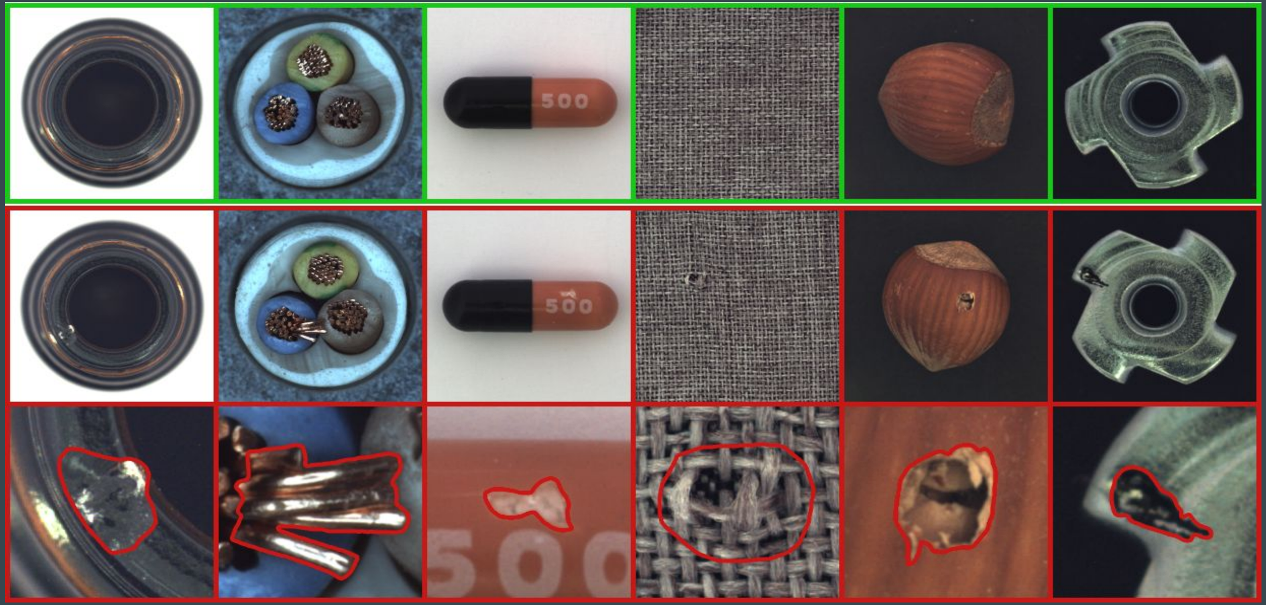
Business Problem

- Small and medium-scale businesses in Germany face challenges in inspecting their products due to resource limitations.
- The German manufacturing sector comprises over 220,000 enterprises, with the majority of them being leaders in niche product lines.
- All of these enterprises encounter quality control issues.

- The software is designed to offer a cost-effective solution targeting quality and reliability concerns.
- It enables small and medium-scale businesses to enhance their manufacturing processes.
- Products can meet required standards without the need for extra human resources or significant investments.

Dataset

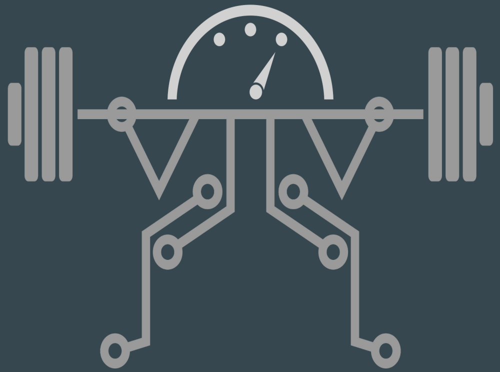
- Benchmark MVTec dataset for industrial defect detection
- Over 5000 high-res images across 15 categories
- Defect-free training images and imbalanced test sets with and without defects for each category



- Bottle
- Cable
- Capsule
- Carpet
- Grid
- Hazelnut
- Leather
- Metal Nut
- Pill
- Screw
- Tile
- Toothbrush
- Transistor
- Wood
- Zipper

Challenges in One-Class Classification with Imbalanced Test Data

- **Imbalanced Training Data**: Limitation to one class during training, hinders differentiation of positive cases.
- **Imbalanced Test Data**: Mix of good (negative) and defective (positive) samples may introduce bias.
- **Limited Positive Examples**: Insufficient positive exposure affects recognition.



- **Increased False Negatives**: Imbalance can lead to false negatives.
- **Generalization Difficulty**: Test set distribution differing from one-class training.
- **Threshold Challenge**: Setting optimal threshold is crucial yet complex.

Evaluation Metric

- Imbalanced train and test data necessitate using the F1 score as the primary evaluation metric.
- F1 score is a harmonic mean of precision and recall.
  - Precision - the proportion of correct positive predictions wrt the number of correct and false positive predictions.
  - Recall - the proportion of correct positive predictions wrt the number of positive cases.

$$\text{F1-Score} = \frac{TP}{TP + \tfrac{1}{2} \cdot (FP + FN)}$$

Why Deep Learning Algorithm?

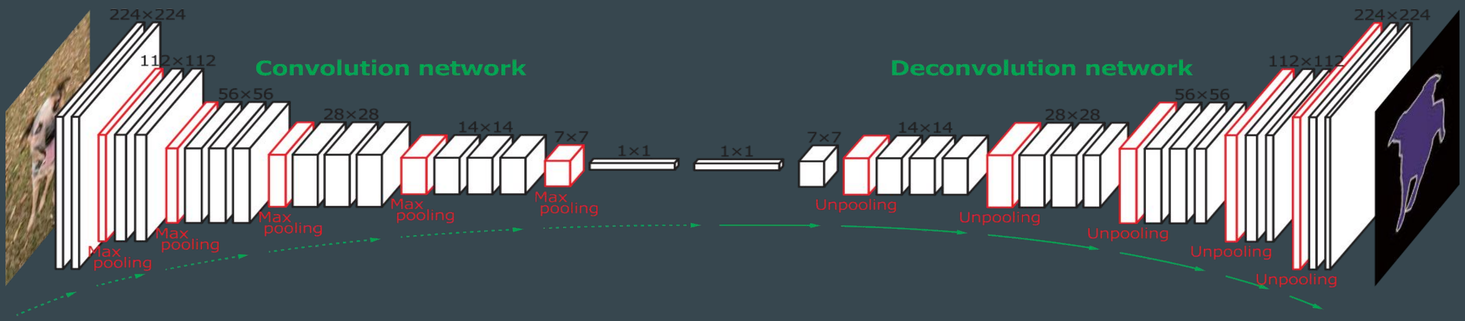
- Classical machine learning algorithms exhibit poor performance.
- Hybrid models, including CNN with Isolation Forest, CNN with Local Outlier Factor, or CNN with a pre-trained model and Isolation Forest, also display subpar performance.
- Among classical algorithms, the one-class Support Vector Machines model with a standard kernel stands out as the best performer.

The best one can get with classical algorithms

PRECISION	0.26
RECALL	0.46
F1-SCORE	0.33

Models

- Simple Stacked Autoencoder
- Complex Stacked Autoencoder
- Complex Stacked Autoencoder -RESNET50



In Search of Defects

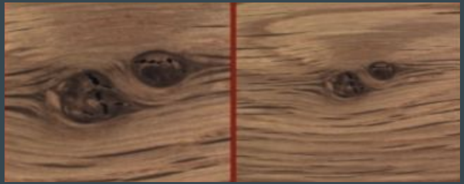
- Reconstructed images are generated based on latent representations [bottleneck]
- The pixel-wise absolute differences between original and reconstructed images are calculated
- Error Image Tensors are vectorized
- Mean (μ) and standard deviation (σ) of absolute differences are calculated for train data

Extreme RGB color combination generated by the defect

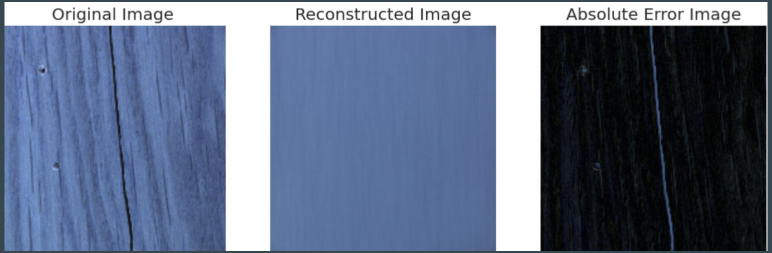
- In vectorized error images for test data we search for extreme deviations
- Extremeness is determined by β - $\mu + \beta \cdot \sigma$
- The GridSearch algorithm determines the optimal level of β for each class, generating the best possible classification while adhering to the constraint that the model predicts a minimum of 30% of TN and TP.

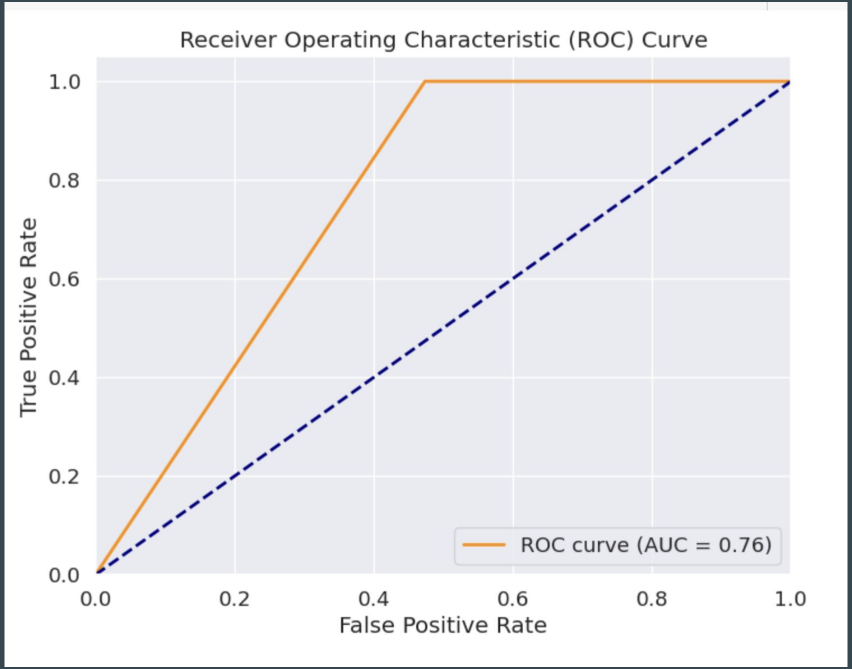
Simple Stacked Autoencoder

- Autoencoder with 3 dense layers
- Nadam as optimizer



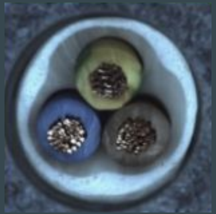
PRECISION	0.87
RECALL	1.00
F1-SCORE	0.93

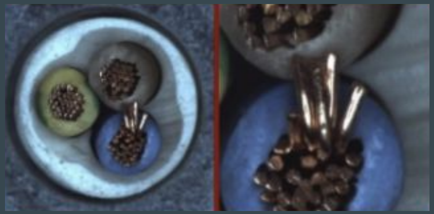




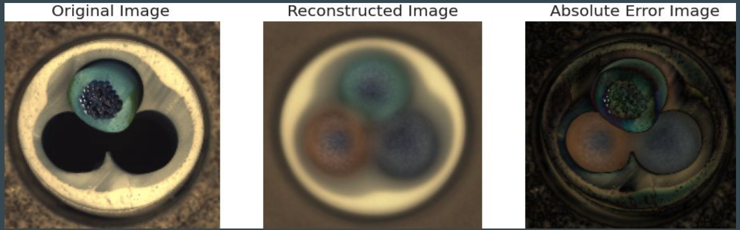
Complex Stacked Autoencoder

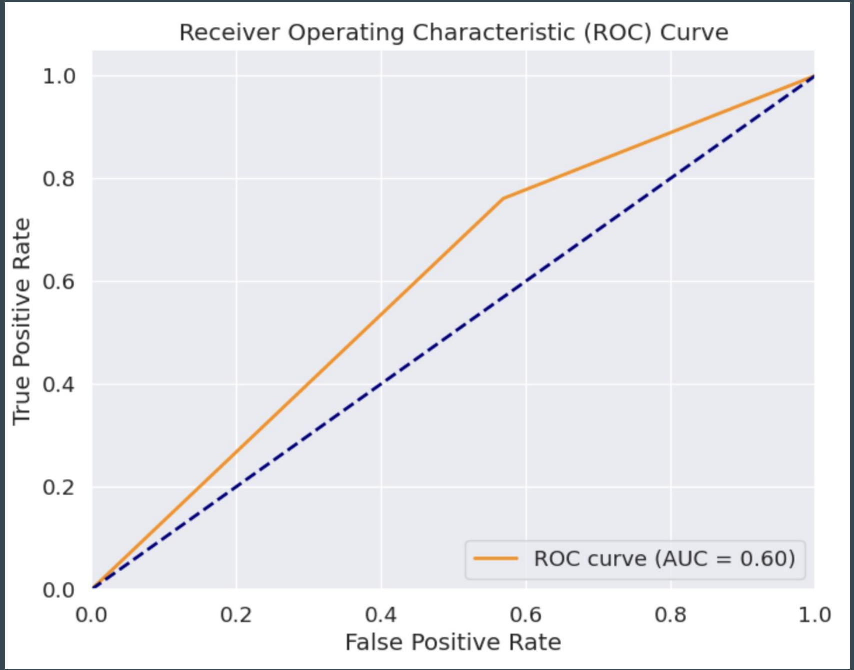
- Conv2D layers added
- MaxPooling2D layers added





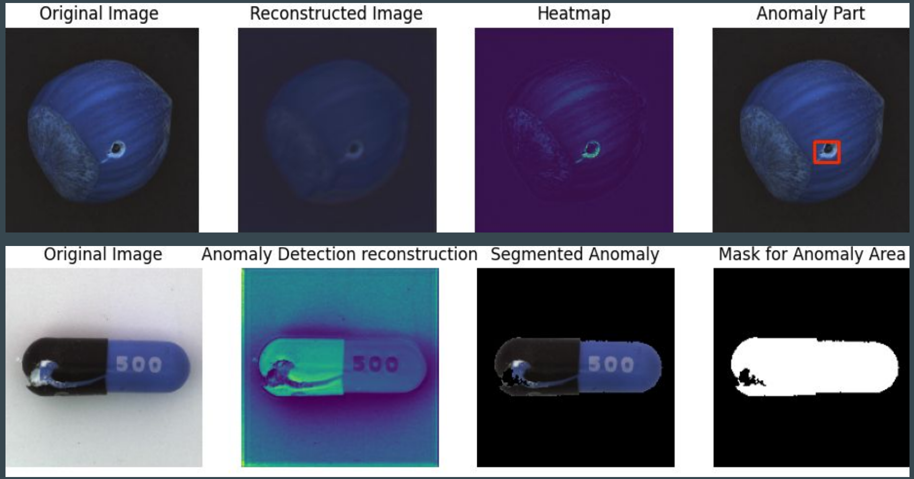
CABLE	Model 1	Model 2
PRECISION	0.70	0.68
RECALL	0.33	0.76
F1-SCORE	0.45	0.72

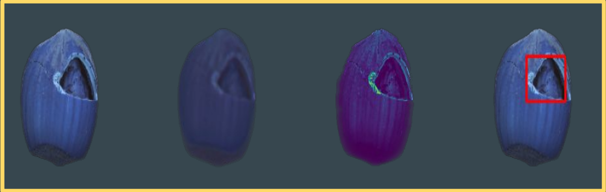


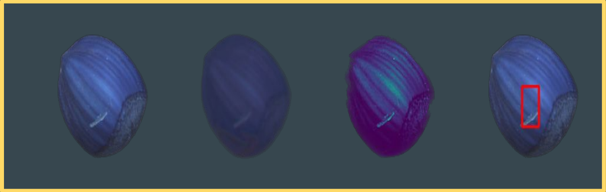


Complex Stacked Autoencoder - RESNET50

- ResNet-50 used as backbone
- Augmented pictures
- Conv2D used
- Batch normalization
- Activations after each batch normalization
- Residual connections added







Overall Results - F1-Scores

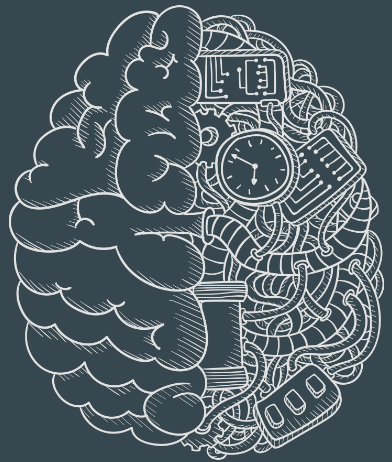
	Simple Stacked Autoencoder	Complex Stacked Autoencoder	Complex-Autoencoder Resnet 50
Wood	93,00%	92,90%	90,90%
Bottle	90,90%	88,40%	-
Toothbrush	89,60%	87,10%	-
Hazelnut	87,20%	87,20%	87,00%
Tile	86,40%	87,90%	84,10%
Zipper	86,40%	81,20%	-
Screw	86,30%	81,90%	85,30%
Metal Nut	86,00%	-	89,40%
Grid	83,30%	75,70%	60,40%
Leather	82,90%	79,10%	85,20%
Pill	80,90%	72,90%	80,00%
Capsule	78,60%	77,60%	-
Transistor	56,30%	64,30%	46,00%
Carpet	54,50%	53,10%	56,80%
Cable	44,40%	71,80%	75,50%

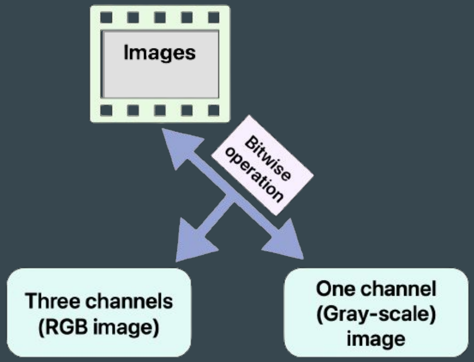
Overall Results - ROCAUC

	Simple Stacked Autoencoder	Complex Stacked Autoencoder	Complex-Autoencoder Resnet 50
Hazelnut	0,77	0,79	0,87
Tile	0,75	0,76	0,67
Pill	0,59	0,63	0,63
Bottle	0,86	0,70	0,60
Cable	0,55	0,60	0,52
Grid	0,68	0,68	0,52
Capsule	0,58	0,61	0,50
Carpet	0,45	0,38	0,50
Leather	0,69	0,68	0,50
Metal Nut	0,50	0,70	0,50
Screw	0,72	0,71	0,50
Toothbrush	0,71	0,74	0,50
Transistor	0,66	0,70	0,50
Wood	0,76	0,78	0,50
Zipper	0,64	0,66	0,50

Anomaly Segmentation and Localization

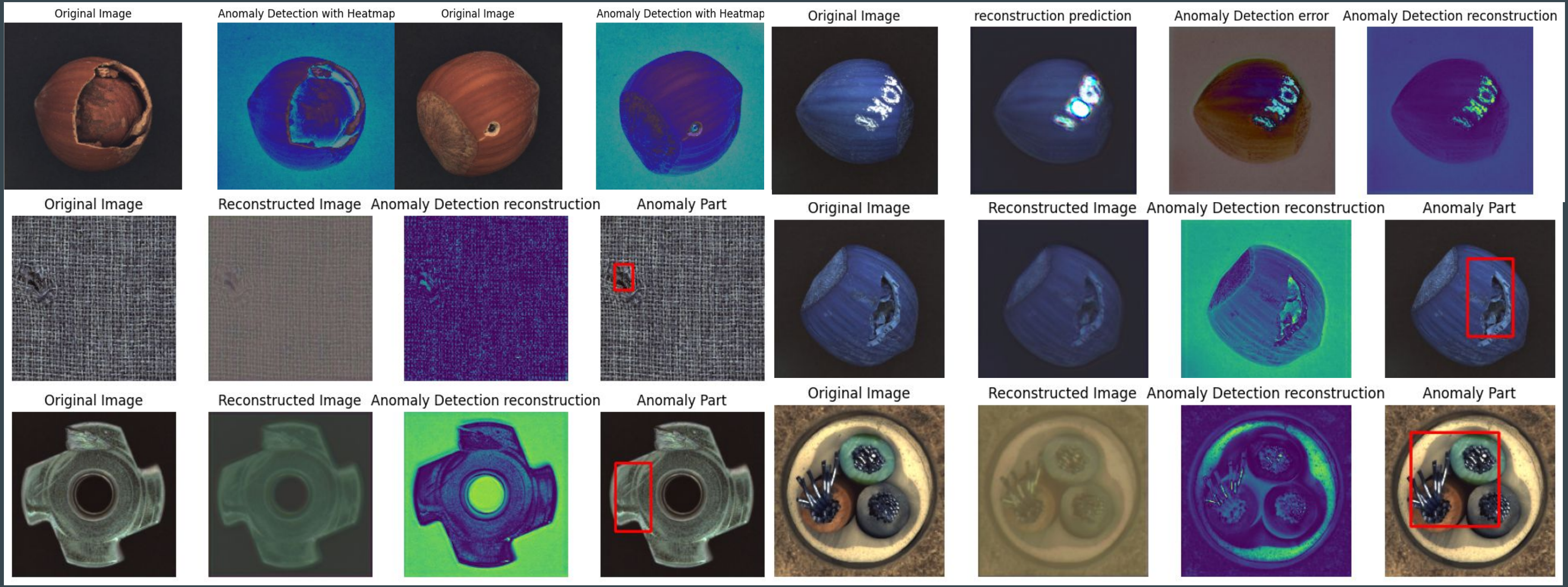
- A complex model is necessary to attain detailed high and low-level features for high-quality reconstructed images.
- Detecting image anomalies requires the reconstruction error to surpass a predefined threshold.
- Higher reconstruction error pixels are identified and localized, with a bounding box drawn around the anomaly region.





- After localization, masking is applied by transferring anomaly information to the bitwise operation that converts RGB to a single channel.

Results - Localization



Results - Segmentation