TrustMAE: A Noise-Resilient Defect Classification Framework using Memory Augmented Auto-Encoders with Trust Regions

臺科大
TAIWAN TECH

Daniel Stanley Tan, 4th Year Ph.D. Student
Advisor: Professor Kai-Lung Hua
Department of Computer Science and Information Engineering
National Taiwan University of Science and Technology

Research Focus

Motivation:

While existing defect detection models do not require detailed labeling of the images, they often assume the input data are free of defective images. As a result, these models can become overly sensitive to defective images (noise) that accidentally leak into the dataset, which frequently occurs in many manufacturing facilities.

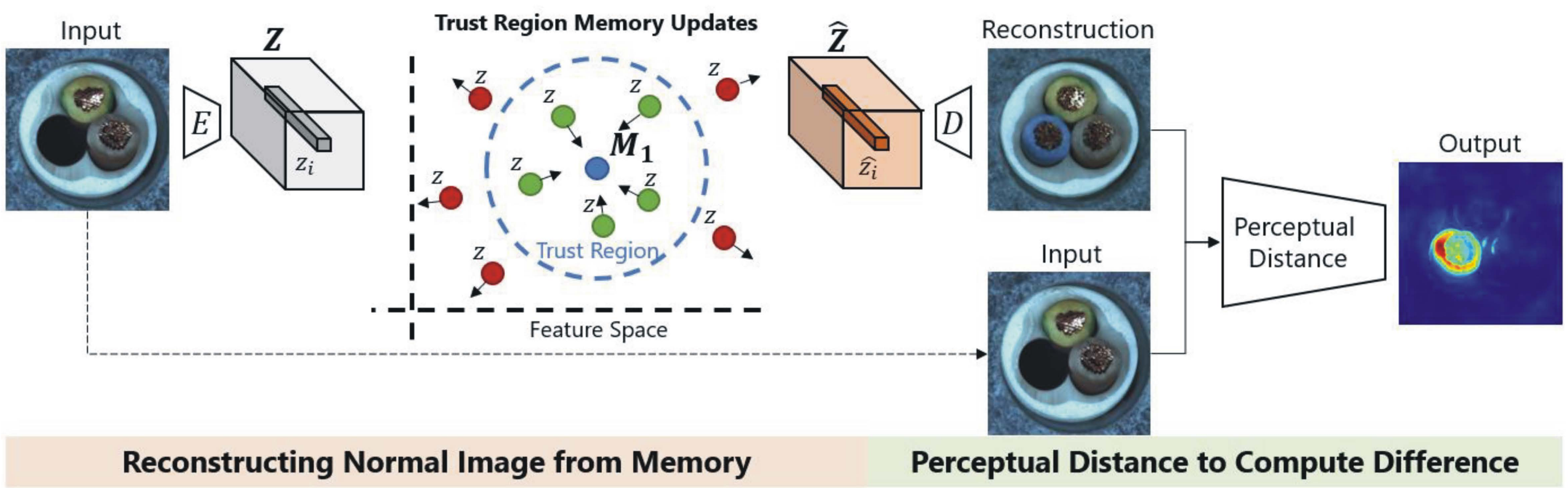
Our Contributions:

- A defect classification framework resilient to noise in the training dataset
- A memory update scheme using trust regions to avoid noise contamination
- A perceptual-based classification method to effectively detect defective regions for highly textured images.

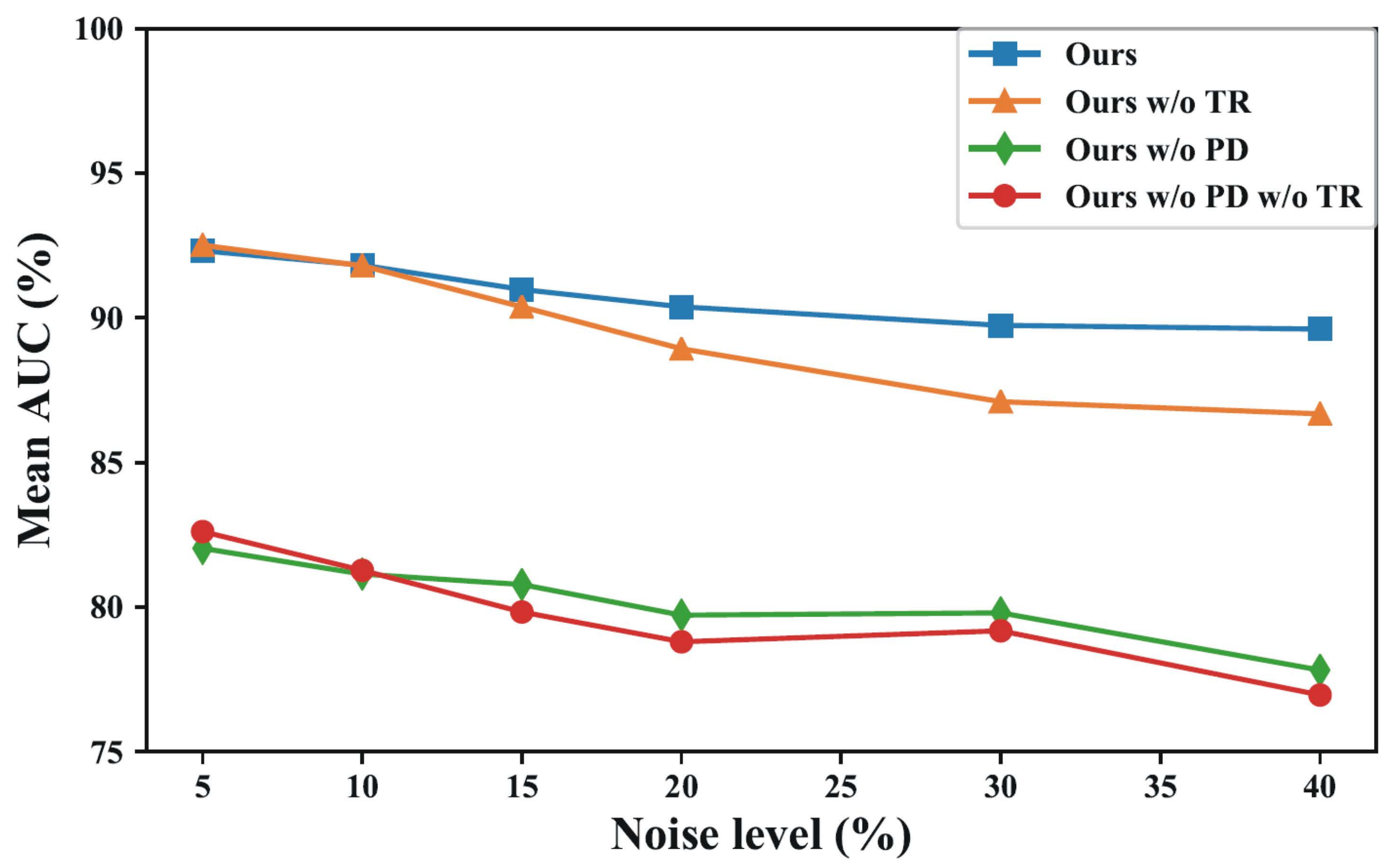
Method

Key Ideas:

- Assuming dataset contains mostly normal images, trust region memory updates prevent defects from being learned by pushing away vectors z that are outside of the trust region, which is defined based on similarity to majority of the data.
- Perceptual distance uses deep features to capture texture and high-level information in computing distance to normal, making it more suitable for texture similarity.

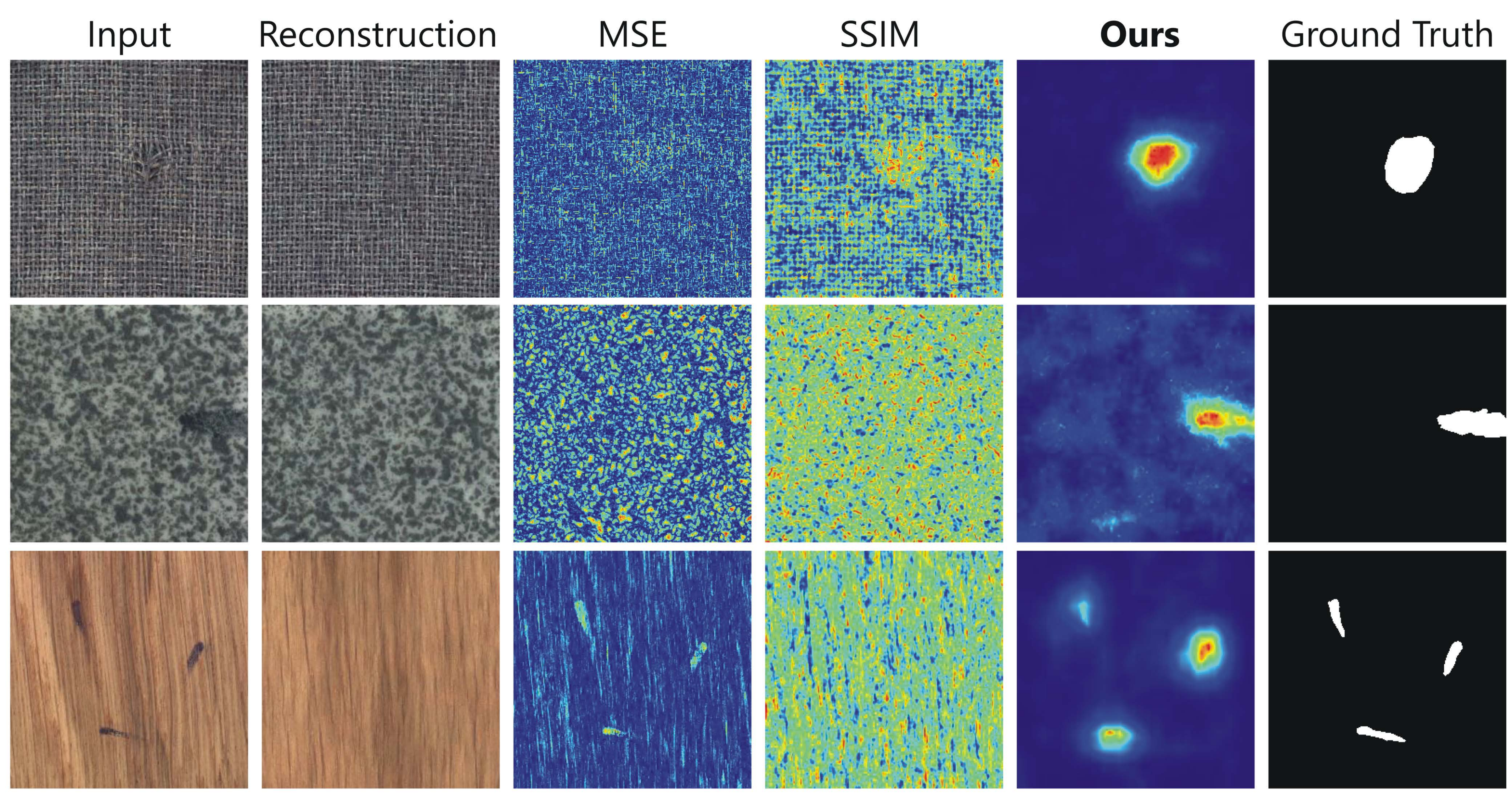


Ablation Study



TR : Trust Region Memory Updates **PD** : Perceptual Distance

Visual Results



Quantitative Results

Table 1: Comparison against baselines in terms of mean area under the receiver operating curve (AUC)

Method	Mean AUC
GeoTrans [1]	67.23%
GANomaly [2]	76.15%
f-Ano-GAN [3]	65.85%
MemAE [4]	81.85%
ARNet [5]	83.93%
TrustMAE (Ours)	**90.78%**

References

[1] Golan, I., & El-Yaniv, R. (2018). Deep anomaly detection using geometric transformations. In *Advances in Neural Information Processing Systems* (pp. 9758-9769).

[2] Akcay, S., Atapour-Abarghouei, A., & Breckon, T. P. (2018, December). Ganomaly: Semi-supervised anomaly detection via adversarial training. In *Asian conference on computer vision* (pp. 622-637). Springer, Cham.

[3] Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., & Schmidt-Erfurth, U. (2019). f-anogan: Fast unsupervised anomaly detection with generative adversarial networks. *Medical image analysis*, *54*, 30-44.

[4] Gong, D., Liu, L., Le, V., Saha, B., Mansour, M. R., Venkatesh, S., & Hengel, A. V. D. (2019). Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1705-1714).

[5] Huang, C., Cao, J., Ye, F., Li, M., Zhang, Y., & Lu, C. (2019). Inverse-transform autoencoder for anomaly detection. *arXiv preprint arXiv:1911.10676*.