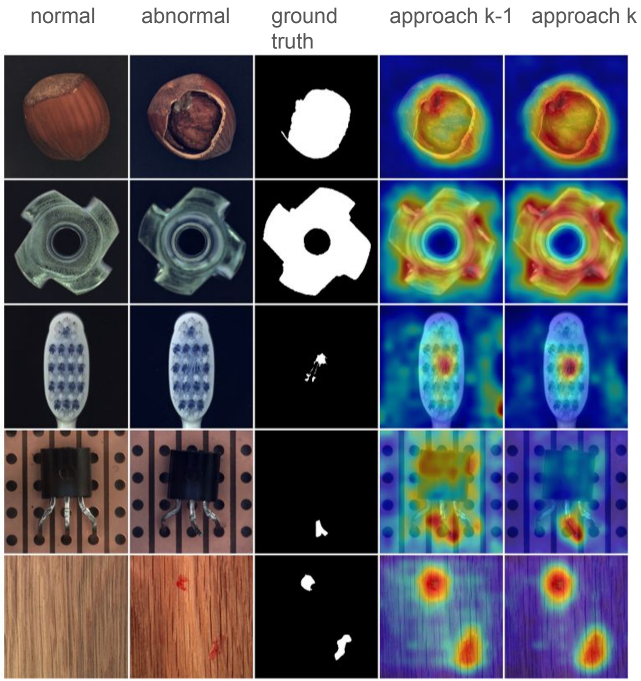
Anomaly Detection in an unsupervised framework



Elaborated by:

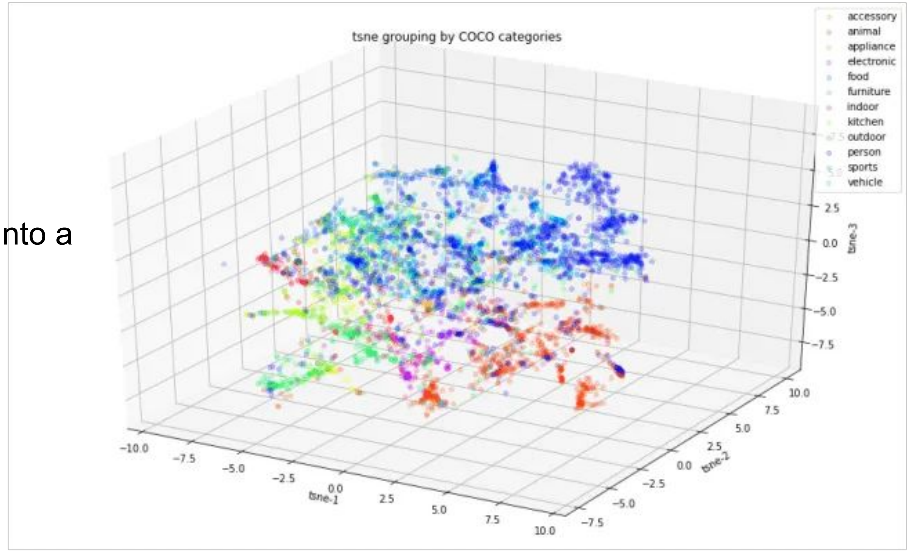
Montassar MHAMDI

Different families of representative methods

Unsupervised AD

- Feature Embedding-based approaches
  - Teacher-Student Architecture
  - One-Class Classification
  - Distribution Map
  - Memory Bank
- Reconstruction-based approaches
  - Autoencoder
  - Generative Adversarial Networks
  - Transformer
  - (Diffusion Model)

Feature Embedding-based approaches



what's a feature embedding ?

A representation of high-dimensional categorical data into a lower-dimensional space.

- ❏ Embedded similarity methods use deep neural networks to extract features that describe a sample.
- ❏ Learn the reference vectors representing normality from a training dataset.
- ❏ Identify anomalies by the distance between the embedding vectors of a test image and the learned reference vectors.

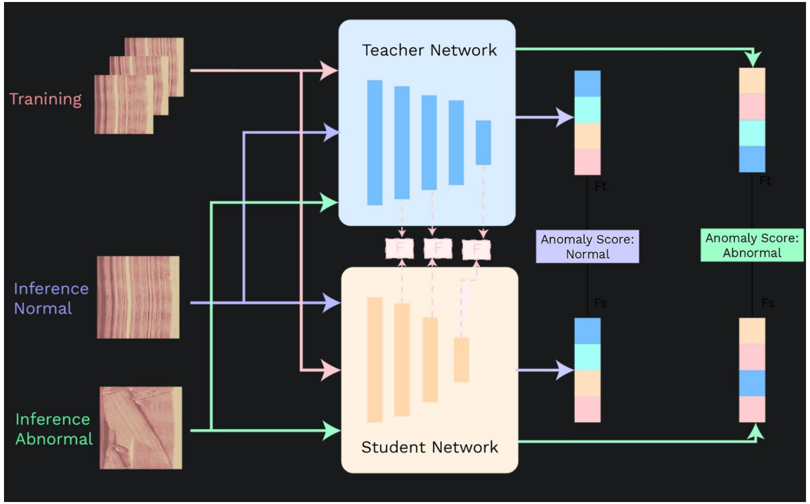
Feature Embedding-based approaches

Teacher-Student Architecture

- The teacher-student methods select partial layers of a backbone network pre-trained on a large scale dataset, as ResNet, VGG or EfficientNet, and use it as a fixed parameter-teacher.

- During training the teacher model guides the student model to learn the characteristics of a normal sample extracting the features that represent them. During inference the features extracted of normal samples from the teacher and student network are comparable but the features extracted for anomalous samples are different.

- Comparing the features maps that generated both networks, these methods can generate score maps to determine whether an image is anomalous or not even at pixel level for anomaly localization. The teacher-student methods have variation in the architectures, some use multiple student networks or teacher-students configuration, use final and internal features to generate the maps and different calculations to define the anomaly score.

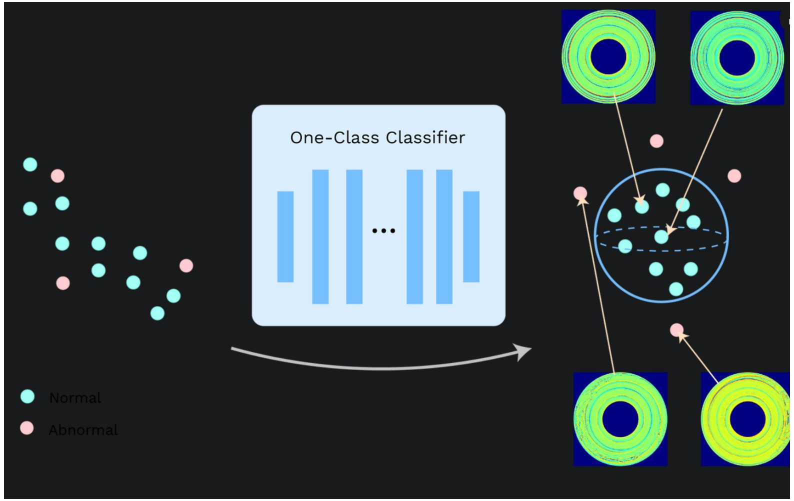


Feature Embedding-based approaches

One-Class Classification

- ❑ Anomaly detection can be tackled with one class classification(OCC) methods. During training, the OCC maps normal samples to a compact representation and finds a boundary that encompasses the normal sample features, usually called hypersphere, that hopefully will provide a good separation for any abnormal features. During inference, the methods determine if a sample contains an anomaly by determining the relative position of the sample features and the hypersphere.

- ❑ These classifiers work under the assumption that the training dataset consists only of normal samples, having abnormal samples mixed in the training dataset may produce detrimental effects on the deep models results. Some methods generate abnormal samples artificially and use them during training to improve the accuracy of the boundary.

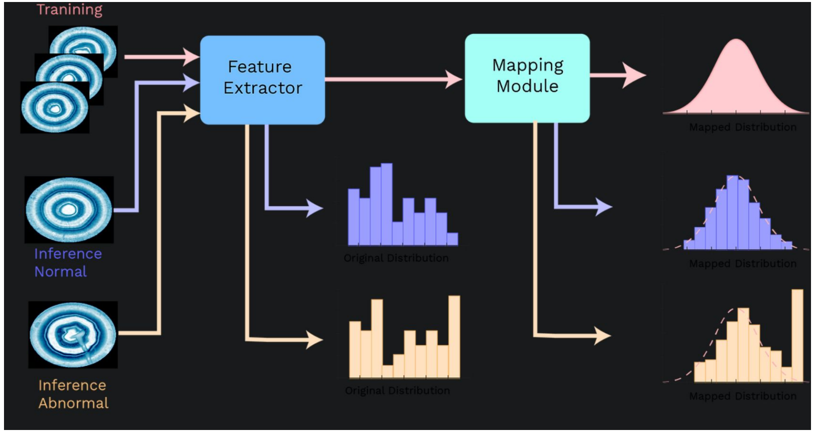


Feature Embedding-based approaches

Distribution Map

- ❏ During training, distribution-map based methods map the input normal images or features to a probability distribution and then, during testing, they judge whether the sample is normal or abnormal by estimating the deviation of the distribution or the likelihood probability of the sample against the established distribution.

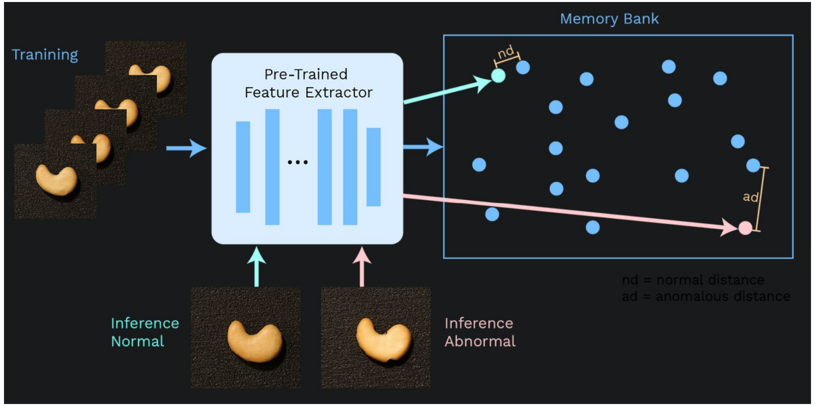
- ❏ In the Distribution Map methods predominate Normalizing Flow(NF) variations because of its strong mapping ability and good performance in anomaly detection tasks. Normalizing Flows are neural networks that are able to learn transformations between data distributions and well defined densities.



Feature Embedding-based approaches

Memory Bank

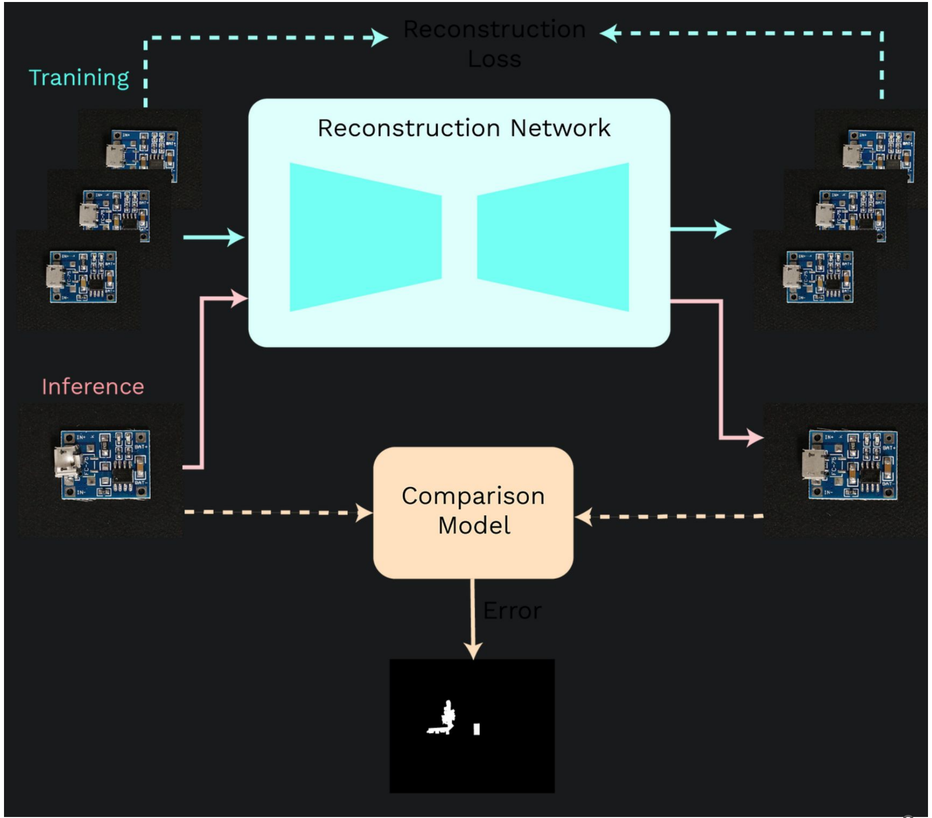
- Memory bank methods rely on robust pre-trained networks and require additional memory space to store the image features. These methods required minimal network training and only required sampling or mapping of the collected features for inference. During inference, the features of the sample image are compared to the features in the memory bank, an image is considered abnormal if the spatial distance between the test features and the nearest normal features in the memory bank is larger than a threshold, otherwise if the distance is small it is considered normal.



Reconstruction-based approaches

- In reconstruction based methods neural network architectures like auto encoders(AE), variational autoencoders(VAE) or generative adversarial networks(GAN) are trained to reconstruct normal images so anomalous images can be identified as those that are not well reconstructed.

- For these methods, the reconstruction error can be used as the anomaly score but some have combined it with additional information from latent space and intermediate activations to improve the anomaly detection score.

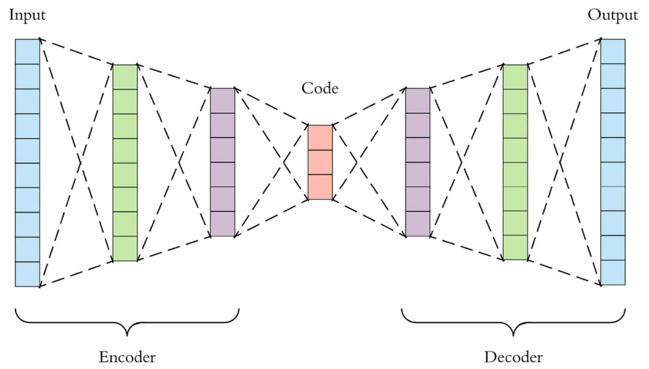


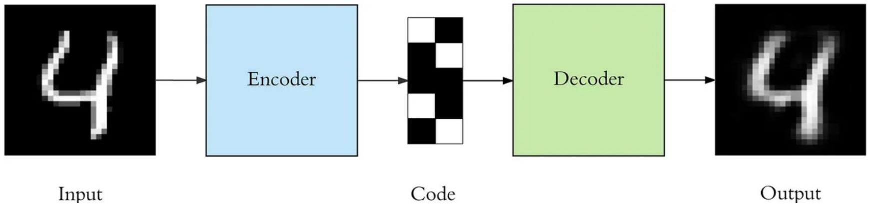
Reconstruction-based approaches

Autoencoder (AE)

- Autoencoder networks are the commonly used method for the anomaly detection in the reconstruction category. The method's variations change how to determine the difference between the reconstructed image and the original image and how to define the anomaly score. Also, some methods to improve the effectiveness of reconstruction use reconstruction at feature level to take advantage of robust pre-trained networks.

- Another technique is to use reconstruction at different scales and patches to take into account detection of anomalies at global structure and in detail. The modification of the autoencoder structure has value to improve the reconstruction capabilities, they have used skipping connections and adding memory modules. On the other hand, another method synthesizes abnormal images and reconstructs them as normal to improve the generalization capacity.



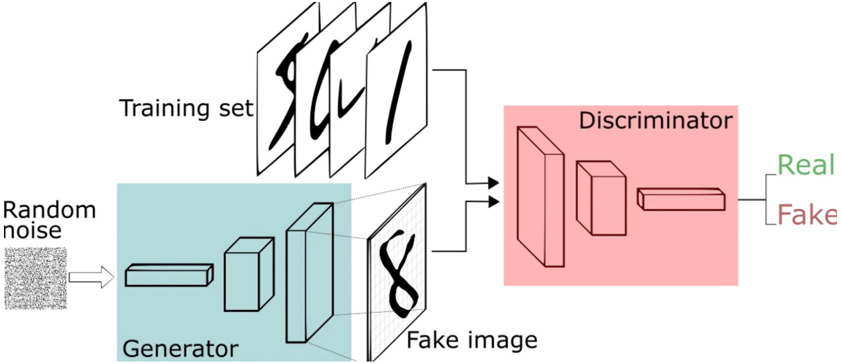


Reconstruction-based approaches

Generative Adversarial Networks (GANs)

- The use of Generative Adversarial Networks(GAN) is also a popular approach for deep anomaly detection. This method targets to learn a latent feature space with a generative network so that the latent space well represents the normality underlying the train data. During inference GAN can reconstruct normal samples from the latent space and differentiate the sample as coming from the true data distribution.

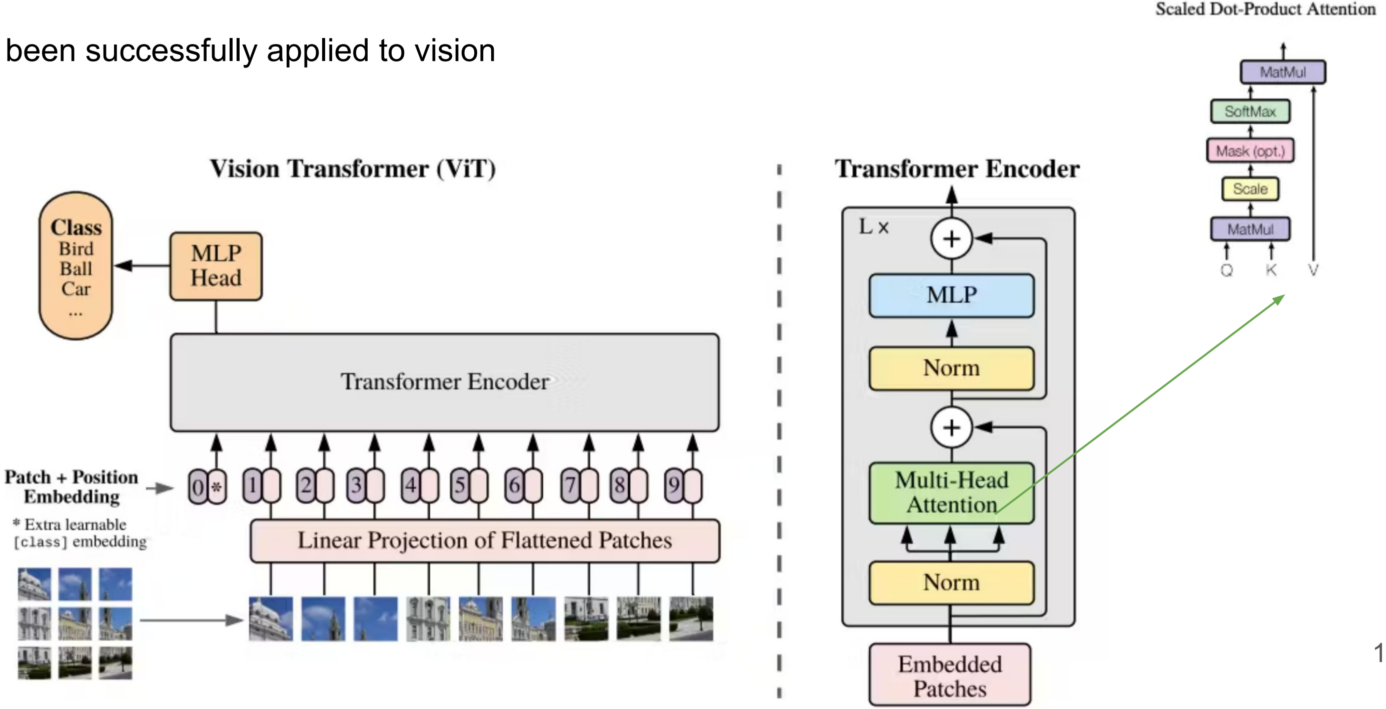
- GANs can capture the complexity and variability of real data but have some drawbacks for anomaly detection, such as mode collapse, where the generator produces only a few modes of the data distribution and ignores the rest. Additionally, **GANs can be difficult to train**, as the generator and the discriminator need to be balanced and synchronized, and **the loss function can be non-convex and non-smooth**.



Reconstruction-based approaches

Transformer

- ❏ Transformers were originally developed for the application of natural language processing (Attention is All You Need 2017, Google).
- ❏ They have recently been successfully applied to vision processing.



Bibliography

Anomaly Detection via Reverse Distillation From One-Class Embedding, **CVPR 2022**, **Teacher-Student**

Revisiting Reverse Distillation for Anomaly Detection, **CVPR 2023**, **Teacher-Student**

SimpleNet: A Simple Network for Image Anomaly Detection and Localization, **CVPR 2023**, **One-Class-Classification**

CFLOW-AD: Real-Time Unsupervised Anomaly Detection With Localization via Conditional Normalizing Flows, **WACV 2022**, **Distribution Map**

PyramidFlow: High-Resolution Defect Contrastive Localization Using Pyramid Normalizing Flow, **CVPR 2023**, **Distribution Map**

Towards Total Recall in Industrial Anomaly Detection, **CVPR 2022**, **Memory Bank**

PNI : Industrial Anomaly Detection using Position and Neighborhood Information, **ICCV 2023**, **Memory Bank**

DRAEM - A Discriminatively Trained Reconstruction Embedding for Surface Anomaly Detection, **ICCV 2021**, **Reconstruction-based**