Improving Modality in Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection

Aryan Solanki
IIT Gandhinagar
23110049@iitgn.ac.in

Nupoor Assudani
IIT Gandhinagar
23110224@iitgn.ac.in

Rishabh Jogani
IIT Gandhinagar
23110276@iitgn.ac.in

ABSTRACT

This project proposes an approach to anomaly detection in industrial image data using a Deep Autoencoding Gaussian Mixture Model (DAGMM) adaptation. Our framework integrates a convolutional autoencoder for dimensionality reduction and a Gaussian Mixture Model for density estimation, optimized jointly to enhance detection accuracy. This method has been used for anomaly detection and classification on medical imagery but hasn't been explored much for industrial uses such as detecting surface and structural defects in manufactured goods.

Additionally, we will explore a cross-modal extension integrating image and text data to correlate visual defects with textual diagnostics, enhancing anomaly detection capabilities.

KEYWORDS

Gaussian Mixture Model, Anomaly Detection, Auto Encoder, Unsupervised Learning, Image Anomaly

1 INTRODUCTION

Industrial image anomaly detection has an important role in quality control. Image data can be used to check for surface and structural defects in manufactured goods such as printed circuit boards.

Traditional methods relying on manual feature extraction struggle with high-dimensional industrial data and subtle defect patterns, while supervised deep learning approaches face limitations due to the scarcity of labeled anomaly samples in real-world production environments.

This work proposes a hybrid architecture combining the density estimation capabilities of Deep Autoencoding Gaussian Mixture Models (DAGMM) with transformer-enhanced feature learning. This implies a custom loss function that jointly trains both the convolutional autoencoder and the Gaussian mixture model.

2 RELATED WORK

2.1 Deep Autoencoding Gaussian Mixture Model (DAGMM) for Anomaly Detection

Unsupervised anomaly detection in high-dimensional data relies on density estimation, but traditional two-step methods of dimensionality reduction followed by density estimation often perform suboptimally due to decoupled learning. The Deep Autoencoding Gaussian Mixture Model (DAGMM) addresses this by jointly optimizing a deep autoencoder for dimensionality reduction and a Gaussian Mixture Model (GMM) for density estimation in an end-to-end manner . DAGMM's compression network (deep autoencoder) produces a low-dimensional representation and reconstruction error, which are then used by an estimation network to predict GMM membership. This joint training allows the autoencoder to learn representations that directly aid density estimation, leading to improved unsupervised anomaly detection.[3]

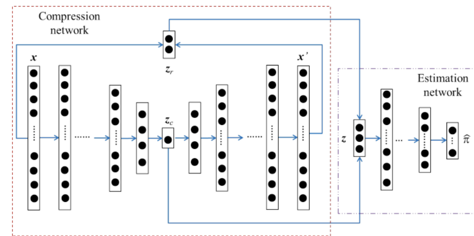

Figure 1: Deep Autoencoding Gaussian Mixture Model Architecture

Algorithm 1 Deep Autoencoding Gaussian Mixture Model (DAGMM)

Require: Dataset $\mathcal{X} = \{x_1, x_2, ..., x_N\}$, number of Gaussian components K, autoencoder parameters θ, threshold τ

Ensure: Anomaly scores for each sample in $\mathcal{X}$

Step 1: Train Autoencoder

Train an autoencoder with encoder function $f_{\theta_e}(x)$ and decoder function $g_{\theta_d}(z)$:

$$z_i = f_{\theta_e}(x_i)$$

$$\hat{x}_i = g_{\theta_d}(z_i)$$

Compute the reconstruction error:

$$r_i = \|x_i - \hat{x}_i\|^2$$

Step 2: Construct Feature Representations

Form feature vector for GMM by concatenating the latent representation and reconstruction error:

$$h_i = [z_i, r_i]$$

Step 3: Fit Gaussian Mixture Model (GMM)

Estimate GMM parameters $\{\pi_k, \mu_k, \Sigma_k\}_{k=1}^K$ using Expectation-Maximization (EM):

$$\gamma_{ik} = \frac{\pi_k \mathcal{N}(h_i \mid \mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(h_i \mid \mu_j, \Sigma_j)}$$

Where $\mathcal{N}(h_i \mid \mu_k, \Sigma_k)$ is the Gaussian density function.

Step 4: Compute Sample Energy

Compute the energy function:

$$E(h_i) = -\log \sum_{k=1}^K \pi_k \mathcal{N}(h_i \mid \mu_k, \Sigma_k)$$

Step 5: Detect Anomalies

Define the anomaly score as:

$$s_i = E(h_i)$$

Classify x_i as an anomaly if $s_i > \tau$

Figure 2: DAGMM ALgorithm (End-to-End)

2.2 Deep convolutional neural network- based anomaly detection for organ classification in gastric X-ray examination

Togo et al. proposed a deep learning-based anomaly detection model for organ classification in gastric X-ray examinations. Their approach, inspired by Deep Autoencoding Gaussian Mixture Model (DAGMM), addresses the challenge of other organs appearing in gastric X-rays, which can introduce noise and degrade classification performance. The model was evaluated on a dataset of gastric X-ray images and compared against other approaches. Results showed that their method outperformed comparison models, achieving 95.6% sensitivity in organ classification. This work demonstrates the potential of deep learning-based anomaly detection in improving organ classification accuracy in gastric X-ray examinations[5].

2.3 Deep Learning for Anomaly Detection: A Survey

Chalapathy & Chawla (2019) present a comprehensive survey of deep learning approaches for anomaly detection, categorizing methods into reconstruction-based (e.g., autoencoders), probabilistic models (e.g., GANs), and hybrid architectures. The analysis highlights the effectiveness of autoencoder-GMM hybrids like DAGMM in handling complex data distributions—a critical requirement for industrial image anomalies. The authors identify key challenges in anomaly detection, including high-dimensional feature learning and unsupervised adaptation, which align with the core problems addressed in industrial inspection systems. While the survey broadly covers medical and cybersecurity applications, it notes the untapped potential of deep anomaly detection in manufacturing quality control. This work provides foundational insights for our integrated autoencoder-GMM approach, particularly in optimizing joint reconstruction-density estimation objectives [1].

Key connections to our work:

(1) Validates the DAGMM architecture's relevance for complex visual data
(2) Supports the need for hybrid models in industrial settings
(3) Highlights computational efficiency challenges addressed in your joint training strategy
(4) Identifies cross-modal detection as an emerging research frontier

3 METHODOLOGY AND OBJECTIVE

1. **Primary Objective**: Explore and enhance the Deep Autoencoding Gaussian Mixture Model (DAGMM) architecture [3] for image anomaly detection.

- Evaluate DAGMM's applicability to image modalities.
- Implement and develop an improved convolutional DAGMM architecture [5].
- Construct an advanced image embedding methodology to optimize unsupervised anomaly detection performance.

2. **Secondary Objective**: Advance cross-modal anomaly detection through integrated image-text representation learning.

- Develop unified anomaly detection across multimodal data.
- Enable mutual information transfer between image and text modalities.

4 DATASETS AND LIBRARIES

We intend to use pytorch and it's libraries. We have chosen the following datasets due to their widespread use in anomaly detection testing. Additionally, we wish to explore general image and text datasets to improve this architecture for cross modality. We are open to explore other anomaly datasets as well.

Dataset	Instances	Anomaly Ratio
VisA (Visual Anomaly Dataset) [4]	10,821	0.12
MVTec AD [2]	5,354	0.16

Table 1: Proposed Datasets for Evaluation

5 PROJECT TIMELINE

(1) **Week 1**: Reproduce DAGMM benchmarks on datasets as per the reference paper [3] and develop convolutional DAGMM implementation from scratch.
(2) **Week 2**: Evaluate proposed image architecture [5] on diverse image datasets for unsupervised anomaly detection.
(3) **Week 3**: Refine architecture for improved image understanding and design cross-modal anomaly detection pipeline.
(4) **Week 4**: Finalize cross-modal experiments and prepare project presentation.

REFERENCES

[1] R. Chalapathy and S. Chawla. 2019. Deep Learning for Anomaly Detection: A Survey. *arXiv:1901.03407*. https://arxiv.org/abs/1901.03407

[2] Paul Bergmann et al. 2021. MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. *Springer Nature*. https://link.springer.com/article/10.1007/s11263-020-01400-4

[3] Zong et al. 2018. DEEP AUTOENCODING GAUSSIAN MIXTURE MODEL FOR UNSUPERVISED ANOMALY DETECTION. In *ICLR*. https://openreview.net/forum?id=BJJLHbb0-

[4] Zou et al. 2021. VisA (Visual Anomaly Dataset). In *SPot-the-Difference Self-Supervised Pre-training for Anomaly Detection and Segmentation*. https://paperswithcode.com/dataset/visa

[5] Haruna Watanabe et al. Ren Togo. 2019. Deep convolutional neural network-based anomaly detection for organ classification in gastric X-ray examination. In *IEEE Global Conference on Life Sciences and Technologies (LifeTech)*. https://www.sciencedirect.com/science/article/pii/S001048252030250X