UDC 004.855.5
DOI: 10.15587/2706-5448.2026.362237

Tymur Dovzhenko, Kamila Storchak

ADVANTAGES OF THE END-TO-END HYBRID AWRED ARCHITECTURE IN TERMS OF THE EFFICIENCY OF VISUAL ANOMALY DETECTION UNDER CONDITIONS OF TRAINING DATA DEFICIENCY COMPARED WITH THE CLASSICAL CNN + ONE-CLASS SVM ENSEMBLE

The object of research is the process of detecting visual anomalies in images under conditions of reduction of the training sample and class imbalance, relevant for visual monitoring systems of IT infrastructure and telecommunication equipment, including recognition of microcracks on printed circuit boards, corrosion on antennas, and damages of fiber-optic lines. The problem lies in the fact that with a small volume of training data two-stage approaches lose stability, reducing defect recognition accuracy. This concerns schemes in which a convolutional autoencoder is combined with an external classifier One-Class SVM. Under such conditions, the latent representation is formed with lower quality, and the ranking of anomalies becomes less reliable.

As an alternative, the Hybrid AWRED v4 architecture was used, in which anomaly detection is performed directly in the space of reconstruction errors without an external classifier. The approach is based on an objective function that combines dynamic weighting and an adaptive cutoff threshold.

The verification was carried out on three datasets of 800, 107, and 54 images. For each dataset, eight runs were performed. On the sample N = 800, the CNN + AWRED architecture showed better Precision, F1-Score, and MCC than the CNN + SVM ensemble. At N = 107, the advantage of the proposed approach was manifested in AUC-ROC and AP.

For the micro-sample N = 54, the threshold metrics of both approaches were close, while AUC-ROC and AP remained higher in the baseline model. This indicates that with such a data volume both approaches approach the limit of their effectiveness without additional expansion of the sample. It was established that Hybrid AWRED reaches the early stopping criterion earlier, and its heat maps form clearer zones in defect areas. The approach is promising for automation of visual control under deficit of training data.

Keywords: *end-to-end architecture, Hybrid AWRED, One-Class SVM, dynamic weighting, class imbalance, visual control.*

Received: 19.02.2026
Received in revised form: 08.05.2026
Accepted: 22.05.2026
Published: 29.05.2026

How to cite
Dovzhenko, T., Storchak, K. (2026). Advantages of the end-to-end Hybrid AWRED architecture in terms of the efficiency of visual anomaly detection under conditions of training data deficiency compared with the classical CNN + One-Class SVM ensemble. Technology Audit and Production Reserves, 3 (2 (89)), 53–59. https://doi.org/10.15587/2706-5448.2026.362237

1. Introduction

The effectiveness of modern convolutional neural networks in computer vision tasks largely depends on the volume and representativeness of the training data. The areas of application of such systems, in particular, automated visual monitoring of the condition of hardware in data centers and telecommunications infrastructure facilities (automatic detection of thermal damage to switching boards, cable insulation defects or wear of external antennas), monitoring of infrastructure facilities and industrial flaw detection, require reliable methods for detecting deviations. In these tasks, the condition of having large data is often not met, since defective samples occur less frequently. This creates a situation of significant class imbalance, where spatial and textural manifestations of anomalies can differ significantly even within the same type of objects. Because of this, the formation of a full-fledged labeled sample for a class of anomalies is difficult, and the use of classical supervised learning approaches is limited.

For this reason, unsupervised anomaly detection methods have become widespread in the practice of infrastructure and visual inspection. One typical approach is to use a convolutional autoencoder trained only on normal samples, followed by an analysis of the reconstruction error. In two-stage schemes, such errors or latent features are passed to external classifiers, such as the One-Class Support Vector Machine (One-Class SVM), Isolation Forest, or other metric algorithms. This scheme can be effective on large samples, but its robustness decreases when the data volume is sharply reduced.

Generalized approaches to anomaly detection using the isolation mechanism were shown in [1]. In it, the authors analyzed the effectiveness of using the Isolation Forest family in unsupervised tasks. However, the conclusion they provided showed that the application

ISSN-L 2664-9969; E-ISSN 2706-5448

to complex visual data has limitations, since it does not take into account the image structure.

Reconstructive models were considered in [2] as one of the basic approaches to anomaly localization in industrial images. The analysis of the work showed that when there is a small number of images, they cannot always clearly distinguish local defects. These conclusions are also confirmed by the results of [3]. In it, the authors concluded that the autoencoder approach is practical for quality control. However, they also noted that additional tools are needed to clearly distinguish anomalous areas. It follows that minimizing only the standard error of reconstruction is not enough to clearly detect small defects. Thus, these limitations gave impetus to the spread of hybrid schemes. In them, a neural network forms features, and an external single-class classifier makes a decision about the presence of an anomaly. In [4], the authors compared the use of single-class methods in industrial tasks. One-Class SVM is considered as one of the basic tools for anomaly detection there. But its result is not stable. Much depends on what features are fed to it as input, what parameters of the model are used and under what conditions it is used.

A scheme that uses a combination of a convolutional autoencoder with a single-class classification [5] can improve the detection of defects in industrial images. However, the problem of two separate stages remains. The model first forms the features, and then another algorithm makes the decision. Because of this, the quality of the entire approach strongly depends on how successful the feature space turned out.

In [6], the authors showed that for single-class models it is very important not only to have well-described normal features, but also to clearly separate anomalous features from them.

In [7, 8], modern deep learning approaches for anomaly detection were reviewed. They showed that the main difficulties are associated with a small number of defective examples, class imbalance and the difficulty of finding small defects. To improve the results, training with a small number of samples, self-supervised learning, weak supervision and reasoning are used, but the quality of the model still largely depends on the formed feature space.

Continuing the analysis of modern research, it can be said that the limitation of reconstructive and two-stage single-class approaches is based on a high dependence on the quality of the latent feature space.

An additional problem is that optimization by the standard mean square error can weaken the local contrast of defects, as a result of which the boundaries of anomalous areas in heat maps become less distinct.

One way to overcome these limitations is to move to end-to-end architectures, in which the anomaly ranking mechanism is built directly into the loss function. The theoretical basis of this approach is the step-by-step learning mode. In [9] it is shown that learning according to the principle of gradual complication contributes to better convergence and stability of optimization, but in itself does not specify a specialized mechanism of anomalous ranking. Therefore, for the defect detection problem, it is advisable to combine step-by-step learning with adaptive error weighting directly in the loss function.

This approach is implemented in Hybrid AWRED. This paper considers the fourth modification of the method, which implements the mechanism of dynamic error weighting within the step-by-step learning mode. The new configuration is aimed at reducing the dependence on the external classifier and transferring the decision-making to the reconstruction error space. This made it possible to focus training on the largest deviations and implement end-to-end ranking of anomalies on microsamples.

The object of research in this paper was the process of detecting visual anomalies in images under conditions of reduced training sample and class imbalance, relevant for visual monitoring systems of IT infrastructure and telecommunications equipment, in particular, the recognition of microcracks on printed circuit boards, corrosion on base station antennas, or damage to fiber-optic communication lines.

The aim of research is to determine the possibility of the end-to-end Hybrid AWRED architecture to provide higher efficiency in detecting visual anomalies with a shortage of training data compared to the classical CNN + One-Class SVM ensemble, which will allow to justify the feasibility of its use in automated visual control systems with limited training resources.

To achieve the aim, the following objectives were solved:

1) to analyze the limitations of the two-stage approach to anomaly detection at different sample scales;

2) to apply the end-to-end Hybrid AWRED v4 architecture with dynamic error weighting without involving an external metric classifier;

3) to experimentally evaluate the quality of ranking and localization of defects on samples of different sizes and compare the results with the basic CNN + One-Class SVM scheme.

2. Materials and Methods

2.1. Methodology, architecture and mathematical apparatus of Hybrid AWRED

The basis of the proposed visual inspection system was a symmetric neural network architecture of the encoder and decoder. This architecture sequentially compressed the input image into the latent space, and then recovered it. In classical two-stage schemes, the latent features were then transferred to the external SVM classifier (Fig. 1).

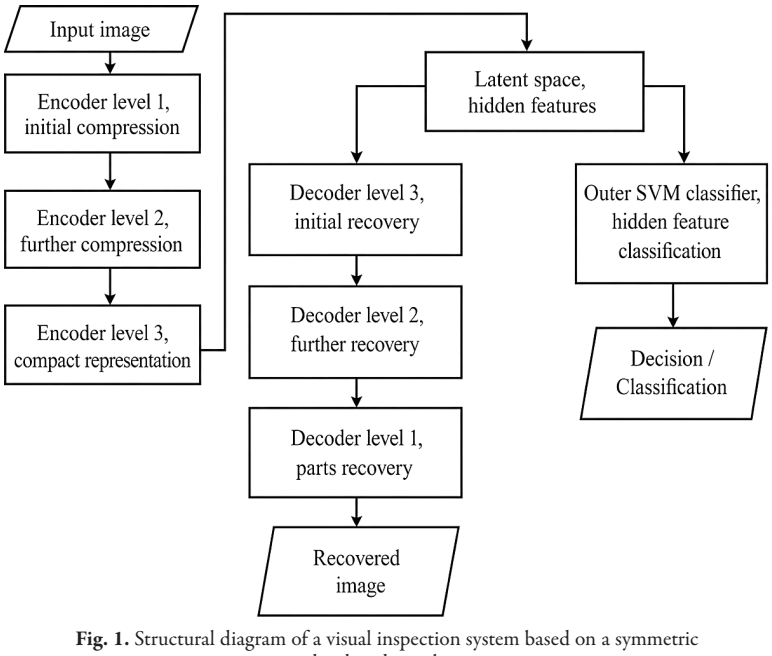

Fig. 1. Structural diagram of a visual inspection system based on a symmetric encoder-decoder architecture

This step was not used in the research. The reason was that the latent space was formed unstable on small samples, and this worsened the further separation of normal and abnormal samples. Therefore, in the developed end-to-end architecture, anomaly detection was performed directly in the reconstruction error space.

To reduce the effect of spatial averaging and not lose local defect features, the proposed model (Hybrid AWRED v4) used a soft clipping mechanism. The basic mathematical apparatus of this mechanism evolved from previous versions of the architecture. In particular, in previous researches by the authors [10–12], early versions of the method

were considered, where the logistic function played an auxiliary role and was used mainly for filtering visual noise before the spatial clustering stage or external threshold clipping.

In this work (version v4), its role was changed. Here, it was directly included in the decision-making scheme and was used as an element of end-to-end anomaly ranking. This made it possible to abandon SVM at the final stage of classification (Fig. 2).

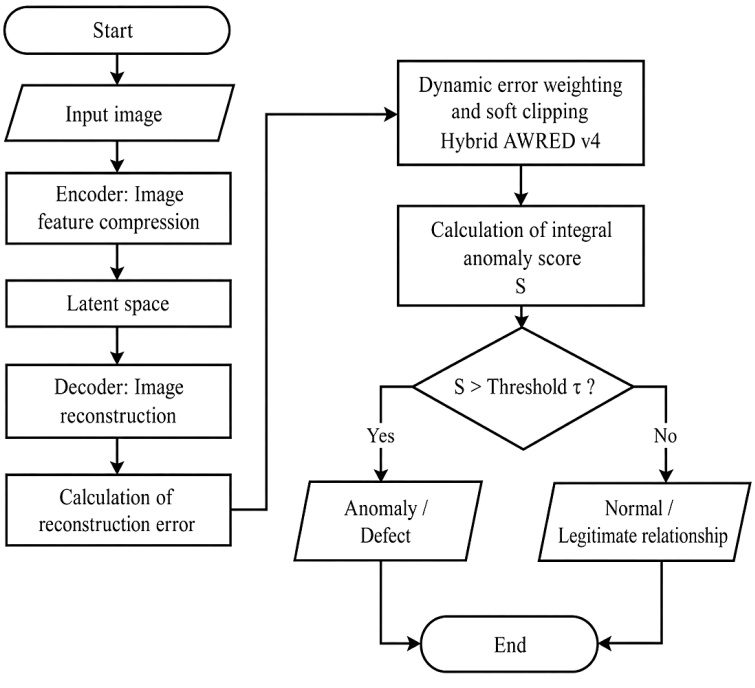

Fig. 2. Block diagram of the end-to-end architecture for visual anomaly detection

The general loss function was as follows

$$L_{total}(\theta,t)=L_{AWRED}(\theta,t)+\lambda_{reg}\|\theta\|_2^2, \quad (1)$$

where λ_{reg} – the L2-regularization coefficient, and the objective function of the proposed algorithm was calculated as the weighted sum of spatial deviations

$$L_{AWRED}(\theta,t)=\frac{1}{N}\sum_{i=1}^{N}w_i(t)\cdot x_i-\hat{x}_i^2. \quad (2)$$

The weight $w_i(t)$ was calculated through the differentiable logistic function

$$w_i(t)=\frac{1}{1+\exp\left(G(t)\cdot\left(e_i-Q_q\right)\right)}, \quad (3)$$

where e_i – the current reconstruction error; Q_q – the empirical quantile of the level q_{tail}.

The strategy for preventing overfitting on microsamples was implemented through a dynamic regularization schedule that defines the penalty steepness parameter G(t) and is based on the idea of curriculum learning [9]:

$$G(t)=\begin{cases}0, & t\le T_{warm},\\ G_{\max}\cdot\dfrac{t-T_{warm}}{T_{ramp}-T_{warm}}, & T_{warm}<t\le T_{ramp},\\ G_{\max}, & t>T_{ramp}.\end{cases} \quad (4)$$

In the context of end-to-end classification under data scarcity, the learning process was divided into three functional stages:

1. Structural absorption phase ($t \le T_{warm}$): $G(t) = 0$, sample weight $w_i \approx 1$. In conditions where the sample size is significantly limited, the network learned the basic topology of the object without applying penalties. This prevented the degradation of the optimization process and false exclusion of complex features.
2. Selective divergence phase ($T_{warm} < t \le T_{ramp}$). The linear increase in the penalty $G(t)$ initiated a smooth stratification of the error space. In this case, the model formed a metric interval between the norm and anomalies ($e_i > Q_q$), minimizing the influence of the latter on the update of the weights.
3. Final ranking phase ($t > T_{ramp}$). In this phase, the maximum penalty $G_{\max}$ was applied. The weight coefficient of defective areas was leveled ($w_i \to 0$). The autoencoder fully performed the function of a metric classifier and the model generated high-contrast maps of deviations, which allowed calculating the integral score (Anomaly Score) without involving external algorithms (in particular, SVM).

The final anomaly score (Anomaly Score) S for the test image was calculated as the aggregated sum of deviations

$$S(X)=\sum_{i=1}^{N}I(e_i>\tau_{test})\cdot e_i, \quad (5)$$

where $I(\cdot)$ – the indicator function, τ_{test} – the threshold that cut off the normal reconstruction noise.

This approach reduced the multidimensional optimization problem to a unidimensional ranking, which eliminated the dependence on SVM.

2.2. Methodology and conditions of experiments

The research materials were open sets of visual data, which were used to test the model's performance under conditions of limited training resources. The data source was the Caltech-101 set, freely available on the Kaggle platform [13].

The assessment of the stability of the models was carried out on three sets of visual data with spatial defects: set A (base scale, airplanes, $N = 800$), set B (limited sample, chandelier, $N = 107$) and set C (microsample, bass, $N = 54$). 8-fold independent testing was used to ensure statistical significance. The training of the neural network part of the model was carried out using the Adam optimizer. [14] was used as a methodological source on modern approaches to optimizing deep learning models.

The choice of methods was determined by the aim of research, which was to evaluate the effectiveness of the end-to-end Hybrid AWRED architecture in conditions of a shortage of training data and its comparison with the basic CNN + One-Class SVM scheme.

Several methods were used in the research. Their choice depended on how strong the class imbalance was and how small the number of reference data was. It was important to check how the model works in conditions of a limited sample.

First, the architectures were compared. For this, Hybrid AWRED was compared with the basic CNN + One-Class SVM scheme. In Hybrid AWRED, ranking was included directly in the training process, while in the CNN + One-Class SVM scheme, features were first formed separately and then the decision was made by an external classifier. Such a comparison made it possible to assess how both approaches remain stable with a small amount of data. The evaluation conditions for them were the same.

Additionally, local density estimation based on k-nearest neighbors was used. This made it possible to refine the description of the feature space. With a small amount of data, conventional distances do

ISSN-L 2664-9969; E-ISSN 2706-5448

not always reflect the placement of samples well. Therefore, the local environment of each example was taken into account separately. This helped to more accurately separate normal samples from abnormal ones.

The hyperparameters q_{tail} and G_{max} were tuned using Bayesian optimization [15]. A complete search for parameters in such a problem would be too expensive. Bayesian optimization allowed to reduce the number of runs.

One-Class SVM was used as the basic reference approach. Its inclusion in the comparison was appropriate, because this method was widely used in unsupervised anomaly detection problems. In such a formulation, the model is trained only on normal samples. Then, the normal class boundary was formed on their basis.

For quantitative analysis of the results, one overall accuracy was not enough. Due to the strong imbalance of classes, it could look high even when some of the defects remained undetected. Therefore, the result was evaluated by several indicators. *AUC-ROC* and *Average Precision* were used to assess the quality of the ranking. They allowed to see how well the model separated defective samples from normal ones even before choosing a specific threshold. After choosing a threshold, *Recall*, *Precision*, and *F1-Score* were additionally analyzed. Recall showed what fraction of real defects the model found. Precision showed how reliable the positive responses were. *F1-Score* reconciled these two estimates.

The Matthews correlation coefficient MCC [16] was calculated separately. It was useful for class imbalance because it took into account all elements of the error matrix. All indicators were calculated using standard formulas.

Numerical metrics show the overall result, but do not explain where exactly the model finds the defect. Therefore, t-SNE was used to analyze the latent feature space. Heat maps of the reconstruction error were also constructed. They made it possible to check whether the model responds to defect areas, and not to random background artifacts.

MATLAB R2025a from MathWorks, USA, was used for the experiments. Deep Learning Toolbox was used to build and train CNN architectures. Statistics and Machine Learning Toolbox was used for One-Class SVM, Bayesian optimization, and metric calculation.

3. Results and Discussion

3.1. Analysis of the limitations of the two-stage approach to anomaly detection at different sample scales

The results of the basic CNN + One-Class SVM scheme and the end-to-end Hybrid AWRED architecture were compared on samples of different sizes. To study anomaly detection, three data scarcity scenarios were created. This made it possible to reproduce different levels of availability of reference samples in production conditions. The obtained data were: sufficient volume ($N = 800$), limited sample ($N = 107$) and critical microsample ($N = 54$). Such a comparison allowed to trace how the stability of the models changed when the training set was reduced. Table 1 shows the averaged results of 8-fold testing on six main metrics for the three scenarios considered.

The data in Table 1 showed that for a sample size of $N = 800$, the CNN + AWRED architecture had higher values for the *Precision* (0.891 vs. 0.832), *F1-Score* (0.509 vs. 0.475) and *MCC* (0.391 vs. 0.332) metrics compared to the baseline architecture. This indicated a more balanced model performance when the data in quantitative terms was sufficient for stable training. At the same time, the *AUC-ROC* of the baseline approach remained slightly higher (0.881 vs. 0.875).

For a limited sample size of $N = 107$, the ratio of results changed. In this mode, CNN + AWRED outperformed CNN + SVM in terms of the integral metrics *AUC-ROC* (0.803 vs. 0.728) and *AP* (0.702 vs. 0.647). This meant that with a smaller amount of training data, the model was better at distinguishing abnormal samples relative to normal ones. However, there was no advantage over CNN + SVM in this scenario in terms of *Recall*, *Precision*, *F1-Score* and *MCC* metrics (Table 1).

For the microsample at $N = 54$, both architectures showed the same values in terms of *Recall*, *Precision*, *F1-Score* and *MCC* metrics. However, the baseline model remained slightly better in terms of *AUC-ROC* and *AP*. The results shown are 0.713 vs. 0.692 and 0.542 vs. 0.513, respectively. Thus, this result can be considered as an empirical limit for both approaches. This amount of data did not provide the models with enough information to support the stable formation of a norm representation.

These results showed that with a decrease in the amount of training data, the two-stage approach gradually lost its stability. For the sample at $N = 800$, the baseline scheme still maintained acceptable quality. At the same time, for $N = 107$, its integral indicators were already noticeably decreasing, and for the sample $N = 54$, both approaches actually reached the limit of efficiency. This confirmed the fact that the lack of a sufficient amount of data is a decisive factor for classical schemes with an external classifier.

3.2. Application of Hybrid AWRED v4 with dynamic error weighting without an external metric classifier

The Hybrid AWRED v4 architecture was evaluated not only by the final metrics. The model training process and the nature of its spatial response were analyzed separately. At the same time, both the dynamics of the loss function, *t-SNE* visualization, and heat maps of the reconstruction error were considered. Such an analysis was necessary in order to show how the model worked without an external metric classifier. And also whether its response was really concentrated in the zones of anomalies.

The adaptation of the model was accompanied by modification of the parameters in the loss space (Fig. 3). At the same time, the evaluation of the optimization dynamics showed that thanks to the weighting mechanism, Hybrid AWRED reached the convergence criterion faster than the base architecture – at approximately the 67th epoch versus the 75th epoch. This can be noted as a sign of faster adaptation to the data structure. Although this does not mean an advantage in all quality metrics.

In Fig. 4 [17], *t-SNE* visualization showed that in the latent space, normal samples formed a more compact region. Anomaly points were located more isolated. This was consistent with the better *AUC-ROC* and *AP* values for the $N = 107$ scenario. There, AWRED had a higher ranking quality. However, it should be noted that *t-SNE* was used as an auxiliary interpretation tool in this work.

Table 1

Comprehensive defect detection efficiency at different sample scales (mean ± Std)

Sampling	Method	AUC-ROC	AP	Recall	Precision	F1-Score	MCC
$N = 800$	CNN + SVM	0.881 ± 0.034	0.818 ± 0.055	0.333 ± 0.031	0.832 ± 0.076	0.475 ± 0.044	0.332 ± 0.076
	CNN + AWRED	0.875 ± 0.038	0.840 ± 0.050	0.356 ± 0.028	0.891 ± 0.071	0.509 ± 0.040	0.391 ± 0.071
$N = 107$	CNN + SVM	0.728 ± 0.106	0.647 ± 0.103	0.284 ± 0.076	0.781 ± 0.209	0.417 ± 0.111	0.265 ± 0.197
	CNN + AWRED	0.803 ± 0.081	0.702 ± 0.099	0.273 ± 0.084	0.750 ± 0.231	0.400 ± 0.123	0.236 ± 0.218
$N = 54$	CNN + SVM	0.713 ± 0.097	0.542 ± 0.054	0.300 ± 0.107	0.750 ± 0.267	0.429 ± 0.153	0.280 ± 0.253
	CNN + AWRED	0.692 ± 0.099	0.513 ± 0.080	0.300 ± 0.107	0.750 ± 0.267	0.429 ± 0.153	0.280 ± 0.253

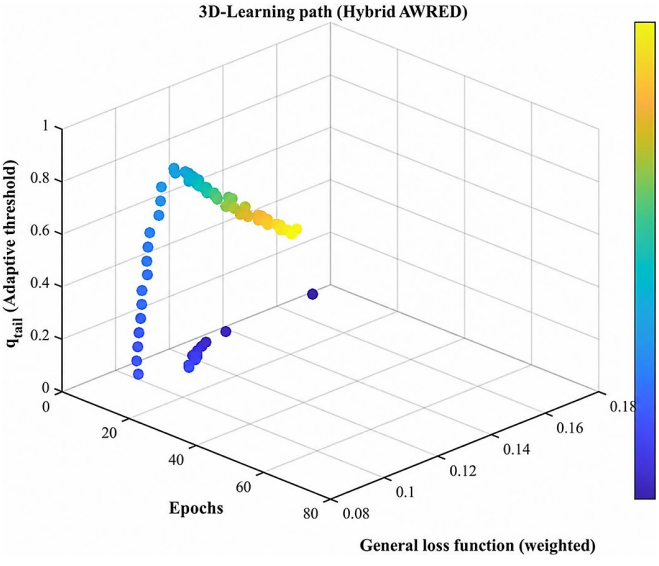

Fig. 3. 3D learning trajectory of Hybrid AWRED architecture

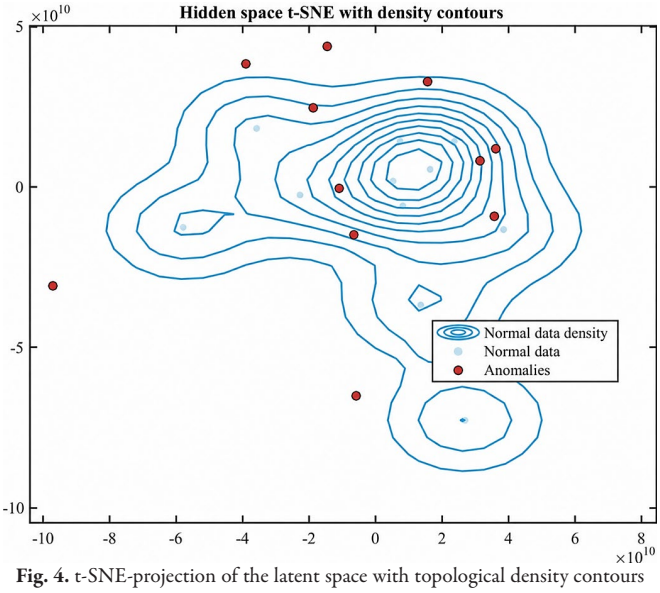

Fig. 4. t-SNE-projection of the latent space with topological density contours

The analysis of the heat maps in Fig. 5 showed that for Hybrid AWRED the response zones in the defect region were more pronounced than for the base architecture. This suggested that the dynamic error weighting helped the model to better focus on local deviations. However, these conclusions regarding spatial localization are mostly qualitative in nature.

Visual analysis showed results that were consistent with numerical metrics. Thus, it indicated the feasibility of switching to an end-to-end architecture in conditions of limited data. However, it should be borne in mind that the obtained localization advantages were preliminary and therefore required further quantitative verification.

The use of Hybrid AWRED v4 showed that the dynamic error weighting mechanism can be integrated directly into the training process. This made it possible to abandon the external metric classifier and maintain the performance of the model even in the absence of sufficient data.

3.3. Experimental evaluation of defect ranking and localization and comparison with CNN + SVM

The advantage of Hybrid AWRED was most noticeable in the quality of anomaly ranking with a limited training resource. Unlike the basic CNN + One-Class SVM scheme, in this architecture the anomaly estimate was formed directly in the reconstruction error space, without a separate external classifier. This is especially important when the training sample size is no longer sufficient for stable formation of the latent representation.

With sufficient data, Hybrid AWRED gave a more balanced result at the final solution level. When the sample size was reduced, the advantage of the proposed method shifted towards integral metrics. That is, in the data deficit mode ($N = 107$), the model better retained the ability to separate anomalous samples from normal ones by the anomaly score, even if this was not accompanied by an advantage in all threshold indicators at the same time.

To clearly present these results, Fig. 4 shows a generalized visualization of the experimental comparison of models.

The panels in Fig. 6, *a*, *b* show the change in the integral metrics *AUC-ROC* and *AP* when the sample size is reduced.

The panels in Fig. 6, *c*, *d* detail the behavior of the models for the transitional sample $N = 107$ through PR curves and the distribution of anomaly scores.

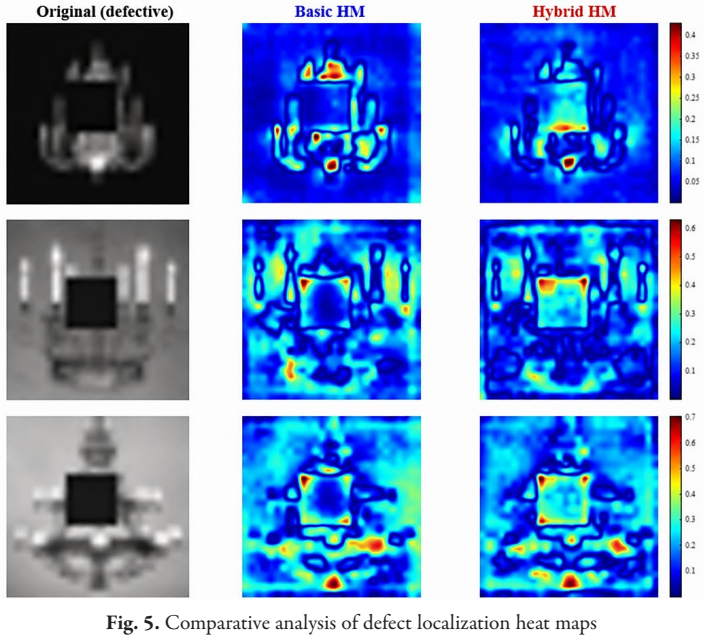

Fig. 5. Comparative analysis of defect localization heat maps

ISSN-L 2664-9969; E-ISSN 2706-5448

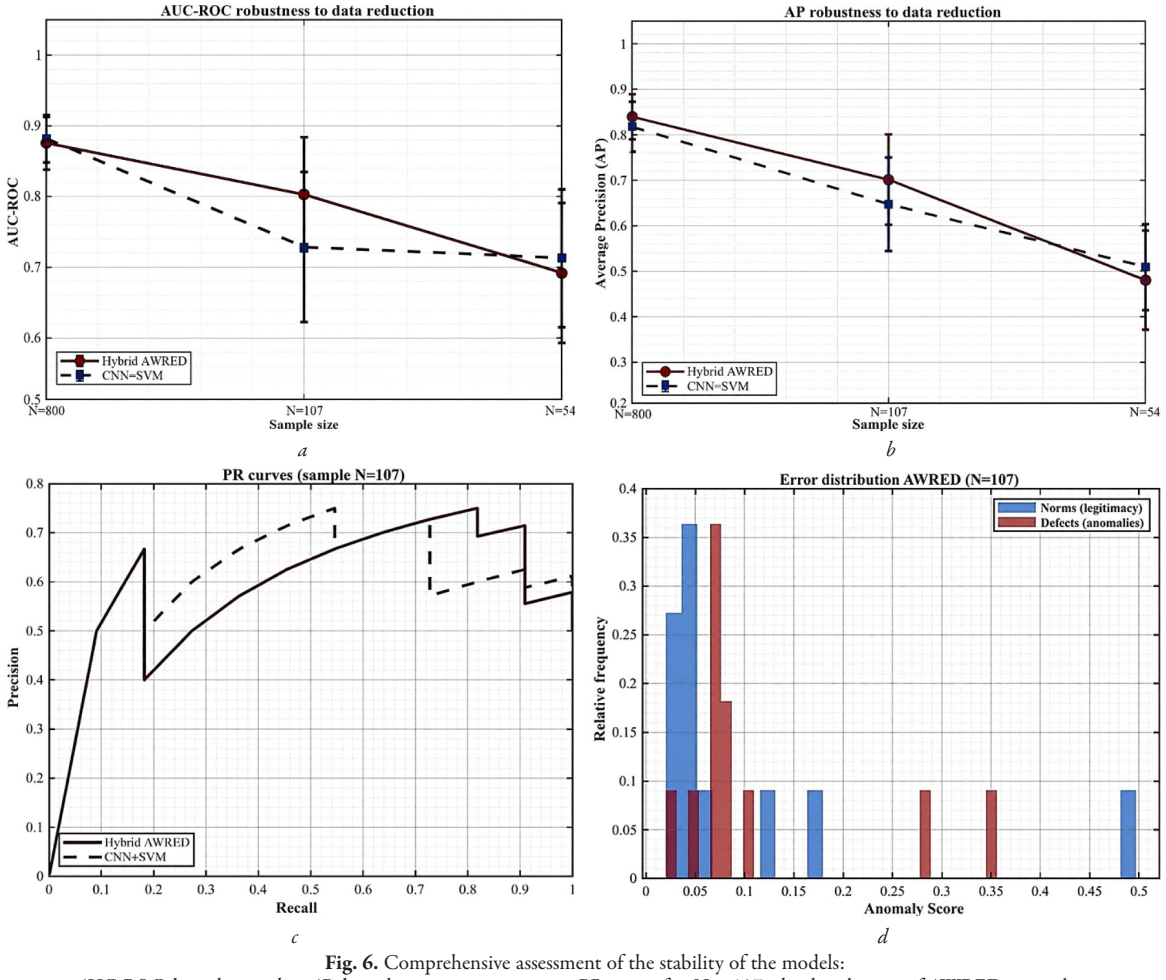

Fig. 6. Comprehensive assessment of the stability of the models:
a – *AUC-ROC* degradation; *b* – *AP* degradation; *c* – comparative *PR* curves for $N = 107$; *d* – distribution of AWRED anomaly scores

Such visualization complemented the numerical results given in Table 1. It was clearly seen that Hybrid AWRED gave a more distinct spatial separation of defective areas on heat maps. In general, the results gave grounds to consider Hybrid AWRED as a promising approach for detecting visual defects with a shortage of training data, although for micro-samples ($N = 54$) an approach to the efficiency limit was observed.

3.4. Discussion of the obtained results

The results obtained during the experiments required a comprehensive discussion to determine their place in the general problems of computer vision. The transition of the advantage from threshold metrics ($N = 800$) to integral indicators of *AUC-ROC* and *AP* ($N = 107$), which is given in Table 1 and illustrated in Fig. 6, *a*, *b*, indicated that under data scarcity conditions, end-to-end ranking turned out to be more stable than the two-stage scheme. In the classical CNN + SVM model, the latent space on small samples degraded, which led to errors of the external classifier. Its rejection in favor of the Hybrid AWRED loss function allowed to preserve the quality of sorting (ranking) of normal and abnormal samples. However, the alignment of the indicators of both models on the sample $N = 54$, which is also visible from the data in Table 1 and Fig. 6, *a*, *b*, means that under critical data scarcity, the neural network lost the ability to learn even the basic structural regularities of the norm.

A distinctive feature of the proposed approach was the shift of decision-making from the latent feature space directly to the reconstruction error space using dynamic weighting and an adaptive threshold. This allowed to obtain a double effect: acceleration of model convergence (67 epochs versus 75 in the base model, Fig. 3) and generation of visually more contrasting and clear zones of response to defects on heat maps without using additional clustering methods (Fig. 5).

However, the research had certain limitations. First, the architecture demonstrated reaching the limit of its efficiency on microsamples of about 50 images, after which the recognition quality significantly degraded regardless of the selected loss function, which was confirmed by the results of Table 1 and Fig. 6, *a*, *b*. Secondly, the conclusions regarding the improvement of spatial localization of defects in this work were based mainly on qualitative visual analysis of heat maps (Fig. 5), without involving quantitative pixel segmentation metrics (such as Intersection over Union).

Further development of the research should be focused on two directions. The first is the combination of the end-to-end Hybrid AWRED architecture with generative methods (e. g., diffusion models) for synthetic expansion of microsamples, which will allow overcoming the efficiency barrier at $N = 54$, identified by the results of Table 1 and Fig. 6, *a*, *b*. The second is the implementation of metrics for quantitative assessment of spatial localization of defects, which will allow mathematically confirming the visual contrast of the created heat maps, shown in Fig. 5.

4. Conclusions

1. The limitations of the two-stage approach to anomaly detection at different sample scales were analyzed. As the amount of training data decreased, the classical CNN + One-Class SVM scheme worked less stably. For a sample of $N = 800$, it still gave an acceptable result, but already at $N = 107$ the integral indicators decreased noticeably. For $N = 54$, both approaches actually reached the limit of their capabilities. This meant that in conditions of data scarcity, the two-stage scheme became unstable, since the quality of the latent representation and subsequent separation already strongly depended on the sample size.

2. The research used the end-to-end Hybrid AWRED v4 architecture with dynamic error weighting without the involvement of an external metric classifier. During training, the difference was noticeable already in the number of epochs. Hybrid AWRED v4 reached convergence at approximately the 67th epoch, the basic configuration – at approximately the 75th. This fact in itself does not mean the complete superiority of the model, but for a small sample it is indicative. In this variant, a separate metric classifier was not added after the CNN features. One-Class SVM remained only in the basic scheme for comparison. The anomaly was taken from the reconstruction error, and not from a separate boundary in the feature space. That is, it was determined by how much the reconstructed image differed from the input one. If the reconstruction error was larger, the sample received a higher anomaly score. This did not simplify the task, but it reduced the number of intermediate stages where instability could appear with a small sample. On *t-SNE* visualization, the anomalous points were separated more noticeably, although such visualization cannot be considered an independent proof of quality. Heat maps gave a more practical sign: in the areas of defects, the response was more contrasting. Therefore, the effect of dynamic error weighting was manifested not only in the metrics. But it should not be exaggerated: in this case, it is primarily about better localization of problem areas, and not about completely eliminating the limitations of a small sample.

3. Comparison with CNN + One-Class SVM showed that the result of Hybrid AWRED depended on the size of the training sample. For $N = 800$, the CNN + AWRED model had better values *Precision* = 0.891 vs. 0.832, *F1-Score* = 0.509 vs. 0.475 and *MCC* = 0.391 vs. 0.332. But for *AUC-ROC*, the basic scheme remained slightly higher: 0.881 vs. 0.875. Therefore, for this sample, it is not possible to speak about the complete superiority of Hybrid AWRED in all indicators. Rather, this indicates a better balance after choosing the threshold.

For $N = 107$, Hybrid AWRED was already better in terms of integral metrics: *AUC-ROC* = 0.803 vs. 0.728 and *AP* = 0.702 vs. 0.647. Here, the quality of anomaly ranking was better preserved with a limited training resource. For a micro-sample of $N = 54$, both architectures gave the same *Recall*, *Precision*, *F1-Score*, and *MCC* values, but for *AUC-ROC* = 0.692 and *AP* = 0.513, Hybrid AWRED was already slightly inferior to the base model, in which these values were 0.713 and 0.542. Therefore, Hybrid AWRED can be considered a promising approach for detecting visual anomalies with a shortage of training data. For very small samples, however, the approach to the efficiency limit was already noticeable, and without data expansion or additional training support, further improving the quality turned out to be a difficult task.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Financing

The research was performed without financial support.

Data availability

Manuscript has no associated data.

Use of artificial intelligence

The authors used Gemini 3 to check the grammar, spelling, and punctuation of individual Ukrainian and English fragments of the manuscript without changing their content. All corrections were checked manually by the authors. The content of the manuscript, the interpretation of the results, and the final conclusions of the article were determined independently by the authors.

Authors' contributions

Tymur Dovzhenko: Conceptualization, Methodology, Software, Data curation, Writing – original draft; ***Kamila Storchak***: Supervision, Formal analysis, Writing – review and editing.

References

1. Cao, Y., Xiang, H., Zhang, H., Zhu, Y., Ting, K. M. (2025). Anomaly Detection Based on Isolation Mechanisms: A Survey. *Machine Intelligence Research, 22 (5),* 849–865. https://doi.org/10.1007/s11633-025-1554-4
2. Tao, X., Gong, X., Zhang, X., Yan, S., Adak, C. (2022). Deep Learning for Unsupervised Anomaly Localization in Industrial Images: A Survey. *IEEE Transactions on Instrumentation and Measurement, 71,* 1–21. https://doi.org/10.1109/tim.2022.3196436
3. Mehta, D., Klarmann, N. (2023). Autoencoder-Based Visual Anomaly Localization for Manufacturing Quality Control. *Machine Learning and Knowledge Extraction, 6 (1),* 1–17. https://doi.org/10.3390/make6010001
4. Paolini, D., Dini, P., Soldaini, E., Saponara, S. (2025). One-Class Anomaly Detection for Industrial Applications: A Comparative Survey and Experimental Study. *Computers, 14 (7),* 281. https://doi.org/10.3390/computers14070281
5. Saeedi, J., Giusti, A. (2022). Anomaly Detection for Industrial Inspection using Convolutional Autoencoder and Deep Feature-based One-class Classification. *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 85–96. https://doi.org/10.5220/0010780200003124
6. Yang, M., Liu, J., Yang, Z., Wu, Z. (2024). SLSG: Industrial image anomaly detection with improved feature embeddings and one-class classification. *Pattern Recognition, 156,* 110862. https://doi.org/10.1016/j.patcog.2024.110862
7. Liu, J., Xie, G., Wang, J., Li, S., Wang, C., Zheng, F., Jin, Y. (2024). Deep Industrial Image Anomaly Detection: A Survey. *Machine Intelligence Research, 21 (1),* 104–135. https://doi.org/10.1007/s11633-023-1459-z
8. Li, Z., Yan, Y., Wang, X., Ge, Y., Meng, L. (2025). A survey of deep learning for industrial visual anomaly detection. *Artificial Intelligence Review, 58 (9).* https://doi.org/10.1007/s10462-025-11287-7
9. Wang, X., Chen, Y., Zhu, W. (2021). A Survey on Curriculum Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44 (9).* https://doi.org/10.1109/tpami.2021.3069908
10. Dovzhenko, T. (2026). Hybrid awred: synergy of adaptive reconstruction and topological clustering for anomaly detection in multimodal data. *Zviazok, 1,* 80–88. https://doi.org/10.31673/2412-9070.2026.017405
11. Dovzhenko, T. P., Zinchenko, O. V. (2026). Robustness of Deep Intrusion Detection Models Under Massive Cyberattacks: Stress-Testing and Architectural Features of Hybrid AWRED. *Telecommunication and Information Technologies, 90 (1),* 199–207. https://doi.org/10.31673/2412-4338.2026.019019
12. Dovzhenko, T. (2026). Topological anchoring and adaptive penalties: the HYBRID AWRED architecture for defect recognition in contaminated visual data. *Connectivity, 180 (2),* 62–71. https://doi.org/10.31673/2412-9070.2026.027603
13. Saha, B. (2022). *Caltech-101 dataset.* Available at: https://www.kaggle.com/datasets/imbikramsaha/caltech-101
14. Belhadri, A., Benchennane, I. (2025). Optimizing deep learning models: A review. *Multiagent and Grid Systems, 21 (2),* 73–95. https://doi.org/10.1177/15741702251370052
15. Wang, X., Jin, Y., Schmitt, S., Olhofer, M. (2023). Recent Advances in Bayesian Optimization. *ACM Computing Surveys, 55 (13s),* 1–36. https://doi.org/10.1145/3582078
16. Chicco, D., Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining, 16 (1).* https://doi.org/10.1186/s13040-023-00322-4
17. Jung, S., Dagobert, T., Morel, J.-M., Facciolo, G. (2024). A Review of t-SNE. *Image Processing on Line, 14,* 250–270. https://doi.org/10.5201/ipol.2024.528

✉ ***Tymur Dovzhenko,*** *PhD, Associate Professor, Department of Software Engineering, State University of Information and Communication Technologies, Kyiv, Ukraine, e-mail: timurdov@ukr.net, ORCID: https://orcid.org/0000-0002-0352-8391*

Kamila Storchak, *Doctor of Technical Sciences, Professor, Head of Department of Information Systems and Technologies, State University of Information and Communication Technologies, Kyiv, Ukraine, ORCID: https://orcid.org/0000-0001-9295-4685*

✉ *Corresponding author*