

Dual-Scoring Knowledge Distillation for Anomaly Detection in Electroluminescence PV Images using Transformer-Guided Efficient-UNet

Tolga Can Ceylan

Abstract

As photovoltaic (PV) systems play a critical role in the global energy supply, accurate and efficient anomaly detection in electroluminescence (EL) images of solar panels has become increasingly important. This paper presents a novel anomaly detection framework that leverages Knowledge Distillation (KD) and integrates reconstruction errors into a Deep Support Vector Data Description (Deep SVDD) module for enhanced performance. The framework utilizes a ResNet-101-based Convolutional Attention Autoencoder as the teacher model and an Efficient-UNet Transformer as the student model. The Efficient-UNet Transformer combines local and global feature extraction capabilities with transformer layers to capture long-range dependencies, making it highly efficient and effective. The teacher and student models' reconstruction error rates are added together and sent to the Deep SVDD module. This lets the system learn small representations of normal data and find different kinds of anomalies without needing large, labeled datasets. Experimental results on both public (ELPV) and private industrial (Factory) datasets demonstrate the effectiveness of the proposed approach. The proposed approach achieved remarkable results, achieving an accuracy of 0.97 on the ELPV dataset and 0.88 on the Factory dataset. These performance metrics represent significant improvements compared to state-of-the-art methods and conventional models.

Keywords: Deep Learning, Solar Panel Defect Detection, Knowledge Distillation, Anomaly Detection

Date of Submission: 02-11-2025

Date of acceptance: 11-11-2025

I. Introduction

The increasing global reliance on solar photovoltaic (PV) systems as a sustainable and clean energy source has heightened the need for effective and efficient monitoring methods to ensure their optimal operation. Problems like micro-cracks, hotspots, and partial shading can greatly reduce the amount of power that is produced, speed up the breakdown of PV panels, cause big financial losses, and shorten the life of systems. Traditional methods for detecting anomalies, such as visual inspections and statistical analyses, often struggle to identify subtle or complex issues effectively. This challenge becomes even more pronounced in large-scale deployments involving diverse and complex datasets. However, recent advancements in deep learning have demonstrated significant potential for enhancing anomaly detection in photovoltaic (PV) systems. Convolutional Neural Networks (CNNs) have been widely employed for their ability to extract features from image data [1], [2], making them suitable for detecting faults in PV panels. However, these methods are often limited in their ability to simultaneously capture both local and global patterns, reducing their robustness against varying environmental conditions and noise.

Moreover, the effectiveness of CNN-based approaches typically relies on large volumes of labelled data, which are often difficult to obtain in real-world PV monitoring scenarios. Transformer models were first created for processing natural language. They have shown a lot of promise in computer vision tasks because they can find long-range dependencies and contextual relationships in data. The self-attention mechanisms in Transformers effectively model complex interactions within image data, making them particularly adept at detecting subtle anomalies that may not be apparent through traditional methods. Recent studies have leveraged Transformer architectures to enhance the accuracy of anomaly detection in PV systems by treating image data points as analogous to words in a language model, thereby capturing intricate dependencies among them [3], [4].

Despite significant advancements, developing a model that is both accurate and lightweight for edge computing environments, such as industrial devices used in real-time photovoltaic (PV) system monitoring, remains a challenge. Models optimized for low power consumption and fast inference times are often necessary in these settings due to their constrained computational

resources. To address this, our study introduces a lightweight framework based on knowledge distillation, where a ResNet-101 convolutional attention autoencoder serves as the teacher model, and an Efficient-UNet Transformer with Deep SVDD-guided anomaly scoring acts as the student. This approach effectively combines the strengths of the UNet architecture and Transformer capabilities, guided by an anomaly scoring network, to enhance the detection of anomalies in electroluminescence (EL) images of solar panels—while being well-suited for deployment on edge devices. The Efficient-UNet Transformer integrates an EfficientNet-based encoder to capture both local and global features with minimal computational overhead. The embedded Transformer layer further enhances the model's ability to learn long-range dependencies, critical for identifying subtle and complex defect patterns in PV cells. Through knowledge distillation, the student model inherits the teacher's learned representations, improving feature extraction while maintaining computational efficiency. Additionally, a Deep SVDD-guided scoring network enables the model to form a compact representation of normal data in latent space. By combining reconstruction loss and Deep SVDD loss, this dual-scoring mechanism provides robust anomaly detection, even under variable environmental conditions. This integrated and lightweight approach improves the robustness, accuracy, and adaptability of anomaly detection in PV systems, making it ideal for edge computing in industrial settings. Validation on benchmark datasets demonstrates the proposed framework's performance, which matches or exceeds state-of-the-art methods while maintaining computational efficiency. This contribution offers a novel and practical solution for PV system monitoring, enhancing the reliability and operational efficiency of solar energy systems in real-world scenarios.

Key contributions of this study are as follows:

- Introduced the Efficient-UNet + Deep SVDD framework tailored for anomaly detection in photovoltaic systems.
- Bridged the server-based teacher model (ResNet-101 Convolutional Attention Autoencoder) with the edge-based student model, ensuring high efficiency and performance.
- Demonstrated the practical application of the proposed methodology by evaluating it on the Factory dataset collected from an industrial factory environment.

These contributions highlight the scalability, efficiency, and practicality of the proposed method for identifying anomalies in PV systems, which

facilitates the integration of machine learning into industrial quality control processes.

1.1 Related Work

Recently, machine learning and deep learning techniques have made big steps forward in finding problems in photovoltaic (PV) systems. These techniques have made fault detection much more accurate and reliable. Anomaly detection in electroluminescence (EL) images is particularly critical for ensuring the quality and reliability of PV modules. A wide range of methods, from traditional image processing to cutting-edge deep learning approaches, have been explored to address this challenge.

Traditional methods like Independent Component Analysis (ICA) and matched filtering have been utilized to detect defects such as cracks, scratches, and finger interruptions in PV modules. Generally, these methods rely on handcrafted features and have limitations in effectively differentiating defect types [4]. Researchers have employed different image processing techniques, like anisotropic diffusion filtering, to classify micro-cracks in EL images [5]. Furthermore, several studies have employed machine learning techniques for anomaly detection in PV systems. Laguna et al. [6] proposed a method based on Recursive Least Squares (RLS) algorithms to detect PV system faults with minimal data requirements. Fadhel et al. [7] used Principal Component Analysis (PCA) combined with K-Nearest Neighbours (KNN) to detect and classify shading faults, demonstrating effectiveness in handling specific PV anomaly types. Chandio et al. [8] focused on a machine learning-based multiclass anomaly detection approach using Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), and Support Vector Machine (SVM) classifiers to enhance detection accuracy in hybrid active distribution networks.

Researchers have extensively explored deep learning methods for anomaly detection in PV systems. Khalil et al. [9] introduced a Transformer-based model for robust fault prediction in PV systems, leveraging attention mechanisms to learn dependencies among data points, which is advantageous when dealing with unclear fault patterns. Hwang et al. [10] proposed a Vision Transformer (ViT) model for detecting faults in PV modules using infrared thermography (IR) images, achieving superior performance over traditional deep learning methods. Chang et al. [11] developed a framework for PV cell anomaly detection using Scale Distribution Alignment Learning and

Multiscale Linear Attention (MLA-SDAL). This framework utilizes multi-head linear attention for efficient feature extraction and employs scale distribution alignment for robust anomaly detection, demonstrating improved adaptability to complex data distributions. Kang et al. [12] applied a weakly supervised learning approach to detect PV cell defects using module-level annotations, significantly reducing the annotation costs associated with traditional cell-level defect detection. In the same way, Tang et al. [13] proposed a deep learning model for automatic defect identification, which demonstrated high accuracy by using data augmentation techniques. Similarly, Su et al. [14] introduced a large-scale open-world dataset for anomaly detection, highlighting the importance of using diverse datasets to enhance model generalization. In another study, Liu et al. [15] proposed an efficient CNN-based detector for photovoltaic module cell defects using EL images, which employs a lightweight model based on EfficientNet-B0 and introduces a Graph Channel Attention Module (GCAM) to enhance feature representation by modelling global information. The use of Contrast Limited Adaptive Histogram Equalization (CLAHE) further improves image contrast, making defect detection more accurate. This method demonstrated superior performance, achieving an accuracy of 97.81%, F1-score of 97.64%, and MCC of 97.32% on the PVEL dataset, outperforming state-of-the-art methods across various metrics. Also, Al-Otum [16] developed a CNN-based deep learning approach for classifying anomalies in EL images of solar PV modules. This study explored three different models: two based on transfer learning with pretrained SqueezeNet and GoogleNet and a lightweight CNN model (LwNet). Experimental validation using the ELPV dataset demonstrated high performance, with average accuracies of 94.6%, 93.95%, and 96.2% for GoogleNet, SqueezeNet, and LwNet, respectively. The LwNet model showed superior classification performance and time-saving efficiency compared to the other models. The SeMaCNN model [17], which combines a semi-supervised learning approach using a PaDiM-based feature extractor and a shallow classifier based on the ResNet18 architecture, achieved an F1 score of 95.8% and precision of 96.9% on a dataset comprising 68,748 EL images from a heterojunction solar cell manufacturing plant. This demonstrates the model's capability to perform well in industrial environments.

Advanced neural network architectures, such as transformer neural networks, have also been employed for anomaly detection in PV systems.

David et al. [18] introduced a Time Series Transformer (TST) for end-to-end islanding detection, simplifying the modelling process by automating feature extraction and outperforming other machine learning methods in terms of accuracy and detection time. Zhao et al. [19] proposed a data-driven solution for anomaly detection and classification in large-scale PV systems that does not require additional equipment or non-SCADA data collection. Their approach integrates a hierarchical context-aware anomaly detection method using unsupervised learning with a multimodal anomaly classification method, demonstrating effectiveness and efficiency in real-world deployments across two large-scale solar farms. In another similar study, Zhao et al. [20] proposed the PV Cell Defects Detection Transformer (PD-DETR) to enhance the detection of small-scale defects in PV cell images, addressing challenges like slow convergence and complex backgrounds. This model combines a hybrid feature module with both one-to-one and one-to-many set matching strategies, achieving a 64.7% accuracy rate on the PVEL-AD dataset, thereby improving defect detection without relying on manual post-processing. Furthermore, Dwivedi et al. [21] proposed an attention-based deep learning model using a Vision Transformer (ViT) to identify surface defects on solar PV panels and wind turbine blades from high-resolution images, demonstrating superior performance over well-known pre-trained models. Hybrid models that integrate deep learning with traditional anomaly detection techniques have also shown promise. CNN models combined with Mahalanobis distance metrics have been useful for dealing with uneven datasets that only have a few labeled examples [22]. Additionally, the application of the CLAHE algorithm, as utilized by Liu et al. [23], has proven useful in enhancing image contrast, thus improving the detection of subtle defects.

Several advanced techniques have emerged to address challenges in PV anomaly detection. Bommies et al. [24] used supervised contrastive learning for anomaly detection in infrared images of PV modules, highlighting the potential of contrastive learning to enhance model robustness and performance. Additionally, Oliveira et al. [25] applied deep learning for anomaly detection in voltage waveform distortion due to geomagnetically induced currents, demonstrating the versatility of deep learning methods in handling diverse anomaly detection tasks. Zarghami et al. [26] employed a concurrent PV production and consumption load forecasting model using a CT-Transformer, illustrating the application of Transformer-based architectures to improve prediction accuracy in PV

systems. Additionally, Tsai et al. [18] focused on islanding detection using Transformer neural networks, underscoring the potential of Transformer-based architectures for real-time anomaly detection in PV systems. Also, in another study automated solar panel monitoring has advanced with the development of a Multi-Branch Spatial Pyramid Dynamic Graph Convolutional Neural Network (MB SPDG-CNN) [27], effectively integrating RGB and thermal imaging for superior fault detection accuracy. Some studies have employed unsupervised learning techniques for anomaly detection by leveraging reconstruction error rates.

Despite advancements in anomaly detection for photovoltaic (PV) systems, significant challenges persist in integrating Transformer models with robust techniques that minimize reliance on labelled datasets. Many existing methods rely on single approaches, such as traditional supervised learning or standard deep learning models, which often require extensive labelled data and struggle to generalize across varying conditions. Addressing these limitations, N. Zhu et al. [28] proposed an unsupervised adversarial training framework with feature reconstruction constraints for anomaly detection in crystalline silicon solar cells, demonstrating strong performance using EL imaging without labelled data. Similarly, C. Shou et al. [29] introduced a hybrid approach combining Generative Adversarial Networks (GANs) and Auto-Encoders (AEs), achieving significant accuracy rates. Their use of Structural Similarity Index (SSIM) and Mean Square Error (MSE) further enhanced detection by addressing image distortions and improving feature representation.

Building on these foundations, our work introduces a novel Knowledge Distillation framework that significantly extends and enhances prior approaches. Unlike existing studies that focus solely on adversarial feature reconstruction or emphasize image distortion metrics, our framework integrates a ResNet-101-based Convolutional Attention Autoencoder as the teacher model and a lightweight Efficient-UNet Transformer as the student model, facilitating efficient knowledge transfer. Moreover, we employ a dual anomaly scoring mechanism that combines reconstruction errors with Deep Support Vector Data Description loss, achieving enhanced robustness and improved accuracy in anomaly detection.

These improvements ensure superior scalability, computational efficiency, and real-time applicability, making our framework particularly

suitable for edge computing in PV system monitoring. Experimental results on public (ELPV) and private (Factory) datasets further validate the framework's effectiveness, demonstrating superior performance compared to existing methods while significantly reducing computational demands, thereby offering a scalable and practical solution for industrial applications.

II. Methodology

The architecture of the proposed knowledge distillation-based approach is detailed in the following subsections.

2.1 Student Model

The student model is based on an Efficient-UNet Transformer, incorporating Deep SVDD-guided anomaly scoring. The subsequent sections further explain these components.

2.1.1. Efficient-UNet Transformer

The proposed architecture, illustrated in Figure 1, employs the EfficientNet model as the backbone for the U-Net encoder, integrating a Transformer module to capture global context within the student model. This U-Net model, enhanced with Efficient Net MBConv blocks, is specifically designed to improve anomaly detection in EL images of solar panels, which is a critical task for ensuring the quality and efficiency of PV systems. The model leverages MBConv blocks, which employ depth wise separable convolutions and squeeze-and-excitation (SE) layers to efficiently capture both local and global features in EL images.

The encoder in the model consists of a series of MBConv blocks that progressively reduce the spatial dimensions of the input images while retaining essential information about potential defects, such as cracks, hotspots, or inactive regions in the solar cells. These blocks include an expansion phase that increases the number of channels, followed by depth wise convolutions to learn spatial patterns, and SE layers that enhance feature selection, making the model particularly sensitive to subtle variations indicative of anomalies. At the core of the model is a high-capacity bottleneck MBConv block, which serves as a feature-rich bridge between the encoder and decoder, designed to capture complex patterns and relationships within the EL images and to ensure that vital information about potential defects is preserved during feature compression.

The input to the model is an EL image of a solar panel cell, denoted as $X \in \mathbb{R}^{H \times W \times C}$, where H and W are the height and width of the image, and

C is the number of color channels (typically 1 for grayscale images). The input image is divided into non-overlapping patches of size $s \times s$. Each patch is represented as $P_i \in \mathbb{R}^{\frac{H}{s} \times \frac{W}{s} \times \frac{C}{s}}$, where i ranges from 1 to $\frac{H \times W}{s^2}$. This step enables multi-scale feature learning by processing the image at various resolutions. The Efficient Net U-Net encoder processes each patch through its network layers, consisting of a series of inverted residual blocks with depth wise separable convolutions and SE modules, and extracts hierarchical features from each patch using convolutional layers with down-sampling operations:

$$f_i = E(P_i), \text{ for } i \in [1, \frac{H \times W}{s^2}]. \quad (1)$$

The output f_i represents local feature maps, capturing fine-grained spatial details within each patch. The encoder progressively reduces the spatial dimensions while increasing the feature depth to extract higher-level features.

To further enhance the model's performance for anomaly detection in EL images, a Transformer block with a Multi-Head Attention (MHA) mechanism is incorporated into the bottleneck. This Transformer block enables the model to capture long-range dependencies and complex relationships in the features extracted by the encoder, which is particularly useful for detecting subtle anomalies in EL images. The Transformer module T is applied to the sequence of encoded features to capture long-range dependencies and contextual information:

$$t_i = T(f_i), \text{ for } i \in [1, \frac{H \times W}{s^2}]. \quad (2)$$

The Transformer with multi-head self-attention mechanism to compute the attention weights for each pair of patches, allowing the model to understand global patterns across the entire image:

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V, \quad (3)$$

Here, Q, K, V represent the query, key, and value matrices derived from the encoded features, and d_k is the dimensionality of the key vectors. The decoder D reconstructs the segmentation map using both the local feature maps f_i from the encoder and the global context-enhanced features t_i from the Transformer:

$$o_i = D(t_i, f_i), \quad (4)$$

The decoder mirrors the encoder's structure, utilizing up-sampling layers and additional MBConv blocks to restore the spatial resolution of the input images. Skip connections between the encoder and decoder layers help retain high-resolution details, which are crucial for accurately identifying the location and extent of anomalies. The final output layer produces a segmentation mask that highlights potential defects in the EL images, enabling precise anomaly detection. By integrating Efficient Net's refined MBConv blocks, the U-Net model balances high accuracy and low computational cost, making it well-suited for real-time monitoring and quality control of solar panels. The detailed architecture of the proposed approach is summarized in Table 1.

Tabel 1. Efficient U-Net with Transformer Architecture and Parameters.

Type	Output Shape	#Param	Connected To
InputLayer	(None, 128, 128, 3)	0	-
MBConv (6× Expansion)	(None, 128, 128, 32)	936	Input
MaxPooling2D	(None, 64, 64, 32)	0	MBConv Block 1
MBConv (6× Expansion)	(None, 64, 64, 64)	16,546	MaxPooling2D_1
MaxPooling2D	(None, 32, 32, 64)	0	MBConv Block 2
MBConv (6× Expansion)	(None, 32, 32, 128)	65,554	MaxPooling2D_2
MaxPooling2D	(None, 16, 16, 128)	0	MBConv Block 3
MBConv (6× Expansion)	(None, 16, 16, 256)	262,912	MaxPooling2D_3
MaxPooling2D	(None, 8, 8, 256)	0	MBConv Block 4
MBConv (6× Expansion)	(None, 8, 8, 512)	1,050,624	MaxPooling2D_4
Reshape	(None, 64, 512)	0	Bottleneck MBConv Block
Transformer Block (MHA)	(None, 64, 512)	526,848	Reshape
Reshape	(None, 8, 8, 512)	0	Transformer Block (MHA)
Conv2DTranspose	(None, 16, 16, 256)	524,544	Transformer Block (Bottleneck)
Concatenate	(None, 16, 16, 512)	0	Up-sample 1, MBConv Block 4
MBConv (6× Expansion)	(None, 16, 16, 256)	1,042,048	Concatenate 1
Conv2DTranspose	(None, 32, 32, 128)	131,200	MBConv Block 5
Concatenate	(None, 32, 32, 256)	0	Up-sample 2, MBConv Block 3
MBConv (6× Expansion)	(None, 32, 32, 128)	262,912	Concatenate 2
Conv2DTranspose	(None, 64, 64, 64)	32,832	MBConv Block 6
Concatenate	(None, 64, 64, 128)	0	Up-sample 3, MBConv Block 2

MBConv (6× Expansion)	(None, 64, 64, 64)	65,554	Concatenate 3
Conv2DTranspose	(None, 128, 128, 32)	8,224	MBConv Block 7
Concatenate	(None, 128, 128, 64)	0	Up-sample 4, MBConv Block 1
MBConv (6× Expansion)	(None, 128, 128, 32)	16,546	Concatenate 4
Conv2D (Output)	(None, 128, 128, 1)	33	MBConv Block 8

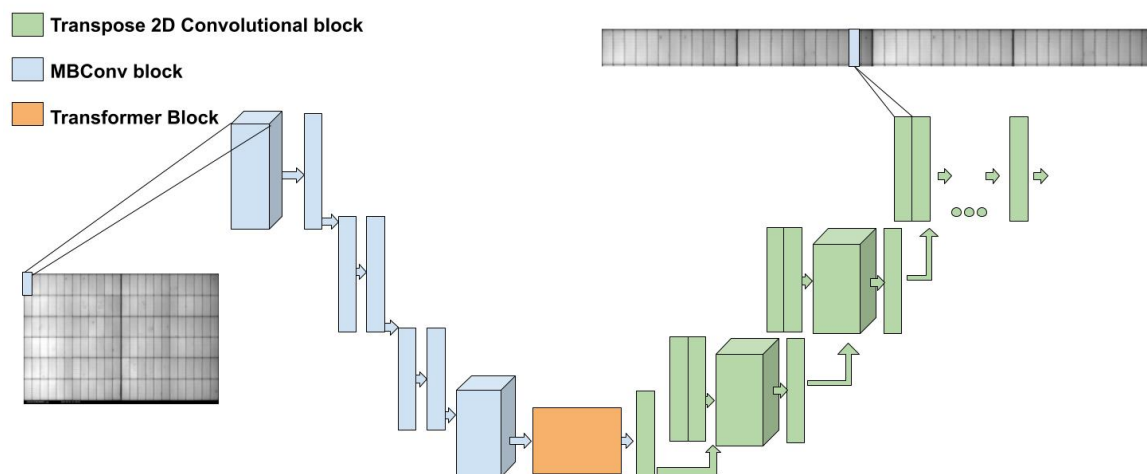


Fig 1. Efficient-UNet Transformer student model architecture.

2.1.2. Deep SVDD-Guided Anomaly Scoring

The Deep Support Vector Data Description (Deep SVDD) is a powerful technique for unsupervised anomaly detection that learns a compact representation of normal data by mapping it into a hypersphere in the latent space. This method is particularly effective for detecting outliers without the need for extensive labelled datasets, which aligns well with real-world scenarios in PV monitoring, where labelled anomalous data can be scarce. Deep SVDD aims to minimize the volume of a hypersphere that encloses most of the normal data points in a feature space, defined by a center c and a radius R . The optimization objective of Deep SVDD is formulated as:

$$\min_{R,W} R^2 + \frac{1}{vN} \max(0, \|f_w(x_i) - c\|^2 - R^2), \quad (5)$$

Where $f_w(x_i)$ is the learned feature extraction function parameterized by weights W , x_i represents an input image, c is the center of the hypersphere in the latent space, typically computed as the mean of the feature vectors obtained during training, v is a regularization parameter that controls the balance between minimizing the volume of the hypersphere and allowing some data points to lie outside of it, and N is the number of training samples. The objective is to find a feature space where normal data points lie close to c , while anomalies are identified by their deviation from this center. The anomaly score for a given input is derived from two

components; firstly, the discrepancy between the input x and its reconstructed version $\hat{x}$ is calculated:

$$\text{Reconstruction Error} = \|x - \hat{x}\|^2, \quad (6)$$

Secondly, the distance of the latent features from the center c :

$$\text{SVDD Score} = \|f_w(x) - c\|^2, \quad (7)$$

The combined anomaly score is formulated as:

$$\text{Anomaly Score}(x) = \|x - \hat{x}\|^2 + \|f_w(x) - c\|^2, \quad (8)$$

High anomaly scores indicate inputs that deviate significantly from normal patterns, suggesting defects or faults.

2.2 Teacher Model

The convolutional attention autoencoder, built on the ResNet-101 backbone, is a sophisticated model designed to capture detailed and hierarchical representations of EL image data, as presented in figure 2. In this framework, the ResNet-101-based encoder processes images through multiple residual blocks, each layer progressively extracting complex features. The deep residual learning structure of ResNet-101 allows the model to capture multi-level representations—from low-level textures to high-level semantic patterns—while using skip connections to maintain gradient flow, which is essential for learning deep, abstract features effectively. This architecture is particularly well-suited to identifying subtle patterns and variations in EL images that might indicate anomalies, a critical requirement in anomaly detection tasks.

Attention mechanisms embedded within the encoder guide the model's focus to the most relevant regions of the image, dynamically assigning weights to areas of interest. These attention layers work by learning to emphasize the parts of the image most likely to contain meaningful information, such as regions with slight irregularities or unique features. This selective focus enables the encoder to capture finer details without being overwhelmed by irrelevant information, improving its ability to detect anomalies. By highlighting these key regions, the attention modules ensure that the encoder emphasizes the features most critical for distinguishing normal from anomalous data.

The decoder portion of the autoencoder reconstructs the encoded features back into the

original image space, attempting to regenerate the input image from its learned representations. The quality of this reconstruction serves as a metric for anomaly detection: any significant discrepancies between the reconstructed and original images indicate potential anomalies. In the context of knowledge distillation, the output from this attention-driven autoencoder serves as a guide for the student model, transferring the valuable representations learned from the teacher model's complex, attention-focused encoding. This process enables the student model to inherit the attention-guided insights of the teacher, empowering it to perform accurate anomaly detection based on learned, nuanced patterns in the data, the architecture of presented teacher model presented in table 2.

Tabel 2. Summary of the Teacher Model Architecture.

Layer Name	Type	Output Shape	Number of Parameters
Input Layer	Input	(224, 224, 3)	0
Encoder	-	-	-
- ResNet101 Base	Convolutional	(7, 7, 2048)	44.5 million (frozen)
- Attention Block	Custom Attention	(7, 7, 256)	527,104
- Conv2D (f, g, h)	Conv2D	(7, 7, 256) each	1,576,960
- Add	Add Layer	(7, 7, 256)	0
- ReLU Activation	Activation	(7, 7, 256)	0
- Conv2D (sigmoid)	Conv2D	(7, 7, 1)	257
- Multiply	Multiply Layer	(7, 7, 256)	0
Decoder	-	-	-
- Conv2DTranspose (1)	Transposed Conv	(14, 14, 256)	590,080
- ReLU Activation	Activation	(14, 14, 256)	0
- Conv2DTranspose (2)	Transposed Conv	(28, 28, 128)	295,040
- ReLU Activation	Activation	(28, 28, 128)	0
- Conv2DTranspose (3)	Transposed Conv	(56, 56, 64)	73,792
- ReLU Activation	Activation	(56, 56, 64)	0
- Conv2DTranspose (4)	Transposed Conv	(112, 112, 32)	18,464
- ReLU Activation	Activation	(112, 112, 32)	0
- Conv2DTranspose (4)	Transposed Conv	(224, 224, 3)	867

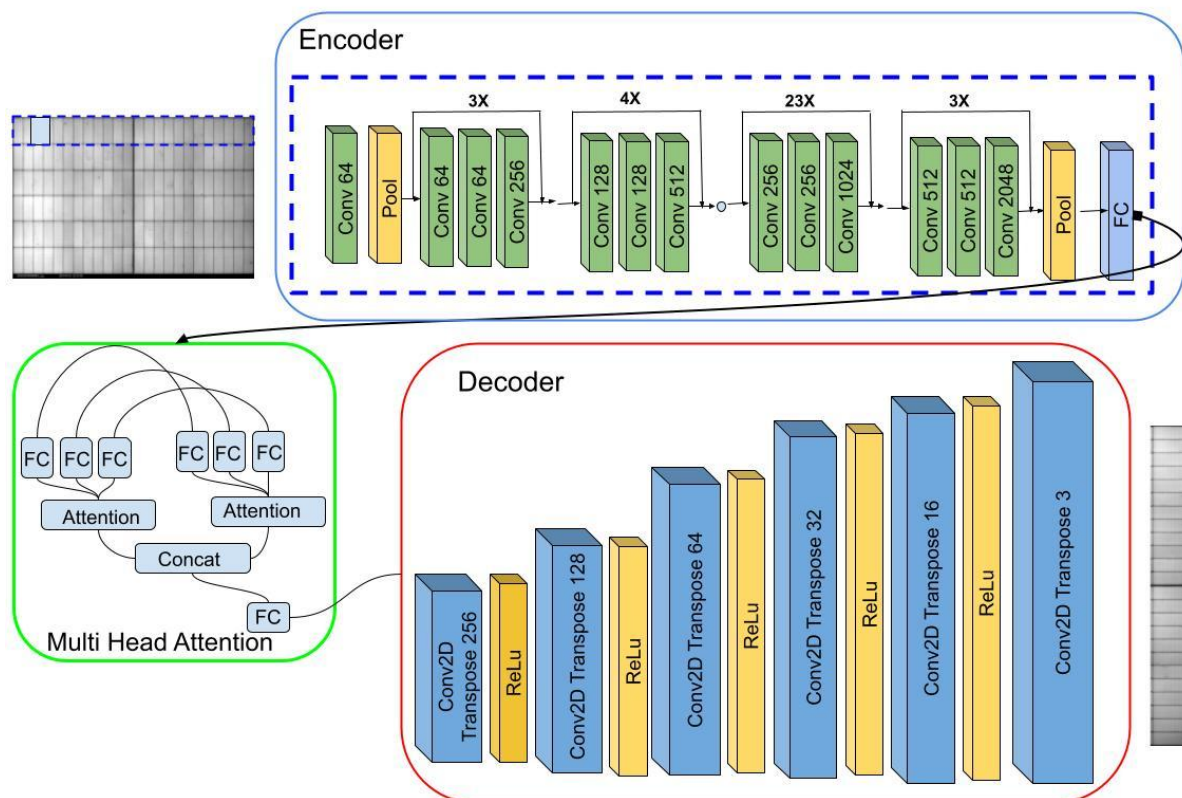


Fig 2. ResNet-101-based convolution encoder teacher model architecture.

2.3 Proposed Approach

The proposed framework for anomaly detection in electroluminescence (EL) images of solar panel cells integrates the strengths of a ResNet-101-based convolutional attention autoencoder as the teacher model and an Efficient-UNet Transformer as the student model through a knowledge distillation (KD) approach, as presented in figure 3. Reconstruction error rates from both models are fed into a Deep SVDD-guided anomaly scoring mechanism, enabling robust and precise anomaly detection. The teacher model captures detailed and hierarchical representations of EL images using its ResNet-101-based encoder, which emphasizes critical regions through residual blocks and attention mechanisms. Its decoder reconstructs these feature representations, with reconstruction errors serving as a key basis for anomaly identification. Through the knowledge distillation process, the teacher model's feature-rich outputs guide the student model, transferring advanced representations to enhance its learning and detection capabilities.

The student model, an Efficient-UNet Transformer, combines EfficientNet and U-Net

architectures to deliver computationally efficient and highly accurate anomaly detection. Its EfficientNet-based encoder uses MBConv blocks with depthwise separable convolutions and squeeze-and-excitation layers to preserve fine-grained spatial details while enabling multi-scale hierarchical feature extraction. At the bottleneck, a Transformer block with multi-head attention captures long-range dependencies and complex relationships, further refining feature representations. The decoder, equipped with up-sampling layers and skip connections, produces high-resolution segmentation masks that precisely localize potential defects in EL images. This design ensures the student model achieves comparable performance to the teacher model while maintaining computational efficiency, making it suitable for edge computing environments. To further enhance detection accuracy, the framework incorporates Deep SVDD-guided anomaly scoring, which maps normal data into a hypersphere in latent space, minimizing the hypersphere's volume to encapsulate normal data points. Anomalies are scored using both reconstruction error and the distance of feature representations from the hypersphere's center. This dual scoring system combines the Efficient-UNet's reconstruction capabilities with Deep SVDD's compact latent representation, enabling the detection of diverse anomaly types with high precision. By

integrating hierarchical representation learning, global context capture, and advanced anomaly scoring, the proposed framework ensures robust and

accurate EL image analysis, supporting effective quality control in photovoltaic systems.

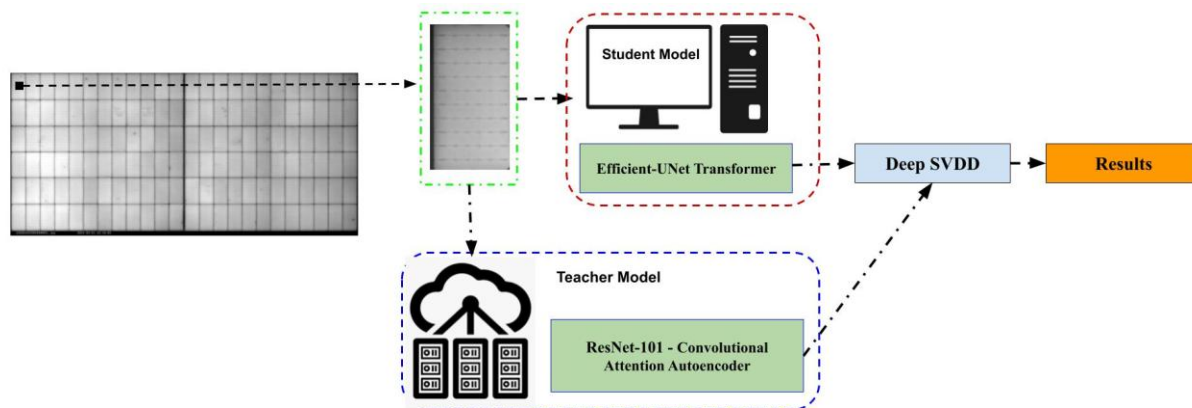


Fig 3. Architecture of proposed framework.

III. Experimental Results

The training and testing configuration of the proposed anomaly detection model for electroluminescence (EL) images is designed for efficient, high-accuracy performance. The model, an Efficient-UNet Transformer Autoencoder, incorporates custom Swish activation and an EfficientNet-based encoder combined with a Transformer bottleneck to capture local and global features for robust anomaly detection. Compiled with the Adam optimizer and Mean Squared Error (MSE) loss, the model is trained using a normalized dataset, rescaled by 1/255. Image data generator facilitates data preprocessing, resizing images to 224×224 pixels with a batch size of 4 to align with memory constraints. The model undergoes 500 training epochs, with a final saved version enabling future deployment. This configuration, utilizing only normal images for training, effectively captures patterns typical of defect-free cells, allowing for anomaly scoring based on deviations in reconstructed images.

3.1 Datasets

In this study, we utilized two datasets for anomaly detection in electroluminescence (EL)

images of solar cells, as presented in figure 4. The primary dataset, Visual Identification of Defective Solar Cells in Electroluminescence Imagery (referred to as the ELPV dataset), consists of 2,624 samples of 300×300-pixel 8-bit grayscale images representing both functional and defective solar cells with varying degrees of degradation. These images, extracted from 44 distinct solar modules, contain annotated defects categorized as intrinsic or extrinsic, both of which contribute to reduced power efficiency in solar modules. The images are standardized for size and perspective, and any distortion from the camera lens used during image acquisition was corrected prior to solar cell extraction. For this study, we classified the samples into two categories: normal and anomalous.

Additionally, in a real-world setting, we collaborated with the PV factory to collect EL images of solar panel cells directly from the production line. This dataset includes 4,000 images—2,000 labeled as normal and 2,000 labeled as anomalous, displaying defects with both micro and macro cracks. This additional dataset offers a practical perspective on defect detection, capturing naturally occurring variations in a production environment and enhancing the robustness of our analysis.

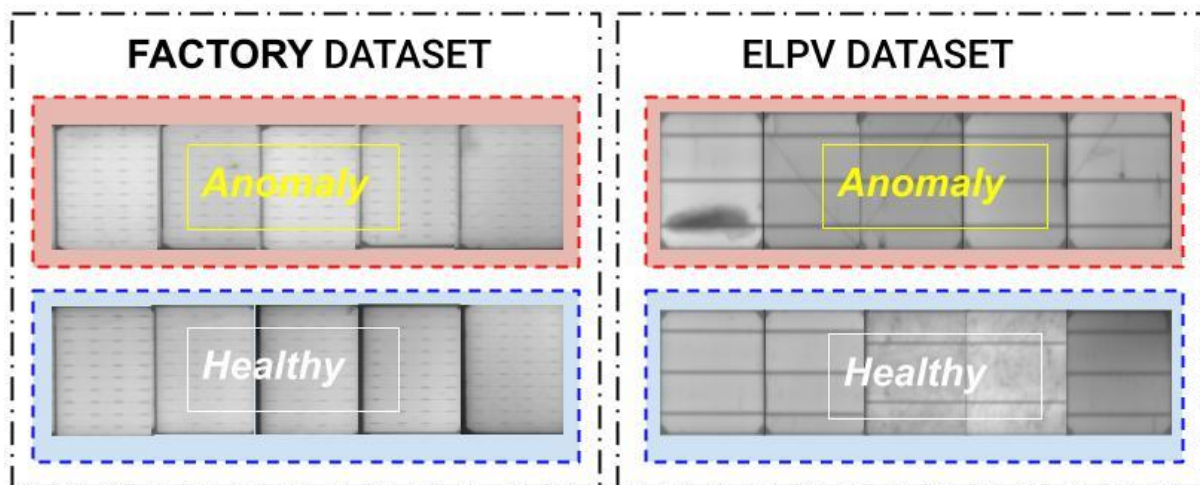


Fig 4. Experimental dataset utilized in this study.

3.2 Ablation study of Student Model

The ablation study of the student model systematically examines the influence of key components—Efficient-UNet, Transformer block, and Deep SVDD-guided anomaly scoring—on anomaly detection performance for electroluminescence (EL) images. Through a series of ablation configurations, we isolate the contributions of each element to understand its effect on detection accuracy, computational efficiency, and robustness. The study includes the following configurations:

- Efficient Net:** This configuration uses only the EfficientNet backbone without the U-Net architecture, Transformer, or Deep SVDD components. By isolating EfficientNet, we establish a baseline that reflects the performance of a simplified, efficient feature extractor, allowing comparison with more complex architectures.
- Baseline Efficient-UNet:** Here, the Efficient-UNet architecture is used alone, without the Transformer block or Deep SVDD anomaly scoring. This configuration provides a direct comparison to the standard U-Net structure for anomaly detection, serving as a foundational model for assessing the benefits of adding more advanced components.
- Efficient-UNet + Transformer:** This setup includes the Transformer block in the Efficient-UNet architecture but excludes Deep SVDD-guided scoring. This configuration isolates the effect of the Transformer, enabling analysis of its contribution to capture long-range

dependencies in EL images and how this influences detection accuracy.

- Efficient-UNet + Deep SVDD:** In this model, Efficient-UNet is combined with Deep SVDD-guided anomaly scoring, but without the Transformer. By isolating Deep SVDD, this setup assesses its ability to enhance anomaly detection through hypersphere-based scoring, examining its effectiveness in distinguishing between normal and defective cells.
- Full Model (Efficient-UNet + Transformer + Deep SVDD):** The complete student model combines Efficient-UNet, a Transformer with multi-head attention, and Deep SVDD-guided anomaly scoring. This full configuration leverages all key components to maximize accuracy, robustness, and efficiency in anomaly detection.

Table 3 presents the performance metrics from a comprehensive ablation study conducted on the Factory datasets to evaluate the impact of different model configurations on anomaly detection. Each model is assessed based on Precision, Recall, F1 Score, and Accuracy, using a fixed Optimal Threshold of 0.168. This evaluation aims to balance detection performance with the computational efficiency required for deployment on resource-constrained edge devices. By isolating and analyzing the contribution of each architectural component, the study provides insights into the effectiveness of various design choices in optimizing anomaly detection systems. The Baseline Efficient model sets a foundational benchmark with a precision of 0.70, recall of 0.84, F1 score of 0.76, and accuracy of 0.68. The introduction of the Efficient-UNet architecture in the Baseline

Efficient-UNet configuration improves recall to 0.85, F1 score to 0.77, and accuracy to 0.69, indicating better detection of anomalies. Incorporating Deep SVDD further enhances precision to 0.72 and achieves the highest accuracy among baseline configurations (0.71), reflecting a more balanced detection performance. The Efficient-UNet + Transformer variant achieves similar results to the baseline Efficient-UNet, with a recall of 0.85 and an F1 score of 0.77, showing its competitive performance.

The Proposed Approach (Student Model) stands out with the highest performance across all

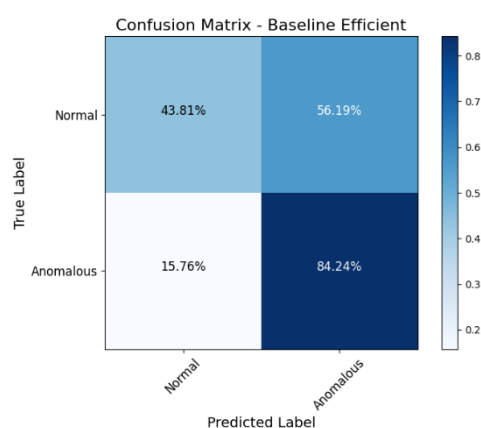
metrics. It achieves a precision of 0.73, recall of 0.86, F1 score of 0.79, and accuracy of 0.72, demonstrating its superiority in accurately detecting anomalies while maintaining robustness. These results underscore the effectiveness of the Student Model, which successfully integrates architectural improvements to achieve the optimal balance of high detection performance and efficiency for edge deployment scenarios. This thorough study highlights the critical role of systematic evaluation in guiding the development of effective and resource-efficient anomaly detection models.

Table 3. Performance metrics of Ablation study of Student Model based on Factory .

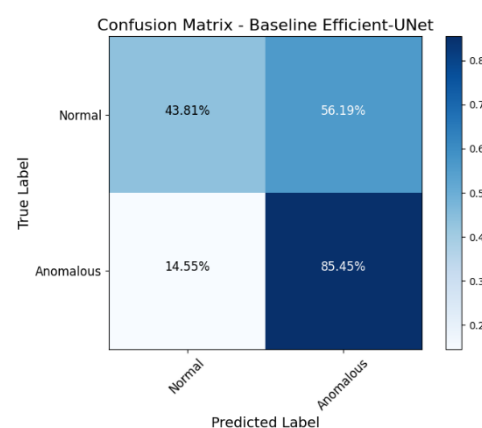
Model	Precision	Recall	F1	Accuracy	ART
Baseline Efficient	0.70	0.84	0.76	0.68	~3 ms
Baseline Efficient-UNet	0.70	0.85	0.77	0.69	~4 ms
Efficient-UNet + Deep SVDD	0.72	0.85	0.78	0.71	~4 ms
Efficient-UNet + Transformer	0.70	0.85	0.77	0.69	~5 ms
Proposed Approach (Student Model)	0.73	0.86	0.79	0.72	~6 ms

Figure 5 showcases confusion matrices for five different model configurations examined during the ablation study of the Student Model. Each matrix presents the distribution of true versus predicted labels for Normal and Anomalous cases. The Baseline Efficient model displays the lowest performance, with a significant misclassification rate for both normal and anomalous cases. Incorporating the Efficient-UNet architecture slightly improves classification accuracy, particularly for anomalous cases. However, there remains considerable room for improvement in both precision and recall. Adding Deep SVDD to the Efficient-UNet architecture further enhances the

model's capability to correctly classify normal cases, showing a more balanced performance. Similarly, the Efficient-UNet + Transformer configuration delivers marginal improvement in identifying anomalous cases while maintaining moderate accuracy for normal cases. Among all configurations, the Proposed Approach (Student Model) demonstrates the best overall performance, achieving the highest classification accuracy for both normal and anomalous cases, with reduced misclassification rates. This highlights the effectiveness of the Student Model in addressing the challenges of anomaly detection.



a) Baseline Efficient



b) Baseline Efficient-UNet

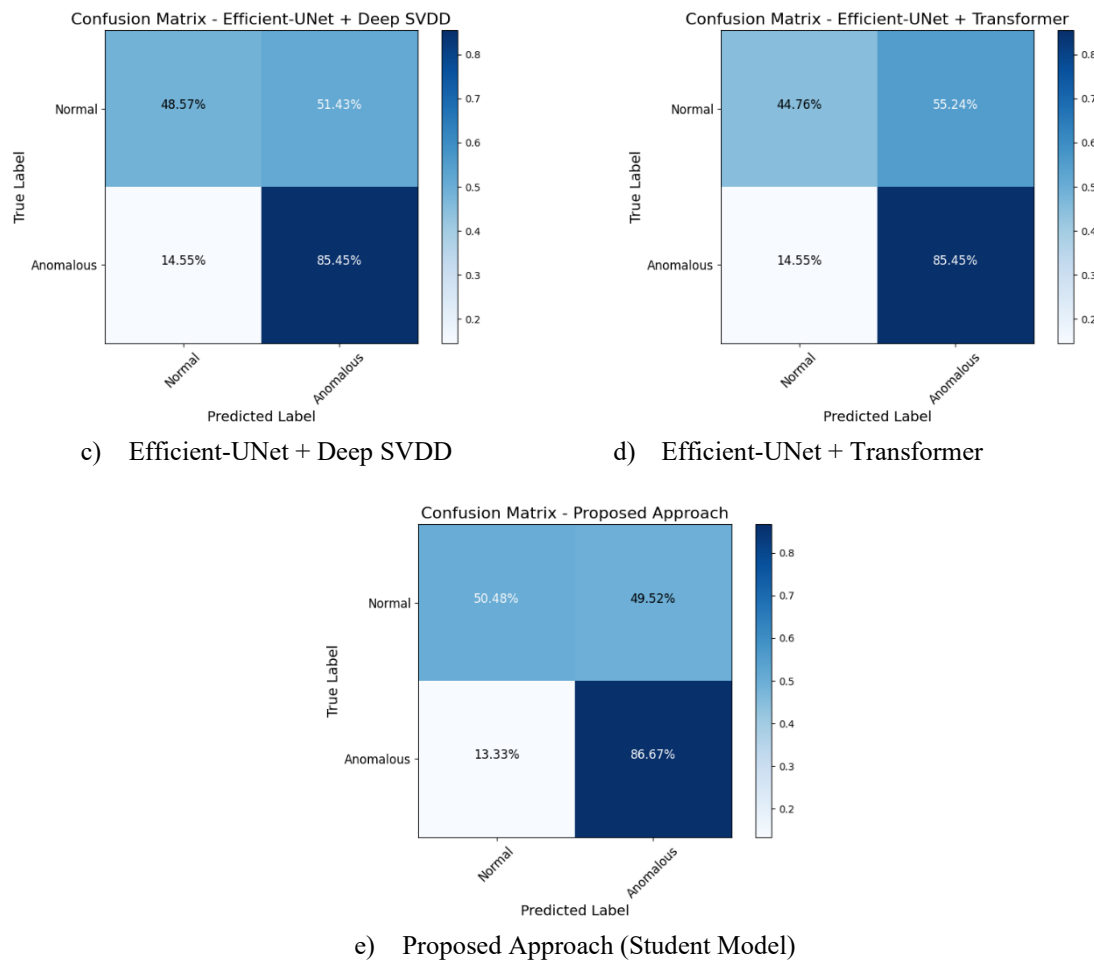


Fig 5. Confusion matrices for various models during the ablation study on Factory dataset.

The ROC curve presented in figure 6 compares the performance of different model configurations in terms of their ability to balance the True Positive Rate (Sensitivity) and the False Positive Rate for anomaly detection. Each curve corresponds to a specific model configuration, with the Area Under the Curve (AUC) serving as a performance metric. Higher AUC values reflect better overall classification performance. The Baseline Efficient model achieves an AUC of 0.72, while the Baseline Efficient-UNet configuration slightly improves upon this with an AUC of 0.73, showcasing the positive impact of integrating the Efficient-UNet architecture. Further enhancements are observed in the Efficient-UNet + Deep SVDD model, which achieves an AUC of 0.75, indicating

superior detection capabilities. Similarly, the Efficient-UNet + Transformer achieves an AUC of 0.74, demonstrating comparable but slightly lower performance. Among all configurations, the Proposed Approach (Student Model) delivers the highest performance, achieving an AUC of 0.78. This highlights its ability to effectively balance true positive and false positive rates, making it the most reliable model for anomaly detection. The dashed diagonal line represents a random guess (AUC = 0.5), providing a baseline for comparison. All tested models significantly outperform this baseline, with the Proposed Approach standing out as the most effective solution, demonstrating its robust and superior classification capabilities.

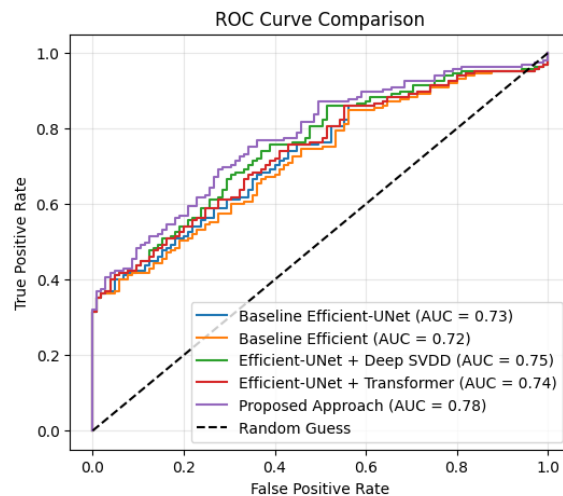


Fig 6. ROC curve comparison of different model configurations from the ablation study on Factory dataset.

3.3 Ablation study of Teacher model

After discussing the Student Model in the previous section, this subsection evaluates the performance of the Teacher Model as part of the knowledge distillation architecture. The presented approach, namely the ResNet-101-based convolutional attention autoencoder, is compared with other widely recognized architectures, including ResNet-101, ResNet-50, Inception V3, and VGG-19 Based on Factory dataset. This comparative analysis aims to highlight the advantages of integrating attention layers within the teacher model. The results presented in table 4

emphasize the positive impact of the attention mechanism in extracting robust features from EL images. These mechanisms help the model focus on the most relevant regions, enhancing the quality of the learned representations. Consequently, this leads to improved classification accuracy and more effective anomaly detection compared to the baseline architectures. By incorporating attention layers, the Teacher Model demonstrates its superiority in feature extraction and classification, which significantly benefits the overall knowledge distillation process and supports the Student Model's performance.

Table 4. Performance Evaluation of Conventional Models for Anomaly Detection on the Factory Dataset.

Model	Accuracy	Precision	Recall	F1	AUC	ART
InceptionV3	0.80	0.73	0.95	0.83	0.79	~8 ms
ResNet50	0.74	0.71	0.84	0.77	0.73	~7 ms
ResNet101	0.85	0.79	0.98	0.87	0.85	~12 ms
VGG19	0.78	0.71	0.96	0.81	0.77	~22 ms
Proposed Approach (Teacher model)	0.88	0.82	0.98	0.89	0.88	~15 ms

Furthermore, the confusion matrices illustrate the performance of different deep learning models—InceptionV3, ResNet50, ResNet101, VGG19, and the Proposed Approach (Teacher model)—on the Factory dataset for anomaly detection in electroluminescence images of solar panels in figure 7. Each matrix visualizes the proportion of true positives, true negatives, false positives, and false negatives as percentages. Across all models, the "Proposed Approach (Teacher

model)" achieved the best performance, with 98.92% accuracy in detecting anomalous images and a notably lower false positive rate (22.77% misclassification of normal images) compared to other models. ResNet101 also performed well with a similar anomalous detection rate (98.92%), but it had a slightly higher false positive rate (27.70%). These results demonstrate the effectiveness of the attention-enhanced ResNet101-based Proposed Approach in minimizing misclassifications while

maintaining high recall.

Other models, such as VGG19 and InceptionV3, showed relatively high anomaly detection rates (96.59% and 95.52%, respectively) but struggled with higher false positive rates (41.37% for VGG19 and 36.43% for InceptionV3), indicating a tendency to over-predict anomalies. ResNet50, despite being computationally efficient, achieved the lowest anomalous detection accuracy

(84.41%) and the highest misclassification of normal images (36.43%), making it less suitable for this task. These results highlight the trade-off between computational efficiency and anomaly detection accuracy, with the Proposed Approach providing a balanced solution by leveraging an attention mechanism to enhance feature representation and classification performance.

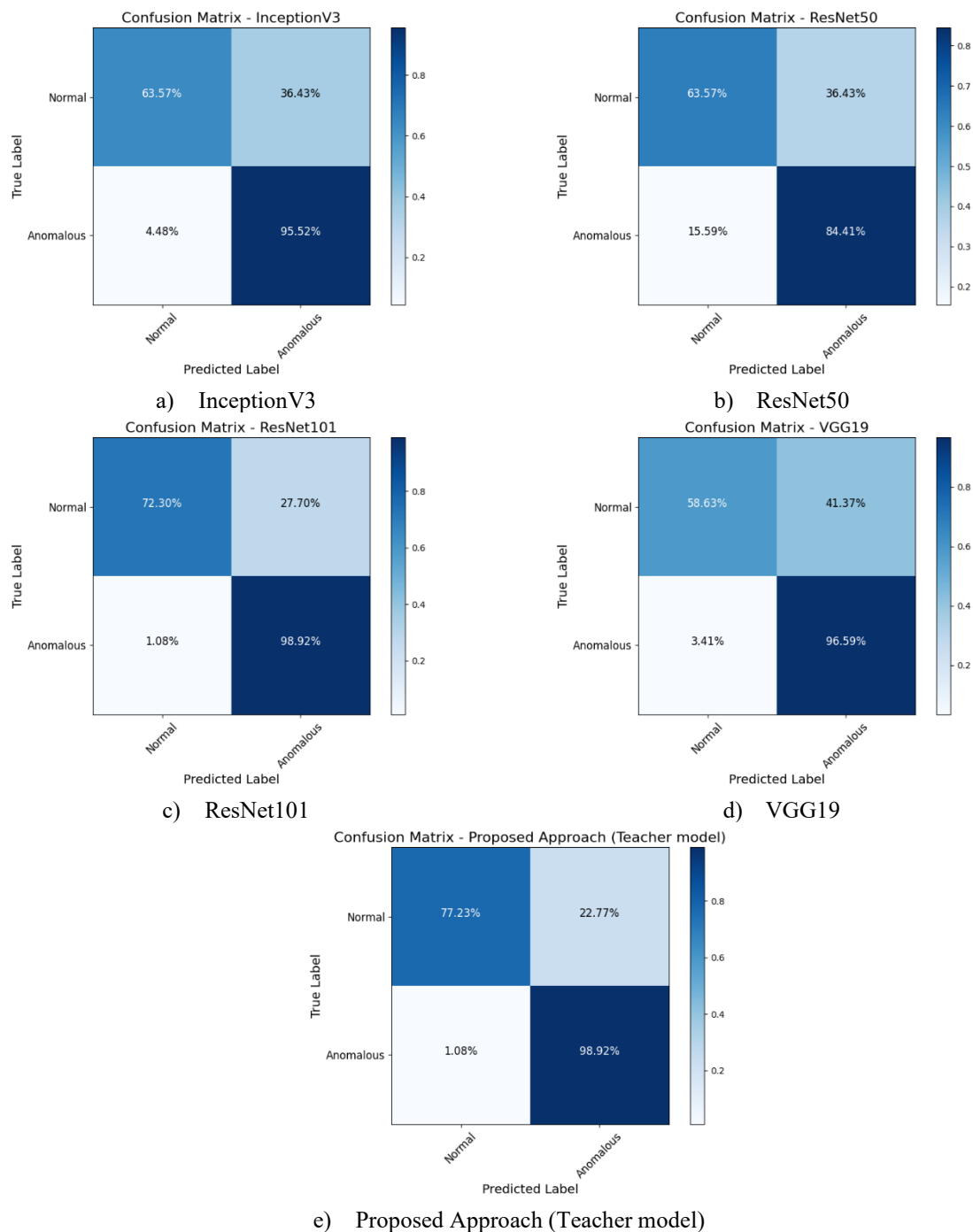


Fig 7. Confusion matrices of various models based on Factory dataset.

Based on the results presented in figure 8, the Knowledge Distillation model (proposed approach) emerges as the most effective for anomaly detection. This is evidenced by its improved separation of normal and abnormal score distributions. The plot illustrates that the normal scores (blue) are sharply concentrated within the lower reconstruction error range (~ 0.0 – 0.2), while the abnormal scores (red) span a broader range (~ 0.2 – 1.0) with minimal overlap. This distinct separation significantly reduces potential misclassifications, making the model highly reliable. In comparison, the Student Model exhibits greater overlap between the normal and abnormal scores, particularly in the range of ~ 0.1 – 0.3 . This overlap increases the probabilities of false positives and false negatives, diminishing its effectiveness. Similarly, the Teacher Model shows high overlap in the range of ~ 0.0 – 0.3 , where many abnormal scores are indistinguishable from normal ones. This

overlap undermines its classification accuracy, despite the Teacher Model's robustness. Quantitatively, the Knowledge Distillation approach achieves a more refined distribution by leveraging the Teacher Model's robustness while enhancing the Student Model's generalization. For instance, the abnormal score peak in the Teacher Model is around ~ 0.2 – 0.3 , overlapping with the normal score tail. In contrast, the Knowledge Distillation model shifts this peak slightly higher, reducing the overlap and the false positive region. Additionally, the normal score concentration in the Knowledge Distillation model remains tightly within ~ 0.0 – 0.1 , compared to the broader spread observed in the Student Model.

These observations underscore that the Knowledge Distillation approach offers the best trade-off between detecting anomalies and maintaining low misclassification rates. Consequently, it is the most effective method among the three models.

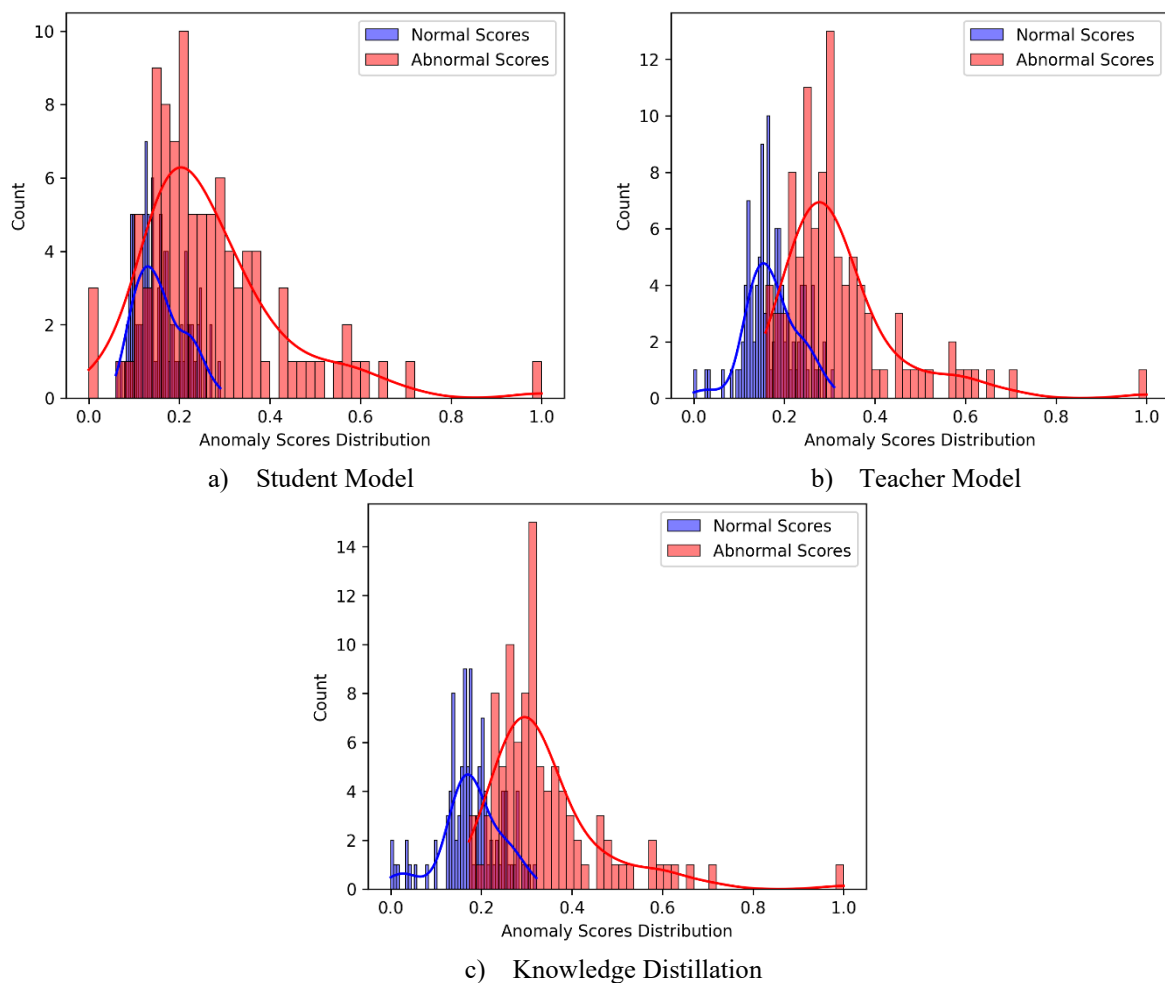


Fig 8. Anomaly score distribution of the proposed approach on Factory Dataset.

3.4 Comparison With State of art Method

To ensure a fair and comprehensive comparison of our proposed model with state-of-the-art approaches, we utilized the ELPV dataset, which is widely recognized for benchmarking performance in this domain. The dataset was pre-processed and partitioned using a consistent train-test split, adhering to established protocols to maintain reproducibility and eliminate biases. By employing the ELPV dataset, we ensure our analysis aligns with the practices of prior studies, facilitating meaningful comparisons with state-of-the-art methodologies while validating the efficacy of the proposed approach.

Table 5 provides a detailed comparison of various models based on accuracy and the number of trainable parameters, illustrating the trade-offs between performance and computational efficiency. SeMaCNN achieves an accuracy of 0.95, showcasing high performance but at the cost of a large parameter count (~23.4M), making it computationally expensive. The Pyramid Dynamic

Graph models provide more lightweight alternatives, with the AlexNet variant achieving an accuracy of 0.88 using ~3.7M parameters, and the VGG16 variant improving accuracy to 0.90 with a reduced parameter count (~2.03M), suitable for resource-constrained scenarios but with a slight trade-off in accuracy. MLA-SDAL and GCGB demonstrate strong accuracies of 0.97 and 0.96 with parameter counts of ~5M and ~15M, respectively, achieving a balance between computational demand and performance. The Efficient-UNet + Transformer + Deep SVDD (Student Model) offers an accuracy of 0.88 with ~4M parameters, reflecting its lightweight design. The ResNet101 Convolutional Attention Autoencoder (Teacher Model) achieves a higher accuracy of 0.97 with just ~3M parameters, highlighting its efficiency as a teacher. Finally, the Proposed Approach (Knowledge Distillation) achieves the same accuracy as the teacher model (0.97) with ~7M parameters, demonstrating the effective transfer of knowledge and performance through distillation while maintaining a reasonable computational cost.

Table 5. Comparison with State-of-the-Art Models Based on ELPV dataset.

Model	Accuracy	Trainable Params
SeMaCNN[17]	0.95	~ 23.4M
Pyramid Dynamic Graph-AlexNet[27]	0.88	~ 3.7 M
Pyramid Dynamic Graph-VGG16 [27]	0.90	~ 2.03 M
MLA-SDAL[11]	0.97	~ 5 M
GCGB [28]	0.96	~ 15 M
GAN-AE [29]	0.90	~ 20 M
Efficient-UNet + Transformer + Deep SVDD (Student Model)	0.88	~ 4 M
Resnet101 Convolutional Attention Autoencoder (Teacher Model)	0.97	~ 3 M
Proposed approach (Knowledge Distillation)	0.97	~ 7 M

IV. Results and Discussion

The experimental results demonstrate the effectiveness of the proposed Knowledge Distillation (KD) framework for anomaly detection in electroluminescence (EL) images of solar panel cells. The KD approach was chosen for its unique ability to balance high detection accuracy with computational efficiency, making it ideal for deployment in resource-constrained environments like the Factory Solar Production Line. By distilling knowledge from the ResNet-101 convolutional attention autoencoder (teacher) into the Efficient-UNet Transformer (student), the framework

achieves superior performance while significantly reducing computational overhead. The student model, guided by the teacher, excels in detecting subtle and complex defect patterns, as shown in figure 8, where key anomaly regions in EL images are correctly identified with high confidence. The lightweight Efficient-UNet Transformer integrates local and global feature learning through its encoder and transformer layers, enabling robust anomaly detection even under challenging conditions in real-world production scenarios.

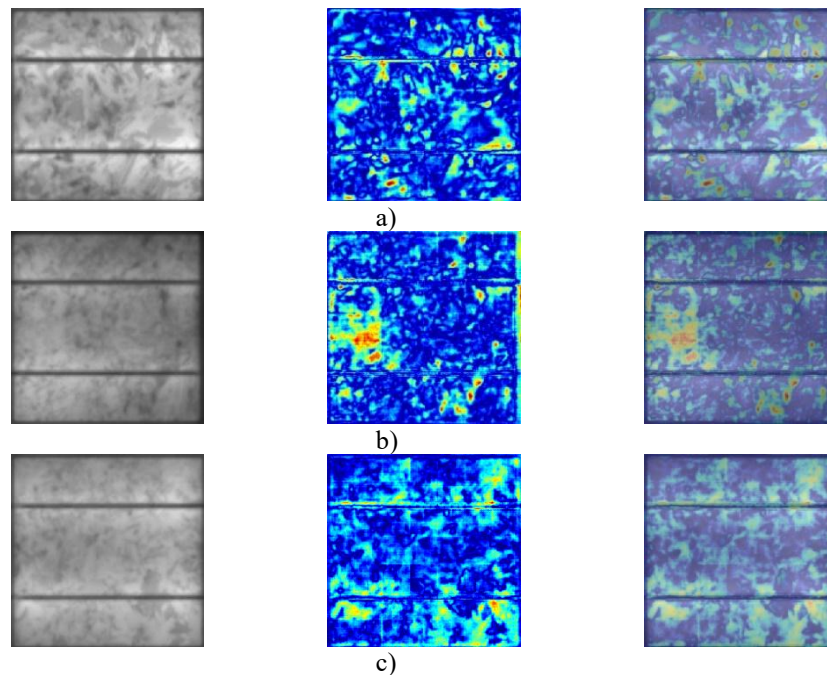


Fig 8. Visual representation of solar panel reconstruction error heatmaps and anomaly scores across ELPV dataset classes (I: 0.1, II: 0.3, III: 0.6).

Table 6 summary of the Performance of the Proposed Approach on the Public ELPV Dataset and the Real-World Industrial Factory Dataset. The results highlight that while the Efficient-UNet + Deep SVDD (Student Model) demonstrates competitive performance, the Knowledge Distillation (KD) approach effectively matches the accuracy of the ResNet-101 Convolutional Attention

Autoencoder (Teacher Model), achieving higher accuracy on both datasets with a more efficient design. Furthermore, the results demonstrate the effectiveness of the edge-based deep learning model, showcasing its ability to achieve high accuracy across both public and industrial datasets, highlighting its potential for real-world applications.

Table 6. Overall Performance of the Proposed Framework on the Factory and ELPV Datasets.

Dataset	Model	Accuracy
Factory	Efficient-UNet + Deep SVDD (Student Model)	0.73
	ResNet-101 – Convolutional Attention Auto encoder (Teacher Model)	0.88
	Knowledge Distillation (KD)	0.88
ELPV	Efficient-UNet + Deep SVDD (Student Model)	0.88
	ResNet-101 – Convolutional Attention Auto encoder (Teacher Model)	0.97
	Knowledge Distillation (KD)	0.97

The deployment of the model in the wild has been illustrated in figure 9, showcasing its application on the Factory Solar Production Line. Key areas include:

- The student model processes EL images in real-time on edge devices installed at inspection stations along the production line.
- The system flags panels with potential anomalies for further inspection, ensuring high-quality standards.

- Anomaly scores are integrated into predictive maintenance pipelines, reducing downtime by identifying early signs of equipment degradation. The adoption of the KD-based framework ensures scalable and efficient deployment across multiple stations without the need for high-performance hardware, making it both cost-effective and reliable for large-scale industrial use. These results underscore the suitability of Knowledge Distillation for advancing the automation and accuracy of PV cell inspection systems in dynamic industrial settings.

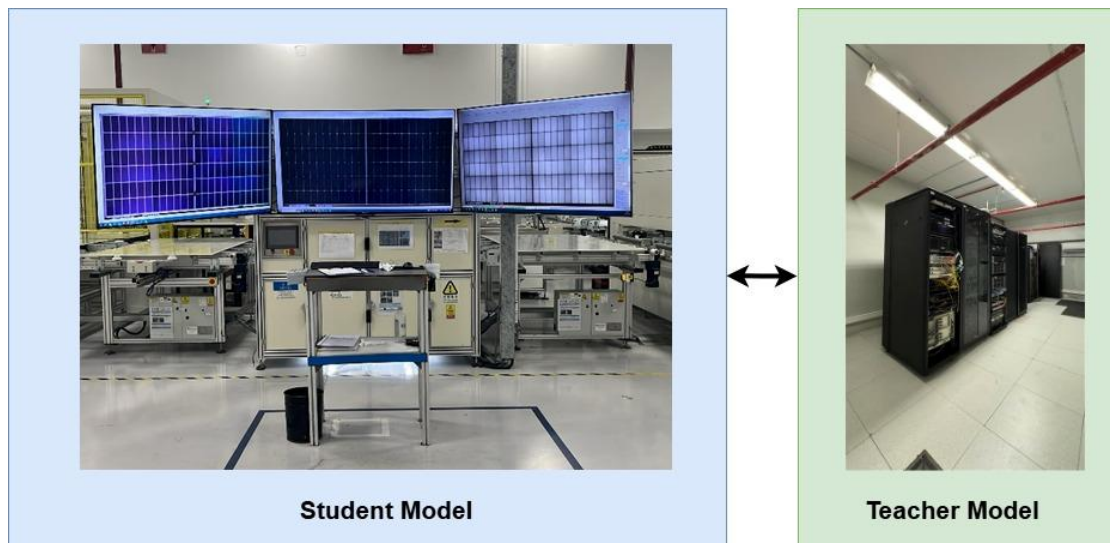


Fig 9. Knowledge Distillation deployment scenario in production lines of Factory .

While the proposed Knowledge Distillation framework demonstrates high accuracy and computational efficiency, there are several areas for potential improvement. One limitation lies in the reliance on teacher-student architecture, which can be sensitive to the quality of knowledge transfer. Future work could explore advanced distillation techniques, such as mutual learning or self-supervised pre-training, to enhance the robustness of the student model. Additionally, the current framework is primarily validated on EL images; expanding its application to other defect detection modalities (e.g., thermographic or visual inspection) could improve versatility. Another drawback is the model's computational demand during training, which, though optimized for inference, may pose challenges for resource-constrained environments. Incorporating model compression techniques, such as pruning or quantization, could further reduce resource requirements without sacrificing accuracy. Lastly, the dependence on a labeled dataset for teacher training highlights a need for exploring semi-supervised or unsupervised approaches to make the framework more applicable to scenarios with limited annotated data. By addressing these limitations, the method can become more scalable and robust for broader industrial applications.

V. Conclusion

In conclusion, the proposed Knowledge Distillation framework successfully combines the strengths of deep learning techniques to deliver an effective solution for anomaly detection in electroluminescence (EL) images of solar panels. By leveraging a ResNet-101-based Convolutional Attention Autoencoder as the teacher model and an Efficient-UNet Transformer with Deep SVDD-

guided anomaly scoring as the student model, the framework achieves high detection accuracy while maintaining computational efficiency. This design makes it well-suited for deployment in resource-constrained environments, such as industrial production lines, where real-time monitoring is essential. The framework's dual anomaly scoring mechanism, combining reconstruction errors and Deep SVDD-based scoring, enhances its ability to detect diverse types of anomalies, even under noisy conditions. The integration of knowledge distillation ensures that the student model inherits the robust feature extraction capabilities of the teacher model, resulting in high performance with significantly reduced resource requirements. Experimental evaluations on both public (ELPV) and real-world industrial (Factory) datasets demonstrate the framework's superiority in balancing accuracy and efficiency, with results matching or exceeding state-of-the-art models. Despite its strengths, the method has limitations, including sensitivity to the quality of knowledge transfer and reliance on labelled data for training the teacher model. Future work could explore self-supervised or semi-supervised learning techniques, as well as model compression methods to further enhance scalability and applicability across diverse defect detection modalities. Overall, the proposed approach represents a significant advancement in automated PV cell monitoring, contributing to improved quality control and operational efficiency in solar energy systems. By addressing the outlined limitations in future research, the framework can be further refined to meet the evolving demands of industrial and real-world applications.

Declaration of competing interest

We declare that there are no conflicts of interest to disclose related to this manuscript.

Data availability

The FACTORY dataset used to support this study's finding has been accessed via email.

References:

- [1] S. Einy, E. Sen, H. Saygin, H. Hivehchi, and Y. Dorostkar Navaei, "Local Binary Convolutional Neural Networks' Long Short-Term Memory Model for Human Embryos' Anomaly Detection," *Sci Program*, vol. 2023, 2023, doi: 10.1155/2023/2426601.
- [2] S. Einy, H. Saygin, H. Hivehch, and Y. Dorostkar Navaei, "Local and Deep Features Based Convolutional Neural Network Frameworks for Brain MRI Anomaly Detection," *Complexity*, vol. 2022, 2022, doi: 10.1155/2022/3081748.
- [3] S. Einy, C. Oz, and Y. D. Navaei, "IoT Cloud-Based Framework for Face Spoofing Detection with Deep Multicolor Feature Learning Model," *J Sens*, vol. 2021, 2021, doi: 10.1155/2021/5047808.
- [4] D. M. Tsai, S. C. Wu, and W. Y. Chiu, "Defect detection in solar modules using ICA basis images," *IEEE Trans Industr Inform*, vol. 9, no. 1, pp. 122–131, 2013, doi: 10.1109/TII.2012.2209663.
- [5] S. A. Anwar and M. Z. Abdullah, "Micro-crack detection of multicrystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique," *EURASIP J Image Video Process*, vol. 2014, 2014, doi: 10.1186/1687-5281-2014-15.
- [6] G. Laguna, P. Moreno, J. Cipriano, G. Mor, E. Gabaldón, and A. Luna, "Detection of abnormal photovoltaic systems' operation with minimum data requirements based on Recursive Least Squares algorithms," *Solar Energy*, vol. 274, May 2024, doi: 10.1016/j.solener.2024.112556.
- [7] S. Fadhel *et al.*, "PV shading fault detection and classification based on I-V curve using principal component analysis: Application to isolated PV system," *Solar Energy*, vol. 179, pp. 1–10, Feb. 2019, doi: 10.1016/J.SOLENER.2018.12.048.
- [8] S. Chandio *et al.*, "Machine Learning-Based Multiclass Anomaly Detection and Classification in Hybrid Active Distribution Networks," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3445287.
- [9] I. U. Khalil, A. ul Haq, and N. ul Islam, "A novel procedure for photovoltaic fault forecasting," *Electric Power Systems Research*, vol. 226, p. 109881, Jan. 2024, doi: 10.1016/J.EPSR.2023.109881.
- [10] H. P. C. Hwang, C. C. Y. Ku, and J. C. C. Chan, "Detection of malfunctioning photovoltaic modules based on machine learning algorithms," *IEEE Access*, vol. 9, pp. 37210–37219, 2021, doi: 10.1109/ACCESS.2021.3063461.
- [11] Z. Chang, A. J. Zhang, H. Wang, J. Xu, and T. Han, "Photovoltaic Cell Anomaly Detection Enabled by Scale Distribution Alignment Learning and Multiscale Linear Attention Framework," *IEEE Internet Things J*, vol. 11, no. 16, pp. 27816–27827, 2024, doi: 10.1109/JIOT.2024.3403711.
- [12] H. Kang, J. Hong, J. Lee, and S. Kang, "Photovoltaic Cell Defect Detection Based on Weakly Supervised Learning With Module-Level Annotations," *IEEE Access*, vol. 12, pp. 5575–5583, 2024, doi: 10.1109/ACCESS.2024.3349975.
- [13] W. Tang, Q. Yang, K. Xiong, and W. Yan, "Deep learning based automatic defect identification of photovoltaic module using electroluminescence images," *Solar Energy*, vol. 201, pp. 453–460, May 2020, doi: 10.1016/J.SOLENER.2020.03.049.
- [14] B. Su, Z. Zhou, and H. Chen, "PVEL-AD: A Large-Scale Open-World Dataset for Photovoltaic Cell Anomaly Detection," *IEEE Trans Industr Inform*, vol. 19, no. 1, pp. 404–413, Jan. 2023, doi: 10.1109/TII.2022.3162846.
- [15] Q. Liu, M. Liu, C. Wang, and Q. M. J. Wu, "An efficient CNN-based detector for photovoltaic module cells defect detection in electroluminescence images," *Solar Energy*, vol. 267, p. 112245, Jan. 2024, doi: 10.1016/J.SOLENER.2023.112245.
- [16] H. Munawer Al-Otum, "Classification of anomalies in electroluminescence images of solar PV modules using CNN-based deep learning," *Solar Energy*, vol. 278, p. 112803, Aug. 2024, doi: 10.1016/J.SOLENER.2024.112803.
- [17] A. Korovin *et al.*, "Anomaly detection in electroluminescence images of heterojunction solar cells," *Solar Energy*, vol. 259, pp. 130–136, Jul. 2023, doi: 10.1016/J.SOLENER.2023.04.059.
- [18] E. David, J. R. Orillaza, and J. R. Pedrasa, "Islanding Detection Using Transformer Neural Networks," in *2024 IEEE 4th International Conference on Power, Electronics and Computer Applications, ICPECA 2024*, Institute of Electrical and

- Electronics Engineers Inc., 2024, pp. 592–596. doi: 10.1109/ICPECA60615.2024.10471069.
- [19] Y. Zhao, Q. Liu, D. Li, D. Kang, Q. Lv, and L. Shang, “Hierarchical anomaly detection and multimodal classification in large-scale photovoltaic systems,” *IEEE Trans Sustain Energy*, vol. 10, no. 3, pp. 1351–1361, Jul. 2019, doi: 10.1109/TSTE.2018.2867009.
- [20] L. Zhao, Y. Wu, and Y. Yuan, “PD-DETR: towards efficient parallel hybrid matching with transformer for photovoltaic cell defects detection,” *Complex and Intelligent Systems*, 2024, doi: 10.1007/s40747-024-01559-0.
- [21] D. Dwivedi, K. V. S. M. Babu, P. K. Yemula, P. Chakraborty, and M. Pal, “Identification of surface defects on solar PV panels and wind turbine blades using attention based deep learning model,” *Eng Appl ArtifIntell*, vol. 131, May 2024, doi: 10.1016/j.engappai.2023.107836.
- [22] J.-H. Kim, S. Yi, and T. Lee, “Semi-orthogonal Embedding for Efficient Unsupervised Anomaly Segmentation.”
- [23] U. Otamendi, I. Martinez, M. Quartulli, I. G. Olaizola, E. Viles, and W. Cambarau, “Segmentation of cell-level anomalies in electroluminescence images of photovoltaic modules,” *Solar Energy*, vol. 220, pp. 914–926, May 2021, doi: 10.1016/J.SOLENER.2021.03.058.
- [24] L. Bommaset *et al.*, “Anomaly detection in IR images of PV modules using supervised contrastive learning,” *Progress in Photovoltaics: Research and Applications*, vol. 30, no. 6, pp. 597–614, Jun. 2022, doi: 10.1002/pip.3518.
- [25] R. A. De Oliveira, A. Malfoy, and S. K. Ronnberg, “Deep Anomaly Detection of Voltage Waveform Distortion in the Transmission Grid Due to Geomagnetically Induced Currents,” *IEEE Trans Instrum Meas*, vol. 73, pp. 1–12, 2024, doi: 10.1109/TIM.2024.3366271.
- [26] M. Zarghami, T. Niknam, J. Aghaei, and A. H. Nezhad, “Concurrent PV production and consumption load forecasting using CT-Transformer deep learning to estimate energy system flexibility,” *IET Renewable Power Generation*, Oct. 2024, doi: 10.1049/rpg2.13050.
- [27] S. Apak and M. Farsadi, “Multi-branch spatial pyramid dynamic graph convolutional neural networks for solar defect detection,” *Computers and Electrical Engineering*, vol. 121, Jan. 2025, doi: 10.1016/j.compeleceng.2024.109872.
- [28] N. Zhu, J. Wang, Y. Zhang, H. Wang, and T. Han, “An Adversarial Training Framework Based on Unsupervised Feature Reconstruction Constraints for Crystalline Silicon Solar Cells Anomaly Detection,” *IEEE Trans Instrum Meas*, 2024, doi: 10.1109/TIM.2024.3462989.
- [29] C. Shou *et al.*, “Defect detection with generative adversarial networks for electroluminescence images of solar cells,” in *Proceedings - 2020 35th Youth Academic Annual Conference of Chinese Association of Automation, YAC 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 312–317. doi: 10.1109/YAC51587.2020.9337676.