Investigating Variational Autoencoders for Data Imbalance Mitigation in Solar Panel Defect Detection

Leonardo Sani

sani.2059860@studenti.uniroma1.it

Simone Mantero

mantero.2014057@studenti.uniroma1.it

1. Introduction & Related Work

The reliability of photovoltaic (PV) systems is strongly influenced by the occurrence of defects, which can lead to efficiency losses and accelerated system degradation [4]. Automated inspection using Electroluminescence (EL) imaging has become a key diagnostic technique. While Convolutional Neural Networks (CNNs) [6] are the standard for analyzing such imagery, their performance is severely bottlenecked by data scarcity. In operational environments, defective samples are naturally rare, creating a class imbalance that leads to model overfitting [4].

Current mitigation strategies on the ELPV dataset [1] typically involve geometric augmentation or cost-sensitive learning [7]. Extending this landscape, we investigate Generative Data Augmentation using Variational Autoencoders (VAEs) [5]. We employ a VAE to generate synthetic defect images and explore this generative approach as an alternative to standard augmentation for training a ResNet18 [3] classifier.

2. Proposed Method Explained

We developed a two-stage pipeline: first, a Variational Autoencoder (VAE) learns the distribution of defective cells to synthesize new samples; second, these synthetic samples are integrated into the training set of a ResNet18 classifier to mitigate class imbalance.

2.1. Generative Augmentation with VAE

We designed a flexible convolutional VAE framework that maintains a lightweight footprint across all experimental configurations. This global efficiency facilitated a broad hyperparameter exploration for every scenario. As a reference point, our optimal model utilizes approximately 4.5 million parameters.

2.1.1. Experimental Scenarios

We defined four experimental scenarios (**A, B, C, D**) distinguished by image resolution (128×128 vs. 256×256) and normalization range ($[-1, 1]$ vs. $[0, 1]$). The first choice dictated the necessity of two additional convolutional layers

for larger inputs, while the second determined the final activation function of the decoder (*tanh* vs. *sigmoid*), which in turn constrained the compatibility of specific reconstruction losses. The hyperparameter search space is summarized in Table 1.

Table 1. VAE Scenarios defined by image size and normalization. **A**: 128^2, $[-1, 1]$; **B**: 256^2, $[-1, 1]$; **C**: 128^2, $[0, 1]$; **D**: 256^2, $[0, 1]$. All sets evaluated the impact of Weight Decay.

Set	Rec. Loss	Latent Dim (z)	β-VAE Strategy
A	MAE, MSE	$64, 128$	$1, < 1$
B	MAE, MSE	$128, 256$	$1, < 1$
C	BCE	$64, 128$	$1, < 1$
D	BCE	$128, 256$	$1, < 1$

2.1.2. Training and Model Selection

To ensure sufficient data density for robust manifold learning, the VAE was trained exclusively on the **geometrically augmented defective subset**. The models were trained using early stopping based on validation loss. Since assessing generation utility via downstream classification for every configuration was inefficient without prior verification of sample quality, we adopted a two-stage qualitative evaluation strategy. First, we conducted a global screening of the hyperparameter space (Scenarios A–D) by inspecting samples from the standard prior ($z \sim \mathcal{N}(0, I)$). This comparative analysis identified **Scenario A** (MSE loss, fixed $\beta = 0.4$, latent dim 128) as the superior candidate, offering the best baseline trade-off between image sharpness and structural coherence. Subsequently, we focused on this candidate to probe latent space continuity via Spherical Linear Interpolation (SLERP) [8], which revealed "ghosting" artifacts indicative of manifold discontinuities. To address this, we refined the regularization strategy by adopting β-annealing ($0 \rightarrow 2$). This dynamic regularization mitigated manifold discontinuities; while it did not completely eliminate ghosting, it significantly improved the semantic consistency of morphing, particularly for transitions between proximal latent vectors.
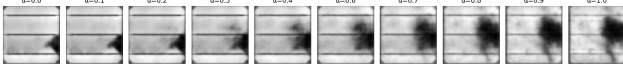
Figure 1. SLERP between two defect samples after β annealing

2.1.3. Controlled Sample Generation

To generate synthetic defects, we performed Spherical Linear Interpolation (SLERP) [8] between pairs of class-consistent samples (strictly mono- or polycrystalline) to minimize unnatural hybrid textures. A critical post-processing step involved filtering ambiguous samples. Since VAEs lose high-frequency detail, subtle defects blur into "healthy-looking" morphologies. We calculated a healthy centroid using encoder-derived μ vectors and discarded proximal interpolations. Although this injects a slight bias towards evident faults, it ensures dataset purity by removing confusing inputs. This strategy allows us to explicitly verify generalization improvements for distinct, *visible* defects.

2.2. Classification Model: ResNet18

We employed a *ResNet18* backbone pre-trained on ImageNet, adapting first convolutional layer to accept single-channel EL inputs by averaging the original RGB weights. The fully connected layer was replaced with a single-neuron output for binary classification. ResNet18 was fine-tuned using the Adam and BCE loss, with hyperparameter tuning based on the validation F1-score. Final training was conducted with a LR=$1e10^{-4}$, a WD=$1e10^{-4}$, and a BS=32.

3. Dataset and Benchmark

3.1. Dataset and Preprocessing

We utilized the ELPV dataset [1], a benchmark collection of $2,624$ samples of functional and defective solar cells characterized by two distinct background textures: Monocrystalline (mono) and Polycrystalline (poly).

The dataset was first partitioned into training (80%), validation (10%), and testing (10%) sets, ensuring that the validation and test sets remained fixed and comprised exclusively of original, real images. To prevent classification bias—where models might rely on background texture rather than defect features—and to stabilize the VAE training, we enforced **equal representation of mono- and polycrystalline cells** within both the **defective and functional subsets** of the training data. This balancing was achieved via **preliminary geometric augmentation** (e.g., rotations, vertical/horizontal flips). Crucially, this geometrically balanced configuration not only facilitates stable VAE learning but also constitutes the **Geometric Baseline** against which we later evaluate the effectiveness of our generative approach.

3.2. Benchmark Strategy

To rigorously evaluate the proposed generative method, we defined five experimental scenarios—three internal implementations and two external references:

1. **Unbalanced (Internal):** Adapted ResNet18 trained on the original imbalanced dataset.
2. **Geometric Baseline (Internal):** Adapted ResNet18 trained on a dataset balanced via standard geometric augmentations.
3. **Proposed VAE Augmentation (Internal):** Adapted ResNet18 trained on a dataset balanced via VAE generations.
4. **ResNet152 (External):** We referenced the ResNet152 baseline reported by Wang et al. [7] to compare our lightweight backbone against deep standard architectures.
5. **State-of-the-Art (External):** We included the best-performing model by Wang et al. [7] to establish the upper bound, which is a significantly more complex and deeper architecture compared to our implementation.

4. Experimental results

Table 2. Performance comparison of Model (IDs in Section 3.2)

Model #	Accuracy	F1 Score	Recall	Precision
1	87.07%	79.52%	79.52%	79.52%
2	**87.83%**	**81.40%**	**84.34%**	78.65%
3	85.93%	76.43%	72.29%	**81.08%**
4	82.21%	77.02%	74.76%	83.46%
5	96.17%	95.57%	95.16%	96.03%

5. Conclusions

We investigated whether the VAE's inherent smoothing could act as regularization, prioritizing robust defect shapes over texture [2]. Our results confirm a clear trade-off: Generative Augmentation (Model 3) surpassed the Geometric Baseline in Precision (81.08% vs 78.65%) but suffered significantly in Recall. This performance shift reflects the architectural constraints discussed: the VAE acts as a low-pass filter, struggling to reconstruct the high-frequency details of minute defects. Consequently, the filtering of ambiguous samples necessary to maintain dataset integrity—biased the training set toward "evident" defects. This forced the classifier into a conservative regime, reducing false positives but missing subtle faults. Ultimately, while standard augmentation remains the robust baseline, effective generative strategies for micro-defects will likely require architectures that preserve high-frequency texture, such as GANs or Diffusion models, rather than standard VAEs.

6. Team Roles

- **Leonardo Sani**: Conceived the project scope and methodology. Responsible for the VAE development, from architectural design to the generation. Developed the data preprocessing and geometric augmentation frameworks. Additionally, designed and implemented the preliminary structural modifications of the ResNet18 classifier to be consistent with the VAE samples.
- **Simone Mantero**: Formalized the modified ResNet18 implementation into a reproducible experimental setup. He was responsible for executing the training, validation, and testing procedures across all the three defined scenarios to gather the final performance metrics.

References

[1] Sergiu Deitsch, Vincent Christlein, et al. Electroluminescence of photovoltaic cells (elpv) dataset. `https://github.com/zae-bayern/elpv-dataset`, 2019. 1, 2

[2] Robert Geirhos, Patricia Rubisch, Claudio Michaelis, Matthias Bethge, Felix A Wichmann, and Wieland Brendel. Imagenet-trained cnns are biased towards texture; increasing shape bias improves accuracy and robustness. In *International Conference on Learning Representations (ICLR)*, 2019. 2

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, Las Vegas, NV, USA, 2016. 1

[4] Usama Hijjawi, Haoyu Yang, Jiangnan Liu, Mohammad Aloqlah, Chris Bingham, and Phil Davies. A review of automated solar photovoltaic defect detection systems: Approaches, challenges, and future orientations. *Solar Energy*, 266:112186, 2023. 1

[5] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *Proceedings of the International Conference on Learning Representations (ICLR)*, Banff, AB, Canada, 2014. 1

[6] D. Matusz-Kalász, I. Bodnár, and M. Jobbágy. An overview of cnn-based image analysis in solar cells, photovoltaic modules, and power plants. *Applied Sciences*, 15(10):5511, 2025. 1

[7] Junjie Wang, Li Bi, Pengxiang Sun, Xiaogang Jiao, Xunde Ma, Xinyi Lei, and Yongbin Luo. Deep-learning-based automatic detection of photovoltaic cell defects in electroluminescence images. *Sensors*, 23(1):297, 2023. 1, 2

[8] Tom White. Sampling generative networks. *arXiv preprint arXiv:1609.04468*, 2016. 1, 2