



Politecnico di Bari

Repository Istituzionale dei Prodotti della Ricerca del Politecnico di Bari

Predictive maintenance for solar photovoltaic systems: deep learning-based anomaly identification using infrared thermography

This is a PhD Thesis

Original Citation:

Predictive maintenance for solar photovoltaic systems: deep learning-based anomaly identification using infrared thermography / Qureshi, Usamah Rashid. - ELETTRONICO. - (2025). [10.60576/poliba/iris/qureshi-usamah-rashid_phd2025]

Availability:

This version is available at <http://hdl.handle.net/11589/286340> since: 2025-04-15

Published version

DOI:10.60576/poliba/iris/qureshi-usamah-rashid_phd2025

Publisher: Politecnico di Bari

Terms of use:

(Article begins on next page)



Politecnico
di Bari

Department of Electrical and Information Engineering
ELECTRICAL AND INFORMATION ENGINEERING Ph.D. Program

SSD: IIND-08/B – ELECTRICAL POWER SYSTEMS

SSD: IBIO-01/A – BIOENGINEERING

Final Dissertation

**Predictive Maintenance for Solar Photovoltaic Systems:
Deep Learning-based Anomaly Identification using
Infrared Thermography**

by

Usamah Rashid QURESHI

Supervisors:

Prof. Massimo La Scala, Ph.D.

Prof. Vitoantonio Bevilacqua, Ph.D.

Dr. Nicola Altini, Ph.D.

Coordinator of Ph.D. Program:

Prof. Mario Carpentieri, Ph.D.



LIBERATORIA PER L'ARCHIVIAZIONE DELLA TESI DI DOTTORATO

Al Magnifico Rettore
del Politecnico di Bari

Il sottoscritto **Usamah Rashid QURESHI** nato a **Sebha, Libia (EE)** il **08/08/1985**

residente a **Bari (BA)** in via **Bartolo Longo n. 35** e-mail usamahrashid.queshi@poliba.it

iscritto al 3° anno di Corso di Dottorato di Ricerca in **Ingegneria Elettrica e dell'Informazione** ciclo **XXXVII**

ed essendo stato ammesso a sostenere l'esame finale con la prevista discussione della tesi dal titolo:

Predictive Maintenance for Solar Photovoltaic Systems: Deep Learning-based Anomaly Identification using Infrared Thermography

DICHIARA

- 1) di essere consapevole che, ai sensi del D.P.R. n. 445 del 28.12.2000, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici conseguenti al provvedimento emanato sulla base di tali dichiarazioni;
- 2) di essere iscritto al Corso di Dottorato di ricerca **Ingegneria Elettrica e dell'Informazione** ciclo **XXXVII**, corso attivato ai sensi del "Regolamento dei Corsi di Dottorato di ricerca del Politecnico di Bari", emanato con D.R. n.286 del 01.07.2013;
- 3) di essere pienamente a conoscenza delle disposizioni contenute nel predetto Regolamento in merito alla procedura di deposito, pubblicazione e autoarchiviazione della tesi di dottorato nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica;
- 4) di essere consapevole che attraverso l'autoarchiviazione delle tesi nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica del Politecnico di Bari (IRIS-POLIBA), l'Ateneo archiverà e renderà consultabile in rete (nel rispetto della Policy di Ateneo di cui al D.R. 642 del 13.11.2015) il testo completo della tesi di dottorato, fatta salva la possibilità di sottoscrizione di apposite licenze per le relative condizioni di utilizzo (di cui al sito <http://www.creativecommons.it/Licenze>), e fatte salve, altresì, le eventuali esigenze di "embargo", legate a strette considerazioni sulla tutelabilità e sfruttamento industriale/commerciale dei contenuti della tesi, da rappresentarsi mediante compilazione e sottoscrizione del modulo in calce (Richiesta di embargo);
- 5) che la tesi da depositare in IRIS-POLIBA, in formato digitale (PDF/A) sarà del tutto identica a quelle **consegnate**/inviata/da inviarsi ai componenti della commissione per l'esame finale e a qualsiasi altra copia depositata presso gli Uffici del Politecnico di Bari in forma cartacea o digitale, ovvero a quella da discutere in sede di esame finale, a quella da depositare, a cura dell'Ateneo, presso le Biblioteche Nazionali Centrali di Roma e Firenze e presso tutti gli Uffici competenti per legge al momento del deposito stesso, e che di conseguenza va esclusa qualsiasi responsabilità del Politecnico di Bari per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;
- 6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi, ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto il Politecnico di Bari ed i suoi funzionari sono in ogni caso esenti da responsabilità di qualsivoglia natura: civile, amministrativa e penale e saranno dal sottoscritto tenuti indenni da qualsiasi richiesta o rivendicazione da parte di terzi;
- 7) che il contenuto della tesi non infrange in alcun modo il diritto d'Autore né gli obblighi connessi alla salvaguardia di diritti morali od economici di altri autori o di altri aventi diritto, sia per testi, immagini, foto, tabelle, o altre parti di cui la tesi è composta.

Bari, 14/04/2025

Firma _____ 

Il sottoscritto, con l'autoarchiviazione della propria tesi di dottorato nell'Archivio Istituzionale ad accesso aperto del Politecnico di Bari (POLIBA-IRIS), pur mantenendo su di essa tutti i diritti d'autore, morali ed economici, ai sensi della normativa vigente (Legge 633/1941 e ss.mm.ii.),

CONCEDE

- al Politecnico di Bari il permesso di trasferire l'opera su qualsiasi supporto e di convertirla in qualsiasi formato al fine di una corretta conservazione nel tempo. Il Politecnico di Bari garantisce che non verrà effettuata alcuna modifica al contenuto e alla struttura dell'opera.
- al Politecnico di Bari la possibilità di riprodurre l'opera in più di una copia per fini di sicurezza, back-up e conservazione.

Bari, 14/04/2025

Firma _____ 



Politecnico
di Bari

Department of Electrical and Information Engineering
ELECTRICAL AND INFORMATION ENGINEERING Ph.D. Program

SSD: IIND-08/B – ELECTRICAL POWER SYSTEMS

SSD: IBIO-01/A – BIOENGINEERING

Final Dissertation

**Predictive Maintenance for Solar Photovoltaic Systems:
Deep Learning-based Anomaly Identification using
Infrared Thermography**

by

Usamah Rashid QURESHI

Referees:

Prof. Francesco Mercaldo, Ph.D.

Prof. Giuseppe Marco Tina, Ph.D.

Supervisors:

Prof. Massimo La Scala, Ph.D.

Prof. Vitoantonio Bevilacqua, Ph.D.

Dr. Nicola Altini, Ph.D.

Coordinator of Ph.D. Program:

Prof. Mario Carpentieri, Ph.D.



UNIONE EUROPEA
Fondo Sociale Europeo



La borsa di dottorato è stata cofinanziata con risorse del
Programma Operativo Nazionale Ricerca e Innovazione 2014-2020, risorse FSE REACT-EU
Azione IV.4 “Dottorati e contratti di ricerca su tematiche dell’innovazione”
e Azione IV.5 “Dottorati su tematiche Green”

Dissertation submitted for the degree of Research Doctorate in Electrical and Information Engineering (XXXVII cycle)

Title:

Predictive Maintenance for Solar Photovoltaic Systems: Deep Learning-based Anomaly Identification using Infrared Thermography

Ph.D. Candidate:

Usamah Rashid Qureshi, Politecnico di Bari (Bari, Italy)

Tutors:

Prof. Massimo La Scala, Ph.D., Politecnico di Bari (Bari, Italy)

Prof. Vitoantonio Bevilacqua, Ph.D., Politecnico di Bari (Bari, Italy)

Dr. Nicola Altini, Ph.D., Politecnico di Bari (Bari, Italy)

External Reviewers:

Prof. Francesco Mercaldo, Ph.D., Università degli Studi del Molise (Molise, Italy)

Prof. Giuseppe Marco Tina, Ph.D., Università di Catania (Catania, Italy)

Last version:

April 11, 2025

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Abstract

The purpose of this Ph.D. thesis is to portray deep learning applications for the Predictive Maintenance of Solar Photovoltaic Systems to identify Infrared Thermographic anomalies.

The research work carried out initiated the conceptualization, design, implementation, and evaluation of a novel Intelligent Fault Detection and Diagnosis (IFDD) System, alternatively realized as an Intelligent Remote Inspection (IRI) System for large-scale (and utility-scale) Solar Photovoltaic (SPV) Arrays having large surface areas, by deploying drone-aided Infrared Thermographic Imaging and process acquired thermal data by employing artificial intelligence (AI) techniques to precisely locate and capture SPV panels' thermal degradation pattern for operational decision-making, diagnostic monitoring and prompt maintenance.

SPV cells are fragile, and temperature is a critical environmental factor for their cascaded and cyclical thermal stresses and irreparable degradation, which are further accelerated by periodical and prolonged heatwave spells and global warming effects. The emergence of AI techniques, particularly deep learning (DL) algorithms leverage the realm of diagnostic monitoring and predictive maintenance (PdM). The thesis focuses on creating accessible, explainable end-to-end pipelines employing the DL framework, and drawing and preparing datasets from public repositories.

Besides, acquiring, preparing, and labeling a quality, calibrated, and noise-free SPV panel's thermal dataset (floating-point temperature intensity values in degrees Celsius) is a challenging task, as SPV energy-producing assets are expected to operate under healthy conditions and achieving targeted output continuously, resulting in fewer faulty labeled instances, whereas a large amount of data is required to train DL models for multiclass classification (or thermal diagnosis).

The thesis work is organized into five chapters: Chapter 1 introduces the objective of Predictive Maintenance (PdM), including Fault Detection and Diagnosis (FDD), Solar Photovoltaic (SPV) energy market, global warming effect, and contribution. Chapter 2 provides a comprehensive overview of the state-of-the-art Artificial Intelligence (AI) techniques in Solar Photovoltaic (SPV) diagnostics, including their application and

implementation challenges encountered while selecting and deploying AI algorithms, various Deep Learning (DL) models employed for diagnosis, explainability methods, and Infrared Radiated Thermographic (IRT) inspection for SPV power systems. Chapter 3 describes a theoretical framework for Convolutional Neural Network (CNN) comprising its architecture, layers, and training process. Chapter 4 describes the contributions proposed in Monitoring Diagnosis for Predictive Maintenance of Solar Photovoltaic Infrared Thermography. Lastly, Chapter 5 summarizes the work accomplished in this thesis and offers insights and considerations towards future works.

Keywords: Radiometric Infrared Thermography, Predictive Maintenance, Intelligent Fault Detection and Diagnosis, Intelligent Inspection, Solar Photovoltaic Systems

Table of Contents

Abstract.....	I
Table of Contents	III
List of Figures.....	V
List of Tables.....	IX
List of Acronyms.....	X
1 Introduction	1
1.1 Contribution.....	4
1.2 Thesis Outline.....	4
2 State of the Art.....	6
2.1 AI in SPV Diagnostics	6
2.1.1 Necessity and Overview.....	6
2.1.2 Specific Applications of AI in SPV Diagnostics	8
2.1.3 AI Techniques and Algorithms for SPV Diagnostics	10
2.1.4 Real-world Implementation and Case Studies	12
2.1.5 Challenges and Limitations of AI in SPV Diagnostics.....	13
2.2 Deep Learning	14
2.2.1 Classification.....	15
2.2.2 Explainable Artificial Intelligence Methods	21
2.3 SPV Power Systems and Intelligent Arial-IRT Inspection	26
2.3.1 SPV Power System Performance Optimization and the Role of Infrared Thermography	26
2.3.2 Standardized Guidelines and the Need for Skilled Expertise	27
2.3.3 Technical Prerequisites for Effective Outdoor-IRT Inspection	28
3 Theoretical Framework	29
3.1 Convolutional Neural Network	29
3.1.1 Input Layer.....	29
3.1.2 Convolutional Layer	31
3.1.3 Pooling Layer.....	37

3.1.4	Activation Function (ReLU)	38
3.1.5	Fully Connected Layer	39
3.1.6	SoftMax Layer	40
3.2	Training CNN	41
3.2.1	Loss Function	42
3.2.2	Backpropagation	44
3.2.3	Stochastic Gradient Descent	47
4	Monitoring Diagnosis for PdM of SPV-IRT	50
4.1	Deep Learning for IRT-Anomaly Diagnosis	50
4.1.1	Related Works	50
4.1.2	Materials	52
4.1.3	Experimental Resources	57
4.1.4	Experimental Workflow	57
4.1.5	Results and Discussion	59
4.1.6	Limitations	70
4.1.7	Conclusion	71
4.2	Deep Transfer Learning for IRT-Anomaly Diagnosis	71
4.2.1	Experimental Workflow	71
4.2.2	Results and Discussion	73
4.2.3	Limitations	76
4.2.4	Conclusion	77
5	Conclusions	79
	My Publications	81
	References	82

List of Figures

Figure 3.1 Architecture of AlexNet with 8 layers, where the first five blocks of layers are convolutional layers, and the rest are fully connected layers.....	30
Figure 3.2 Sample input layer: A digit 3 and 8 image with size 28×28 on the left and the corresponding pixel value on the image.....	30
Figure 3.3 Demonstration of input map and a filter: A 5×5 images with a single channel on the left and a 3×3 filter on the right. Note that each value from the image or filter is a pixel value.	32
Figure 3.4 Demonstration of padding: A matrix (or image) with zero-padding of size 1 ($P = 1$) (shaded).....	33
Figure 3.5 Demonstration of stride: The graph represents a movement of a filter with stride ($S = 2$) from the first position (a) to the next position (b).	34
Figure 3.6 Feature map extraction when kernel of 3×3 convolve an input tensor of 5×5 with stride 1.	34
Figure 3.7 Examples of how kernels in convolution layers extract features from an input tensor are shown. Multiple kernels work as different feature extractors, such as a horizontal edge detector (top), a vertical edge detector (middle), and an outline detector (bottom). Note that the left image is an input, those in the middle are kernels, and those in the right are output feature maps.	36
Figure 3.8 Result of reshaping: The red area refers to the first filter area in Channel 1 of the input image, the green area refers to the first filter area in Channel 2 of the input image. Similarly, the blue area comes from the first filter area of Channel 3 of the image	37
Figure 3.9 Demonstration of max pooling operation: A sample max pooling operation of an input feature map of size 5×5 with filter size ($R = 2$), stride ($S = 1$) and no padding resulted in an output of size 4×4	38
Figure 3.10 Three most commonly used activation functions in the hidden layers: ReLU, tanh, and sigmoid.....	39
Figure 3.11 Demonstration of ReLU operation: A ReLU operation on a sample matrix with input on the left and output on the right.	39
Figure 3.12 Example of loss computation: A demonstration of how the loss being computed from an input x_i of size 3×1	43

Figure 3.13 Example of backpropagation process: An example of finding $\partial y/\partial x$ (2) by applying the backpropagation algorithm. The blue number indicates the forward direction, while the red number indicates the backward direction started from the end to the beginning.	45
Figure 3.14 Example of backpropagation process in neural network: Acyclic computation graph of the network model $f(x) = g(W^2 \times g(W^1 \times x + b^1) + b^2)$, where g is a ReLU function ($f(x) = \max(0, x)$), $W = [1, -1]$, $b = [1, 5]$ and $x = 3$. The forward propagation computes values from input to output (blue) and the backpropagation is performed from the end to the beginning by applying the chain rule and the derivative rules (red)	46
Figure 3.15 Effects of various learning rates on convergence	49
Figure 4.1 Raw aerial IRT-grayscale images of the SPV array having a visually defective SPV panel labeled as (a) hotspot effect, (b) patchwork pattern, (c) faulty substring.....	53
Figure 4.2 Extracted SPV panel labeled as heated Junction Box: Original and magnified region of 2D radiometric thermal datapoints—temperature intensity matrix [$^{\circ}\text{C}$] (jet colormap as pseudo-color applied for visual depiction).....	53
Figure 4.3 Original SPV panel image of 100×60 pixels (left) and zero-padded 106×66 pixels (right) having maximum temperature [$^{\circ}\text{C}$] with jet colormap as pseudo-color applied for visual depiction (red shades corresponds to higher temperature values, whereas blue to lower ones).	54
Figure 4.4 Original dataset of 6-class SPV panels with jet colormap applied for pseudo-color visual depiction (max temperature [$^{\circ}\text{C}$]). Color bar represents temperature values corresponding to colors (higher temperature values in red shades, lower temperature values in blue shades). Data augmentation applied for visual diversity (multistring, substring, and hotspot).	55
Figure 4.5 Raincloud plot of dataset (Good and Faulty class)	56
Figure 4.6 Raincloud plot of dataset (5-Faulty class).....	56
Figure 4.7 Experimental workflow. First, the dataset is acquired with a drone-aided IRT of SPV arrays. Then, an ensemble of CNN models is realized via a stratified 4-fold cross-validation. Finally, quantitative results are determined, and explainability techniques are used to unveil the mechanisms underlying the diagnostic process. Heatmaps have a jet colormap, with red representing higher temperature (for radiometric images) or activation values (for CNN explanations), and blue depicting lower temperature or activation values.	58

Figure 4.8 Explainable deep learning - Activation Maximization visualization from the last convolutional layer for the 6-class classification. Results for each iteration of the stratified 4-fold cross-validation are presented. Jet pseudo-color (red corresponds to higher activation values, whereas blue to lower ones) is used for visual depiction.	61
Figure 4.9 Explainable deep learning — SmoothGrad visualization from the last convolutional layer for the 6-class classification. Results for each iteration of the stratified 4-fold cross-validation are presented. Jet pseudo-color (red corresponds to higher activation values, whereas blue to lower ones) is used for visual depiction.	62
Figure 4.10 Explainable deep learning — Grad-CAM visualization from the last convolutional layer for the 6-class classification. Results for each iteration of the stratified 4-fold cross-validation are presented. Jet pseudo-color (red corresponds to higher activation values, whereas blue to lower ones) is used for visual depiction.	63
Figure 4.11 Embedding with t-SNE (cluster map) of complete input radiometric dataset images. Each data-point is depicted as a star-shaped point. Kindly note that some points overlap.	64
Figure 4.12 Embedding with UMAP (feature clustering) of test dataset predictions (Hellinger metric)—ensembled. Each data point is depicted as a star-shaped point. Kindly note that some points overlap.	65
Figure 4.13 Confusion matrices of class-weighted test data for each iteration of the stratified 4-fold cross-validation (6-class).	66
Figure 4.14. Confusion matrix of class-weighted test data ensemble model (6-class).	66
Figure 4.15 Categorical loss and accuracy of the training and validation of class-weighted test dataset for 6-class output, for each iteration of the stratified 4-fold cross-validation. .	67
Figure 4.16 Confusion matrices of SMOTE test data for each iteration of the stratified 4-fold cross-validation (6-class).	69
Figure 4.17 Confusion matrix of SMOTE test data—ensemble model (6-class).	69
Figure 4.18. Categorical loss and accuracy of the training and validation of SMOTE test dataset for a 6-class output, for each iteration of the stratified 4-fold cross-validation.	70
Figure 4.19. Experimental workflow (simplified block diagram).....	72
Figure 4.20 Explainable deep learning – Guided SmoothGrad visualizations from the last convolutional layer for 6-class classification (jet colormap as pseudo-color applied for visual depiction). Red corresponds to higher activation values, whereas blue to lower ones.	75

Figure 4.21 Embedding with UMAP (feature clustering) of test dataset predictions (Hellinger Metric). Each data point is depicted as a star-shaped point. Kindly note that some points are overlapped..... 76

List of Tables

Table 2.1 AI techniques and its applications.....	11
Table 4.1 List of DL Models	72
Table 4.2 DL Models Performance	74

List of Acronyms

AC	Alternating Current
AI	Artificial Intelligence
ANNs	Artificial Neural Networks
BN	Batch Normalization
CNNs	Convolutional Neural Networks
CV	Computer Vision
DBNs	Deep Belief Networks
DC	Direct Current
DL	Deep Learning
DT	Decision Trees
DTL	Deep Transfer Learning
FDD	Fault Detection and Diagnosis
GA	Genetic Algorithm
GANs	Generative Adversarial Networks
GBP	Guided Backpropagation
GFI	Ground-Fault Interrupters
GPU	Graphical Processing Unit
Grad-CAM	Gradient-Weighted Class Activation Mapping
IFDD	Intelligent Fault Detection and Diagnosis
IRI	Intelligent Remote Inspection
IRT	Infrared Radiated Thermography/Thermographic
I-V	Current-Voltage
KNN	K-Nearest Neighbor

LIST OF ACRONYMS

LCOE	Levelized Cost of Electricity
LR	Logistic Regression
ML	Machine Learning
MPP	Maximum Power Point
MPPT	Maximum Power Point Trackers
NAS	Neural Architecture Search
NDT	Non-Destructive Testing
NLP	Natural Language Processing
NN	Neural Networks
O&M	Operations and Maintenance
OCP	Overcurrent Protection
PdM	Predictive Maintenance
PSO	Particle Swarm Optimization
P-V	Power-Voltage
RAM	Random Access Memory
RBM	Restricted Boltzmann Machines
ReLU	Rectified Linear-Activation Unit
RF	Random Forest
RGB	Red, Green, and Blue
RL	Reinforcement Learning
RoI	Return on Investment
RUL	Remaining Useful Life
SGD	Stochastic Gradient Descent
SME	Stochastic Neighbor Embedding
SMOTE	Synthetic Minority Over-Sampling Technique
SPV	Solar Photovoltaic

SVM	Support Vector Machine
TL	Transfer Learning
t-SNE	t-distributed Stochastic Neighbor Embedding
UMAP	Uniform Manifold Approximation and Projection
USD	US dollar
VGG	Visual Geometry Group
ViTs	Vision Transformers
XAI	Explainable Artificial Intelligence

Chapter I

Introduction

Predictive Maintenance (PdM) is condition-based maintenance (EN 13306:2017) [1] that harnesses advanced machine learning (ML) and artificial intelligence (AI) techniques to proactively identify and address failure symptoms and anomalies, ensuring the smooth operation of critical systems. These tools excel in detecting subtle patterns and deviations in complex data, offering precise insights into anomalies and potential performance issues [2]. At the heart of an effective PdM strategy lies operational fault detection and diagnosis (FDD), a cornerstone for maintaining modern industrial and renewable energy systems' efficiency, quality, and reliability.

Traditional FDD approaches, which rely heavily on human expertise and diagnostic experience, face scalability challenges in large-scale factories and electric power plants where process variables proliferate [3]. In Solar Photovoltaic (SPV) systems, conventional FDD methods such as overcurrent protection (OCP) devices and ground-fault interrupters (GFI) often fall short. Factors like low solar irradiance, the nonlinear output characteristics of SPV panels, the presence of maximum power point trackers (MPPT) in inverters, and high fault impedances hinder the accurate detection of specific faults.

To address these limitations, the focus is shifting towards intelligent fault detection and diagnosis (IFDD) techniques. AI-driven methods, particularly those leveraging ML and deep learning (DL), are emerging as powerful alternatives to conventional FDD approaches [4]. These intelligent systems, designed for tasks like diagnosing SPV infrared radiated thermographic (IRT) images, can analyze complex and dynamic patterns, adapt to variable operating conditions, and detect thermal anomalies with high precision, even under challenging scenarios. Further, the large-scale (and utility-scale) SPV electric power systems occupy large surface areas of SPV arrays, where intelligent remote inspection (IRI) employing drone-aided IRT imaging and advanced AI techniques makes diagnostic monitoring efficient, accurate, and speedy. The transition to IFDD not only enhances

anomaly classification capabilities but also drives significant advancements in PdM for renewable energy technologies. This integration of AI-based solutions heralds a new era of operational reliability, efficiency, and innovation, underscoring the critical role of intelligent systems in the sustainable energy landscape.

In the SPV energy market as of 2023, the global renewable energy landscape witnessed a significant milestone with the addition of 576 gigawatts of net renewable power generating capacity. SPV systems dominated this growth, accounting for 78% (447 gigawatts) of the newly installed capacity [5], reflecting their pivotal role in the transition to sustainable energy. Within the European Union (EU27), SPV installation capacity surged by 53.124 gigawatts, bringing the cumulative installed capacity to 256.911 gigawatts. This achievement positioned SPV to exceed 44% of the EU's total renewable electricity generation share, highlighting its critical contribution to the region's energy mix [6], [7], [8].

From an economic perspective, investments in SPV power generation reached unprecedented levels, totaling USD 480 billion globally in 2023, surpassing all other electricity generation technologies, including fossil fuels (oil, natural gas, coal) and alternative renewables (hydro, wind, nuclear), which collectively accumulated USD 413 billion [9]. Given the 25-year economic lifespan of SPV panels, utility-scale SPV energy projects have become increasingly cost-efficient, with those commissioned in 2022 boasting a global weighted average total capacity installation cost of USD 876/kW (4.4% < USD 917/kW in 2021), and the levelized cost of electricity (LCOE) for SPV electricity generation also declined, reaching USD 0.049/kWh (3.9% < USD 0.051/kWh in 2021) [10].

Meanwhile, the escalating impact of climate change, particularly the intensifying effects of global warming, has led to more frequent, prolonged, and severe heatwaves that begin earlier, persist longer, and affect wider regions concurrently [11], [12]. Recent examples highlight the intensity of the situation: temperatures reached a staggering 48.2°C on 24 July 2023 in Lotzorai and Jerzu (Sardinia, Italy), while Agadir (Morocco) recorded a record-breaking 50.4°C on 11 August 2023 [13]. These extreme heat events not only strain human health and ecosystems but also place immense stress on critical electric power infrastructure. During heatwaves, the demand for electricity surges dramatically as households and businesses rely heavily on refrigeration, air conditioning, and cooling systems to mitigate the scorching temperatures. This increased load pushes energy systems to their limits, creating significant risks of systematic overloading and supply disruptions precisely when electricity is most critical for essential services [14], [15], [16]. The cascading effects of such

failures are profound: infrastructure breakdowns during heat extremes can escalate local energy crises into widespread emergencies, further amplifying the heatwave's burden. The consequences extend beyond power outages, jeopardizing vital humanitarian and health operations. For instance, the interruption of electricity-dependent cold storage facilities can compromise the preservation of essential life-saving operations and supplies, including medications, and vaccines [12]. Such disruptions not only delay response efforts but also intensify the vulnerabilities of communities already reeling from extreme heat, underlining the urgent need for resilient energy systems capable of withstanding the growing challenges posed by climate change.

The delicate nature of SPV cells [17] makes temperature a critical environmental variable that significantly accelerates degradation mechanisms, particularly those influenced by permeation processes such as chemical reactions and diffusion. SPV modules are designed with specific operating conditions, typically maintaining a nominal operating cell temperature of 42 to $43 \pm 2^\circ\text{C}$ [18], [19], [20], [21]. However, internal cell temperatures can deviate from ambient conditions, primarily driven by incident solar irradiance, wind speed, and other environmental factors. Thermal degradation in SPV modules arises from three primary causes. First, *deterioration processes* such as thermo-chemical reactions [22], [23] and thermo-mechanical stresses [24], [25], directly impact the structural and chemical integrity of the module. Second, *cyclic aging conditions* [26], [27], including diurnal and seasonal temperature variations, induce wear over time amplifying the aging effects. Third, *non-uniform sunlight passage* caused by environmental factors [28], [29] such as partial shadowing, soiling, dusting, cracked glass, or biological debris like bird droppings disrupt the thermal uniformity of the module. These factors exacerbate heat accumulation, leading to hotspots and accelerated degradation.

The thermal conductivity and geometric configuration of surrounding materials, wind flow, and installation design influence heat dissipation within SPV modules. Discrepancies in the thermal expansion coefficients of module materials further contribute to thermal stresses. These stresses result in differential expansion and contraction, which can compromise the mechanical stability of critical electrical components such as cells, solder joints, and interconnect ribbons. This may lead to issues like deformation, delamination, and cell cracking. Additionally, the cyclic thermal stresses caused by temperature fluctuations can result in fatigue-induced failures that may propagate across various components [30], [31], [32], [33]. Thermal-related failures in SPV modules often cascade, resulting in

irreversible damage. Such failures diminish module efficiency and lifespan, underscoring the critical need for proactive thermal management. Here, thermographic inspection plays a pivotal role in this context by providing a non-invasive method to assess the heat signature and temperature profile of SPV panel surfaces. By enabling the detection of hotspots and temperature anomalies, thermography facilitates the classification, prediction, and mitigation of potential defects, helping to ensure the reliability and longevity of SPV energy systems in the field.

1.1 Contribution

Within the PdM framework of renewable energy systems, my primary objective focused on the extraction of quantitative thermal information (floating-point temperature intensity numerical values in degrees Celsius) and analysis of drone-aided radiometric IRT images for diagnostic monitoring of SPV electric power systems. Initially, a highly imbalanced six-class radiometric IRT dataset (thermal matrix) of SPV panels was prepared based on IEC TS 62446-3:2017 standards [34] from raw grayscale images of SPV arrays to visually identify and label thermal faults, and then a customized DL ensemble model based on a convolutional neural network (CNN) was developed for thermal fault multiclass-classification (or thermal anomaly diagnosis) and compared the results with pre-trained CNN models. Further, explainable artificial intelligence (XAI) techniques are applied to elaborate the dataset pattern and prediction outcome.

1.2 Thesis Outline

The thesis work is organized as follows: This Chapter introduces the objective of PdM, including FDD, SPV energy market, global warming effect, and the contribution to the thesis.

Chapter 2 provides a comprehensive overview of the state-of-the-art AI techniques in SPV diagnostics, including their application and implementation challenges encountered while selecting and deploying AI algorithms, various DL models employed for diagnosis, explainability methods, and IRT inspection for SPV power systems.

Chapter 3 describes a theoretical framework for CNN comprising its architecture, layers, and training process.

Chapter 4 describes the contributions proposed in Monitoring Diagnosis for PdM of SPV-IRT.

Lastly, Chapter 5 summarizes the work accomplished in this thesis and offers insights and considerations towards future works.

Chapter II

State of the Art

2.1 AI in SPV Diagnostics

Solar energy is a pivotal renewable energy source, attracting increasing investments from businesses and governmental bodies worldwide [35]. This surge in adoption of AI techniques underscores the critical need for ensuring the sustained efficiency and operational longevity of SPV energy systems through robust fault detection and diagnostic mechanisms [36]. As the global community increasingly transitions towards sustainable energy solutions, SPV power systems have become a cornerstone of this movement. The escalating installed capacity of SPV energy installations necessitates the development and implementation of efficient and reliable diagnostic methodologies to guarantee optimal performance and maximize the return on investment (RoI). Traditional diagnostic approaches, often reliant on manual inspections and basic electrical measurements, are increasingly proving inadequate to address the inherent complexities of the vast scale of SPV energy system installations [37].

2.1.1 Necessity and Overview

Necessity: The growing dependence on solar energy as a primary contributor to the utility energy mix highlights the importance of advanced diagnostic tools in preventing significant energy losses and maintaining the stability of the electrical grid [38]. Undetected faults within SPV systems can lead to substantial reductions in energy generation, impacting both the economic viability of SPV energy projects and the overall reliability of the power supply. Consequently, the integration of AI into SPV diagnostics emerges as a crucial advancement in this field. This integration offers the potential to move beyond the limitations of traditional methods, providing sophisticated capabilities for analyzing system performance, identifying anomalies, and predicting potential failures with greater speed and accuracy [37]. The economic and environmental ramifications of overlooked malfunctions

within SPV arrays underscore the significant role that AI can play in ensuring the continued success and expansion of SPV power generation.

The broader movement towards a cleaner energy future is inextricably linked to the efficiency and dependability of SPV power systems. AI-driven diagnostic techniques hold the potential to significantly accelerate this transition by minimizing system downtime and maximizing the energy yield from both existing and future SPV installations [36], [37]. By enabling proactive and even predictive maintenance strategies, AI contributes directly to the overall sustainability and effectiveness of SPV energy systems as a key technology in mitigating climate change and meeting the world's growing energy demands [39].

Overview: AI, with its advanced analytical and ML capabilities, offers a comprehensive suite of tools that can fundamentally optimize the performance of SPV systems, enhance their maintenance and safety protocols, and ultimately maximize the production of clean energy. Specifically, within the realm of SPV diagnostics, AI is being applied to address some of the most critical challenges associated with these systems. These challenges include accurately tracking the maximum power point (MPP) of SPV modules, reliably forecasting the energy produced by the system, precisely estimating the parameters of equivalent circuit models for SPV modules, and effectively detecting faults within SPV modules or individual cells [35].

The versatility inherent in AI technologies allows for their application across a wide spectrum of crucial aspects in SPV system management. This extends beyond mere fault detection to encompass the overall optimization of system design, improved maintenance scheduling, enhanced information security, real-time performance monitoring, dynamic adjustment of SPV panel orientation, and the implementation of intelligent energy management systems. This broad applicability signifies a shift towards a more holistic and integrated approach to managing and securing SPV energy assets throughout their entire lifecycle [40].

The integration of AI into SPV diagnostics represents a fundamental change in how these systems are managed. Moving away from traditional *reactive* maintenance approaches, AI enables the transition towards *proactive* and even *predictive* management strategies for SPV energy assets. This shift leads to significant improvements in overall system efficiency and a substantial reduction in operational costs. By leveraging the power of data analysis and ML, AI enables operators to anticipate and address potential issues before they escalate, ensuring the continuous and optimal performance of SPV power plants.

2.1.2 Specific Applications of AI in SPV Diagnostics

2.1.2.1 Fault Detection and Classification

AI plays a crucial role in analyzing the operational data of SPV systems to identify faults in panels, inverters, or other critical components at an early stage. This capability is particularly important considering the inherent challenges in fault detection within SPV arrays due to their current-limiting nature and nonlinear output characteristics. To address this, AI techniques often involve comparing the performance of potentially faulty modules with accurate models, sometimes referred to as their “factory fingerprint,” by analyzing their current-voltage (I-V) and power-voltage (P-V) curves using algorithms such as artificial neural networks (ANN) [41].

Various AI techniques are employed for the purpose of fault detection in SPV systems, including classical machine learning (ML) methods like K-nearest neighbor (KNN) and random forest (RF), as well as more advanced DL models such as CNNs. These algorithms are adept at processing and analyzing large volumes of data to discern complex patterns and anomalies that may indicate the presence of a fault. Furthermore, AI can be utilized to detect and categorize a wide range of fault types, encompassing physical issues such as cracks, corrosion, and discoloration; electrical problems including direct current (DC) and alternating current (AC) faults; and environmental factors like hotspots and partial shading [37].

The ability of AI to learn intricate patterns from extensive datasets enables the detection of subtle and gradually developing faults that might escape the notice of traditional diagnostic methods. This early identification allows for timely intervention, preventing further damage to the system and minimizing potential energy losses. Moreover, the accurate classification of faults by AI not only helps in pinpointing the specific problem but also contributes to a deeper understanding of the underlying causes. This knowledge facilitates more targeted maintenance interventions and can inform improvements in the design and operational strategies of future SPV energy installations.

2.1.2.2 Performance Prediction and Optimization

AI is instrumental in analyzing both historical and real-time data to optimize the sizing and configuration of SPV systems, thereby enhancing their design and overall energy production [42]. AI algorithms can also effectively forecast energy demand and dynamically adjust the output of SPV panels to meet these needs. By analyzing comprehensive weather

data, including factors like cloud cover, temperature, and wind patterns, AI can accurately predict the energy output of SPV installations and make necessary adjustments to panel orientation to maximize energy capture. ML models further contribute to performance analysis by comparing the expected power output with the actual generated power, highlighting any unusual deviations that may indicate underlying issues.

The accurate prediction of energy generation through AI enables more effective grid management and facilitates a smoother integration of solar power into the existing energy infrastructure. This enhanced predictability contributes to making solar energy a more reliable and stable source of power. Furthermore, the optimization capabilities of AI extend beyond merely adjusting individual panels. AI can optimize the entire energy system, encompassing energy storage solutions and distribution networks, leading to a more efficient utilization of generated energy and a reduction in overall energy waste. This holistic approach ensures that SPV power systems are designed and operated to meet expected energy needs effectively.

2.1.2.3 Anomaly Detection and Diagnosis

AI algorithms possess the capability to detect subtle anomalies and deviations in the performance of SPV systems that might not be readily apparent to human operators. This is facilitated by the real-time monitoring and analysis of system performance data enabled by AI. By continuously scrutinizing various operational parameters, AI can identify patterns or irregularities that signify potential faults or inefficiencies within the system.

The early detection of anomalies through AI intervention is critical in preventing minor issues from escalating into more significant failures [36], [37]. This proactive approach can substantially reduce system downtime and associated repair costs. Moreover, by analyzing the patterns of detected anomalies over extended periods, AI can provide valuable insights into the degradation patterns of different SPV components. This information can then be used to formulate effective maintenance schedules and optimize component replacement strategies, ultimately extending the lifespan of the SPV system and maximizing its overall RoI.

2.1.2.4 Predictive Maintenance and Remaining Useful Life Estimation

AI plays a pivotal role in predicting potential failures within SPV systems, allowing for the proactive planning of preventative maintenance activities. By analyzing historical performance data and identifying recurring patterns, AI algorithms can forecast when a

particular SPV panel or system component is likely to experience a failure or require maintenance. This predictive capability extends to energy storage systems as well, where AI models can analyze battery charge and discharge cycles to estimate their remaining useful life (RUL) [43], [44], [45].

The transition from reactive maintenance, where issues are addressed only after they occur, to predictive maintenance, facilitated by AI, results in significant cost savings. This is achieved by avoiding unplanned system downtime and optimizing the timing of maintenance interventions, ensuring that maintenance is performed only when necessary. Furthermore, the accurate prediction of the remaining lifespan of critical components enables better long-term financial planning for SPV energy projects. This includes reliable and accurate budgeting for future replacements and optimized management of operational expenditures throughout the entire lifecycle of the SPV power system installation.

2.1.3 AI Techniques and Algorithms for SPV Diagnostics

ML and DL techniques are particularly well-suited for applications in SPV fault detection due to their inherent capability to process and analyze large volumes of complex data [46]. These AI approaches enable the identification of intricate patterns and subtle anomalies that may indicate underlying faults or performance issues. A diverse range of AI techniques is currently employed in SPV diagnostics, spanning from classical ML methods to more advanced DL architectures and even emerging technologies like quantum circuits [37].

Traditional ML algorithms such as Logistic Regression (LR), Decision Trees (DTs), and Support Vector Machines (SVMs) are also utilized for various diagnostic tasks, including fault classification. Neural networks (NNs) play a significant role in real-time data analysis, enabling predictive maintenance capabilities within SPV energy installations. Furthermore, optimization algorithms like Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) are employed in Maximum Power Point Tracking (MPPT) and hold potential for optimizing diagnostic processes as well [37], [47].

The selection of a specific AI technique is often dictated by the nature of the diagnostic problem, the characteristics and volume of the available data, and the desired balance between accuracy and interpretability. Different algorithms possess unique strengths and weaknesses when it comes to handling various types of faults and the complexities inherent in SPV system data. For instance, some techniques excel at classification tasks, while others

are particularly effective in analyzing image data or optimizing system parameters. The following table provides a comparative overview of some of the key AI techniques used in SPV diagnostics:

Table 2.1 AI techniques and its applications

AI Technique	Application in SPV Diagnostics	Advantages	Limitations	References
Artificial Neural Network (ANN)	Fault detection, performance prediction, model comparison	Effective for nonlinear data, can learn complex relationships	Can be computationally intensive, requires large data sets	[41], [48]
Convolutional Neural Network (CNN)	Fault detection and classification using thermal images or visual data	Excellent for image recognition and analysis	Requires large labeled image datasets, can be less interpretable	[37]
Support Vector Machine (SVM)	Fault classification	Effective in high-dimensional spaces, robust against overfitting in some cases	Can be computationally expensive for large datasets, kernel selection is crucial	[36], [46]
K-Nearest Neighbors (KNN)	Fault classification	Simple to implement, no training phase	Computationally expensive for large datasets, sensitive to feature scaling	[36]
Random Forest (RF)	Fault classification, performance prediction	Robust to noise and outliers, handles large feature sets effectively	Can be less interpretable than Decision Trees	[36]
Logistic Regression (LR)	Binary fault classification	Simple and efficient, provides probability estimates	Limited to linear relationships	[36]
Decision Trees (DT)	Fault classification	Highly interpretable and visualizable	Can be prone to overfitting	[36], [49]
Particle Swarm Optimization (PSO)	Maximum Power Point Tracking (MPPT), potentially fault diagnosis optimization	Effective in finding optimal solutions in complex search spaces	Can converge slowly in some cases	[47]

Genetic	Maximum Power	Effective in	Can be	[47]
Algorithm (GA)	Point Tracking (MPPT), potentially fault diagnosis optimization	exploring large solution spaces	computationally expensive and may not guarantee the global optimum	

Emerging as a promising direction in the field are hybrid AI models that integrate different techniques [50], [51]. These approaches aim to capitalize on the strengths of individual algorithms, potentially leading to even higher levels of accuracy and robustness in the performance optimization and diagnosis of faults within SPV energy systems. By combining the capabilities of various AI methodologies, researchers are striving to develop more comprehensive and effective diagnostic solutions for the SPV energy industry.

2.1.4 Real-world Implementation and Case Studies

2.1.4.1 AI-powered Drones for Solar Farm Inspection

A significant real-world application of AI in SPV diagnostics involves the use of AI-powered drones for the inspection of solar PV farms. These drones, equipped with high-resolution cameras, can efficiently survey extensive SPV panel arrays to identify various issues such as defects, accumulated debris, and shading problems that could negatively impact energy production. Furthermore, infrared drones integrated with AI algorithms can detect thermal anomalies like hotspots and poor electrical connections at an early stage of their development [52], [53]. This early detection is crucial for preventing minor issues from escalating into more significant problems [54], [55].

The speed and efficiency of drone-based inspections are particularly noteworthy. Drones can inspect entire SPV power plants within a few hours, a task that would take significantly longer using traditional manual methods. The raw data collected by these drones, including visual and thermal imagery, is then analyzed by specialized AI software. This software can automatically identify and geo-locate faulty modules within the SPV energy plant, providing operators with precise information for targeted maintenance interventions [56], [57], [58], [59].

The case studies have demonstrated the substantial benefits of employing AI-powered drones for SPV farm inspections. For instance, one study reported achieving 97% operational efficiency along with significant cost savings compared to manual inspection

methods. The combination of drone technology and AI not only enhances the speed and efficiency of inspections but also improves their accuracy, enabling proactive maintenance and minimizing downtime for large-scale SPV energy installations. This approach is driving a transition towards more automated and data-driven maintenance practices within the SPV power industry, ultimately leading to improved asset management and reduced operational expenditures [60], [61], [62].

2.1.4.2 AI-based Software and Platforms for Performance Monitoring and Fault Detection

Several companies are at the forefront of developing and implementing AI-based software and proprietary platforms for the real-world performance monitoring and fault detection in SPV systems. Companies such as Mipu (Italy) and Sun'Agri (France) utilize AI to predict energy production from SPV installations with high accuracy and optimize the overall efficiency of these systems [63], [64], [65]. Globally, SmartHelio offers AI-powered diagnostic solutions that boast high accuracy rates in both fault detection and classification [66], [67]. Their platform focuses on automating traditionally manual diagnostic tasks, providing actionable insights to asset managers in significantly reduced timeframes.

A compelling case study from SmartHelio illustrates the transformative impact of AI in SPV diagnostics. For a megawatt-scale SPV farm, their AI-powered diagnostic process reduced the analysis time from an astounding 358 hours to a mere 3 hours [66]. Furthermore, the platform successfully identified hard-to-detect faults, such as subtle damage to DC wiring and module mismatches, which are often overlooked during conventional manual inspections. The emergence and successful implementation of these specialized AI tools and platforms underscore the growing recognition of the immense value that AI brings to optimizing the management of SPV assets, providing operators with the critical insights needed to enhance efficiency and minimize operational costs.

2.1.5 Challenges and Limitations of AI in SPV Diagnostics

2.1.5.1 Dataset Scarcity

While the application of AI in SPV diagnostics offers numerous benefits, several challenges and limitations need to be addressed. One primary limitation lies in the substantial amount of data and the significant computational resources required to effectively train AI models [35]. Furthermore, a notable challenge in the field is the scarcity of publicly available

datasets specifically designed for SPV fault detection [36], [37]. This lack of standardized data hinders the development, validation, and benchmarking of AI models, making it difficult to compare the performance of different approaches objectively and potentially limiting the broader adoption of these technologies.

2.1.5.2 Computational Needs and Weather Fluctuation

The complexity of some advanced AI methods can also pose a challenge, as certain algorithms may be too computationally intensive or too intricate for practical, real-time implementation in all scenarios [36], [37].

Another difficulty arises in accurately distinguishing between an actual fault within the SPV system and a temporary reduction in energy production due to external factors such as fluctuating weather conditions [68]. The intricate nature of large-scale SPV systems, with their multitude of interconnected components and potential points of failure, also presents a considerable challenge for AI algorithms to effectively monitor and diagnose issues.

2.1.5.3 Data Security and AI Model Transparency

Ensuring the privacy and security of the vast amounts of sensitive energy data collected by AI-powered diagnostic systems is another critical concern that needs to be carefully managed [62], [67]. Developing robust and reliable AI algorithms tailored for solar PV panel optimization and diagnostics requires specialized expertise in both AI and SPV energy technologies, which can be a limiting factor for researchers and organizations.

Additionally, the effectiveness of AI analysis can be hampered by older or inconsistent data management systems that may lack the standardized, clean datasets necessary for accurate model training and performance [66], [67]. Finally, some sophisticated AI models, particularly deep learning (DL) architectures, can operate as “black boxes”, meaning their decision-making processes are not easily interpretable. This lack of transparency can be a limitation in critical applications like fault diagnosis, where understanding the reasoning behind a prediction is essential for building trust and verifying the accuracy of the results [36].

2.2 Deep Learning

Within the domain of ML methodologies rooted in ANNs, DL has emerged as a cornerstone, emphasizing the development of computational models capable of processing raw data through multiple layers of abstraction. These techniques enable the automated

discovery of complex patterns, making them especially valuable in scenarios where designing handcrafted features is either infeasible or inefficient. Deep learning demonstrates remarkable potential in fields such as computer vision, intelligent diagnosis, and speech recognition, where the end-to-end learning of hierarchical features is pivotal [69]. By automatically extracting representations at various levels of granularity, these methods eliminate the dependency on manual feature engineering, thus allowing for more robust and adaptable solutions.

Prominent DL architectures including CNN's [70], Autoencoders, Deep Belief Networks (DBN's), and Restricted Boltzmann Machines (RBM's) have proven their versatility across a wide range of big data applications. These include fault detection and diagnosis [71], [72], [73], pattern recognition, natural image processing, natural language processing (NLP), computer vision, and recommender systems [74], [75]. Their ability to model intricate relationships and extract meaningful insights from vast and unstructured data underscores their transformative role in modern data-driven technologies, unlocking new opportunities for innovation and efficiency.

2.2.1 Classification

Classification is the fundamental computer vision task related to pattern recognition and image classification. Image classification involves assigning a single label to an entire image, typically representing the most prominent object within the frame. The methodologies for achieving image classification differ significantly depending on the approach employed. The workflows for image classification vary depending on the chosen classification algorithm. In traditional approaches, the process is twofold: handcrafted features are first extracted from the image (e.g. edges, textures, or shapes), and these features are then fed into a statistical classifier, such as a SVM or DT, which assigns the label based on predefined rules or learned patterns.

Modern advancements, driven by DL have revolutionized this process by enabling end-to-end classification. DL models such as CNNs bypass the need for manual feature extraction or learning hierarchical representations of the input data directly from the raw image. These models take the image as input and produce the label in a single integrated workflow, offering superior performance and adaptability, particularly in complex or large-scale datasets. This paradigm shift highlights the transformative impact of DL on computer vision tasks.

2.2.1.1 Convolutional Neural Networks

CNNs, first introduced by LeCun *et al.* [76], have become indispensable in the domain of natural image classification, particularly over the last decade. A typical CNN architecture comprises two main components, each serving a distinct purpose. The *upstream feature extraction module* forms the foundation of the CNN. This module leveraged convolutional layers, pooling layers, and other specialized architectural enhancements such as skip connections, dense connections, inception modules, and batch normalization. These elements work in tandem to extract meaningful patterns and features from raw input images, enabling the model to learn hierarchical representations. This process is critical for capturing both low-level details (e.g. edges and textures), and high-level abstractions (e.g. object shapes and structures). The *downstream classification module* is simpler by design, serving to translate the extracted features into actionable outputs. It employs a conventional ANN that uses the final feature map as input to perform the classification task, assigning labels to images based on learned patterns.

An effective strategy for leveraging CNN's is transfer learning (TL) or deep transfer learning (DTL), wherein pre-trained DL models are utilized new tasks. Transfer learning is especially advantageous in domains such as fault diagnosis, where datasets are often limited. Pre-trained CNNs, typically trained on large-scale datasets like ImageNet [77], [78] offer robust feature extraction capabilities, as their layers have already learned to capture essential structural and semantic patterns present in images. By fine-tuning these pre-trained DL models, researchers can achieve state-of-the-art performance in data-scarce applications, enhancing their diagnostic accuracy and reliability. From the wide array of CNN architectures available, each with unique advantages for feature extraction and end-to-end multiclass classification (or thermal fault diagnosis), we focus on the following selected pre-trained and customized DL models applied in this study to demonstrate their practical utility:

- 1) **ConvNeXtTiny** [79] (2022, 109.42 MB size, 28.6 Million Parameters) is a modern CNN architecture that draws inspiration from the principles of Vision Transformers (ViTs) while retaining the simplicity and efficiency characteristic of traditional convolutional designs. By incorporating innovative design elements such as *inverted bottlenecks*, *depth-wise convolutions*, and *layer normalization*, ConvNeXtTiny bridges the gap between CNN robustness and the scalability advantages of transformers. This architecture seeks to harness the strengths of both paradigms: the localized feature extraction and computational efficiency of CNNs, alongside the global attention and

adaptability typically associated with transformers. The inclusion of techniques like inverted bottlenecks enhances the efficiency of information flow, depth-wise convolutions reduce parameter overhead, and layer normalization stabilizes training dynamics. It is particularly notable for aligning convolutional architectures closer to the performance benchmarks of transformers, all while maintaining computational simplicity. This makes it a promising model for resource-constrained environments or applications requiring both robustness and scalability, offering a compelling fusion of CNN reliability and transformer-like flexibility.

- 2) **DenseNet121** [80] (2017, 33 MB size, 8.1 Million Parameters, 242 Layers Depth) represents a paradigm shift in CNN design by introducing *dense connectivity*, where each layer is directly connected to all subsequent layers in a feed-forward manner. This innovative approach effectively mitigates the *vanishing gradient problem* by ensuring stronger gradient flow during backpropagation, thereby enhancing the model's learning capability. The dense connectivity pattern promotes *feature reuse*, as each layer receives inputs from all preceding layers and passes its own feature maps to all subsequent layers. This not only reduces redundancy but also allows the network to learn richer feature hierarchies. As a result, DenseNet121 achieves remarkable *parameter efficiency*, requiring fewer parameters compared to traditional architectures while maintaining or even exceeding their performance. Such a compact design is particularly advantageous in *resource-constrained scenarios*, where computational power or memory is limited. Hence, without demanding excessive resources, it makes a compelling choice for tasks that prioritize both accuracy and efficiency.
- 3) **EfficientNetV2S** [81] (2021, 88 MB size, 21.6 M Parameters) builds on the strengths of its predecessor by introducing a more refined and efficient approach to scaling neural networks, combining *neural architecture search (NAS)* with *compound scaling*. This progressive learning strategy carefully adjusts network dimensions: depth, width, and resolution, ensuring optimal balance between accuracy and computational efficiency. The key innovation in EfficientNetV2S is its *enhanced speed and memory efficiency*, achieved through advanced architectural improvements like fused convolutional layers and optimized training methods. These enhancements not only accelerate inference but also reduce memory consumption, making it particularly well-suited for real-time applications. Its unique architectural design serves as a robust framework for deployment in *resource-limited environments* such as mobile devices and edge computing platforms,

wherein scenarios demand both high accuracy and low latency, paving the way for scalable and accessible AI solutions.

- 4) **InceptionV3** [82] (2016, 92 MB size, 23.9 M Parameters, 189 Layers Depth) advances the Inception architecture by integrating *factorized convolutions* and *aggressive dimensionality reduction*, achieving a highly optimized balance between computational efficiency and accuracy. By breaking down larger convolutions into smaller, sequential operations, such as 3×3 and 5×5 kernels, the model reduces computational complexity while maintaining its ability to capture intricate spatial patterns. One of the defining features of InceptionV3 is its use of *multi-sized convolutional filters*, enabling the network to extract features at multiple scales simultaneously. This multi-level feature extraction is particularly advantageous for complex image classification tasks, where understanding patterns across various granularities is essential for accurate predictions. The model's *modular design philosophy* promotes architectural flexibility making it both computationally economical and structurally deep. This hybrid approach not only ensures scalability for large datasets but also facilitates interpretability and adaptability across diverse applications, achieving high performance with reduced computational overhead.
- 5) **InceptionResNetV2** [83] (2017, 215 MB size, 55.9 M Parameters, 449 Layers Depth) represents a significant leap in DL architecture by merging the *multi-scale feature extraction power of inception modules* with the *stability and efficiency of residual connections*. This hybrid design enables the training of substantially deeper networks while mitigating performance degradation – a common challenge in very deep architectures. The inception modules in this architecture excel at capturing multi-level features by leveraging diverse filter sizes, ensuring that both fine-grained and coarse-grained patterns are detected. Residual connections, on the other hand, facilitate smoother gradient flow during backpropagation by bypassing unnecessary layers, effectively countering the vanishing gradient problem. This combination provides a dual advantage: *efficient feature capture across multiple scales* and *stabilized learning dynamics* that allow for unprecedented depth and complexity in network design. InceptionResNetV2 is particularly well suited for tasks requiring high accuracy and detailed feature extraction, exemplifying how hybrid models push the boundaries of both scalability and performance in DL.
- 6) **MobileNetV2** [84] (2018, 14 MB size, 3.5 M Parameters, 105 Layers Depth) revolutionizes lightweight neural networks by introducing two core innovations:

inverted residuals and *linear bottlenecks*. These mechanisms dramatically reduce the model's computational and memory requirements, making it highly efficient without compromising performance. The *inverted residual structure* connects bottleneck layers with expansion layers, allowing feature maps to pass through narrower representations and encouraging efficient feature reuse. Meanwhile, the *linear bottleneck layers* preserve crucial information by avoiding unnecessary non-linearities in low-dimensional feature spaces, thereby enhancing representational fidelity. The architecture also leverages *depth-wise separable convolutions*, a key factor in achieving its remarkable efficiency. By separating spatial and channel-wise computations, minimizes redundancy, striking an ideal balance between model compactness and accuracy. This design is suited for *mobile and embedded device applications*, enabling real-time inference with minimal hardware requirements, where it serves as a lightweight yet powerful model in resource-constrained environments.

- 7) **NASNetMobile** [85] (2018, 23 MB size, 5.3 M Parameters, 389 Layers Depth) emerges as a groundbreaking architecture derived from *Neural Architecture Search (NAS)*, exemplifying the potential of automated model design tailored for *mobile and low-power environments*. By employing *Reinforcement Learning (RL)* to systematically explore and optimize network configurations, it identifies architectures that balance precision and efficiency without human bias or manual trial-and-error. The model's design prioritizes *modularity and scalability*, allowing it to adapt seamlessly to resource-constrained scenarios while maintaining high performance. Its ability to autonomously select optimal structures ensures that computational resources are allocated judiciously, making it an ideal candidate for *personalized and domain-specific applications* where efficiency is paramount. The model symbolizes a pivotal shift in DL, where the *automation of CNN design* reduces development time and democratizes access to cutting edge architectures. By combining automation with precision, it sets the stage for future advancements in intelligent and adaptive network customization for diverse use cases.
- 8) **ResNet50V2** [86] (2016, 98 MB size, 25.6 M Parameters, 103 Layers Depth) enhances the foundational ResNet architecture by introducing critical refinements in the *residual block design*, such as reordering batch normalization (BN) and activation layers. This adjustment improves both *training stability* and *convergence speed*, making the model more robust in learning complex features. The core innovation of *deep residual learning* remains central to ResNet50V2, enabling efficient gradient flow through the use of *identity mappings*. By bypassing layers with residual connections, the architecture

mitigates the *performance degradation* typically encountered in very deep networks, ensuring scalability without sacrificing accuracy. Its refined design demonstrates a balance between depth and efficiency, making it well-suited for diverse applications, from computer vision tasks to transfer learning (TL) scenarios. Its emphasis on optimizing gradient flow and training dynamics ensures that it continues to serve as a benchmark for reliable and scalable deep network architectures.

- 9) **VGG16** [87] (2015, 528 MB size, 138.4 M Parameters, 16 Layers Depth) stands as a milestone in deep learning architectures, celebrated for its *simplicity and depth* achieved through the consistent use of 3×3 *convolutional layers* stacked sequentially. This uniformity in design streamlines feature extraction, making it both intuitive and effective for a wide range of image classification tasks. While its architecture involves a relatively *large parameter size*, VGG16 demonstrates the power of deep, uniform convolutional layers to capture hierarchical representations without the need for complex modular components. Its straightforward design has not only set a foundation for subsequent advancements but also inspired investigations into *parameter reduction techniques* to address computational efficiency challenges. Its legacy lies in its *methodical simplicity*, offering a clear framework that balances depth and design clarity while serving as a stepping-stone for exploring optimization strategies in well-structured and modern convolutional networks.
- 10) **Xception** [88] (2017, 88 MB size, 22.9 M Parameters, 81 Layers Depth) or Extreme Inception elevates the concept of *depth-wise separable convolutions* to its logical extreme, crafting a fully convolutional architecture devoid of any intermediate fully connected layers. This streamlined structure emphasizes *efficiency* by minimizing computational redundancy, while simultaneously preserving a *high degree of representational power* for complex feature extraction. The model's simplicity lies in its modular and systematic design, which leverages *separable convolutions* to independently learn spatial and channel-wise features. This approach not only optimizes computational efficiency but also significantly enhances the network's capacity to model high-dimensional feature spaces. By integrating efficiency and depth, this architecture is especially adept at balancing the computational load with the need for fine-grained feature learning, making it a prime candidate for tasks requiring extreme optimization in resource-constrained environments.
- 11) **Custom CNN** [89] (2024, 1.80 MB size, 0.47 M Parameters, 5 Layers Depth) and **Ensembled CNN** [89] (2024, 7.19 MB size, 1.88 M Parameters, 5 Layers Depth) are

designed to strike a balance between efficiency and performance, leveraging a modular design composed of *three sequential blocks of layers*: Conv2D, Batch Normalization, and MaxPooling2D. The innovation lies in *progressively increasing kernel sizes* 3×3 , 5×5 , and 7×7 , allowing the model to capture multi-scale features effectively, accommodating intricate patterns within the input data. The architecture concludes with *two dense layers* for robust decision-making and a *Global-Max-Pooling2D layer*, which reduces the spatial dimensions while retaining the most significant feature activations. This design ensures a lightweight and computationally efficient model ideal for scenarios with limited computational resources.

2.2.1.2 Performance Evaluation Metrics

The performance evaluation metrics provide a comprehensive quantitative evaluation of the DL model predictions to perform multiclass classification, including accuracy, precision, recall, F1 score, and training time (in seconds). While the *Confusion Matrix* is used for the tabulated classification between Ground Truths (True Label) and Predictions (Predicted label). These metrics are important in determining how well a model performs in classification tasks, where the balance among true positives, true negatives, false positives, and false negatives is crucial.

Learning curves served as a visual narrative to track and assess the learning progress and performance of DL models over successive epochs or training iterations. These curves provide a behavioral illustration, highlighting how the model evolves in its capacity to minimize error and improve predictive accuracy. A well-fit model is depicted when training and validation curves exhibit consistent improvement, reflecting effective learning from the training data and the ability to generalize to unseen validation data. Conversely, divergent trends – such as overfitting (training improvement without validation improvement) or underfitting (poor performance on both sets) – indicate potential issues with the model or the training process. By offering real-time insights into training dynamics, learning curves become an indispensable tool for diagnosing model behavior, fine-tuning hyperparameters, and ensuring the model achieves an optimal balance among fit, complexity, and generalizability.

2.2.2 Explainable Artificial Intelligence Methods

The growing focus on explainability and interpretability in DL models has catalyzed the development of explainable XAI techniques [90]. These techniques are pivotal for

unraveling the decision-making processes of complex, often opaque, DL systems, thereby bridging the divide between their technical intricacies and practical transparency [91]. Broadly, XAI methods can be classified into two complementary categories [92], [93]: perceptive explainability and mathematical interpretability. *Perceptive explainability* provides intuitive, visual insights into model behavior, such as saliency maps, feature attributions, or attention heat maps. These visualizations highlight critical features or regions within the input data that influence the model's predictions, aligning closely with human cognition. This intuitive connection enables immediate comprehension of the model's decision-making process, fostering confidence in its outcomes. *Mathematical interpretability*, on the other hand, examines the logical and mathematical constructs underlying the model's representation of input data. This approach investigates how features are clustered, weighted, or transformed, offering a more structured understanding of the mechanisms governing the model's internal processes. By exposing the mathematical underpinnings, it ensures greater accountability and insights into the reproducibility of model behavior. The synergistic interplay between these approaches provides a holistic framework for understanding DL models, addressing both practical transparency and technical depth. By fostering trust and accountability, XAI not only enhances the acceptance and reliability of AI systems but also lays the groundwork for further refinement, innovation, and ethical deployment across diverse domains.

2.2.2.1 Perceptive Explainability

The primary objective of XAI methods within the realm of perceptive explainability is to provide clear, intuitive visual representations of the most influential features driving a model's predictions. These methods emphasize identifying and analyzing the relevance of specific regions or elements within the input data that significantly impact the model's decision-making process. By visually mapping the contribution of specific features, perceptive explainability enables a deeper understanding of feature-level classification behavior. This approach is particularly useful in analyzing the relevance of specific regions or elements in the input data, shedding light on the model's reasoning. Through such visual insights, perceptive explainability fosters interpretability, trustworthiness, and transparency, allowing users to better comprehend how and why a model arrives at its predictions. This, in turn, enhances confidence in AI-driven systems and supports their refinement and responsible deployment across various applications. The perceptive explainability methods applied here are:

- 1) The **Activation Maximization** [94], [95] method proposed by Erhan et al. in 2009, aims to optimize and maximize the activation of a specific neuron in a neural network by framing it as an optimization problem. Here, input data point x (or radiometric thermal matrix in this work), neural network parameter θ (including both weights and biases), and neuron activation $h_{ij}(\theta, x)$ (the activation of neuron i in layer j). The goal is to find the input x^* that maximizes the activation, defined as:

$$x^* = \operatorname{argmax}_x h_{ij}(\theta, x) \quad (2.1)$$

For generality, $h(\bullet)$ can be used to denote the activation instead of $h_{ij}(\bullet)$, when specific indexes i, j are not relevant.

To find x^* , gradient ascent is used with an update rule:

$$x_{t+1} = x_t + \varepsilon_1 \nabla_x h(\theta, x_t) \quad (2.2)$$

where: x_t (input at iteration t), ε_1 (step size or learning rate), $\nabla_x h(\theta, x_t)$ (gradient of the activation with respect to the input). For initialization, the choice of initial input x_0 is critical to avoid convergence to local optima that may not yield meaningful representations. There are two options for x_0 : First, the random noise (often used to explore the input space broadly), and second, the real data point (in this work – radiometric thermal matrix serves as initialization, simplifying the optimization process and helping to avoid uninformative local optima). By starting with a radiometric thermal matrix as x_0 , the gradient ascent process is guided toward meaningful representations aligned with the input space. This ensures that the resulting maximized activations are interpretable and directly tied to the context of the data such as radiometric thermal matrix features.

- 2) **SmoothGrad method:** Simonyan et al. [96] introduced *image-specific class saliency visualization*, a technique designed to identify the most influential regions in an image that drive a deep CNN's predictions. This method provides insight into the internal decision-making of DL models by highlighting image regions critical for classification. In *linear models*, important features are typically identified using large weight values, directly indicating feature significance. In *highly nonlinear models* like deep CNNs, weight magnitudes alone are insufficient due to complex interactions between layers. To approximate a CNN's behavior, the first-order Taylor expansion defines the saliency

map. This *gradient-based saliency map* highlights the most influential pixels in the image. However, early implementations of this method often suffered from high noise, making the visualizations difficult to interpret. To enhance interpretability, Smilkov et al. [97] proposed *SmoothGrad*, which reduces noise in saliency maps by averaging multiple perturbed versions of the input image. The advantages of SmoothGrad are: a) reduces noise in saliency maps, b) enhances visual coherence, making the maps easier to interpret, and c) preserves important features by averaging multiple perspectives. By leveraging SmoothGrad, CNN explanations become more stable and visually meaningful, helping users better understand model behavior in tasks such as image classification and fault detection. Here, the 25 smooth samples are applied in this work for computational purposes. Further, taking a hybrid approach, the addition of the Guided Backpropagation (GBP) [98] method to retain high-resolution details with SmoothGrad (preserving class-specific focus), complements the existing *gradient-based saliency map* for enhanced and sharper visualizations, and this merger is known as *Guided SmoothGrad*.

- 3) **Guided Backpropagation (GBP) method** [98]: GBP is the *gradient-based saliency visualization technique* introduced by Springenberg et al., designed to improve the interpretability of deep CNN predictions by selectively modifying the gradient flow during backpropagation. The key idea is that the standard backpropagation computes the gradient of the class score with respect to the input image, this can include both positive and negative gradients. GBP modifies the gradient flow by blocking negative gradients, allowing only positive signals to pass, which enhances the clarity of feature visualization. During standard backpropagation, gradients can be positive (exciting neurons) or negative (suppressing neurons). GBP blocks negative gradients during the backward pass, preventing the influence of features that reduce the activation of the target class. The result is a more focused, interpretable saliency map highlighting features that positively impact the classification decision. GBP preserves fine textures and spatial details making it useful for tasks like object localization and fine-grained feature attribution. However, GBP does not preserve class-discrimination information, meaning it highlights all edges and textures in an image without focusing on specific class features. To address this, GBP is often combined with other XAI methods like SmoothGrad [97], which preserves class-specific information (provides class-discriminative heatmaps), while GBP highlights and retains fine-grained details, and this combination is called *Guided SmoothGrad*.

- 4) **Grad-CAM** (Gradient-weighted Class Activation Mapping) [99]: Grad-CAM generates visual explanations for deep CNN decisions by leveraging gradient information flowing through a specific convolutional layer. It helps interpret which parts of an input image influence the model's prediction by overlaying a heatmap that highlights the most relevant regions. Grad-CAM focuses on a particular convolutional layer (typically the last one before classification, but it's not a necessary option). Its advantage is class-specific interpretability, which highlights only the relevant features that contribute to the decision for a specific class. However, it has limitations, like low-resolution details cannot be captured, as compared to pixel-wise saliency methods like Guided Backpropagation (GBP), and focuses only on high-level features, ignoring fine-grained textures.

2.2.2.2 Mathematical Interpretability

Mathematical interpretability methods seek to explain how DL models process information by analyzing their internal mathematical structures. These methods reveal how NNs cluster, organize, and utilize features for decision-making, helping researchers evaluate the relevance and meaningfulness of learned representations. For feature space visualization, it provides an understanding how models encode and separate data in high-dimensional space. The mathematical interpretability methods employed in this research work are t-distributed Stochastic Neighbor Embedding (t-SNE) [100] and Uniform Manifold Approximation and Projection (UMAP) [101].

- 1) **t-SNE** (t-distributed Stochastic Neighbor Embedding) [100]: The t-SNE is a nonlinear dimensionality reduction technique designed for visualizing high-dimensional data in two or three dimensions. It improves upon traditional Stochastic Neighbor Embedding (SNE) by addressing key optimization challenges and enhancing interpretability. The advantages are: 1) traditional SNE uses a Gaussian to model similarities in lower dimensions, which can lead to poor separation. The t-SNE replaces this with a heavy-tailed Student's t-distribution, reducing distortions in global structure. 2) it solves the *crowding problem*, high-dimensional data often clusters tightly, making visualization difficult. The t-SNE spreads points out more evenly in the low-dimensional space, preventing overcrowding. 3) it optimizes *cost function* – unlike SNE, which minimizes a Kullback-Leibler divergence asymmetrically, t-SNE symmetrizes this function, leading to more stable training and better visualization.

2) **UMAP** (Uniform Manifold Approximation and Projection) [101]: is a nonlinear dimensionality reduction technique designed for high-dimensional data visualization and feature clustering. It is particularly effective for understanding complex structures and datasets and is often used due to its faster computation and better preservation of global structure. The core assumptions of UMAP: 1) manifold hypothesis – data points lie on a low-dimensional manifold within a high-dimensional space (i.e. uniform distribution of data points). 2) topological structure preservation – the algorithm maintains the relationships between points in the original space. 3) local connectivity – nearby points remain connected when mapped to a lower-dimensional space. UMAP works in two steps: Step-1, learning the high-dimensional manifold structure – constructs a fuzzy topological representation of the data using k-NN and fuzzy simplicial sets, Step-2, mapping the data to a lower-dimensional space – optimizes the layout to preserve local and global structure while ensuring separation between clusters.

2.3 SPV Power Systems and Intelligent Aerial-IRT Inspection

The sustainable and uninterrupted operation of large-scale and utility-scale SPV electric power systems is paramount in achieving targeted clean energy generation goals. However, the performance and longevity of these systems are inherently susceptible to degradation mechanisms, particularly thermal stresses which can significantly impair energy output efficiency [102]. To address this challenge, advanced diagnostic monitoring – specifically, intelligent aerial IRT inspection has emerged as a transformative approach with profound techno-economic implications.

The effectiveness of ANN-based diagnostic monitoring (thermographic diagnosis) in SPV power systems is fundamentally contingent upon the availability, diversity, and quality of datasets. High-quality (and high-fidelity) thermographic data acquisition is essential for training robust predictive models capable of accurately identifying anomalies, degradation patterns, thermal signatures, and potential failures in SPV arrays. Given the complexity and variability of real-world operational conditions, ensuring comprehensive data coverage across different environmental settings, seasonal variations, and panel configurations is critical for improving generalization and reliability in ANN-driven diagnostics.

2.3.1 SPV Power System Performance Optimization and the Role of Infrared Thermography

Periodic inspections and proactive operational maintenance are essential for early defect

detection in SPV modules. Infrared thermographic analysis, as a non-destructive testing (NDT) methodology, enables the identification of thermal anomalies and stress-induced degradation patterns that may not be detectable through conventional monitoring techniques [103]. According to EN 13306:2017, the degradation state of SPV panels due to thermal stress compromises the system's ability to meet projected energy output levels [1]. Without timely intervention, this thermal-based deterioration can escalate, leading to energy losses, increased operational costs, and reduced system reliability. To mitigate these risks, PdM strategies have gained prominence, allowing for data-driven decision-making that optimizes both technical performance and financial viability [2]. Unlike reactive or scheduled maintenance, PdM leveraged real-time and historical data analytics to forecast potential failures, thereby minimizing downtime and enhancing overall system resilience [104]. This paradigm shift aligns with the broader objective of intelligent asset management (or intelligent asset health management), wherein operational efficiency is maximized through advanced sensor-driven diagnostics and computational modeling.

2.3.2 Standardized Guidelines and the Need for Skilled Expertise

The implementation of aerial IRT inspections is governed by stringent technical standards to ensure accuracy, consistency, and regulatory compliance. The International Electrotechnical Commission (IEC TS 62446-3:2017) provides comprehensive technical specifications for outdoor infrared thermographic analysis, outlining protocols for evaluation, supervision, monitoring, diagnostics, inspection, and maintenance in SPV installations [34]. Adherence to these guidelines ensures the reliability of thermographic data, which is crucial for identifying defective modules, mitigating power losses, and informing maintenance interventions. However, the efficacy of IRT-based diagnostics is contingent upon the expertise of trained personnel. Proper certification and technical proficiency in thermographic interpretation are essential for accurate anomaly detection and meaningful data extraction [105], [106], [107]. Misinterpretation of thermal patterns or incorrect calibration of infrared sensors can lead to false positives (incorrect defect detection) or false negatives (missed faults), potentially undermining the effectiveness of the PdM framework. Thus, continued investment in professional training programs and certification initiatives is indispensable for maximizing the potential of IRT technologies in SPV system diagnostics.

2.3.3 Technical Prerequisites for Effective Outdoor-IRT Inspection

IRT inspections of SPV modules require strict adherence to technical guidelines, precise calibration of equipment, and optimal environmental conditions to ensure accurate and reproducible thermal imaging [108]. Several key parameters must be carefully controlled to maximize diagnostic precision and minimize potential errors:

- 1) **Solar Irradiance Requirements:** A minimum solar irradiance of 500W/m^2 is necessary to generate sufficient thermal contrast on SPV panels. Lower irradiance levels may obscure thermal anomalies, reducing the efficacy of defect detection.
- 2) **Infrared Detector Specifications:** The use of uncooled microbolometer detectors is recommended, particularly those with high sensitivity in the $8\text{-}14\ \mu\text{m}$ long-wavelength infrared (LWIR) band. This spectral range is optimal for capturing radiometric emissions from SPV glass surfaces while mitigating interference from atmospheric absorption.
- 3) **Thermal sensitivity (NETD – noise equivalent temperature difference):** The thermal imaging system must exhibit a thermal sensitivity and (NETD) of $\leq 0.08\ \text{K}$, enabling it to discern minute temperature variations on the glass surface of SPV modules. This level of sensitivity is essential for detecting early-stage defects, such as microcracks, hotspots, delamination, and electrical faults, before they escalate into significant performance losses.
- 4) **Camera Viewing Angle Optimization:** To reduce SPV module glass reflections and optimize relative emissivity (high), the camera viewing angle should be maintained between 5° to 60° (where 0° represents the perpendicular axis to the SPV panel surface). This angular range serves as a critical balance between minimizing reflection artifacts and capturing precise temperature differentials across the SPV module.

Chapter III

Theoretical Framework

3.1 Convolutional Neural Network

CNNs resemble the way that human beings do visual processing. Feature extraction, one of the most important aspects of the network, can be done effectively in CNN. Furthermore, the max pooling layer (see Section 3.1.3 Pooling Layer) of CNN is powerful in reducing the dimension of the image without affecting the feature representations [109].

Figure 3.1 shows the architecture of AlexNet, which demonstrates a general structure of CNN. Each of the small sections, which contains certain tasks, is called a layer. Firstly, there are three main types of layers in CNN, which are convolutional layer, max pooling layer and the fully connected layer. Secondly, layers inside the network are connected with each other and use the output of the previous as its input so as neuron behavior. Furthermore, the main task for the first two layers, convolutional layer and max pooling layer, is to achieve feature extraction, while the last layer is considered as classification which normally has one or more fully connected layer [110], [111]. In particular, the last fully connected layer maps the features into an actual object with probability output. Lastly, similar to AlexNet, a classic CNN consists of blocks of layers, where each block contains several sub-layers such as convolutional layer, rectified linear unit (ReLU) layer or max pooling layer.

3.1.1 Input Layer

In a general CNN, like AlexNet, the input to the first layer is not the image itself, but a so-called channel [112]. Channel refers to a certain component of an image. For example, an RGB image has three red, green, and blue channels. There are $F_0 = 3$ channels to the first layer, where each channel corresponds to each of the RGB colors (Figure 3.2). Each of the RGB channels, together, forms the first input layer (1-channel grayscale image is duplicated thrice).

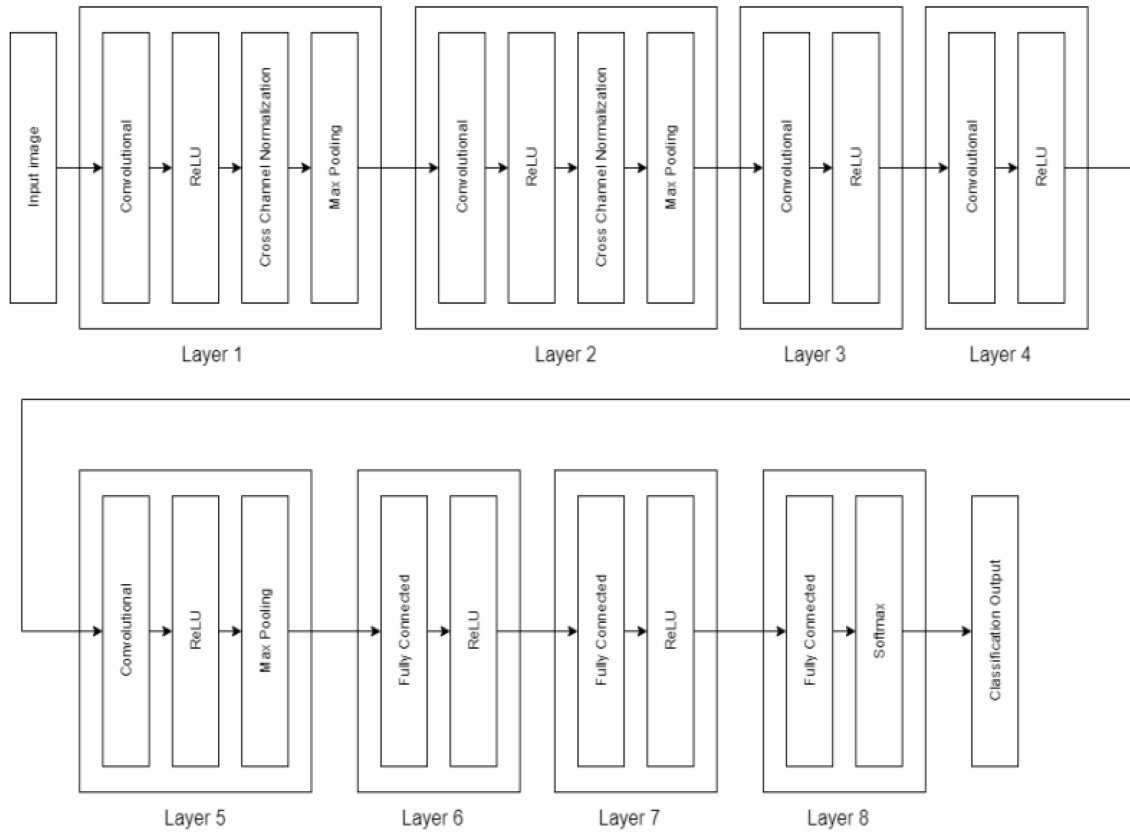


Figure 3.1 Architecture of AlexNet with 8 layers, where the first five blocks of layers are convolutional layers, and the rest are fully connected layers.

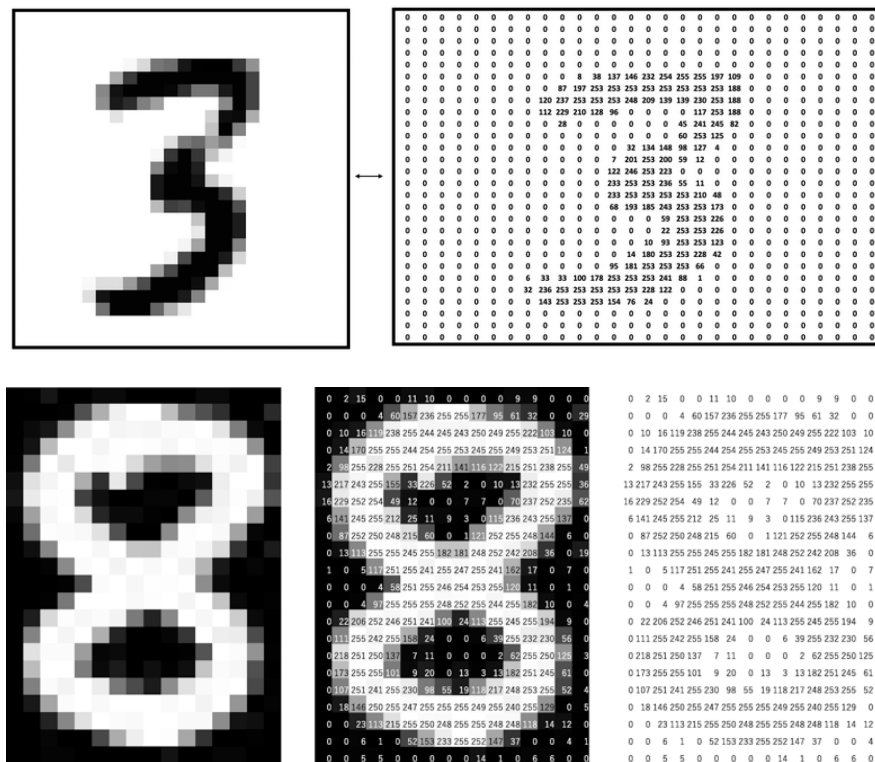


Figure 3.2 Sample input layer: A digit 3 and 8 image with size 28×28 on the left and the corresponding pixel value on the image.

3.1.2 Convolutional Layer

Convolutional layer, the most characteristic layer in CNN, provides the essential functionality to feature extraction. Like its name the operation convolves (in traditional signal processing this operation would technically be a cross-correlation) a feature of an input layer with two learnable parameters, weight and bias, to generate the output feature map [112]. The operation is done by a mathematical linear operation, called convolution product. Let us recall the convolution of two integrable functions f and g denoted as $f * g$, is defined as the following:

$$(f * g)(t) \triangleq \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \quad (3.1)$$

where t is an input variable and $t \in \mathbb{R}$.

In order to obtain the most appropriate output *feature map*, the value of the weight and bias are what need to be discovered in the later optimization process. The feature map refers to the output feature for one given *filter*.

First, all the variables are under real number field (\mathbb{R}), and we define some of the important terminologies below. *Filter* (or kernel) is a three-dimensional matrix of size $K \times K \times F$ where $K \times K$ refers to the single aspect of a channel and F refers to the number of channels. The channel is also known as the *depth of the image* since the number of channels in a filter is equal to the number of channels for the input image. The depth of the filters is the number of filters, which also refers to the number of output feature maps from the convolutional operation. For example, in the first convolutional layer of the AlexNet, given the $227 \times 227 \times 3$ input image, the filter is of size $11 \times 11 \times 3$ with the depth of the filters equal to 96. Note that the number of channels in the filter is equal to the number of channels for the input image. Then, the convolutional layer produces 96 output feature maps ($55 \times 55 \times 96$). The output dimension is computed via Equation (3.6) below.

It is worth mentioning that the convolutional operation is a combination of four summations. From inner to outer, the first two sums over one channel of a *filter* ($K \times K$), the third one sums over the *channel* (F), and the last one sums over the *depth of the filters* (D). Regarding how the operation applies to an image, we first consider $F = 1$, $D = 1$, and start from an image of size 5×5 and a filter (or kernel) of size 3×3 (Figure 3.3). The convolution is computed by sliding the filter over the image, starting from the top left corner. Then, the first value of the output matrix is obtained by summing the product of the corresponding

pixel value and the filter value. Similarly, the final output is computed by moving the filter from left to right and from top to bottom.

I ₁₁	I ₁₂	I ₁₃	I ₁₄	I ₁₅
I ₂₁	I ₂₂	I ₂₃	I ₂₄	I ₂₅
I ₃₁	I ₃₂	I ₃₃	I ₃₄	I ₃₅
I ₄₁	I ₄₂	I ₄₃	I ₄₄	I ₄₅
I ₅₁	I ₅₂	I ₅₃	I ₅₄	I ₅₅

K ₁₁	K ₁₂	K ₁₃
K ₂₁	K ₂₂	K ₂₃
K ₃₁	K ₃₂	K ₃₃

Figure 3.3 Demonstration of input map and a filter: A 5×5 images with a single channel on the left and a 3×3 filter on the right. Note that each value from the image or filter is a pixel value.

In other words, we can define the operation mathematically for one single channel as follow [113]:

$$O(i, j) = \sum_{k=1}^K \sum_{l=1}^K I(i+k-1, j+l-1) K(k, l) \quad (3.2)$$

where $K \times K$ is the size of the filter, O denotes output, and $O(i, j)$ is the value of the i -th row and j -th column of the output matrix. Furthermore, i is running from 1 to $W - K - 1$ and j is running from 1 to $H - K + 1$ where $W \times H$ is the size of the image [113]. For example, the $O(1, 1)$ value from the above example is

$$\begin{aligned} O(2, 3) &= \sum_{k=1}^3 \sum_{l=1}^3 I(2+k-1, 3+l-1) K(k, l) \\ &= \sum_{k=1}^3 \sum_{l=1}^3 I(1+k, 2+l) K(k, l) \\ &= I_{23}K_{11} + I_{24}K_{12} + I_{25}K_{13} + \\ &\quad I_{33}K_{21} + I_{34}K_{22} + I_{35}K_{23} + \\ &\quad I_{43}K_{31} + I_{44}K_{32} + I_{45}K_{33} + \end{aligned} \quad (3.3)$$

where we have set $I_{ij} = I(i, j)$ for $1 \leq i, j \leq 5$ and $K_{kl} = K(k, l)$ for $1 \leq k, l \leq 3$.

As a result, a 4×4 matrix is produced from the above example. Now, if we consider a normal image with RGB colors, which means $F = 3$, $D = 1$, the final matrix is formed by summing all the corresponding value from each of the channel output matrices. The result can be defined as

$$O'(i, j) = \sum_{v=1}^F O_v(i, j) \quad (3.4)$$

where O_v denotes the output matrix from the v -th channels, F is the number of channels and O' is the output feature map.

3.1.2.1 Padding and Stride

As we can notice, the outer pixel value of an image is only involved once in one of the operations, whereas the inner pixel value of an image is involved multiple times. Moreover, compared to the original image, the dimension of the output feature map has been reduced. For example, in the above example, the original image has size 5×5 , while the output has size only 4×4 . Padding (P) is added around the images with value 0 to preserve the width and height of the original image [112] (Figure 3.4).

If the original input has size $W \times H$, then, after adding the padding, the output has a size

$$(W + 2P) \times (H + 2P) \quad (3.5)$$

where each P corresponds to the padding size in each direction of the input image.

0	0	0	0	0	0	0
0	I_{11}	I_{12}	I_{13}	I_{14}	I_{15}	0
0	I_{21}	I_{22}	I_{23}	I_{24}	I_{25}	0
0	I_{31}	I_{32}	I_{33}	I_{34}	I_{35}	0
0	I_{41}	I_{42}	I_{43}	I_{44}	I_{45}	0
0	I_{51}	I_{52}	I_{53}	I_{54}	I_{55}	0
0	0	0	0	0	0	0

Figure 3.4 Demonstration of padding: A matrix (or image) with zero-padding of size 1 ($P = 1$) (shaded).

Another important term to be introduced here is called stride (S), which is the distance between the movement of the filter (Figure 3.5 and Figure 3.6). In general, the stride value for the convolutional layer is 1 and it is widely used in the max pooling layer (see Section 2.1.3 Pooling Layer) [110].

In conclusion, if the size of the filter is $K \times K$, then the output width (W_0) and height (H_0) are

$$W_0 = \frac{W + 2P - K}{S} + 1 \quad \text{and} \quad H_0 = \frac{H + 2P - K}{S} + 1 \quad (3.6)$$

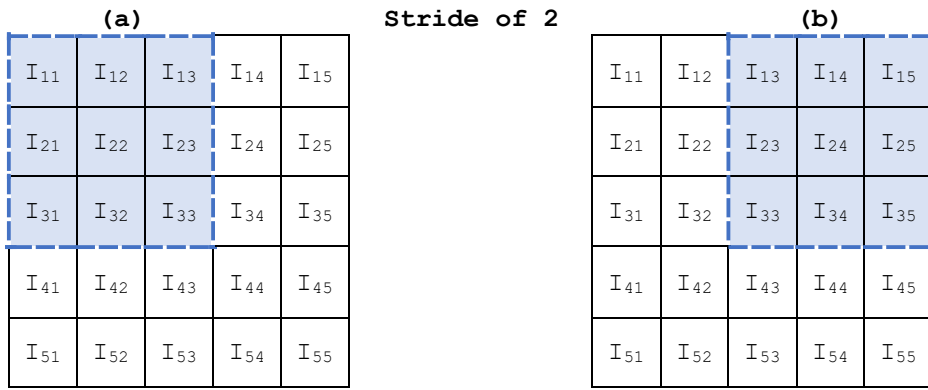


Figure 3.5 Demonstration of stride: The graph represents a movement of a filter with stride ($S = 2$) from the first position (a) to the next position (b).

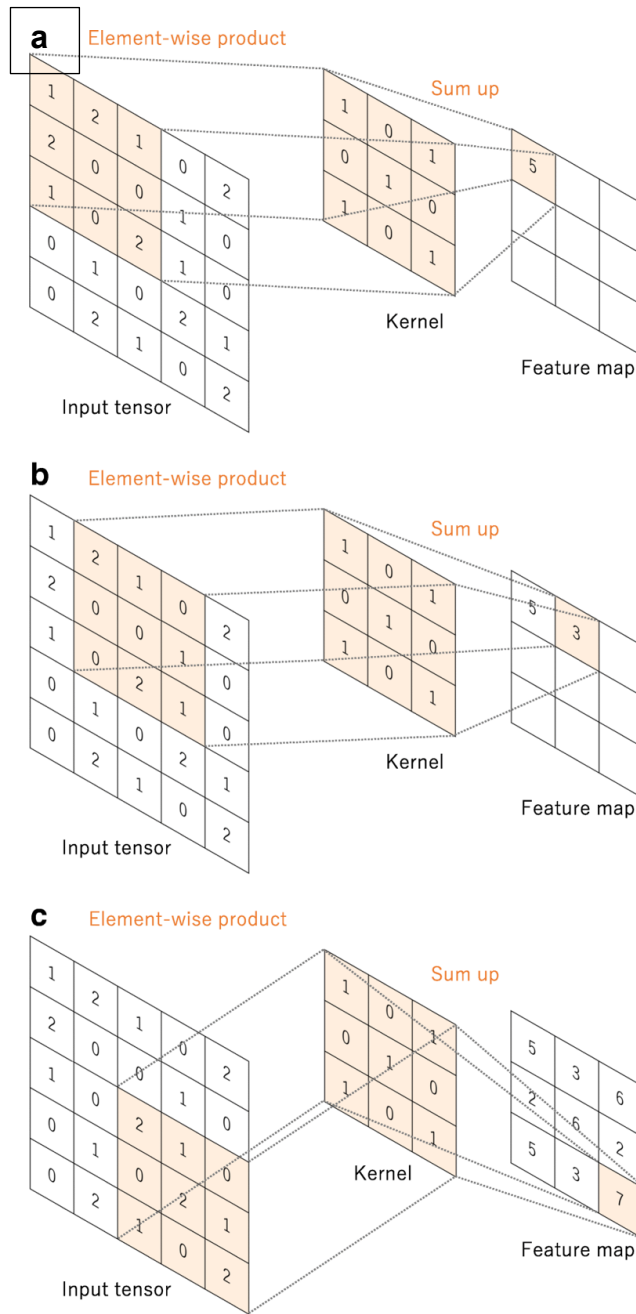


Figure 3.6 Feature map extraction when kernel of 3×3 convolve an input tensor of 5×5 with stride 1.

Also, in order to preserve the width and the height of the original image, common padding is

$$P = \frac{K-1}{2} \quad (3.7)$$

where K is the dimension of the filter.

Figure 3.7 shows a general convolution operation applied to a real image with size 32×32 , a filter of size 3×3 , no padding and 1 be the stride value. In realistic, there is more than one filter applied to the input to extract more features. For example, in AlexNet, the first convolution layer applies 96 filters with size $11 \times 11 \times 3$ and a stride of 4 to the input image with size $227 \times 227 \times 3$, where the 3 represents the RGB colors [114].

3.1.2.2 Parameters

In general, the convolutional operation is a matrix multiplication which can be expressed as a single neuron function without an activation function

$$f(\mathbf{x}) = W' \times \mathbf{x} + \mathbf{b} \quad (3.8)$$

where $\mathbf{x} \in \mathbb{R}^n$, $f(\mathbf{x}) \in \mathbb{R}^m$, $W' \in \mathbb{R}^{m \times n}$ is a weight matrix and $\mathbf{b} \in \mathbb{R}^m$ is a bias. The high-dimensional filter is reshaped into a 2D weight matrix for the image input. A simple example is used here to demonstrate the intuition behind it. Consider an image with the following settings, $W = 4$, $H = 4$, $K = 2$, $F = 3$, $D = 2$, and $S = 1$. In other words, the input RGB image is of size $4 \times 4 \times 3$ ($W \times H \times F$), the filter is of size $2 \times 2 \times 3$ ($K \times K \times F$) with the depth of the filters (D) equal to 2, and the stride (S) is 1 (Figure 3.8)

The conversion result of the filter and the input image is provided in Figure 3.7. The first column of the reshaped x is formed by merging the filter area from each channel. In other words, the first filtered areas for the input image are

$$\begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 17 & 18 \\ 21 & 22 \end{bmatrix}, \begin{bmatrix} 33 & 34 \\ 37 & 38 \end{bmatrix} \quad (3.9)$$

from channel 1, channel 2, and channel 3, respectively. Then we flattened the matrix by row into the first column of reshaped x .

$$[1 \ 2 \ 5 \ 6]^T, [17 \ 18 \ 21 \ 22]^T, [33 \ 34 \ 37 \ 38]^T \quad (3.10)$$

$$[1 \ 2 \ 5 \ 6 \ 17 \ 18 \ 21 \ 22 \ 33 \ 34 \ 37 \ 38]^T \quad (3.11)$$

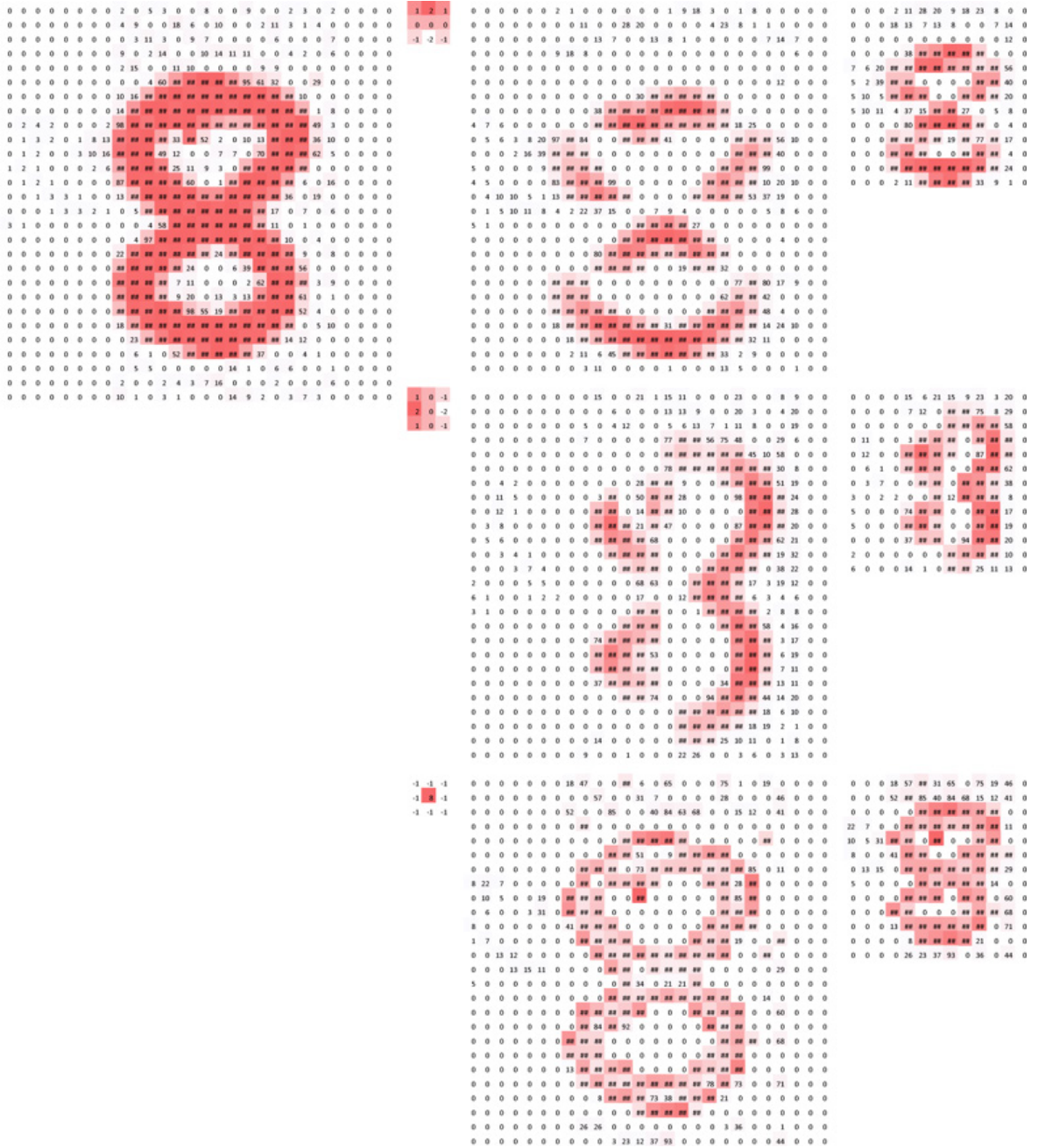


Figure 3.7 Examples of how kernels in convolution layers extract features from an input tensor are shown. Multiple kernels work as different feature extractors, such as a horizontal edge detector (top), a vertical edge detector (middle), and an outline detector (bottom). Note that the left image is an input, those in the middle are kernels, and those in the right are output feature maps.

Later, we combine every separated vector (Equation (3.10)) into one vector (Equation (3.11)) to be the first column of the output matrix. The same method applies to the rest of the columns. Note that since the stride is 1 ($S = 1$), we have 9 columns in total for the reshaped x by taking the filter from left to right and top to bottom. Similarly, the new weight matrix is obtained from the filter where the number of rows is the depth of the filters ($D = 2$). Each row is formed by concentrating the flattening form of the individual channel matrix.

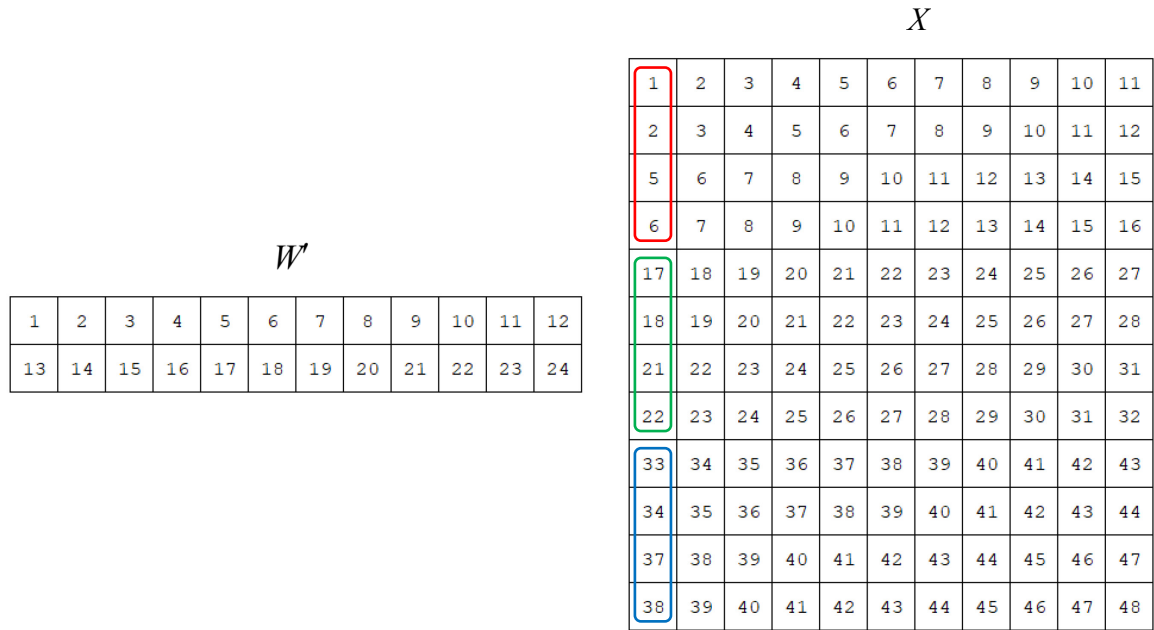


Figure 3.8 Result of reshaping: The red area refers to the first filter area in Channel 1 of the input image, the green area refers to the first filter area in Channel 2 of the input image. Similarly, the blue area comes from the first filter area of Channel 3 of the image

The total number of values from all the learnable parameters is a critical measure of the complexity of a CNN [111]. The number of values in this learnable parameter, W' , is

$$\text{Number of values} = (K \times K \times F + 1) \times D \quad (3.12)$$

where K is the size of the individual channel, F is the number of channels, D is the depth of the filters and 1 refers to a single bias for the current filter [115]. For example, the first convolution layer in AlexNet contains

$$\text{Number of values} = (11 \times 11 \times 3 + 1) \times 96 = 34,944 \quad (3.13)$$

Moreover, AlexNet architecture has around 60 million values to be determined at the end of the networks, which is trained in two Graphical Processing Units (GPUs) [114].

3.1.3 Pooling Layer

The pooling operation, or sub-sampling layer, integrates a down sampled operation to the input feature map, which is also called dimension reduction operation [110]. It is worth pointing out that the pooling operation does not require to train more parameters. In general, there are two kinds of pooling operations, one is called average pooling, and the other one is called max pooling. Similar to convolutional layer, the pooling layer also involves hyperparameters including filter size (R), stride (S) and padding (P) [110].

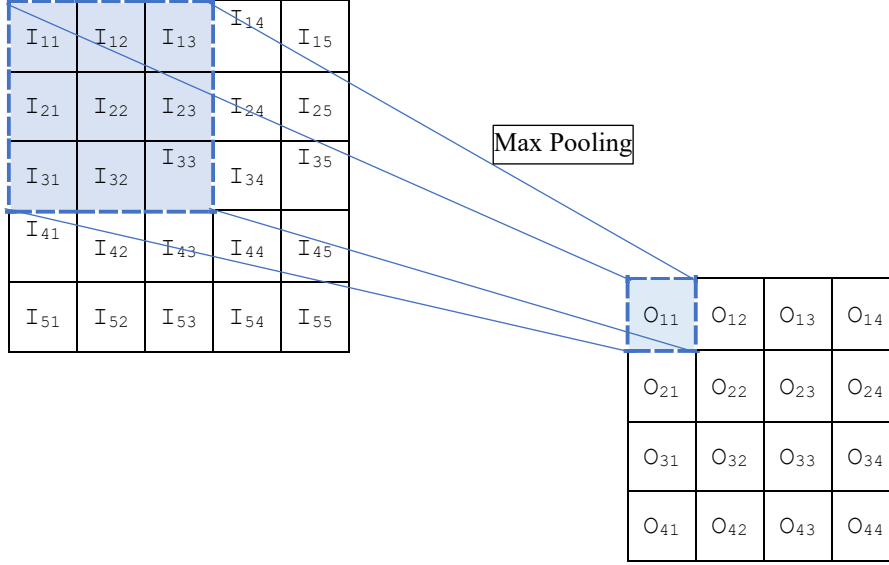


Figure 3.9 Demonstration of max pooling operation: A sample max pooling operation of an input feature map of size 5×5 with filter size ($R = 2$), stride ($S = 1$) and no padding resulted in an output of size 4×4 .

Max pooling, one of the most populate types of pooling operation, is essentially a max function, which filters the input map and takes the maximum value from the filtered region (Figure 3.9). The operation is defined by

$$O(i, j) = \max_{i \leq l \leq i+R-1, j \leq k \leq j+R-1} \{I(l, k)\} \quad (3.14)$$

where $O(i, j)$ is the value of the i -th row and j -th column of the output matrix. As an example (Figure 3.9), with a filter size equal to 2 ($R = 2$) and stride equal to 1 ($S = 1$), an output feature matrix of size 4×4 is computed by taking the maximum value inside the filtered space. The first value is computed by

$$O_{11} = \max \{I_{11}, I_{12}, I_{21}, I_{22}\} \quad (3.15)$$

where we have set $O_{ij} = O(i, j)$ and $I_{lk} = I(l, k)$. Compared to the input map, the dimension of the output map is reduced by 1.

3.1.4 Activation Function (ReLU)

Rectified linear-activation unit (ReLU), a nonlinear activation function, acts mathematically as a biological neuron [110]. Although the traditional activation functions, such as sigmoid (logistic) or hyperbolic tangent (tanh) [116], when perform the same task, it has been observed that a CNN with ReLU as an activation function trains much faster than a network with hyperbolic tangent [114].

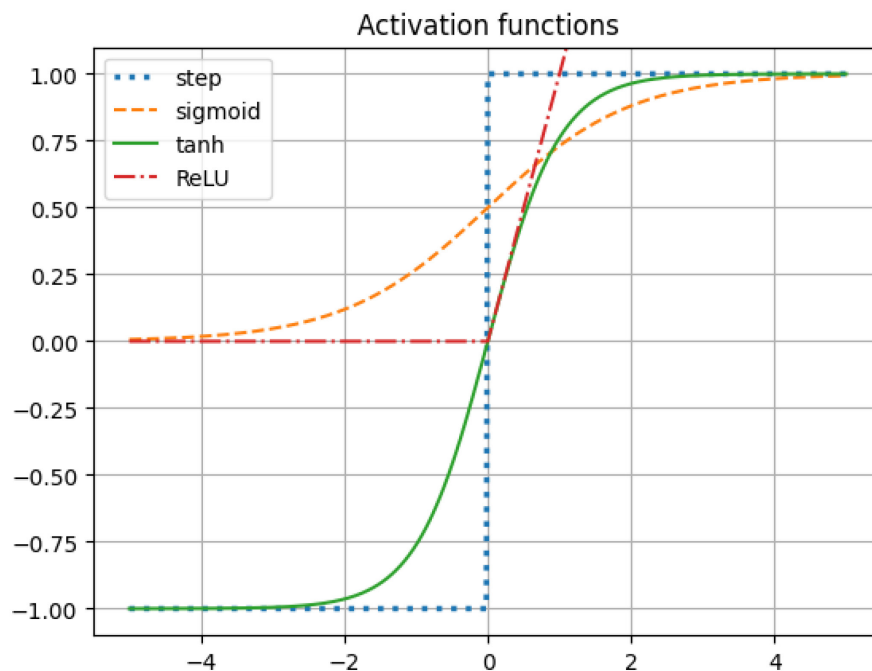


Figure 3.10 Three most commonly used activation functions in the hidden layers: ReLU, tanh, and sigmoid

Furthermore, ReLU is a very simple function, which is defined by

$$f(x) = \max(0, x) \tag{3.16}$$

where the function is linear for $x > 0$ and 0 when x is negative (Figure 3.11).

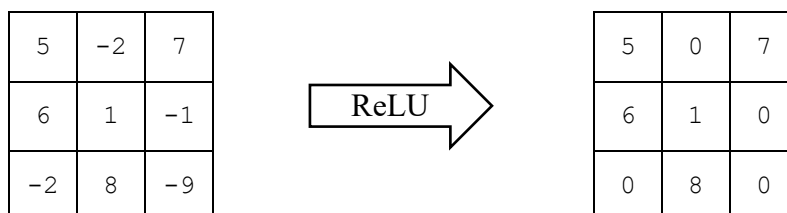


Figure 3.11 Demonstration of ReLU operation: A ReLU operation on a sample matrix with input on the left and output on the right.

In other words, the function changes all the negative value with 0 and the positive value remains the same. Note that same as pooling layer, there are no parameters needed to be trained.

3.1.5 Fully Connected Layer

After the previous operation (pooling or convolution), a fully connected layer, like its name, connects all the input feature maps from the previous layers to a larger feature map. In other words, the operation transforms the multi-dimensional input matrix, possibly of size $K \times K \times F$, from the previous layer to a vector (one-dimensional array) [110]. Note that there

is more than one fully connected layers in general, and each of the layers is followed by an activation function. In particular, the last fully connected layer inside a CNN connects all the features to the actual predicted objects (such as a cat or dog), which will provide a score for each object.

The operation inside this layer is relatively simple as well, which simply flattens the output (such as input map of size $6 \times 6 \times 256$) from the previous layer as input to a one-dimensional vector (a 9216×1 vector) [114]. Note that for the first fully connected layer in CNN, the input feature map is of size $K \times K \times F$, which is transformed into one huge vector by taking it row by row. The transformation is similar to the conversion of filters into a weight matrix in the convolutional layer. Then, the input vector \mathbf{x} is multiplied by a weight matrix (W) plus a bias (\mathbf{b}). The operation is the same as a neuron function in Equation (3.17) with a weight matrix W , a bias vector \mathbf{b} , and an activation function g .

$$f(\mathbf{x}) = g(W \times \mathbf{x} + \mathbf{b}) \quad (3.17)$$

Considering the first fully connected layer in AlexNet with a transformed vector of size 9216×1 , we initialize a weight matrix of size 4096×9216 , which would have an output vector of size 4096×1 [114].

Then, the output is treated as an input to the next fully connected layer, which is the last fully connected layer in AlexNet and maps to 1000 distinct classes.

3.1.6 SoftMax Layer

In this layer, the input one-dimensional array is translated into a vector with probability for each class via a SoftMax function [117]. The SoftMax function is defined as

$$\sigma(\mathbf{y})_j = \frac{e^{y_j}}{\sum_{j=1}^K e^{y_j}} \quad (3.18)$$

where K is the number of classes, $\mathbf{y} = [y_1, \dots, y_j, \dots, y_K] \in \mathbb{R}^K$ and j refers to the j -th index in the vector \mathbf{y} . Note that

$$\sum_{j=1}^K \sigma(\mathbf{y})_j = 1 \quad (3.19)$$

which makes sense since the sum of all the probability is 1. For example, if the input is a vector of form $[1, -2, -1]$ and there are 3 classes in total, then the output is

$$\left[\frac{e}{e+e^{-2}+e^{-1}}, \frac{e^{-2}}{e+e^{-2}+e^{-1}}, \frac{e^{-1}}{e+e^{-2}+e^{-1}} \right] \quad (3.20)$$

which is [0.884, 0.042, 0.114].

3.2 Training CNN

The ambition of CNN is to classify new data into proper classes with a high score or high probability. As a result, a strong accurate model is vital, which is able to identify all the training datasets with a low error rate. Intuitively, minimizing the differences between output result and given true label would be one of the options, which consists of minimizing the mean square error (loss function) between the predicted result and the actual label. However, since the classification output is a probability result from SoftMax layer in AlexNet, the other common method is introduced instead, which is to minimize the cross-entropy equation (loss function).

In CNN, training is a process of finding the optimal parameters in the convolutional layers and the fully connected layers. Training is accomplished by applying a gradient descent method, one of the techniques to solve the minimization problem and is commonly used in neural networks [110], [111], [114]. In particular, backpropagation is used in the gradient descent method to find the partial derivative with respect to the learnable parameters since, generally, there are multiple layers inside a network, which means that chain rules are applied multiple times when computing the gradients (see Section 3.2.2 Backpropagation).

One of the important measures of the generated model is the loss value, which is a measure of the distance between the predicted scores (or probability) and the truth labels. It is also a measure of the quality of the solution to the optimization problem. We want to minimize the loss as closer to zero as possible. Training is completed by multiple iterations, while the learnable parameters are updated during each iteration to minimize the loss value. Before the first iteration of the training, weight initialization is done by providing small non-zero random numbers to the weight matrix to ensure the parameters get updated correctly . Each input data is passed into the network and output a loss value from the last layer, which is called forward propagation [110]. After updating the result with the correct value, backpropagation is used to find the partial derivation of each of the weight matrix. Finally, parameters are updated according to the learning rate (hyperparameter) and the derivative of its, which finishes the current iteration [117]. Then, the process enters the next iteration with the new parameter (or weight matrix) [117]. Before discussing how to implement the

network, several definitions are required.

3.2.1 Loss Function

In ML, the loss function, also known as cost function, is a function which outputs a loss value to measure how well the model is. Two types of loss functions are commonly adopted. One is the mean square error, which is typically employed to a continued value such as linear regression, and the other one is cross-entropy frequently addressed in multiclass classification because of its probability output [110]. In ML, we define the loss function as

$$\mathcal{L}(f, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2 \quad (3.21)$$

where $f(x_i)$ is the output vector from the last layer in the network, x_i is the input vector, N is the number of training examples and $\mathbf{y} = [y_1, \dots, y_i, \dots, y_N] \in \mathbb{R}^N$ is the actual output, defined as follows

$$y_i = \begin{cases} 1, & \text{if } x \in i\text{-th class} \\ 0, & \text{otherwise} \end{cases}, \text{ for } 1 \leq i \leq N \quad (3.22)$$

is the true label vector. The cross-entropy of the discrete probability distribution q relative to a correct discrete probability distribution p is defined by [117]:

$$H(p, q) = -\sum_{\mathbf{x}} p(\mathbf{x}) \log(q(\mathbf{x})) \quad (3.23)$$

where $\mathbf{x} = [x_1, \dots, x_n] \in \mathbb{R}^n$ is an input to the probability distribution function. When applying the cross-entropy to a loss function, the cross-entropy loss has the form [117]:

$$\mathcal{L}_i = -\log \left(\frac{e^{f_{y_i}}}{\sum_i e^{f_j}} \right), \text{ or equivalently } \mathcal{L}_i = -f_{y_i} + \log \left(\sum_i e^{f_j} \right) \quad (3.24)$$

where f_i denotes the j -th element of the output vector f from the SoftMax layer. Note that the formula is derived from Equation (3.23) where

$$q = \frac{e^{f_{y_i}}}{\sum_j e^{f_j}} \quad (3.25)$$

which is the predicted probability coming from the SoftMax layer, and $p = 1$ which means the correct probability for this j -th element is 1 [117].

Figure 3.12 describes an extremely simple example of a network with only two layers: the fully connected layer and the SoftMax layer. Given an input data x_i with only three values, an intermediate weight matrix W and a bias b , an output vector is produced from the layer and is set as an input to the SoftMax layer, which outputs a probability for each class in a vector form. After that, the cross-entropy loss is calculated by the natural log and based on the given true predicted label ($y_i = 1$), which is equal to 0.03. Note that the goal is to leave the loss as smaller as possible, then the probability would close to 1.

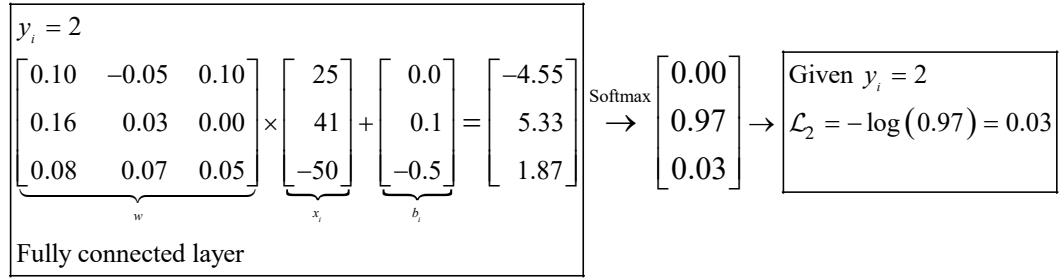


Figure 3.12 Example of loss computation: A demonstration of how the loss being computed from an input x_i of size 3×1 .

In general, given a training set of n data (such as images) and m number of actual classes, then the general loss function for a multiclassification problem is [117]:

$$\mathcal{L} = \sum_{i=1}^N \mathcal{L}_i \quad (3.26)$$

where

$$\mathcal{L}_i = -\sum_{k=1}^m [y_{ik} \cdot \log(f_{ik})] \quad (3.27)$$

Note that \mathcal{L}_i is essentially the cross-entropy loss defined before, but in a more general form, and f_{ik} is the output from the SoftMax function, and

$$y_{ik} = \begin{cases} 1, & \text{if } x \in k\text{-th class} \\ 0, & \text{otherwise} \end{cases} \quad (3.28)$$

In particular, when $K = 2$, we have:

$$\mathcal{L}_i = -\sum_{k=1}^m [y_{ik} \cdot \log(f_{ik}) + (1 - y_{ik}) \cdot \log(1 - f_{ik})] \quad (3.29)$$

in which the conditional distribution of the target classes given input data x is a Bernoulli distribution of the form:

$$p(y_i | x_i, w) = [f_i(x_i)]^{y_i} \cdot [1 - f_i(x_i)]^{1-y_i} \quad (3.30)$$

where the result is 1 if $f_i \geq 0.5$ and 0 if $f_i < 0.5$. Then, the loss function is simply the negative log of the conditional distribution given above, which makes sense since we want to maximize the above conditional probability which is equivalent to minimize the negative log of the function. To minimize such general form of the loss function, we apply the optimization algorithm called Stochastic Gradient Descent (SGD) [110], [111].

3.2.2 Backpropagation

Backpropagation is an algorithm for efficiently finding the gradient of a function with respect to each learnable parameter especially in NNs, since each neural network model can be represented by computation graphs [111]. With that graph, the result can be deduced from the top layer to the bottom very quickly because of the chain rule.

First, we define the function

$$y = f(x) = g\left(W^L \cdot \dots \cdot g\left(W^2 \cdot g\left(W^1 \cdot x + b^1\right) + b^2\right) + \dots + b^L\right) \quad (3.31)$$

where $g(\bullet)$ is the activation function, each L refers to the layer number, W^i is the weight matrix and b_i refers to bias for $1 \leq i \leq L$. Note that the above structure is a typical form of the output from a NN. To demonstrate how the backpropagation process, we consider the composite function

$$y = f(g(x)) \quad (3.32)$$

Then, if we calculate the derivative with respect to the x , by the chain rule, we have

$$\frac{\partial y}{\partial x} = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial x} \quad (3.33)$$

In order to find $\partial y / \partial x$, we first calculate $\partial f / \partial g$ with the value g we know from the forward direction, Then, we obtain the result by multiplying it with $\partial g / \partial x$.

For the value of

$$f(x) = x^3 \quad g(x) = x^2 \quad x = 2 \quad (3.34)$$

we can find the derivative of the f with respect to x by applying the backpropagation algorithm (Figure 3.13).

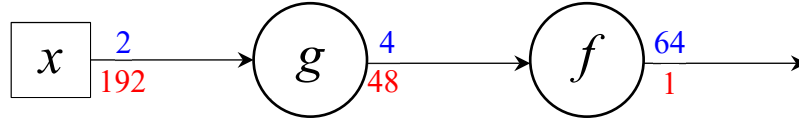


Figure 3.13 Example of backpropagation process: An example of finding $\partial y/\partial x(2)$ by applying the backpropagation algorithm. The blue number indicates the forward direction, while the red number indicates the backward direction started from the end to the beginning.

After the corresponding values in the forward direction are successfully calculated, backpropagation is used to compute the derivative. Since $\partial f/\partial f = 1$, the last value is always 1. We then have

$$\frac{\partial f}{\partial g} = 3g^2 \quad \left. \frac{\partial f}{\partial g} \right|_{g=4} = 48 \quad \left. \frac{\partial y}{\partial x} \right|_{x=2} = 192 \quad (3.35)$$

Another example of the backpropagation process is demonstrated below, which is done in a NN with only four layers. There are two sets of fully connected layer and ReLU layer, one set followed the other set. Then, we can represent the network model by

$$f(x) = g(W^2 \cdot g(W^1 \cdot x + b^1) + b^2) \quad (3.36)$$

where the inner layer $g(W^1 \cdot x + b^1)$ represents the first fully connected layer with weight matrix W^1 and bias b^1 , and the outer layer $g(W^2 \cdot x + b^2)$ represents the second fully connected layer with weight and bias, W^2 and bias b^2 respectively. Note that this is a composition function for two layers and there is no learnable parameter from the ReLU layers. Consider the case that $W^1 = 1$, $W^2 = -1$, $b^1 = 1$, $b^2 = 5$ and $x = 3$, Figure 3.3 illustrates the computation graph in a forward direction and the corresponding derivation in the backward direction. In particular, for the ReLU function, $f(x) = \max(0, x)$, the derivative function, denoted as $F(x)$, is defined as follows

$$F(x) = \begin{cases} \frac{\partial f}{\partial x}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases} \quad (3.37)$$

Since f is not differentiable at $x = 0$.

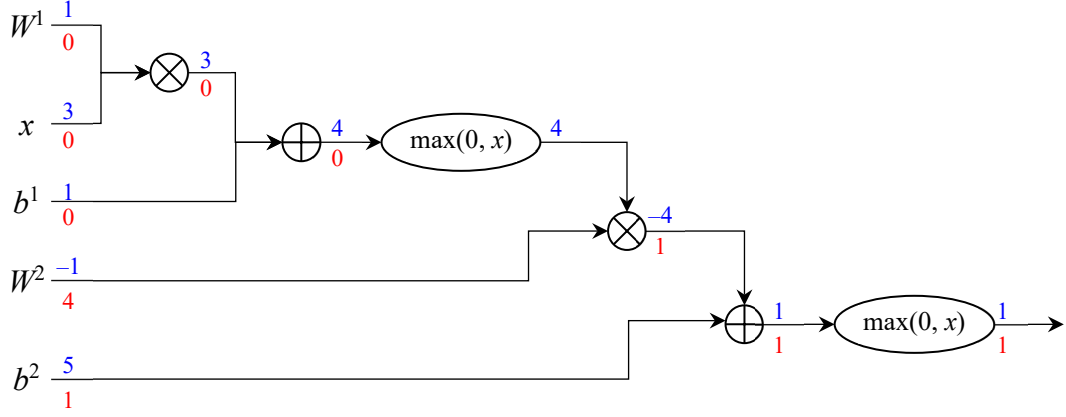


Figure 3.14 Example of backpropagation process in neural network: Acyclic computation graph of the network model $f(x) = g(W^2 \times g(W^1 \times x + b^1) + b^2)$, where g is a ReLU function ($f(x) = \max(0, x)$), $W = [1, -1]$, $b = [1, 5]$ and $x = 3$. The forward propagation computes values from input to output (blue) and the backpropagation is performed from the end to the beginning by applying the chain rule and the derivative rules (red)

In general, the backpropagation is done by the following pseudocode [111].

Backpropagation pseudocode

Input: A network with L layers, the activation function g_l the output hidden layer

$$f_l = g_l(W \times f_{l-1} + b_l)$$

Compute the gradient $\delta = \frac{\partial \mathcal{L}_i}{\partial y}$

For $i = L$ to 0 **do**

 Calculate gradient for present layer

$$\frac{\partial \mathcal{L}_i}{\partial W_l} = \frac{\partial \mathcal{L}_i}{\partial g_l} \cdot \frac{\partial g_l}{\partial W_l} = \delta \frac{\partial g_l}{\partial W_l}$$

$$\frac{\partial \mathcal{L}_i}{\partial b_l} = \frac{\partial \mathcal{L}_i}{\partial g_l} \cdot \frac{\partial g_l}{\partial b_l} = \delta \frac{\partial g_l}{\partial b_l}$$

 Apply gradient descent using $\frac{\partial \mathcal{L}_i}{\partial W_l}$ and $\frac{\partial \mathcal{L}_i}{\partial b_l}$

 Back-propagate gradient to the lower layer

$$\delta = \frac{\partial \mathcal{L}_i}{\partial g_l} \cdot \frac{\partial g_l}{\partial g_{l-1}} = \delta \frac{\partial g_l}{\partial g_{l-1}}$$

End

3.2.3 Stochastic Gradient Descent

Stochastic Gradient Descent (SGD) is among the most fundamental and widely adopted optimization algorithms [111], [118]. The primary objective of SGD is to efficiently minimize an objective function by iteratively adjusting model parameters based on the gradient of the loss function. Unlike standard Batch Gradient Descent, which computes the gradient over the entire dataset before updating parameters, SGD makes incremental updates by evaluating a small, randomly selected batch of training samples in each iteration. This stochasticity introduces computational efficiency, allowing for faster convergence and better generalization, especially when dealing with large-scale datasets.

In SGD, there is a trade-off between computational speed and convergence stability. A key advantage of SGD is its ability to significantly reduce training time without compromising the accuracy of the final model. Since training data within the same class often exhibits correlated patterns, using a subset of the data to approximate the full gradient does not significantly distort the optimization path. Instead, it enables faster learning cycles, leveraging the computational benefits of matrix-vector operations, particularly when batch sizes are chosen as powers of 2 (e.g., 32, 128 or 256), aligning with hardware-efficient operations in deep learning frameworks. The learnable parameter for the k^{th} iteration to update the rule governing SGD follows the expression:

$$w_{k+1} = w_k - \eta \frac{1}{N} \sum_{i=1}^N \frac{\partial \mathcal{L}_i}{\partial w_k} \quad (3.38)$$

where:

- w represents the model parameters,
- η is the hyperparameter or learning rate,
- N is the batch size, and
- \mathcal{L}_i denotes the loss function for an individual training sample.

This formulation ensures that the model iteratively updates its parameters in a direction that minimizes the average loss over the selected batch. The *backpropagation algorithm* computes these gradients and keeps adjusting weights accordingly. A structured pseudocode representation of SGD provides a clear blueprint for its implementation in practical settings.

Algorithm: Stochastic Gradient Descent

Input:

- Loss function \mathcal{L}_i ,
- Learning rate η ,
- Dataset X, y ,
- Model $f(w, x)$ where $x \in X$

Output:

- Optimized parameters w that minimize L_i

Procedure:

1. Shuffle the dataset (X, y) .
 2. Partition the dataset into K mini-batches $B = [B_1, B_2, \dots, B_K] = X$
 3. For each batch B_j containing (x, y_i) samples:
 - Compute the gradient of the loss function
$$w_{k+1} = w_k - \eta \frac{1}{N} \sum_{i=1}^N \frac{\partial \mathcal{L}_i(f(w, x), y_i)}{\partial w_k}$$
 4. Repeat until convergence
-

This iterative process ensures that model parameters continually adapt to minimize the loss function effectively.

The role of the learning rate in convergence dynamics: Among the hyperparameters influencing SGD's performance, the learning rate η is arguably the most critical. It governs the step size in the parameter space and directly impacts the convergence behavior of the model.

- If η is too large: The optimization may diverge, oscillating around the minimum without settling.
- If η is too small: The convergence is guaranteed but at the cost of increased training time.

To balance this trade-off a common approach is learning rate decay, where η is reduced over time following:

$$\eta_t = \eta_0 \beta^{\lfloor \frac{t}{\epsilon} \rfloor} \quad (3.39)$$

where η_0 is the initial learning rate, β is the decay factor (often set to 0.1), and t represents the iteration step. In practice, reducing the learning rate by a factor of 10 every few epochs is an effective strategy to fine-tune convergence.

Visualizing SGD convergence behavior: The effectiveness of different learning rates can be observed through loss versus epoch graphs. These curves in plot illustrate how the loss function evolves over successive iterations, highlighting [119]:

- Rapid convergence at an optimal learning rate,
- Unable to converge and stuck at a high learning rate,
- Divergence at excessively high learning rates, and
- Slow but stable convergence at lower learning rates.

Thus, selecting an appropriate learning rate is both an art and a science, requiring empirical tuning based on the dataset and model complexity.

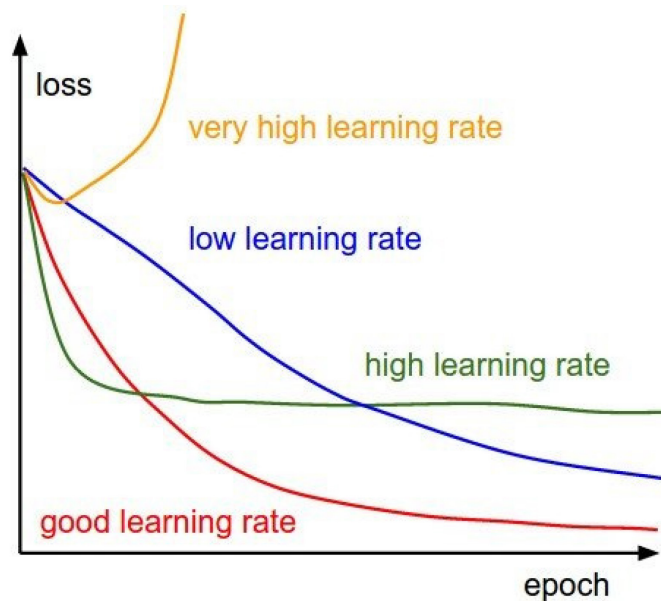


Figure 3.15 Effects of various learning rates on convergence

Chapter IV

Monitoring Diagnosis for PdM of SPV-IRT

4.1 Deep Learning for IRT-Anomaly Diagnosis

In the evolving landscape of PdM for SPV energy systems, the demand for intelligent remote aerial IRT anomaly diagnosis has become increasingly critical. As large-scale SPV power plant installations continue to expand, ensuring operational reliability, efficiency, and longevity requires advanced diagnostic methodologies capable of early-fault detection and proactive maintenance strategies. Integrating DL algorithms into SPV-IRT image analysis presents a significant leap toward sustainable and intelligent energy management.

Unlike traditional rule-based diagnostic approaches, DL-based systems are trained to recognize complex thermal signatures employing computer vision (CV) techniques to decode intricate thermal defect patterns that may not be immediately apparent to human inspectors, reducing downtime and maintenance costs. Meanwhile, AI-based thermographic diagnostics can enhance fault detection precision by filtering noise, reducing false positives, and improving classification accuracy, making them reliable for large-scale autonomous inspections.

4.1.1 Related Works

DL methodologies particularly leveraging CNNs have been extensively explored for the classification of thermal anomalies in IRT images of SPV panels. In this context, various state-of-the-art models and algorithms have been employed to enhance diagnostic accuracy and streamline fault detection processes. Some of them are mentioned here:

Haidari et al. utilized infrared thermal images categorized into three distinct classes – healthy, hotspot, and hot substring – each represented by a resolution of 70×50 pixels. The preprocessing pipeline involved medium blurring and grayscale conversion, implemented via Adobe Photoshop[®], to refine image quality and reduce noise. Subsequently, a pre-trained VGG16 model was deployed for classification, yielding fault detection accuracy of 98%

[120].

Le and others conducted an in-depth investigation into fault classification of SPV panels using grayscale IRT images spanning 12 classes, each with a resolution of 24×40 pixels. To enhance model generalization and mitigate data imbalance, an extensive preprocessing pipeline was implemented, incorporating data augmentation techniques such as horizontal and vertical flipping, as well as $\pm 20\%$ size translation. Additionally, an optional synthetic minority oversampling technique (SMOTE) was employed to address class imbalance by generating synthetic samples for underrepresented categories. Leveraging the power of transfer learning, the pre-trained ResNet model was utilized and an ensemble strategy was adopted by integrating 15 models through various combination methods. This ensemble learning approach significantly bolstered classification robustness culminating in an 86% fault classification accuracy [121].

Korkmaz and Acikgoz utilized grayscale IRT images across 12 classes, each with a resolution of 24×40 pixels. To enhance model robustness and mitigate data imbalance, an augmented oversampling strategy was employed, incorporating $\pm 30\%$ brightness variation, two instances of image reversal, and a single 180° rotation. This preprocessing approach aimed to generate a more diverse and representative data set for improved generalization. A pre-trained AlexNet model was deployed, leveraging multiscale TL with a novel three-parallel-branch architecture applied specifically to 11 selected classes. This innovative strategy facilitated hierarchical feature extraction at multiple spatial scales, capturing intricate patterns in SPV anomalies with greater precision. The proposed model demonstrated a fault classification accuracy of 93.51%, underscoring the potential of multi-scale TL in enhancing SPV thermal defect diagnosis [122].

Duranay investigated grayscale IRT images spanning 12 classes, each initially sized 24×40 pixels for intelligent thermal fault diagnosis. To enhance feature representation and improve model performance, the images underwent a meticulous preprocessing pipeline, including resizing to 224×224 pixels and the extraction of exemplary patches at 28×28 and 56×56 resolutions. The study employed a pre-trained EfficientNet B0 model for feature extraction, capitalizing on its superior balance between computational efficiency and representational power. To refine the extracted features, neighborhood component analysis (NCA) was utilized as a feature selection technique, optimizing the most relevant downstream classification. An SVM classifier was then applied to leverage the selected features for robust fault classification. To ensure model reliability and generalization, a

rigorous 10-fold cross-validation strategy was incorporated. This integrated approach achieved an anomaly classification accuracy of 93.93%, highlighting the effectiveness of combining DL-based feature extraction, statistical feature selection, and traditional ML classifiers. The findings reinforce the potential of hybrid-AI frameworks in enhancing PDM strategies for intelligent fault diagnosis [123].

Pamungkas et al. conducted a multiclass classification study leveraging grayscale IRT images across 12 distinct anomaly classes, each originally sized at 24×40 pixels. To enhance model generalization and address class imbalances, a robust preprocessing pipeline was implemented, incorporating augmentation techniques such as geometric transformations alongside synthetic data generation via generative adversarial networks (GANs). For feature extraction and classification, the pre-trained coupled UdenseNet model was employed, capitalizing on its densely connected convolutional layers and upsampling capabilities to capture intricate spatial and textual patterns within the grayscale IRT images. This sophisticated hybrid approach facilitated a more refined representation of fault characteristics, enabling improved thermal anomaly detection in SPV panel inspections. By integrating augmentation strategies with a DL architecture optimized for feature preservation and hierarchical learning, the study achieved a fault classification accuracy of 95.72% [124].

4.1.2 Materials

4.1.2.1 Dataset Preparation and Visualization

A customized, severely imbalanced 6-class radiometric IRT dataset (temperature intensity matrix) was meticulously curated to facilitate the precise thermal fault diagnosis of SPV panels. Unlike publicly available aerial grayscale IRT SPV array dataset [125], [126], which often lack structured categorization and labeled defect information, this dataset was systematically refined through rigorous visual assessment. Each defective SPV panel was identified and classified in accordance with International Electrotechnical Commission standards (IEC TS 62446-3:2017) [34], ensuring a standardized and reliable reference framework. The samples of defects are shown in Figure 4.1.

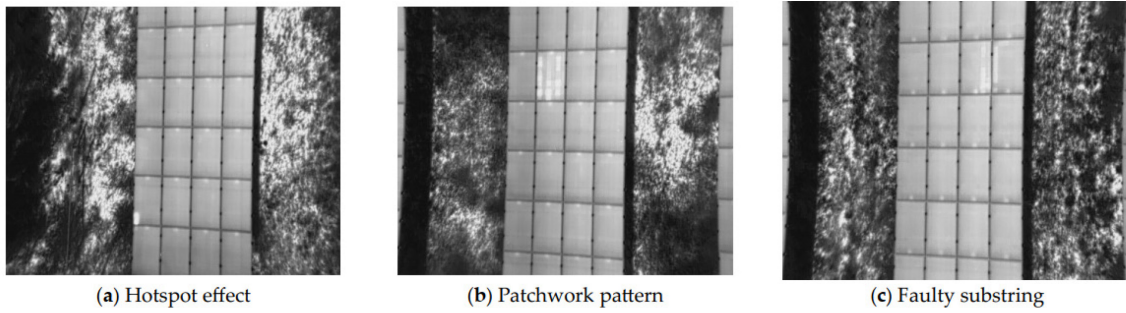


Figure 4.1 Raw aerial IRT-grayscale images of the SPV array having a visually defective SPV panel labeled as (a) hotspot effect, (b) patchwork pattern, (c) faulty substring.

The radiometric dataset construction is a challenging task to identify the grayscale image pixel position coordinates that are truly aligned with spatial temperature intensity values. This involves the extraction and resizing of individual SPV panel images from raw aerial grayscale IRT SPV array images and transforming them into radiometric temperature matrices with spatial dimensions of 60×100 pixels. Each panel was numerically represented by floating-point temperature intensity values in degrees Celsius as shown in Figure 4.2. To maintain structural consistency, zero-padding was applied, extending the images to 66×106 pixels as portrayed in Figure 4.3.

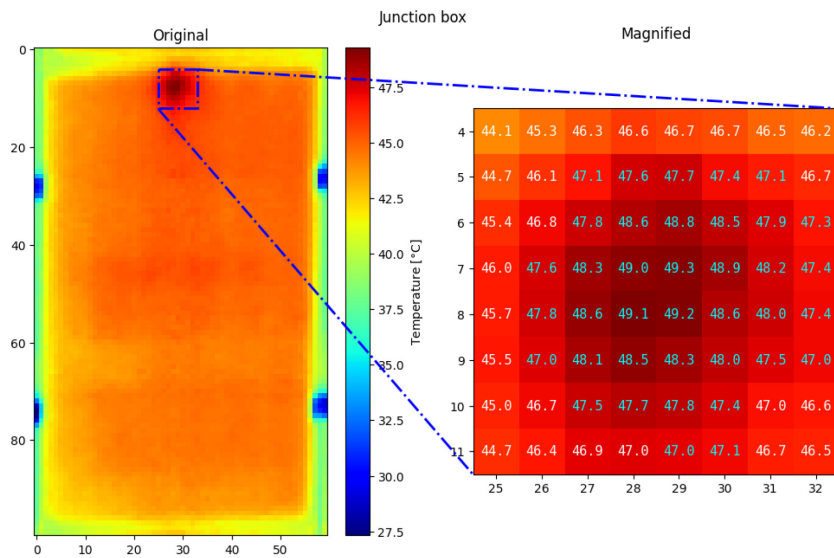


Figure 4.2 Extracted SPV panel labeled as heated Junction Box: Original and magnified region of 2D radiometric thermal datapoints—temperature intensity matrix [$^{\circ}\text{C}$] (jet colormap as pseudo-color applied for visual depiction).

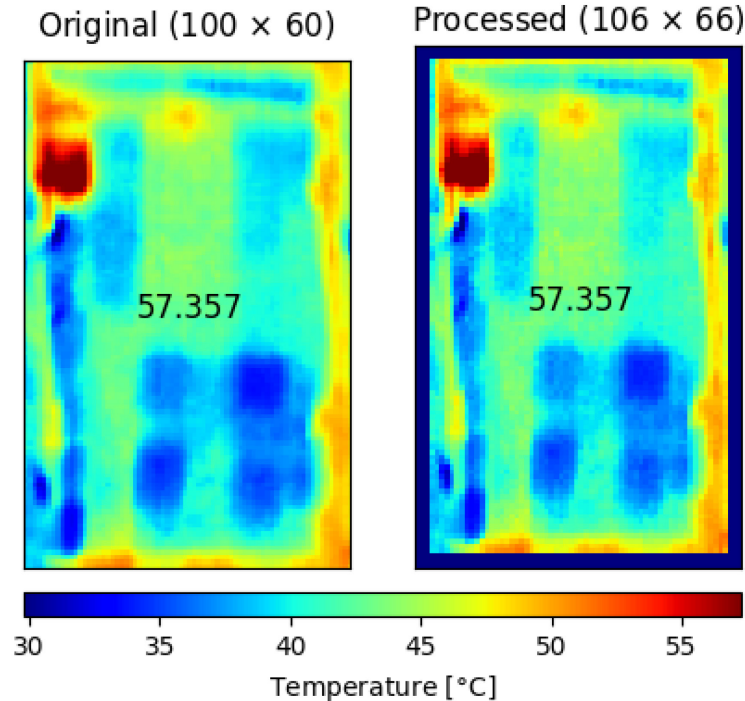


Figure 4.3 Original SPV panel image of 100×60 pixels (left) and zero-padded 106×66 pixels (right) having maximum temperature [$^{\circ}\text{C}$] with jet colormap as pseudo-color applied for visual depiction (red shades corresponds to higher temperature values, whereas blue to lower ones).

Further, a notable challenge in this radiometric dataset was the severe class imbalance; to train, tune, and validate the DL model for multiclass classification. Of the 2672 instances, an overwhelming majority 2647 belonged to the “Good” class (i.e., normal panels with no detected anomalies). In contrast, only 25 instances were categorized into five distinct “Faulty” classes, each containing just five samples. These faulty classifications encompassed critical defect types: faulty multistring, heated junction box, faulty substring, patchwork pattern, and hotspot affect – each visually distinguishable via pseudo-color thermal mapping as shown in Figure 4.4.

The offline data augmentation was implemented, incorporating horizontal and vertical flipping to enhance the visual diversity of the underrepresented faulty classes, i.e., without enhancing or balancing minority instances quantity. Furthermore, min-max normalization was employed to rescale the temperature intensity values into the $[0, 1]$ range, optimizing the dataset for subsequent DL-based analysis. The distribution of spatial temperature intensity values across different classes was effectively visualized using raincloud plots, offering an intuitive representation of thermal variations among normal and defective panels as shown in Figure 4.5 and Figure 4.6.

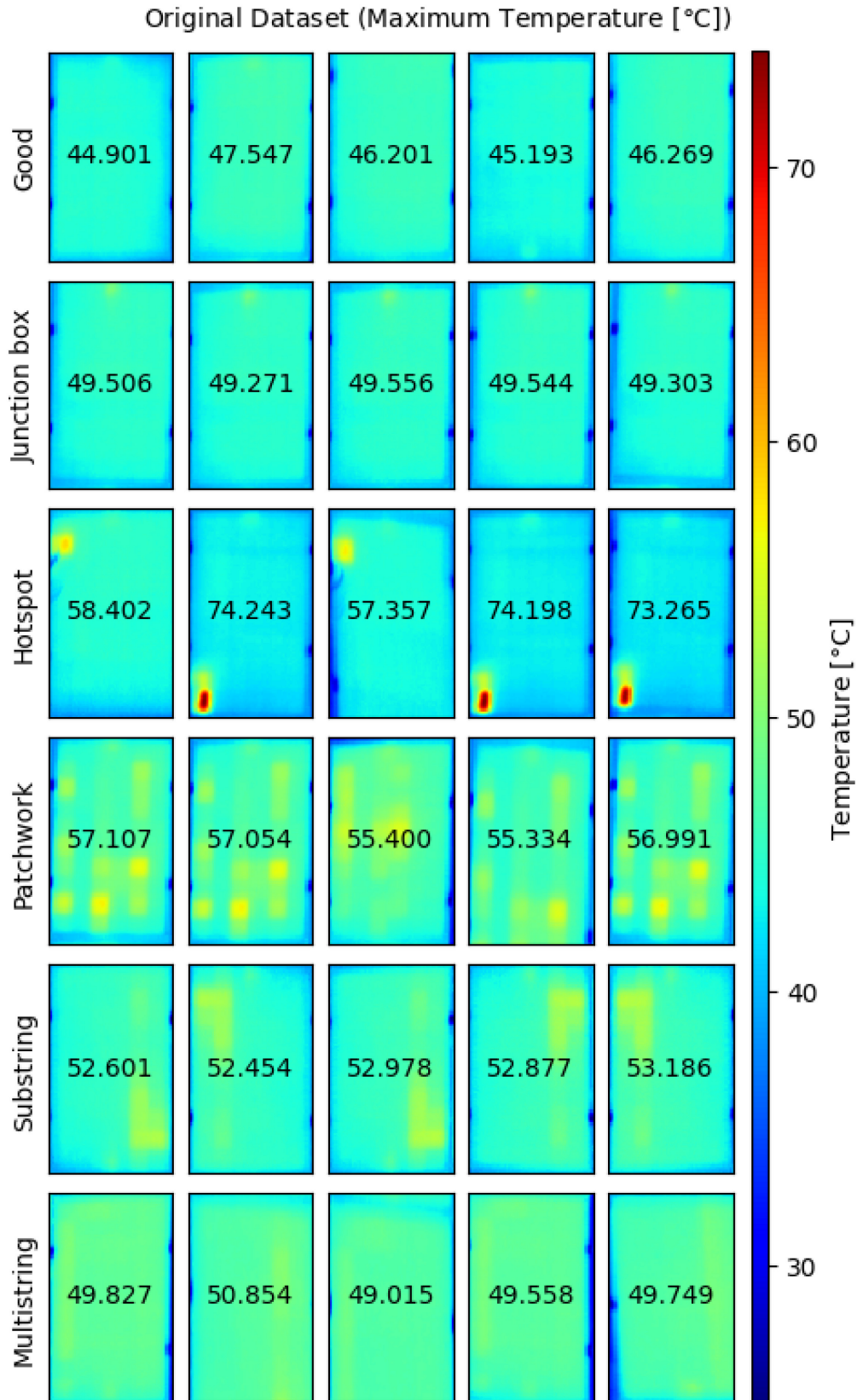


Figure 4.4 Original dataset of 6-class SPV panels with jet colormap applied for pseudo-color visual depiction (max temperature [°C]). Color bar represents temperature values corresponding to colors (higher temperature values in red shades, lower temperature values in blue shades). Data augmentation applied for visual diversity (multistring, substring, and hotspot).

This radiometric thermal dataset serves as a crucial foundation for advancing AI-driven PdM strategies in SPV energy systems, enabling the development of robust DL models capable of detecting intricate thermal anomalies with precision and reliability.

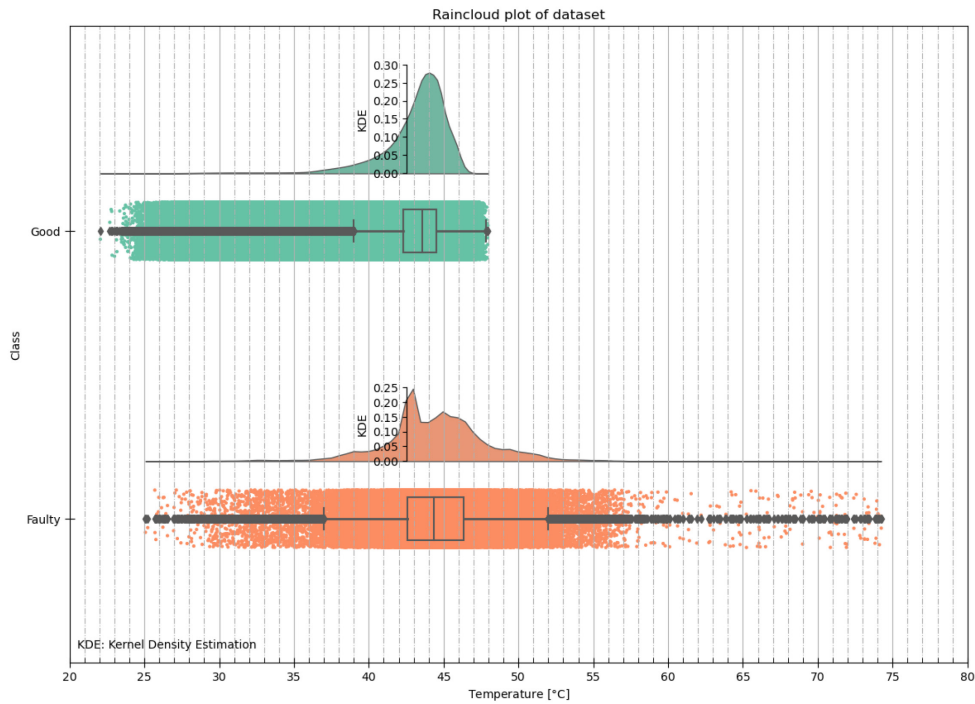


Figure 4.5 Raincloud plot of dataset (Good and Faulty class)

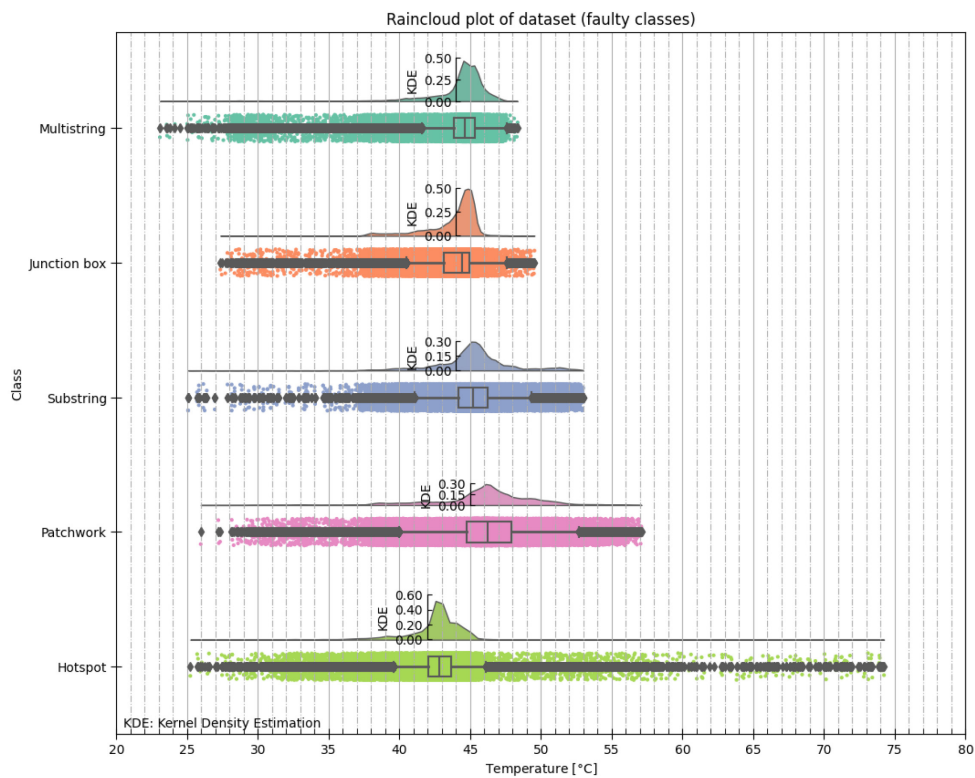


Figure 4.6 Raincloud plot of dataset (5-Faulty class)

4.1.3 Experimental Resources

The experiments were performed on a computer server with 50 CPUs of Intel® (Santa Clara, California, United States of America) Xeon® Gold 6130 Processors with 22 M Cache, 2.10 GHz. The server was equipped with 100 GB of memory (RAM). Further, the paid version resources of Google Colab (available online at <https://colab.research.google.com/>, last accessed on 30 January 2025) were also utilized.

The software environment is based on Python (v. 3.11.11) and Keras (v. 3.5.0) frameworks with TensorFlow (v. 2.17.1) at the backend to apply pre-trained DL models (online available: <https://keras.io/api/applications/>, last accessed on 30 January 2025). The performance evaluation and analyses employed the scikit-learn (v. 1.6.1) library whereas SMOTE resampling exploited the imbalanced-learn (v. 0.10.1) library. The explainable predictive visualizations exploited the tf-keras-vis (v. 0.8.7) library. For extracting and preprocessing floating-point temperature data from radiometric JPG (rJPG) images of solar photovoltaic arrays, the FLIR® Thermal Studio Suite thermography software (v. 2.0) was employed [127], [128].

4.1.4 Experimental Workflow

To systematically assess the operational health of SPV arrays, a drone-assisted IRT imaging system was deployed to remotely capture high-resolution grayscale thermal images. These raw grayscale thermographic images were then processed into a 2-dimensional radiometric thermal dataset (temperature intensity matrix), which was subsequently fed into a *custom CNN model* for automated anomaly detection and classification (radiometric thermal fault diagnosis). The experimental workflow developed in this study is shown in Figure 4.7.

1) Model Training via Ensemble Learning and Cross-Validation: To enhance the robustness and generalization of the diagnostic system, an *ensemble of four custom CNN models* was constructed using 4-fold cross-validation. Each iteration of the cross-validation process was used to train and validate a unique model within the ensemble. The dataset was partitioned as follows:

- 20% of the dataset (534 samples, including 5 anomalous instances, i.e. only 1-instance of each fault) were set aside as a final hold-out test set, ensuring an unbiased evaluation.

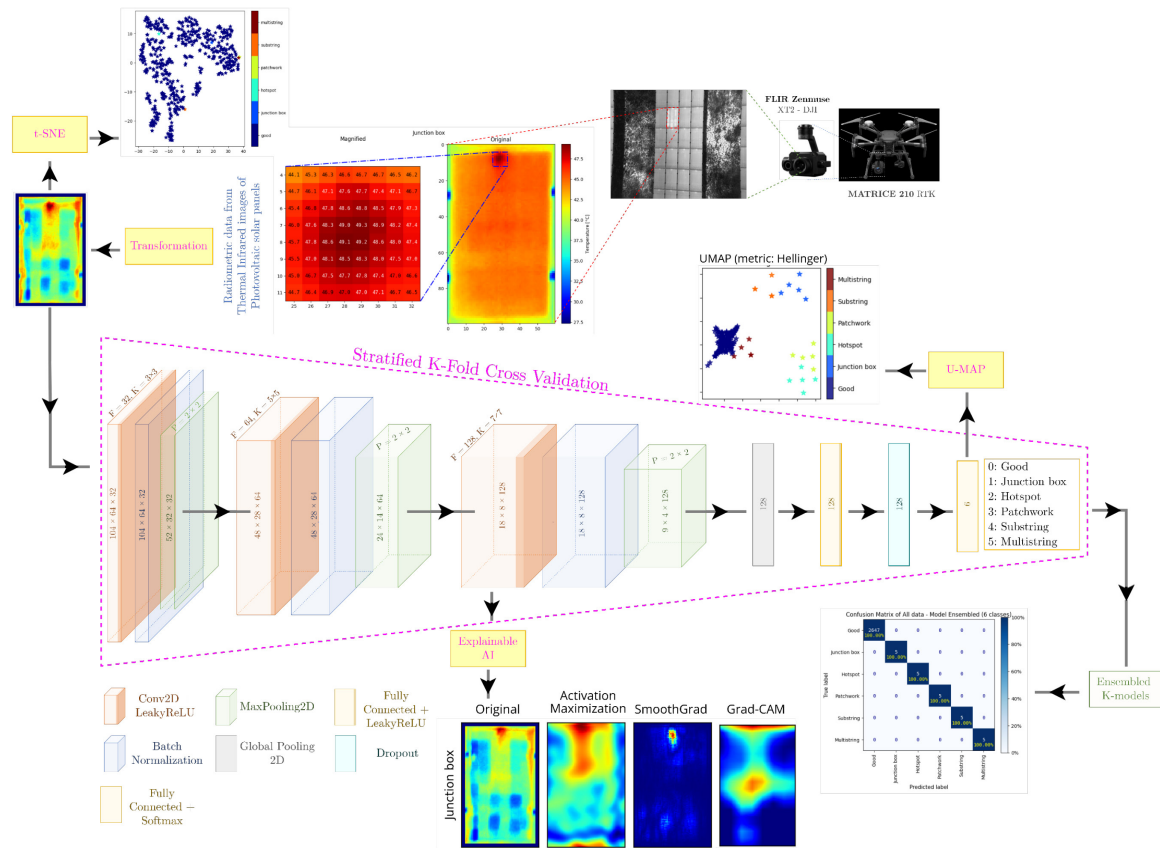


Figure 4.7 Experimental workflow. First, the dataset is acquired with a drone-aided IRT of SPV arrays. Then, an ensemble of CNN models is realized via a stratified 4-fold cross-validation. Finally, quantitative results are determined, and explainability techniques are used to unveil the mechanisms underlying the diagnostic process. Heatmaps have a jet colormap, with red representing higher temperature (for radiometric images) or activation values (for CNN explanations), and blue depicting lower temperature or activation values.

- During each cross-validation iteration, the remaining 80% of the dataset was split into:
 - 60% for training (1604 samples, including 15 anomalies, i.e. 3-instances of each fault)
 - 20% for validation (534 samples, including 5 anomalies, i.e. only 1-instance of each fault)

After completing the training cycles, the outputs of the 4-CNN models were averaged to form an ensemble prediction, which was subsequently tested on the hold-out dataset.

2) Addressing Class Imbalance via SMOTE: Given the extreme class imbalance, where defective panels were severely underrepresented, an additional experiment was conducted using the SMOTE to generate a more balanced dataset. The data set was resampled to achieve class parity, after which the following partitions were applied:

- 20% (3182 samples) was allocated as the final hold-out test set

- The remaining 80% was used in each cross-validation iteration with:
 - 60% (9544 samples) for training
 - 20% (3181 samples) for validation

By ensuring an even class distribution, this experiment aimed to assess whether data augmentation and oversampling techniques could improve the model's ability to detect rare radiometric thermal faults in SPV panels.

3) Explainable AI for Model Interpretability: Beyond achieving high performance anomaly diagnosis, the study incorporated XAI techniques with two the objectives:

- **Mathematical Interpretability:** To analyze and cluster the learned feature representations, offering insights into how different radiometric fault categories are distinguished at the feature level.
- **Perceptive Explanation Mapping:** To highlight the most influential radiometric datapoint regions contributing to classification decisions, using visual explanation techniques such as saliency maps and class activation heatmaps.

4.1.5 Results and Discussion

This section presents a comprehensive analysis of the experimental findings, highlighting the classification (radiometric thermal diagnosis) accuracy of the proposed model, the insights derived from XAI techniques, and the key trends observed during the training and validation process. By integrating multiple perceptive explainability methods, including Activation Maximization, SmoothGrad, and Grad-CAM, we provide a deeper understanding of how the DL model identifies and classifies radiometric thermal anomalies in SPV panels.

1) Insights from Perceptive Explainability Techniques: To ensure a robust evaluation across different cross-validation iterations, the results of various XAI techniques are compared.

- *Activation Maximization:* The maximum saliency regions vary across different cross-validation models, indicating that distinct parts of the radiometric thermal matrix are highlighted in each iteration as shown in Figure 4.8. This variation is not a major concern as CNNs are inherently translation-invariant, meaning the model can recognize patterns regardless of their spatial location within the data (radiometric thermal matrix).

- *SmoothGrad*: The application of SmoothGrad on the radiometric thermal matrix reveals consistent patterns across the trained models. Notably, all models correctly identified high-saliency regions near hotspots and junction box areas as shown in Figure 4.9, demonstrating the network's ability to focus on critical thermal anomalies.
- *Grad-CAM* (Gradient-weighted Class Activation Mapping): Grad-CAM visualizations extracted from convolutional layers reveal highly activated regions near hotspot areas (or high temperature locations) as shown in Figure 4.10. The observed class activation maps correspond closely to the final classification decisions, reinforcing the model's reliability in diagnosing radiometric thermal faults in SPV panels.

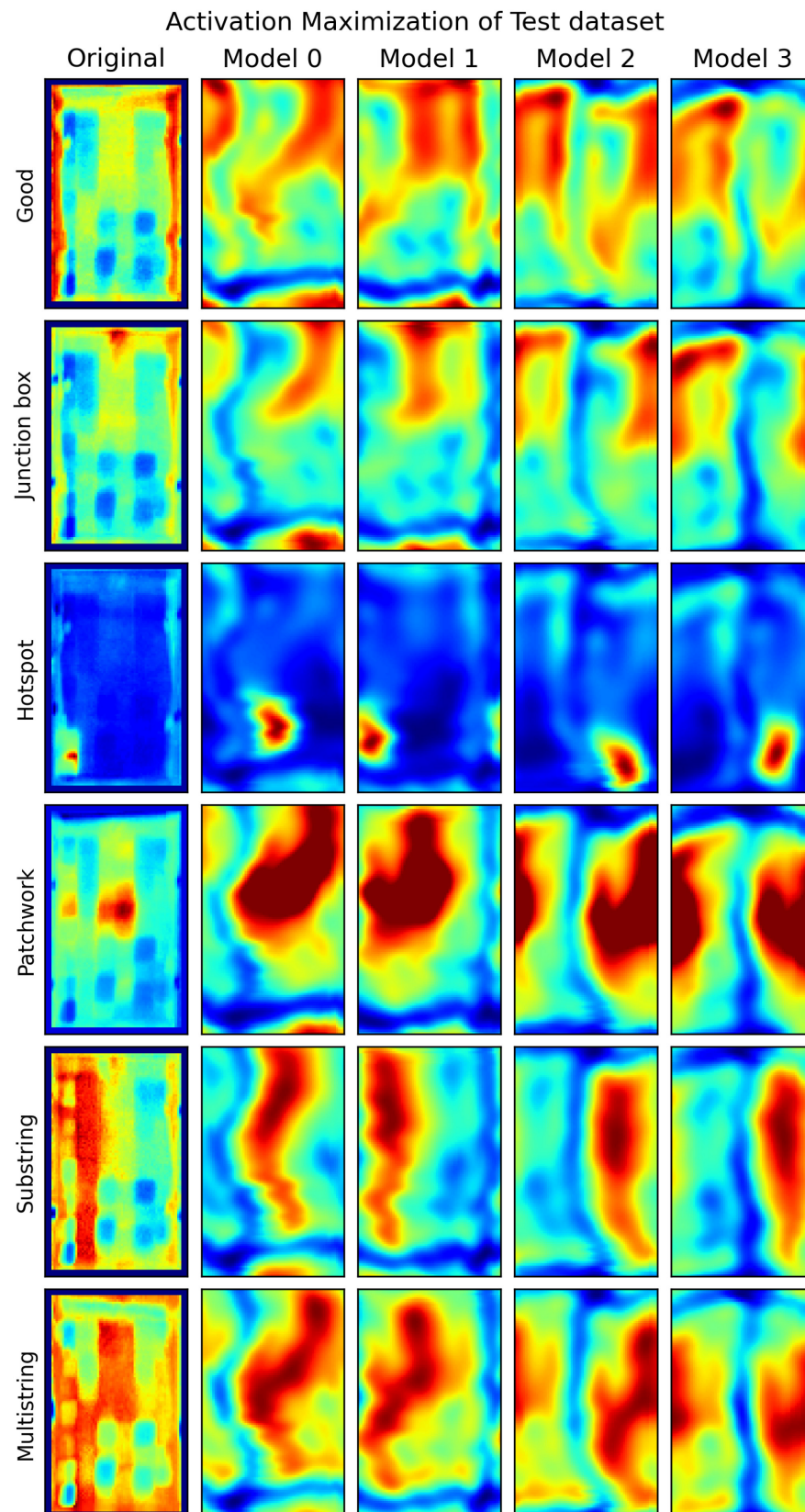


Figure 4.8 Explainable deep learning - Activation Maximization visualization from the last convolutional layer for the 6-class classification. Results for each iteration of the stratified 4-fold cross-validation are presented. Jet pseudo-color (red corresponds to higher activation values, whereas blue to lower ones) is used for visual depiction.

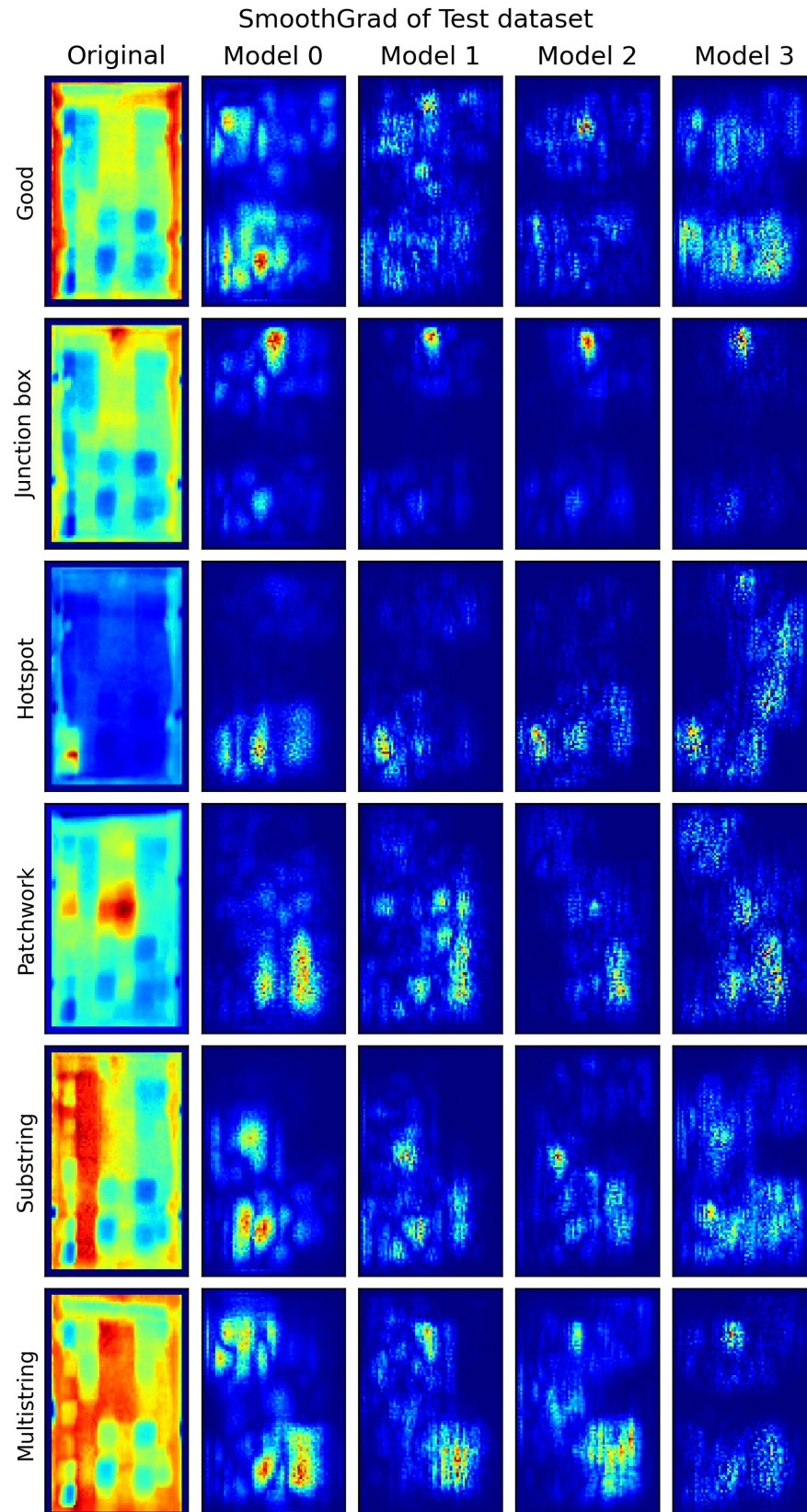


Figure 4.9 Explainable deep learning — SmoothGrad visualization from the last convolutional layer for the 6-class classification. Results for each iteration of the stratified 4-fold cross-validation are presented.

Jet pseudo-color (red corresponds to higher activation values, whereas blue to lower ones) is used for visual depiction.

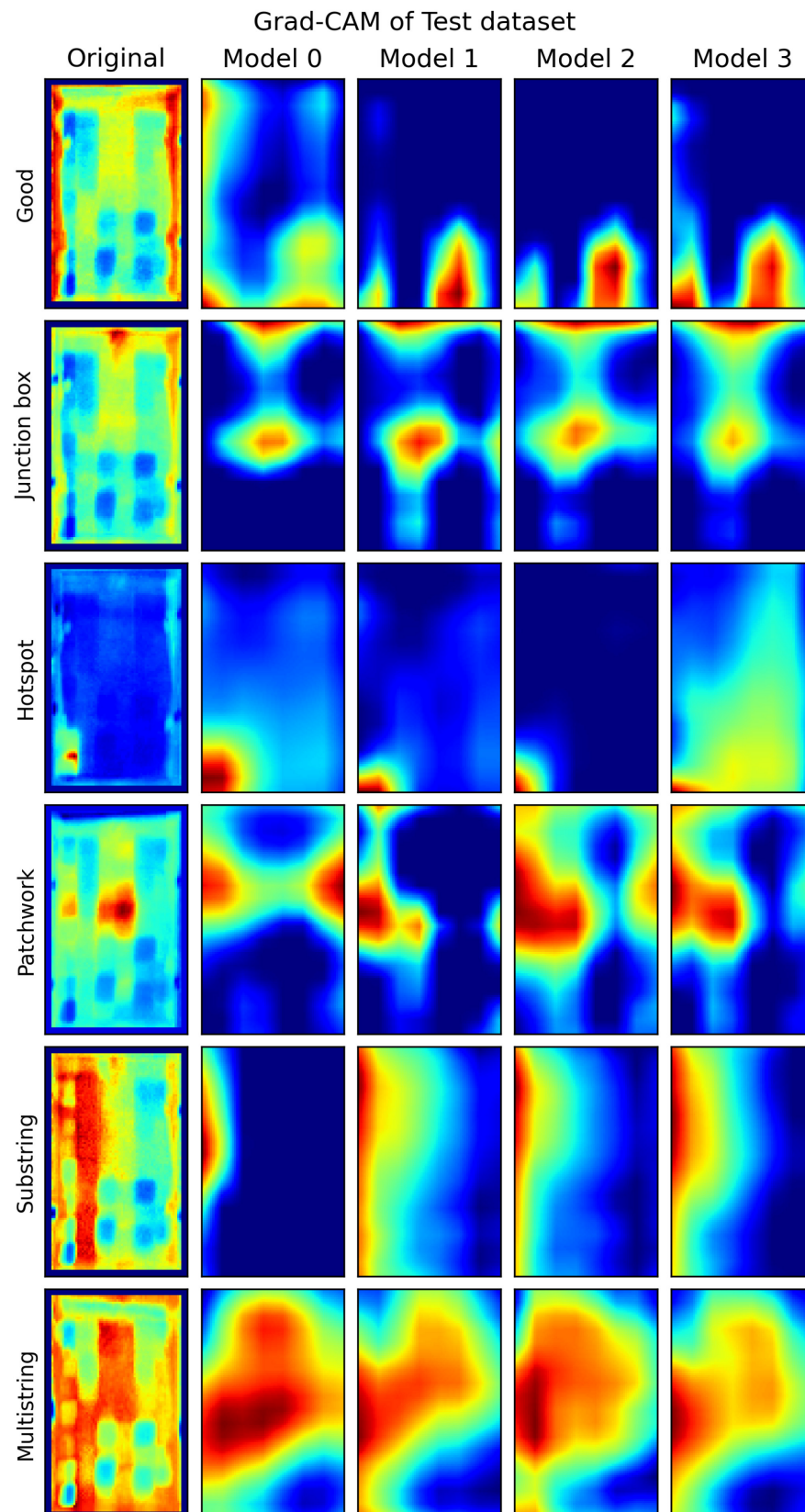


Figure 4.10 Explainable deep learning — Grad-CAM visualization from the last convolutional layer for the 6-class classification. Results for each iteration of the stratified 4-fold cross-validation are presented.

Jet pseudo-color (red corresponds to higher activation values, whereas blue to lower ones) is used for visual depiction.

2) Feature Space Visualization and Clustering Analysis: To further investigate how the model structures and organizes learned representations, we employed dimensionality reduction techniques to visualize feature embeddings:

- t-SNE (t-distributed Stochastic Neighbor Embedding) was used to depict the input data features of healthy versus faulty panels in a low-dimensional space.
- UMAP (Uniform Manifold Approximation and Projection), using the *Hellinger metric*, was applied to visually represent the CNN's predicted feature embeddings.

The results indicate that:

- In the original feature space, no clear separation is observed among healthy and faulty samples as shown in Figure 4.11.
- After training, distinct clusters emerge, indicating that the model successfully learns meaningful feature representations that differentiate among various types of SPV panel's radiometric thermal anomalies as shown in Figure 4.12.

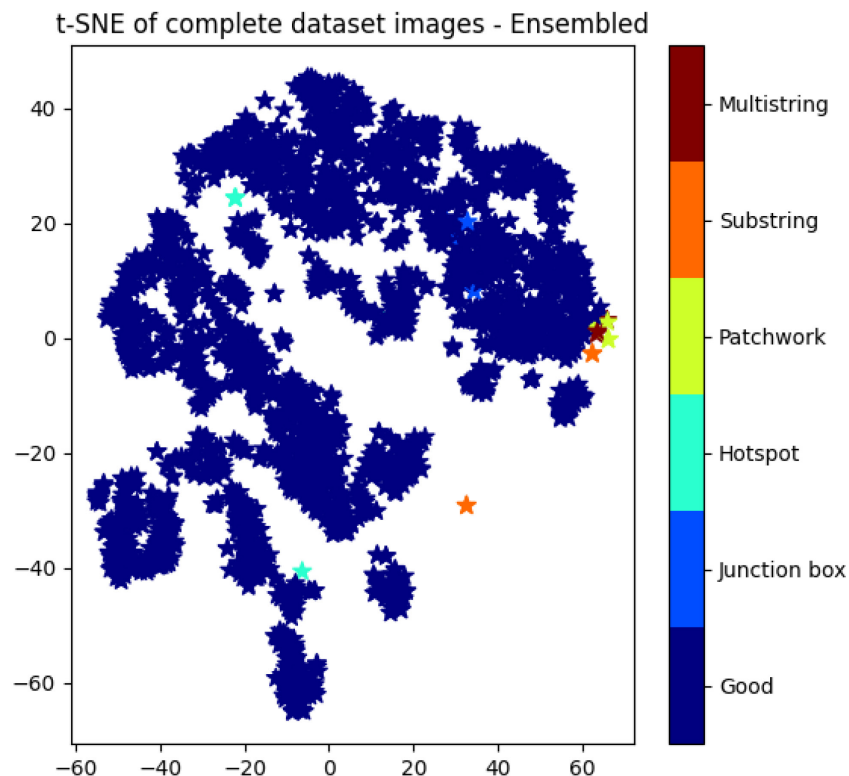


Figure 4.11 Embedding with t-SNE (cluster map) of complete input radiometric dataset images. Each data-point is depicted as a star-shaped point. Kindly note that some points overlap.

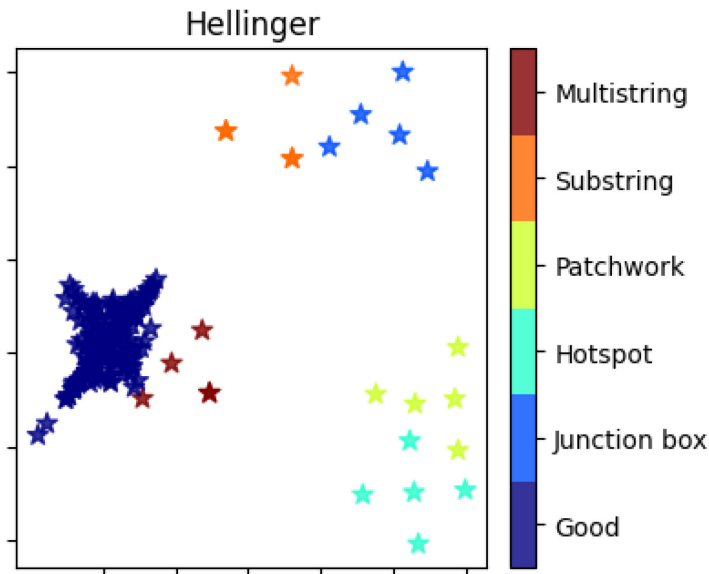


Figure 4.12 Embedding with UMAP (feature clustering) of test dataset predictions (Hellinger metric)—enssembled. Each data point is depicted as a star-shaped point. Kindly note that some points overlap.

3) Classification Performance and Cross-Validation Results:

- The stratified 4-fold cross-validation confusion matrices, presented in Figure 4.13 confirm the model's exceptionally high classification performance. The CNN models achieved an impressive $99.81 \pm 0.15\%$ accuracy (mean \pm standard deviation) across validation sets.
- However, it is important to note that the dataset was highly imbalanced, with relatively few anomaly samples, causing the accuracy to be skewed towards high values. To deal with the imbalance dataset, the *categorical cross-entropy* with *class-weights* are applied to test dataset.
- The ensemble of all four cross-validation models achieved a perfect 100% accuracy as illustrated in Figure 4.14, further reinforcing the model's effectiveness.

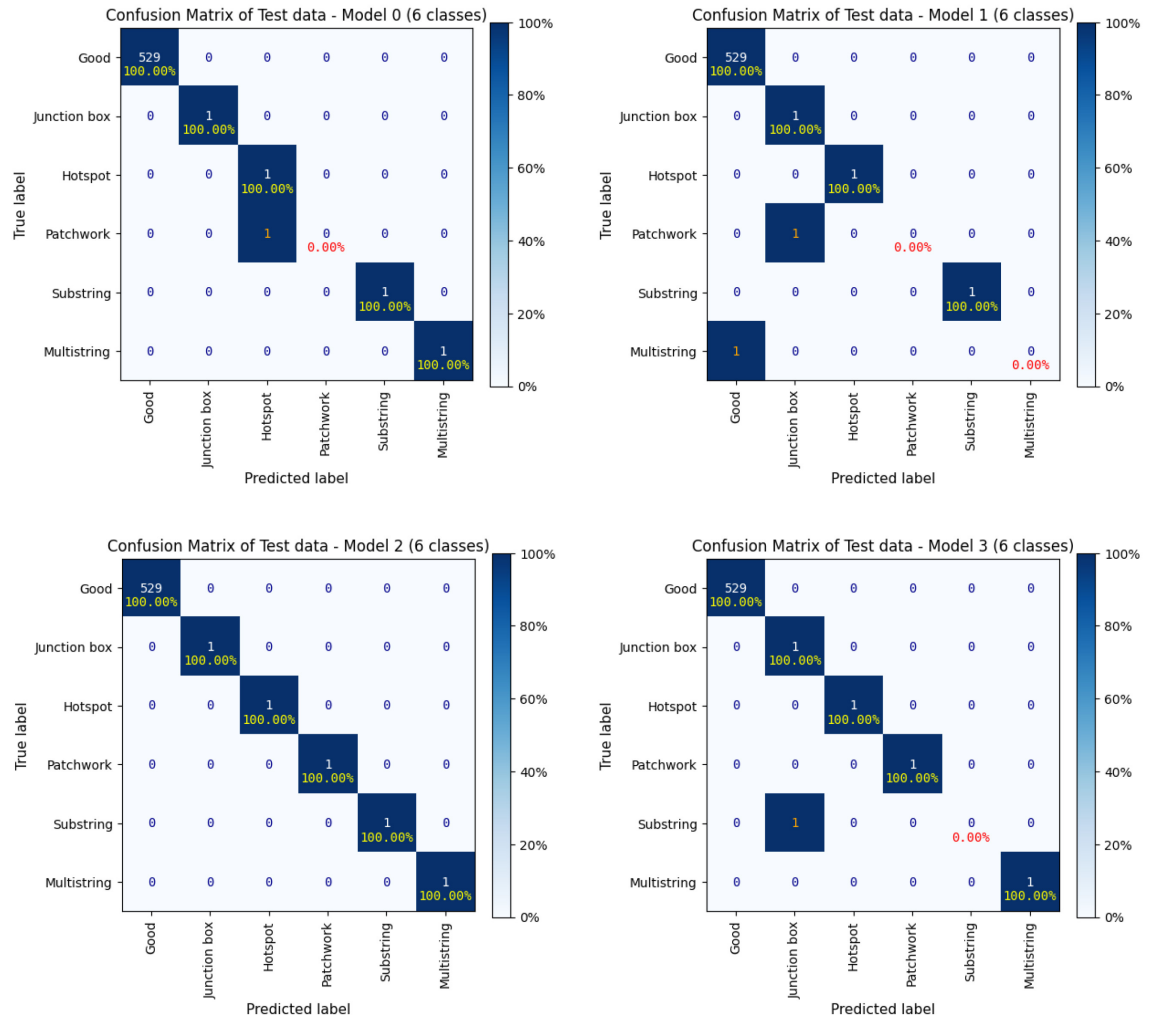


Figure 4.13 Confusion matrices of class-weighted test data for each iteration of the stratified 4-fold cross-validation (6-class).

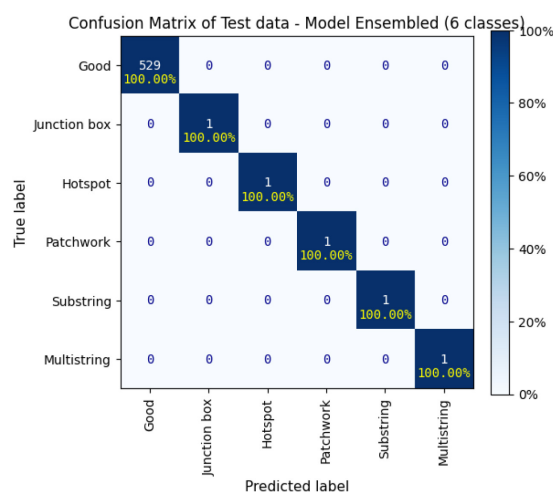


Figure 4.14. Confusion matrix of class-weighted test data ensemble model (6-class).

4) Training Dynamics and Learning Curves: The learning curves, shown in Figure 4.15, reveal that:

- The training process exhibited inherent noise, primarily due to the severely imbalanced data set.
- Fluctuations and skews in learning curves were observed across the 4-fold cross-validation models, reflecting challenges in optimizing the network under class-imbalanced conditions.

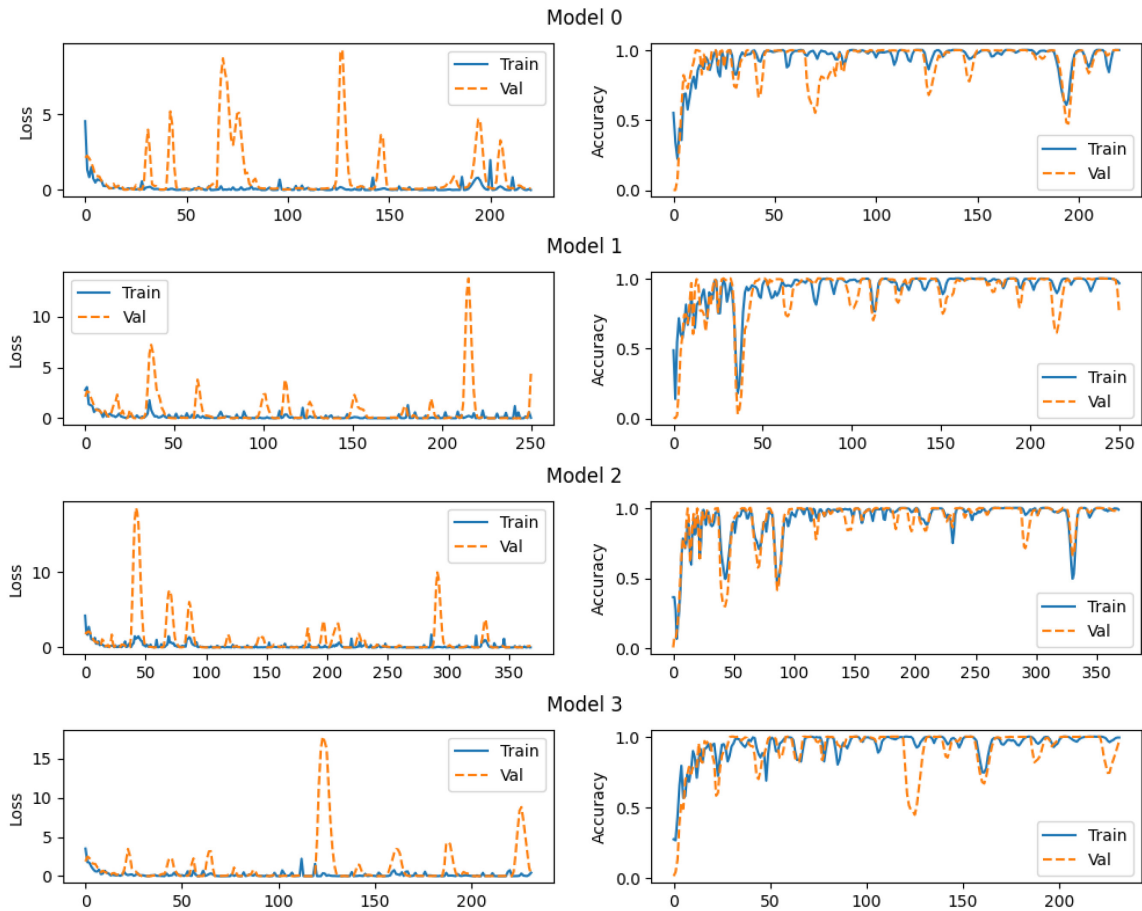


Figure 4.15 Categorical loss and accuracy of the training and validation of class-weighted test dataset for 6-class output, for each iteration of the stratified 4-fold cross-validation.

5) Addressing Class Imbalance in DL for PdM: One of the fundamental challenges in applying DL models to imbalanced datasets is the risk of the model disproportionately favoring the majority class, often at the expense of the minority class. In the context of PdM for SPV energy systems, this issue is particularly critical, as faults and anomalies – which are scarce in healthy operations – must be accurately identified to ensure system reliability.

To mitigate class imbalance, we applied the SMOTE for radiometric data augmentation,

resampling, and enhancing the minority classes. By synthetically generating new instances of radiometric faulty samples, SMOTE improves the model's ability to recognize and classify anomalies effectively.

- *Model Performance and Cross-Validation Insights:* The stratified 4-fold cross-validation results confirm that the resampled radiometric dataset led to significant improvements in fault classification:
 - Each individual model achieved an exceptional accuracy of $99.98 \pm 0.03\%$ (mean \pm standard deviation) across validation sets as portrayed in Figure 4.16.
 - The ensemble model obtained by aggregating the outputs from cross-validation models, further improved performance, achieving 100% accuracy on the final hold-out test set as shown in Figure 4.17.
- *Error Analysis and Learning Curve Trends:* While the overall classification performance was outstanding, Model-2 misclassified two samples from the majority (healthy) class in the test dataset.

Despite these misclassifications, the learning curves exhibited stable trends Figure 4.18, indicating that the training process converged efficiently and the model maintained robust generalization ability across different cross-validation iterations.

4.1 DEEP LEARNING FOR IRT-ANOMALY DIAGNOSIS

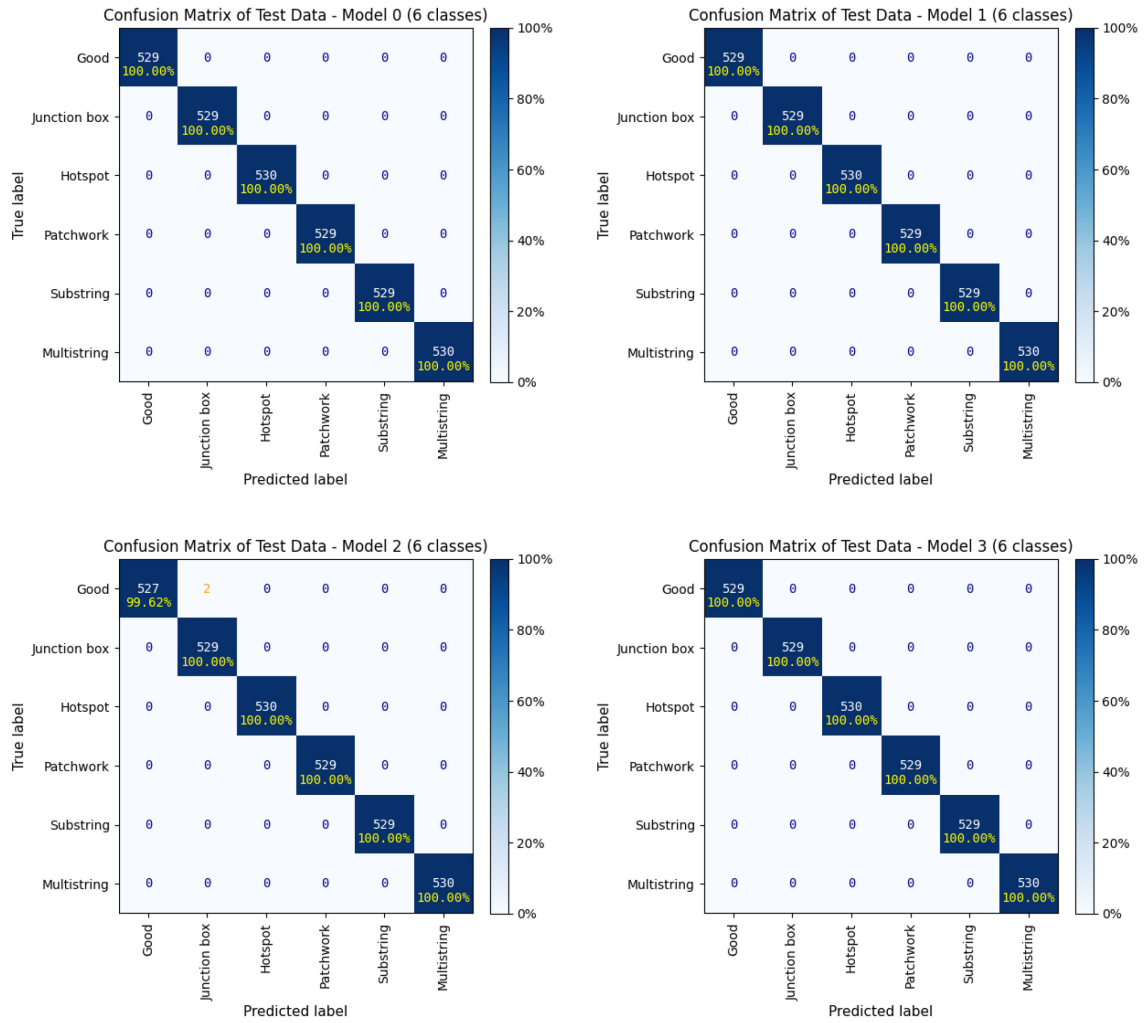


Figure 4.16 Confusion matrices of SMOTE test data for each iteration of the stratified 4-fold cross-validation (6-class).

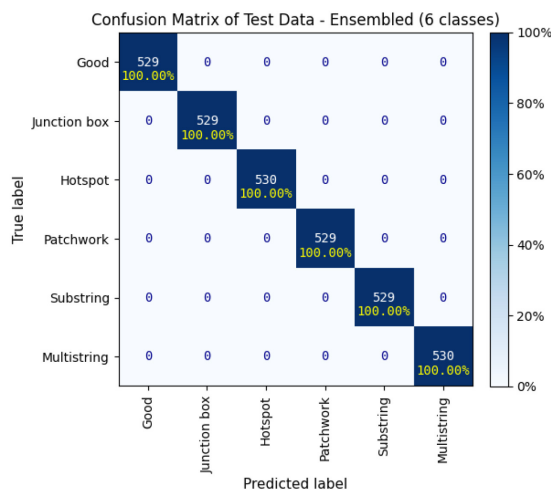


Figure 4.17 Confusion matrix of SMOTE test data—ensemble model (6-class).

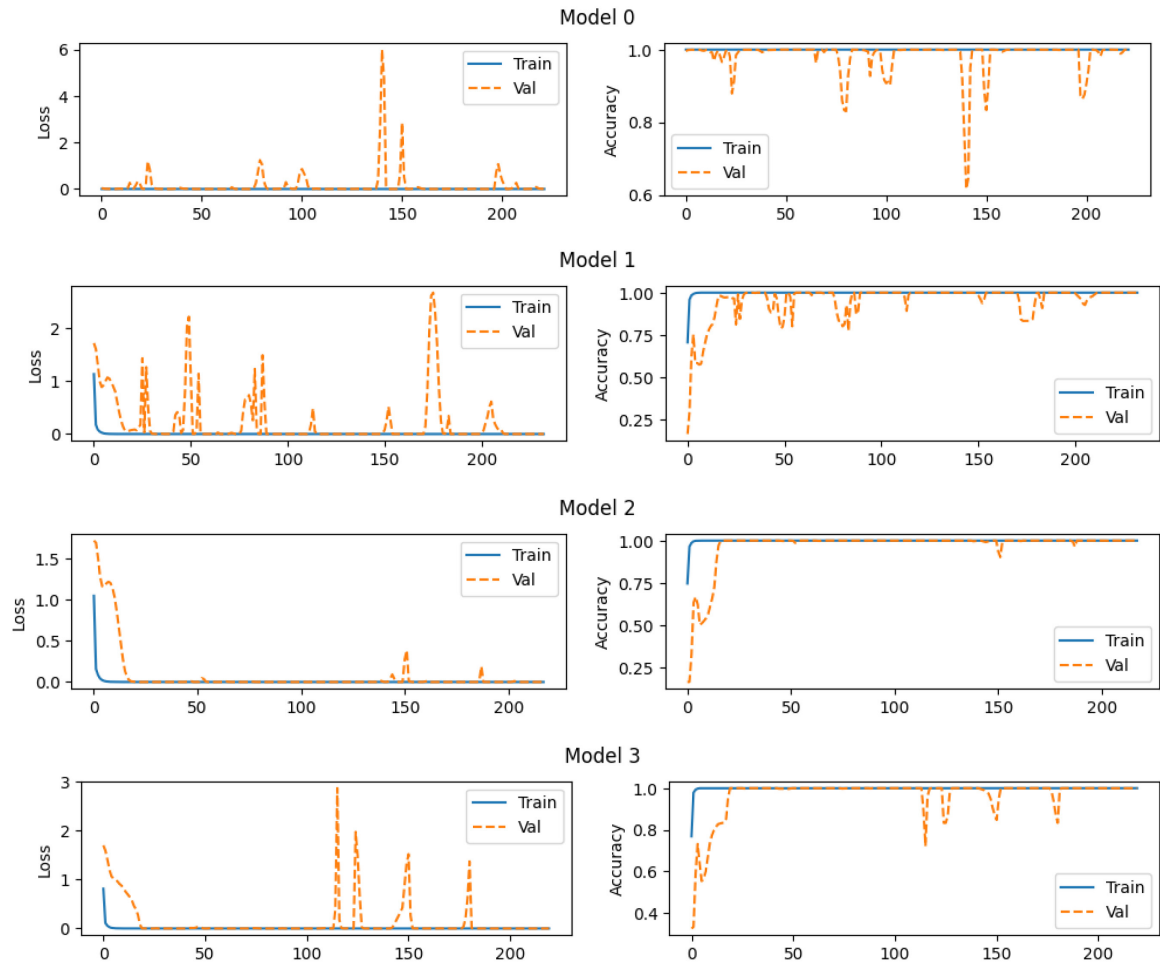


Figure 4.18. Categorical loss and accuracy of the training and validation of SMOTE test dataset for a 6-class output, for each iteration of the stratified 4-fold cross-validation.

4.1.6 Limitations

In real-world SPV energy systems, operational conditions are predominantly healthy, ensuring a continuous and optimal power output. As a result, thermal fault occurrences are rare, leading to an inherent scarcity of faulty instances in available radiometric thermal imaging dataset. This imbalance poses a significant challenge for training DL models, as they require sufficient labeled faulty samples to accurately classify and diagnose defects.

The limited availability of labeled faulty data presents a bottleneck for developing robust multiclass classification models. To address this issue, we focused on designing a customized, lightweight CNN specifically tailored for:

- Single-channel radiometric thermal dataset, representing temperature intensity matrices.
- Highly imbalanced class distributions, ensuring efficient diagnosis of rare faulty instances.

4.1.7 Conclusion

These results validate the effectiveness of our custom (ensembled) CNN-based diagnostic framework for automated SPV panel radiometric thermal anomaly diagnosis. By incorporating XAI methodologies, the study enhances model explainability and interpretability, allowing us to rationalize classification decisions and understand the learning feature structures. While the model achieves near-perfect accuracy, future work should focus on addressing class-imbalance through more faulty data acquisition and addition of other variety of labeled fault data classes, to ensure greater generalizability in real-world applications.

4.2 Deep Transfer Learning for IRT-Anomaly Diagnosis

This study explored the effectiveness of pre-trained DL models, applying a deep transfer learning (DTL) approach to conduct multiclass classification (or diagnosis) of radiometric IRT anomalies in SPV panels. The methodology aimed to enhance the model's ability to generalize despite the presence of severely imbalanced class distributions in real-world datasets.

To validate the proposed approach, we employed a highly imbalanced 6-class thermographic radiometric dataset, where each SPV panel's thermal profile was represented as floating-point temperature intensity values (in degrees Celsius). This dataset provided a challenging yet realistic scenario for assessing the effectiveness of pre-trained DL models in radiometric fault classification and thermal anomaly detection.

4.2.1 Experimental Workflow

The experimental workflow (Figure 4.19) for the proposed DTL-based radiometric thermal fault classification in SPV power systems comprised four primary steps:

- 1) Input Preparation: Constructing the radiometric IRT-SPV dataset, ensuring proper preprocessing to feed for DTL model training.
- 2) Model Selection: Identifying ten pre-trained DTL models as tabulated in Table 4.1 for radiometric thermal fault classification (temperature intensity matrix diagnosis) and compare them with our custom CNN (ensembled) model.
- 3) Fine-tuning: Modifying the DTL models by incorporating additional layers to enhance classification accuracy.

- 4) Performance Evaluation: Assessing classification metrics to validate model effectiveness.

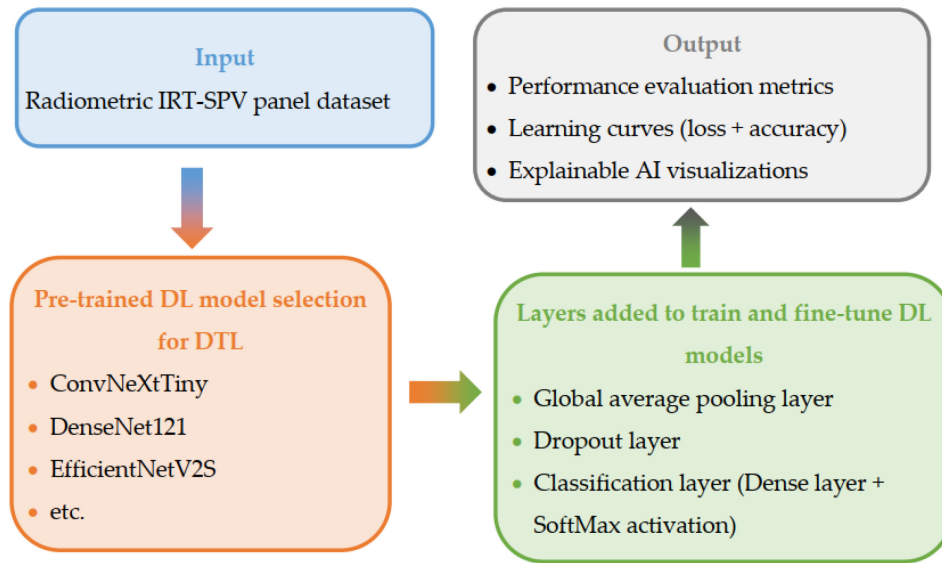


Figure 4.19. Experimental workflow (simplified block diagram)

Table 4.1 List of DL Models

Model	Year	Size (MB)	Parameters	Depth
ConvNeXtTiny [79]	2022	109.42	28.6 M	-
DenseNet121 [80]	2017	33	8.1 M	242
EfficientNetV2S [81]	2021	88	21.6 M	-
InceptionV3 [82]	2016	92	23.9 M	189
InceptionResNetV2 [83]	2017	215	55.9 M	449
MobileNetV2 [84]	2018	14	3.5 M	105
NASNetMobile [85]	2018	23	5.3 M	389
ResNet50V2 [86]	2016	98	25.6 M	103
VGG16 [87]	2015	528	138.4 M	16
Xception [88]	2017	88	22.9 M	81
Custom CNN [89]	2024	1.80	471,302	5
Ensembled Model [89]	2024	7.19	1,885,208	5

- a) **Challenges in Handling Highly Imbalanced Dataset:** The input dataset used in this study followed a 60%-20%-20% split for training, validation, and testing respectively. However, given the extreme class imbalance, only a single data instance was available for testing in each of the five minority classes, making multiclass classification a highly complex task. This scenario is akin to “finding a needle in a haystack”, demanding robust pre-trained DL models with advanced generalization capabilities.

b) Deep Transfer Learning Architecture and Hyperparameter Tuning:

- Each pre-trained DL model was carefully selected and customized for DTL by incorporating: 1) Global Average Pooling / flatten layer for dimensionality reduction, 2) Dropout layer (0.25) to mitigate overfitting, 3) Fully Connected Dense Layer (SoftMax Activation) for 6-class radiometric thermal fault diagnosis.
- To optimize training performance, the following hyperparameters were applied: Batch size 512, Early stopping: patience of 200 epochs, Learning rate 0.0001, Loss function: categorical cross-entropy, Class weighting strategy: addressing severe class imbalance

4.2.2 Results and Discussion

This section presents the experimental findings of the selected DTL models and the custom CNN (ensembled) model, focusing on their classification performance and training-validation behavior through learning curves. To assess model effectiveness, four key performance evaluation metrics were employed:

- Accuracy: Measures the overall correctness of predictions.
- Precision: Evaluates the proportion of correctly identified positive cases.
- Recall (Sensitivity): Determines the ability to detect actual positive cases.
- F1 Score: Balances precision and recall for a holistic performance assessment.

These metrics provide crucial insights into the models' ability to handle true positives, true negatives, false positives, and false negatives, ensuring robust and reliable multiclass classification of SPV panel radiometric thermal anomalies.

- 1. Performance of Pre-trained DTL models vs. Custom CNN:** The results reveal that among the selected pre-trained DL models (Table 4.2), the InceptionV3, MobileNetV2, and VGG16 architectures demonstrated an impressive accuracy of 0.998 in classifying radiometric thermal anomalies (temperature intensity matrix diagnosis). However, the custom CNN (ensembled) model achieved a perfect classification accuracy of 1.000, highlighting its ability to outperform standard DTL models in this specific dataset.

Table 4.2 DL Models Performance

Model	Accuracy	Precision	Recall	F1 Score	Training Time (sec)
ConvNeXtTiny [79]	0.994	0.989	0.994	0.992	1068
DenseNet121 [80]	0.994	0.993	0.994	0.993	742
EfficientNetV2S [81]	0.996	0.993	0.996	0.995	2894
InceptionV3 [82]	0.998	0.996	0.998	0.997	326
InceptionResNetV2 [83]	0.994	0.993	0.994	0.993	537
MobileNetV2 [84]	0.998	0.999	0.998	0.998	1095
NASNetMobile [85]	0.996	0.996	0.996	0.996	1907
ResNet50V2 [86]	0.996	0.993	0.996	0.995	511
VGG16 [87]	0.998	0.996	0.998	0.997	2529
Xception [88]	0.996	0.993	0.996	0.995	1733
Custom CNN (ensembled) [89]	1.000	1.000	1.000	1.000	204

2. Computational Efficiency and Training Time Analysis: Beyond classification accuracy, computational efficiency was also analyzed:

- InceptionV3 exhibited the least training time (326 seconds) among pre-trained DTL models.
- The custom CNN (ensembled) model demonstrated an even faster training time (204 seconds), indicating a highly optimized and lightweight network architecture.

This finding underscores the potential of custom CNN architectures in achieving both high classification accuracy and computational efficiency, making them viable for SPV arrays diagnostic monitoring applications.

3. Explainability through XAI Techniques: To enhance model perceptive explainability, Guided SmoothGrad, an advanced explainability method that merges SmoothGrad with Guided Backpropagation (GBP) was applied. The saliency visualizations revealed that:

- Pre-trained DL models effectively highlighted high-temperature anomaly regions in *patchwork* and *multistring* fault cases as shown in Figure 4.20.
- Feature space embeddings, visualized using UMAP (Hellinger Metric) demonstrated distinct clustering as shown in Figure 4.21, reinforcing the models' ability to differentiate among healthy and faulty radiometric SPV panels.

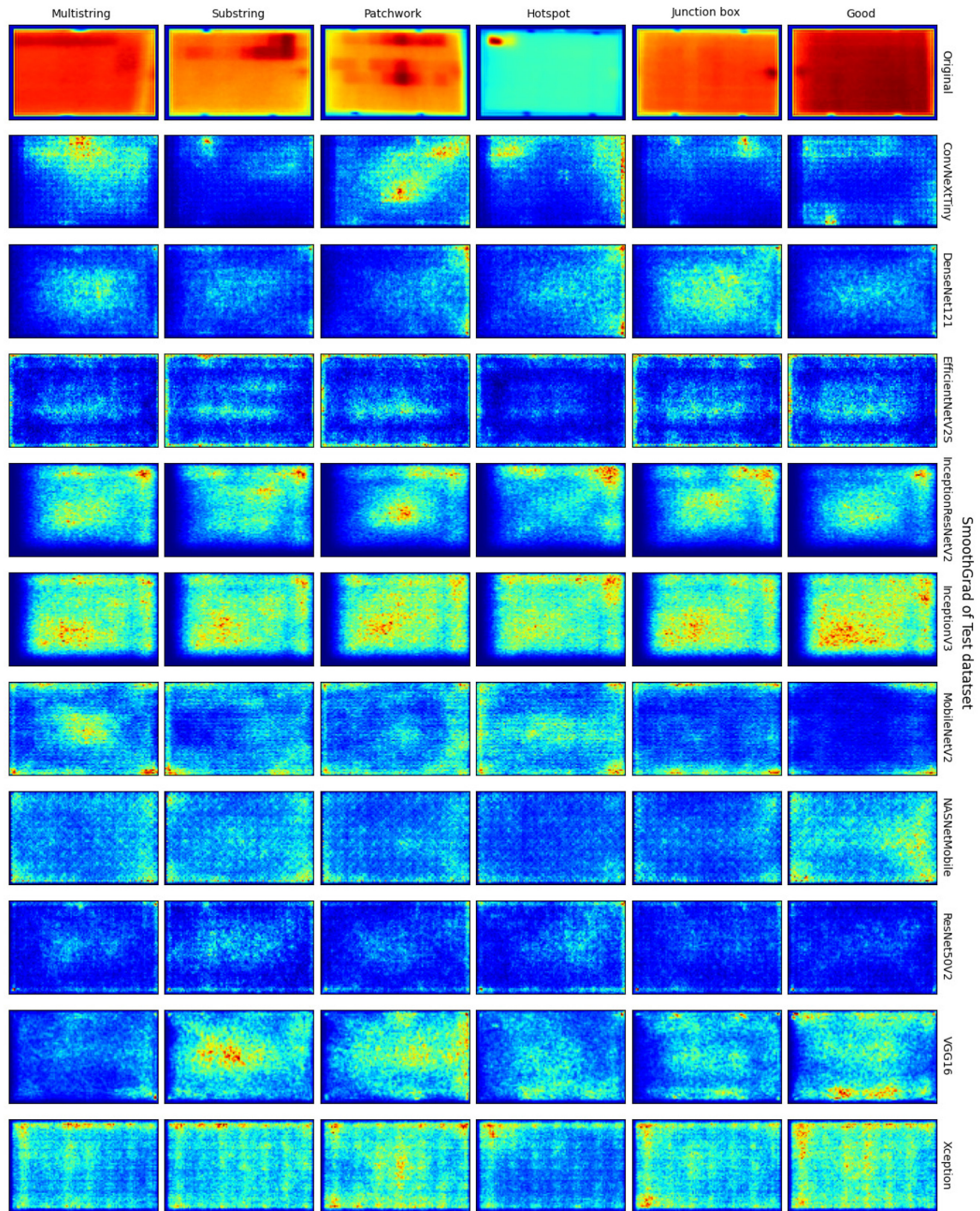


Figure 4.20 Explainable deep learning – Guided SmoothGrad visualizations from the last convolutional layer for 6-class classification (jet colormap as pseudo-color applied for visual depiction). Red corresponds to higher activation values, whereas blue to lower ones.

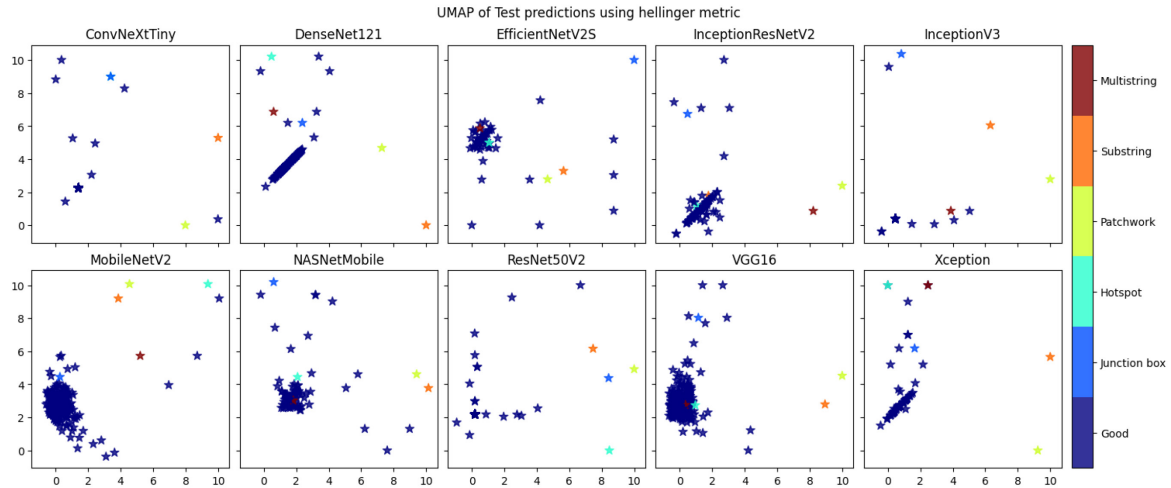


Figure 4.21 Embedding with UMAP (feature clustering) of test dataset predictions (Hellinger Metric). Each data point is depicted as a star-shaped point. Kindly note that some points are overlapped.

4.2.3 Limitations

This case study presents two primary limitations that affect the training and validation processes of DTL models for multiclass radiometric thermal anomaly classification (temperature intensity matrix diagnosis) in SPV energy systems.

- 1) **Data Scarcity and Imbalanced Class Distribution:** One of the fundamental challenges encountered in this study was the limited number of available instances in minority classes. Since SPV panels are typically designed to operate under optimal conditions, occurrences of thermal anomalies are scarce, leading to an intrinsically imbalanced dataset.
 - DTL models require a substantial amount of training data across all classes to generalize effectively. However, the lack of sufficient faulty instances resulted in suboptimal classification accuracy for thermal anomalies.
 - Conversely, the custom CNN (ensembled) model demonstrated superior adaptability to the radiometric thermal dataset's imbalance, thereby outperforming pre-trained DL models.
 - This limitation emphasizes the need for larger, more diverse datasets to enhance model generalizability and mitigate bias toward majority class predictions.
- 2) **Absence of Synthetic Data Augmentation (Resampling Techniques):** Unlike previous studies that leverage synthetic over-sampling (e.g. SMOTE, GAN-based augmentation), this research intentionally omitted these techniques to examine the inherent learning capabilities of pre-trained DL models on original radiometric thermal dataset.

- By working with floating-point temperature intensity values, the study aimed to assess how well DTL architectures handle real-world, unaltered radiometric thermographic data.
- While this approach provides valuable insights, future studies could explore hybrid methodologies combining resampling techniques with DTL models to further improve diagnostic accuracy and model robustness.

4.2.4 Conclusion

The findings of this study reinforce the efficacy of DTL models in performing multiclass classification of 6-class radiometric thermal anomalies in SPV panels. Given that pre-trained DL architectures are originally trained on standardized image datasets, their adaptation to radiometric thermal datasets (floating-point temperature intensity matrix) presents both opportunities and challenges.

1) Performance Benchmarking of DTL Models vs. Custom CNN: The study highlights key observations regarding model performance:

- Among the selected DTL models, InceptionV3, MobileNetV2, and VGG16 exhibited high classification accuracy (0.998), demonstrating their robustness in 6-class radiometric thermal fault diagnosis.
- InceptionV3 proved to be the fastest DTL model (326 seconds training time), indicating its computational efficiency in handling the radiometric thermographic dataset.
- However, the custom CNN (ensembled) model outperformed all pre-trained DL models, achieving:
 - Perfect accuracy (1.000) in multiclass classification.
 - The shortest training time (204 seconds), underscoring its efficiency in learning from the imbalanced 6-class dataset.

These results suggest that while DTL models offer substantial accuracy, a tailored CNN architecture optimized for the dataset can provide even greater diagnostic precision and computational efficiency.

2) XAI Techniques – Enhancing Interpretability and Trust in DL models: A major contribution of this research is the incorporation of XAI techniques – such as Guided SmoothGrad and UMAP – to provide a rationale for model decision-making.

- Guided SmoothGrad (a fusion of SmoothGrad and Guided Backpropagation) effectively visualized high-temperature feature regions, aiding in model explainability.
- UMAP (Uniform Manifold Approximation and Projection) revealed clear clustering of predicted features, offering deeper insights into how different anomaly types are distinguished.

By integrating XAI techniques, this study not only enhances the transparency of SPV arrays' radiometric thermal anomaly diagnosis models but also fosters greater cognitive understanding, bridging the gap between AI-driven predictions and human decision-making in PdM frameworks.

3) Strategic Significance for SPV Energy Systems: The implications of this research extend beyond AI-based thermographic vision inspection to real-world SPV energy system management:

- **Early Fault Diagnosis:** The proposed AI-driven framework enables proactive identification of potential thermal faults, ensuring timely maintenance and reducing energy output degradation.
- **Prolonged Asset Lifespan:** By diagnosing latent defects, the system extends economic life of SPV panels, mitigating capital-intensive losses.
- **Climate Impact and Humanitarian Efforts:** In the wake of global warming and the urgent transition to renewable energy, such AI-powered PdM solutions can:
 - Safeguard investments in large-scale and utility-scale SPV power plants.
 - Support humanitarian initiatives by ensuring sustainable solar power for energy-insecure regions.

Chapter V

Conclusions

The research works carried out for this thesis were devoted to the conceptualization, implementation, and evaluation of novel systems of intelligent diagnostic monitoring that can support the Operations and Maintenance (O&M) team in routine practice, offering an objective methodology to identify thermographic anomalies in SPV energy systems.

In the proposed diagnostic pipelines, different objectives were considered, including enhancement of the radiometric thermographic fault detection performance, usability of the methodology, utilization of DL models, and feasibility of the integration into the PdM practice. Moreover, the investigation of principles behind the developed systems, by exploiting perceptive explainability and mathematical interpretability techniques, has been carried out to ensure trust and transparency of the models' decision-making process.

Chapter 2 delves into the analysis of AI techniques in SPV diagnostics, methodologies employed in this thesis, in particular, CNNs for image classification, performance evaluation metrics, XAI methods, standardized guidelines, and technical prerequisites for intelligent aerial IRT inspection of SPV Power Systems.

Chapter 3 elaborates a theoretical framework for CNN comprising a CNN architecture, layers, and training process.

Chapter 4 features two distinctive case studies (custom CNN and DTL) for monitoring diagnosis within PdM framework, representing a comprehensive workflow, based on end-to-end pipelines designed and implemented for radiometric anomaly diagnosis. Initially, a customized, severely imbalanced 6-class radiometric IRT dataset (floating-point temperature intensity matrix in degrees Celsius) was prepared to facilitate the precise thermal fault diagnosis of SPV panels. Then, a lightweight custom CNN (ensembled) model was developed to classify radiometric thermal anomalies with near-perfect accuracy (1.000). Further, the performance evaluation of the developed custom CNN (ensembled) model is compared with the selected ten pre-trained DL models (deep transfer learning - DTL), and the classification predictions are visualized with XAI techniques.

In conclusion, this thesis showcases the significance of intelligent aerial thermographic inspection in PdM framework. In the domain of PdM, the integration of intelligent fault diagnosis presents a transformative approach, offering a cost-effective, time-efficient, and automated solution for monitoring large-scale SPV energy systems. By leveraging remote aerial IRT imaging with DL, ensemble learning, data augmentation, and XAI-driven explanations, the proposed framework advances the state-of-the-art maintenance operations in the future, which can shift from reactive to proactive strategies, thereby enhancing reliability and sustainability in SPV power generation.

My Publications

- 1) **Usamah Rashid Qureshi**, Aiman Rashid, Nicola Altini, Vitoantonio Bevilacqua and Massimo La Scala. Explainable Intelligent Inspection of Solar Photovoltaic Systems with Deep Transfer Learning: Considering Warmer Weather Effects using Aerial Radiometric Infrared Thermography. *Electronics* **2025**, 14, 755.
<https://doi.org/10.3390/electronics14040755>
- 2) **Usamah Rashid Qureshi**, Aiman Rashid, Nicola Altini, Vitoantonio Bevilacqua and Massimo La Scala. Radiometric Infrared Thermography of Solar Photovoltaic Systems: An Explainable Predictive Maintenance Approach for Remote Aerial Diagnostic Monitoring. *Smart Cities* **2024**, 7, 1261–1288.
<https://doi.org/10.3390/smartcities7030053>

References

- [1] The British Standards Institution (BSI), “BS EN 13306:2017 Maintenance - maintenance terminology,” BSI Standards Limited, London, United Kingdom, Jan. 2018. Accessed: Apr. 24, 2023. [Online]. Available: <https://knowledge.bsigroup.com/products/maintenance-maintenance-terminology?version=standard>
- [2] K. Keisang, T. Bader, and R. Samikannu, “Review of Operation and Maintenance Methodologies for Solar Photovoltaic Microgrids,” *Front Energy Res*, vol. 9, p. 730230, Nov. 2021, doi: 10.3389/FENRG.2021.730230/BIBTEX.
- [3] J. Yu and Y. Zhang, “Challenges and opportunities of deep learning-based process fault detection and diagnosis: a review,” *Neural Computing and Applications 2022 35:1*, vol. 35, no. 1, pp. 211–252, Nov. 2022, doi: 10.1007/S00521-022-08017-3.
- [4] A. Abubakar, C. F. M. Almeida, and M. Gemignani, “Review of Artificial Intelligence-Based Failure Detection and Diagnosis Methods for Solar Photovoltaic Systems,” *Machines 2021, Vol. 9, Page 328*, vol. 9, no. 12, p. 328, Dec. 2021, doi: 10.3390/MACHINES9120328.
- [5] SolarPower Europe, “Global Market Outlook for Solar Power 2024-2028,” Jun. 2024. Accessed: Jul. 10, 2024. [Online]. Available: https://api.solarpowereurope.org/uploads/Global_Market_Outlook_for_Solar_Power_2024_a083b6dcd5.pdf
- [6] EurObserv'ER, “Photovoltaic Barometer,” Apr. 2024. Accessed: Jul. 09, 2024. [Online]. Available: <https://www.eurobserv-er.org/photovoltaic-barometer-2024/>
- [7] IRENA, “Renewable Capacity Statistics 2024,” International Renewable Energy Agency, Abu Dhabi, 2024. Accessed: Jul. 09, 2024. [Online]. Available: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Mar/IRENA_RE_Capacity_Statistics_2024.pdf?rev=a587503ac9a2435c8d13e40081d2ec34

- [8] S. Brown and D. Jones, “European Electricity Review 2024,” Ember, London Fields, 2024. Accessed: Jul. 23, 2024. [Online]. Available: <https://ember-climate.org/insights/research/european-electricity-review-2024/>
- [9] IEA, “World Energy Investment 2024,” 2024. Accessed: Jul. 17, 2024. [Online]. Available: <https://www.iea.org/reports/world-energy-investment-2024/overview-and-key-findings>
- [10] IRENA, “Renewable Power Generation Costs in 2022,” Abu Dhabi, 2023. Accessed: Jul. 11, 2024. [Online]. Available: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2023/Aug/IRENA_Renewable_power_generation_costs_in_2022.pdf
- [11] WMO, “More Extreme Heat Demands Coordinated Action,” Jun. 2024. Accessed: Jul. 15, 2024. [Online]. Available: <https://reliefweb.int/report/world/more-extreme-heat-demands-coordinated-action>
- [12] IFRC and UN OCHA, “Extreme Heat: Preparing for the Heatwaves of the Future,” Oct. 2022. Accessed: Jul. 16, 2024. [Online]. Available: <https://www.ifrc.org/document/extreme-heat-preparing-heat-waves-future>
- [13] WMO, “State of the Global Climate 2023 (WMO-No. 1347).” Accessed: Jul. 15, 2024. [Online]. Available: <https://library.wmo.int/records/item/68835-state-of-the-global-climate-2023>
- [14] IEA, “Unlocking Smart Grid Opportunities in Emerging Markets and Developing Economies,” Paris, France, Jun. 2023. Accessed: Jul. 23, 2024. [Online]. Available: <https://www.iea.org/reports/unlocking-smart-grid-opportunities-in-emerging-markets-and-developing-economies>
- [15] IEA, “Electricity 2024: Analysis and forecast to 2026,” Paris, France, Jan. 2024. Accessed: Jul. 23, 2024. [Online]. Available: https://managenergy.ec.europa.eu/publications/electricity-2024-analysis-and-forecast-2026_en
- [16] T. A. Brás, S. G. Simoes, F. Amorim, and P. Fortes, “How much extreme weather events have affected European power generation in the past three decades?,” *Renewable and Sustainable Energy Reviews*, vol. 183, p. 113494, Sep. 2023, doi: 10.1016/j.rser.2023.113494.

-
- [17] G. R. Venkatakrishnan *et al.*, “Detection, location, and diagnosis of different faults in large solar PV system—a review,” *International Journal of Low-Carbon Technologies*, vol. 18, pp. 659–674, Feb. 2023, doi: 10.1093/IJLCT/CTAD018.
- [18] Panasonic, “530-550 Watt (Anchor by Panasonic) 144 Half-cell Mono PERC Module Datasheet,” LT, 2023. Accessed: Oct. 09, 2023. [Online]. Available: https://lssth.panasonic.com/products/energy/product_finder/img/solar002.pdf
- [19] Solar Electric UK, “Glass/Glass Model No STKM-72-330 (330Wp Monocrystalline 72 Cell Module) Datasheet,” 2023. Accessed: Nov. 09, 2023. [Online]. Available: <http://www.solarelectricuk.com/datasheets/Glass-Model-STKM-72-330-Monocrystalline-Cell-Module.pdf>
- [20] Solar Electric UK, “Glass/Glass Model No STKP-72-310 (310Wp Polycrystalline 72 cell module) Datasheet,” 2023. Accessed: Nov. 09, 2023. [Online]. Available: <http://www.solarelectricuk.com/datasheets/Glass-Model-STKP-72-310-Polycrystalline-Cell-Module.pdf>
- [21] Trina Solar, “Tallmax Framed 252 Layout Mono Half-cut (470-490 Watt, 20.8% Efficiency) Datasheet,” 2022. Accessed: Nov. 09, 2023. [Online]. Available: https://static.trinasolar.com/sites/default/files/DT-M-0007%20E%20Datasheet_TallmaxM_DE15V%28II%29_NA_EN_2022_A_web.pdf
- [22] M. Owen-Bellini *et al.*, “Advancing reliability assessments of photovoltaic modules and materials using combined-accelerated stress testing,” *Progress in Photovoltaics: Research and Applications*, vol. 29, no. 1, pp. 64–82, Jan. 2021, doi: 10.1002/pip.3342.
- [23] A. Omazic *et al.*, “Relation between degradation of polymeric components in crystalline silicon PV module and climatic conditions: A literature review,” *Solar Energy Materials and Solar Cells*, vol. 192, pp. 123–133, Apr. 2019, doi: 10.1016/J.SOLMAT.2018.12.027.
- [24] P. Nivelles, J. A. Tsanakas, J. Poortmans, and M. Daenen, “Stress and strain within photovoltaic modules using the finite element method: A critical review,” *Renewable and Sustainable Energy Reviews*, vol. 145, p. 111022, Jul. 2021, doi: 10.1016/J.RSER.2021.111022.

- [25] Y. Du, L. Wang, and W. Tao, "Modeling, imaging and resistance analysis for crystalline silicon photovoltaic modules failure on thermal cycle test," *Eng Fail Anal*, vol. 118, p. 104818, Dec. 2020, doi: 10.1016/J.ENGFAILANAL.2020.104818.
- [26] J. Kim, M. Rabelo, S. P. Padi, H. Yousuf, E. C. Cho, and J. Yi, "A review of the degradation of photovoltaic modules for life expectancy," *Energies (Basel)*, vol. 14, no. 14, Jul. 2021, doi: 10.3390/en14144278.
- [27] G. C. Eder *et al.*, "Climate specific accelerated ageing tests and evaluation of ageing induced electrical, physical, and chemical changes," *Progress in Photovoltaics: Research and Applications*, vol. 27, no. 11, pp. 934–949, Nov. 2019, doi: 10.1002/pip.3090.
- [28] K. Hasan, S. B. Yousuf, M. S. H. K. Tushar, B. K. Das, P. Das, and M. S. Islam, "Effects of different environmental and operational factors on the PV performance: A comprehensive review," *Energy Sci Eng*, vol. 10, no. 2, pp. 656–675, Feb. 2022, doi: 10.1002/ESE3.1043.
- [29] M. Santhakumari and N. Sagar, "A review of the environmental factors degrading the performance of silicon wafer-based photovoltaic modules: Failure detection methods and essential mitigation techniques," Aug. 01, 2019, *Elsevier Ltd.* doi: 10.1016/j.rser.2019.04.024.
- [30] M. Aghaei *et al.*, "Review of degradation and failure phenomena in photovoltaic modules," *Renewable and Sustainable Energy Reviews*, vol. 159, p. 112160, May 2022, doi: 10.1016/J.RSER.2022.112160.
- [31] M. Dhimish and A. Alrashidi, "Photovoltaic Degradation Rate Affected by Different Weather Conditions: A Case Study Based on PV Systems in the UK and Australia," *Electronics 2020, Vol. 9, Page 650*, vol. 9, no. 4, p. 650, Apr. 2020, doi: 10.3390/ELECTRONICS9040650.
- [32] T. Rahman *et al.*, "Investigation of Degradation of Solar Photovoltaics: A Review of Aging Factors, Impacts, and Future Directions toward Sustainable Energy Management," *Energies 2023, Vol. 16, Page 3706*, vol. 16, no. 9, p. 3706, Apr. 2023, doi: 10.3390/EN16093706.

-
- [33] C. Sun, Y. Zou, C. Qin, B. Zhang, and X. Wu, "Temperature effect of photovoltaic cells: a review," *Advanced Composites and Hybrid Materials* 2022 5:4, vol. 5, no. 4, pp. 2675–2699, Aug. 2022, doi: 10.1007/S42114-022-00533-Z.
- [34] IEC, "IEC TS 62446-3:2017 Photovoltaic (PV) systems - Requirements for testing, documentation and maintenance - Part 3: Photovoltaic modules and plants - Outdoor infrared thermography," Geneva, Switzerland, Jun. 2017. Accessed: Feb. 03, 2024. [Online]. Available: <https://webstore.iec.ch/en/publication/28628>
- [35] H. F. Mateo Romero *et al.*, "Applications of Artificial Intelligence to Photovoltaic Systems: A Review," *Applied Sciences* 2022, Vol. 12, Page 10056, vol. 12, no. 19, p. 10056, Oct. 2022, doi: 10.3390/APP121910056.
- [36] A. Thakfan and Y. Bin Salamah, "Artificial-Intelligence-Based Detection of Defects and Faults in Photovoltaic Systems: A Survey," *Energies* 2024, Vol. 17, Page 4807, vol. 17, no. 19, p. 4807, Sep. 2024, doi: 10.3390/EN17194807.
- [37] M. Islam, M. R. Rashel, M. T. Ahmed, A. K. M. K. Islam, and M. Tlemçani, "Artificial Intelligence in Photovoltaic Fault Identification and Diagnosis: A Systematic Review," *Energies* 2023, Vol. 16, Page 7417, vol. 16, no. 21, p. 7417, Nov. 2023, doi: 10.3390/EN16217417.
- [38] S. Rahman *et al.*, "Analysis of Power Grid Voltage Stability with High Penetration of Solar PV Systems," *IEEE Trans Ind Appl*, vol. 57, no. 3, pp. 2245–2257, May 2021, doi: 10.1109/TIA.2021.3066326.
- [39] F. Martinez-Gil, C. Sansom, A. Fernández-García, A. Alcayde-García, and F. Manzano-Agugliaro, "Maintenance techniques to increase solar energy production: A review," *Energy Nexus*, vol. 17, p. 100384, Mar. 2025, doi: 10.1016/J.NEXUS.2025.100384.
- [40] Y. Ledmaoui, A. El Maghraoui, M. El Aroussi, and R. Saadane, "Review of Recent Advances in Predictive Maintenance and Cybersecurity for Solar Plants," *Sensors* 2025, Vol. 25, Page 206, vol. 25, no. 1, p. 206, Jan. 2025, doi: 10.3390/S25010206.
- [41] A. A. Al-Katheri, E. A. Al-Ammar, M. A. Alotaibi, W. Ko, S. Park, and H. J. Choi, "Application of Artificial Intelligence in PV Fault Detection," *Sustainability* 2022, Vol. 14, Page 13815, vol. 14, no. 21, p. 13815, Oct. 2022, doi: 10.3390/SU142113815.

- [42] K. Ukoba, K. O. Olatunji, E. Adeoye, T. C. Jen, and D. M. Madyira, "Optimizing renewable energy systems through artificial intelligence: Review and future prospects," *Energy and Environment*, vol. 35, no. 7, pp. 3833–3879, Nov. 2024, doi: 10.1177/0958305X241256293/ASSET/8D92F57B-7887-4AB6-9318-235015A49E74/ASSETS/IMAGES/LARGE/10.1177_0958305X241256293-FIG7.JPG.
- [43] J. N. C. ; Domathoti, B. ; Santibanez Gonzalez, A. Di Vaio, J. N. Chandra Sekhar, B. Domathoti, and E. D. R. Santibanez Gonzalez, "Prediction of Battery Remaining Useful Life Using Machine Learning Algorithms," *Sustainability 2023, Vol. 15, Page 15283*, vol. 15, no. 21, p. 15283, Oct. 2023, doi: 10.3390/SU152115283.
- [44] Z. Mustaffa and M. H. Sulaiman, "Battery remaining useful life estimation based on particle swarm optimization-neural network," *Cleaner Energy Systems*, vol. 9, p. 100151, Dec. 2024, doi: 10.1016/J.CLES.2024.100151.
- [45] P. ; Sharma, B. J. A. Bora, P. Sharma, and B. J. Bora, "A Review of Modern Machine Learning Techniques in the Prediction of Remaining Useful Life of Lithium-Ion Batteries," *Batteries 2023, Vol. 9, Page 13*, vol. 9, no. 1, p. 13, Dec. 2022, doi: 10.3390/BATTERIES9010013.
- [46] A. Kumar, A. K. Dubey, I. Segovia Ramírez, A. Muñoz del Río, and F. P. García Márquez, "Artificial Intelligence Techniques for the Photovoltaic System: A Systematic Review and Analysis for Evaluation and Benchmarking," *Archives of Computational Methods in Engineering*, vol. 31, no. 8, pp. 4429–4453, Dec. 2024, doi: 10.1007/S11831-024-10125-3/FIGURES/6.
- [47] J. Hu *et al.*, "A Comprehensive Review of Artificial Intelligence Applications in the Photovoltaic Systems," *CAAI Artificial Intelligence Research*, vol. 3, p. 9150031, Dec. 2024, doi: 10.26599/AIR.2024.9150031.
- [48] M. Bacha, A. Terki, and M. Bacha, "Comparative study of real-time photovoltaic fault diagnosis using artificial intelligence: Fuzzy logic and neural network approaches," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 46, no. 1, pp. 13536–13560, Dec. 2024, doi: 10.1080/15567036.2024.2406424.

-
- [49] F. Harrou, Y. Sun, B. Taghezouit, and A. Dairi, “Artificial Intelligence Techniques for Solar Irradiance and PV Modeling and Forecasting,” *Energies* 2023, *Vol. 16*, Page 6731, vol. 16, no. 18, p. 6731, Sep. 2023, doi: 10.3390/EN16186731.
- [50] Z. K. Fan, A. Setianingrum, K. L. Lian, and S. Suwarno, “A Hybrid Approach for Photovoltaic Maximum Power Tracking under Partial Shading Using Honey Badger and Genetic Algorithms,” *Energies* 2024, *Vol. 17*, Page 3935, vol. 17, no. 16, p. 3935, Aug. 2024, doi: 10.3390/EN17163935.
- [51] K. H. Chao and M. N. Rizal, “A Hybrid MPPT Controller Based on the Genetic Algorithm and Ant Colony Optimization for Photovoltaic Systems under Partially Shaded Conditions,” *Energies* 2021, *Vol. 14*, Page 2902, vol. 14, no. 10, p. 2902, May 2021, doi: 10.3390/EN14102902.
- [52] A. K. Vidal De Oliveira, M. Aghaei, and R. Rütther, “Automatic Inspection of Photovoltaic Power Plants Using Aerial Infrared Thermography: A Review,” 2022, doi: 10.3390/en15062055.
- [53] L. Cardinale-Villalobos, C. Meza, A. Méndez-Porras, and L. D. Murillo-Soto, “Quantitative Comparison of Infrared Thermography, Visual Inspection, and Electrical Analysis Techniques on Photovoltaic Modules: A Case Study,” *Energies* 2022, *Vol. 15*, Page 1841, vol. 15, no. 5, p. 1841, Mar. 2022, doi: 10.3390/EN15051841.
- [54] I. Polymeropoulos, S. Bezyrgiannidis, E. Vrochidou, and G. A. Papakostas, “Enhancing Solar Plant Efficiency: A Review of Vision-Based Monitoring and Fault Detection Techniques,” *Technologies* 2024, *Vol. 12*, Page 175, vol. 12, no. 10, p. 175, Sep. 2024, doi: 10.3390/TECHNOLOGIES12100175.
- [55] F. Cannella and G. Di Francia, “PV Large Utility Visual Inspection via Unmanned Vehicles,” *MESA 2024 - 20th International Conference on Mechatronic, Embedded Systems and Applications, Proceedings*, 2024, doi: 10.1109/MESA61532.2024.10704867.
- [56] S. R. Joshua, S. Park, and K. Kwon, “EfficientNetB0 for AI-Powered Solar Panel Maintenance: Advanced Fault Detection,” *International Conference on ICT Convergence*, pp. 301–306, 2024, doi: 10.1109/ICTC62082.2024.10827751.

- [57] O. K. Pal, M. S. H. Shovon, M. F. Mridha, and J. Shin, “A Comprehensive Review of AI-enabled Unmanned Aerial Vehicle: Trends, Vision , and Challenges,” Oct. 2023, Accessed: Apr. 03, 2025. [Online]. Available: <https://arxiv.org/abs/2310.16360v1>
- [58] O. K. Pal, M. S. H. Shovon, M. F. Mridha, and J. Shin, “In-depth review of AI-enabled unmanned aerial vehicles: trends, vision, and challenges,” *Discover Artificial Intelligence 2024 4:1*, vol. 4, no. 1, pp. 1–24, Dec. 2024, doi: 10.1007/S44163-024-00209-1.
- [59] K. Telli *et al.*, “A Comprehensive Review of Recent Research Trends on Unmanned Aerial Vehicles (UAVs),” *Systems 2023, Vol. 11, Page 400*, vol. 11, no. 8, p. 400, Aug. 2023, doi: 10.3390/SYSTEMS11080400.
- [60] PADCON Services, “The Use of AI Drones for Large-Scale Solar Inspection - Padcon.” Accessed: Apr. 03, 2025. [Online]. Available: <https://www.padcon.com/2024/05/23/the-use-of-ai-drones-for-large-scale-solar-inspection/>
- [61] MapperX, “Thermal Drone - AI Software for Solar Power Plant Inspection and Examination - MapperX.” Accessed: Apr. 03, 2025. [Online]. Available: <https://mapperx.com/en/thermal-drone-2/>
- [62] Techstack, “Optimizing Solar Panel Efficiency and Placement Using AI-Driven Project Management Tools - Techstack.” Accessed: Apr. 03, 2025. [Online]. Available: <https://tech-stack.com/blog/solar-panel-efficiency/>
- [63] eMergy Tech, “Artificial Intelligence and Photovoltaics - eMergy Tech.” Accessed: Apr. 03, 2025. [Online]. Available: <https://www.emergytech.it/en/artificial-intelligence-and-photovoltaics-what-are-the-applications/>
- [64] Mipu, “Artificial Intelligence Applied to the Photovoltaic Sector: Technologies and New Prospects - Mipu.” Accessed: Apr. 03, 2025. [Online]. Available: <https://mipu.eu/topic/lintelligenza-artificiale-applicata-al-settore-fotovoltaico/>
- [65] Sun’Agri, “Technology - Sun’Agri.” Accessed: Apr. 03, 2025. [Online]. Available: <https://sunagri.fr/en/our-solution/technology/>
- [66] SmartHelio, “From Hours to Minutes: How AI is Revolutionizing Solar Plant Diagnostics - SmartHelio.” Accessed: Apr. 03, 2025. [Online]. Available:

- <https://smarthelio.com/from-hours-to-minutes-how-ai-is-revolutionizing-solar-plant-diagnostics/>
- [67] SmartHelio, “Top 5 AI Tools to Boost Solar Plant Efficiency and Reduce Costs - SmartHelio.” Accessed: Apr. 03, 2025. [Online]. Available: <https://smarthelio.com/top-5-ai-tools-to-boost-solar-plant-efficiency-and-reduce-costs/>
- [68] E. H. Sepúlveda-Oviedo, L. Travé-Massuyès, A. Subias, M. Pavlov, and C. Alonso, “Fault diagnosis of photovoltaic systems using artificial intelligence: A bibliometric approach,” *Heliyon*, vol. 9, no. 11, Nov. 2023, doi: 10.1016/J.HELIYON.2023.E21491/ASSET/A63AAAD3-24D1-47AC-9C12-A456A80464E2/MAIN.ASSETS/GR003_LRG.JPG.
- [69] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature* 2015 521:7553, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [70] Y. LeCun *et al.*, “Backpropagation Applied to Handwritten Zip Code Recognition,” *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989, doi: 10.1162/NECO.1989.1.4.541.
- [71] N. V. S and V. Sugumaran, “Fault diagnosis of visual faults in photovoltaic modules: A Review,” *Int J Green Energy*, vol. 18, no. 1, pp. 37–50, Jan. 2021, doi: 10.1080/15435075.2020.1825443.
- [72] C. Qian, J. Zhu, Y. Shen, Q. Jiang, and Q. Zhang, “Deep Transfer Learning in Mechanical Intelligent Fault Diagnosis: Application and Challenge,” *Neural Process Lett*, vol. 54, no. 3, pp. 2509–2531, Jun. 2022, doi: 10.1007/S11063-021-10719-Z/FIGURES/6.
- [73] B. Li, C. Delpha, D. Diallo, and A. Migan-Dubois, “Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: A review,” *Renewable and Sustainable Energy Reviews*, vol. 138, Mar. 2021, doi: 10.1016/J.RSER.2020.110512.
- [74] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications,” *Neurocomputing*, vol. 234, pp. 11–26, Apr. 2017, doi: 10.1016/J.NEUCOM.2016.12.038.

- [75] I. Goodfellow, Y. Bengio, and A. Courville, *Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning*, no. 1. MIT Press, 2016.
- [76] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.
- [77] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, Miami, FL, USA, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.
- [78] O. Russakovsky *et al.*, “ImageNet large scale visual recognition challenge,” *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.
- [79] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “A ConvNet for the 2020s,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA: IEEE, Jun. 2022, pp. 11966–11976. doi: 10.1109/CVPR52688.2022.01167.
- [80] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” in *30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Honolulu, HI, USA, Aug. 2016, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.
- [81] M. Tan and Q. V. Le, “EfficientNetV2: Smaller Models and Faster Training,” in *38th International Conference on Machine Learning (ICML 2021)*, Virtual, Jul. 2021, pp. 10096–10106. Accessed: Jul. 25, 2024. [Online]. Available: <http://arxiv.org/abs/2104.00298>
- [82] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.
- [83] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning,” in *31st AAAI Conference on*

-
- Artificial Intelligence (AAAI 2017)*, San Francisco, CA, USA, Feb. 2017, pp. 4278–4284. doi: 10.1609/aaai.v31i1.11231.
- [84] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.
- [85] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, “Learning Transferable Architectures for Scalable Image Recognition,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, Jun. 2018, pp. 8697–8710. doi: 10.1109/CVPR.2018.00907.
- [86] K. He, X. Zhang, S. Ren, and J. Sun, “Identity Mappings in Deep Residual Networks,” in *14th European Conference on Computer Vision (ECCV 2016 - Part IV)*, Amsterdam, The Netherlands, Oct. 2016, pp. 630–645. doi: 10.1007/978-3-319-46493-0_38.
- [87] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” in *3rd International Conference on Learning Representations (ICLR 2015)*, San Diego, CA, USA, May 2015. Accessed: Aug. 16, 2024. [Online]. Available: <https://arxiv.org/abs/1409.1556v6>
- [88] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” in *30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, Honolulu, HI, USA, Jul. 2017, pp. 1800–1807. doi: 10.1109/CVPR.2017.195.
- [89] U. R. Qureshi, A. Rashid, N. Altini, V. Bevilacqua, and M. La Scala, “Radiometric Infrared Thermography of Solar Photovoltaic Systems: An Explainable Predictive Maintenance Approach for Remote Aerial Diagnostic Monitoring,” *Smart Cities 2024, Vol. 7, Pages 1261-1288*, vol. 7, no. 3, pp. 1261–1288, May 2024, doi: 10.3390/SMARTCITIES7030053.
- [90] D. Minh, H. X. Wang, Y. F. Li, and T. N. Nguyen, “Explainable artificial intelligence: a comprehensive review,” *Artificial Intelligence Review 2021 55:5*, vol. 55, no. 5, pp. 3503–3568, Nov. 2021, doi: 10.1007/S10462-021-10088-Y.

- [91] E. Mohamed, K. Sirlantzis, and G. Howells, “A review of visualisation-as-explanation techniques for convolutional neural networks and their evaluation,” *Displays*, vol. 73, p. 102239, Jul. 2022, doi: 10.1016/J.DISPLA.2022.102239.
- [92] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, “Explainable AI: A Review of Machine Learning Interpretability Methods,” *Entropy 2021, Vol. 23, Page 18*, vol. 23, no. 1, p. 18, Dec. 2020, doi: 10.3390/E23010018.
- [93] E. Tjoa and C. Guan, “A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI,” *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 11, pp. 4793–4813, Nov. 2021, doi: 10.1109/TNNLS.2020.3027314.
- [94] D. Erhan, Y. Bengio, A. C. Courville, and P. Vincent, “Visualizing Higher-Layer Features of a Deep Network,” in *ICML 2009 Workshop on Learning Feature Hierarchies, Montréal, Canada.*, 2009.
- [95] A. Nguyen, J. Yosinski, and J. Clune, “Understanding Neural Networks via Feature Visualization: A survey,” 2019, [Online]. Available: <https://arxiv.org/pdf/1904.08939>
- [96] K. Simonyan, A. Vedaldi, and A. Zisserman, “Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps,” in *2nd International Conference on Learning Representations (ICLR 2014) - Workshop Track*, Banff, AB, Canada, Apr. 2014. doi: 10.48550/arXiv.1312.6034.
- [97] D. Smilkov, N. Thorat, B. Kim, F. Viégas, and M. Wattenberg, “SmoothGrad: removing noise by adding noise.” Accessed: Feb. 03, 2025. [Online]. Available: <https://arxiv.org/pdf/1706.03825>
- [98] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, “Striving for Simplicity: The All Convolutional Net,” in *3rd International Conference on Learning Representations (ICLR 2015) - Workshop Track*, San Diego, CA, USA, May 2014. doi: 10.48550/arXiv.1412.6806.
- [99] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” in *Proceedings of the IEEE International Conference on Computer Vision*, Institute of Electrical and Electronics Engineers Inc., Dec. 2017, pp. 618–626. doi: 10.1109/ICCV.2017.74.

-
- [100] L. Van Der Maaten and G. Hinton, “Visualizing Data using t-SNE,” *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.
- [101] L. McInnes, J. Healy, and J. Melville, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,” *ArXiv e-prints*, Feb. 2018, doi: 10.48550/arXiv.1802.03426.
- [102] L. B. Bosman, W. D. Leon-Salas, W. Hutzler, and E. A. Soto, “PV system predictive maintenance: Challenges, current approaches, and opportunities,” Mar. 02, 2020, *MDPI AG*. doi: 10.3390/en13061398.
- [103] J. Ramirez-Vergara, L. B. Bosman, E. Wollega, and W. D. Leon-Salas, “Review of forecasting methods to support photovoltaic predictive maintenance,” *Clean Eng Technol*, vol. 8, Jun. 2022, doi: 10.1016/J.CLET.2022.100460.
- [104] A. Livera, M. Theristis, L. Micheli, E. F. Fernandez, J. S. Stein, and G. E. Georghiou, “Operation and Maintenance Decision Support System for Photovoltaic Systems,” *IEEE Access*, vol. 10, pp. 42481–42496, 2022, doi: 10.1109/ACCESS.2022.3168140.
- [105] M. Fidali, “Thermographic Criteria of Evaluation of Technical Condition of Machinery and Equipment,” *Measurement Automation Monitoring*, vol. 61, no. 06, 2015.
- [106] ISO 18436-7:2008(E), “Condition monitoring and diagnostics of machines - Requirements for qualification and assessment of personnel - Part 7: Thermography (ISO 18436-7),” 2008. [Online]. Available: www.iso.org
- [107] The British Institute of Non-Destructive Testing, “Specific Requirements for Qualification and PCN Certification of Condition Monitoring and Diagnostic Personnel for Infrared Thermography (CM/GEN APPENDIX B, Issue 7),” Northampton, Jan. 2022. [Online]. Available: www.bindt.org,
- [108] FLIR Commercial Systems B.V., “Thermal Imaging Cameras: A Fast and Reliable Tool for Testing Solar Panels (Technical Note),” 2018. Accessed: Nov. 09, 2023. [Online]. Available: http://support.flir.com/appstories/AppStories/Electrical&Mechanical/Testing_solar_panels_EN.pdf
- [109] K. Fukushima, “Neocognitron: A Hierarchical Neural Network Capable of Visual Pattern Recognition,” vol. 1, pp. 119–130, 1988.

- [110] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.
- [111] Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, S. Nasrin, and V. K. Asari, “The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches,” *ArXiv*, Sep. 2018, Accessed: Jan. 20, 2025. [Online]. Available: <https://arxiv.org/abs/1803.01164>
- [112] T. Epelbaum, “Deep learning: Technical introduction,” 2017, Accessed: Feb. 20, 2025. [Online]. Available: <https://arxiv.org/abs/1709.01412>
- [113] R. Fisher, S. Perkins, A. Walker, and E. Wolfart, “Hypermedia Image Processing Reference (Glossary: Convolution),” 2003. Accessed: Feb. 02, 2025. [Online]. Available: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/convolve.htm>
- [114] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” 2012. Accessed: Feb. 01, 2025. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
- [115] S. Mallick and S. Nayak, “Number of Parameters and Tensor Sizes in a Convolutional Neural Network (CNN).” Accessed: Feb. 01, 2025. [Online]. Available: <https://learnopencv.com/number-of-parameters-and-tensor-sizes-in-convolutional-neural-network/>
- [116] K. Chen, “Hidden Layer Activation Functions,” 2021. Accessed: Feb. 01, 2025. [Online]. Available: <https://kinder-chen.medium.com/hidden-layer-activation-functions-6fd65489ed25>
- [117] S. J. D. Prince, *Understanding Deep Learning*. The MIT Press, 2023. Accessed: Feb. 01, 2025. [Online]. Available: <https://udlbook.github.io/udlbook/>
- [118] S. Chan, “Gradient Descent,” 2020. Accessed: Feb. 25, 2025. [Online]. Available: https://engineering.purdue.edu/ChanGroup/ECE595/files/Lecture05_descent.pdf
- [119] A. Rakhecha, “Understanding Learning Rate,” 2019. Accessed: Feb. 15, 2025. [Online]. Available: <https://medium.com/data-science/https-medium-com-dashingaditya-rakhecha-understanding-learning-rate-dd5da26bb6de>

- [120] P. Haidari, A. Hajiahmad, A. Jafari, and A. Nasiri, "Deep learning-based model for fault classification in solar modules using infrared images," *Sustainable Energy Technologies and Assessments*, vol. 52, p. 102110, Aug. 2022, doi: 10.1016/J.SETA.2022.102110.
- [121] M. Le, L. Van Su, N. Dang Khoa, V. D. Dao, V. Ngoc Hung, and V. Hong Ha Thi, "Remote anomaly detection and classification of solar photovoltaic modules based on deep neural network," *Sustainable Energy Technologies and Assessments*, vol. 48, p. 101545, Dec. 2021, doi: 10.1016/J.SETA.2021.101545.
- [122] D. Korkmaz and H. Acikgoz, "An efficient fault classification method in solar photovoltaic modules using transfer learning and multi-scale convolutional neural network," *Eng Appl Artif Intell*, vol. 113, p. 104959, Aug. 2022, doi: 10.1016/j.engappai.2022.104959.
- [123] Z. B. Duranay, "Fault Detection in Solar Energy Systems: A Deep Learning Approach," *Electronics (Switzerland)*, vol. 12, no. 21, Nov. 2023, doi: 10.3390/electronics12214397.
- [124] R. F. Pamungkas, I. B. K. Y. Utama, and Y. M. Jang, "A Novel Approach for Efficient Solar Panel Fault Classification Using Coupled UDenseNet," *Sensors 2023, Vol. 23, Page 4918*, vol. 23, no. 10, p. 4918, May 2023, doi: 10.3390/S23104918.
- [125] Q. Wang, K. Paynabar, and M. Pacella, "Online automatic anomaly detection for photovoltaic systems using thermography imaging and low rank matrix decomposition," *Journal of Quality Technology*, vol. 54, no. 5, pp. 503–516, 2021, doi: 10.1080/00224065.2021.1948372.
- [126] Q. Wang, K. Paynabar, and M. Pacella, "Drone-based Thermography Image Dataset of Photovoltaic Systems." Accessed: Nov. 15, 2023. [Online]. Available: https://tandf.figshare.com/articles/dataset/Online_automatic_anomaly_detection_for_photovoltaic_systems_using_thermography_imaging_and_low_rank_matrix_decomposition/15123655
- [127] Teledyne FLIR LLC, "FLIR® Thermal Studio Suite (Analysis and Reporting Software)." Accessed: Oct. 31, 2023. [Online]. Available: <https://www.flir.com/products/flir-thermal-studio-suite/?vertical=condition%20monitoring&segment=solutions>

REFERENCES

- [128] FLIR® Systems Inc., “Radiometric JPEG images.” Accessed: Oct. 31, 2023. [Online]. Available: https://flir.custhelp.com/app/answers/detail/a_id/1729/~radiometric-jpeg-images