**Comp9517 report**

**Enhancing cell Image Classification via Convolutional Networks with Reinforcement Learning**

**Mingyuan ZHAO**

**Raymond Li**

**University of New South Wales, information technology**

**November 2023**

**TABLE OF CONTENTS**

[CHAPTER 1. Introduction 1](#_Toc151055945)

[1.1 Background and Purpose 1](#_Toc151055946)

[CHAPTER 2. Literature Review 2](#_Toc151055947)

[2.1 Introduction 2](#_Toc151055948)

[2.2 Technology Application 3](#_Toc151055949)

[CHAPTER 3. Methodology 4](#_Toc151055950)

[3.1 Research Framework 4](#_Toc151055951)

[3.2 Objective Function and Constraints 4](#_Toc151055952)

[CHAPTER 4. Experimental Results and Discussion 5](#_Toc151055953)

[4.1 Data Sources and Collection 5](#_Toc151055954)

[4.2 Implementation of the Model 5](#_Toc151055955)

[4.3 Experimental Results and Analysis 6](#_Toc151055956)

[CHAPTER 5. Conclusion 9](#_Toc151055957)

[5.1 Summary of Research Results 9](#_Toc151055958)

[5.2 Limitations and Future Works 9](#_Toc151055959)

[REFERENCES 10](#_Toc151055960)

# Introduction

## Background and Purpose

Due to defects detected that may reduce the power efficiency of solar panels, it is crucial to use Electroluminescence (EL) imaging technology for high-resolution scanning to visualize these defects. These EL images need to be analyzed using computer vision techniques to effectively detect and classify defects. The project aims to develop and test computer vision methods to predict the health status of PV cells in solar component EL images.

**1.2 Research Questions**

How can computer vision techniques be effectively utilized for automated analysis of Electroluminescence (EL) images of solar panels to accurately predict the health status of solar components and detect potential defects?

This question focuses on the development and testing of new computer vision technologies to improve the efficiency and accuracy of monitoring the state of solar panels, especially in terms of damage and defect detection.

Dataset Overview: The dataset contains different types of lattice images, each with an associated label indicating its category. Label information is stored in labels.txt, including image paths, numerical labels, and type labels (mono or poly).

**1.3 Research Objectives**

This project aims to establish a deep learning model for four-class classification of lattice images. The images first go through a preprocessing stage, including the removal of extraneous parts and noise reduction, followed by classification using a Convolutional Neural Network (CNN). Additionally, 【RL (Reinforcement Learning) is used to attempt classification from another perspective, comparing the advantages and disadvantages of different classification methods.】

# Literature Review

## Introduction

In the field of computer vision, enhancing image preprocessing and classification accuracy is a current research focus. Fourier Transform, as a powerful tool for image noise reduction, has been extensively studied in its application to image preprocessing. By converting images from the spatial domain to the frequency domain, Fourier Transform can reveal the periodic structures and noise components in images. Convolutional Neural Networks (CNN) have become the standard tool for handling complex image recognition tasks. CNNs learn the feature representation of images through a series of convolutional layers, with each layer designed to extract more abstract features. Reinforcement Learning (RL) has shown tremendous potential in optimizing parameter selection and network structure design for CNNs. Particularly, the application of RL in Automated Machine Learning (AutoML) provides new avenues for finding optimal network architectures and hyperparameters. Through reinforcement learning algorithms, positive control over the CNN training process can be achieved. This method searches for strategies to improve network performance through trial and error, optimizing learning rate scheduling, preventing overfitting, and enhancing the model's generalization ability to unseen data【supplementary instruction】

## Technology Application

The [1] showcases CNN's capability in efficiently learning discriminative features from training samples and achieving significant classification results in medical imaging.[2] integrated CNN-RNN framework effectively captures both the image features and label dependencies, significantly enhancing the performance of multi-label classification tasks by learning a joint low-dimensional image-label embedding. [3] focus is on analyzing different learning rate settings and optimization algorithms to optimize the parameters for image classification. [4] develops a novel learning rate scheduler using Reinforcement Learning (RL), enabling efficient and automatic adjustment of the learning rate during CNN training.

# Methodology

## Research Framework

【The drawing needs redrawn】

## Objective Function and Constraints

Preprocessing: The Sobel operator is used to detect and remove black lines in the image, followed by noise reduction using Fourier Transform. These steps aim to reduce interference and noise in the images for more accurate classification.

CNN Model Construction: A CNN model is built using the Keras framework, which includes multiple convolutional layers, batch normalization layers, pooling layers, and fully connected layers. The model aims to learn to extract features from the processed images and perform effective classification.

Training and Validation: Data augmentation techniques are used to expand the training dataset and improve the model's generalization capability. Callback functions are utilized to save the best model during training and to evaluate model performance on the validation set.

【The RL part, doesn't have to be at the end of this paragraph】

# Experimental Results and Discussion

## Data Sources and Collection

Main from：ELPV Dataset. A Benchmark for Visual Identification of Defective Solar Cells in Electro-luminescence Imagery. https://github.com/zae-bayern/elpv-dataset

## Implementation of the Model

In our experiments, we focus on using Convolutional Neural Networks (CNN) for image recognition. The CNN architecture in this experiment includes three convolutional layers and an output layer. The first convolutional layer is configured with eight 3x3 convolution kernels, followed by batch normalization. After convolution, a 2x2 max pooling layer with a stride of 2 is applied. In the CNN, each convolutional layer is equipped with activation functions like ReLU and Sigmoid.

During CNN training, overfitting is a common issue, leading to models that are overly complex and tend to learn noise rather than patterns from the training data. To mitigate this, methods like dropout are used. In our model, this is achieved by adding a Dropout layer to prevent overfitting.

Another challenge in CNN model training is adjusting the learning rate, which significantly impacts model performance. Too high a learning rate can lead to overshooting and convergence to suboptimal solutions, while too low a learning rate can lead to slow training or non-convergence of the model. To 日历

描述已自动生成address this, we implemented the ADAM algorithm in our CNN model training to ensure stability and convergence, thereby achieving optimal performance and accuracy.

In the experiment, we also added a batch normalization layer, followed by a convolutional layer with 32 filters, and applied batch normalization again. The backend of the network includes a flattening layer, a fully connected layer with 64 neurons, and an output layer. The output layer uses a four-class sigmoid activation function.

## Experimental Results and Analysis

图片包含 图表

描述已自动生成All image’s result

图片包含 日历

描述已自动生成poly cell’s result

mono cell’s result

According to the results, the accuracy of poly is slightly lower than mono.

Our Convolutional Neural Network (CNN) was tested on a multi-class image classification task, with results displayed in the provided confusion matrix and accompanying performance metrics. This CNN model aims to classify images into one of four categories, represented on the axes of the confusion matrix as classes 0, 1, 2, and 3.

Performance Metrics: The confusion matrix is a key tool in classification tasks, showing the frequency of correct predictions by the CNN for each category versus cases of misclassification. The model reported an accuracy of about 65% on the test set, with a loss of approximately 1.17.

Breakdown of the Confusion Matrix: Analyzing the confusion matrix, we observed the following:

Class 0 (Normal): The model showed strong performance, with 218 true positives and 88 false negatives. This indicates that while the model is good at identifying this category, it tends to misclassify other categories as normal. Class 1 (0.3 Damage): The

model struggled significantly, with all 36 instances being misclassified. This suggests that the model may not have effectively learned to distinguish features of this category or was overshadowed by features of other categories. Class 2 (0.6 Damage): Similar to Class 1, all 9 instances were incorrectly classified, further indicating that the model failed to capture the defining features of this particular category. Class 3 (Full Damage): Here, the model again performed well, correctly identifying 119 instances, but still misclassifying 20 as Class 0. This may suggest some common features between Classes 0 and 3, or a bias towards Class 0. Analysis: The significant confusion between certain categories, especially between Classes 0 and 3, and the complete misclassification of Classes 1 and 2, can be attributed to several factors. These might include an imbalance in the dataset, as it is evident that the 0.6 category (Class 2) is underrepresented in the overall dataset (only 106 out of 2624 images), leading to a bias as the model received more training on instances of Classes 1 and 0. Additionally, the model may not have learned distinctive enough features for some categories.

【This is the result of pure CNN before using RL, and the result after RL should be better than this】

# Conclusion

## Summary of Research Results

From the confusion matrix, it is clear that the current configuration of the model is more proficient at detecting extreme conditions (normal and fully damaged) rather than intermediate conditions (0.3 and 0.6 damage). The high loss value also indicates that there is significant room for improvement in the model's ability to generalize from training data to unseen data.

【Insert RL's overall enhancement ability for the model here】

## Limitations and Future Works

Limitation: Data Imbalance: The disparity in the model's performance across different categories may suggest an issue of data imbalance. Future models could benefit from a more evenly distributed sample across categories.

Overfitting: Despite attempts to mitigate overfitting through techniques such as dropout, the model may still be overly complex for the provided training data. This could be addressed by collecting more data or implementing more complex regularization techniques.

Future Work: Data Augmentation: Exploring more complex data augmentation strategies could enrich the dataset, especially for underrepresented categories, without the need for additional data collection.

Ensemble Methods: Combining predictions from multiple models often results in better performance than any single model, which might be particularly effective in dealing with categories that the current model struggles with.

# REFERENCES

[1] Li, Qing, Weidong Cai, Xiaogang Wang, Yun Zhou, David Dagan Feng, and Mei Chen. "Medical image classification with convolutional neural network." In 2014 13th international conference on control automation robotics & vision (ICARCV), pp. 844-848. IEEE, 2014.

[2] Wang, Jiang, Yi Yang, Junhua Mao, Zhiheng Huang, Chang Huang, and Wei Xu. "Cnn-rnn: A unified framework for multi-label image classification." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2285-2294. 2016.

[3] Guo, Tianmei, Jiwen Dong, Henjian Li, and Yunxing Gao. "Simple convolutional neural network on image classification." In 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), pp. 721-724. IEEE, 2017.

[4] Wen, Long, Xinyu Li, and Liang Gao. "A new reinforcement learning based learning rate scheduler for convolutional neural network in fault classification." IEEE Transactions on Industrial Electronics 68, no. 12 (2020): 12890-12900.

[5] 1] W. Voker, “CNN with OpenCV,” Kaggle. [online]. Available: <https://www.kaggle.com/code/winvoker/cnn-with-opencv>