SOILINGNET: CONVOLUTIONAL NEURAL NETWORKS FOR ANALYSIS OF SOILING DEFECTS IN PHOTOVOLTAIC PANELS

Thomas Gust
Seattle Academy of Arts and Sciences
Seattle, WA, USA
thomasgust@seattleacademy.org

Ivan Rodriguez
Brown University
Providence, RI, USA
ivan.rodriguez5@brown.edu

ABSTRACT

Solar panel soiling is an important problem for the energy sector as we transition away from fossil fuels. Previous methods have aimed to analyze and combat soiling through the use of classical computer vision techniques or weakly supervised deep learning. In our study, we present a two-step, fully supervised deep learning approach for analyzing solar panel soiling on a per-panel level. From a single RGB image of a solar panel, combined with the environmental factor of solar irradiance, our method produces predictions for soiling type, soiling location, and soiling severity. We also introduce a first-of-its-kind dataset labeled with ground truth semantic segmentation maps for the purpose of solar panel soiling analysis. This dataset contains 1104 samples. We find that our model can achieve Jaccard indices in excess of 70 when predicting semantic segmentation maps and top-1 accuracy's in excess of 92% when predicting soiling severity. Our method outperforms previous techniques in the domain of solar panel soiling analysis.

***K*eywords** Artificial Intelligence · Solar Panel Soiling · Convolutional Neural Networks · Solar Panels · Photovoltaic Soiling · Computer Science · Semantic Segmentation · Classification

1 Introduction

The rapid growth of the photovoltaic industry over the past decade has been nothing short of amazing. Solar panels provide what is perhaps humanity's most realistic option for generating carbon-free electricity. Although revolutionary to the energy sector, large and mid-sized solar farms are not without their issues. Maintenance of solar panels can often be complicated, costly, and resource intensive. The phenomenon of solar panel soiling, or the accumulation of particulate matter across the surface of a photovoltaic cell, has been shown to decrease panel energy output by over 25% in real-world studies in very dirty conditions [7]. This establishes PV soiling as a serious problem for the solar panel industry, reducing worldwide revenue by millions of dollars and worldwide PV output by almost 5% [13].

While a human can determine the soiling type and, to some degree, the severity of the soiling impact on a PV cell, this would be absurdly inefficient. Because of this, most large-scale solar farms clean their panels at set intervals. This method also introduces large resource inefficiencies, often using massive amounts of water in excess of 10 billion gallons per year [8]. More targeted cleaning strategies have the potential to dramatically reduce water usage in solar panel cleaning by prioritizing resources where they are actually needed.

Furthermore, without detailed soiling information for individual solar panels, it can be hard to enact effective countermeasures. Some types of soiling, such as bird droppings, may require special attention when cleaning. Different types of soiling can also take more or fewer resources to remove; for example, dust that has been exposed to solar panels for longer can often be more difficult to remove due to the process of dust cementation [3]. This creates a compelling argument for systems that can provide a more in-depth analysis of solar panel soiling because that information is critical in a more efficient cleaning process.

In our study, we aim to develop a deep learning-based system for monitoring, analyzing, and ultimately combating solar panel soiling. This system has 2 primary objectives; the first objective is to produce a segmentation map containing information on soiling type and location; the second objective is a predicted value for the severity or impact of the soiling. If this information can be reliably and accurately produced, it can greatly improve the efficiency and effectiveness of the solar farm cleaning process.

One of the greatest hurdles to developing machine learning systems for solar panel analysis is the lack of labeled datasets large enough to train segmentation models [14]. In this paper, we present a dataset containing images of solar panels with semantic segmentation labels, allowing for the fully supervised training of segmentation models. From a single RGB image of a solar panel and a solar irradiance reading, our system produces a detailed semantic segmentation map to represent the soiling type and location while also predicting the severity of the soiling impact.

2 Related Work

Many techniques exist, both in literature and in practice, for analyzing solar panel soiling. Most of these techniques rely on traditional computer vision algorithms, utilizing IR and RGB images as inputs [1, 2, 11, 3]. While methods such as histogram matching or color filtering are reasonably effective at determining soiling location but are completely incapable of differentiating between different kinds of soiling. More recently, many papers have aimed to take a more intelligent approach to the problem, utilizing modern deep-learning techniques.

These papers always attempt to circumvent the lack of a dataset for solar panel soiling segmentation. While DeepSolarEye [14] introduced the first large-scale dataset of solar panel images labeled with power loss (or soiling impact), it did not tackle the problem of segmentation with human-generated masks, but rather through weakly supervised learning techniques. DeepSolarEye [14] trained a network to predict power loss and then analyzed how each pixel influenced the predicted power loss. While circumventing the need for a traditional labeled image and mask dataset, it also introduces potential inefficiencies. For example, training a weakly supervised model in this manner can possibly result in soiling patches with less influence being undetectable. Having a labeled dataset also makes the process of determining soiling type much simpler.

Deep learning-based semantic segmentation models, specifically fully convolutional networks, have been proven to be powerful tools, used for everything from biomedical image segmentation [15, 4] to self-driving cars [5, 16]. Up until this point, the lack of large labeled datasets for these models has prevented them from being used in the domain of solar panel defect analysis [14]. In our study, we solve this problem by introducing a human-labeled dataset of solar panel soiling.

3 Dataset

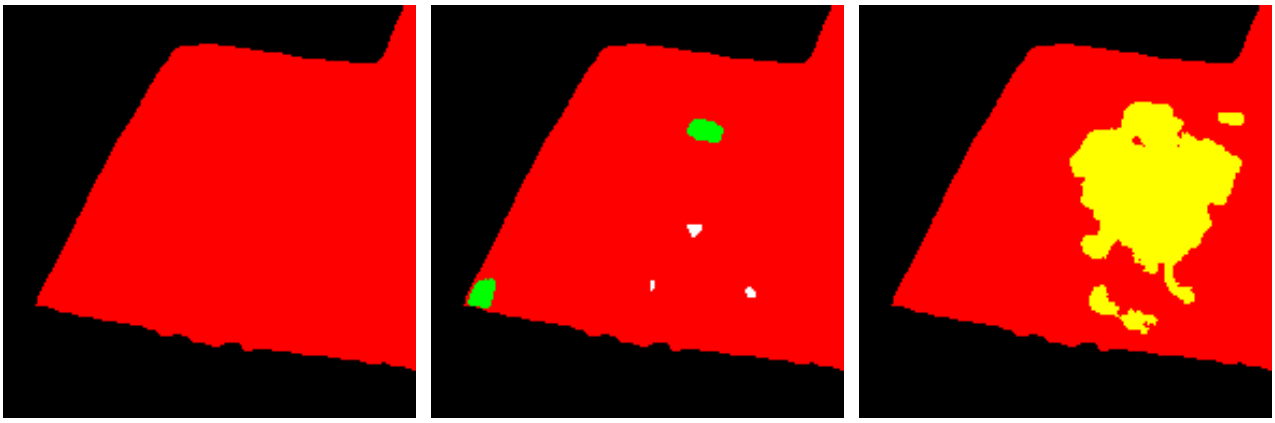
DeepSolarEye [14], introduced the first large-scale solar panel soiling dataset. This dataset was comprised of 45,754 images of clean and soiled solar panels annotated with solar irradiance and soiling impact. While powerful, this dataset was not annotated with segmentation masks for any of the images. We introduce a human-labeled dataset of solar panel images paired with semantic

segmentation masks. Our dataset contains 8 classes, 6 of which are soiling classes, and is expanded to over 2000 images through the use of data augmentation.

3.1 Semantic Segmentation Labels

Previous works have utilized weakly supervised learning techniques to build segmentation models with only input images and soiling severity labels. This method circumvents the need for semantic segmentation labels but also places a limit on the accuracy of these models. Furthermore, these methods employ webly supervised learning techniques to predict soiling category [6], an additional step that can further reduce model performance and add inefficiencies. We circumvent these issues by creating human-labeled semantic segmentation labels for a subset of the solar panel soiling dataset introduced in DeepSolarEye [14]

In our study, we utilize the modern computer-assisted labeling tool, segments.ai, to efficiently create a set of semantic segmentation labels for our dataset. Of our initial dataset of 45,754 images, we take 184 and produce ground truth segmentation masks. This relatively small dataset contains multiple samples representing every kind of soiling and distribution contained within the larger dataset. We then apply dataset augmentation discussed in section 3.3 to expand our dataset to 1104 images. This process of data augmentation also increases the diversity of our dataset.



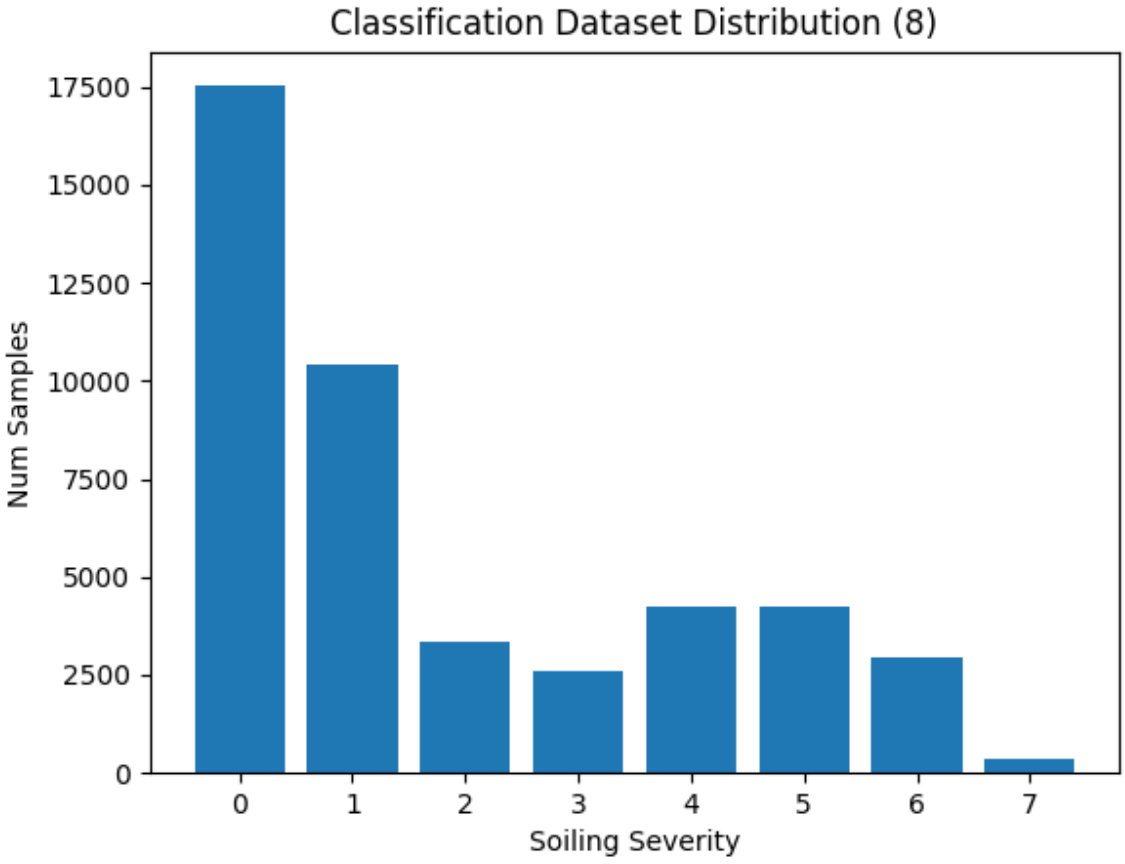
Our newly labeled dataset contains 8 unique classes. These classes are background, clean solar panel surface, red, brown, light gray, bird poop, snow, and white speck. These classes are somewhat arbitrary and are defined by subjective visual appearance. The source dataset [14] utilizes substances such as dust and talcum powder to emulate different kinds of solar panel soiling with different distributions and consistencies. Because of this, classes such as bird poop and snow are not perfectly representative of their in-field counterparts. This is not necessarily an issue due to recent advancements in transfer learning and knowledge distillation. By applying these techniques, it is possible to fine-tune semantic segmentation models to new classes with minimal training data.

3.2 Dataset Binning

Traditionally, deep learning techniques, particularly convolutional neural networks, perform much better on classification problems than regression problems. It is easier for a model to predict a discrete class than to predict some sort of continuous output. In our dataset, soiling impact is recorded as a decimal. In order to solve the problem of one of our output values being continuous, we partition the soiling impact variable into discrete classes.

In order to better understand the performance of our method, we create multiple datasets with different levels of detail in their partition. We create datasets with 4, 8, 12, and 16 discrete solar

panel soiling severity classes. In order to achieve this, we use the simplest method of dataset binning, equal width binning. Equal width binning splits the dataset into n-bins with an equal range or width. An image of our dataset distribution when it is binned with 8 severity classes can be found below. See Appendix A for figures representing different configurations of the dataset



(d) Soiling Severity Dataset Distribution

3.3 Data Augmentation

Data augmentation, or the process of artificially increasing the number of samples in a dataset by editing existing samples, has been proven to be a powerful tool in preventing model over-fitting and increasing accuracy [17]. In order to increase the diversity of our semantic segmentation dataset, we employ a variety of data augmentation techniques. We apply 5 different augmentations to every image in our dataset. These transforms are as follows, a random 90-degree rotation, grid distortion, horizontal flipping, vertical flipping, and Gaussian blur. Each of these augmentations is applied once to every image in our dataset, thus multiplying the number of samples in our semantic segmentation dataset by 6. This provides a significant increase in both the size and diversity of our dataset. Our new dataset contains 1104 image and segmentation mask pairs.

(e) Gaussian Blur Example

(f) Vertical Flip Example

(g) Grid Distortion Example

4 Our Method

We propose a two-stage method for defect analysis of PV cells. The first stage is a fully convolutional neural network (FCN) trained to produce soiling segmentation masks on RGB images of solar panels. The second stage consists of a classification network, which utilizes the mask produced by the FCN, the original solar panel image, and a solar irradiance reading from near the solar panel as input. The learning objective for this network is to classify the power loss or soiling impact of the panel. In our study, we evaluate multiple FCN model architectures and multiple architectures for power loss prediction. Evaluated segmentation models include U-Net [15], SegNet [4], and FCN-32[12]. We evaluated the performance of a variety of CNN architecture for the soiling impact classification learning objective. Additionally, we also evaluated the performance of a simple, fully connected network on the classification task to provide a baseline for our other models.

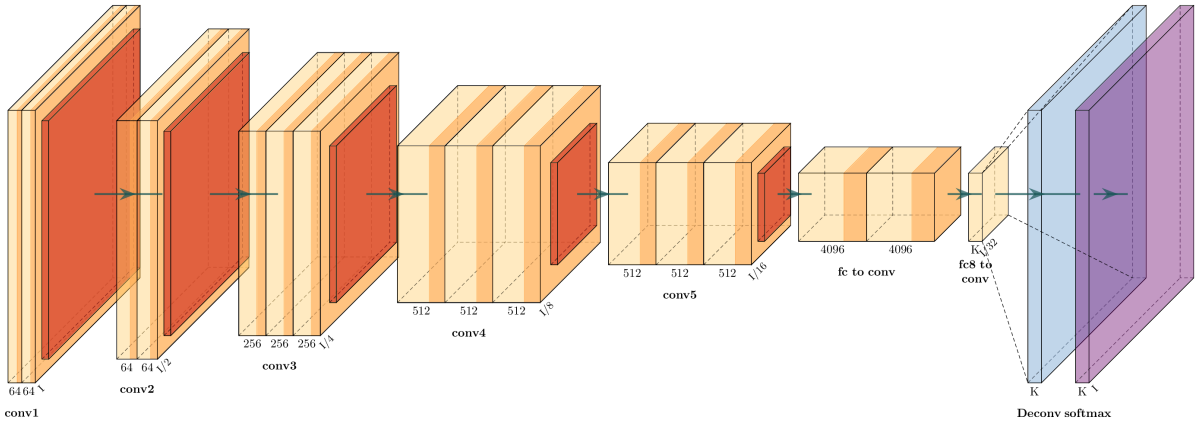
4.1 Segmentation Model Architectures

We aim to find a model architecture for semantic segmentation that effectively balances the trade-off between computational efficiency and accuracy. All models we analyze are FCNs. FCNs utilize convolutional down-sampling, typically seen in traditional image classification networks, with convolutional up-sampling. FCN models come in many different varieties, but most follow the basic pattern of down-sampling the input image into a bottleneck and then up-sampling the feature representation of the bottleneck into an image with a number of channels equal to the number of soiling types. To produce a prediction that can be visualized by humans, the softmax activation function is applied to the model's outputs to create per-pixel probabilities for each segmentation class; then, a pixel-wise max of these probabilities creates the final semantic segmentation prediction. In our study, we analyze the effectiveness of multiple different FCN models in a new domain, the semantic segmentation of photovoltaic soiling.

4.1.1 FCN-32

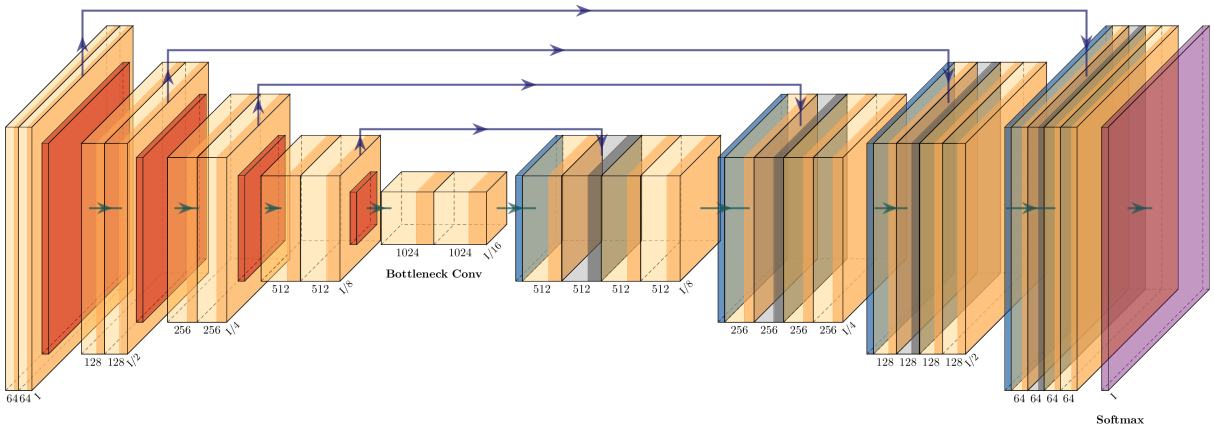
Among the first fully convolutional architectures ever proposed for image segmentation, FCN-32 is a good baseline for segmentation model effectiveness. The architectures encoder block, or convolutional downsampler, is similar to any other convolutional neural network architecture. In opposition to other fully convolutional architectures, the FCN-32 decoder consists of only a single layer, upsampling the feature representation of the bottleneck into a segmentation prediction. It is noteworthy that the FCN-32 architecture lacks any type of residual connection, feature representations or pooling indices are never re-introduced at a later stage in the network. Due to its simplicity,

in our study, we use the FCN-32 architecture as our segmentation baseline model. A simplified outline of the FCN-32 architecture can be seen below.



4.1.2 U-Net

The U-Net FCN architecture [15], first introduced in 2015 by Ronneberger et al, has been the go-to method for semantic segmentation of images across all fields of computer vision. Originally developed as an improvement to some of the first FCNs [12] introduced in 2014 by Long et al, the U-Net intended to solve the problems of information loss and dataset size that had plagued early FCN architectures. Convolutional encoders (or down-samplers) can never perfectly encode their inputs in a lossless way due to pooling operations. Due to this, convolutional decoders had a fundamental limit to their accuracy because there would also be some amount of information loss throughout the network. U-Nets, aim to solve this problem by reintroducing feature maps from each level of the convolutional encoder to the respective layer of the convolutional decoder. In this way, U-Net provides dramatically better semantic segmentation accuracy than its predecessors, at the cost of lower efficiency, due to the computational expense of re-introducing entire feature maps from past layers of the network. We utilize a variation of the U-Net architecture known as the ResNet-UNet, which utilizes the ResNet50 [10] CNN architecture as a convolutional encoder for the U-Net. A diagram of a generic U-Net model can be seen below.

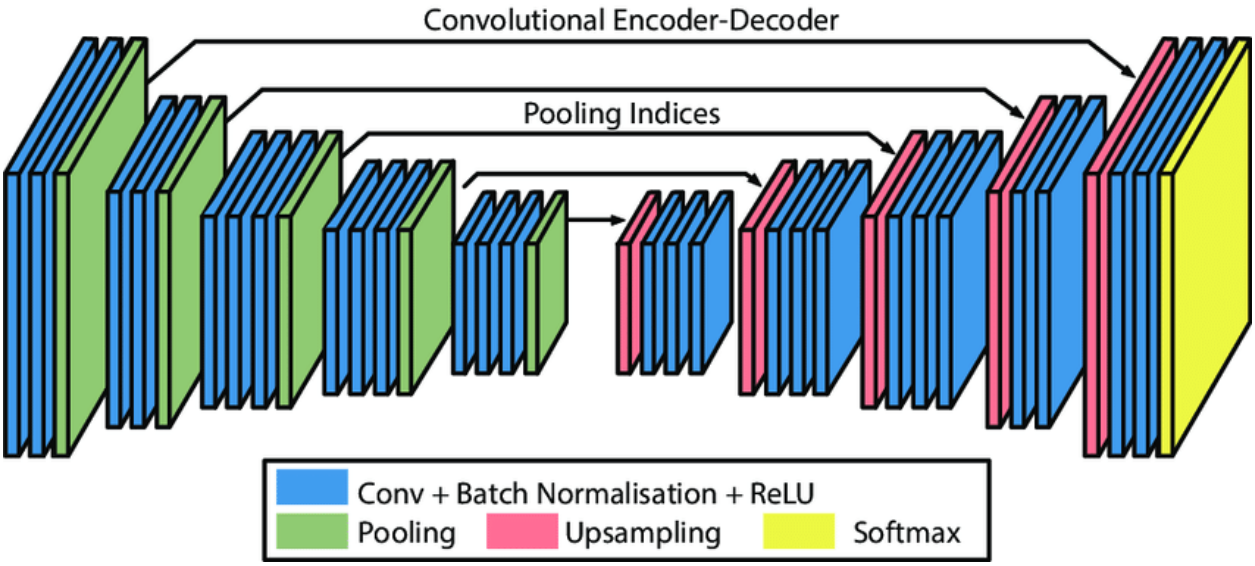


Another key feature of U-Nets, and most other FCN models that came after it, is its ability to train with very few training examples. Its symmetric architecture has proven to be very efficient.

Only a few hundred, not thousands, of semantic segmentation labels, are required to effectively train; this feature of U-Net is very important for our domain, where only limited data is available.

4.1.3 SegNet

SegNet [4] is an architecture designed to solve one of the critical inefficiencies of the U-Net system. Where U-Net reintroduces entire feature representations from its encoder into the decoder using residual connections between convolutional layers, SegNet only reintroduces pooling indices. By only moving pooling indices from the compression path to the expansion path, SegNet has the potential to use dramatically less memory. This property of SegNets is very important for our unique task. When deployed in the field, it is not guaranteed that our system will have vast computational resources. Because of this, model efficiency and size are very important when selecting a proper architecture.



The SegNet aims to eliminate needless computational inefficiencies and is often able to achieve accuracies and Jaccard indices equal to or surpassing those of UNet models. For the purposes of our study, we use a variant of the classical SegNet architecture known as the ResNet-SegNet, which utilizes the ResNet50 [10] CNN architecture as a convolutional encoder for the SegNet. In our results section, we see that SegNet performs better than other segmentation models in our domain.

4.2 Classification Model Architecture

In our study, we aim to effectively predict the soiling impact on photovoltaic cell power output, not just soiling type and distribution. In order to do this, we employ a simple convolutional neural network which is trained on binned versions of the dataset introduced by DeepSolarEye [14].

The model consists of only two convolutional layers followed by two feedforward or linear layers. The input is a four-dimensional array representing a concatenation of a solar panel image and a predicted segmentation map. At the beginning of the linear component of the model, the on-site solar irradiance measurement is concatenated to the flattened feature representation generated by the convolutional layers.

4.3 Segmentation Model Training

The training process for each of our semantic segmentation models is identical. We optimizer our segmentation models with the Adadelta optimizer and the pixel-wise categorical cross-entropy loss function. The segmentation masks within our dataset contain a high degree of class imbalance,

this is due to background and clean panel classes being much more common than any soiling class. In order to combat this problem, we utilize the common technique of weighting our loss function. The losses for background and clean panel classes are kept the same, whereas the losses for classes representing soiling are multiplied by 3. This technique of a weighted categorical cross-entropy loss ensures the model does not get stuck, only predicting clean panel and background classes.

We train each segmentation model for 100 epochs with an optimizer learning rate of 0.023. During model training, we track the metrics of loss, accuracy, and mean intersection over union (otherwise known as the Jaccard index). Graphs of these values can be found below in Appendix B

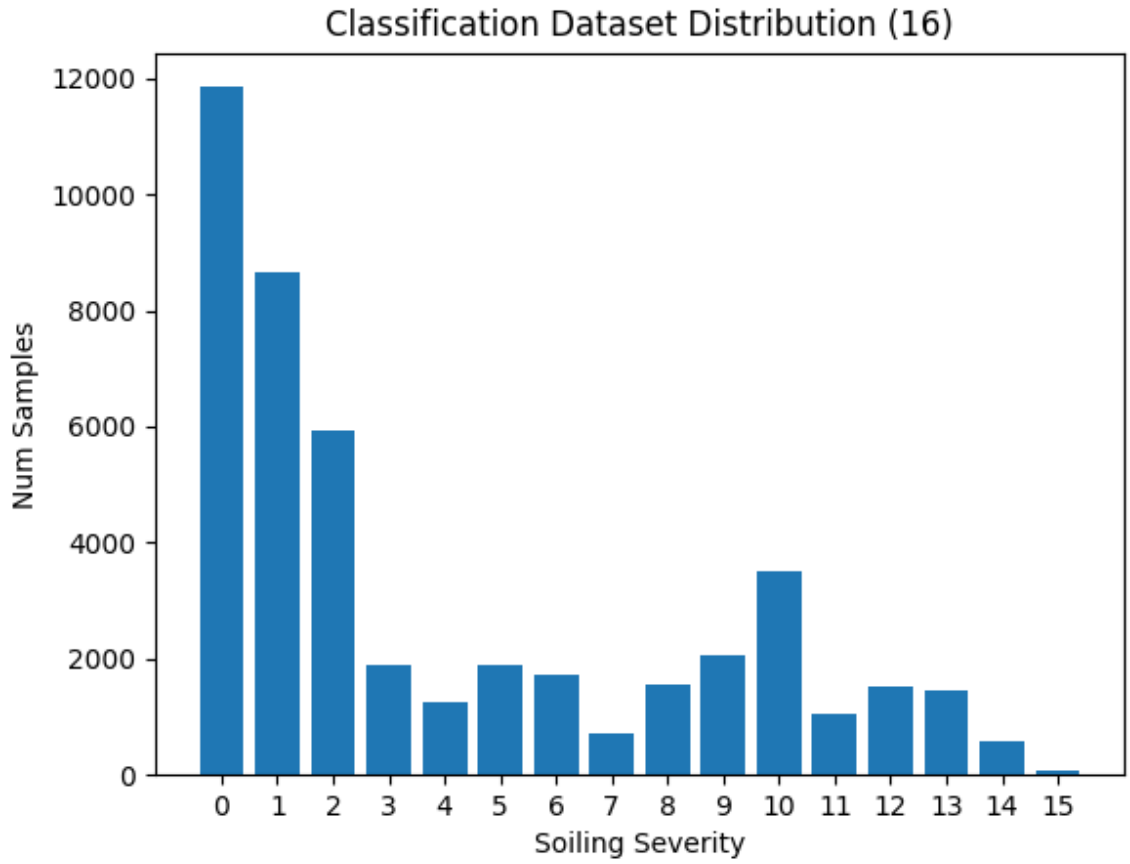
4.4 Classification Model Training

Classification training is dependent on a segmentation model having been previously trained. For our classifier model, the input is an image of a solar panel, a predicted dust mask of the solar panel, and a solar irradiance reading from near the panel location. The training process is similar to that of any other image classification model, but we introduce critical changes, such as dataset oversampling, that are required for high performance on our unique dataset.

Dataset loading is similar to the dataset loading for segmentation model training; the only change is swapping semantic segmentation labels for semantic segmentation predictions from one of our previously trained segmentation models. Instead of the comparatively small dataset used for segmentation model training, we can now have access to the full dataset containing 45k+ annotated solar panel images. Instead, to verify our classification models are not overfitting, we train on every third image within our wider dataset.

4.4.1 Dataset Oversampling

As discussed in section3, the original dataset of 45,754 images is incredibly imbalanced in terms of soiling impact classes.



As seen in the figure above 1h, the dataset is heavily skewed towards classes on the lower end of the soiling impact severity spectrum. Representation is especially low for moderately soiled images, which are likely to appear often in real-world circumstances. This presents a problem for our classification models. When training any of our classification models on this dataset directly, accuracy would never exceed a certain limit, usually as low as 42% for a dataset with two classes. This is due to the model only predicting one class due to skew within the dataset. At the start of training, the model will quickly reach a false local minimum in the loss function, never reaching a more global minimum.

To combat this problem, we apply dataset oversampling to our dataset. Instead of loading every training sample once per epoch in random order, we apply weights to the training samples in our dataset. The number of each soiling severity class is counted across the dataset; the corresponding weight for this class is

$$\frac{1}{N_{SAMPLES}} \tag{1}$$

. For example, in a dataset with two classes, one containing 80 images and another containing 20, the corresponding weights would be

$$\frac{1}{80} \tag{2}$$

, and

$$\frac{1}{20} \tag{3}$$

These class weights are then applied to our data loader, ensuring that the model will encounter all soiling severity classes in roughly equal number. Additionally, we artificially multiply the sampler weights for the first and second classes by 1.5 due to their high prevalence in our dataset and in the real world. This solves the problem of classification models only learning to predict one class.

4.4.2 Training Loop

The training loop for our classification model is not unlike that of any other image classification model. We train each model for a total of 100 epochs. For each epoch, we iterate through our data loader, utilizing the dynamically generated class weights created by the sampler. The fetched training samples are passed through our model, which generates predictions on the input data. Model loss is subsequently computed with the categorical cross-entropy loss function. Classification model parameters are then adjusted through the use of the Adam optimizer off of the computed loss. We also log metrics such as loss and top-1 accuracy. These metrics can be found below in 7. At the end of the training process, we also evaluate a top-2 accuracy to understand how often our model is off by more than one soiling severity class.

Our initial learning rate at epoch 0 is (1e-4), or 0.0001. We apply a learning rate scheduler to this value at each epoch, multiplying the current learning rate by 0.985. The function for our learning rate at any epoch can be defined as

$$\alpha(n) = 0.0001 * 0.985^n \tag{4}$$

. Our learning rate at epoch 100 is, therefore, 0.00002206089. We find that a batch size of 64 performs well with different dataset configurations.

5 Results

In this section, we discuss the overall effectiveness of our method both subjectively and objectively. We individually explore the effectiveness of both our segmentation models and our classification models. We find that our method achieves record-breaking segmentation and classification accuracies for the task of analyzing solar panel soiling.

5.1 Segmentation Models Results

We find that our method achieves record-breaking Jaccard indices on our dataset; this is expected due to us introducing ground truth training labels rather than using techniques such as weakly supervised learning. Our most effective model, SegNet, is often able to achieve Jaccard indices in excess of 70 on images it has not seen before. In this section, more statistics and insights for each of our segmentation models.

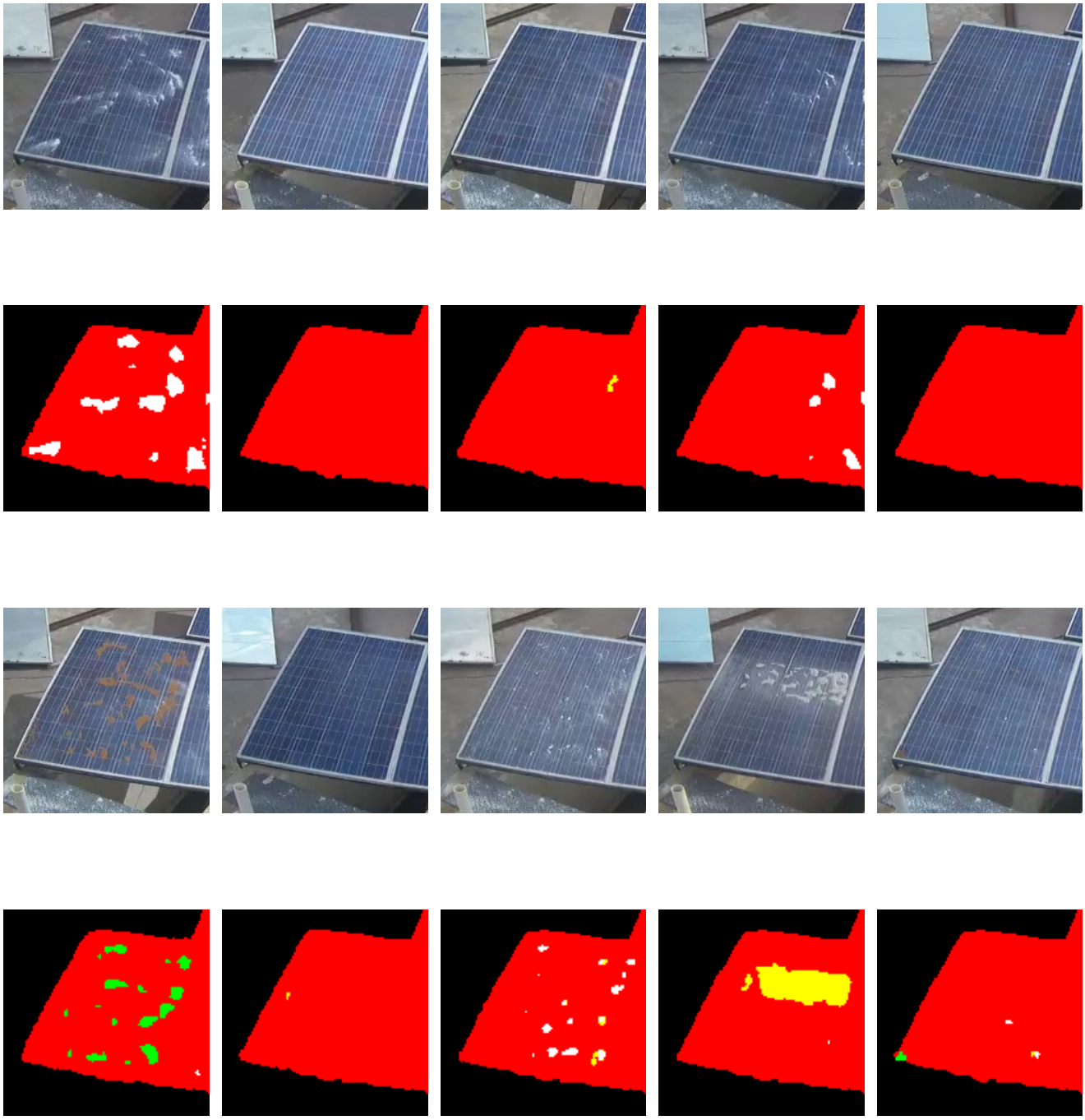
5.1.1 Objective Analysis

Table 1: Segmentation Model Evaluation

Name	Test Split	Train Jaccard Index	Dataset Jaccard Index	Parameters
FCN-32	20.0%	57.38	54.46	70.02 million
U-Net	20.0%	82.01	71.81	16.38 million
SegNet	20.0%	**97.73**	**72.35**	**14.87 million**

5.1.2 Subjective Analysis

SegNet is both the most accurate and efficient model in terms of Jaccard indices and model size. We subjectively analyze SegNet by testing it with images from our wider dataset that it has never seen before. For more predictions from each model, look below in Appendix D



SegNet is able to accurately identify soiling patches across a variety of different input images. SegNet can sometimes become confused between similar soiling classes, as seen above but is consistently able to locate a variety of soiling patches.

5.2 Classification Models

We find that our system is able to classify soiling severity with a high degree of accuracy, able to achieve top-1 accuracies of over 90% in some configurations of our dataset. We find that our method for power loss classification is capable of meeting and exceeding the soiling classification accuracies of other soiling severity classification methods. Furthermore, our method is rarely wildly wrong, with top-2 classification accuracy's in excess of 89%, even in dataset configurations with 16 classes. In the sections below, we provide objective and subjective analyses of our classification models performance on different dataset configurations.

5.2.1 Analysis

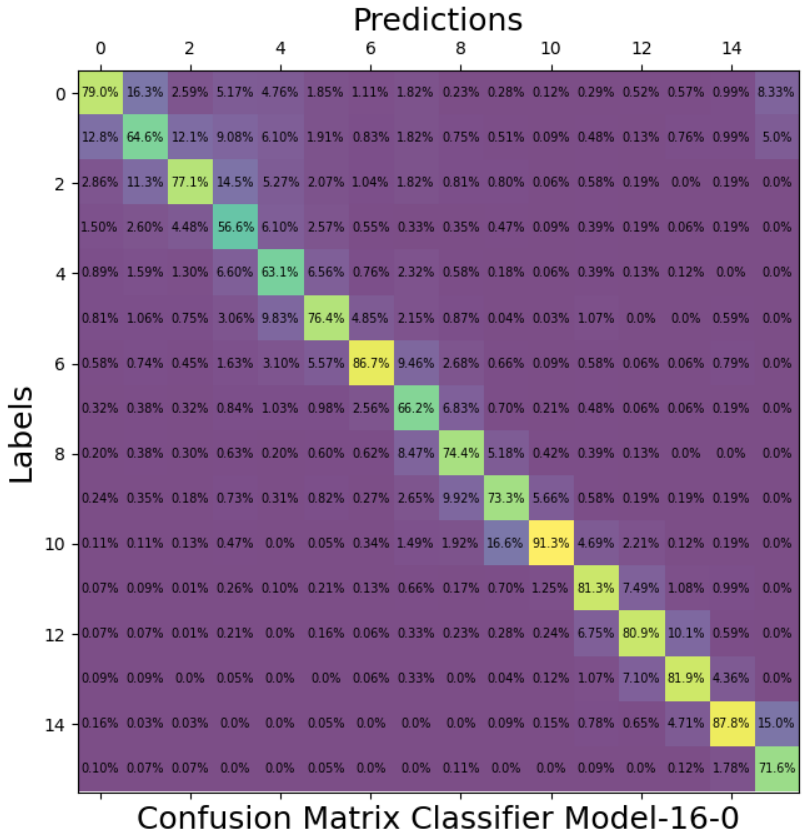
After training for 100 epochs, we evaluate each of our classification models with the metrics of top-1 accuracy and top-2 accuracy. Though every model was only trained on every third image, the metrics below are computed across the entire dataset. This serves as a way to ensure our models are not severely overfitting to their training data.

Table 2: Classifier Model Evaluation

Name	Classes	Top-1 Accuracy	Top-2 Accuracy
Classifier-4	4	92.47%	98.66%
Classifier-8	8	84.01%	95.03%
Classifier-12	12	80.21%	91.92%
Classifier-16	16	75.33%	89.38%

These results are expected. We see that our models achieve lower accuracy's as the number of classes in the dataset increases. It can also be seen that every model has a very high top-2 accuracy; this is important because it verifies that each model is rarely wildly wrong in its power loss predictions.

In order to better understand where our classification models are incorrect, we generated confusion matrices for each model. Each of these confusion matrices can be found below in Appendix E.



In the figure above, we can see that most error is concentrated toward the lower end of our dataset. This is largely due to the similarity between images at the lower end of our dataset. There is only a minute visual difference between an image with a soiling severity of 1 and a soiling severity of 2 in a dataset of 16 classes. Furthermore, most errors are under-classifications of soiling severity rather than over-classifications.

6 Business Applications

As discussed in the introduction, analysis of solar panel soiling can streamline the process of cleaning solar farms. By gathering detailed information about solar panel soiling at a per-panel level, many new possibilities are opened to large-scale solar farms. By observing the data generated by our model over a long period of time, trends of where an accumulation of solar panel soiling is most intense can be discovered and analyzed; this would allow more long-term countermeasures to be introduced. Combating soiling on a per-panel level also opens the door for businesses to save money in a variety of ways. By cleaning each panel individually, businesses can highly increase the efficiency of their cleaning resources, both cleaning agents and personnel. Information from soiling severity classification models can be used to corroborate power loss measurements from previously installed sensors. Deep learning-based monitoring techniques have the potential to reduce the 10 billion gallons of water used every year for solar panel cleaning [8]. Enriched data also has the potential to inform solar panel cleaning robots such as those discussed in Hashim et al [9].

Due to the rapidly growing nature of the solar panel market, reducing soiling power loss by even 2-3% has the potential to save solar panel farms millions of dollars worldwide. This reduction in

solar panel soiling power loss also has the potential to increase the power outputs of large-scale solar farms, particularly in solar panel soiling-rich environments such as deserts.

7 Conclusion

By providing in-depth, highly accurate information about solar panel solar soiling, SoilingNet opens the doors for the solar panel cleaning technologies of the future. In our study, we present the first-ever fully supervised learning approach to monitoring and analyzing solar panel soiling. From a single RGB image of a solar panel, combined with environmental factors, our method is able to accurately predict important characteristics such as soiling type, soiling location, and soiling impact. We also construct a dataset containing 1104 pairs of solar panel images and soiling segmentation masks. This dataset can be used and expanded upon in future studies to create even more powerful solar panel soiling analysis systems.

References

[1] Mohammadreza Aghaei, Sonia Leva, and Francesco Grimaccia. "PV power plant inspection by image mosaicing techniques for IR real-time images". English. In: *2016 IEEE 43rd Photovoltaic Specialists Conference, PVSC 2016*. 43rd IEEE Photovoltaic Specialists Conference, PVSC 2016, PVSC 2016 ; Conference date: 05-06-2016 Through 10-06-2016. United States: Institute of Electrical and Electronics Engineers, Nov. 2016, pp. 3100–3105. DOI: 10.1109/PVSC.2016.7750236. URL: https://www.ieee-pvsc.org/PVSC43/.

[2] Mohammadreza Aghaei et al. "IR real-time Analyses for PV system monitoring by Digital Image Processing Techniques". In: June 2015. DOI: 10.1109/EBCCSP.2015.7300708.

[3] Tamadher M.A Alnasser et al. "Impact of dust ingredient on photovoltaic performance: An experimental study". In: *Solar Energy* 195 (2020), pp. 651–659. ISSN: 0038-092X. DOI: https://doi.org/10.1016/j.solener.2019.12.008. URL: https://www.sciencedirect.com/science/article/pii/S0038092X1931223X.

[4] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. *SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation*. 2015. DOI: 10.48550/ARXIV.1511.00561. URL: https://arxiv.org/abs/1511.00561.

[5] Senay Cakir et al. *Semantic Segmentation for Autonomous Driving: Model Evaluation, Dataset Generation, Perspective Comparison, and Real-Time Capability*. 2022. DOI: 10.48550/ARXIV.2207.12939. URL: https://arxiv.org/abs/2207.12939.

[6] Xinlei Chen and Abhinav Gupta. *Webly Supervised Learning of Convolutional Networks*. 2015. DOI: 10.48550/ARXIV.1505.01554. URL: https://arxiv.org/abs/1505.01554.

[7] R. R. Cordero et al. "Effects of soiling on photovoltaic (PV) modules in the Atacama Desert". In: *Scientific Reports* 8.1 (2018), p. 13943. ISSN: 2045-2322. DOI: 10.1038/s41598-018-32291-8. URL: https://doi.org/10.1038/s41598-018-32291-8.

[8] World Economic Forum. *This is how solar panels can be kept clean - without using water*. 2022. URL: https://www.weforum.org/agenda/2022/03/solar-panels-waterless-clean.

[9] Nurhasliza Hashim et al. "Study on Solar Panel Cleaning Robot". In: *2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*. 2019, pp. 56–61. DOI: 10.1109/I2CACIS.2019.8825028.

[10] Kaiming He et al. *Deep Residual Learning for Image Recognition*. 2015. DOI: 10.48550/ARXIV.1512.03385. URL: https://arxiv.org/abs/1512.03385.

[11] Chris Henry et al. "Automatic Detection System of Deteriorated PV Modules Using Drone with Thermal Camera". In: *Applied Sciences* 10.11 (2020). ISSN: 2076-3417. DOI: 10.3390/app10113802. URL: https://www.mdpi.com/2076-3417/10/11/3802.

[12] Jonathan Long, Evan Shelhamer, and Trevor Darrell. *Fully Convolutional Networks for Semantic Segmentation*. 2014. DOI: 10.48550/ARXIV.1411.4038. URL: https://arxiv.org/abs/1411.4038.

[13] Mohammad Reza Maghami et al. “Power loss due to soiling on solar panel: A review”. In: *Renewable and Sustainable Energy Reviews* 59 (2016), pp. 1307–1316. ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2016.01.044. URL: https://www.sciencedirect.com/science/article/pii/S1364032116000745.

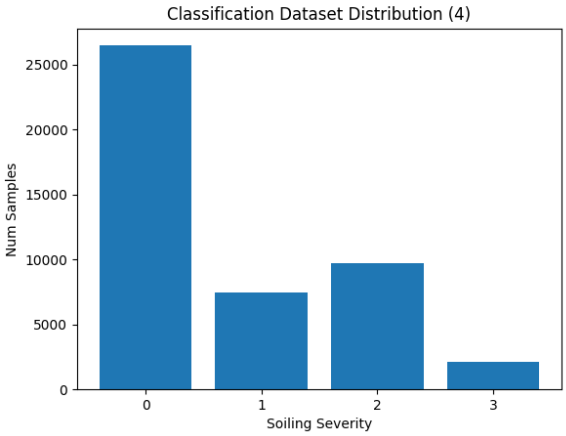
[14] Sachin Mehta et al. *DeepSolarEye: Power Loss Prediction and Weakly Supervised Soiling Localization via Fully Convolutional Networks for Solar Panels*. 2017. DOI: 10.48550/ARXIV.1710.03811. URL: https://arxiv.org/abs/1710.03811.

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. DOI: 10.48550/ARXIV.1505.04597. URL: https://arxiv.org/abs/1505.04597.

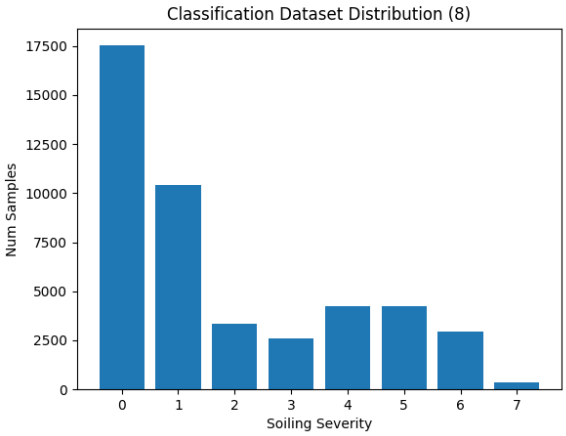
[16] Qusay Sellat et al. “Intelligent Semantic Segmentation for Self-Driving Vehicles Using Deep Learning”. In: *Computational Intelligence and Neuroscience* 2022 (2022), p. 6390260. ISSN: 1687-5265. DOI: 10.1155/2022/6390260. URL: https://doi.org/10.1155/2022/6390260.

[17] Connor Shorten and Taghi M. Khoshgoftaar. “A survey on Image Data Augmentation for Deep Learning”. In: *Journal of Big Data* 6.1 (2019), p. 60. ISSN: 2196-1115. DOI: 10.1186/s40537-019-0197-0. URL: https://doi.org/10.1186/s40537-019-0197-0.

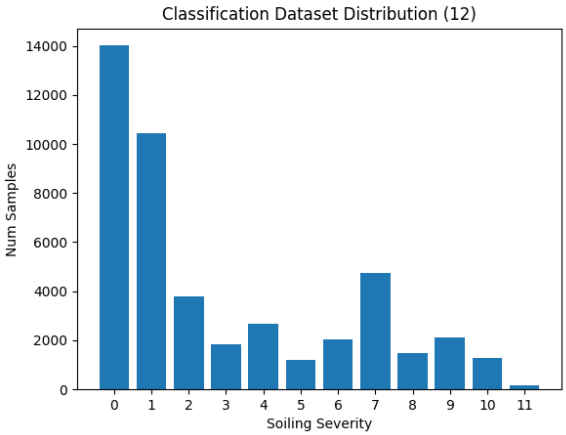
Appendix A



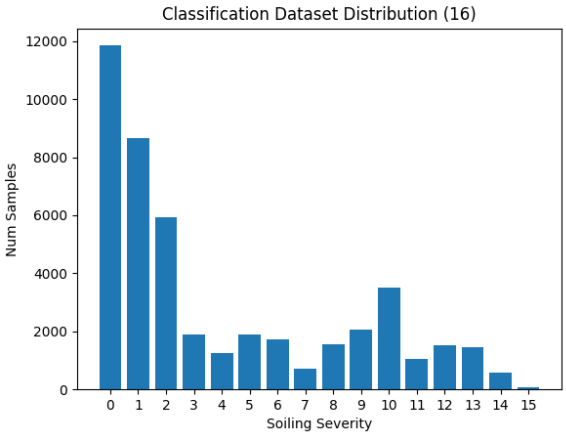
(ad) Soiling Severity Distribution (4 classes)



(ae) Soiling Severity Distribution (8 classes)

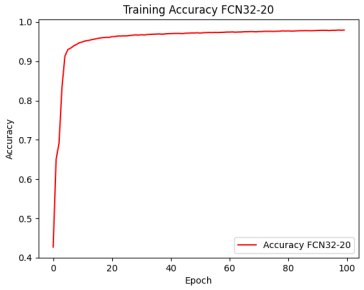


(af) Soiling Severity Distribution (12 classes)

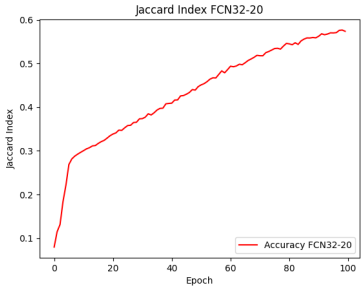


(ag) Soiling Severity Distribution (16 classes)

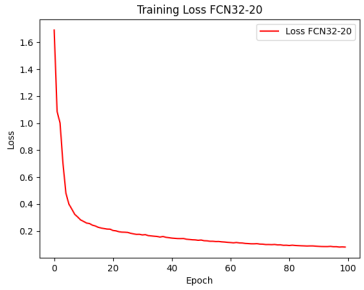
Appendix B



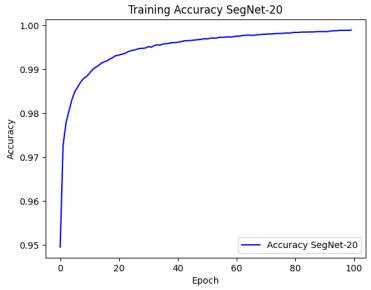
(ah) Training Accuracy FCN-32 model



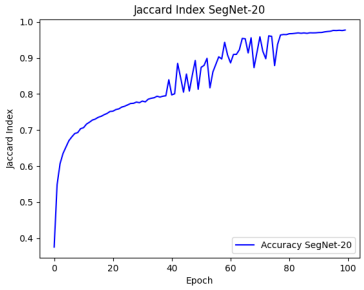
(ai) Training Jaccard Index FCN-32 model



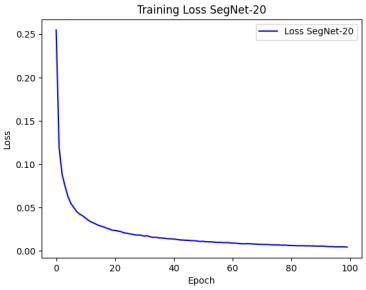
(aj) Training Loss FCN-32 model



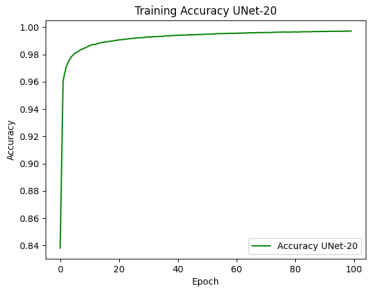
(ak) Training Accuracy SegNet model



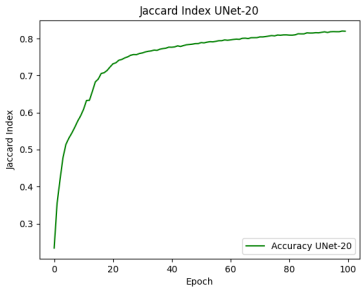
(al) Training Jaccard Index SegNet model



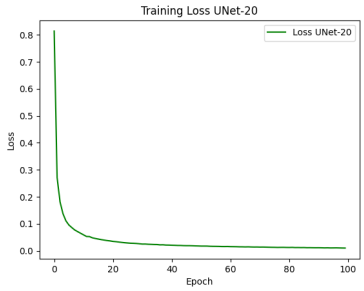
(am) Training Loss SegNet model



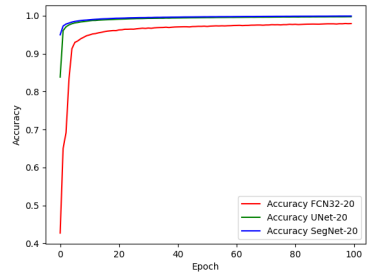
(an) Training Accuracy U-Net model



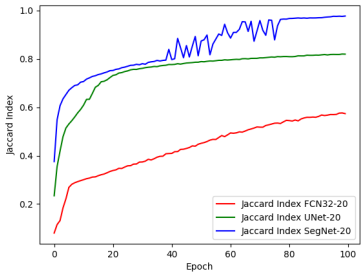
(ao) Training Jaccard Index U-Net model



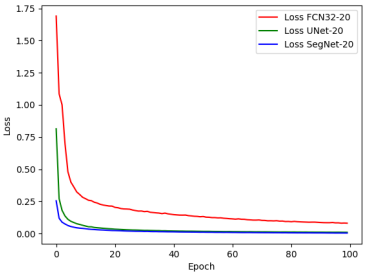
(ap) Training Loss U-Net model



(aq) Train Accuracy Segmentation Models

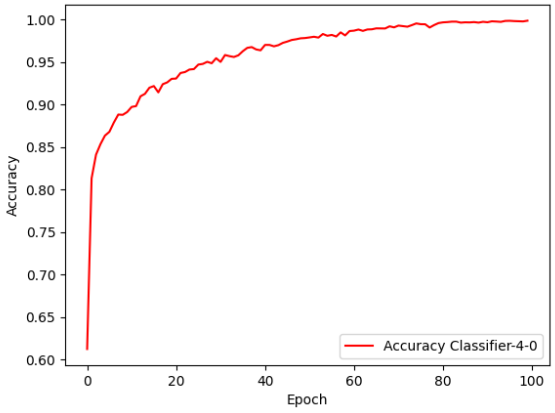


(ar) Train Jaccard Index Segmentation Models

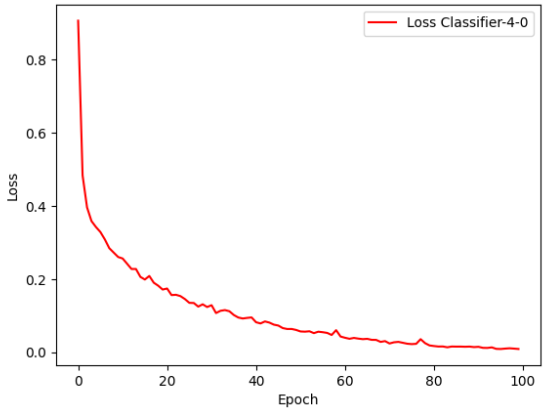


(as) Train Loss Segmentation Models

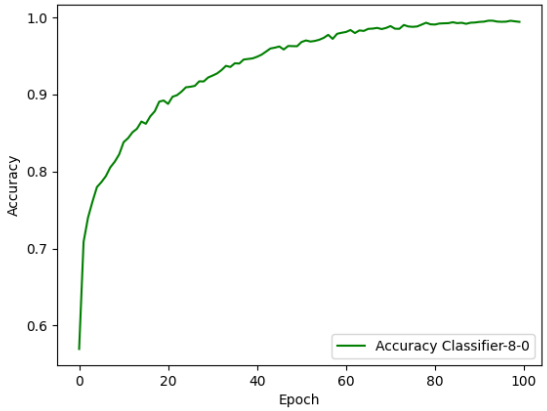
Appendix C



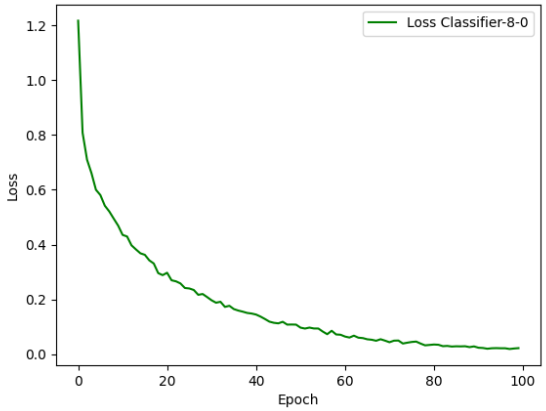
(at) Training Accuracy Classifier 4



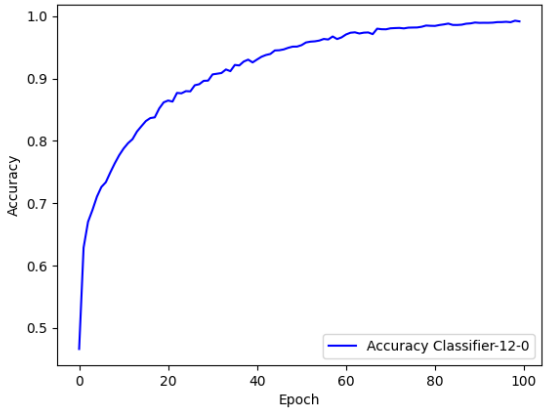
(au) Training Loss Classifier 4



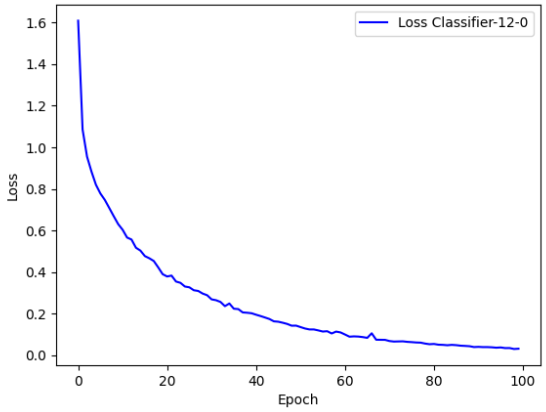
(av) Training Accuracy Classifier 8



(aw) Training Loss Classifier 8



(ax) Training Accuracy Classifier 12



(ay) Training Loss Classifier 12

Appendix D

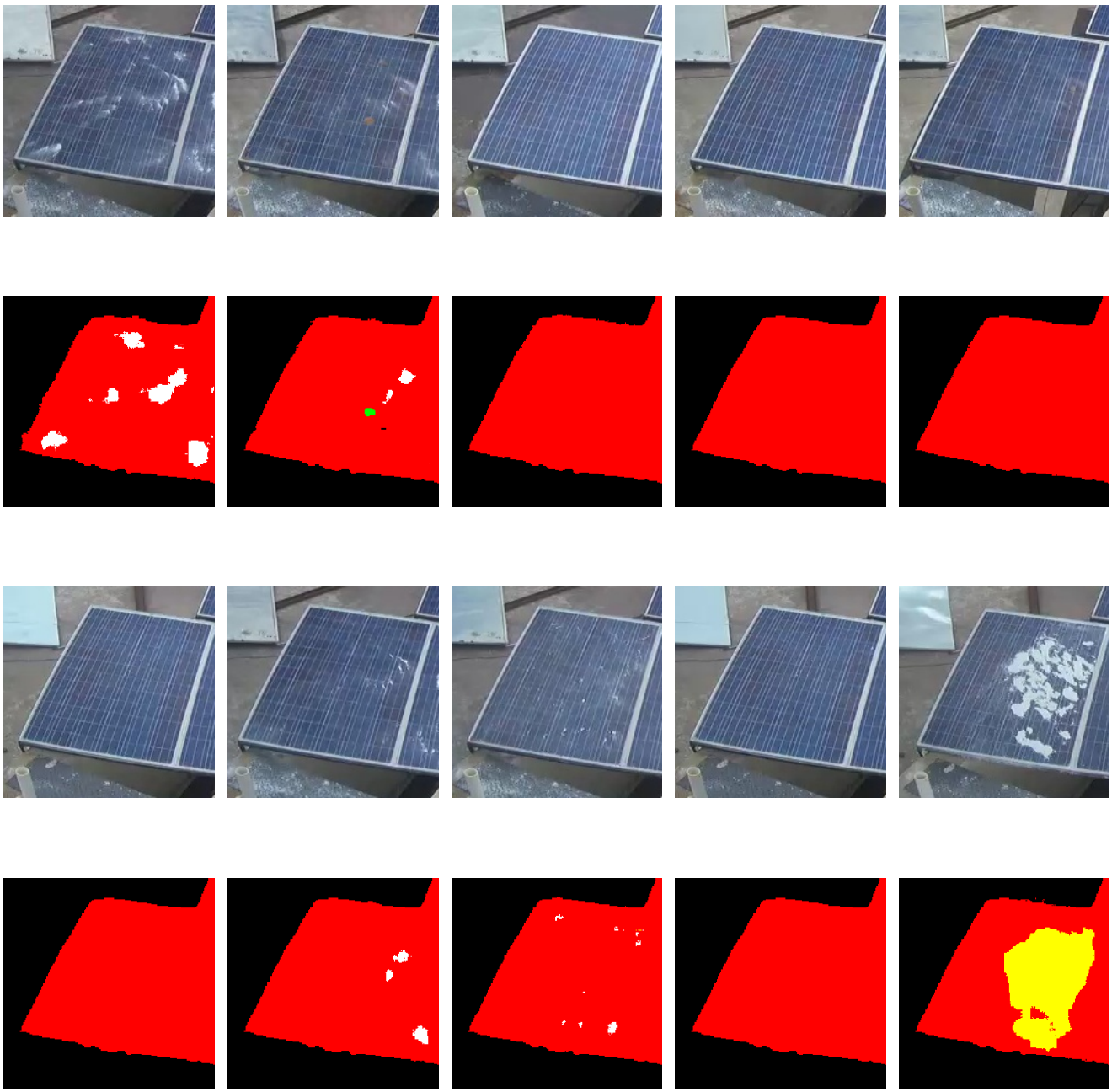


Figure 1: Prediction Examples FCN-32

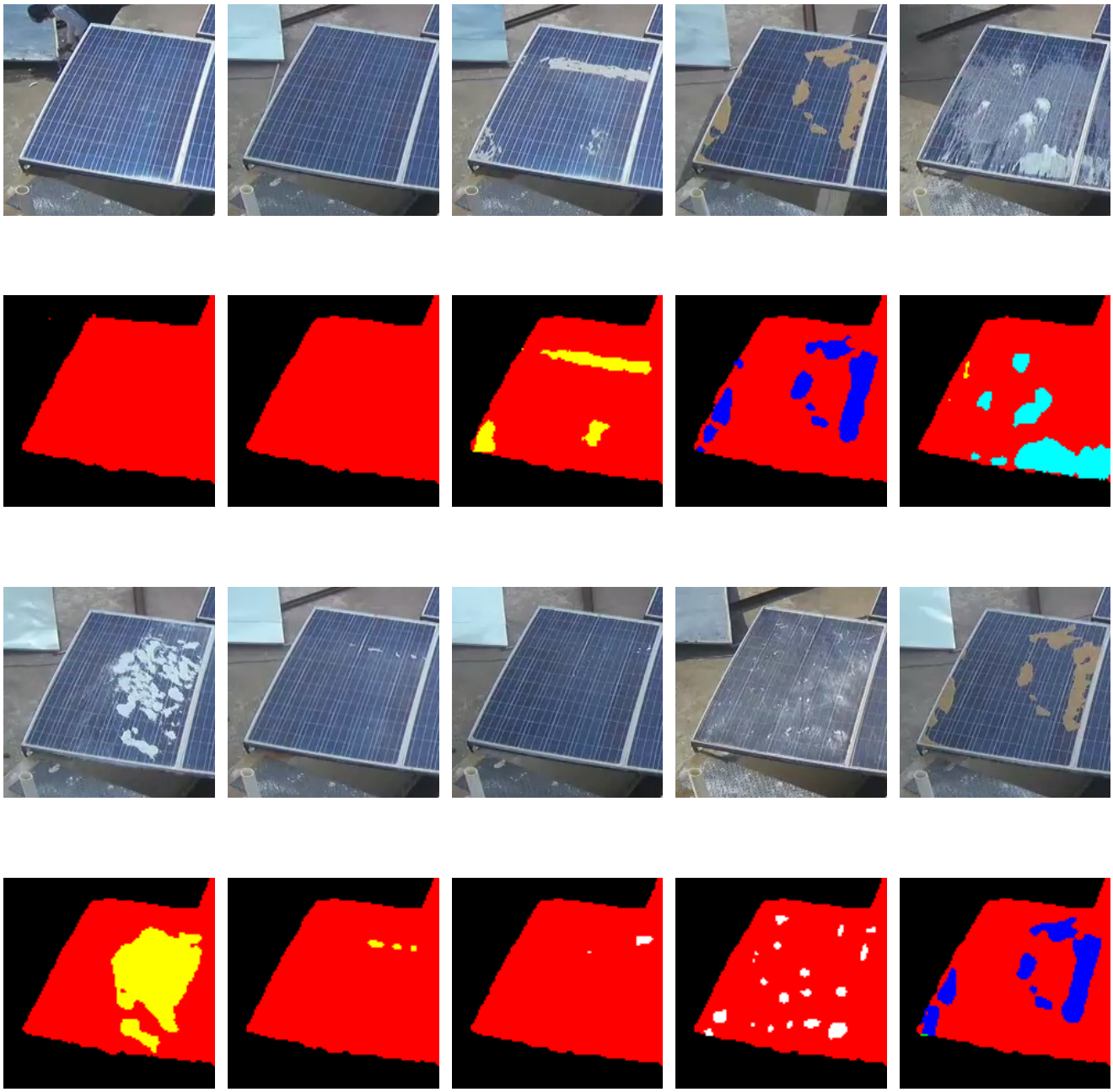


Figure 2: Prediction Examples U-Net

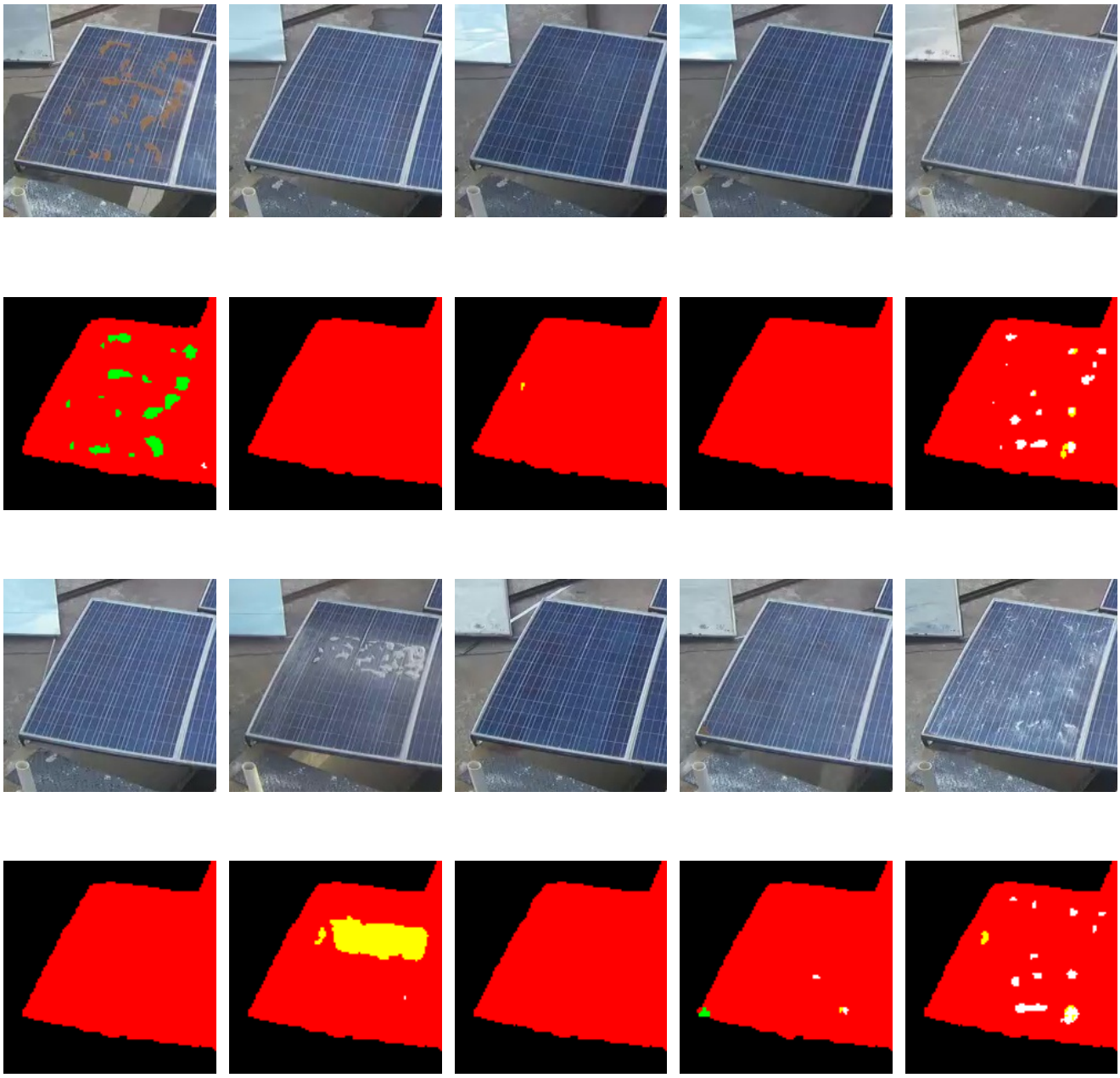
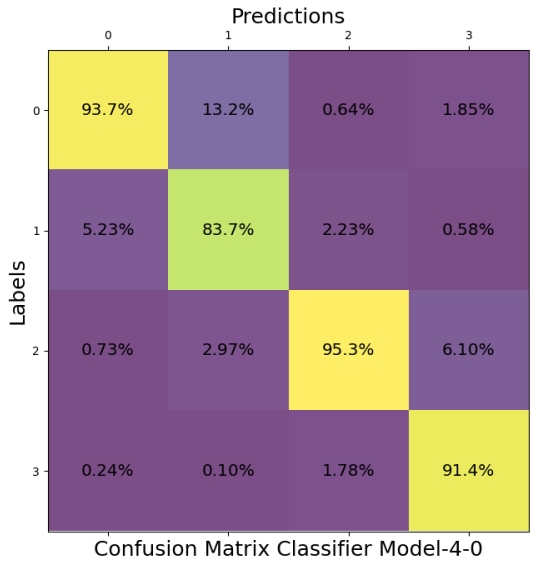
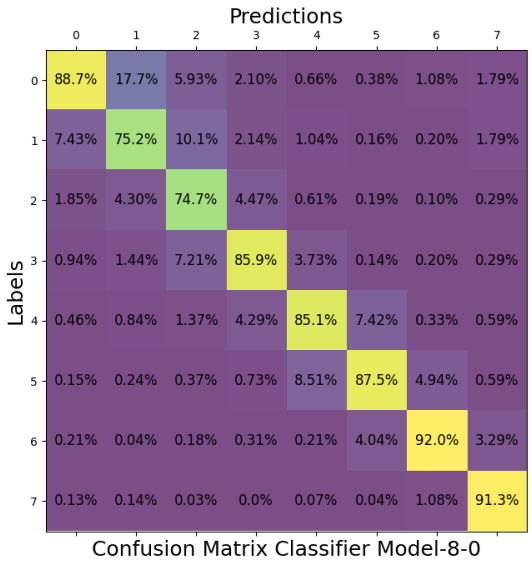


Figure 3: Prediction Examples SegNet

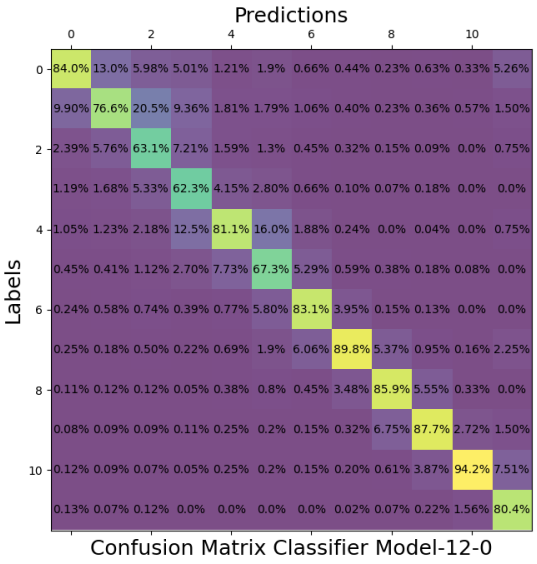
Appendix E



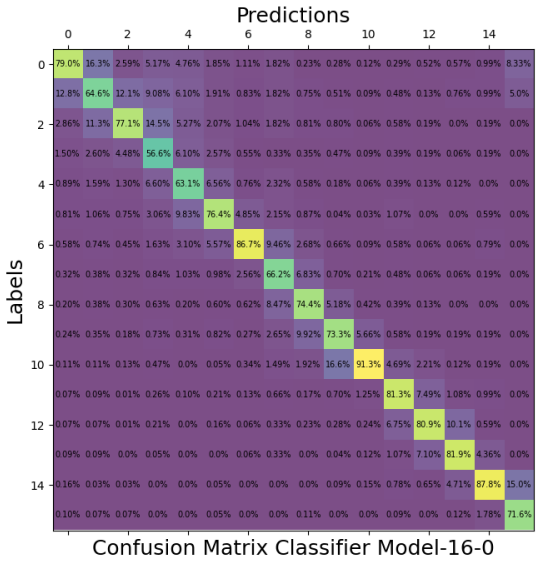
(a) Confusion Matrix Classification Model 4 Classes



(b) Confusion Matrix Classification Model 8 Classes



(c) Confusion Matrix Classification Model 12 Classes



(d) Confusion Matrix Classification Model 16 Classes

Figure 4: Confusion Matrices Classification Models