Machine Learning System for Avian-Solar Interactions: An Improved Model to Classify Bird's Activity

By

Jianhao Zhang

Thomas Harmon

Krystal Zhuo

Camille Xia

Advisor: Igor Yakushin, PhD

A Capstone Project

Submitted to the University of Chicago in partial fulfillment

of the requirements for the degree of

Master of Science in Analytics

Division of Physical Sciences

December 2022

Abstract

Current methods for monitoring avian activity at solar panel facilities are unable to capture the full complexity of avian-solar interactions. Data for this capstone is collected using 4K HD cameras placed at various solar panel sites. Videos are broken down into single images which are labeled by hand. The final training dataset is comprised of single images, an activity label, and associated metadata. This data is used to train a two-dimensional convolutional neural network which will provide single image classifications for bird activity. Previous work has been done to identify videos that contain birds, and collect metadata such bird speed, X and Y coordinates of the bird, and within image area of the bird.

Keywords: supervised machine learning, avian species, solar energy, classification, CNN, TensorFlow, Keras

Executive Summary

At present, surveyors traverse solar panel facilities counting the number of bird carcasses to provide an estimate of mortality rates. This method is labor intensive and fails to quantify other forms of avian-solar interaction. Utilizing convolutional neural networks for image classification, Argonne can decrease labor costs and gain a complete understanding of the avian impact of solar panel facilities. Data is collected via 4K HD cameras placed at solar panel facilities across geographic regions in the United States. A process is in place to determine when a bird enters camera view, track the bird, collect frame by frame metadata, and provide a single activity label to each video track. However, several activities can occur within a single track, thus Argonne now seeks to provide a multilabel to each track by analyzing the activity present in each image of the track. 74,011 images are manually labeled to create a training dataset. A two-dimensional convolutional neural network that utilizes a pretrained ResNet50 model is trained on this manually labeled dataset. The final product is a multilabel for each track based on the single labels for the images comprising the track. This research allows Argonne Research Laboratory to quantify bird activities at various solar panel facilities.

Table of Contents

List of Figures

List of Tables

Introduction

In the last decade, according to Solar Industry Research Data, solar energy has experienced an average annual growth of 33%. The infrastructure required to harness solar energy has unknown effects on surrounding wildlife, and in particular, birds. Current data collection methods are labor intensive, error prone, and provide insufficient volume to fully describe avian-solar interactions. Argonne National Laboratory aims to bridge the knowledge gap by classifying bird activities in and around solar panel facilities utilizing 4K HD cameras and machine learning. This will provide a reliable and cost-effective mechanism to assess the impact of solar energy collection on avian species.

Problem Statement

There is currently no infrastructure in place that can provide a comprehensive description of avian-solar interactions. To address this, Argonne has pioneered an end-to-end project comprised of several parts. Monitoring a sample of solar panel facilities using 4K HD video cameras. Detecting birds when they come into the camera's field of view. Classifying bird activities around solar panels and providing a multilabel to each video track based on the activity labels of its component images. Pending model performance, Argonne can create a notification system to alert facility staff of collisions. Our specific project focuses on task 3: Classifying bird activities. This research is necessary to assess the environmental impact of large-scale solar panel facilities.

Financial Benefits

The main financial benefit of this model is a decrease in labor costs. The average solar panel field in the United States is 100 acres and requires three employees for site monitoring. With an estimated base salary of $15 per hour, labor costs would be approximately $93,600 per year per 100 acres. Conversely, 30 cameras are needed per 100 acres and each camera costs $1500. This brings the total camera costs to $45,000, saving $48,600 per 100 acres. There are roughly 14,000,000 acres of solar panel filled land in the U.S. according to the conservative numbers from NREL's land use report. Thus, it is estimated that the whole project will save approximately $6.8 billion per year.

Research Goals

Our aim is to create a two-dimensional convolutional neural network that can classify images of birds into one of nine activity categories and then provide a track multilabel based on the labels of the images making up the track. The CNN inputs are single images and metadata representing the (x,y) position of the bird within the image, the speed of the bird, and the area of the bird within the image.

Scope

This model is intended to generalize for use at solar panel facilities outside of those represented in the training dataset. Argonne has made a concerted effort to include facilities from varying geographic regions. However, each site in the training dataset has unique background features present in many of the tracks such as water towers or buildings. This may decrease the generalization power of the model for certain avian activity classifications where the background is more present such as “flying over background”. It may be less of an obstacle for activity classifications where the background is less present such as “Flying over reflection” or “Sit on panel”.

Background

Theoretically, solar energy has the potential to fulfill the energy demands of the entire world with 50,000 exajoules readily harvestable (Kabir et al., 2017). Topaz Solar Farm in San Luis Obispo County, California, is the largest solar panel farm in the United States, housing more than 8 million photovoltaic solar modules over 4,700 acres and generating 550 megawatts of energy. Due to a rapid decrease in the costs associated with manufacturing solar panels, large scale solar farms are becoming more feasible (Grossmann et al., 2014). However, the implications of large solar panel fields on surrounding wildlife, particularly birds, has yet to be fully quantified. Current methods for monitoring bird behavior include surveyors traversing solar panel fields and counting the number of bird carcasses to estimate mortality. This method collects insufficient information to determine how the birds are dying and lacks the ability to quantify any other form of avian-solar interaction.

Business Partner

Argonne National Laboratory is a U.S Department of Energy multidisciplinary science and research institution founded at the University of Chicago in 1940. Argonne's Environmental Science Division (EVS) delivers pioneering research in fundamental and applied environmental science to better understand and solve global, national, regional, and local environmental challenges. Classifying avian-solar Interactions is a project funded by the U.S Department of Energy Solar Energy Technologies Office with the aim of better understanding how avian life interacts with solar panel fields. The project lead is Biophysical Scientist, Dr. Yuki Hamada and our main point of contact is Principal Software Engineer, Adam Szymanski.

Concurrent Research at Argonne

It is important to note that there is a contemporaneous capstone also in progress with Argonne National Laboratory. This project also aims to provide multilabel classifications to tracks. Our projects differ in the unit of analysis. While this project provides a track multilabel by aggregating single image classifications, the other project will provide multi-labels by analyzing the track in its entirety.

Literature Review

Two-dimensional convolutional neural networks are a leading methodology for image classifications, winning numerous computer vision competitions (Hussain et al., 2018). There are several options available for implementing a CNN architecture. Create a custom architecture, utilize a pretrained model in a process known as transfer learning, or combine both approaches to leverage the sophistication of pretrained models while adjusting them for the use case at hand (Shaha & Pawar, 2018). A variety of pretrained CNN models are available including VGG-16, ResNet50, and SE-ResNet50 (Theckedath & Sedamkar, 2020). ResNet50 is CNN with 50 layers and was pretrained on 1.28 million images from the ImageNet database (Deep Residual Learning for Image Recognition). ImageNet contains over 1000 classes of images including a variety of bird species, making it an ideal choice for our use case (Fox, E., & Guestrin, C. (n.d.). Coursera Machine Learning Specialization.

The training dataset provided by Argonne contains large class imbalances. Traditional mitigation strategies such as synthetic oversampling have limited success due to the complex data structures of deep learning models. Deep over-sampling methodologies build upon traditional over-sampling approaches, proving more successful in dealing with class imbalance in convolutional neural networks (Ando & Huang, 2017).

Goals of Analysis

The goal of this research is to design a model framework that can assign single image inputs into one of nine activity categories. First, a training dataset must be created as the data provided is not suitable for this problem in its raw form. While each track has a label, this label is not necessarily the label for each frame within the track. Therefore, a substantial amount of time must be spent relabeling individual frames from tracks to create a training dataset. This is accomplished using Pigeon API. The activity labels are as follows:

- Fly over sky – bird flying with a clear sky background and no other buildings, trees, or facilities, etc.
- Fly over other backgrounds - bird flying with backgrounds other than sky, solar panel, and ground
- Fly over reflection - bird's reflections projected on the solar panel but no actual bird appearing in the frame
- Fly with solar panel - bird flying with solar panel in the background
- Fly under solar panel - bird flying with a position under solar panel
- Fly with ground - bird flying with ground as a background
- Sit on panel – a bird perching on the solar panel
- Sit on the ground – a bird perching on the ground
- Sit in background - a bird perching on backgrounds other than solar panel and ground

Once the images are labeled, they are then divided into different activity groupings. Using image and metadata inputs, a 2-dimensional convolutional neural network is built to classify single images. Lastly, the model is tested and validated on unseen data, and multi-labels are assigned to tracks based on the labels of the tracks constituent frames.

Data

Data Source

Argonne provides 9000 hours of daytime 4K HD video from various solar panel facilities and associated metadata. Each video clip is 5-minutes and represents a unique track of the moving object. Using the existing object detection and tracking algorithm, the video footage is processed, extracted, and cropped to form groups of images as the object moves through a camera's field of view. This data is stored in the midway 2 cluster at RCC. Data is organized via a csv file containing a directory path, label, and metadata for each frame as shown in Figure 1.

Figure 1. *Sample Track Information*

	Directory	Label	x	y	speed	area	obj_id	frame	image_count
0	/project2/msca/projects/AvianSolar/ImageDatase...	Flying with solar panel	832	1134	66.79106	12769	40273	2184	28.0
1	/project2/msca/projects/AvianSolar/ImageDatase...	Flying with solar panel	1692	1330	89.06198	13447	40273	2196	28.0
2	/project2/msca/projects/AvianSolar/ImageDatase...	Flying under solar panel	2988	1787	147.78356	13340	40273	2207	28.0
3	/project2/msca/projects/AvianSolar/ImageDatase...	Flying over reflection	1598	1310	84.22580	13221	40273	2195	28.0
4	/project2/msca/projects/AvianSolar/ImageDatase...	Flying under solar panel	3149	1856	161.16776	12656	40273	2208	28.0

It is important to note that the existing label is not suitable for our project. This label is from a prior Argonne project where the unit of analysis was the entire track. Thus, this track label is not necessarily the label for each image contained in the track. A total of

457,687 unlabeled images are available. A single image has the shape of 200 pixels by 200 pixels with 3 color (RGB) channels. Each pixel value ranges from 0 to 255. Metadata for each image has the following features:

- Bird’s X-axis coordinate in the 2D video frame
- Bird’s Y-axis coordinate in the 2D video frame
- Bird’s relative speed at each coordinate
- Bird’s relative size at each coordinate

Table 1. *Sample Feature Inputs*

Track ID	Image ID	X	Y	Speed	Area	Label
1	1	3600	489	20.87	12430	Fly over above
1	2	3581	486	20.86	12430	Fly over above
1	3	3545	484	20.84	12210	Fly through
...	...	...	...	...	...	...

Descriptive Analysis

74,011 images are manually labeled with label distributions shown in Table 2. Most of the image labels are flying over sky, flying over other background, flying with solar panel, and flying over reflection. This imbalanced dataset will lead the model to predict well on the majority labels but not on the minority labels. Thus, under-sampling and over-sampling techniques are desired to create a balanced dataset.

Table 2. *Label Distribution*

Label	Count	Percentage
Fly over sky	34080	46.05%
Fly over other background	12382	16.73%
Fly with solar panel	10449	14.12%
Fly over reflection	9538	12.89%
Fly with ground	2541	3.43%
Sit on panel	2203	2.98%
Fly under solar panel	1389	1.88%
Sit in background	1039	1.40%
Sit on ground	390	0.53%

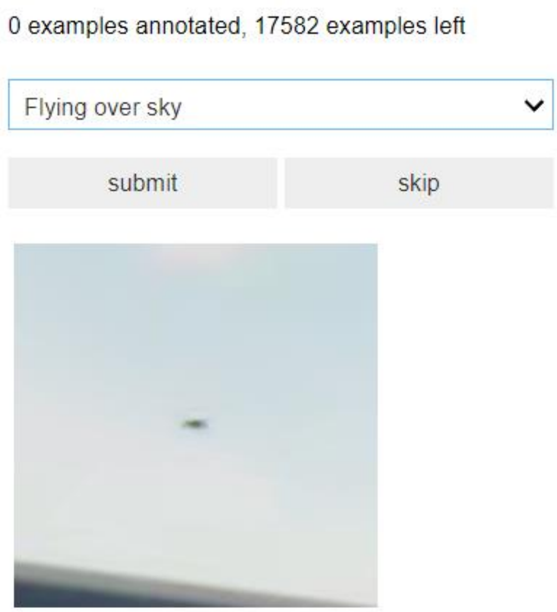
Methodology

Relabeling

The traditional approach of assigning one label for an entire track does not capture whenever multiple activities are present in one track. Our approach is to process image by image, changing the unit of analysis from a track to a single image.

Pigeon API is utilized to manually label images. It takes image path and label options as inputs and displays the image within Jupyter Notebook. When a label is chosen, the image path along with the label are stored in a tuple and the next image will show. The images are displayed in order within each track. This approach is more efficient as images in series tend to be similar. The labels, image path and frame metadata are then merged and stored.

Figure 2. *Pigeon API Demonstration*



Feature Engineering

In this project, bird's activity labels are imbalanced. Training a classification model on imbalanced data will cause the model to be biased, leading to high predictive accuracy for the majority classes and poor predictive power on the minority classes. To address the data imbalances, different activities groups are created by merging similar labels into a single label as shown in Table 3. In this way, the minority class can be joined with other classes to slightly relieve the imbalance and the classification model's performance can be examined with different label groups to better understand the capability and limitation of the model. Under-sampling using random sampling without replacement is also employed to further balance the newly merged labels.

Table 3. *Activity Group*

Activity Group	Group Label
Original	• Fly over sky, Fly over other backgrounds, Fly with solar panel, Fly over reflection, Fly under solar panel, Fly with ground, Sit on panel, Sit in background, Sit on the ground
Group 1	• Flying (Fly over sky, fly over other backgrounds, fly with solar panel, fly over reflection, fly under solar panel, Fly with ground) • Non-Flying (Sit on panel, sit in background, sit on the ground)
Group 2	• With panel (Flying with solar panel, flying over reflection, sit on panel, flying under solar panel) • Without panel (Flying over sky, flying over other backgrounds, flying with shadow on the ground, sit in background, sit on the ground)
Group 3	• Background (Flying over sky, flying over other backgrounds, sit in background) • Panel (Flying with solar panel, flying over reflection, sit on panel, flying under solar panel) • Ground (Flying with shadow on the ground, sit on the ground)
Group 4	• Background (Flying over sky, flying over other backgrounds, sit in background) • Fly with panel (Flying with solar panel, flying over reflection, flying under solar panel) • Interact with panel (Sit on panel) • Ground (Flying with shadow on the ground, sit on the ground)

Label encoding is used to convert categorical labels to integer labels, and then one-hot encoding is used to convert the integer labels into a vector representation. The size of the vector varies depending on the number of labels in each label group. This transformed vector label becomes the target variable for the classification model. In addition, the metadata (x, y, position, speed) is also standardized.

Data Pipeline

Due to memory constraints, it is infeasible to read all images into pixels and stack them in a single NumPy array. Therefore, a functional TensorFlow API, tf.data, is used to build out the data pipeline following the ETL structure. This API reads images from files and feeds the model in batch sizes without leveraging RAM to store the data. This data pipeline extracts images from Midway2 cluster at RCC, metadata, and labels from the csv file. The images are then decoded into pixel arrays in the shape of (200, 200, 3) and resized into (224, 224, 3) using zero padding to fit the default input shape of Resnet 50. Resnet 50’s input preprocessing function is used to adjust the images to the required format. Finally, all pixel arrays, metadata, and labels are converted into tensor objects to be fed into the model. Cache, prefetch, and batch techniques are utilized to increase the speed and efficiency of data ingestion during training stages.

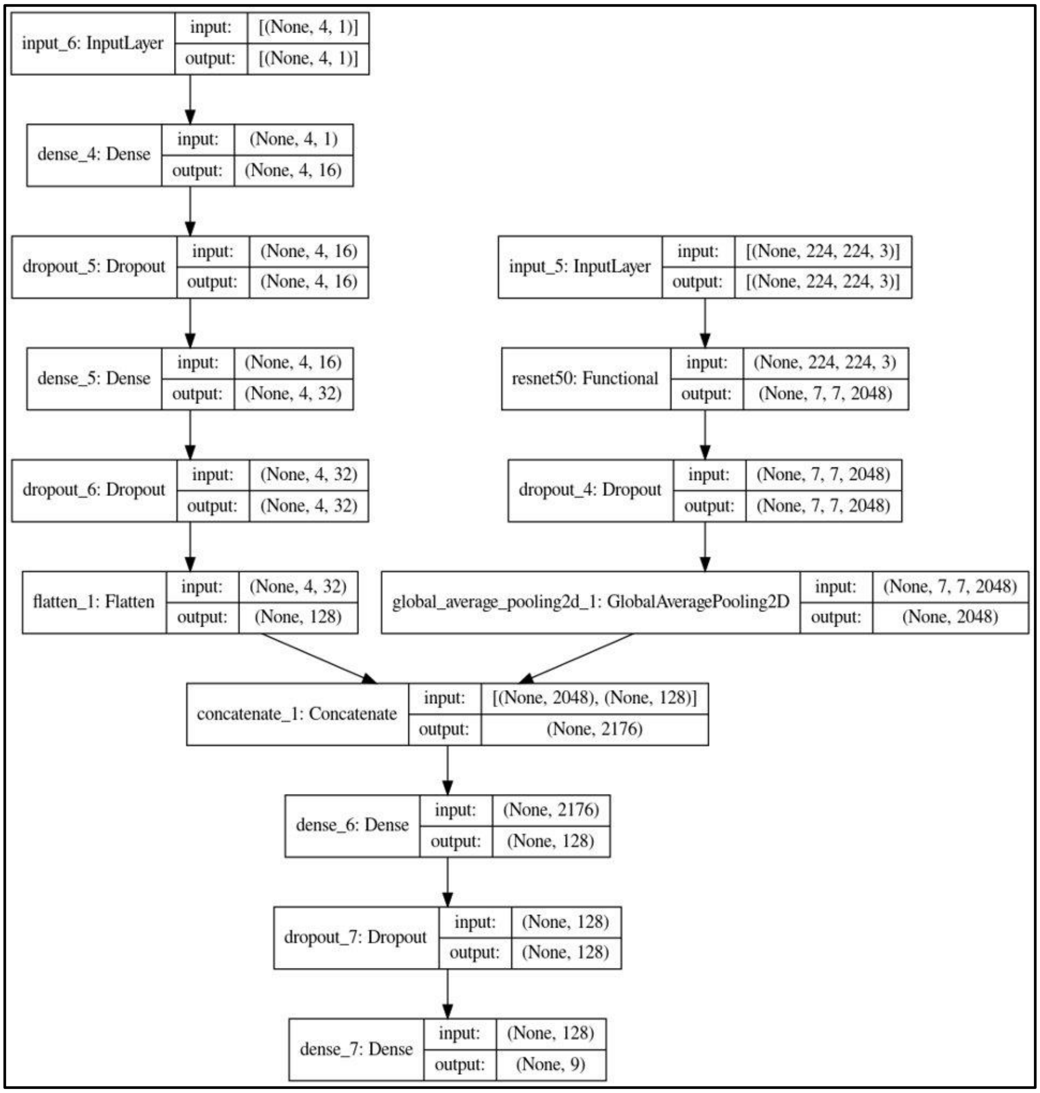
Training/Validation/Test Split

Data is split into train, test, and validation in an 80:10:10 ratio. Two options are available to split the dataset, either by image ID or track ID. Images within tracks can be nearly identical. Therefore, splitting by image ID will result in nearly identical images falling into all of training, validation, and test sets. This method will produce misleadingly high accuracy and have poorly predictive power on new data. However, splitting by track ID will assign all images in a track to one of the training, validation, and test sets. This prevents data leakage and results in a model with greater generalizing power.

Model Framework

A 2-Dimensional Convolution Neural Network (CNN) is chosen as the image classification algorithm. It takes images and metadata as combined inputs and outputs a predicated activity label of each image. Figure 3 describes the model framework with layers, input, and output dimensions of each layer specified. Transfer Learning is adopted to mitigate the effects of insufficient amount training data as well as to save training time. Resnet 50, which has been pretrained on over 1 million images and can predict over 1000 object categories, is used as the first branch of the CNN model. It is composed of 48 convolutional layers along with 1 max pooling and 1 average pooling layer. The residual learning block in Resnet 50 has a skip connection feature, which allows the layer inputs to learn an identity function to prevent vanishing gradient and ensure higher layers perform better than the lower layers. The output of Resnet 50 has dimensions of (7, 7, 2048). This is then connected to a dropout layer and a 2D global average pooling layer. The second branch of the CNN consists of multiple dense layers, a dropout layer, and a flatten layer to fit the metadata inputs. These two branches are then concatenated together into a fully connected layer to output the predicted results.

Figure 3. *CNN Model Layers*



Categorical accuracy and categorical loss entropy are used to evaluate the model performance for the training, validation, and test set as shown in table 4. In addition, learning curves for both categorical accuracy and categorical loss entropy against training epochs are shown in the Appendix. Overfitting exists when training for Original Group, Group 1 and Group 2 as suggested by the gaps between train accuracy and validation accuracy. On the other hand, train and validation accuracy for Group 3 and Group 4 seem to converge during the training and validation stage.

Table 4. *Model Validation Results*

	Train-Loss	Train-Accuracy	Val-Loss	Val-Accuracy	Test-Loss	Test-Accuracy
Original	0.72	76.09	1.115	67.78	0.84	71.38
Group 1	0.14	94.90	0.387	84.12	0.27	88.44
Group 2	0.19	92.31	0.284	88.41	0.19	92.84
Group 3	0.10	96.57	0.174	94.79	0.19	94.70
Group 4	0.37	86.21	0.322	87.67	0.49	81.65

Findings

To examine the generalization power of the model, the full unbalanced dataset is used as the test dataset for the trained model to predict. Accuracy, precision, recall, and f1 score are accessed and discussed for all the label groups. The confusion matrix and classification report for all the label groups are also provided in the Appendix.

The model achieved its best performance on group 3 (Background vs Panel vs Ground) with accuracy of 0.94 and weighted average f1 score of 0.94 and its lowest performance on the original 9 labels with accuracy of 0.75 and weighted average f1 score of 0.77. The CNN model takes into account 150,528 pixels (224 x 224 x 3) for each image, a majority of which describe the image's background. Thus, when predicting labels, the model relies mainly on recognizing the background present in the image. Group 3 has the most distinct set of backgrounds and thus has the highest performance metrics. However, when the model attempts to predict the full 9 labels, classes such as “sit on panel”, “fly over reflection” and “fly with solar panel” are inaccurately predicted due to all images having solar panels as the main background.

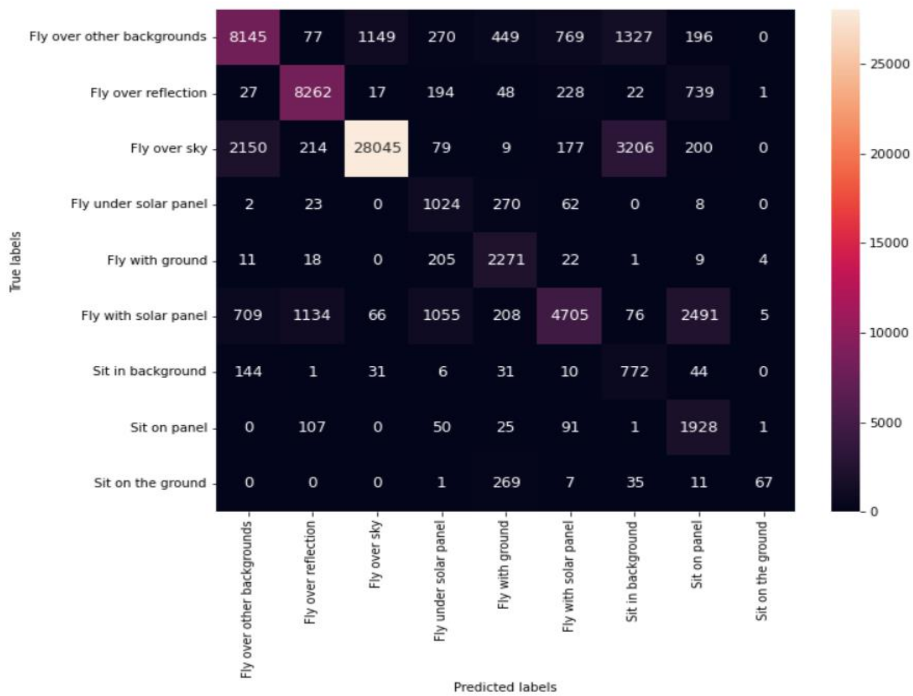
The binary label groups, Group 1 and Group 2, also underperform in comparison to Group 3 with accuracy of 0.85 and 0.89 and weighted average f1 score of 0.89 and 0.89. The flying category in Group 1 consists of “fly over sky”, “fly with solar panel”, and “fly with ground”, while the non-flying category in Group 1 consists of “sit in background”, “sit on panel”, and “sit on the ground”. Labels “fly over sky” and “sit in background” have similar backgrounds leading to lower performance in this grouping. Performance of Group 2 is slightly better than Group 1 as it predicts whether a solar panel exists in the image.

Table 5. *Model Performance Metrics by Label Grouping*

	Accuracy	Precision – MA	Precision – WA	Recall – MA	Recall – WA	F1 – MA	F1 – WA
Original	0.75	0.63	0.83	0.69	0.75	0.58	0.77
Group 1	0.85	0.62	0.96	0.89	0.85	0.64	0.89
Group 2	0.89	0.87	0.92	0.92	0.89	0.88	0.89
Group 3	0.94	0.82	0.95	0.95	0.94	0.87	0.94
Group 4	0.87	0.71	0.92	0.85	0.87	0.74	0.88

Figure 4 describes the model’s classification on the full 9 labels. Most of the misclassifications occur between classes where the backgrounds are similar ex. “Sitting on panel” vs “flying with panel” vs “flying over reflection”. Since these classes have very similar backgrounds, the model has difficulty differentiating them. In a similar manner, a majority of the “sit on ground” cases are predicted as “fly with ground”.

Figure 4. *Confusion Matrix for Original Labels*



Discussion

There are several limitations and routes for improving the current methodology and findings. First, when manually labeling, humans consider the temporal nature of images, using knowledge of previous images to inform labeling choices for subsequent ones, whereas the CNN model cannot. Second, the training dataset was not exhaustive. There were 457,687 images available, however, due to time constraints only 74,011 were manually labeled. To address data scarcity and class imbalances, image augmentation is explored as a method to up-sample minority classes. However, if an image is transformed, its metadata must be manipulated in the same degree. Pending an appropriate pipeline for image and metadata augmentation, the training dataset could be increased, and class imbalances more thoroughly addressed. Third, the model takes 150,528 pixels as inputs, within image area of the bird, speed, x position, and y position.

A majority of the pixels, and thus a majority of the model inputs, pertain to the image background. However, the speed of the bird is equally important and could aid in differentiating between classes with similar backgrounds such as “fly over panel” and “sit on panel”. By tuning the weight of each model input the model could achieve better class separation. Lastly, there is also no set method in place for transforming single image labels into multi-labels. Available options include a majority vote, proportion of track spent performing each activity, or a list of each label as they occur in the track.

Conclusion

The 2-dimensional convolutional neural network can predict group 3 labels (background vs panel vs ground) with sufficient accuracy to begin implementation. All Python scripts and manually labeled data are available to Argonne via a private GitHub repository. The information in this report allows Argonne to decide which label groupings will have the most significant business impact while balancing model performance metrics.

Recommended next steps for Argonne are to begin using the CNN to predict group 3 labels. In this manner, Argonne can narrow their video footage to tracks where birds interact with solar panels. Further, given additional training data, the model architecture described in this report can be retrained to differentiate between activity classes where birds interact with solar panels.

Reference

Grossmann, Wolf D., et al. "Solar Electricity Generation across Large Geographic Areas, Part II: A Pan-American Energy System Based on Solar." *Renewable and Sustainable Energy Reviews*, vol. 32, 2014, pp. 983–993.

Kabir, Ehsanul, et al. "Solar Energy: Potential and Future Prospects." *Renewable and Sustainable Energy Reviews*, Pergamon, 6 Oct. 2017

Law.lclark.edu. "100 Megawatts per Day: Solar Power on the Rise." • *Green Energy Institute • Lewis & Clark*, Law.lclark.edu, 3 Nov. 2014

Topaz Solar Farms. Edited by Paulette Rush, *BHE Renewables*, https://www.bherenewables.com/topaz_solar.aspx.

Oberhaus, Daniel. "Why Do Solar Farms Kill Birds? Call in the AI Bird Watcher." *Wired*, Conde Nast, 10 Aug. 2020, https://www.wired.com/story/why-do-solar-farms-kill-birds-call-in-the-ai-bird-watcher/.

Ramyachitra, D. and Parasuraman Manikandan. "IMBALANCED DATASET CLASSIFICATION AND SOLUTIONS: A REVIEW." (2014).

Stefan Kahl, Connor M. Wood, Maximilian Eibl, Holger Klinck, "BirdNET: A deep learning solution for avian diversity monitoring", Ecological Informatics, Volume 61, 2021, 101236, ISSN 1574-9541, https://doi.org/10.1016/j.ecoinf.2021.101236.

Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, W. Philip Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique", Journal of Artificial Intelligence Research, vol.16, 2002, pp.321–357

H. Roberts and A. Segev, "Animal Behavior Prediction with Long Short-Term Memory," 2020 IEEE International Conference on Big Data (Big Data), 2020, pp. 4157-4164, doi: 10.1109/BigData50022.2020.9378184.

Sener, Ozan, and Silvio Savarese. “Active Learning for Convolutional Neural Networks: A Core-Set Approach.” ArXiv.org, 1 June 2018, https://arxiv.org/abs/1708.00489.

“Solar Industry Research Data.” SEIA. Accessed October 28, 2022. https://www.seia.org/solar-industry-research-data.

M. Shaha and M. Pawar, "Transfer Learning for Image Classification," *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2018, pp. 656-660, doi: 10.1109/ICECA.2018.8474802.

Theckedath, D., Sedamkar, R.R. Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks. *SN COMPUT. SCI.* **1**, 79 (2020). https://doi.org/10.1007/s42979-020-0114-9

Ando, S., Huang, C.Y. (2017). Deep Over-sampling Framework for Classifying Imbalanced Data. In: Ceci, M., Hollmén, J., Todorovski, L., Vens, C., Džeroski, S. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2017. Lecture Notes in Computer Science, vol 10534. Springer, Cham. https://doi.org/10.1007/978-3-319-71249-9_46

Hussain, Mahbub, et al. *A Study on CNN Transfer Learning for Image Classification*, Springer.

Appendix

Figure A 1. *Original Confusion Matrix*

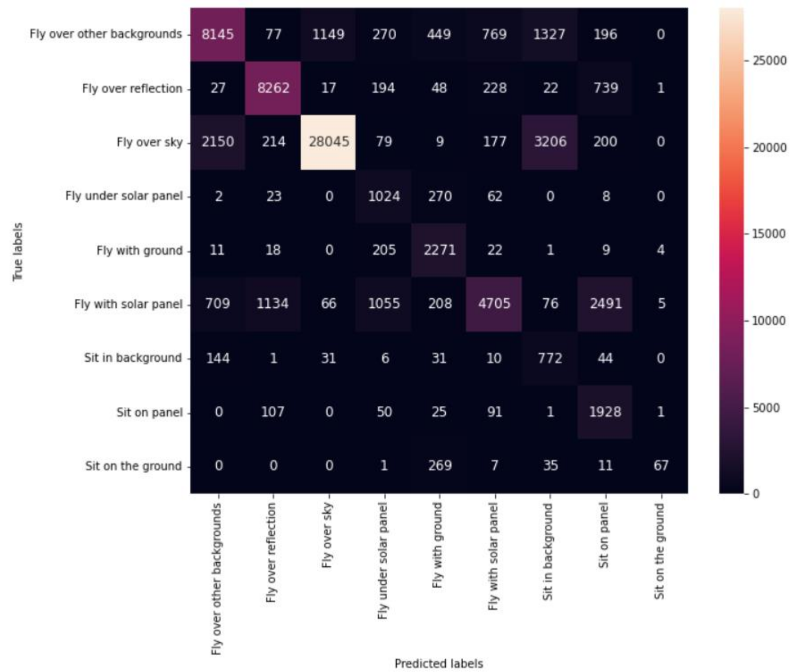


Figure A 2. *Group 1 Confusion Matrix*

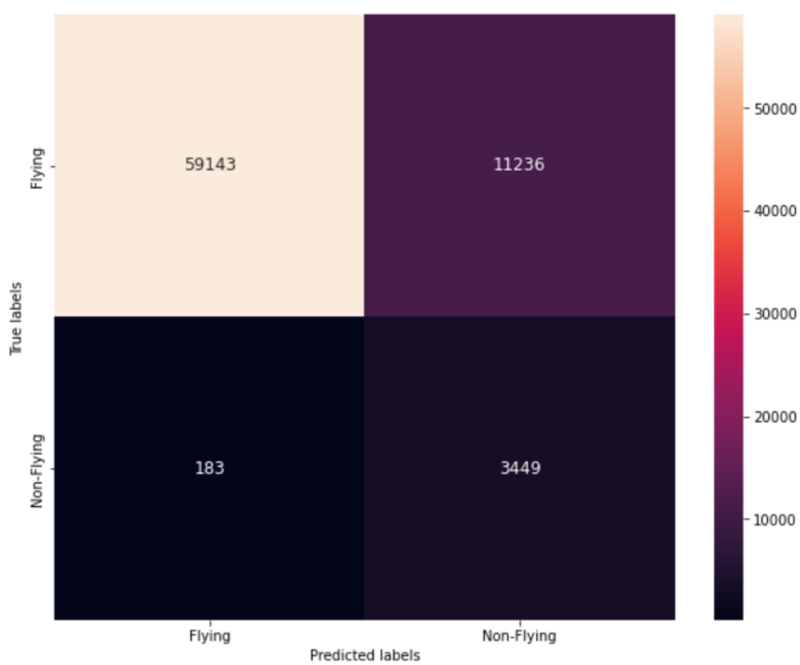


Figure A 3. *Group 2 Confusion Matrix*

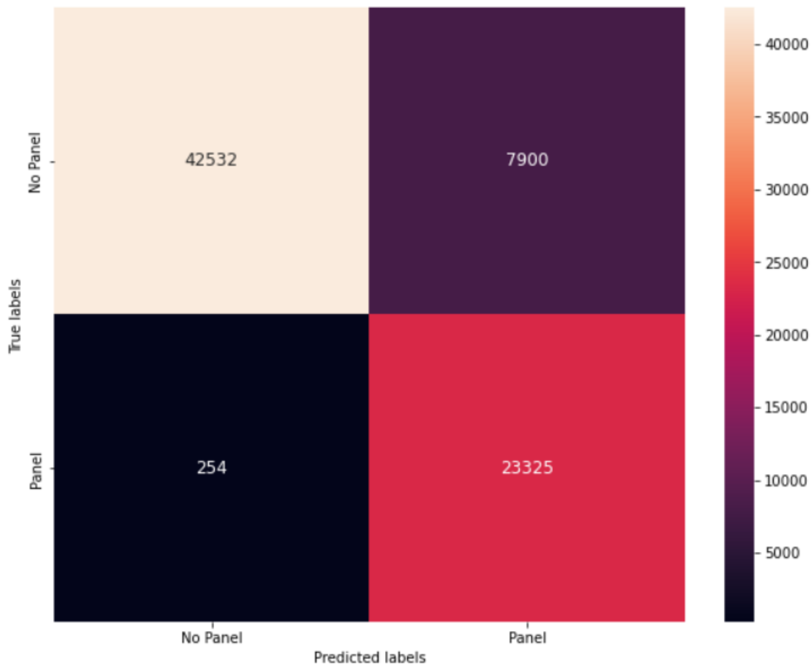


Figure A 4. *Group 3 Confusion Matrix*

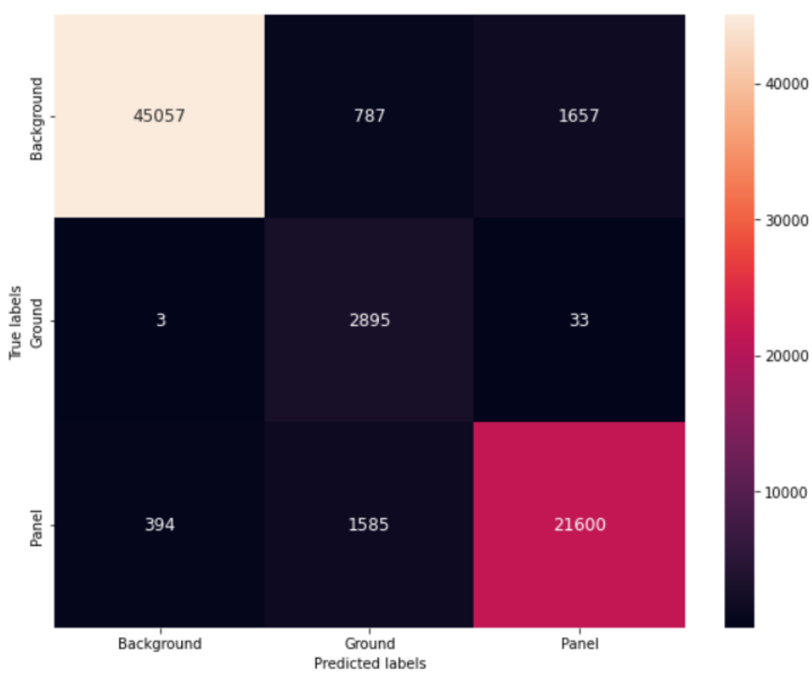


Figure A 5. *Group 4 Confusion Matrix*

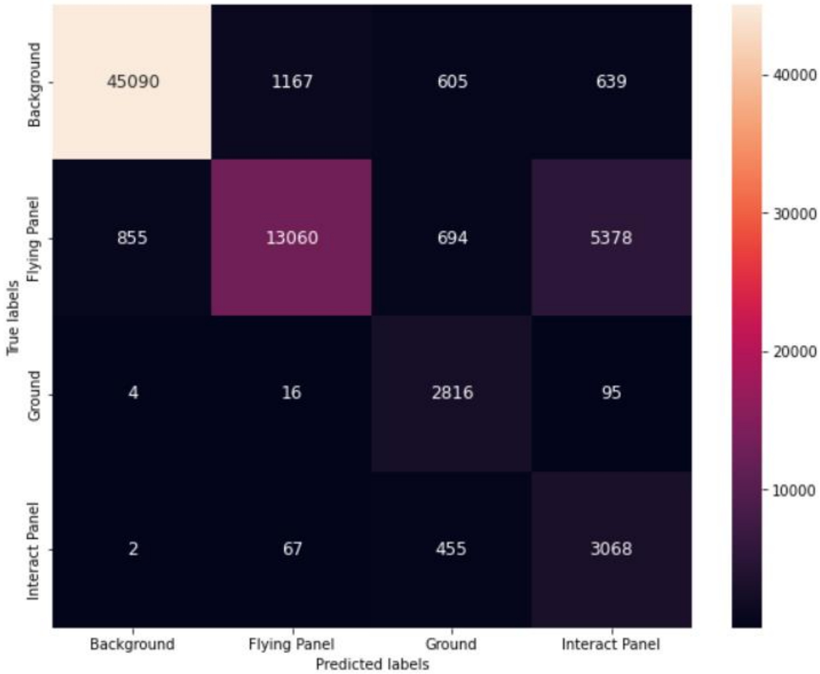


Figure A 6. *Original Learning Curve*

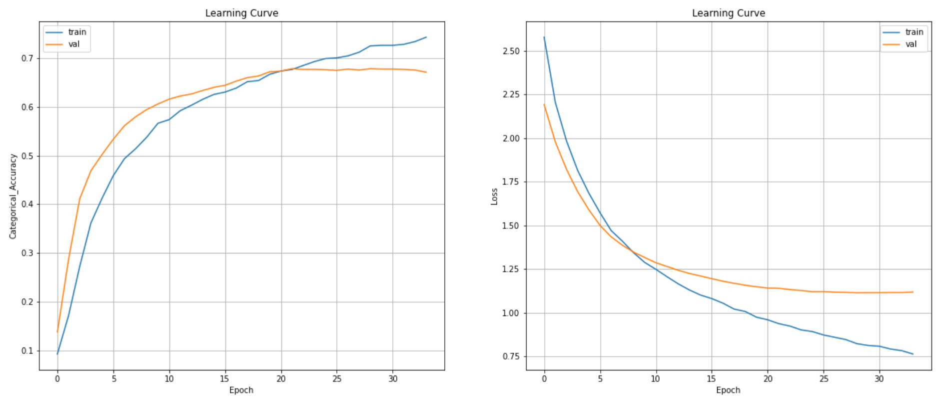


Figure A 7. *Group 1 Learning Curve*

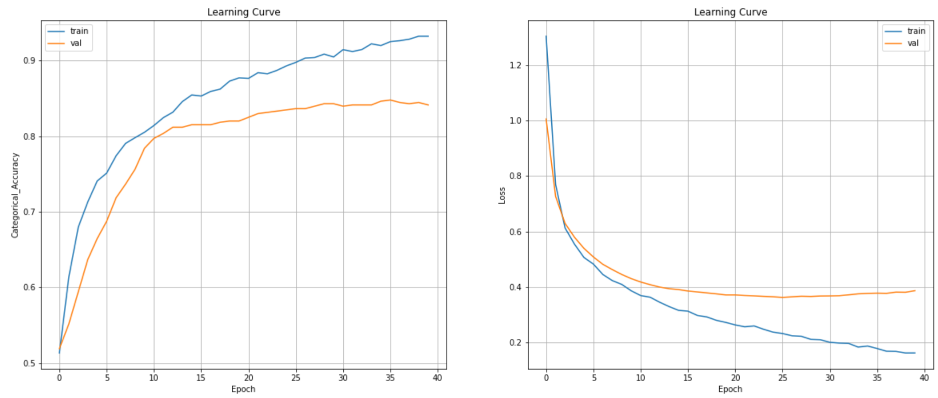


Figure A 8. *Group 2 Learning Curve*

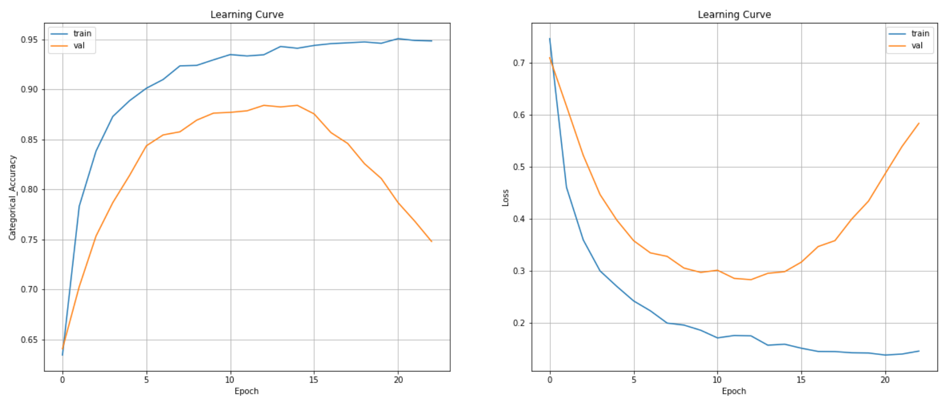


Figure A 9. *Group 3 Learning Curve*

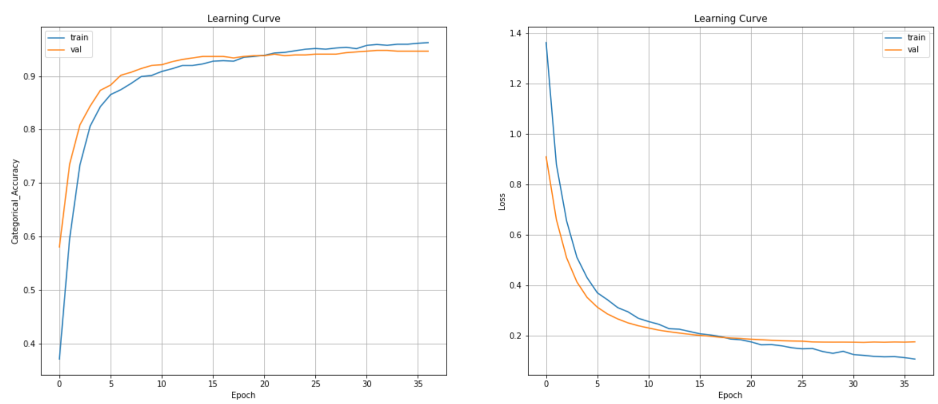


Figure A 10. *Group 4 Learning Curve*

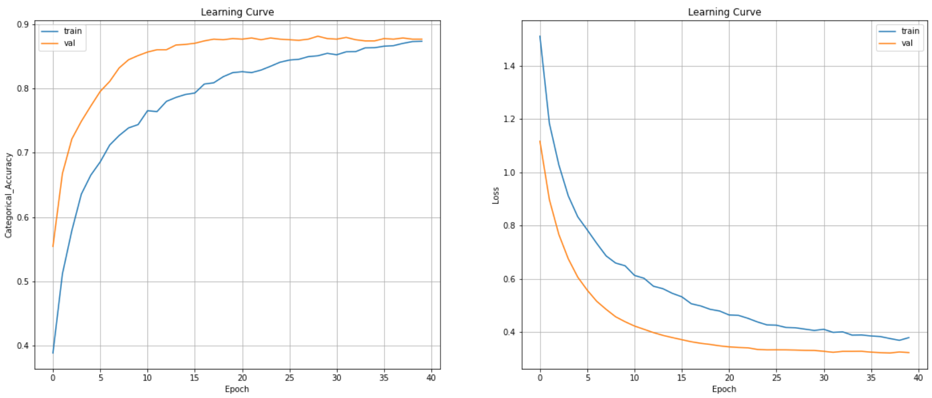


Figure A 11. *Original Classification Report*

	precision	recall	f1-score	support
0	0.73	0.66	0.69	12382
1	0.84	0.87	0.85	9538
2	0.96	0.82	0.88	34080
3	0.36	0.74	0.48	1389
4	0.63	0.89	0.74	2541
5	0.77	0.45	0.57	10449
6	0.14	0.74	0.24	1039
7	0.34	0.88	0.49	2203
8	0.86	0.17	0.29	390
accuracy			0.75	74011
macro avg	0.63	0.69	0.58	74011
weighted avg	0.83	0.75	0.77	74011

Figure A 12. *Group 1 Classification Report*

	precision	recall	f1-score	support
0	1.00	0.84	0.91	70379
1	0.23	0.95	0.38	3632
accuracy			0.85	74011
macro avg	0.62	0.89	0.64	74011
weighted avg	0.96	0.85	0.89	74011

Figure A 13. *Group 2 Classification Report*

	precision	recall	f1-score	support
0	0.99	0.84	0.91	50432
1	0.75	0.99	0.85	23579
accuracy			0.89	74011
macro avg	0.87	0.92	0.88	74011
weighted avg	0.92	0.89	0.89	74011

Figure A 14. *Group 3 Classification Report*

	precision	recall	f1-score	support
0	0.99	0.95	0.97	47501
1	0.55	0.99	0.71	2931
2	0.93	0.92	0.92	23579
accuracy			0.94	74011
macro avg	0.82	0.95	0.87	74011
weighted avg	0.95	0.94	0.94	74011

Figure A 15. *Group 4 Classification Report*

	precision	recall	f1-score	support
0	0.98	0.95	0.96	47501
1	0.91	0.65	0.76	19987
2	0.62	0.96	0.75	2931
3	0.33	0.85	0.48	3592
accuracy			0.87	74011
macro avg	0.71	0.85	0.74	74011
weighted avg	0.92	0.87	0.88	74011