Computational Imaging For Long-Term Solar Irradiance Forecasting

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Department of Electrical and Computer Engineering

Leron K. Julian

B.S., Computer Science, Morehouse College

M.S., Electrical & Computer Engineering, Carnegie Mellon University

Carnegie Mellon University

Pittsburgh, PA

August 2024

Acknowledgments

I would first like to thank God for the blessings and opportunities given to me thus far in my life.

Thank you to my mother Judy Julian, my grandmother Philippia George, and my family for the constant support and love while I stayed in school for all these years. The random phone calls, food, and goodies that they sent to me will never go unappreciated. Without them, I would not be where I am today.

Thank you to my friends who kept me sane while living in Pittsburgh. My brain really needed a break from research sometimes and these guys really helped. Natalia, Akeem, David, Claude, NyJaee, Moses, Brandon, and the many others I interacted with during my time here! Thank you to my girlfriend, Erin, for the unconditional love, support, and inspiration.

Academically, I will forever be grateful to those who have mentored and guided me throughout my scholastic journey. It's amazing to look back and realize how much I did not know when I thought I knew it all. Thank you to my advisor, Aswin Sankaranarayanan, who allowed me to join his amazing lab and gain a rich understanding about the world of vision. Aswin will never seek to amaze me with how brilliant of a person he his while balancing such a busy life! Thank you to my committee: Ashok Veeraraghavan, Soummya Kar, and Deva Ramanan. I would also like to thank Prof. Shawn Blanton who connected me with Aswin and CMU. Thank you to my fellow lab-members – Rick, Anqi, Michael, Byeongjoo, Wei-Yu, Yi, Vishwanath, Vijay, Yash, Jean-Michael, Jeremy, Carlos, Yingsi, Haejoon, Sagnik, Tyler, Natalie – whose brilliant minds are such an inspiration. Grateful to have had the opportunity to work with them all these years!

I could never forget the person who had the vision of me pursuing graduate school - Dr. Brock Mayers, who since high school, saw a vision of who I could be and guided me to pathways to reach that vision. Grateful to say that I made it. Thank you to my undergraduate advisor Dr. Kinnis Gosha, who allowed me to gain the basis and interest to do research.

Lastly, I would like to acknowledge my friend and brother, Tyrone Holmes who is not with us today. I wish you were here to share this momentous moment with me but you will always be remembered.

I would also like to thank the National Science Foundation under the CAREER award CCF-1652569, the EAGER award 2235063, the GEM Fellowship and Fritsch Family Fellowship for assisting in funding my PhD.

Abstract

The intermittency of solar energy, due to occlusion from cloud cover, is one of the key factors inhibiting its widespread use in commercial, residential, and utility-scale settings. Hence, real-time forecasting of solar irradiance for grid-connected photovoltaic (PV) systems is necessary to mitigate voltage fluctuations, limited time to adjust between energy sources and ultimately, energy disruptions.

Images of the sky provide rich context of cloud patterns around a localized PV site and if these images are captured in sequence, provide a copious amount of data for inference. The stochastic nature of the trajectory of cloud patterns are very difficult to forecast which in-turn makes foresight into when solar energy will decrease, difficult. However, by leveraging learning-based frameworks coupled with other sources of data such as global horizontal irradiance (GHI), we open the world of solar irradiance forecasting to the field of computational imaging to increase forecasting accuracy and present an exciting advance to state-of-the-art methods.

Limitations of traditional solar irradiance forecasting methods using sky images initially stem from limitations of the imaging systems themselves. Existing imagers provide non-uniform spatial resolution of the sky with a higher detail and resolution at the zenith and significantly lower-resolution near the horizon – severely limiting long-term prediction. To that end, we make the following contributions to the theory, hardware, and algorithms of computational imaging for long-term solar irradiance forecasting are made.

Initially, we present a learning-based framework that forecasts future sky image frames with higher precision than previous methods. Our key contribution within this work is the derivation of an optimal warping algorithm that counters the adverse effects of non-uniform spatial resolution present in traditional sky imagers. We show that by warping these images to a new space, the model more accurately determines cloud evolution for longer time horizons. Secondly, we present a catadioptric imaging system that maintains wide-angle imagery and uniform spatial resolution of the sky without the need of any warping-based algorithms. This catadioptric system optically redistributes pixels without the need of any digital warping which preserves resolution and accurate pixel information. Finally, we present both learning and non-learning based algorithms that exploit the benefit of our catadioptric imaging system which achieves accurate long-term prediction of solar irradiance and sun occlusion by clouds.

Together, these contributions provide a fundamental advance to solar irradiance forecasting using core computational imaging.

Contents

List of Figures

List of Tables

1 Introduction

The sky is the daily bread of the eye.

— Ralph Waldo Emerson

As the increase of concern in global warming and environmental pollution increases, there has been a large push towards utilizing more clean energy solutions within the electricity grid. Solar energy is a viable solution due to the abundant and available sun which can be harnessed as a clean energy source. Therefore, at this turning point of energy generation, solar energy is stemming its way towards the forefront of this push.

In the first quarter of 2024 alone, the United States (US) solar energy market has installed 11.8 gigawatts-direct current (GWdc) capacity with the utility-scale segment attributing to a large portion of that total at 9.8 GWdc; the second best quarter for this segment. PV energy accounts for 75% (Figure 1.1) of electricity-generating capacity additions within this quarter and will continue to rise within the years to come [Davis *et al*., 2024].

Solar energy is generated from solar irradiance, the output of light energy from the sun measured at a location on Earth, which is converted to usable energy through the use of photovoltaic (PV) devices. Increasing the penetration of solar generation into the electricity grid, however, has been deemed difficult due to intermittent weather changes around a localized PV site. In particular, cloud occlusion of the sun, which scatters or absorbs the sun's rays, drastically decreases the amount of available solar irradiance that reaches PV sites at ground level [Alados-Arboledas *et al*., 1995, Wielicki *et al*., 1995]. Cloud cover and factors such as shape, vertical depth, and trajectory are stochastic by nature and influenced by complex global interactions that are very difficult to forecast. This random nature of cloud factors lead to large fluctuations in the availability of solar irradiance and as a result, the output power is prone to fluctuations and uncertainly which poses a challenge for grid operators when balancing the demand for energy versus consumption. Figure 1.2 shows how significant solar irradiance can fluctuate within a short window of

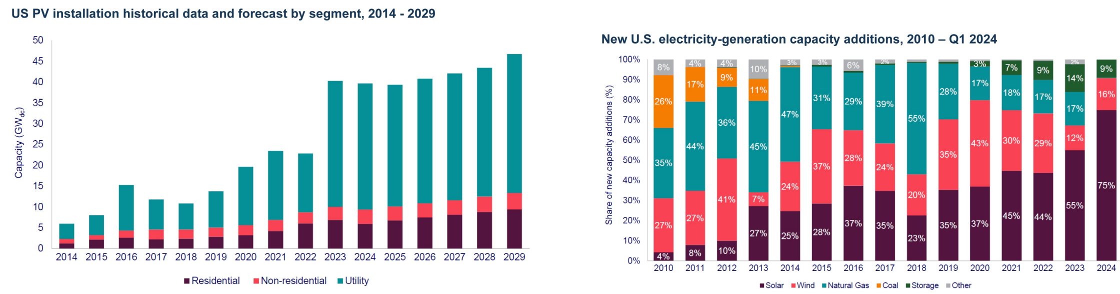


Figure 1.1: **Photovoltaic availability.** From left to right: Total US PV installations from 2014 to present along with future predictions to 2029. US electricity-generation additions from 2010 to Q1 2024 [Davis *et al.*, 2024].

time due to cloud occlusion.

Failure to maintain stable operation across the grid leads to voltage and frequency fluctuations, limited time to adjust between energy sources, and ultimately energy disruptions [Diagne *et al.*, 2013, Impram *et al.*, 2020, Infield and Freris, 2020, Kumar *et al.*, 2016]. In fact, studies suggest that many of the available regulation equipment is not capable of mitigating the adverse effects of PV intermittency. The development of complex power inverters with the use of sophisticated forecasting and monitoring systems is needed to maintain grid stability while incorporating PV power [Eftekharnejad *et al.*, 2013]. Furthermore, downstream tasks for grid components such as the *charging* and *discharging* of a connected solar battery for energy storage, can be improved through optimal scheduling via irradiance forecasting. This can lead to improved self consumption maximization of these components. As one can see - solar energy is invariably intermittent, posing a significant challenge in its full integration and wide-spread usage within the electricity grid [S. Sayeef and Rowe, 2012, Sovacool, 2009]. How do we increase the use of solar generation into the electricity grid while preventing system downtime due to random weather fluctuations? My thesis presents various computational imaging and learning-based methods that answers this question in an attempt to accurately forecast solar irradiance.

Images of the sky provide rich context of cloud patterns around a localized PV site and if these images are captured in sequence, provide a copious amount of data for inference. These images are traditionally captured by ground-based cameras with a large field-of-view (FOV) via a fisheye lens or hemispherical mirror. These imagers provide non-uniform spatial resolution of the sky with a higher detail and resolution at the zenith and significantly lower-resolution near the horizon. As a result the apparent motion of clouds within these images are non-uniform; large at the zenith and compressed at the periphery (Figure 1.3).

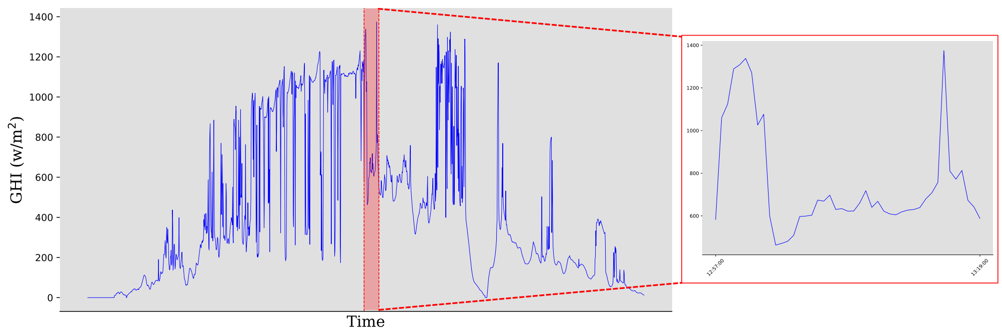


Figure 1.2: **GHI variability.** On the left, we show GHI values over a full day. The excerpt figure on the right shows the variability of GHI within a short time window. In less than a minute, GHI decreases about 1000 W/m^2 and remains low for a few minutes. The inability to forecast this phenomena can lead to balancing issues between demand and consumption and ultimately energy disruptions.

The motion of clouds at the periphery is vital because it builds on the future evolution of their trajectory. Optical flow and learning-based algorithms therefore suffer when attempting to make a prediction of cloud evolution due to this non-linear apparent motion present in traditional imagers. The affect of this is shown in previous works where accurate predictions are only made within very short forecasting windows (e.g. a few minutes) [Andrianakos *et al.*, 2019a, Kato and Nakagawa, 2020, Le Guen and Thome, 2020, Nie *et al.*, 2024]. To that end, this thesis aims to increase long-term prediction of solar irradiance forecasting using computational imaging.

1.1 Thesis Contribution

This thesis makes the following contributions to improve solar irradiance forecasting:

- **Precise Forecasting of Sky Images Using Spatial Warping (Chapter 3)**. To address the limitations of traditional sky imagers, why not bring the original image to a new space where we can obtain uniform apparent motion? That's exactly what this chapter proposes. We present a spatial warping algorithm that takes the image and warps it to a new space that ensures that clouds further from the zenith of the hemispherical mirror have similar apparent motion to those in the periphery. By doing this during training time, we are able to show that we can obtain longer forecasting of cloud evolution.

- **A Catadioptric Sky Imaging System (Chapter 4)** Digitally warping traditional sky images does indeed help in forecasting, but we can take this one step further and *optically* warp these images instead.

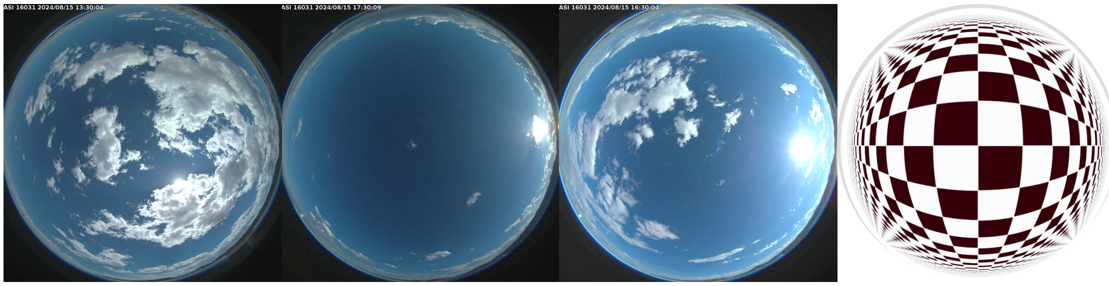


Figure 1.3: **Sky image samples.** (First 3) samples of sky images captured from the National Renewable Energy Laboratory (NREL) whole sky imager (WSI) [Andreas and Stoffel, 1981]. (Last) simulated image of a hemispherical imager with a large checkerboard placed in its FOV. This image shows the resulting affect of the non-linearity present when using these imagers.

In this chapter, A novel computational sky imaging system is developed and deployed that utilizes a catadioptric system of a hyperboloidal-shaped mirror. The benefits of this system with the optimal mirror shape profile is that it provides a uniform spatial resolution of the sky (for each height), over the entire field of view of the device. We are able to use these captured images to build a large dataset of images for inference.

- **Algorithms For Long-Term Forecasting (Chapter 5)** We show the benefits of utilizing our hyperboloidal-shaped imaging system for long-term solar irradiance forecasting via a suite of learning and non-learning based algorithms.

1.2 Thesis Structure

The remainder of this thesis is structured in the following way: In chapter 2, I delve into the topic thoroughly by presenting the background of solar irradiance forecasting. This includes discussing traditional data capture devices, forecasting methods, as well as limitations and key challenges in long-term forecasting of solar irradiance. Chapter 3-5 discussed the presented contributions. In chapter 6, I expound on the results of the contributions that is presented within this thesis. I also present possible future directions to take within the scope of solar irradiance forecasting using sky images.

2 Background

Within this thesis, I focus on the main contribution of the intermittency in solar generation which is cloud occlusion of the sun. By turning this into an imaging problem, we can use periodically captured images of the sky to gain inference on atmospheric factors that affect the amount of solar irradiance reached at ground-level PV systems. This chapter discusses the basics of sky imagery along with traditional methods used to analyze sky conditions and perform accurate forecasts of solar irradiance.

2.1 The Importance of Solar Irradiance Forecasting

Before introducing methods that use information from sky images to accurately forecast solar irradiance, we need to discuss why this forecasting is necessary.

Meeting electricity grid demands. Balancing demand versus consumption is vital for electricity grid operators to ensure no system downtime and power outages for consumers. Integrating solar energy into the grid challenges this requirement due to the intermittency of solar power which is influenced by cloud occlusion of the sun. Fluctuations in solar output can lead to suboptimal grid decisions which can be mitigated by irradiance forecasting that provides critical insights into solar energy's impact on the grid. In-turn, this allows for better management of electricity loads, storage, and backup generation. Also, by knowing the availability of power long into the future, operators are able to ramp up other energy sources to meet demand [Gowrisankaran *et al*., 2016].

Alternative energy. Other forms of power (e.g. gas and hydrothermal turbines) require tens of minutes, if not hours, to warm up for full use and leaving them running, even in a low-powered state, is wasteful. Without solar forecasting, sufficient reserve capacity must be maintained to cover significant fluctuations in solar generation. This may lead to operating conditions where conventional generators are maintained and operated at inefficient partial load states to cover the loss of unscheduled generation capacity [West *et al*., 2014]. Therefore, accurate forecasting knowledge of available power is necessary to balance demand and consumption for grid operators [Dixon *et al*., 2022].

Stability. Solar power brings new challenges within the electricity grid. These challenges: poor voltage regulation, long recovery times from voltage dips, and protection failures affect the stability of the grid and is caused by the intermittency of solar power. While energy storage techniques such as battery storage, ultra-capacitors, fuel cells, and flywheels help smooth this voltage fluctuations within the grid, they also raise costs substantially. A proactive approach is to use solar forecasts to manage fluctuations and limit the need for extensive storage capacity [Saleh *et al*., 2018]. Using irradiance forecasts to dynamically adjust inverter output has been shown to improve voltage regulation in a smart grid environment [Ghosh *et al*., 2017].

Storage. Indeed, the question of *"why not store unused energy in batteries?"* pops up. It is a valid solution which, if it was a complete solution would null this thesis, can be tricky when integrating within the electricity grid. Take a mobile phone with a battery for example. The battery discharges, you charge it again, and the cycle repeats for the life of the device. Over time, the quality of the battery decreases and it either does not hold full charge, it discharges quickly, or both. Recently, with the goal prolonging battery life in mind, mobile phone manufacturers have developed software for optimized charging strategies based on when users charge their phone and how they use it. With the image of a very relatable example in your head, adapt that same principle to PV sites with extra energy stored in batteries, increased by a scale factor of 10. Solar irradiance forecasting is needed for optimized charging and discharging of PV components, prolonging their longevity, and for the overall preservation of resources [Berrada and Loudiyi, 2016, K *et al*., 2023, Modi *et al*., 2024].

On the aside, for large-scale utility grids, storing large amounts of energy requires large batteries which is expensive and takes up a significant amount of space. For example, a large-sized battery storage system providing 500 Megawatts (MW) of power for 4 hours has a storage capacity of 2,000 Mega-watt-hours (MWh). The average battery cost of a system this size is $470 per kilowatt hour (kWh) resulting in a total of about $900 million. A significant cost to deploy and maintain long-term.

2.2 Sky Imaging Systems

Acquisition of sky images for solar irradiance forecasting comes in 2-fold. They be can captured from a top-down view using a satellite or a bottom-up view using ground based imagers looking up towards the sky. There are advantages and disadvantages of each acquisition approach. Satellite-based imagery has a

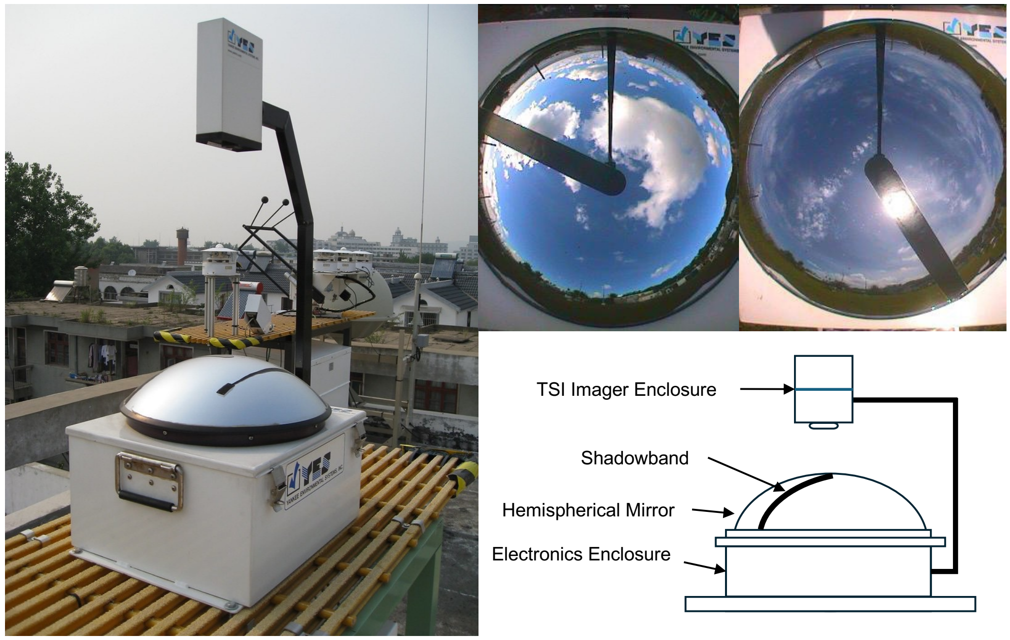


Figure 2.1: **Total Sky Imager.** (Left) a Total Sky Imager (TSI) [Victor, 2005]. (Right-top) sample images from a TSI. (Right-bottom) Specifications of the TSI.

larger spatial coverage, often over kilometer (km) scales, with constant persistence. However, they suffer from low temporal resolution; often capturing images every 30 minutes. There is also a data latency when sending captured images to be processed on Earth. Deploying and maintaining satellites are costly as well, although there are some publicly available datasets for use. Compared to satellites, ground-based imagery is cost-effective, has a high temporal resolution, and captures localized data around a PV site to accurately forecast irradiance. However, ground-based sky imagers have limited coverage determined by the shape of the lens; limiting prediction ranges. They are also high susceptible to local weather conditions such as rain and snow which inhibits any form of acquisition. For this thesis, I focus on ground-based imagery for accurate localized solar irradiance forecasts.

Ground-based devices used to capture images of the sky - often dubbed *Sky Imagers* - encompass a fisheye lens [Cazorla *et al*., 2008, Kleissl *et al*., [n.d.]] or a catadioptric *(see below)* combination of cameras and mirrors [Victor, 2005]. These are wide FoV images captured in regular intervals. The imagers also have additional features such as a shadow-band to prevent over-saturation by the sun in the image and prevent damage to the camera sensor. Figure 2.1 shows an example of a Total Sky Imager (TSI).

Catadioptric Imaging. Catadioptric imagers which utilizes the reflective nature of mirrors during the image acquisition process are designed such that their shape can achieve various tasks. Baker and Nayar have extensively studied the family of these shapes [Baker and Nayar, 1998]. The shape of the mirror (hemispherical, planer, parabolic, etc.) affects the effective viewpoint, FOV, resolution, and manufacturability of the mirror. Overall, a catadioptric setup is a great way to increase the customizability of a sky imaging setup.

Key Challenges. An apparent limitation of these imagers is the non-linear fisheye distortion introduced which affects clouds motion estimates for trajectory predictions. The further-out clouds attenuate evolution over time and their prior motion estimates attribute to longer forecasting horizons. These sky imagers, either from a traditional RGB camera with a fisheye lens or a catadioptric setup, capture images in regular intervals. The need of the fisheye lens/hemispherical mirror is to capture a wide angle (about $180°$) sky image with a single shot. These existing setups are not sufficient for accurate long-term prediction of cloud evolution due to the limitation of the imager.

Obtaining a $180°$ (FoV) photograph results in a highly nonlinear mapping between the sky and the image. As shown in Figure 1.3, objects in the periphery are heavily compressed compared to those at the zenith. This makes motion modeling near the horizon fragile to small perturbation. This problem is exacerbated by optical flow estimation, which is hard to perform on cloud imagery that lack high-contrast textures. The resulting flow estimates are inherently fragile, especially near the horizon. While enforcing smoothness priors on the flow estimates often leads to robustness especially at the zenith, they tend to make the flow at the horizon nearly zero. Hence, the nonlinear spatial resolution is not conducive for predicting cloud evolution over longer time horizons

By addressing how these images are captured along with improving methods for predicting sun occlusion and irradiance values, we can make better long-term predictions.

2.3 Irradiance

Irradiance is the measure of energy per unit area expressed in watts per meter-squared (W/m^2) on the ground. Irradiance can be scattered through the atmosphere or through clouds before reaching the ground. This is referred to as diffuse irradiance (DI). The irradiance directly coming from the sun is called direct normal irradiance (DNI) and is a measure of the beam radiation through a plane perpendicular to the direction of the sun. The total amount of irradiance from the sun that hits the ground is global horizontal irradiance (GHI). GHI is the summation of diffuse and direct irradiance incident on a horizontal surface

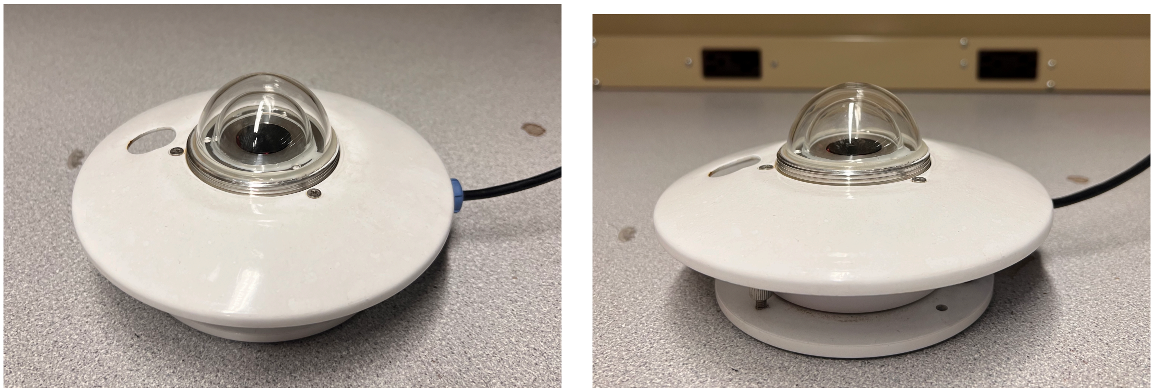


Figure 2.2: **Pyranometer.** A thermopile pyranometer which measures GHI in units of watts per meter-squared (W/m^2).

represented via equation 2.1 where z is the solar zenith angle [Stein *et al.*, 2012].

$$GHI = DI + DNI \cdot \cos(z) \tag{2.1}$$

GHI is measured using a pyranometer (Figure 2.2). A pyranometer is used on a planar surface and measures GHI (direct and diffuse) within a wavelength range of 0.3 μm and 3 μm. The thermopile pyranometer used for this thesis is a sensor based on thermopiles designed to measure the broad band of the solar radiation flux density from a 180° FOV. The difference between the pyranometer output and PV output should be noted. GHI measures the available solar energy, while PV output measures the electrical energy produced via solar energy. PV output is affected by factors such as temperature, panel orientation, and efficiency. Compared to Figure 1.2 that shows the variability of GHI under cloudy conditions, Figure 2.3 shows GHI under a clear sky day.

Now that the foundation of sky imaging has been laid, we go into methods of forecasting both sky images and GHI along with the methods used to increase predictions.

2.4 Sky Image Forecasting

Non-learning forecasting. Sky image forecasting is the process of predicting future sky image(s) based on a sequence of past frames. These algorithms seek to understand and predict the evolution and trajectory of clouds within these images. Non-learning based image forecasts primarily use motion estimation techniques such as optical-flow based methods, block matching, particle image velocimetry (PIV), or correlation-based methods [Chow *et al.*, 2015, 2011, Chu *et al.*, 2013, Huang *et al.*, 2013]. Solely using

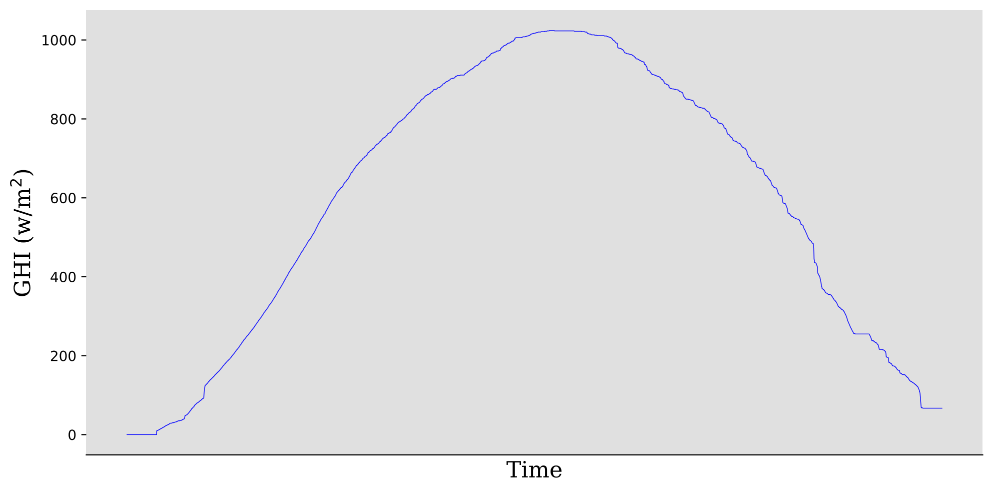


Figure 2.3: **Clear sky GHI.** We GHI values over a full day with minimum to no clouds. Notice how smooth the curve of GHI values are when clouds are not present.

motion estimation techniques to model cloud dynamics has immediate consequences due to the variability and constant changing of shape of clouds which makes forecasting their trajectory difficult. For example, accurate short-term forecasts, approximately 1 min, using block-matching based estimation, become increasingly difficult when complex cloud dynamics are involved. Predicting long-term poses an even greater challenge and results in severe degradation. More recent methods have seen better success by incorporating deep learning methods, coupled with motion estimation, to model cloud dynamics and evolution in sky images.

Learning-based forecasting. Learning-based forecasting methods improve upon traditional motion-based estimation by tasking machine and deep learning frameworks to learn a representation capable of forecasting solar irradiance using sky images. In particular, the model f_θ learns a mapping between sky images at past time instances $[I_{t-N}, \ldots, I_{t-1}, I_t]$ and future sky image frames $\left[\hat{I}_{t+1}, \ldots, \hat{I}_{t+N}\right]$ or irradiance values $\left[\hat{G}_{t+1}, \ldots, \hat{G}_{t+N}\right]$. Initial works used a convolutional neural network (CNN) as this framework [Crisosto *et al.*, 2021, Feng *et al.*, 2022, Jiang *et al.*, 2020, Paletta and Lasenby, 2020a,b, Sun *et al.*, 2019] to predict a sky image frame or solar irradiance value. Recurrent models (e.g. RNN and LSTM) have also been used to assist in modeling the temporal dynamics between the sequence of frames [Kato and Nakagawa, 2020]. More recent works have explored using a generative adversarial network (GAN) [Andrianakos *et al.*, 2019b, Nie *et al.*, 2023], modeling the physical dynamics to enhance cloud motion analysis [Le Guen and Thome, 2020], and even transformer models [Demir *et al*., 2022, Liu *et al*., 2023, Mercier *et al*., 2023, Pospichal *et al*., 2022, Zhang *et al*., 2024]. Each of these models have their unique benefits when forecasting irradiance or a subsequent sky image. Due to the lack of a unified dataset or benchmarks for solar irradiance forecasting, it can be difficult comparing models and prediction results – especially when these models are trained on sky images captured from different regions. Each location has their own unique set of weather pattens which may not be transferable to a new location when using a learned model.

Nowcasting is a popular approach used in forecasting solar irradiance. It works by using a learning framework to initially forecast a sky image frame. That predicted image is then passed through another model that estimates the solar irradiance value based on the current cloud conditions within the image [Gao and Liu, 2022, Nouri *et al*., 2023, Song *et al*., 2022].

Sky image understanding. To aid in forecasting solar irradiance, there are numerous methods that seek to understand the semantics and dynamics within the scene to increase accuracy. Segmentation is a widely used method to differentiate between pixels corresponding to a cloud, the sky, or the sun [Dev *et al*., 2019, Li *et al*., 2011, Liu *et al*., 2015, Xie *et al*., 2019]. These segmented pixels assist in modeling the motion of clouds in a sequence of frames and remove unnecessary information within the sky image. However, segmentation is often difficult due to thin and opaque clouds which is difficult to differentiate between sky and cloud pixels. Also the decision on how to handle overlapping clouds at different heights makes segmentation even more difficult.

Sun location identification within sky images is also necessary for some tasks. This can be done by using intensity based identification where the brightest pixel is classified as the location of the sun [Paletta and Lasenby, 2020c]. Using the camera information along with the location of the sun in the real-world relative to the sky imager is another method. However, this method cannot be used when information about the imager is unavailable. Using infrared cameras is also a promising approach due to the imagers ability to overcome saturation by the sun [Niccolai and Nespoli, 2020].

2.5 GHI Forecasting

Some methods only use GHI and bypass incorporating sky imagery into the forecasting pipeline [Almarzooqi *et al*., 2024, Alzahrani *et al*., 2017, Jailani *et al*., 2023]. This essentially becomes a time-series forecasting task where the goal is to predict a future GHI value given only past GHI values. This is a very difficult task because there is no notion of cloud trajectory within the learning pipeline. If you do not know that a cloud is headed towards the sun via an image, it will not be possible to forecast a decrease in GHI solely using past GHI values. Therefore, incorporating sky images into the task of solar irradiance forecasting is essential for accurate long-term prediction.

This thesis will address these traditional limitations by presenting methods in the subsequent chapters that increases the long-term forecasting of solar irradiance.

3 Precise Forecasting of Sky Images Using Spatial Warping

Prediction is very difficult, especially if it's about the future.

— Niels Bohr

In chapter 2, we discussed methods to capture sky images and the limitations that prevented long-term forecasting when using traditional imagers. For example, the Total Sky Imager (TSI) is one approach for predicting solar irradiance by monitoring the movement of clouds around a particular site. It captures of a hemispherical image of the sky with a 180° FOV using a catadioptric system of a single RGB camera observing the sky through a curved mirror. Figure 3.2 provides examples of such images.

These TSIs capture images periodically, every 30 seconds, and as a result, these images can be stacked together to create a time-lapse of historical cloud cover data around a particular site. Such TSIs, coupled with a pyranometer that measures GHI, have been utilized in recent studies to nowcast and forecast solar irradiance [Al-lahham *et al*., 2020, Dev *et al*., 2016, Siddiqui *et al*., 2019].

It is important to note that the definitions of short and long-term prediction of cloud movement in sky-images is subjective to the sampling period, T_0, at which the images are captured. Here, T_0 is set as $T_0 = 30$ seconds. Therefore, short-term prediction is quantitatively defined as predicting 30 seconds in the future, whereas long-term prediction is anything greater than 30 seconds. Overall, however, both short and long-term prediction is difficult due to the constant changing shape of clouds. Accurate prediction greater than short-term is exacerbated by distortions introduced by the hemispherical mirror used to capture the wide FOV sky image. Specifically, in a typical image obtained from a TSI, objects near the horizon are spatially-compressed and hence, appear much smaller at the horizon than when they are at the zenith. Due to this non-linear mapping produced by hemispherical mirrors, uniform physical motion of clouds leads to apparent motion of varying magnitude on the image plane. This in turn affects the accuracy of motion estimates for cloud movement tracking as the apparent motion at the horizon is extremely small and overwhelmed by the larger optical flow induced by clouds at the zenith. For forecasting longer time

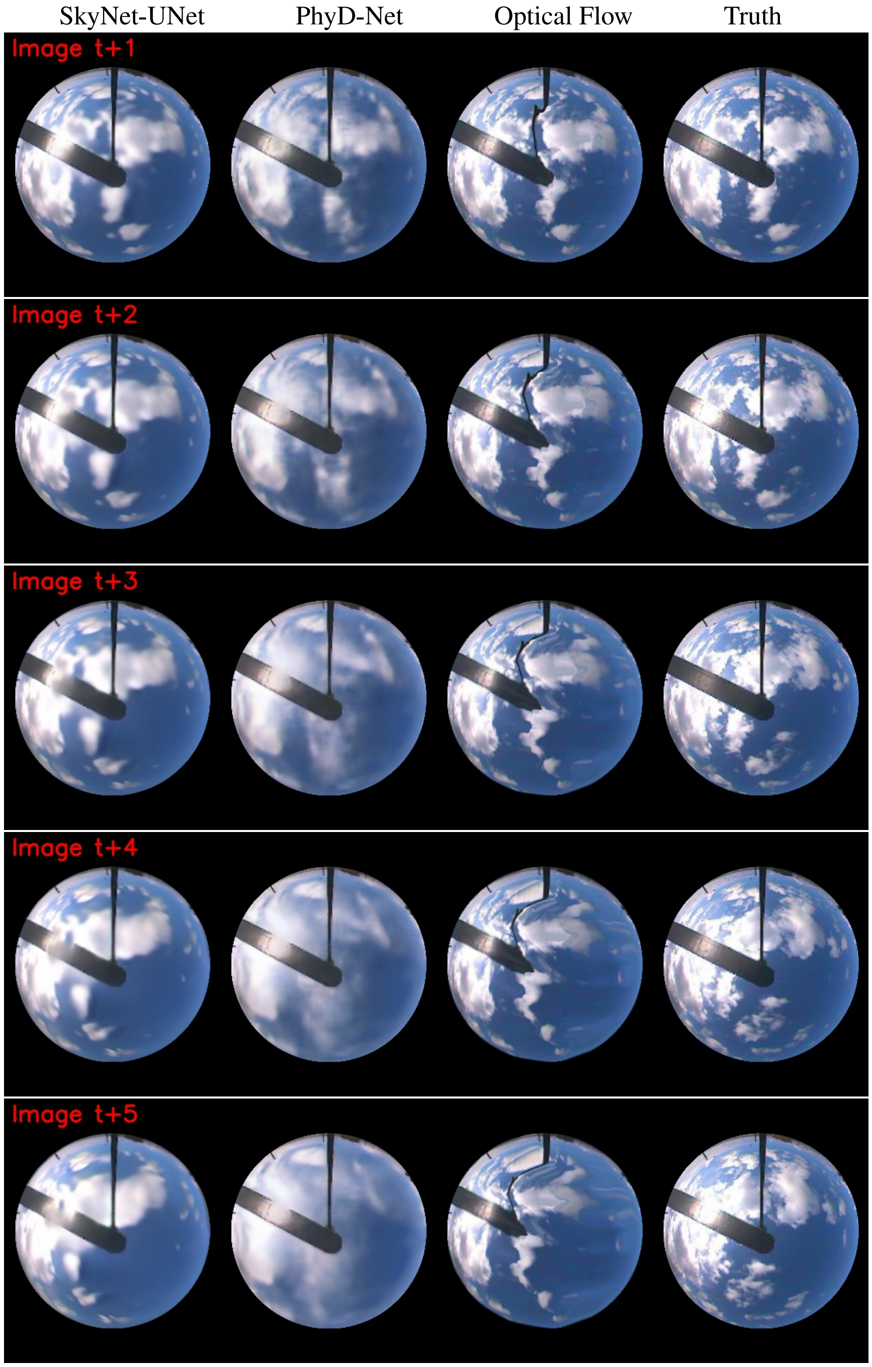


Figure 3.1: **Predicting future sky images.** Left-to-right: Results from SkyNet-UNet (the proposed technique), PhyD-Net-Dual [Le Guen and Thome, 2020], optical flow, and ground truth images. The methods take in as input images $[I_{t-5}, I_{t-3}, I_{t-1}, I_t]$, and predict future frames $[\widehat{I}_{t+1}, \widehat{I}_{t+2}, ..., \widehat{I}_{t+5}]$ (From top-to-bottom).

horizons, the small movement of these clouds are what determines how the texture of clouds evolve over time.

We attempt to counter this problem by warping the original image to a different space where the apparent motion is uniformly preserved both at the zenith and the horizon. This allows us to achieve longer forecasting times when modeling cloud evolution in sky images.

Contributions. In this chapter, we propose *SkyNet*, which focuses on improving sky-image prediction. The main contributions for this chapter are as follows:

- *Cloud forecasting via spatially warped images.* Our primary contribution is in showing that spatially-warping the sky images during training facilitates longer-forecasting of cloud evolution. This counters the adverse affects of resolution loss near the horizon.
- *Incorporating larger temporal context.* We adapt prior work on future frame prediction in videos [Liu *et al*., 2018] to the case of sky images. Here, to increase precision in forecasting, we go beyond two input frames to usher in a larger temporal context. Specifically, to predict the image at time $t + 1$, we take in as input four input images spanning $\{t - 5, t - 3, t - 1, t\}$.
- *Training and validation.* We train and evaluate our approach on a large dataset of sky images and demonstrate the ability to accurately forecast sky image frames with higher resolution metrics than previous cloud forecasting methods. We further use the forecasted sky-images to evaluate our results on estimating the GHI value for future time instances.

The accuracy of the proposed SkyNet predictions are shown in Figure 3.1 where we show our ability to predict sky images for future time-instants to $t + 5$. Each frame denotes a time lapse of 30 seconds in this dataset and hence, we predict up to two and a half minutes into the future.

3.1 Prior Work

We discuss prior work in modeling cloud dynamics with the goal of predicting solar irradiance.

3.1.1 Modeling Sky Evolution Using Optical Flow

Early works for cloud motion modeling such as those of Ai *et al*. [Ai *et al*., 2017a] and Jayadevan *et al*. [Jayadevan *et al*., 2012] use a grid or block-based optical flow methods to model cloud velocity and motion. This motion estimation method involves constructing a set of grid elements across the sky image in which the direction and velocity is then found between the correlation of grid blocks between adjacent frames. Although accurate for short-term cloud movement prediction at approximately 1 min, this technique becomes increasingly difficult when complex cloud dynamics are involved. Clouds are not uniform objects and their shape and trajectory can shift drastically. Also, clouds can overlap at different heights with different velocities and trajectories at each height. Therefore, this method becomes less accurate and difficult to forecast for long-term time horizons.

Recent advances of optical flow techniques have improved upon block-based motion estimation. Differential methods for optical flow estimation such as the Lucas-Kanade and Horn-Schunck are common and popular techniques for estimating cloud motion [Chang *et al.*, 2017, El Jaouhari *et al.*, 2015, Tiwari *et al.*, 2019, Zhang *et al.*, 2019].

Solely using optical flow to model cloud dynamics maintains consequences of the variability and constant changing of shape of clouds; making forecasting their trajectory difficult. More recent methods have seen better success by incorporating deep-learning methods with optical flow and other variables to model cloud dynamics and evolution in sky images.

3.1.2 Modeling Sky Evolution Using Deep Learning

Many of the works that utilize learning frameworks tasks a model to learn cloud dynamics from a previous representation of images to then predict a subsequent sky image. This predicted sky image is then used to predict solar irradiance at that time instance.

Kato and Nakagawa [Kato and Nakagawa, 2020] use a convolutional long short-term memory network (LSTM) with optical flow vectors and past sky images as input to generate a predicted sky image by extrapolating the flow vectors with the input images. Andrianakos *et al.* [Andrianakos *et al.*, 2019b] utilize a generative adversarial network (GAN) for sky image prediction to counter the adverse blurry image effects of using traditional mean squared error loss (MSE) for image prediction. Le Guen and Thome [Le Guen and Thome, 2020] incorporate physical knowledge in deep models based on PhyDNet [Guen and Thome, 2020] that exploits physical dynamics to enhance cloud motion modeling.

Deep neural networks are currently the most recent methods for modeling cloud dynamics in sky image frames and provide some of the best prediction results. However, precise forecasting of future sky image frames for longer time horizons is hindered by artifacts induced by the imager. Fisheye camera lenses and hemispherical mirrors, commonly used for capturing sky images due to their wide angle FOV, compress the imagery near the horizon which affects the prediction of cloud evolution when forecasting sky images. To counter this, we present a uniform warping scheme on the captured images to ensure that clouds further from the zenith of the hemispherical mirror have similar apparent motion to those in the periphery, so as to ensure accurate forecasting.

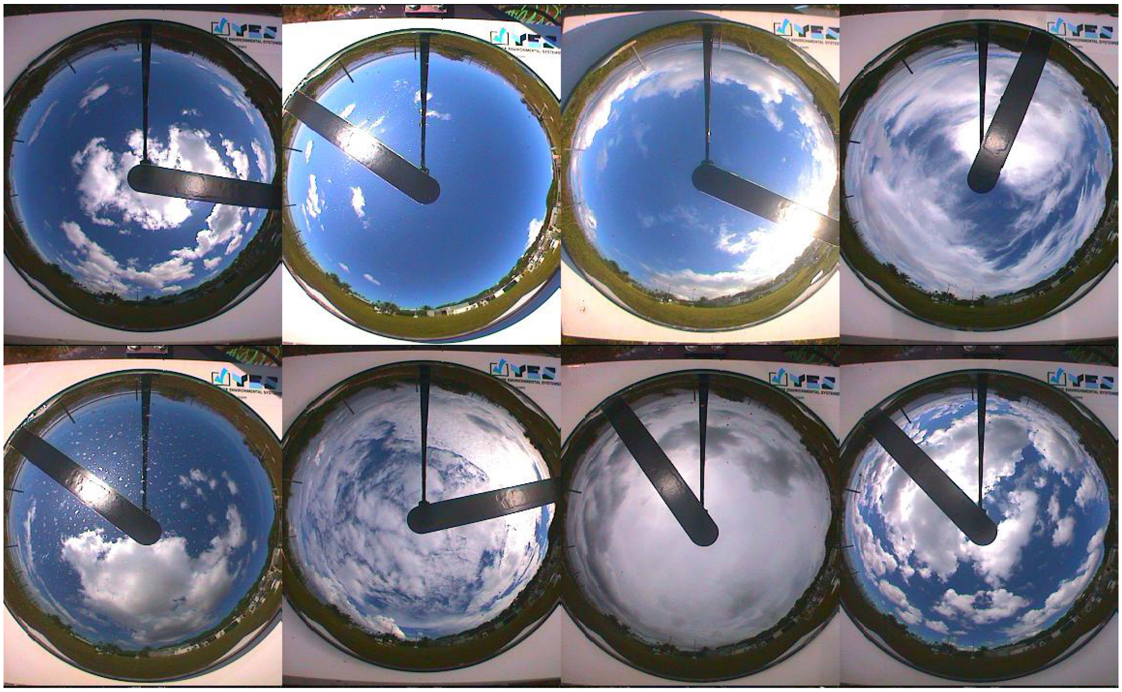


Figure 3.2: **Sample TSI Images.** Sample images captured by a TSI [Victor, 2005] across various conditions.

3.2 Background and Problem Setup

We begin by describing how sky images are captured along with the basic notation of how cloud occlusion relates to the amount of solar radiation being received at a site. We follow this by deriving the proposed uniform warping method for sky images.

3.2.1 Total Sky Imagers and Solar Irradiance

A TSI provides a time-lapse video sequence from an RGB camera that observes the sky via a hemispherical mirror [Victor, 2005]. Generally, these systems are deployed to capture imagery of the sky at regular intervals for applications such as solar irradiance forecasting and visualizing cloud dynamics. To prevent damage of the camera sensor from direct expose of the sun, the TSI typically includes a mechanical arm that travels along the path of the sun throughout the images to occlude direct exposure. Figure 3.2 shows some images from the TSI.

The RGB image captured from the TSI provides a sky map from which we can identify the location of clouds, their movement over time, and even a crude understanding of their absorption properties by associating the cloud cover at a time instance with it's associated GHI value when using a ground-based pyranometer. Suppose that we have a solar panel collocated with the TSI, denoted by the location x. If the area of this panel is A in m^2, then the radiant flux $\Phi(t)$ measured at time t is given as:

$$\Phi(t) = A \int_{\lambda} Q(\lambda) E_x(\lambda, t) d\lambda. \tag{3.1}$$

where $Q(\lambda)$ is the quantum efficiency of the panel, and $E_x(\lambda, t)$ is the spectral irradiance at the location of the panel, at the wavelength λ and time t, expressed in the units of $J/(nm \cdot m^2)$. This spectral irradiance can be related to the spectral radiance $L_x(\omega, \lambda, t)$ — the flux at a point x along a direction ω in the units of $J/(nm \cdot m^2 \cdot Sr)$. Therefore, the radiant flux $\Phi(t)$ can now be written as:

$$\Phi(t) = A \int_{\lambda \in \Lambda} Q(\lambda) \left[\int_{\omega \in \Omega} L_x(\omega, \lambda, t) \max(0, \mathbf{n}^T \omega) d\omega \right] d\lambda. \tag{3.2}$$

The set Ω defines the solid angle over which light is received at the solar panel and $\mathbf{n}$ is the surface normal, or the orientation of the solar panel, in the same coordinates as ω. As an approximation, this integral can be written as the occlusion map of the clouds multiplied by the spectral radiance from sunlight as well as skylight, which can be pre-measured.

3.2.2 Problem Definition

The goal of this chapter is to provide a framework for short term prediction of sky images. Specifically, the TSI takes an image every T_0 seconds to provide a time lapse video.[1] For simplicity of notation, we denote this time-lapse video as a collection of frames $\{\ldots, I_{t-1}, I_t, I_{t+1}, \ldots\}$, where t is an integer-valued index for the sequence, keeping in mind that any two successive images are obtained T_0 seconds apart by the TSI.

Given $\{\ldots, I_{t-2}, I_{t-1}, I_t\}$, the past and current images in time lapse sequence at a time instant t, our goal is to predict $\{I_{t+1}, I_{t+2}, \ldots\}$, the images in the time lapse sequence for the next few instants. Since clouds often move fast, there is little correlation between images taken at sufficiently far away time instances; hence, we can restrict the time horizon of images that we consider both for the input images (from the past) as well as the predicted output images (of the future). Hence, our objective can be refined to using the image set $\{I_{t-T_p}, \ldots, I_{t-1}, I_t\}$ to predict the image set $\{I_{t+1}, I_{t+2}, \ldots, I_{t+T_f}\}$, where the choice of the input time horizon T_p and output time horizon T_f are discussed later.

[1] For the dataset that we work with, this sampling period $T_0 = 30$ seconds. This choice balances the need to monitor fast moving clouds, that would benefit from shorter sampling period, and the size of the dataset, which scales inversely with T_0.

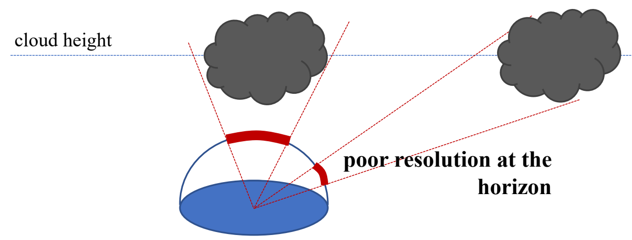


Figure 3.3: **Cloud resolution.** A cloud subtends a smaller angle when it is further away from the zenith. This results in the nonlinear spatial warping and poses critical challenges for effective forecasting of cloud movement.

Challenges. Modeling the evolution of the sky and predicting images at future time instants faces challenges that stem from the clouds themselves as well as features induced by the imager. Clouds are amorphous, lacking the rich features that are prized in traditional motion modeling and flow estimation. Such domain-specific concerns can be handled by using learning techniques that implicitly build a prior for the underlying imagery. However, even when using sophicated learning techniques, there are significant challenges that arise from the spatial distortions introduced by the TSI.

Getting a 180° FOV photograph with a TSI results in a highly nonlinear mapping between the sky and the image as is seen in Figure 3.3. The effect of this distortion is easily seen in Figure 3.2. An immediate consequence of this nonlinear warping is that motion near the horizon is not easily observable; for the same amount of cloud movement, the *perceived optical flow* on the image plane of the camera is significantly smaller at the horizon. This makes motion modeling near the horizon fragile to small perturbation. This problem is exacerbated by optical flow estimation, which is hard to perform on cloud imagery that lack high-contrast textures. The resulting flow estimates are inherently fragile, especially near the horizon. While using enforcing smoothness priors on the flow estimates often leads to robustness especially at the zenith, they tend to make the flow at the horizon nearly zero. Hence, the nonlinear spatial resolution is not conducive for predicting cloud evolution over longer time horizons.

Solution outline. To address these challenges in motion estimation, and provide a framework for precise prediction of sky images, we make two modifications to traditional ideas in future frame prediction.

- *Optimal spatial warping.* First, under a simple model of image formation, we propose a warping of the TSI image so as to preserve motion of clouds over the spatial field. This serves to amplify motion near the horizon that is otherwise small. We describe this in Section 3.3.

- *Multi-image prediction.* Second, since the image after warping is still smooth, we use multiple frames from the past to stabilize motion estimates. We perform this by adapting prior work on two-frame activity prediction. This is described in Section 3.4.

3.3 Optimal Warping of Sky Images

The warping of sky images is necessary because the apparent motion of clouds around the periphery of the hemispherical mirror will be much smaller than when at the zenith. As a result, we can only get good optical flow estimates at the zenith at the cost of poor optical flow estimates elsewhere. For better long-term prediction of cloud evolution and as a result, better long-term prediction of solar irradiance, we spatially warp the images so that the apparent motion is more uniform. We design a warping scheme so that over a specific site, we can achieve uniform optical flow.

Image formation model. We model the imager as being an orthographic camera observing the sky through a spherical mirror of radius R_m. The optical axis of the camera points is normal to the ground, and the optical center is aligned to the center of the hemispherical mirror. We model the ground as being planar, an assumption that is reasonable given that the radius curvature of the earth is couple of orders of magnitude larger than the geographic region we can image with the TSI. Given this, we adopt a world coordinate system whose origin is at the center of the spherical mirror. The xy coordinate plane is aligned with the ground plane and the z axis is pointing towards the sky and hence, the optical axis of the camera is aligned to $[0, 0, -1]^\top$. Figure 3.4 provides a schematic of this setup.

Suppose that a cloud at $X_c = [x_c, y_c, h]^\top$ maps to image pixel coordinates $[u_c, v_c]^\top$. We now seek to estimate the relationship between these quantities. We first move from cartesian coordinates on the ground plane to polar coordinates, which allows us to exploit the rotational symmetry of the mirror about the z-axis. With this, we can write the cartesian coordinates of the cloud as

$$X_c = [x_c, y_c, h]^\top = \begin{bmatrix} \rho\cos\theta & \rho\sin\theta & h \end{bmatrix}^\top. \tag{3.3}$$

and that of the image pixel coordinates as

$$[u_c, v_c]^\top = \begin{bmatrix} s\cos\theta & s\sin\theta \end{bmatrix}^\top. \tag{3.4}$$

where (ρ, θ) and (s, θ) are polar coordinates for ground and image plane location, respectively, for the cloud. Note that we have effectively used the rotational symmetry of the mirror in insisting both the clouds and its corresponding camera pixel subtend the same angle θ in polar coordinates.

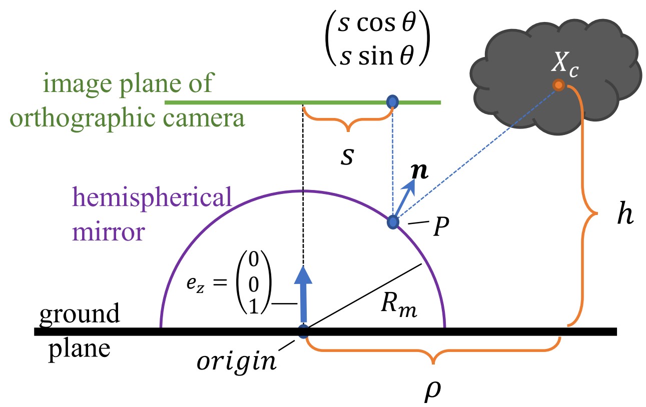


Figure 3.4: **Warping scheme.** Overview of how the 3D position of a cloud in the world space gets mapped to a point on the image plane using a hemispherical mirror.

Let's denote P as the point on the mirror that reflects the cloud to its corresponding image plane pixel. Given that the camera is orthographic, we can derive the P to be

$$P = \begin{bmatrix} s\cos\theta & s\sin\theta & \sqrt{R_m^2 - s^2} \end{bmatrix}^T. \tag{3.5}$$

This comes from the fact that the point P is on a sphere of radius R_m. We can now enforce Snell's laws of reflection to relate ρ and s to each other; thereby getting the functional relationship between the position of the cloud in world coordinates to its location on the image plane. Specifically, we can write the surface normal at P, which is simply a unit norm vector oriented along P, to be equal to the average between the line produced by the point P and the vertical line at e_z:

$$\left(\frac{X_c - P}{\|X_c - P\|} + e_z\right) \cdot \frac{1}{2} = \frac{P}{\|P\|}. \tag{3.6}$$

Noting that $\|P\| = R_m$, the radius of the sphere, we can express the 3rd coordinate of (3.6) as

$$\frac{1}{2} \cdot \left[\frac{h - \sqrt{R_m^2 - s^2}}{\|X_c - P\|} + 1\right] = \left[\frac{\sqrt{R_m^2 - s^2}}{R_m}\right]. \tag{3.7}$$

We can now solve for $\|X_c - P\|$ to get

$$\|X_c - P\| = \gamma(s) = \frac{h - \sqrt{R_m^2 - s^2}}{\frac{2\sqrt{R_m^2 - s^2}}{R} - 1} \approx \frac{h}{\frac{2\sqrt{R_m^2 - s^2}}{R_m} - 1}. \tag{3.8}$$

Now plugging $\gamma(s) = \|X_c - P\|$ back into (3.6), we get

$$\frac{X_c - P}{\gamma(s)} + e_z = \frac{2P}{R_m}, \tag{3.9}$$

from which we can obtain an expression for X_c as

$$X_c = \gamma(s)\left[\frac{2P}{\|P\|} + e_z\right] + P. \tag{3.10}$$

Therefore, we can expand (3.10):

$$\begin{pmatrix} \rho\cos\theta \\ \rho\sin\theta \\ h \end{pmatrix} = \gamma(s)\left[\frac{2}{R_m}\begin{pmatrix} s\cos\theta \\ s\sin\theta \\ \sqrt{R_m^2 - s^2} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} s\cos\theta \\ s\sin\theta \\ \sqrt{R_m^2 - s^2} \end{pmatrix}\right]. \tag{3.11}$$

From the top two rows of the previous equation, we can relate ρ to s as follows:

$$\begin{aligned}\rho &= \left(\frac{2\gamma(s)}{R_m} + 1\right)s = \left(\frac{2h}{2\sqrt{R_m^2 - s^2} - R_m} + 1\right)s \\ &\approx \frac{2hs}{2\sqrt{R_m^2 - s^2} - R_m}.\end{aligned}$$

Instead of modeling ρ directly, we can model $\frac{\rho}{h}$ which gives us height invariance

$$\widetilde{\rho} = \frac{\rho}{h} = \frac{2s}{2\sqrt{R^2 - s^2} - R}. \tag{3.12}$$

Therefore, it does not matter the height at which the cloud is and therefore, we do not need to specify h.

Remarks. While there is a significant distortion of the sky in the image plane of the camera, we can undo this distortion by redefining the image in terms of $\widetilde{\rho}$ instead of s. That is, using the expression in equation (3.12), we can map the image plane from $(s\cos\theta, s\sin\theta)$ to $(\widetilde{\rho}\cos\theta, \widetilde{\rho}\sin\theta)$. This has the benefit of normalizing the observed optical flow so that it no longer suffers from the spatial distortion. More specifically, in the transformed coordinates, the observed flow magnitude is the same, immaterial of where the motion occurs in the field of view of the device.

Inverting the warping function. We can invert the relationship between $\widetilde{\rho}$ and s in equation (3.12), using some algebraic manipulation to get the following inverse relationship.

$$s = \frac{-R_m\widetilde{\rho} + R_m\widetilde{\rho}\sqrt{1 + 3(1 + \widetilde{\rho}^2)}}{2(1 + \widetilde{\rho}^2)}. \tag{3.13}$$

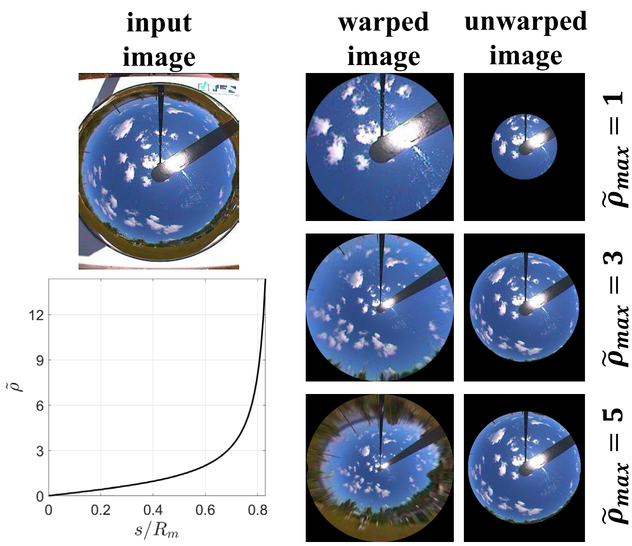


Figure 3.5: **Radial distance of warp.** We show how the appearance of an input image changes under the proposed warping. The plot on the left visualizes how we map from radial distances on the image to radial distances in the proposed representation. Choosing different values of the range of $\widetilde{\rho}$ produces different FOVs and associated distortions. This is visualized in the center column. The right column shows the image after inverting the warp to obtain the original image.

Implementation details. To implement the warping function, we need to find the radius of the mirror R_m. We observe that for an orthographic camera and a planar ground, the horizon maps to a circle with a radius of $R_m/\sqrt{2}$. We use this to estimate R_m in pixel count. The other important parameter that we need to set is the maximum value of $\widetilde{\rho}$. Setting $\widetilde{\rho} \in [0, \widetilde{\rho}_{\max}]$ defines the FOV of the device to be restricted to $\pm \tan^{-1} \widetilde{\rho}_{\max}$; for example, choosing $\widetilde{\rho} \in [0, 1]$ corresponds to a FoV of the sky of $\pm 45^{\circ}$. Figure 3.5 shows warped and unwarped images for different values of this range. A small value of this range leads to poor coverage of the sky and a large value has textures that are extremely blurred due to the nonlinearity of the warp as well as the incorporation of trees and buildings. For all of our experiments, we choose a range for $\widetilde{\rho} \in [0, 3]$ corresponding to a FoV of $\pm 71^{\circ}$ which provided a good balance between coverage and distortions. Finally, to avoid loss of information, we upsample the image dimensions by a factor of three.

Figure 3.5 also shows how the warping can be inverted so as to revert back to the original image space. It is worth observing how the relative sizes of clouds at the zenith and horizon changes in the warped space. This warping ensures that motion magnitude is preserved and spatially-invariant. Next, we

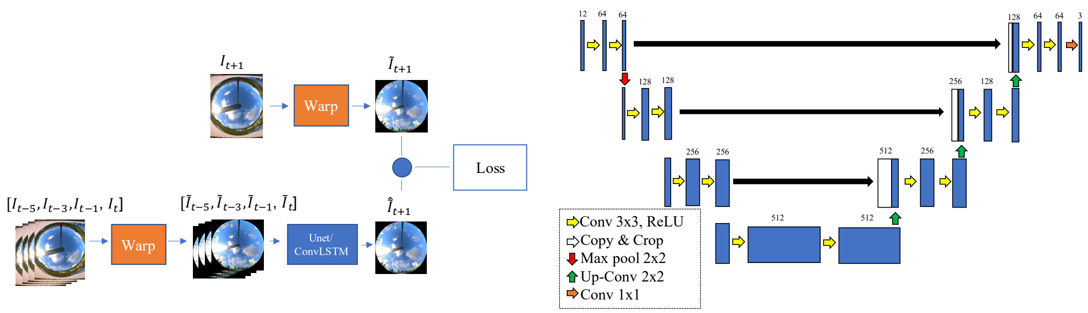


Figure 3.6: **SkyNet Learning Pipeline.** (Left) The proposed SkyNet forecasting method that incorporates the U-Net [Ronneberger *et al.*, 2015] or ConvLSTM [SHI *et al.*, 2015] neural network architecture used to forecast a subsequent sky-image frame. (Right) Diagram of U-Net architecture used for SkyNet.

look at a learning framework for multi-image prediction.

3.4 SkyNet

To forecast sky images, we learn a deep neural network that we refer to as *SkyNet* that takes in as input multiple images and predicts the next frame.

Network input. As mentioned earlier, using multiple images to forecast provides us robust estimates of slow moving clouds as well as to combat the distortions introduced by the imager. To faciliate this, we use the stack of images $\{I_{t-5}, I_{t-3}, I_{t-1}, I_t\}$ to predict the image at I_{t+1}. This choice reflects the need to have a long time horizon in the past, but given the redundancy, dropping some of the intermediate images help alleviate training time. The images are warped using the approach described in Section 3.3.

Network architecture. We consider 2 network architectures for our SkyNet Model. Initially, our first network architecture, SkyNet-UNet, adapts the future frame prediction model proposed for activity forecasting in [Liu *et al.*, 2018]. The backbone of this architecture is a U-Net [Ronneberger *et al.*, 2015] that takes in the input images to predict the image at the next time instant in the time lapse. Our second network architecture, SkyNet-LSTM, performs the same task as our initial network architecture, however, incorporates a convolutional long short-term memory network (ConvLSTM) [SHI *et al.*, 2015]. Both architectures incorporate the same loss function further described below. Figure 3.6 shows the end-to-end training pipeline for SkyNet.

SkyNet-UNet starts with the number of input channels representing the number of time steps being considered. For each layer of the encoder, the number of channels are doubled until the bottleneck of the architecture which has 512 layers. The decoder, with skip connections between the encoder, brings the number of channels down to the target image size.

SkyNet-LSTM incorporates a similar encoder-decoder architecture using 2 ConvLSTM cells for the encoder.

Loss functions. The forecasting model enforces the predicted frames to be close to their ground truth in the spatial space as well as enforcing the optical flow between the predicted frames to be close to their optical flow ground truth as well. This is done by imposing a combination of penalties as network loss functions between the predicted frame $\widehat{I}_{t+1}$ and ground truth I_{t+1}. The network is trained using three loss functions based on intensity, gradient, and motion. The intensity loss ensures that pixels in the RGB space are similar by minimizing the ℓ_2 distance between $\widehat{I}$ and I:

$$L_{int}(\widehat{I}, I) = \|\widehat{I} - I\|_2^2. \tag{3.14}$$

When forecasting frames using the standard Mean Squared Error (MSE) loss function, the predicted images are blurry. This is due to the fact that MSE generates the expected value of all the possibilities for each pixel independently which causes a blurry Image. Therefore, the gradient loss is used to sharpen the predicted image:

$$L_{gd}(\widehat{I}, I) = \sum_{i,j} \| \mid \widehat{I}_{i,j} - \widehat{I}_{i-1,j} \mid - \mid I_{i,j} - I_{i-1,j} \mid \|_1 + \| \mid \widehat{I}_{i,j} - \widehat{I}_{i,j-1} \mid - \mid I_{i,j} - I_{i,j-1} \mid \|_1, \tag{3.15}$$

where i and j are the spatial indices of the image.

To predict an image with the correct motion, we place a loss on the optical flow field generated by the predicted image and the input image. We employ a pre-trained CNN [Hui *et al.*, 2018] for the optical flow estimation. Denoting f_{op} as the optical flow network used, the motion penalty is expressed as:

$$L_{op} = \|f_{op}(\widehat{I}_{t+1}, I_t) - f_{op}(I_{t+1}, I_t)\|_1. \tag{3.16}$$

The three functions above are combined to define the overall loss function as:

$$L = \lambda_{int} L_{int}(\widehat{I}_{t+1}, I_{t+1}) + \lambda_{gd} L_{gd}(\widehat{I}_{t+1}, I_{t+1}) + \lambda_{op} L_{op}(\widehat{I}_{t+1}, I_{t+1}, I_t). \tag{3.17}$$

We define λ_{int}, λ_{gd}, and λ_{op} as 0.5, 0.001, 0.01 respectfully.

Training Details. Our implementation of the models are in Python using the PyTorch framework [Paszke *et al.*, 2017]. Training until convergence ends around 40 epochs with a learning rate of 0.001 using Adam optimization [Kingma and Ba, 2015] and a batch size of 4. We run all of our experiments on 3 NVIDIA GeForce RTX 2080 Ti GPUs which takes about 1.5 hours per epoch to train.

Long-Term Forecasting. To forecast a sky image frame longer into the future, we implement a simple recursive method. Once we have a prediction for $\widehat{I}_{t+1}$, to predict the image at time $t+2$, we use the image set $\{I_{t-4}, I_{t-2}, I_t, \widehat{I}_{t+1}\}$; that is, we use the predicted image at $t+1$ to recursively predict the next image in the sequence. We can repeat this multiple times to increase the time horizon of the predictions.

3.5 Experiments

We compare our method to prior deep-learning approaches to model cloud evolution in sky images along with the benefit of warping the sky images to achieve better long-term prediction.

3.5.1 Sky-Image Dataset

We use a publicly available dataset of TSI images for training and evaluation. The source of the dataset is a TSI located on the Nauru Island and available for download at the Atmospheric Radiation Measurement facility [Victor, 2005]. Images in the dataset were captured over a duration spanning November 2002 to September 2013. Each successive image pairs are 30 seconds apart and are at a resolution of 352×288 pixels. In total, the dataset includes $4,272,938$ images. However, for our study, we utilize a subset of the available data as our primary train and test sets. We utilize $42,171$ images from the year 2002 for training and validation and a disjoint set of $5,271$ images from 2003 for testing. Figure 3.2 shows sample images from the dataset.

3.5.2 Comparison to Previous Methods

Figure 3.1 provides qualitative comparison between the SkyNet predictions, as well as basic optical flow-based prediction using a constant velocity model, and the PhyD-Net-Dual approach [Le Guen and Thome, 2020]. As is seen in Figure 3.1, SkyNet predictions are of a significantly higher quality than the competitors. We provide quantitative evaluation in the form of Peak-signal-to-noise-ratio (PSNR) for these competing methods in Figure 3.7. Here, we also compare with a version of the SkyNet models without the optimal warping applied to it to study the influence that the warping function has. We observe that there

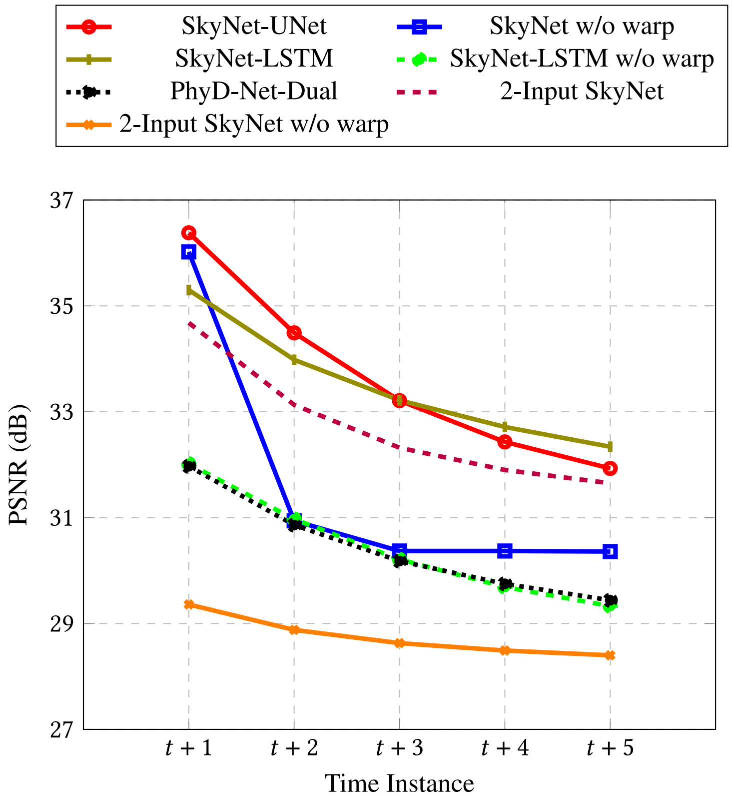


Figure 3.7: **Forecasting results.** Performance of various methods for sky image forecasting for time a time horizon of $t + 1$ to $t + 5$.

is a significant drop in performance when forecasting without the warping function; especially looking beyond the first predicted image.

We also compare against a two-frame version of SkyNet-UNet, both with and without spatial warping, to test the effectiveness of using a larger time horizon. In this version, we only provide $\{I_{t-1}, I_t\}$ as inputs to the network. As we expect, the performance of the prediction drops by a small amount when given a smaller past horizon, and by a larger amount when we disable spatial warping.

It should also be noted from Figure 3.7 that although the SkyNet-UNet model performs the best at time instance t+1, as the images are forecasted longer into the future, the SkyNet-LSTM model outperforms all other tested models. This may be attributed to the long-term memory units of the convLSTM network.

PSNR values in dB					
	$\widehat{I}_{t+1}$	$\widehat{I}_{t+2}$	$\widehat{I}_{t+3}$	$\widehat{I}_{t+4}$	$\widehat{I}_{t+5}$
1K Dataset	31.24	31.05	30.94	30.88	30.84
10K Dataset	32.16	31.71	31.4	31.19	31.05
100K Dataset	33.2	32.64	32.28	32.03	31.84

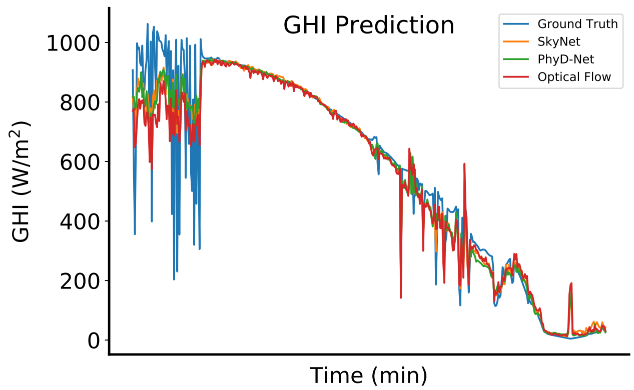
Table 3.1: **Affect of dataset size on results.** Comparison of PSNR for various dataset sizes of 1K, 10K, and 100K samples.

3.5.3 Dataset Size Dependent Results

We also experimented with the affect that the size of dataset had on prediction results. As shown in Table 3.1, as we increase the amount of training data from 1000, 10000, and 100000 respectfully, the model performance increases. This is due to the fact that more data allows the model to generalize to unseen samples. At the same time, although we train on a small subset of the sky-image dataset, there are upwards of millions of images in total that can be utilized for training. Therefore, with the right amount of training, our method can be improved even further.

3.5.4 GHI

Prediction of cloud movement in a subsequent image is only one step in precise prediction of solar irradiance. Therefore, using the predicted frames, we calculate GHI values similar to [Al-lahham *et al*., 2020] in order to validate our results on accurately forecasting solar irradiance. GHI values were captured at the same site that the TSI images were taken. We use a random forest (RF) ensemble model that, when trained with ground truth GHI values and sky-images, predicts a GHI value for that time instance. Table 3.2 shows predicted GHI values captured each minute on a day's worth of data. Given 4 previous sky-image frames at time instants $\{t-5, t-3, t-1, t\}$ as input, the predicted image frame at $\widehat{I}_{t+1}$ is used as inference to compute GHI. This is repeated for all time instances throughout the specified day. Table 3.2 shows comparison signal-to-noise (SNR) metrics for longer term prediction. Time instances after $t+1$ are recursively forecasted using predicted frames.



GHI prediction performance (dB)				
	$t+1$	$t+3$	$t+5$	$t+7$
SkyNet	**18.45**	**16.24**	**14.40**	**13.27**
PhyD-Net	17.54	15.27	13.66	12.62
Optical Flow	17.53	15.76	14.34	13.83

Table 3.2: **SkyNet GHI predictions.** (Top) Predicted GHI values captured each minute on August 6, 2003. Each minute interval of GHI is predicted using 4 previous time instances. The table below shows comparison metrics. (Bottom) Comparison of the signal-to-noise-ratio for GHI. Due to the fact that ground truth GHI values are captured each minute, we must predict every other subsequent image frame. For example $t+5$ represents a 5-min ahead forecasting time using 4 previous time instance frames. These values are averages over 5 days from 08/06/2003 to 08/11/2003.

3.6 Discussion

Limitations. Although SkyNet improves upon previous works modeling cloud dynamics, our method has limitations. First, due to the fact that we are using a learning-based algorithm, we are restricted to modeling clouds in the image intensity space where physical factors are not measured. Second, our model is also dataset dependent, inferring that sky images captured using a different camera than a TSI will require retraining on that camera specific dataset.

Overall, within this chapter, we presented SkyNet which improved sky-image prediction to model cloud dynamics with higher spatial and temporal resolution than previous works. Our method handles distorted clouds near the horizon of the hemispherical mirror by partially warping the sky images during training to facilitate longer forecasting of cloud evolution. Although our method performs well, the textures are still blurred near the horizon which is hard to undo and further degrades when predicting longer into the future. Possible future works can plan to move away from the RGB image space and capture the 3D distribution of clouds. This will allow the understanding of the absorption and reflectance properties of clouds across a large scale to better attenuate how they affect the amount of solar radiation being received at the ground. Also, we would like to develop computational imaging approaches that captures wide-angle FOV images without the expense of objects being distorted near the horizon. In chapter 4, we actually present this follow-up work of a deployed computational sky imaging setup and present algorithms in chapter 5 for accurate prediction of solar irradiance.

Above all, we believe the methods presented in this chapter is the first step toward precise prediction of solar irradiance to enable the widespread use of solar power both commercially and residentially.

4 A Catadioptric Sky Imaging System

In the previous chapter, we discussed an initial warping method to combat the non-linear spatial resolution stemming from traditional sky imagers. Digital warping does help alleviate some of the underlying challenges in non-uniform flow estimation, but it is fundamentally limited by the loss of resolution at image formation. Adding more pixel by using more cameras or even a higher resolution sensor can be an effective approach, but comes with increased costs. Further, direct imaging of the sky needs to be done with some care, given that the potential damage to the sensor caused by a focused image of the sun. We instead pose a different question: is it possible to *optically redistribute* the pixels in a wide FOV camera so that resolution is uniform for a cloud as it traverses the field of a sensor?

In this chapter, we take a different approach than before and *optically* warp the scene by designing a catadioptric system that provides a uniform spatial resolution of the sky (for each height), over the entire field of view of the device. We achieve this by imaging the sky through a mirror whose shape is designed to provide the aforementioned property. This design also has the added benefit of making motion of the clouds equally perceptible, be it at the zenith or the horizon. As a result of using this mirror shape in a catadioptric setup, our ability to estimate cloud trajectory is improved over traditional methods even when a cloud is farther away. This improves long-term cloud evolution prediction and as a result, prediction of when a cloud will occlude the sun.

Contributions. We present a method that advances long-term forecasting of cloud evolution and enables predictions far beyond previous works into the 10s of minutes. Our contributions for this chapter are as follows:

- *Imaging system for whole sky imager*. We have designed and deployed a novel sky imager comprising of a catadioptric system with an adapted hyperboloidal mirror [Baker and Nayar, 1998] to capture and analyze sky images with the eventual goal of improved solar forecasting.

- *Dataset of sky images*. Using our imaging system, we have captured high dynamic range sky images across a period spanning many months. The dataset also provides time-synchronized ground solar irradiance captured using a pyranometer.
- *Predicting from spatial-temporal slices.* We have a novel prediction algorithm that uses estimated wind velocity to identify an informative 2D space-time slice of the imagery; this allows us to ignore the clouds that are unlikely to occlude the sun at our vantage point. More importantly, it significantly simplifies the resulting prediction problem, which we perform using a learned-network.

This *long-term prediction* is an order of magnitude improvement over previous methods such as in chapter 3. That method relied on a similar premise, but with digital warping that showed prediction results that only span 2-3 minutes.

4.1 Prior Work

Large Field-of-View Sky Imagers. As previously stated, a limitation of sky imagers is the non-linear fisheye distortion introduced which affects clouds optical flow estimates for trajectory prediction. The further-out clouds attenuate evolution over time and their prior motion estimates attribute to longer forecasting horizons. To combat this issue, many works have attempted to spatially warp these images to achieve uniform apparent motion and limit the apparent fisheye distortion [Paletta *et al*., 2022, Rajagukguk *et al*., 2021, Richardson *et al*., 2017]. In chapter 3, we have even shown that spatially warping these images achieves longer forecasting horizons. Limiting the long-term prediction accuracy when using digital warping is the loss of resolution of pixels at the periphery [Eising *et al*., 2008, Ishii *et al*., 2003]. Think of it as digital zoom versus optical zoom. Periphery pixels are stretched out and then interpolated. Therefore, true pixel values at these locations are absent. For learning and optical flow based methods, these pixel values are essential for accurate long-term prediction.

Predicting Cloud Movement. Modeling cloud evolution through tracking and forecasting clouds solely using sky images are achieved using mainly two methods. Initial works utilized motion based methods [Ai *et al*., 2017b, Chang *et al*., 2017, El Jaouhari *et al*., 2015, Jayadevan *et al*., 2012] using pairs of subsequent sky images to forecast cloud trajectory. Overall, this method is inadequate for long-term prediction due to the variability of cloud shapes and trajectory between image captures making it difficult to forecast.

More recent works have seen greater success using deep learning based methods to predict a subsequent sky image for a future time instance [Julian and Sankaranarayanan, 2021, Nie *et al*., 2023, Paletta *et al*., 2022, Sun *et al*., 2014, Wei *et al*., 2023]. Learning-based methods can be further improved for more accurate long-term forecasting by addressing the fisheye distortion of the scene introduced by traditional large-FOV imagers. However, in this chapter, we show that optical warping leads to even greater success when coupled with learning-based predictions.

Photovoltaic Power Output Prediction. Directly predicting photovoltaic power output has been a direction taken by previous works. These studies either take a statistical approach to predicting future irradiance values from past values [Alzahrani *et al*., 2017, Sharifzadeh *et al*., 2019] or finding the relationship between an associated sky image and its irradiance value; also known as nowcasting [Zhen *et al*., 2020], [Zhang *et al*., 2018]. However, these works do not take the future state of cloud patterns into consideration which directly influences the amount of irradiance received at the ground. Therefore, by having an accurate method of predicting the future distribution of clouds, a better estimate of future irradiance can be obtained.

Computational Imaging for Atmospheric Tomography. Many works attempt to image clouds as a true three-dimensional (3D) volumetric matter rather than two-dimensional (2D) beings in images. These works fall under the category of solutions with the goal of tomographic reconstruction. They consider the heterogeneous multi-scattering media and reconstruct the full volumetric field using distributed ground-based camera systems [Aides *et al*., 2020, Holodovsky *et al*., 2016, Mejia *et al*., 2018, Veikherman *et al*., 2014], or airborne imagery [Diner *et al*., 2018, 2013, Martonchik *et al*., 1998]. Although beneficial for true analysis of solar irradiance transmission through atmospheric media, we believe that simple 2D RGB images are sufficient for long-term prediction of sun occlusion by cloud for this application.

Catadioptric Imaging Systems. Catadioptric imaging which utilizes the reflective nature of mirrors during the acquisition process are designed such that their shape can achieve various tasks [Baker and Nayar, 1998]. In particular, the hyperboloidal shape provides practical wide-angle imaging with minimal distortion and solves to the aforementioned challenges. This selected mirror shape will be further discussed in the subsequent sections.

4.2 Problem Definition

In this section, we introduce the problem of sky imaging by stating the desired specifications and discussing the gaps between these requirements and current wide FOV imagers.

4.2.1 Design Specifications

Prediction of cloud movement requires precise estimates of their velocities when they are at the periphery of the field of view. For example, a cloud with a typical height of 1km and a velocity of 50 km/hr, will cover an angle of $\tan^{-1}(25/1) = 87°$, in a camera's view, over half an hour. If the sun is at the zenith, to provide a reliable prediction half-an-hour in the future, we need a $174°$ field of view, as well as the ability to precisely sense motion when the cloud is near the horizon. This provides us with our design specifications: an imager with extremely large field of view approaching $180°$, while providing the ability to detect and estimate cloud motion over the entire field. We interpret the second part of the specification as one of providing uniform spatial resolution for the cloud as it appears and traverses the FOV of the imager. While uniform resolution by itself is not a necessary condition (for example, we could ask for higher resolution at the horizon over the zenith), it allows for a robust solution that can also accommodate cloud creation events within the FOV.

4.2.2 Gaps in Current Sky Imagers

Current wide FOV imagers can be built with a fisheye lens or more commonly with a catadioptric system where the sky is imaged through a hemispherical mirror. Such traditional sky imagers are not conducive for long-term prediction due to their lack of resolution at the periphery of the imager. Specifically, in such systems, an object placed at the zenith of the sky will appear to have a larger total spatial extent as opposed to the same object at the horizon. Figure 4.1 visualizes this circumstance via a large checkerboard placed above a simulated hemispherical mirror. The checkerboard, which has a length and width of 50 km, is placed 2 km high above the mirror where each square is uniformly spaced at 1 km per square space. This hemispherical setup shows that squares at the zenith of the imager appear larger compared to squares at the periphery, despite their physical dimensions being the same. This compression of the squares at the horizon translates to poor localization of clouds at the horizon in the world. Another related factor is our ability to estimate motion. In current sky imagers, motion of clouds appear to be non-uniform despite their physical speed being largely the same (since clouds are driven by wind), with large apparent motion at the zenith and significantly smaller ones at the horizon. In more practical terms, the imagery in current images only allow for precise estimates of cloud velocity only after it is significantly away from the horizon; in turn, this limits the time horizon over which sun occlusions can be predicted.

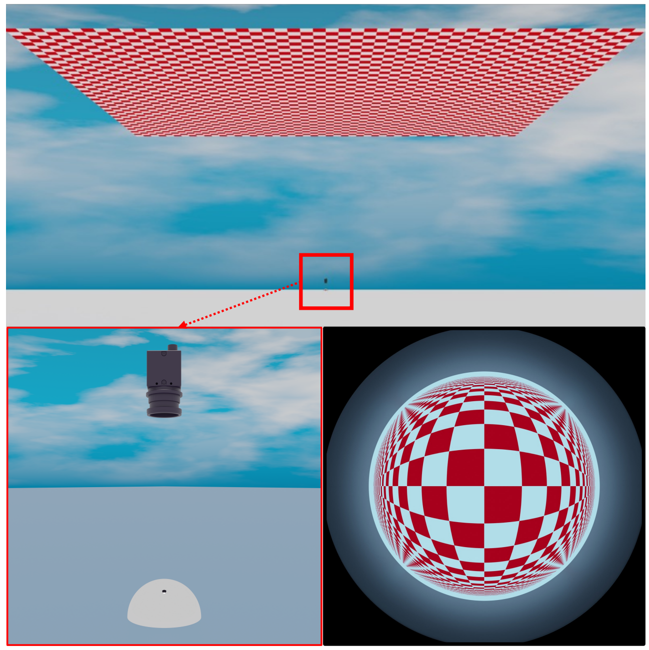


Figure 4.1: **Simulated hemispherical imaging setup.** A rendering obtained using Blender to visualize the non-uniform resolution of a hemispherical mirror. (Top) We create a scene consisting of a checkerboard with a length and width of 50 km, placed 2 km high above the ground. Each square on the checker board has physical extent of 1 km. (Bottom-left) Our imaging system consists of a pinhole camera observing the sky or the checkerboard indirectly through a hemispherical mirror. (Bottom-right) The image observed on the camera has high resolution at the zenith of the image and significantly lower resolution at the periphery.

4.2.3 Solution Outline

Our goal is to address the limitations of current sky imagers which achieve a large viewing angle at the cost of two issues that limit long-term cloud motion prediction: lack of resolution at the horizon, and non-uniformity of motion. How can the problem of non-linear motion and lack of pixel resolution be circumvented?

Our approach relies on the insight that we can redesign the mirror used in a sky imager to spatially redistribute the pixels with the eventual goal of having the same spatial resolution on a cloud over the FOV of the imaging system—immaterial of whether the cloud is at the horizon or at the zenith. This allows for early detection of clouds, as well as simplifies the motion estimation problem since the clouds largely translate over the field of view.

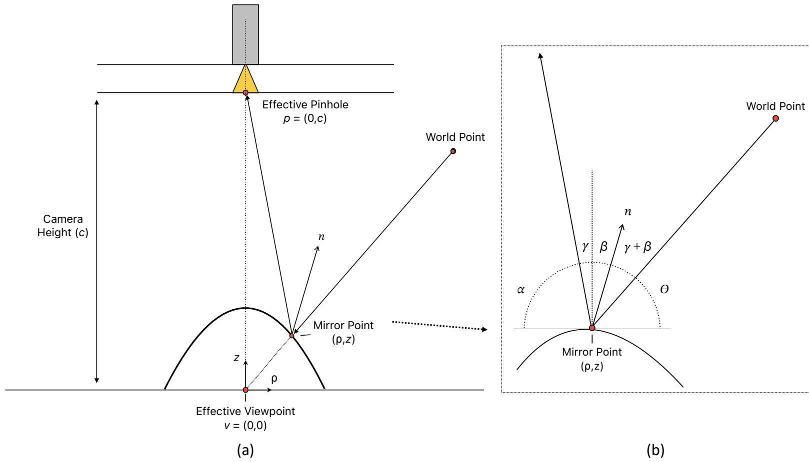


Figure 4.2: **Catadioptric setup.** Visualization of the equations presented in section 4.3

The second part of our solution, discussed in chapter 5, is an algorithmic technique for prediction over time-horizons of tens of minutes. Part of the challenge here is the high-dimensionality of the input image which makes any learning-based solution hard to implement due to compute and memory requirements, as well as the need for a large amount of input data. To simplify this problem, we argue that cloud motion due to wind is largely translational; hence, to predict the occlusion of the sun as well as solar irradiance at future time instants, it is sufficient if we look at a spatial slice through the sun that is parallel to the wind direction. While this likely misses out on predicting irradiance due to indirect skylight, it has all the relevant information for predicting direct sunlight which is the dominant term in the overall irradiance.

Finally, we build and deploy a test bed, and collect a dataset of sky images over a period of months. We evaluate our algorithms over this dataset, as well as a synthetic counterpart, in Section 5.

4.3 Mirror Design

We frame the problem of mirror design as one that *flattens* the sky image formed on the sensor. Figure 4.2 illustrates the relevant variables. Here, we delve into the derivation of the general mirror shape profile, which share the same goals as [Baker and Nayar, 1998], followed by the specific solution for the hyperboloidal case.

For a catadioptric system, it is desirable to have a single center of projection, also called the *fixed viewpoint constraint*, for geometrically corrected images. This constraint requires that in 3D-space, the catadioptric sensor only measures the intensity of light passing through a single point. This constraint

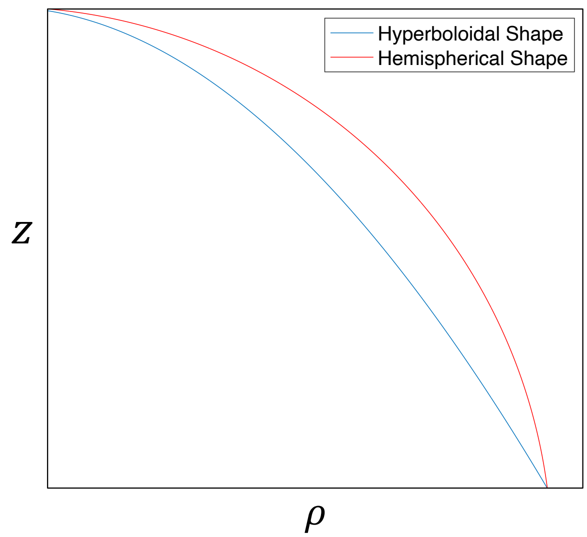


Figure 4.3: **Mirror shape comparison.** We show the plotted equation of the line forming the hyperboloidal mirror shape. We also show the hemispherical shape as a comparison.

is laxed for the direction of light passing through this point. The location at which this 3D point is sampled is known as the *effective viewpoint*. As shown in Figure 4.2(a), this constraint also requires that each reflected ray of light that passes through the pinhole of the camera would have passed through the effective viewpoint if not reflected by the mirror.

With Figure 4.2(a) and 4.2(b) as a visualization of the formation for these equations, we can now go into the derivation of the fixed viewpoint constraint. The effective viewpoint $v = (0, 0)$ lies at the origin of the z, ρ-axis in the 2D cartesian coordinate frame. The effective pinhole $p = (0, c)$ is located at the camera height c. With this, the goal is to find the 2D mirror profile

$$z(\rho) = z(x, y). \tag{4.1}$$

Due to the fact that an incoming angle θ from the world intersects the mirror at a point (ρ, z), we assume that it also passes through the origin at v and reflects at an angle α. γ and β are the angles between the z-axis of the reflected and the normal at a point (ρ, z) respectfully. Shown in Figure 4.2(b), this results in the constraint that

$$\theta + \alpha + 2\beta + 2\gamma = 180^\circ. \tag{4.2}$$

Based on the law of reflection, $\gamma = 90^\circ - \alpha$, which can be plugged into (4.2) to reach

$$2\beta = \theta + \alpha. \tag{4.3}$$

Taking the tan of both sides and using the *sum of angles* and *double angle* trigonometric identites

$$\frac{2\tan\beta}{1-\tan^2\beta} = \frac{\tan\theta - \tan\alpha}{1+\tan\theta\tan\alpha}. \tag{4.4}$$

We can now also define the relationship with these angles and the current geometry:

$$\tan\theta = \frac{z}{\rho}. \tag{4.5}$$

$$\tan\alpha = \frac{c-z}{\rho}. \tag{4.6}$$

$$\tan\beta = \frac{dz}{d\rho}, \tag{4.7}$$

which can be plugged into equation (4.4) to get the fixed viewpoint constraint:

$$\frac{2\left(\frac{dz}{d\rho}\right)}{1-\left(\frac{dz}{d\rho}\right)^2} = \frac{\left(\frac{z}{\rho}\right)-\left(\frac{c-z}{\rho}\right)}{1+\left(\frac{z}{\rho}\cdot\frac{c-z}{\rho}\right)}. \tag{4.8}$$

The fixed viewpoint constraint can be rearranged into a $1st$-order differential equation in which the solution can be found by solving it as a quadratic to obtain the expression for the surface slope. We can ultimately find the general solution of the fixed viewpoint constraint in the form of two equations

$$\left(z-\frac{c}{2}\right)^2 - \rho^2\left(\frac{k}{2}-1\right) = \frac{c^2}{4}\left(\frac{k-2}{k}\right)\ (k \geq 2)\,. \tag{4.9}$$

$$\left(z-\frac{c}{2}\right)^2 + \rho^2\left(1+\frac{c^2}{2k}\right) = \left(\frac{2k+c^2}{4}\right)\ (k > 0)\,, \tag{4.10}$$

where $k = 2\exp^C > 0$, a constant, and C is the constant of integration. As a result, equation (4.9) and (4.10) form the mirror shape profiles as a 2-parameter function of k and c. A number of solutions can be created to construct the mirror profile using one of the two equations, however for the hyperboloidal solution, equation (4.9) is used with values of $k > 2$ and $c > 0$.

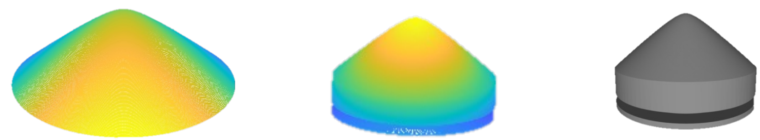


Figure 4.4: **Simulated Hyperboloidal Mirror.** (Left-to-right): Surface plot of hyperboloidal shape as a line spun around it's axis. Surface plot of final mirror shape. 3D mesh of mirror shape.

4.3.1 Hyperboloidal Solution with Perspective Camera

In the 2D (z, ρ) cartesian frame, we initially specify the parameters critical for the mirror development. Our basic setup is that of a pinhole camera with a sensor size of $w = 12.5$mm, placed at a distance $c = 1$m from the mirror, with a field of view of $f_c = 200$mm, and mirror curvature of $k = 38$. These choices are based on design considerations for the final implementation where we need the camera to be sufficiently far away so as to avoid blocking a significant portion of the FOV. The long focal length also allows us to effectively mimic the pinhole camera with a lens-based counterpart. The mirror has a shape $z = f(\rho)$, where ρ is the radial distance over the ground plane. We make an additional assumption that the cloud is at some height h; the exact height of the clouds do not play an actual role as we will assume that $h \gg c$ and so only the tangent of the angle subtended by the cloud at the mirror matters. With this, we formulate the mirror design as one of designing the profile $f(\cdot)$ such that the effective sky to sensor mapping is a *scaling operation over the desired field of view*. Effectively, we are scaling the FOV of the camera—which is $\theta_{cam} = 3.58°$—by a constant spatial factor to achieve a target FOV $\theta_{target} = 170°$ and focal length $f_t = 546.8mm$. The detailed derivation of the hyperboloidal solution starts by modeling rays coming from the world. Every incident ray coming from the world hits the mirror at an angle

$$\theta_o = atan\left(\rho_{vec} \cdot \frac{w_c}{f_t}\right), \tag{4.11}$$

with the desired outgoing ray being reflected at the mirror surface at a point P and reaching the pinhole of the camera at an angle

$$\theta_i = atan\left(\rho_{vec} \cdot \frac{w_c}{f_c}\right). \tag{4.12}$$

As a result,

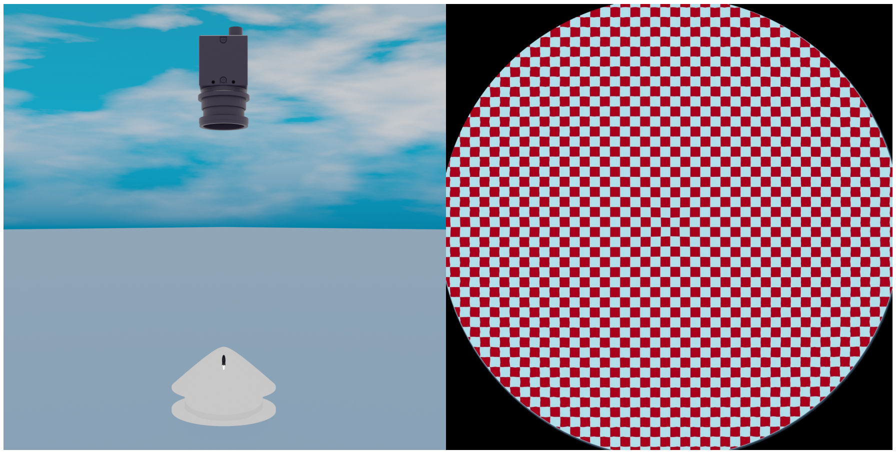


Figure 4.5: **Simulated hyperboloidal imaging setup.** Simulated case visualizing the uniform resolution of our hyperboloidal mirror. (Left) The same parameters as Figure 4.1 with the proposed mirror replacing the hemispherical mirror. (Right) Observe how the checkerboard resolution is spatially uniform—a consequence of the system acting as overall scaling operation.

$$\theta_n = \frac{(\theta_o - \theta_i)}{2}, \tag{4.13}$$

where ρ_{vec} is the radial distance along the ρ-axis to the point at which the reflected ray reaches the sensor. Now, we can combine the equations to formulate the desired results. The point P at the intersection of a ray and the mirror surface can be represented as the 2D parameter:

$$\left[-z(\rho) \cdot \left(\frac{\rho_{vec} \cdot w_c}{f_c}\right), z(\rho)\right]. \tag{4.14}$$

The gradient at this point can be represented as the 2D parameter:

$$\left[-\nabla z(\rho) \cdot \left(\frac{\rho_{vec} \cdot w_c}{f_c}\right) - z(\rho) \cdot \left(\frac{w_c}{f_c}\right), \nabla z(\rho)\right]. \tag{4.15}$$

With these points, we can set

$$\tan\theta_n = \frac{-\nabla z(\rho)}{-z(\rho) \cdot \left(\frac{\rho_{vec} \cdot w_c}{f_c}\right) - z(\rho) \cdot \left(\frac{w_c}{f_c}\right)}, \tag{4.16}$$

which can be simplified to reach,

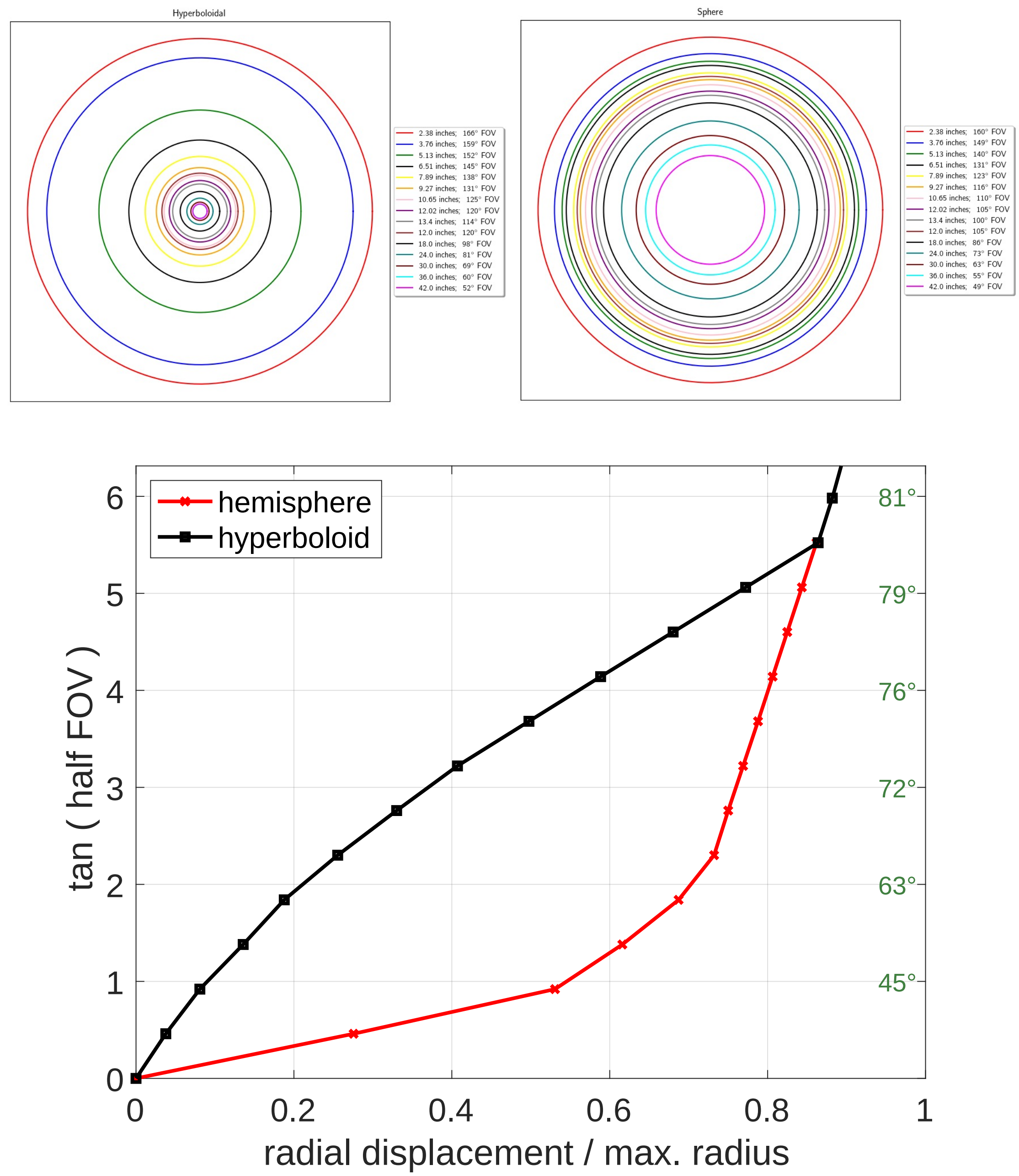


Figure 4.6: **FOV of real setup.** (Top) By placing a light source at various heights along the frame of our real setup, we are able to measure the FoV that the hyperboloidal (left) and hemispherical (right) mirror sees. Within the legend, the left column represents the height of the light source from the base of the system and the right column represents each FOV. (Bottom) Field of view of the sky (in tangent of angle) as a function of radial distance from center for the hyperboloidal and hemispherical mirrors in our deployed system.

$$\frac{\nabla z}{z} = -\frac{\left(\frac{w_c}{f_c}\right) \cdot \tan\theta_n}{\left(\frac{\rho_{vec} \cdot w_c}{f_c}\right) \cdot \tan\theta_n - 1}. \tag{4.17}$$

We can integrate the right hand side of the equation (4.17) such that

$$\ln z(\rho) - \ln z(0) = \int_0^\rho (\cdot)\, d\rho. \tag{4.18}$$

This equation can be solved using the cumulative sum of all values. With this, we can formulate the range for ρ as

$$\left[\min\left(-\rho_{vec} \cdot z_0 \left(\frac{w_c}{f_c}\right)\right), \max\left(-\rho_{vec} \cdot z_0 \left(\frac{w_c}{f_c}\right)\right)\right]. \tag{4.19}$$

$$z_0 = H \cdot \exp\left(\int_0^\rho (\cdot)\, d\rho\right), \tag{4.20}$$

where $H = 1$, the camera height in meters. To determine the mirror shape, we use a numerical procedure where we solve for the axial profile $f(\cdot)$ by densely ray tracing over the image plane. With this, we also have the constraint that the ray — after mirror reflection — behaves like a pinhole camera with the target field of view. This provides a constraint on the derivative of f (since the surface normal is determined by the normal). Integrating this derivative provides us with the desired shape. A visualization of the resulting mirror shape is shown in Figure 4.4.

In Figure 4.5, we use Blender to render an identical setup as Figure 4.1, but with the hyperboloidal shaped mirror. Our mirror achieves a uniform image of the checkerboard while maintaining a large FOV, showing that we are able to image the sky with uniform resolution. As is to be expected, this design also enables uniform motion estimates throughout the whole FOV of the imager.

For the real setup, we measure the FOV by placing a light source at various distances from both mirrors. We show the comparison between the hemispherical and hyperboloidal mirror in Figure 4.6.

4.4 Testbed

We now describe our imaging setup in the context of our hyperboloidal-based mirror.

4.4.1 Simulation Setup and Dataset

We initially evaluate and report results of our setup and methods on simulated data which achieves the idealized scenario of a real-world setup with known parameters. The platform used to develop our simulated

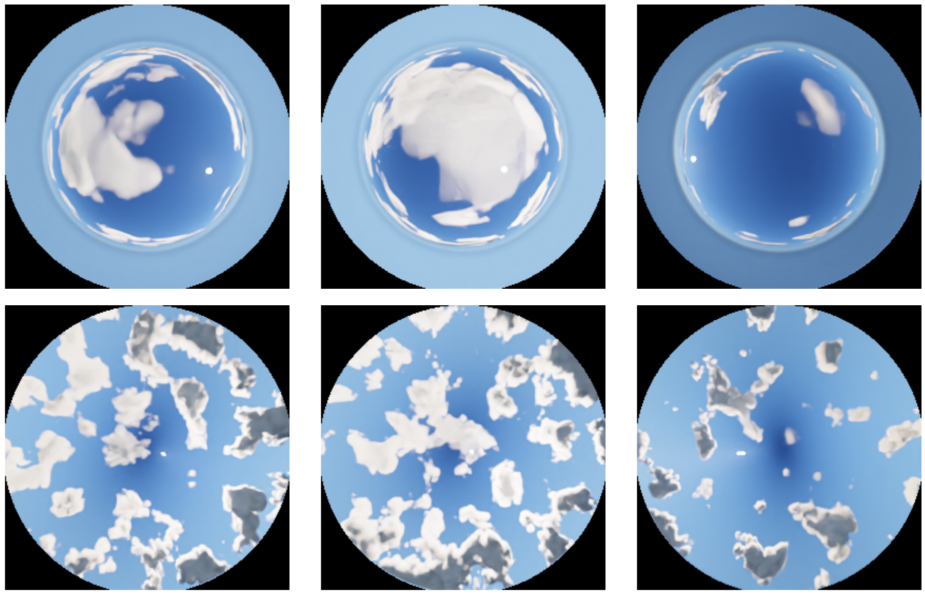


Figure 4.7: **Simulated dataset samples.** Various images captures from the synthetic dataset. (Top) Captures from hemispherical setup. (Bottom) Captures from hyperboloidal setup. Each column is captured at the same time instant

data is Blender which uses the same catadioptric setup as in our real-world data. In Blender, the Pure-Sky Pro package which simulates an array of cloud formations inspired by [lin, [n.d.]] is used. Although not modeled as mathematically in-depth as a large-eddy simulation [Mason, 1994], Pure-Sky Pro is accurate to the scale of this simplified simulated application. The package does allow for the modification of cloud dynamics such as how warm/cold air affects cloud evolution.

Using the computer generated hyperboloidal-mirror shape, as shown in Figure 4.5 placed with a reflective mirror material property, we capture simulated data on various cloud scenes with a sampling period of $T_0 = 30$ seconds. We also capture the same data using a hemispherical mirror, as in Figure 4.1, with the same parameters. These cloud scenes include randomized cloud parameters across a 28 day period from 8AM to 5PM based on real-world factors such as wind, hot/cold air patterns, and sunlight. Images of simulated data for both mirror setups are shown in Figure 4.7. Even in the simulated case, the benefits of our hyperboloidal imaging setup are clearly shown compared to traditional hemispherical imagers.

Physical prototype. To develop a physical prototype for our mirror, we began by computationally plotting the equation of a line for the mirror shape as described in Section 4.3 and visualized in Figure 4.3. Due to rotational symmetry, a surface of revolution can be formed about the vertical z-$axis$ to create the 3D

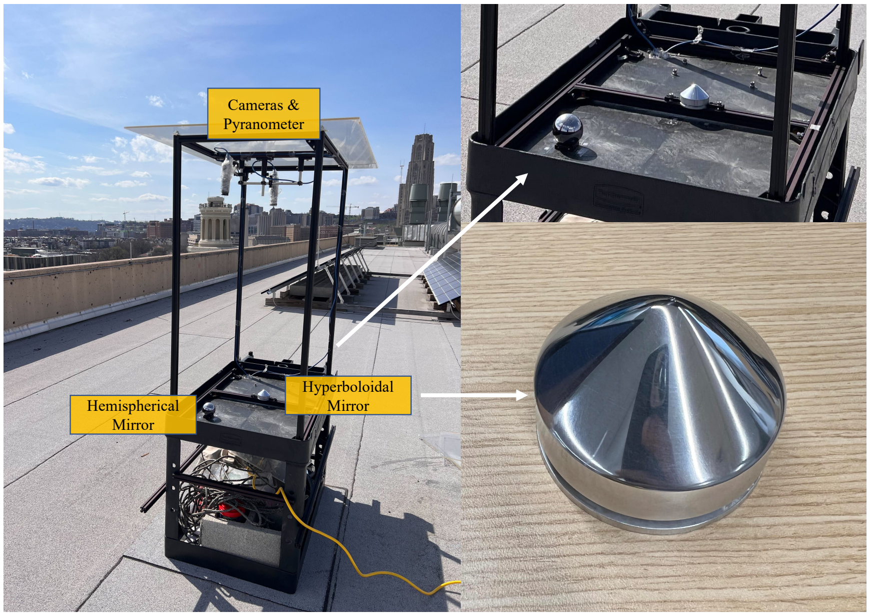


Figure 4.8: **Deployed testbed.** Our deployed test bed with a detailed visualization of mirror placement of the hyperboloidal mirror and hemispherical mirror.

shape of the hyperboloidal mirror. Now, armed with the mirror shape as a 3D surface, we can export the 3D shape to a format suitable G-code that can be used to fabricate metal using a Computer Numerical Control (CNC) machine. We experimented with 2 types of metal – aluminum and steel – and nylon polymer (plastic). The choice of material affected how obtainable it would be to get the surface reflective enough to be used in a catadioptric imaging setup. Steel is harder than aluminum which makes obtaining a reflective surface via polishing, grinding, or material subtraction more difficult than aluminum.

We initially attempted to obtain a reflective surface for the metal pieces by manually grinding the surface using sandpaper ranging form coarse-to-fine grit. After, we applied a polishing agent to buff the surface and reach the desired result. This method allowed us to obtain a reflective surface on the aluminum piece, but when placed in direct sunlight for image acquisition, the resulting image displayed severe artifacts which made it difficult for any inference. We deduced that grinding the surface of the metal using sand paper created micro-grooves within the surface. When imaged under bright scenes such as sunlight, this created a diffraction grating affect that oversaturated the whole image. We were then tasked with going back and figuring out another method of obtaining a reflective surface on the metal without any material subtraction methods such as sanding.

After much research, we found a solution that obtained a reflective mirror surface via a chemical

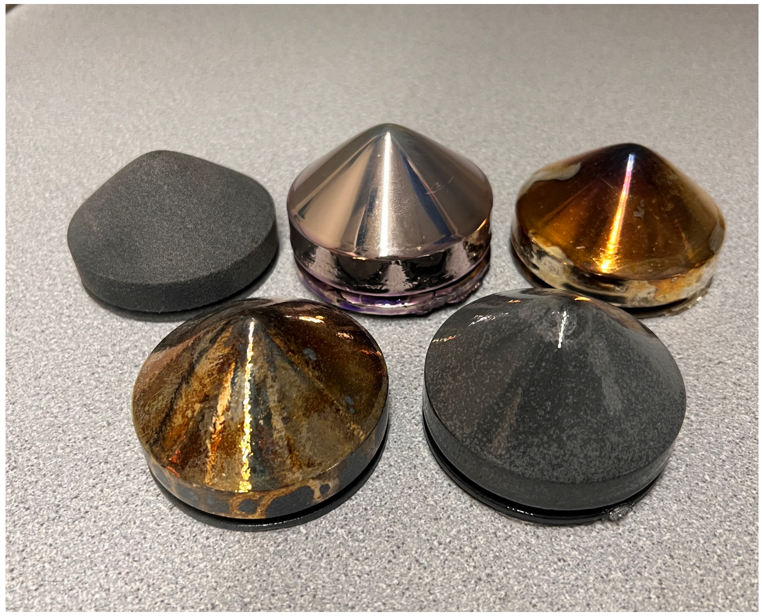


Figure 4.9: **Sample failed mirrors.** Examples of failed attempts when applying the reflective chemical deposition on the surface of steel, aluminum, and plastic pieces.

deposition process that is commonly used to coat diffuse plastic and metallic surfaces [sil, [n.d.]]. This process was very tedious and required precise mixing of numerous chemicals to obtain the desired results. Figure 4.9 shows failure cases on aluminum, steel, and plastic which was a cause of incorrect application. After many trials, we were able to coat the aluminum mirror with this chemical process and proceeded with our proposed imaging setup.

For image acquisition in our proposed solution, we utilize an RGB camera mounted on a cuboidal frame above the mirror. The mirror itself lies on the horizontal axis of the frame and coupled with a mini PC, captures sky images with a sampling period $T_0 = 30$ seconds. To minimize any nearby building occlusion, our imaging device is placed on a building roof and captures data continuously during daylight with a frequency of T_0. We also included a second-connected system with a hemispherical mirror for evaluating improvements of our proposed method. To handle the large dynamic range of the sky, due to the sun, we capture images using exposure bracketing and fuse them to get a single HDR image. The top of our system also includes a pyranometer that captures solar irradiance in the form of global horizontal irradiance (GHI) with the same T_0 interval. This setup is shown in Figure 4.8.

Comparison To Other Sky Imager Systems To highlight the benefits of our single-viewpoint catadioptric system with our mirror shape, we note the key differences of other possible solutions. In terms of imaging setup - one could utilize a multiple camera system that images the sky in different directions and composites the individual images together [Brown and Lowe, 2007, Dong *et al*., 2022, Nghonda Tchinda *et al*., 2023]. This, in theory, should limit the distortion caused by a fisheye lens. However, this increases costs due to multiple cameras being involved, another step needed to stitch images together, and the lack of mobility due to recalibration of the camera clusters if moved. Our system overcomes the challenges faced by other solutions by utilizing a simple and inexpensive single camera approach.

4.4.2 Dataset Collection

Figure 4.10 shows a gallery of real images captured from our setup. These images are captured from our hyperboloidal mirror and a hemispherical mirror at the same time instant. Similar to the simulated data, the real images achieve the benefits of using the hyperboloidal shaped mirror. We are able to see more clouds within a single capture and the motion is more translational through time. Imagery from October 20th, 2023 to March 5th, 2024 is collected for this dataset. We excluded days that were entirely cloudy or completely clear skied, so as to remove scenarios where GHI is nearly constant over the entire day. This left us with 76 days worth of data with most days having partly cloudy conditions.

4.4.3 Pre-processing

Before we can apply learning-based techniques on this dataset, we need to perform certain operations on it. In particular, knowledge of the sun as well as the wind velocity at each frame is helpful for the algorithms we describe in the subsequent chapter.

Sun localization. As a crude pre-processing step, we use the shortest exposure in our HDR stack to estimate the location of the sun. This makes identification easier as the sun is most likely the brightest object within the scene at a low exposure. However, this technique fails when the sun is occluded by clouds or when there are other bright reflections from other objects in the scene. To get a robust estimate, we pool the data across multiple days of maximal saturation and reject outliers using RANSAC. This provides a sun estimate as a function of daytime where occluded sun estimates are filled by fitting a polynomial function over sun locations identified by maximal saturation. Of course this will fail for incorrect predictions; therefore manual identification is required for some cases.

On an aside, the location of the sun in absolute angular coordinates with respect to the zenith of the sky can be analytically computed given the latitude and longitude of the testbed. We can in principle map such elevation and azimuthal position of the sun to the image plane coordinates via a calibration procedure. We

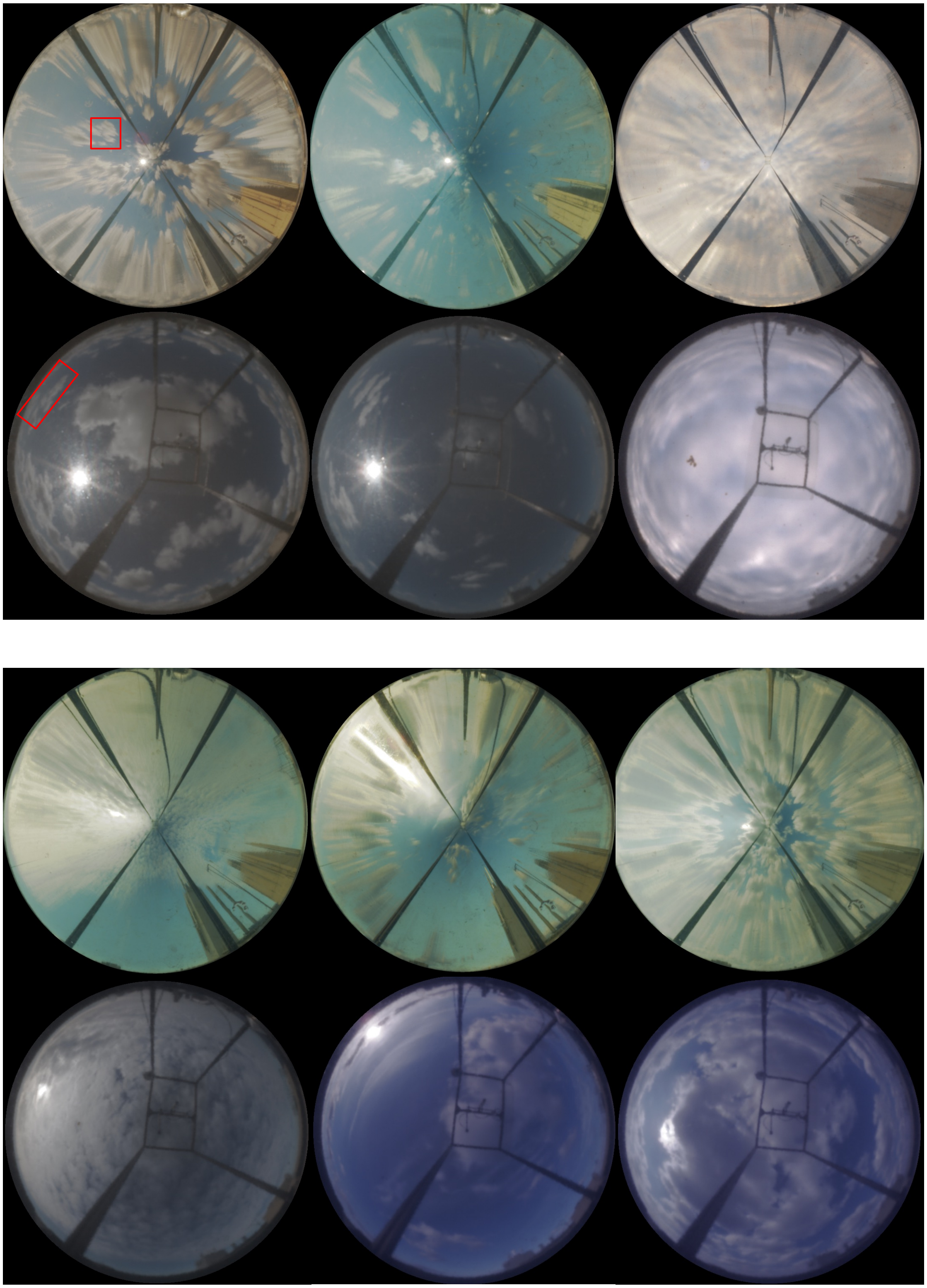


Figure 4.10: **Real dataset gallery.** Real images captured from various dates and weather conditions. (Odd-rows) Images from the proposed hyperboloidal setup. (Even-rows) Images from the hemispherical setup. Each column is captured at the same time instant. The cropped cloud in the red box in both images show the benefit of our mirror being able to image a cloud much further out.

opted for a simpler approach which does not require what we felt was a complicated calibration problem that had to account for the mirror shape.

Wind velocity estimation. Another useful information for the learning-based formulation that we will present next is the direction of wind velocity. A challenge here is that clouds are largely featureless, which makes traditional optical flow techniques fragile. Further, there are features in our field of view which are constant, for example buildings at the periphery and the frame used to hold the cameras. These static features bias the optical flow estimates especially since they are also high contrast ones.

To overcome these effects we use the mask to suppress the static regions and run the optical flow technique proposed by Liu [Liu, 2009] with very a strong weight associated with the spatial regularization term. Finally we use an aggressive temporal median filter on the estimated optical flow across frames to ensure a smooth flow field.

4.5 Discussion

Leveraging the benefits of our hyperboloidal-shaped mirror in a catadioptric imaging system, we are able to image the sky with uniform spatial resolution over the wide FOV of the imager. Now, we can exploit these benefits by discussing algorithms that can make accurate predictions over longer forecasting windows compared to traditional hemispherical imagers. We present these algorithms and results in the subsequent chapter.

5 Algorithms For Long-Term Forecasting

In the previous chapter, we presented the design and deployment of a hyperboloidal-shaped mirror in a catadioptric sky imaging setup to optically redistribute pixels over the FOV. In this chapter, we present algorithms that exploit the benefits of this optimal setup for accurate long-term irradiance forecasting. By using our mirror design, we show that we can see further out clouds, maintain resolution, and limit non-linear apparent motion which therefore improves predictions over the traditional hemispherical setup.

The sudden rise and fall of received solar irradiance at the ground within a short period of time is crucial information for electricity grid operators to mitigate disruptions in power output. This event, also called a ramp event (RE) [CUI, 2017, Godfrey *et al.*, 2010], is influenced by the occlusion state of the sun by a cloud. This is necessary to forecast and can be achieved through prediction of cloud trajectory. Thus, we show the benefits of utilizing our hyperboloidal-based imaging system and focus on a suite of algorithms that can exploit these benefits to better predict RE's.

5.1 Space-Time-Slice Image

The key benefit of having uniform apparent motion of clouds when using our system is that we are able to linearly back-trace the trajectory of clouds through time to attenuate a cloud's projected path toward the sun. We can interestingly use this fact and state that the only part of the image that is important is the sun and clouds that are moving towards it. We can disregard other parts of the image and simplify this problem. This will be done through the use of *space-time-slices* which summarizes cloud patterns throughout a day by simply using a single image (see Figure 5.1). A space-time-slice image takes a narrow-band slice of a full sky image and horizontally stacks the slices through time to form a single image. Each slice is taken from a distinct image at time instance t where the horizontal x-axis of the final space-time image represents time and the vertical y-axis represents space. We briefly describe how we create the space-time image.

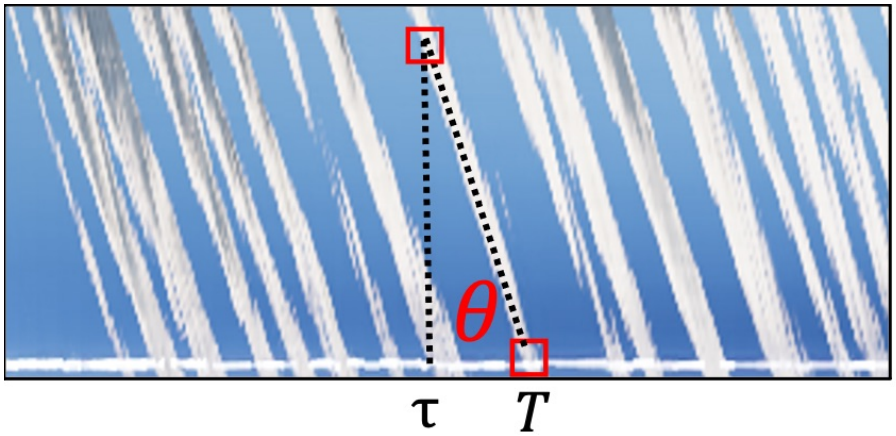


Figure 5.1: **Simulated space-time-slice.** We show that a cloud at time instance τ will occlude the sun at time instance T based on some angle θ. If we take use this fact, we can take this cloud at τ and warp it to where it will be in the future at T+N. This is the intuition behind our non-learning approach and how we warp the images as input to our learning-based models.

For each time instant for a single day, the x and y coordinate of the sun is initially identified. Next, the general direction of cloud motion ($\hat{\theta}$) through time is obtained. We take a sun-centered slice of the image in the direction of cloud motion for each time instant and horizontally concatenate these images through time to formulate the space-time image. Figure 5.3 visualizes each described state and shows the resulting space-time images.

For the case of the simulated sky images, we are benefited by having the ground truth sun location, direction of cloud motion, along with the binary sun occlusion state. Therefore, we have the necessary parameters for an ideal space-time-slice image. Figure 5.3 shows this space-time image and compares the image obtained from our system to the hemispherical-based system. Our system clearly achieves the desired linear apparent motion and is the ideal case for predicting sun occlusion discussed in section 5.2.1.

For the real images, however, ground-truth parameters are not given and are estimated. Sun location and general cloud direction is identified using the methods described in sub-section 4.4.3. In the real-world environment, cloud motion direction and velocity is not static and changes through time. We use an aggressive temporal median filter on the estimated optical flow across frames to ensure a smooth flow field and select a single ($\hat{\theta}$) for the whole day. We present sample space-time slice images for the real case in Figure 5.9.

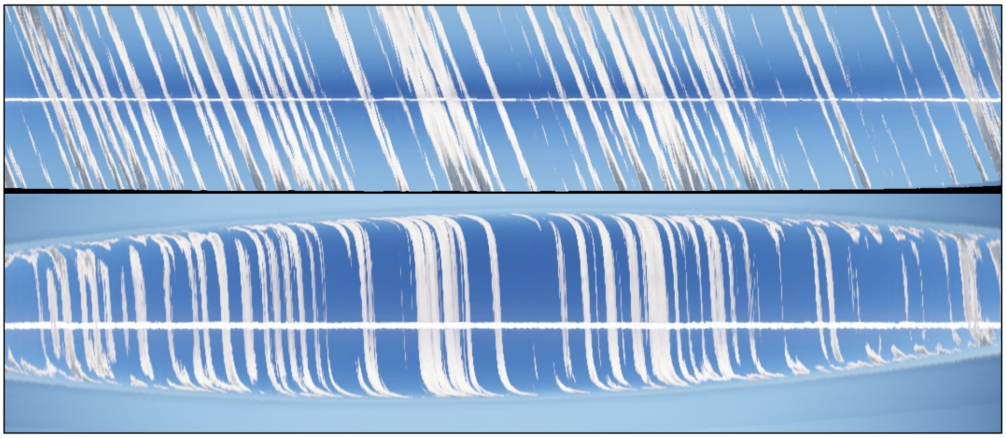


Figure 5.2: **Simulated space-time-slice comparison.** (Top) Space-time-slice image produced from hyperboloidal setup. (Bottom) Space-time-slice image produced from hemispherical setup. Notice the linear streaks present in the hyperboloidal case which can be exploited via our algorithms in Section 5.2 for accurate occlusion and irradiance prediction.

5.2 Real Setup

We now present non-learning and learning based algorithms that exploit the advantages of our optimal mirror setup.

5.2.1 Non-Learning Occlusion Prediction

Given that the ground-truth parameters for the simulated images are available, we are able to perform non-learning based sun occlusion and show the benefits of our setup.

Back Projected Sun Occlusion Prediction

Looking at the space-time image produced by our hyperboloidal shaped mirror, the linear streaks of clouds whose trajectory through time occludes the sun at the center of the space-time image, can easily be seen. Due to the non-linearity of the apparent motion within the hemispherical images, these space-time images do not have the same effect (Figure 5.2). We can use this and make the assumption that a cloud that occluded the sun at a time instant T is the same cloud that is at a location v on the image such that:

$$v = \tau \tan \theta \tag{5.1}$$

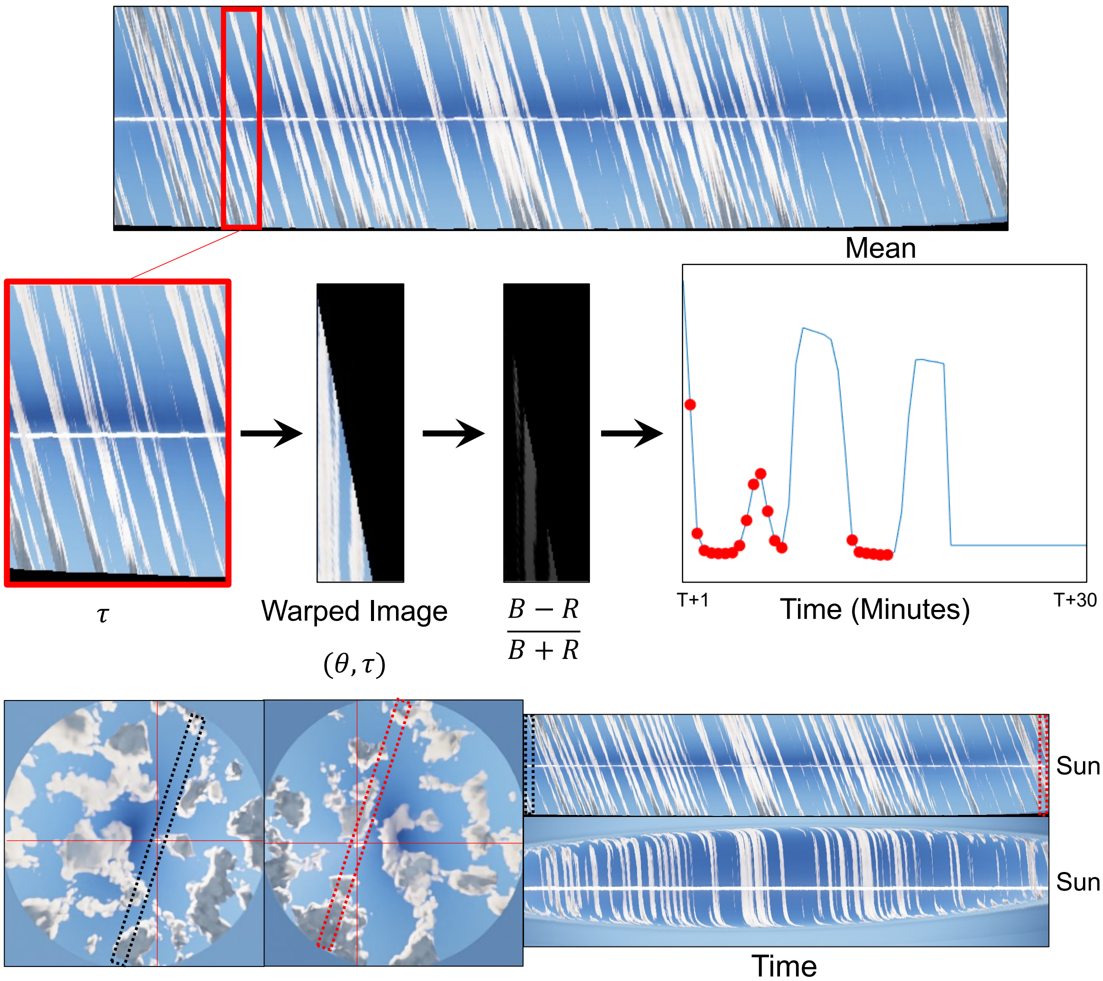


Figure 5.3: **Space-time-slice image creation.** (Top) Space-time image. (Middle-left-to-right) We visualize the non-learning steps for occlusion prediction. We take some window out of the original space-time image, warp it based on the optimal $\hat{\theta}$, convert it to a new color space based on [Li *et al.*, 2011] and plot the mean through time. Red points show ground truth occlusion states. (Bottom-left-to-right) We show how space-time-slices are extracted based on the wind direction. The bottom far right shows the space-time image produced by the hyperboloidal mirror on the top and the hemispherical mirror on the bottom.

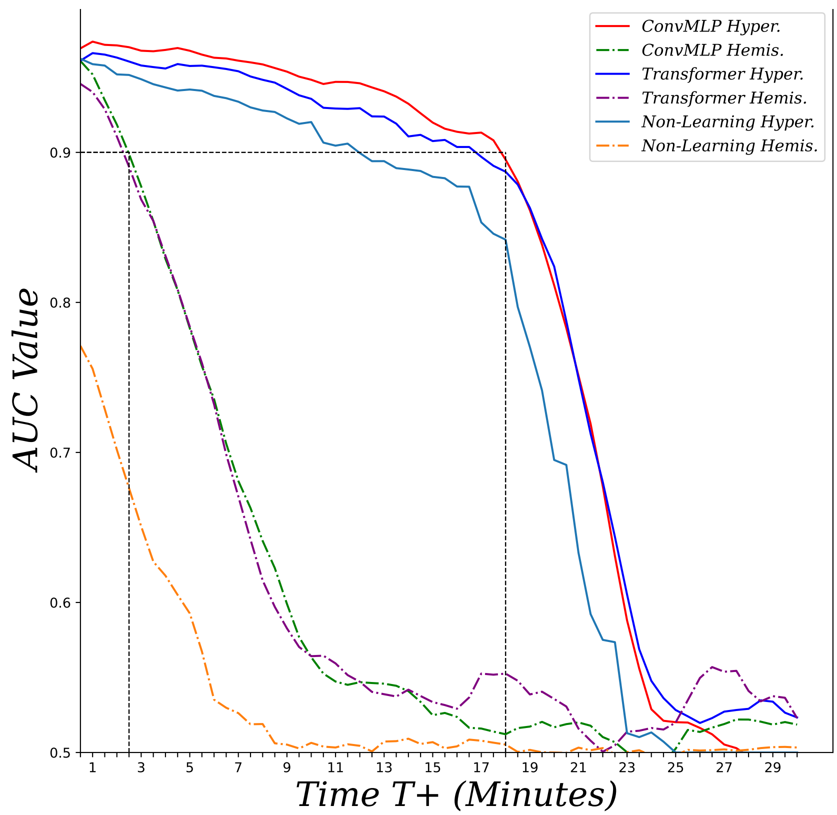


Figure 5.4: **Forecasting results.** We compare learning and non-learning based approaches for binary sun occlusion states on the simulated data. Using 90% as an AUC value for confident predictions, this plot clearly visualizes the stark difference in prediction that we are able to obtain. Looking at the learning-based approaches, for the hemispherical-based mirror, we are limited to a confident forecasting window of around 3 minutes. Compared to hyperboloidal mirror which is able to achieve confident forecasts around 18 minutes into the future.

Where $\tau = (T - t)$ is the time displacement from T along the horizontal x-axis. If we sweep over a range of τ and obtain the slice warped to the current location of the sun at T based on Equation 5.1, this results in a new image which, although seemingly meaningless, actually obtains information about the future sun occlusion states.

We take this warped image, and then convert it to a new color space that can easily contrast cloud versus non-cloud pixels [Li *et al.*, 2011]. The mean of this image along the y-axis is computed and used to identify the binary sun occlusion state. As shown in Figure 5.3, dips in the plotted mean relate to sun occlusion states.

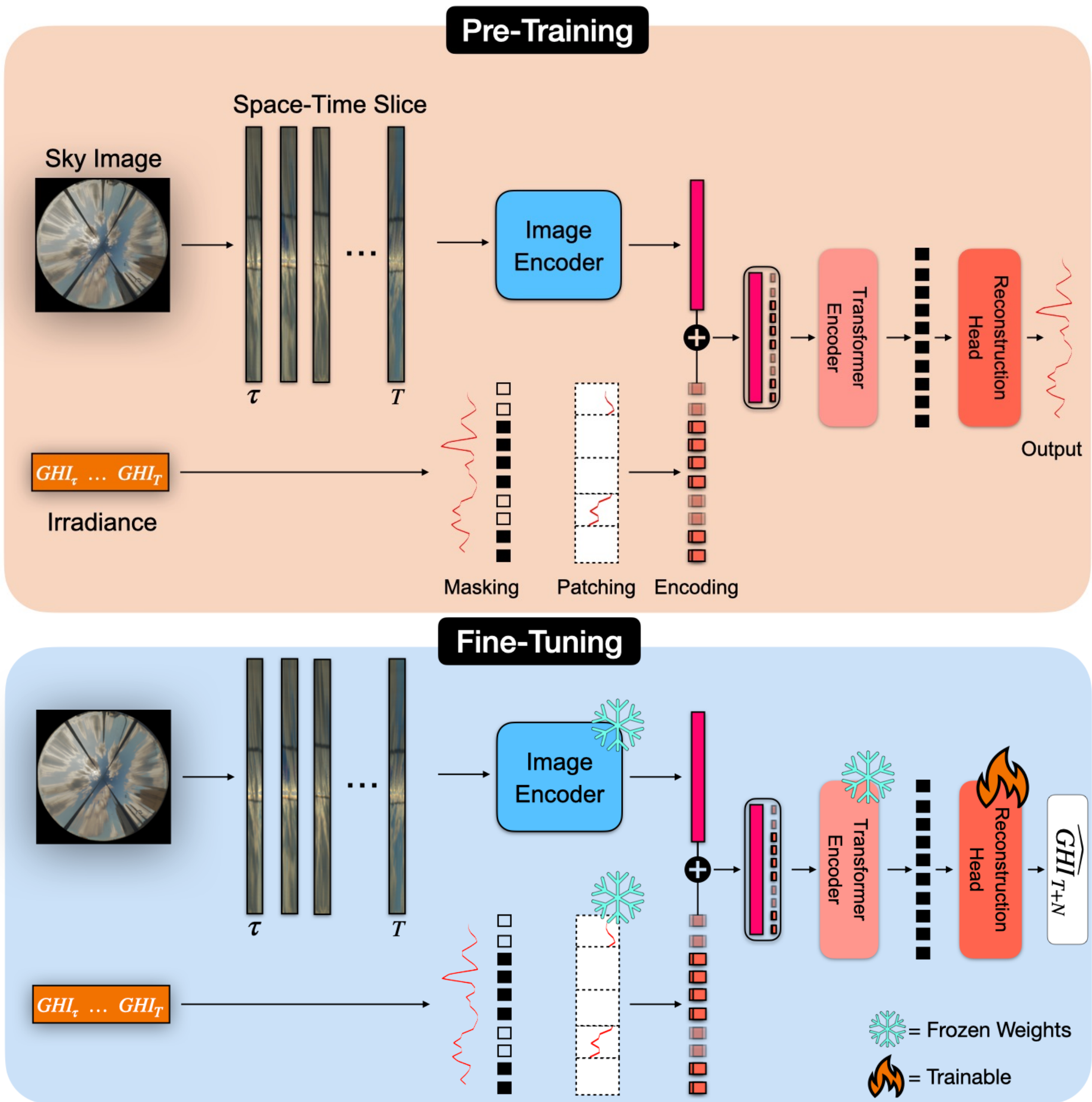


Figure 5.5: **Summary of proposed method.** We present a computational imaging system based on a catadioptric combination of mirrors and cameras. We initially capture a set of sky images and corresponding irradiance values, extract spatio-temporal slices from each image (reference section 5.1) and forecast future irradiance values via a 2-part learning pipeline that pre-trains on reconstruction followed by fine-tuning for irradiance forecasting. The benefit of our hyperboloidal-based mirror which delivers wide-angle imagery with uniform spatial resolution of the sky over its field of view enables more accurate prediction over a longer time horizon than traditional hemispherical imagers.

It should be noted that to find the optimal $\hat{\theta}$, we sweep over a range of empirically chosen $\hat{\theta} = [60°, 85°]$. Instead of computing the mean, we use the standard deviation and attribute the lowest value of the standard deviation to the optimal $\hat{\theta}$ and warp the image by that value using equation 5.1.

To obtain metrics in terms of accuracy of the sun occlusion state, we employ the receiver operating characteristic (ROC) curve to obtain the optimal threshold for deciding the occlusion state. We then look at accuracy through time for future time steps using the area under the curve (AUC) which provides an accumulated measurement of performance across all classification thresholds.

5.2.2 Learning-Based Occlusion Prediction

We believe that tasking a learning-based system to learn the dependencies of the space-time image to predict a future sun occlusion state will yield better results than a non-learning approach.

Neural Occlusion For Simulated Images

Non-learning based prediction of sun occlusion is limited by the dimensionality of the space-time image. As a result, the time prediction to $T + N$ of a sun occlusion is capped at a certain value of N which is based on $\hat{\theta}$. Using a simple CNN-MLP, we are able show that a learning-based method can learn the dependency between spatial cloud locations at τ along with $\hat{\theta}$ to predict the sun occlusions state for a future time instant. Our model is fed the warped image consisting of τ space-time-slices and produces a $(1 \times T + N)$ vector which are the binary occlusion state predictions from $T + 1$ to $T + N$. Our model is trained end-to-end using binary cross entropy as the loss function. We present comparable results to the non-learning approach for both mirrors in Figure 5.4.

Neural Occlusion Prediction For Real Images

For the real images, we utilize a different approach. Instead of directly predicting the binary occlusion state of the sun, we predict the solar irradiance at a future time instant which directly correlates to sun occlusion. Global Horizontal Irradiance (GHI), the total amount of solar irradiance received at a location horizontal to the Earth's surface is measured using a pyranometer in the units of watts per meters-squared $\left(W/m^2\right)$. As shown in Figure 5.10 a decrease in GHI directly correlates to the occlusion of the sun by a cloud and therefore can be used as a prediction method for real images. GHI is predicted as opposed to the direct sun occlusion state due to the fact that we do not have the ground truth occlusion state for the real images. Overall, GHI is a better value to predict than occlusion state due to the fact binary occlusion state has no notion of *how* occluded the sun is by a cloud. GHI provides this quantifiable intensity value.

Our learning pipeline for forecasting GHI is 2-fold and very similar to [Goswami *et al*., 2024] being that we first pre-train our model on masked-input reconstruction followed by fine-tuning for prediction. Our model architecture uses a transformer encoder [Raffel *et al*., 2019] and instead of a transformer-based decoder, we use a simple lightweight reconstruction head for pre-training and a forecasting head for fine-tuning. Both of these heads are small multi-layer perceptrons (MLPs) consisting of linear and dropout layers.

Pre-training. During pre-training, we use information from our space-time slices along with their associated GHI values as input $\{\tau;\ G_\tau \dots G_T\}$. The input space-time slice image is sent to a small image encoder consisting of 5 convolution layers where each layer is followed by a ReLU and 2D Maxpool; except for the last layer. This results in a latent embedding (K) consisting of 2-channels that encapsulates the information from the space-time slice image. Concurrently, for the associated GHI values, we employ masked-input reconstruction where, during training, 25% of the input is masked-out and replaced with a learnable mask embedding. Theoretically, we treat this input GHI as a time series data where information about the current cloud conditions are added via the space-time slice images. The masked GHI is fed into a patch embedding layer similar to [Dosovitskiy *et al*., 2021] and the resulting latent embedding (I) is concatenated with K and passed into the transformer encoder. The encoder passes its learned output to the reconstruction head which reconstructs the original masked input.

Overall, the goal of this pre-training is to allow the model to learn a representation of the original GHI with cloud information present in the space-time slice image. Figure 5.5 presents a visual of the described steps.

Fine-tune forecasting. The goal of fine-tune the model to the task of forecasting is to use the pre-trained learned representation of a space-time slice image and its associated GHI value. During fine-tune training for forecasting, we replace the reconstruction head with a forecasting head. The forecasting head is again a lightweight MLP consisting of a dropout and linear layer. Every other weight parameter of the model is frozen during fine-tuning except for the forecasting head. The model takes the same input of the space-time slice image and its associated GHI values $\{\tau;\ G_\tau \dots G_T\}$. However, instead of reconstructing the original input, the model predicts the GHI at future time instances: $[\widehat{G}_{T+1} \dots \widehat{G}_{T+N}]$. The goal of this is to limit the amount of trainable parameters for the task of forecasting all while using the high-level features and learned weights from the encoder. Both pre-trained and fine-tuned models utilize mean-squared error (MSE) as loss functions.

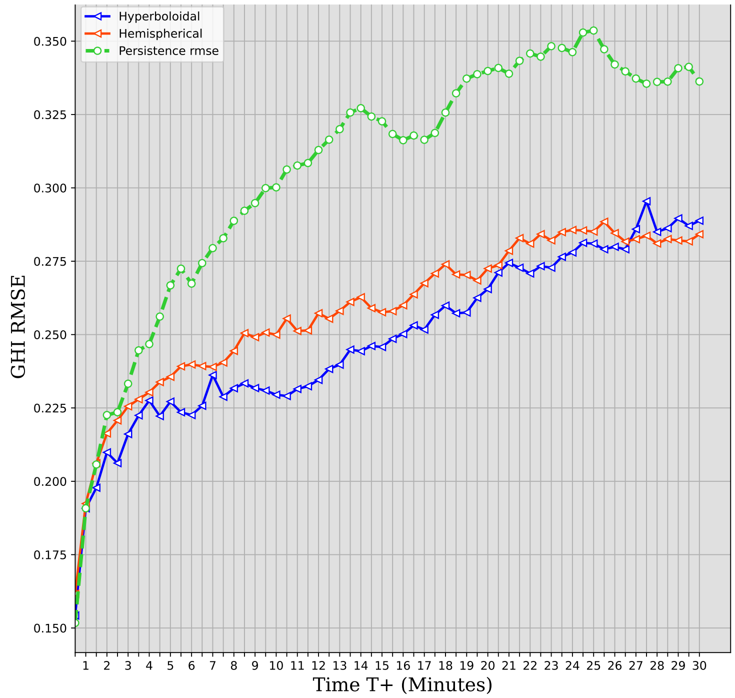


Figure 5.6: **Real forecasting comparison.** This figure shows the outperformance of the hyperboloidal mirror. Our mirror is able to achieve lower forecasting error longer into the future compared to the traditional hemispherical mirror. Both models are trained using the model described in section 5.3.2 which is fed 30 minutes of past data to predict 30 minutes of GHI. As a baseline comparison, we compare both imaging setups to the persistence model which states that the irradiance value will remain unchanged over the forecasting horizon ($\widehat{GHI}_{t+\Delta t} = GHI_t$). Our model still outperforms the persistence model longer into the future.

5.3 Evaluation

In this section, we present results from the above algorithms for both the simulated and real data.

5.3.1 Simulations

For experiments on the non-learning based sun occlusion prediction for the simulated data, we use 100 minutes ($\tau = 200$) of past data to predict 30 minutes ($N = 60$) of sun occlusion state values, in 30 second intervals. Using the algorithms expressed in section 5.2, we achieve promising results.

As seen in Figure 5.4 we able to obtain reliable occlusion predictions up to $\approx$ 18 minutes into the future compared to the hemispherical mirror that is only able to maintain solidified predictions to ≈ 3 minutes. Learning-based methods for occlusion prediction provide the best results, providing greater improvement over the back projected method. For the hyperboloidal mirror we benefit with even greater accuracy, longer through time, all while still substantially outperforming predictions obtained using the hemispherical mirror. For comparison, we experimented using a Transformer-based architecture on the simulated data with the same inputs. Simulated results clearly show the benefit of using a hyperboloidal mirror setup coupled with a learning-based system for long-term prediction of sun occlusion by a cloud. We now present real-world results and show the benefits of using our system.

5.3.2 Real Data

For experiments using real data, we pass 30 minutes ($\tau = 60$) of data to predict 30 minutes ($N = 60$) of GHI, in 30 second intervals. Although as not ideal as in the simulated case, results from the real case still achieve GHI prediction performance over the hemispherical mirror. For our accuracy metric, we use the normalized root mean-squared error (nRMSE):

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N} (\hat{y}_n - y_n)^2}{N}}. \tag{5.2}$$

$$nRMSE = RMSE / \sqrt{\frac{\sum_{n=1}^{N} y_n^2}{N}}, \tag{5.3}$$

where $\hat{y}_n$, y_n is the predicted GHI and true GHI, respectfully. A lower value equates to a more accurate prediction.

We present results of prediction values in Figure 5.6 which shows that we are able to predict GHI longer into the future with lower error compared to the hemispherical mirror. We also present sample GHI predictions in Figure 5.7.

As a baseline comparison, we compare results from our hyperboloidal mirror and hemispherical mirror to the naive persistence model. Persistence says the current value will remain the same for future values, i.e.:

$$\widehat{GHI}_{t+\Delta t} = GHI_t. \tag{5.4}$$

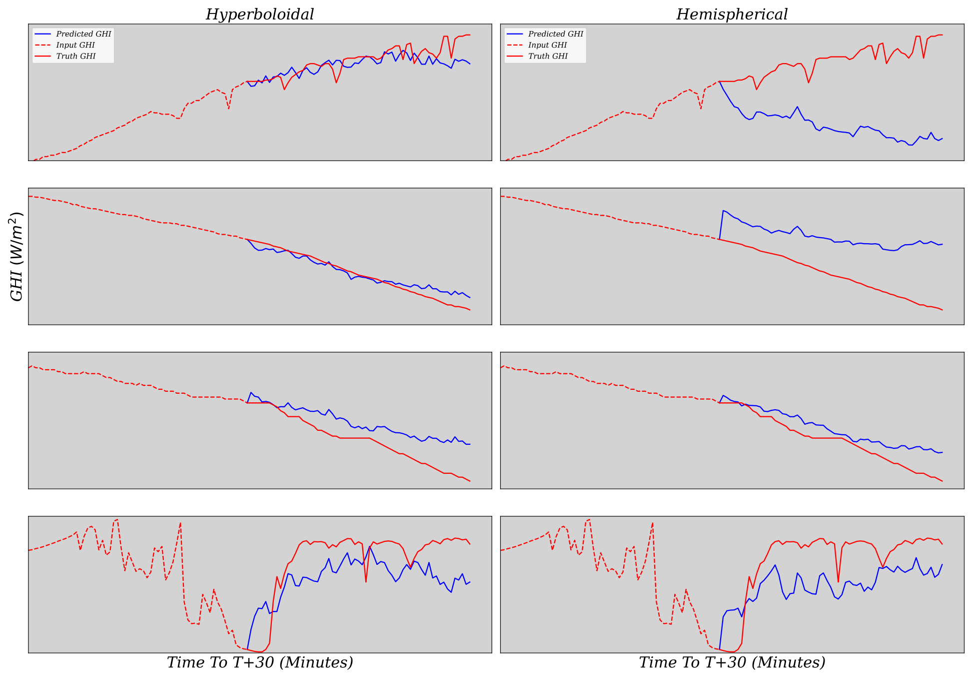


Figure 5.7: **Sample GHI Predictions.** We present sample GHI predictions for the hyperboloidal mirror and the hemispherical mirror. (Left) we input 8 minutes of data into the model to predict 10 minutes of GHI. (Right) We input 30 minutes of data to predict 30 minutes of GHI. For the hyperboloidal mirror, we not only predict the GHI trend close to the ground truth but we are also able to predict the sharp event of when GHI decreases.

5.4 Model Ablation

In Table 5.1, we present a brief ablation study that shows how different model architectures and variations to the input affect prediction results. The Transformer model is the proposed architecture in this Chapter. We wanted to *feed* the model the prediction information and exploit the benefits of our hyperboloidal setup. "Combined" concatenates the space-time-slice image from the hyperboloidal and the hemispherical mirror in the channel dimension. ”Combined” concatenates the space-time-slice image from the hyperboloidal and the hemispherical mirror in the channel dimension. The hyperboloidal imager captures long-term predictions compared to the hemispherical imager. What information can be learned if we combined both images in an attempt to improve prediction results? That is what this input variation attempted to understand. The Seq2Seq model exploits recurrence in the form of gated recurrent units (GRUs) to model the sequential aspect of the space-time-slice images along with the associated GHI. When com-

Model	Data	1 min	5 min	10 min	15 min	20 min	30 min
Persistence	GHI alone	0.191	0.267	0.300	0.323	0.340	0.336
Seq2Seq	GHI alone	**0.169**	0.233	0.265	0.287	0.304	0.307
Seq2Seq	Hyperboloidal + GHI	0.179	0.227	0.248	0.260	0.277	0.277
Seq2Seq	Hemispherical + GHI	0.185	0.227	0.239	0.255	0.271	0.270
Transformer	Hyperboloidal + GHI	0.191	0.227	**0.230**	**0.246**	0.266	0.289
Transformer	Hemispherical + GHI	0.192	0.236	0.250	0.258	0.273	0.284
Transformer	Combined + GHI	0.183	**0.223**	0.239	0.252	**0.259**	**0.268**
Transformer	Hyperboloidal Warped + GHI	0.189	0.233	0.266	0.294	0.303	0.319
Transformer	Hemispherical Warped + GHI	0.190	0.236	0.253	0.294	0.310	0.341

Table 5.1: **Model ablation.** We present comparison results on various models trained and tested on the same datasets. The Transformer model is the proposed architecture in Chapter 5. "Combined" concatenates the space-time-slice image from the hyperboloidal and the hemispherical mirror in the channel dimension. The Seq2Seq model exploits recurrence in the form of gated recurrent units (GRUs) to model the sequential aspect of the space-time-slice images along with the associated GHI.

paring these models using nRMSE as a prediction metric, no architecture clearly stands out. This could be due to each model learning and focusing on different aspects of the input, leading to variations of the predicted outputs. In Figure 5.8, we question if pre-training is necessary. Therefore, we skip pre-training the model on masked-input reconstruction and only train our model end-to-end on the task of forecasting future GHI. The input is the same space-time-slice image along with the associated GHI values with the goal of now predicting future GHI values.

5.5 Discussion

This chapter argues for a novel system that brings core computational imaging techniques to a compelling problem in renewable energy. This chapter provides a pathway to improve the time horizon over which we can reliably forecast solar irradiance; specifically, over conventional wide FOV systems, we can improve predictions from minutes to tens of minutes. We expect such a prediction framework to be of wide interest in the solar photovoltaics community, where resource allocation and energy dispatch is often done in the absence of such predictive analytics. Finally, on a broader scale, we hope the techniques suggested in this

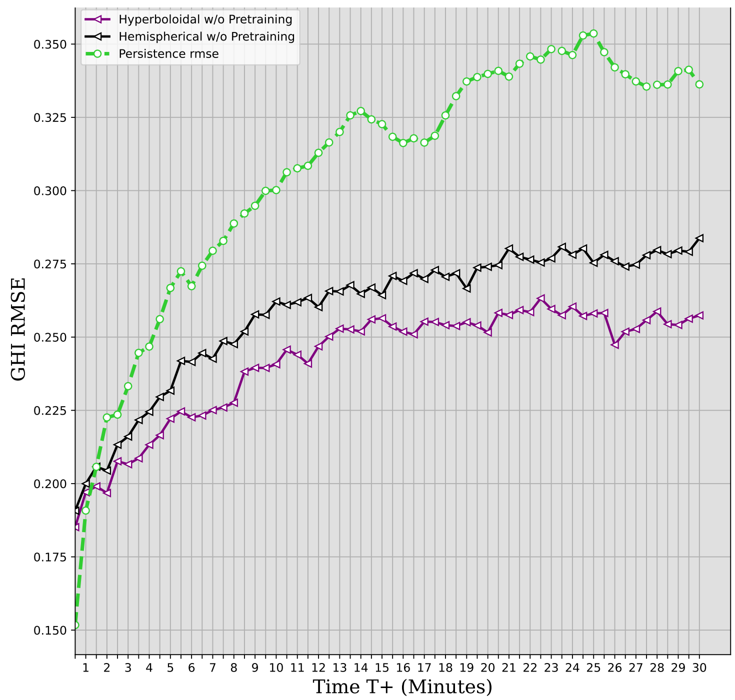


Figure 5.8: **No pre-training forecasting comparison.** This figure presents results from our learning pipeline without pre-training i.e. we only train end-to-end on forecasting future GHI values given a space-time-slice image and associated past GHI values as input. Our hyperboloidal mirror is still able to achieve lower forecasting error longer into the future compared to the traditional hemispherical mirror and the persistence model.

chapter continue to incite the interest in applications that lie at the intersection of imaging and climate change.

Cloud formation and disappearance. One of the factors that we fail to consider in this chapter is that clouds form and disappear based on changes in humidity, temperature, and pressure. This violates the slicing model used in this chapter, in part because clouds can appear in the middle of the field of view, or disappear as it traverses the field. In our testbed, this happens frequently at a particular spot that is over a water body, a few miles from the deployed system. The spatial consistency of the cloud formation suggests that statistical models that have a better understanding of how clouds form and terminate, augmented with other sources of data such as weather, humidity and the geographic layout of the surrounding regions, might have a better chance in handling the effects.

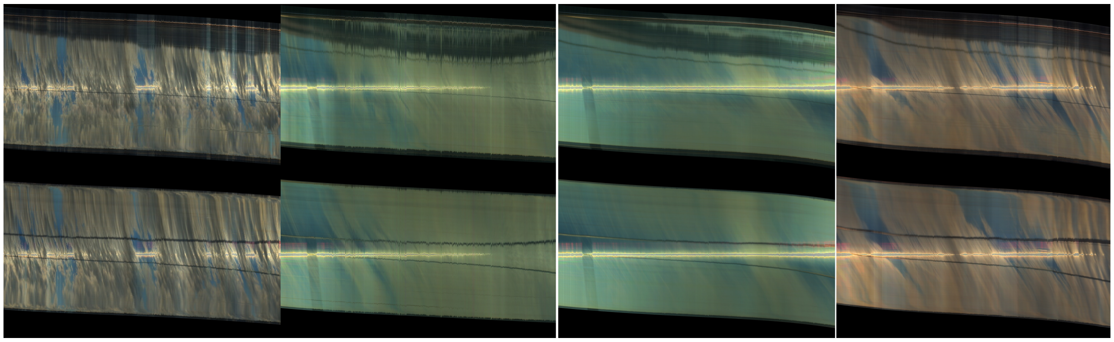


Figure 5.9: **Real space-time-slice images.** We show resulting space-time-slice images from captured from our real setup. (Top) a result of selecting $\theta = 65°$ while the bottom is $\theta = 85°$. To forecast GHI, we select θ $[65°, 85°]$ in $+5°$ increments.

Self-occlusion by clouds. Another factor that we fail to consider is that clouds have vertical extent, and hence a cloud closer to the camera may block one that is further away. This shows itself in the form of radial streaks in Figures 4.7 and 4.10. This is a hard problem to resolve in the absence of additional view points. It is likely that a multi-camera version of our system with a baseline in kilometers will be able to reason such occlusions and handle them effectively.

Incorporating other data sources. The techniques proposed in this chapter will also benefit from other richer sources of data such as satellite imagery, weather prediction, and humidity measurements. Such sources of data are often publicly available; however, each of them have unique features that need to be accounted for. For example, satellite data that provides very large spatial extent, has very poor temporal resolution, often in minutes, if not hours. Further, the ground spatial resolution of such data is in meters, which may not be sufficient for the kind of prediction we envision. Wind velocity, which is often available from weather data, is something we can benefit from. However such measurements are often made at ground level, and at very sparse locations, which limits their utility, since atmospheric wind velocities differ from ground measurements. Yet the role that temperature, humidity, and more broadly the weather play in forecasting cannot be denied. Translating such models to near-future time horizons and higher precision that is demanded for solar forecasting is an interesting approach for subsequent research.

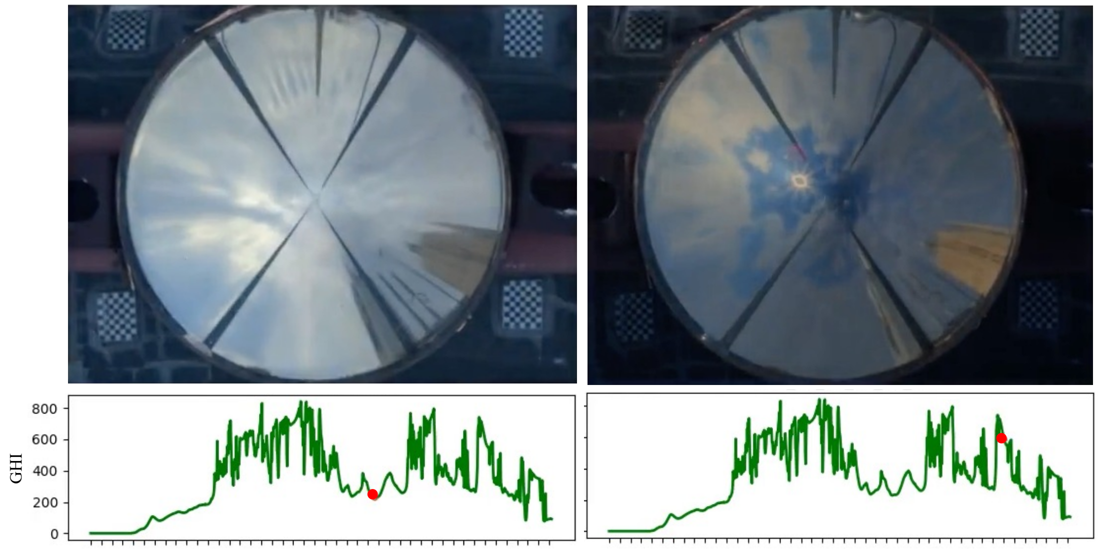


Figure 5.10: **GH with images.** We show sample GHI values based on cloud conditions

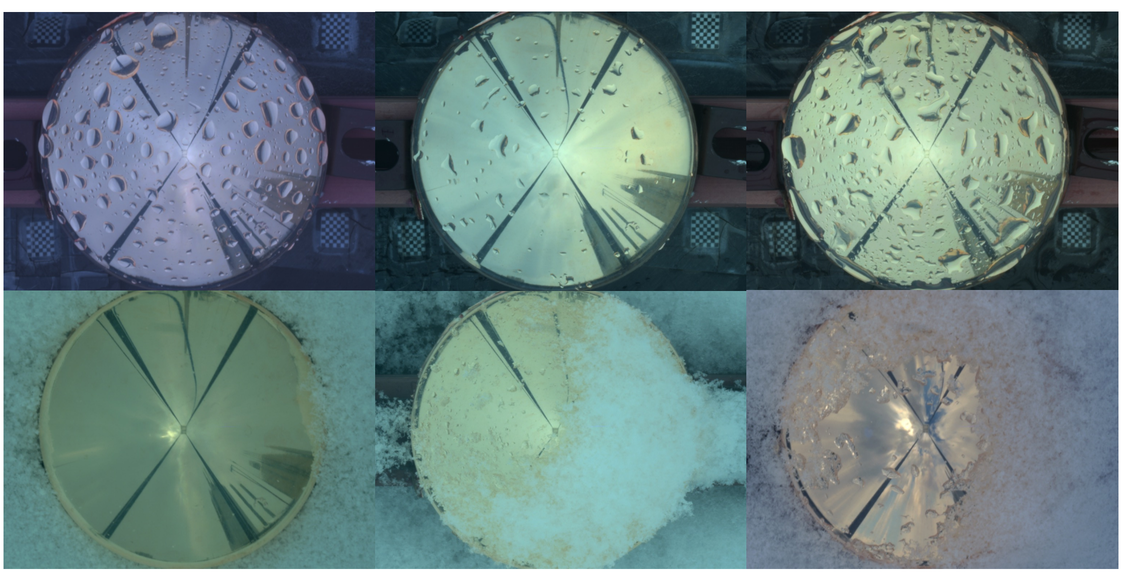


Figure 5.11: **Adverse weather conditions.** We show examples of adverse weather conditions captured from our real setup of rain (top) and snow (bottom).

6 Conclusion

6.1 Thesis Contribution

Solar irradiance forecasting is they key to increasing the penetration of solar energy into the electricity grid at all scales. By framing this forecasting challenge as an imaging problem, we can directly observe the primary factor limiting solar irradiance—cloud occlusion of the sun—and make accurate forecasts of future conditions. The information within these images provide us with rich information necessary to understand how the influence of cloud cover affects the amount of irradiance received at the ground. By leveraging learning-based algorithms and novel imaging systems we can increase this prediction even further

Overall, we have contributed to solar irradiance forecasting by applying a computational imaging approach. Through the work of this thesis:

- We achieved longer forecasting times when modeling cloud evolution in sky images by initially spatially-warping sky images during training; facilitating to longer-forecasting of cloud evolution. We showed that warping these images countered the adverse affects of resolution loss near the horizon in traditional sky imagers. We also learned a deep neural network that took multiple sky images in as input and predicted the next frame accurately. With this, we were able to demonstrate the ability to accurately forecast sky image frames with higher resolution metrics than previous cloud forecasting methods.

- We presented a different approach of tackling the problem of non-linear apparent motion present in traditional hemispherical sky imagers. By *optically redistributing* the pixels in a wide FOV catadioptric imaging system, we were able to image the sky such that we maintained uniform spatial resolution over the entire FOV of the device. As a result, the motion of the clouds are equally perceptible, be it at the zenith or the horizon.

- Using our deployed catadioptric imaging system, we were able to capture months of data which enabled us to create a large testbed for inference. With our dataset of real and simulated sky images captured from our hyperboloidal setup, we show that our ability to estimate cloud trajectory is improved over traditional methods even when a cloud is farther away. Via learning and non-learning algorithms, we show that our system provides precise prediction of sun occlusion and solar irradiance over a time horizon of tens of minutes (30 minutes for the simulated data, and 10-20 mins for the real system).

6.2 Future Work

The work of this thesis provides a fundamental advance to solar irradiance forecasting using core computational imaging. Further work in this space can build upon the presented ideas in the following areas.

6.2.1 Improved Mirror Fabrication

It is necessary to develop a catadioptric imaging system in which the mirror that is used is highly reflective and free from imperfections on the surface. Within this thesis, I attempted to achieve these goals using 2 methods: material subtraction and a chemical addition process. Material subtraction attempts to make the diffuse metallic surface as reflective as possible by polishing until the desired reflection is obtained. This can be done through sanding by hand or with a machine such as a lathe or grinder. A major limitation when using these methods is that they leave scratches and imperfections on the surface of the mirror. Utilizing this in an imaging setup in which the sun is very bright creates strong artifacts in the resulting image which limits inference tasks. Also, removing too much material from the surface is a problem which could overall alter the shape of the mirror; removing the tight tolerances required during fabrication. Therefore, more advanced material subtraction methods are necessary. One promising process that can meet the requirements of obtaining a reflective surface without imperfections is diamond turning. Diamond turning is a machining technique that using a single crystal of diamond to produce mirror surfaces on metals of complex geometries.

Another option is applying complex chemical deposition to the surface of the material which will remove the need for any material subtraction. There are various processes depending on the surface type of the mirror. Future work can streamline the mirror fabrication process and pave the way to the deployment of multiple distributed catadioptric sky imaging systems.

6.2.2 Emphasis On Horizon Resolution

In our catadioptric mirror design that utilized a hyperboloidal shaped mirror, we chose a mirror design that prioritized resolution everywhere within the FOV. This ensured that we did not lose prediction over small time forecasts. The question arises – *why not place a larger emphasis on the periphery?* If the clouds that are further out at the horizon better determine what its evolution will lead to right above you, we should put a larger emphasis on that location. Future work could develop imaging systems that focus on clouds at the horizon in an effort to increase forecasting times even greater.

6.2.3 Expanded Sources of Data

Forecasting solar irradiance can further be improved by incorporating a larger gamut of data values instead of solely using sky images and previous GHI values. The amount of irradiance that reaches the ground is influenced by complex global interactions that include water vapor, pollution, elevation, wind, temperature, air pressure, etc. Also, incorporating other streams of visual data such as satellite imagery, would allow knowledge over a larger spatial context of cloud information. Coupled with ground-based imagery, for a fine spatial context, the addition could lead to coarse-to-fine representation of cloud patterns over a region. For example, The National Weather Service (NWS) provides publicly accessible real-time satellite feeds of visible and radar imagery useful for this task. Infrared, polarized, and hyper-spectral imagery can also be used to image the sky in a larger wavelength spectrum. All-in-all incorporating a larger stream of data into the forecasting pipeline, especially for learning-based models, will lead to more accurate predictions across a longer time range.

6.2.4 Distributed Imaging Systems

By itself, a single viewpoint does not provide sufficient information to reason about the 3D structure of the clouds that is required to forecast over a larger area. 2D images treat clouds as planar objects and disregard their 3D and volumetric scattering-tomographic nature. Having multiple sky imaging systems deployed with a ground baseline of many miles can provide a multi-view imaging system for accurate 3D reconstructions of clouds and sky model. With this, we could image over a larger spatial context while maintaining localized information, model the clouds to measure the appearance of absorption and scattering, and reconstruct the cloud as a four dimensional field using implicit neural networks.

Another interesting way to scale the sensing platform to large spatial regions is to exploit publicly-available video streams. Most cities today have a rich set of real-time web feeds for monitoring traffic and city streets. Many of them capture a small portion of the sky with nearly half of these videos having viable information about the sky. We can expand the imaging baseline to many tens of miles without having to deploy new cameras. Even cameras that have no sky information have value in the form of providing a measurement of solar irradiance simply by their relative brightness, which can be used for evaluation. These camera streams also provide a scalable solution to other cities that might not have dedicated sky imaging systems deployed. Distributed sky imaging systems are a necessary future step to take; allowing for multiple views and enhanced inference of the atmosphere.

6.2.5 Solar Irradiance Forecasting For Downstream Tasks

The impact of solar irradiance forecasting on downstream tasks is an interesting future direction of research to consider. For example, one could study the role that forecasting plays in energy management for buildings equipped with PV systems and storage. Forecasting irradiance also has the ability to significantly enhance home energy management systems to predictively control building loads with respect to household-level objectives (e.g., optimally scheduling appliances and battery charging to maximize the use of solar generation) as well as grid-level objectives (e.g., regulating voltage fluctuations at the distribution system or minimizing transformer loading). Incorporating irradiance forecasting is essential for a closed looped system of energy management incorporating solar.

6.3 Conclusion

We aim that this thesis brings light to the possibilities that imaging and learning can bring to the world of solar irradiance forecasting. 2D imaging is just a single modality to improve forecasting. 3D reconstruction to understand their radiative transfer and distributed imaging systems to understand their trajectory over large scales can open up to the full understanding of clouds and how they traverse our atmosphere. Overall, the world needs more renewables for energy generation and, by bridging the gap between the unknown future of cloud states and solar generation, full and efficient integration is possible.

Bibliography

[n.d.]. https://cloudatlas.wmo.int/fr/descriptions-of-clouds.html.

[n.d.]. https://angelgilding.com/media/wysiwyg/PDF/How_To_Silver_3-D_Glass_Objects.pdf

2017. Characterizing and analyzing ramping events in wind power, solar power, load, and netload. *Renewable Energy* 111 (2017), 227–244. https://doi.org/10.1016/j.renene.2017.04.005

Yiyang Ai, Yonggang Peng, and Wei Wei. 2017a. A model of very short-term solar irradiance forecasting based on low-cost sky images. *AIP Conference Proceedings* 1839, 1 (2017), 020022.

Yiyang Ai, Yonggang Peng, and Wei Wei. 2017b. A model of very short-term solar irradiance forecasting based on low-cost sky images. *AIP Conference Proceedings* 1839, 020022. https://doi.org/10.1063/1.4982387

Amit Aides, Aviad Levis, Vadim Holodovsky, Yoav Y. Schechner, Dietrich Althausen, and Adi Vainiger. 2020. Distributed Sky Imaging Radiometry and Tomography. In *2020 IEEE International Conference on Computational Photography (ICCP)*. 1–12. https://doi.org/10.1109/ICCP48838.2020.9105241

Anas Al-lahham, Obaidah Theeb, Khaled Elalem, Tariq A. Alshawi, and Saleh A. Alshebeili. 2020. Sky Imager-Based Forecast of Solar Irradiance Using Machine Learning. *Electronics* 9, 10 (2020).

L. Alados-Arboledas, J. Vida, and F. J. Olmo. 1995. The estimation of thermal atmospheric radiation under cloudy conditions. *International Journal of Climatology* 15, 1 (1995), 107–116.

Ameera M Almarzooqi, Maher Maalouf, Tarek H M El-Fouly, Vasileios E Katzourakis, Mohamed S El Moursi, and Constantinos V Chrysikopoulos. 2024. A hybrid machine-learning model for solar irradiance forecasting. *Clean Energy* 8, 1 (01 2024), 100–110.

Ahmad Alzahrani, Pourya Shamsi, Cihan Dagli, and Mehdi Ferdowsi. 2017. Solar Irradiance Forecasting Using Deep Neural Networks. *Procedia Computer Science* 114 (2017), 304–313.

Afshin Andreas and Thomas Stoffel. 1981. *NREL Solar Radiation Research Laboratory (SRRL): Baseline Measurement System (BMS); Golden, Colorado (Data)*. NREL Report DA-5500-56488. National Renewable Energy Laboratory (NREL). https://doi.org/10.5439/1052221

George Andrianakos, Dimitrios Tsourounis, Spiros Oikonomou, Dimitris Kastaniotis, George Economou, and Andreas Kazantzidis. 2019a. Sky Image forecasting with Generative Adversarial Networks for cloud coverage prediction. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*. 1–7.

George Andrianakos, Dimitrios Tsourounis, Spiros Oikonomou, Dimitris Kastaniotis, George Economou, and Andreas Kazantzidis. 2019b. Sky Image forecasting with Generative Adversarial Networks for cloud coverage prediction. In *International Conference on Information, Intelligence, Systems and Applications (IISA)*.

S. Baker and S.K. Nayar. 1998. A theory of catadioptric image formation. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*. 35–42.

Asmae Berrada and Khalid Loudiyi. 2016. Operation, sizing, and economic evaluation of storage for solar and wind power plants. *Renewable and Sustainable Energy Reviews* 59 (2016), 1117–1129.

Matthew A. Brown and David G. Lowe. 2007. Automatic Panoramic Image Stitching using Invariant Features. *International Journal of Computer Vision* 74 (2007), 59–73.

Alberto Cazorla, F. Olmo, and Lucas Arboledas. 2008. Using a Sky Imager for aerosol characterization. *Atmospheric Environment* 42 (04 2008), 2739–2745.

Ming-Ching Chang, Yi Yao, Guan Li, Yan Tong, and Peter Tu. 2017. Cloud tracking for solar irradiance prediction. In *2017 IEEE International Conference on Image Processing (ICIP)*. 4387–4391.

Chi Wai Chow, Serge Belongie, and Jan Kleissl. 2015. Cloud motion and stability estimation for intra-hour solar forecasting. *Solar Energy* 115 (2015), 645–655.

Chi Wai Chow, Bryan Urquhart, Matthew Lave, Anthony Dominguez, Jan Kleissl, Janet Shields, and Byron Washom. 2011. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy* 85, 11 (2011), 2881–2893.

Yinghao Chu, Hugo T.C. Pedro, and Carlos F.M. Coimbra. 2013. Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning. *Solar Energy* 98 (2013), 592–603.

Cristian Crisosto, Eduardo W. Luiz, and Gunther Seckmeyer. 2021. Convolutional Neural Network for High-Resolution Cloud Motion Prediction from Hemispheric Sky Images. *Energies* 14, 3 (2021).

Michelle Davis, Sylvia Leyva Martinez, Zoë Gaston, Sagar Chopra, Caitlin Connelly, Kaitlin Fung, Max Issokson, Elissa Pierce, Amanda Colombo, Colin Silver, Tyler Thompson, Forrest Levy, and Justin Baca. 2024. U.S. Solar Market Insight: Q2 2024 Executive Summary. `https://www.woodmac.com/industry/power-and-renewables/us-solar-market-insight/`

A. Demir, L. Gutierrez, A. Namin, and S. Bayne. 2022. Solar Irradiance Prediction Using Transformer-based Machine Learning Models. In *2022 IEEE International Conference on Big Data (Big Data)*. IEEE Computer Society, 2833–2840.

Soumyabrata Dev, Atul Nautiyal, Yee Hui Lee, and Stefan Winkler. 2019. CloudSegNet: A Deep Network for Nychthemeron Cloud Image Segmentation. *IEEE Geoscience and Remote Sensing Letters* 16 (2019), 1814–1818.

S. Dev, F. M. Savoy, Y. H. Lee, and S. Winkler. 2016. Estimation of solar irradiance using ground-based whole sky imagers. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*.

Maimouna Diagne, Mathieu David, Philippe Lauret, John Boland, and Nicolas Schmutz. 2013. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews* 27 (2013), 65–76.

David J. Diner, Stacey W. Boland, Michael Brauer, Carol Bruegge, Kevin A. Burke, Russell Chipman, Larry Di Girolamo, Michael J. Garay, Sina Hasheminassab, Edward Hyer, Michael Jerrett, Veljko Jovanovic, Olga V. Kalashnikova, Yang Liu, Alexei I. Lyapustin, Randall V. Martin, Abigail Nastan, Bart D. Ostro, Beate Ritz, Joel Schwartz, Jun Wang, and Feng Xu. 2018. Advances in multiangle satellite remote sensing of speciated airborne particulate matter and association with adverse health effects: from MISR to MAIA. *Journal of Applied Remote Sensing* 12, 4 (2018), 042603. `https://doi.org/10.1117/1.JRS.12.042603`

D. J. Diner, F. Xu, M. J. Garay, J. V. Martonchik, B. E. Rheingans, S. Geier, A. Davis, B. R. Hancock, V. M. Jovanovic, M. A. Bull, K. Capraro, R. A. Chipman, and S. C. McClain. 2013. The Airborne Multiangle SpectroPolarimetric Imager (AirMSPI): a new tool for aerosol and cloud remote sensing.

Atmospheric Measurement Techniques 6, 8 (2013), 2007–2025. https://doi.org/10.5194/amt-6-2007-2013

Ben Dixon, María Pérez-Ortiz, and Jacob Bieker. 2022. Comparing the carbon costs and benefits of low-resource solar nowcasting. arXiv:2210.04554

Yanmei Dong, Mingtao Pei, Yuwei Wu, and Yunde Jia. 2022. Stitching images from a conventional camera and a fisheye camera based on nonrigid warping. *Multimedia Tools and Applications* 81 (05 2022). https://doi.org/10.1007/s11042-022-12236-0

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. 2021. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ICLR* (2021).

Sara Eftekharnejad, Vijay Vittal, Gerald Thomas Heydt, Brian Keel, and Jeffrey Loehr. 2013. Impact of increased penetration of photovoltaic generation on power systems. *IEEE Transactions on Power Systems* 28, 2 (2013), 893–901.

Ciarán Eising, Martin Glavin, Edward Jones, and Patrick Denny. 2008. Review of geometric distortion compensation in fish-eye cameras. *IET Conference Publications*, 162 – 167. https://doi.org/10.1049/cp:20080656

Z. El Jaouhari, Y. Zaz, and L. Masmoudi. 2015. Cloud tracking from whole-sky ground-based images. In *International Renewable and Sustainable Energy Conference (IRSEC)*.

Cong Feng, Jie Zhang, Wenqi Zhang, and Bri-Mathias Hodge. 2022. Convolutional neural networks for intra-hour solar forecasting based on sky image sequences. *Applied Energy* 310 (2022), 118438.

Huiyu Gao and Miaomiao Liu. 2022. Short-term Solar Irradiance Prediction from Sky Images with a Clear Sky Model. In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 3074–3082.

Shibani Ghosh, Saifur Rahman, and Manisa Pipattanasomporn. 2017. Distribution Voltage Regulation Through Active Power Curtailment With PV Inverters and Solar Generation Forecasts. *IEEE Transactions on Sustainable Energy* 8, 1 (2017), 13–22.

Tim Godfrey, Sara Mullen, David W Griffith, Nada Golmie, Roger C Dugan, and Craig Rodine. 2010. Modeling smart grid applications with co-simulation. In *2010 first IEEE International conference on smart grid communications*. IEEE, 291–296.

Mononito Goswami, Konrad Szafer, Arjun Choudhry, Yifu Cai, Shuo Li, and Artur Dubrawski. 2024. MOMENT: A Family of Open Time-series Foundation Models. In *International Conference on Machine Learning*.

Gautam Gowrisankaran, Stanley S. Reynolds, and Mario Samano. 2016. Intermittency and the Value of Renewable Energy. *Journal of Political Economy* 124, 4 (2016), 1187–1234.

Vincent Le Guen and Nicolas Thome. 2020. Disentangling Physical Dynamics From Unknown Factors for Unsupervised Video Prediction. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*. IEEE, 11471–11481.

Vadim Holodovsky, Yoav Y. Schechner, Anat Levin, Aviad Levis, and Amit Aides. 2016. In-situ multi-view multi-scattering stochastic tomography. In *2016 IEEE International Conference on Computational Photography (ICCP)*. 1–12. `https://doi.org/10.1109/ICCPHOT.2016.7492869`

Hao Huang, Jin Xu, Zhenzhou Peng, Shinjae Yoo, Dantong Yu, Dong Huang, and Hong Qin. 2013. Cloud motion estimation for short term solar irradiation prediction. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. 696–701.

Tak-Wai Hui, Xiaoou Tang, and Chen Change Loy. 2018. LiteFlowNet: A Lightweight Convolutional Neural Network for Optical Flow Estimation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 8981–8989.

Semich Impram, Secil Varbak Nese, and Bülent Oral. 2020. Challenges of renewable energy penetration on power system flexibility: A survey. *Energy Strategy Reviews* 31 (2020), 100539.

David Infield and Leon Freris. 2020. *Renewable energy in power systems*. John Wiley & Sons.

C. Ishii, Y. Sudo, and H. Hashimoto. 2003. An image conversion algorithm from fish eye image to perspective image for human eyes. In *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003)*, Vol. 2. 1009–1014 vol.2.

Nur Liyana Mohd Jailani, Jeeva Kumaran Dhanasegaran, Gamal Alkawsi, Ammar Ahmed Alkahtani, Chen Chai Phing, Yahia Baashar, Luiz Fernando Capretz, Ali Q. Al-Shetwi, and Sieh Kiong Tiong.

2023. Investigating the Power of LSTM-Based Models in Solar Energy Forecasting. *Processes* 11, 5 (2023).

Vijai Thottathil Jayadevan, Jeffrey J. Rodriguez, Vincent P.A. Lonij, and Alexander D. Cronin. 2012. Forecasting solar power intermittency using ground-based cloud imaging. In *World Renewable Energy Forum (WREF)*.

Huaiguang Jiang, Yi Gu, Yu Xie, Rui Yang, and Yingchen Zhang. 2020. Solar Irradiance Capturing in Cloudy Sky Days–A Convolutional Neural Network Based Image Regression Approach. *IEEE Access* 8 (2020), 22235–22248.

Leron Julian and Aswin C. Sankaranarayanan. 2021. Precise Forecasting of Sky Images Using Spatial Warping. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*. 1136–1144.

Iyswarya Annapoorani K, Rajaguru V, Vedanjali N, and Pappula Rajasri. 2023. Solar forecasting for a PV-battery powered DC system. *Heliyon* 9, 10 (2023), e20667.

T. Kato and T. Nakagawa. 2020. Short-Term Sky Image Prediction using LSTMs from the Past Frames of Sky Images and Synthesized Images by Optical Flow Extrapolation1. In *IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.).

J. Kleissl, B. Urquhart, M. Ghonima, E. Dahlin, A. Nguyen, B. Kurtz, C. W. Chow, and F. A. Mejia. [n.d.]. University of California, San Diego (UCSD) Sky Imager Cloud Position Study Field Campaign Report. ([n. d.]).

Varun Kumar, AS Pandey, and SK Sinha. 2016. Grid integration and power quality issues of wind and solar energy system: A review. In *2016 International conference on emerging trends in electrical electronics & sustainable energy systems (ICETEESES)*. IEEE, 71–80.

V. Le Guen and N. Thome. 2020. A Deep Physical Model for Solar Irradiance Forecasting with Fisheye Images. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*.

Qingyong Li, Weitao Lu, and Jun Yang. 2011. A Hybrid Thresholding Algorithm for Cloud Detection on Ground-Based Color Images. *Journal of Atmospheric and Oceanic Technology* 28, 10 (Oct. 2011), 1286–1296.

Qingyong Li, Weitao Lyu, and Jun Yang. 2011. A Hybrid Thresholding Algorithm for Cloud Detection on Ground-Based Color Images. *Journal of Atmospheric and Oceanic Technology* 28 (10 2011), 1286–1296.

Ce Liu. 2009. *Beyond pixels: exploring new representations and applications for motion analysis*. Ph.D. Dissertation. Massachusetts Institute of Technology.

Jingxuan Liu, Haixiang Zang, Lilin Cheng, Tao Ding, Zhinong Wei, and Guoqiang Sun. 2023. A Transformer-based multimodal-learning framework using sky images for ultra-short-term solar irradiance forecasting. *Applied Energy* 342 (2023), 121160.

Shuang Liu, Linbo Zhang, Zhong Zhang, Chunheng Wang, and Baihua Xiao. 2015. Automatic Cloud Detection for All-Sky Images Using Superpixel Segmentation. *IEEE Geoscience and Remote Sensing Letters* 12, 2 (2015), 354–358.

W. Liu, D. Lian W. Luo, and S. Gao. 2018. Future Frame Prediction for Anomaly Detection – A New Baseline. In *2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

J.V. Martonchik, D.J. Diner, R.A. Kahn, T.P. Ackerman, M.M. Verstraete, B. Pinty, and H.R. Gordon. 1998. Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. *IEEE Transactions on Geoscience and Remote Sensing* 36, 4 (1998), 1212–1227. https://doi.org/10.1109/36.701027

P. J. Mason. 1994. Large-eddy simulation: A critical review of the technique. *Quarterly Journal of the Royal Meteorological Society* 120, 515 (1994), 1–26. https://doi.org/10.1002/qj.49712051503

Felipe A Mejia, Ben Kurtz, Aviad Levis, Íñigo de la Parra, and Jan Kleissl. 2018. Cloud tomography applied to sky images: A virtual testbed. *Solar Energy* 176 (2018), 287–300. https://doi.org/10.1016/j.solener.2018.10.023

Thomas M. Mercier, Tasmiat Rahman, and Amin Sabet. 2023. Solar Irradiance Anticipative Transformer. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2023), 2065–2074.

Nilesh Modi, Marta Val Escudero, Ken Aramaki, Xiaoyao Zhou, and Pauli Partinen. 2024. High Inverter-Based Resource Integration: The Experience of Five System Operators. *IEEE Power and Energy Magazine* 22, 2 (2024), 78–88.

Erman Nghonda Tchinda, Maximillian Kealoha Panoff, Danielle Tchuinkou Kwadjo, and Christophe Bobda. 2023. Semi-Supervised Image Stitching from Unstructured Camera Arrays. *Sensors* 23, 23 (2023). https://doi.org/10.3390/s23239481

Alessandro Niccolai and Alfredo Nespoli. 2020. Sun Position Identification in Sky Images for Nowcasting Application. *Forecasting* 2, 4 (2020), 488–504. https://www.mdpi.com/2571-9394/2/4/26

Yuhao Nie, Eric Zelikman, Andea Scott, Quentin Paletta, and Adam Brandt. 2023. SkyGPT: Probabilistic Short-term Solar Forecasting Using Synthetic Sky Videos from Physics-constrained VideoGPT. arXiv:cs.CV/2306.11682

Yuhao Nie, Eric Zelikman, Andea Scott, Quentin Paletta, and Adam Brandt. 2024. SkyGPT: Probabilistic ultra-short-term solar forecasting using synthetic sky images from physics-constrained VideoGPT. *Advances in Applied Energy* 14 (2024), 100172.

Bijan Nouri, Stefan Wilbert, Niklas Blum, Yann Fabel, Elke Lorenz, Annette Hammer, Thomas Schmidt, Luis F. Zarzalejo, and Robert Pitz-Paal. 2023. Probabilistic solar nowcasting based on all-sky imagers. *Solar Energy* 253 (2023), 285–307.

Quentin Paletta, Anthony Hu, Guillaume Arbod, and Joan Lasenby. 2022. ECLIPSE: Envisioning CLoud Induced Perturbations in Solar Energy. *Applied Energy* 326 (2022), 119924.

Quentin Paletta and J Lasenby. 2020a. Convolutional Neural Networks applied to sky images for short-term solar irradiance forecasting.

Q. Paletta and J. Lasenby. 2020b. Convolutional Neural Networks Applied to Sky Images for Short-Term Solar Irradiance Forecasting. *EU PVSEC* (2020), 1834 – 1837.

Quentin Paletta and Joan Lasenby. 2020c. A Temporally Consistent Image-based Sun Tracking Algorithm for Solar Energy Forecasting Applications. *ArXiv* abs/2012.01059 (2020).

Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in PyTorch. (2017).

Jiri Pospichal, Martin Kubovcik, and Iveta Dirgova Luptakova. 2022. Solar Irradiance Forecasting with Transformer Model. *Applied Sciences* 12, 17 (2022).

Colin Raffel, Noam M. Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2019. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *J. Mach. Learn. Res.* 21 (2019), 140:1–140:67. `https://api.semanticscholar.org/CorpusID:204838007`

Rial A. Rajagukguk, Raihan Kamil, and Hyun-Jin Lee. 2021. A Deep Learning Model to Forecast Solar Irradiance Using a Sky Camera. *Applied Sciences* 11, 11 (2021).

Walter Richardson, Hariharan Krishnaswami, Rolando Vega, and Michael Cervantes. 2017. A Low Cost, Edge Computing, All-Sky Imager for Cloud Tracking and Intra-Hour Irradiance Forecasting. *Sustainability* 9, 4 (2017). `https://doi.org/10.3390/su9040482`

O. Ronneberger, P.Fischer, and T. Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 234–241.

D. Cornforth T. Moore S. Percy J. Ward A. Berry S. Sayeef, S. Heslop and D. Rowe. 2012. Solar intermittency: Australia's clean energy challenge. characterizing the effect of high penetration solar intermittency on Australian electricity networks. *CSIRO* (2012), 195.

Mojtaba Saleh, Lindsay Meek, Mohammad A. S. Masoum, and Masoud Abshar. 2018. Battery-Less Short-Term Smoothing of Photovoltaic Generation Using Sky Camera. *IEEE Transactions on Industrial Informatics* 14, 2 (2018), 403–414.

Mahdi Sharifzadeh, Alexandra Sikinioti-Lock, and Nilay Shah. 2019. Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian Process Regression. *Renewable and Sustainable Energy Reviews* 108 (2019), 513–538. `https://doi.org/10.1016/j.rser.2019.03.040`

Xingjian SHI, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun WOO. 2015. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In *Advances in Neural Information Processing Systems*, Vol. 28. Curran Associates, Inc.

Talha Ahmad Siddiqui, Samarth Bharadwaj, and S. Kalyanaraman. 2019. A Deep Learning Approach to Solar-Irradiance Forecasting in Sky-Videos. *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* (2019), 2166–2174.

Shaojian Song, Zijun Yang, HuiHwang Goh, Qingbao Huang, and Gang Li. 2022. A novel sky image-based solar irradiance nowcasting model with convolutional block attention mechanism. *Energy Reports* 8 (2022), 125–132. ICPE 2021 - The 2nd International Conference on Power Engineering.

Benjamin K. Sovacool. 2009. The intermittency of wind, solar, and renewable electricity generators: Technical barrier or rhetorical excuse? *Utilities Policy* 17, 3 (2009), 288–296.

Joshua S. Stein, Clifford W. Hansen, and Matthew J. Reno. 2012. Global horizontal irradiance clear sky models : implementation and analysis. `https://api.semanticscholar.org/CorpusID:128920440`

Shanhui Sun, Jan Ernst, Archana Sapkota, Eberhard Ritzhaupt-Kleissl, Jeremy Wiles, Joachim Bamberger, and Terrence Chen. 2014. Short term cloud coverage prediction using ground based all sky imager. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. 121–126. `https://doi.org/10.1109/SmartGridComm.2014.7007633`

Yuchi Sun, Vignesh Venugopal, and Adam R. Brandt. 2019. Short-term solar power forecast with deep learning: Exploring optimal input and output configuration. *Solar Energy* 188 (2019), 730–741.

S. Tiwari, R. Sabzehgar, and M. Rasouli. 2019. Short Term Solar Irradiance Forecast based on Image Processing and Cloud Motion Detection. In *IEEE Texas Power and Energy Conference (TPEC)*.

Dmitri Veikherman, Amit Aides, Yoav Schechner, and Aviad Levis. 2014. Clouds in the Cloud. 659–674. `https://doi.org/10.1007/978-3-319-16817-3_43`

Donna Flynn Victor, Morris. 2005. Total Sky Imager (TSI) Handbook. Atmospheric Radiation Measurement (ARM) user facility.

Liang Wei, Tingting Zhu, Yiren Guo, Chao Ni, and Qingyuan Zheng. 2023. CloudpredNet: An Ultra-Short-Term Movement Prediction Model for Ground-Based Cloud Image. *IEEE Access* PP (01 2023), 1–1. `https://doi.org/10.1109/ACCESS.2023.3310538`

Sam West, Daniel Rowe, Saad Sayeef, and Adam Berry. 2014. Short-term irradiance forecasting using skycams: Motivation and development. *Solar Energy* 110 (11 2014), 188–207.

Bruce A. Wielicki, Robert D. Cess, Michael D. King, David A. Randall, and Edwin F. Harrison. 1995. Mission to Planet Earth: Role of Clouds and Radiation in Climate. *Bulletin of the American Meteorological Society* 76, 11 (1995), 2125 – 2154.

Wanyi Xie, Dong Liu, Ming Yang, Shaoqing Chen, Benge Wang, Zhenzhu Wang, Yingwei Xia, Yong Liu, Wang Yiren, and Chaofang Zhang. 2019. SegCloud: a novel cloud image segmentation model using deep Convolutional Neural Network for ground-based all-sky-view camera observation. (2019).

Hanjin Zhang, Bin Li, Shun-Feng Su, Wankou Yang, and Liping Xie. 2024. A Novel Hybrid Transformer-Based Framework for Solar Irradiance Forecasting Under Incomplete Data Scenarios. *IEEE Transactions on Industrial Informatics* 20, 6 (2024), 8605–8615.

Jinsong Zhang, Rodrigo Verschae, Shohei Nobuhara, and Jean-François Lalonde. 2018. Deep photovoltaic nowcasting. *Solar Energy* 176 (2018), 267–276. `https://doi.org/10.1016/j.solener.2018.10.024`

S. Zhang, Z. Dong, X. Yang, S. Chai, Z. Xu, and D. Qi. 2019. Intrahour Cloud Tracking Based on Optical Flow. In *Chinese Control Conference (CCC)*.

Zhao Zhen, Jiaming Liu, Zhanyao Zhang, Fei Wang, Hua Chai, Yili Yu, Xiaoxing Lu, Tieqiang Wang, and Yuzhang Lin. 2020. Deep Learning Based Surface Irradiance Mapping Model for Solar PV Power Forecasting Using Sky Image. *IEEE Transactions on Industry Applications* 56, 4 (2020), 3385–3396. `https://doi.org/10.1109/TIA.2020.2984617`