Robust Heading Estimation from Polarization Images by Deep Neural Networks

Yadala Chanchu, L.S.; Zuidgeest, R.F.G.; Stam, D.M.; de Croon, G.C.H.E.

Publication date
2024

Document Version
Final published version

Published in
IMAV 2024: Proceedings of the 15th annual International Micro Air Vehicle Conference and Competition

Citation (APA)
Yadala Chanchu, L. S., Zuidgeest, R. F. G., Stam, D. M., & de Croon, G. C. H. E. (2024). Robust Heading Estimation from Polarization Images by Deep Neural Networks. In T. Richardson (Ed.), *IMAV 2024: Proceedings of the 15th annual International Micro Air Vehicle Conference and Competition: September 16-20, 2024 Bristol, United Kingdom* (pp. 56-63). Article IMAV2024-6

Robust Heading Estimation from Polarization Images by Deep Neural Networks

L. Yadala Chanchu[*], R.F.G. Zuidgeest[*], D.M. Stam[†], G.C.H.E de Croon[‡]

ABSTRACT

Heading estimation is vital for the autonomous flight of unmanned aerial vehicles. Magnetometers are typically used for this purpose, but they are not robust to electro-magnetic interferences. As a promising alternative, we investigate the insect-inspired solution of skylight polarization sensing. In particular, we develop a robust polarization compass for azimuth estimation. Two datasets are created - one based on a Mie scattering simulation, and one containing real-world pictures captured with a polarization camera under a variety of weather conditions. We employ the ResNet-18 model, which is trained and tested on both datasets separately. The trained model is robust to different weather conditions, and is able to directly analyze maps in the instrumental plane. The median error on the (mostly cloudy) real-world images of 4.30 degrees makes it a promising new method for the navigational toolkit of UAVs. We publish the real-world polarization dataset as open access data, in order to facilitate improvements by the community.

1 INTRODUCTION

The rise of autonomous systems has brought with it the challenge of navigating effectively and robustly in the multitude of environments these systems can be used. As the demand for unmanned aerial vehicles (UAVs) continues to grow in a variety of sectors, challenges associated with reliable navigation under non-ideal conditions become more relevant. UAVs need to navigate effectively also under adverse environmental conditions which can include limited or blocked visibility and disrupted satellite signals.

Conventionally, drone navigation is done through the use of GPS, inertial measurement units (IMUs), and magnetometers [1], but these methods have significant shortcomings. GPS relies on clear signal communication to satellites, which is not possible in the case of large obstructions. On the other hand, IMUs will accumulate errors while the drone is flying [2]. Magnetometers are sensitive to external influences like electrical cables or large metal structures such as ships [3]. For this reason, skylight polarization is being heavily researched as an alternative navigation method.

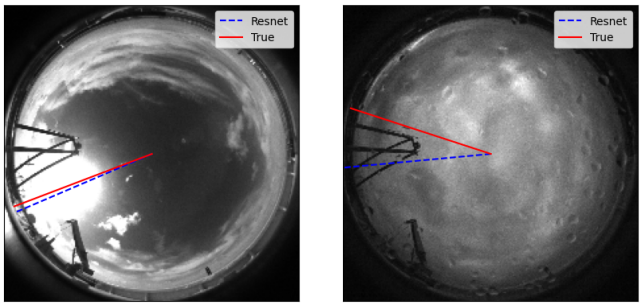


Figure 1: Sun direction estimation by a deep neural network based on polarization images. Red lines are ground-truth sun directions based on location and time. Blue dashed lines represent the predictions of the trained ResNet model. Predictions shown for (left) a clear sky and (right) an overcast sky.

Work on skylight polarization has mainly focused on azimuthal angle determination in clear sky conditions to test their efficacy. Traditional hand-made methods are often verified with captured clear sky images or simulations based on Rayleigh scattering. Newer models based on machine learning (ML) are predominantly validated on simulation datasets, which might not reflect the real world imperfections that occur - especially with the presence of heavily overcast skies. Related work has utilised many different approaches, from exploiting the anti-symmetry present [4] along the solar meridian in Angle of Polarisation (AOP) maps to ML-based techniques which experiment with different data label encodings [5]. However these methods focus on performance under ideal conditions, typically under near-ideal sky conditions or simulation data.

It has been observed that non-ideal environmental conditions like cloud cover significantly reduce the efficacy of present methods to determine heading [6]. In this study, the aim is to further investigate how a machine learning-based approach behaves under less-than-ideal sky conditions and compare performance when trained on simulated images.

2 BACKGROUND

Polarization preliminaries

Polarization cameras capture polarized light intensity in four different directions using a polarization array, shown in

[*]Faculties of EEMCS and ME at Delft University of Technology, equal contribution
[†]Leiden Observatory
[‡]Faculty of AE at Delft University of Technology, contact: g.c.h.e.decroon@tudelft.nl

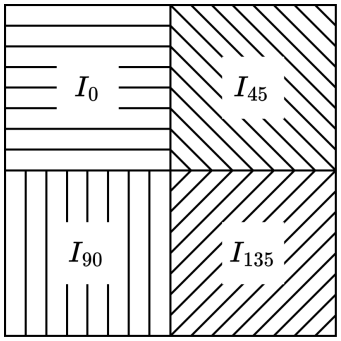


Figure 2: The polarization array of a single pixel.

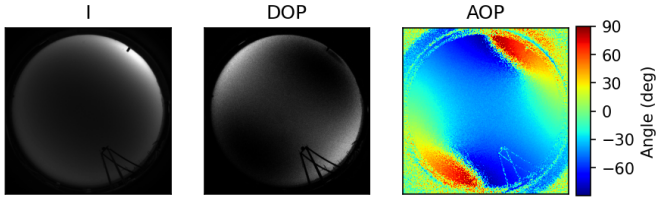


Figure 3: Intensity map with corresponding DOP and AOP.

figure 2. This polarization information is then translated to the Stokes parameters I, Q and U using equation 1. These values describe the polarization state of electromagnetic radiation [6]. There is a fourth parameter, V, describing circular polarization and is negligible for sunlight in the earth's atmosphere.

$$\begin{cases} I = \frac{1}{2}\left(I_0 + I_{45} + I_{90} + I_{135}\right) \\ Q = I_0 - I_{90} \\ U = I_{45} - I_{135} \end{cases} \tag{1}$$

With the Stokes parameters of all pixels, polarization maps can be created. Two common maps are degree of polarization (DOP) and angle of polarization (AOP) maps. Examples of these maps are shown in figure 3. The formula for calculating these maps is shown in equation 2.

$$\begin{cases} DOP = \frac{\sqrt{Q^2+U^2}}{I} \\ AOP = \frac{1}{2}\arctan\left(\frac{U}{Q}\right) \end{cases} \tag{2}$$

There are two reference frames that can be used for the angle of polarization (see figure 5): the instrumental frame (left, fig. 5) - representing the polarisation directly from the measured input and the global frame (right, fig. 5) - representing the polarisation angle with respect to the solar meridian. Conventionally, navigation algorithms rely on the use of an AOP map in the global reference plane. In this reference frame, the solar meridian is a straight line - the axis of central symmetry. This reference frame can be identified by the characteristic '8-shape' visible in the AOP map. AOP-based navigation is a popular choice, as the AOP map is more robust to cloud cover and other aerosols [7].

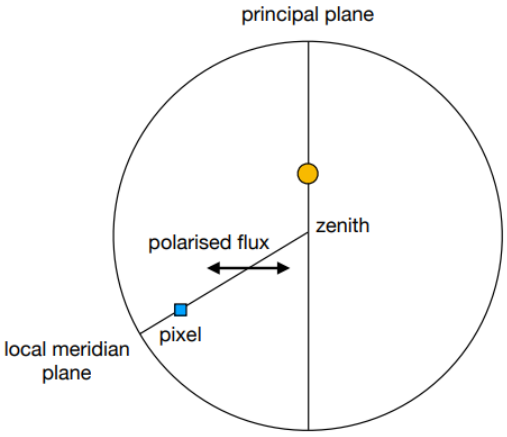


Figure 4: The 2 reference planes projected on a 2D plane, $\beta = 120$ deg (measured clockwise).

The problem arises in obtaining maps in the global reference plane (also known as the principal plane). To transition from the local (instrumental) plane, the stokes parameters are adjusted according to equation 3. A variable rotation Mueller matrix is applied to each pixel, with angle β measured clockwise in degrees from the pixel to the solar meridian [8]. Figure 4 schematically shows how the transformation is carried out. Then the AOP is updated with formula 2. The problem with this transformation, however, is that the location of the solar meridian must be known beforehand. This means that this approach is not usable for applications where the solar meridian is unknown to begin with.

$$\begin{bmatrix} I' \\ Q' \\ U' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(2\beta) & sin(2\beta) \\ 0 & -sin(2\beta) & cos(2\beta) \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \end{bmatrix} \tag{3}$$

Non-ML Algorithms

Several hand-made approaches have been proposed to find the solar meridian in the global plane AOP map. A commonly used baseline is a least squares fit. One such method

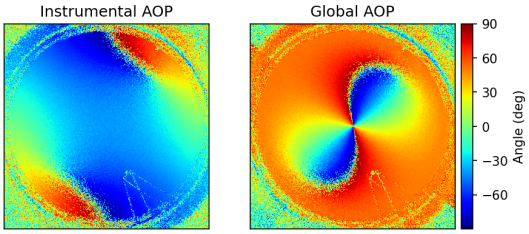


Figure 5: The two reference planes for AOP maps.

tries to find the optimal fit with a symmetry scanning approach of randomly sampled pixel pairs [7]. Other methods are based on fits with a moment of inertia approach [4] or comparing the rotation of pixel rings of sequential AOP maps [9].

DOP-based navigation do not rely on fitting a symmetry line. For example, SkyPole is a novel method that is able to find the celestial pole from a sequence of DOP measurements [10]. In contrast, we are interested in a method that maps a single image to the sun direction.

ML-Based Methods

ML models are able to learn what polarization features are useful for decoding orientation information.

One such approach is based on the intensity map, DOP and AOP [5]. These maps are fed to a fully connected (FC) network, which then outputs a heading estimation. Preprocessing of images with neural networks is also being researched. PCA-Net has showed to be useful in denoising real-life AOP maps [11].

3 METHODOLOGY

In this article, we investigate the use of deep neural networks for mapping sky polarization images to sun direction directly. In this section, we explain the neural network model and dataset.

3.1 ResNet Model

We utilized the ResNet-18 model, which has been pretrained on the ImageNet dataset, due to its capability to effectively learn and extract significant features from a variety of images automatically [12]. Polarization information from the sky can contain subtle patterns and gradients that are crucial for determining heading direction, and ResNet-18's deep architecture can effectively capture these features.

Special care must be taken in the way the azimuth angle is encoded. Previous research has shown that the best neural network performance is achieved when azimuth estimation is transformed to multi-class classification, as opposed to a regression task [5]. The proposed encoding method is known as exponential encoding, $N_e(k)$.

$$\begin{cases} N_e(k) = m^{|i|}, & 0 < m < 1 \\ i = k - \frac{\phi}{j} \end{cases}$$

Where $k \in \left[0, \frac{360}{j}\right]$ denotes the index of the output node, $0 < \phi < 360$ is the encoded angle in degrees, j is the resolution, and m is the constant determining the exponential decay. In essence, i describes how many indices a node k is located away from the encoded angle. Exponential encoding ensures that nodes coherently share the information of orientational mapping. Figure 6 shows an example of an exponential encoding probability density function with a resolution of $j = 0.1$ deg. For this resolution, the last fully connected layer of the ResNet model is altered to contain 3600 output nodes.

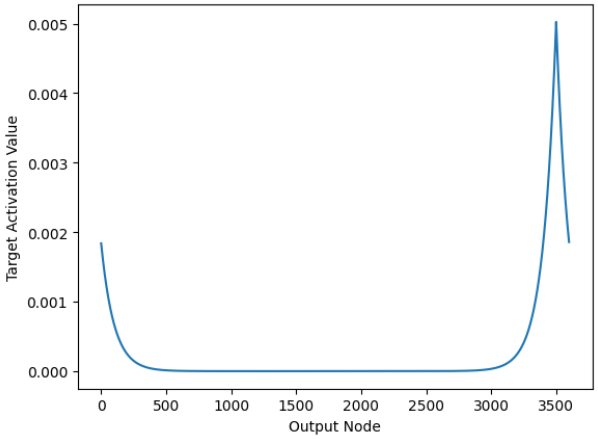


Figure 6: Exponential encoding ($m = 0.99$, $\phi = 350.0$) normalized with a Softmax

Note that by normalizing the exponential encoding to resemble a probability mass function (using a Softmax), the cross-entropy loss metric can be used during training.

3.2 Setup and Dataset

We used two different methods to gather the data required to train the model. The first method creates a synthetic dataset, based on a theoretical model for polarization. This model takes into account Mie scattering and is therefore more realistic than the standard Rayleigh model [13].

The simulation outputs Gaussian quadrature points for a map of the sky hemisphere. With the help of the Python OpenCV library, the I_0, I_{45}, I_{90}, I_{135} intensity maps are acquired and scaled to 224 x 224 pixels. To simulate the information loss that a cloudy atmosphere would introduce, it is possible to randomly obscure information in elliptical areas. This approach mimics the loss of information a cloud can introduce, and is used in the PSNS dataset [14]. Between 1 and 5 ellipsoidal clouds are randomly added to the images, with center pixel locations drawn from $\mathcal{N}(112, 28^2)$, axis lengths of $\mathcal{N}(224/5, (224/25)^2)$ pixels, and a random integer rotation of $\mathcal{U}(0, 90)$ degrees.

The second method captures real-world data with a physical setup, located on the rooftop of the faculty of aerospace engineering at TU Delft, in Delft, the Netherlands. Real-world pictures were captured with the Sony XCG-CP510 camera and the FE185C057HA-1 Fujifilm fisheye lens. The experimental setup is shown in figure 7. The images were collected from 8:00 until 19:00 with picture intervals of 10 minutes to get a varied dataset. At the time of writing, the setup has collected 2618 images over the course of 4 months, from January 2024 until April 2024*. The camera was regularly

*With some breaks between due to overheating and exposure issues

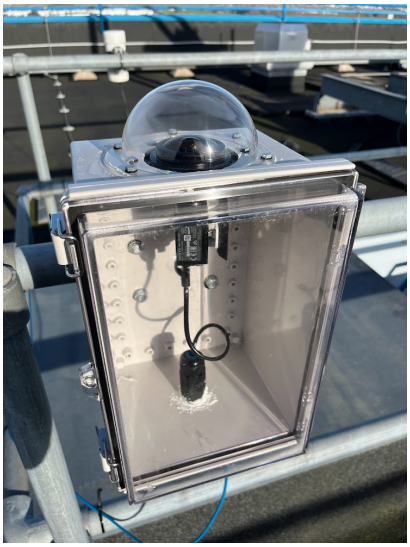


Figure 7: Set-up as mounted on the rooftop

rotated during this period. More details about this dataset can be found in the Appendix.

Rotation of the camera is necessary, as there are fixed obstructions on the rooftop which might provide the model information about the camera orientation. This provides the model with information. By rotating the camera, it is made more difficult for the neural network to infer the sun's azimuthal angle.

Images are captured using the open-source Aravis viewer [15] (version 0.8.30). Utilizing the process of demosaicing, we proceed to extract the I_0, I_{45}, I_{90}, I_{135} intensity maps from the collected raw images.

Each instance in the dataset consist of the Stokes parameters I, Q, U. Equation 1 shows that the Stokes parameters can be retrieved using minimal pre-processing of the original intensity maps. Only +/- operations and rescaling by a factor 1/2 is necessary. For both the simulation and real data, the maps have been saved as 3 x 224 x 224 tensors. The 3 channels are for each Stokes' parameter respectively. This data shape is required to be used as input to ResNet-18 and many other standard convolution neural networks.

In order to investigate the effect of noise generated by clouds on the final error of the trained model, we used a 3x3 kernel on the AOP maps, which computes the standard deviation of the pixels contained in the kernel. This gives an estimate for the noisiness of an image, which corresponds well to how "cloudy" an image is, since clouds introduce noise into the AOP mapping [16]. Figure 8 shows this rudimentary approach can successfully classify cloudiness, although outliers still exist.

4 RESULTS

For ResNet, hyperparameter tuning is run for the real data with a train/test split of 85/15. The result of 5-fold cross-

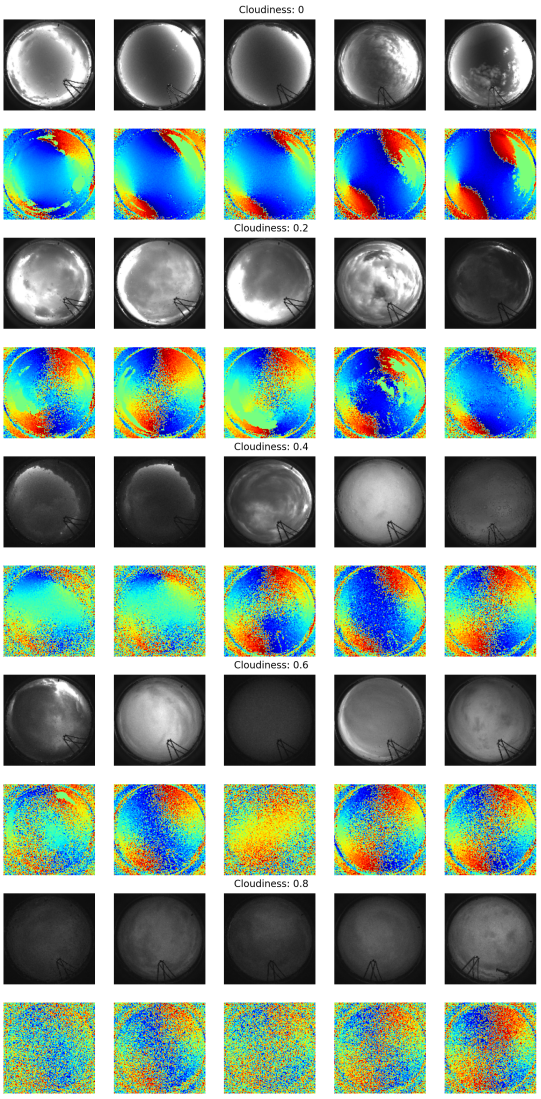


Figure 8: Example intensity maps, sorted by their cloudiness values

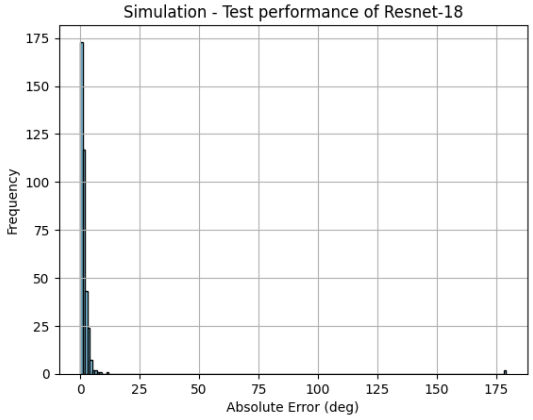


Figure 9: ResNet simulation error distribution

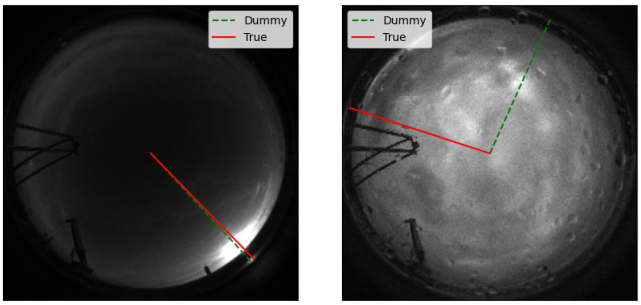


Figure 10: Brightest spot prediction on (left) a clear sky and (right) an overcast sky

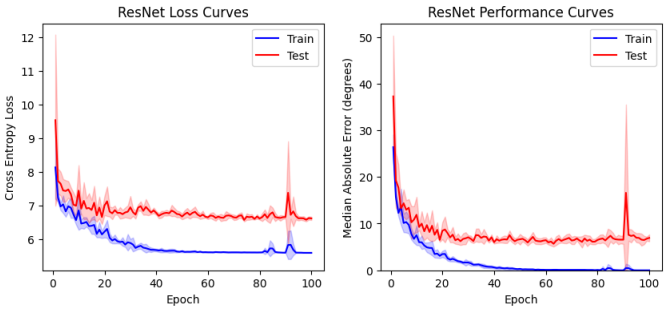


Figure 11: ResNet training curves on the real images

validation on the training set is shown in table 1. The selected parameters of lr = 5e-3 and m = 0.98 are used for the next subsections. The optimizer used is Adam with default PyTorch settings [17].

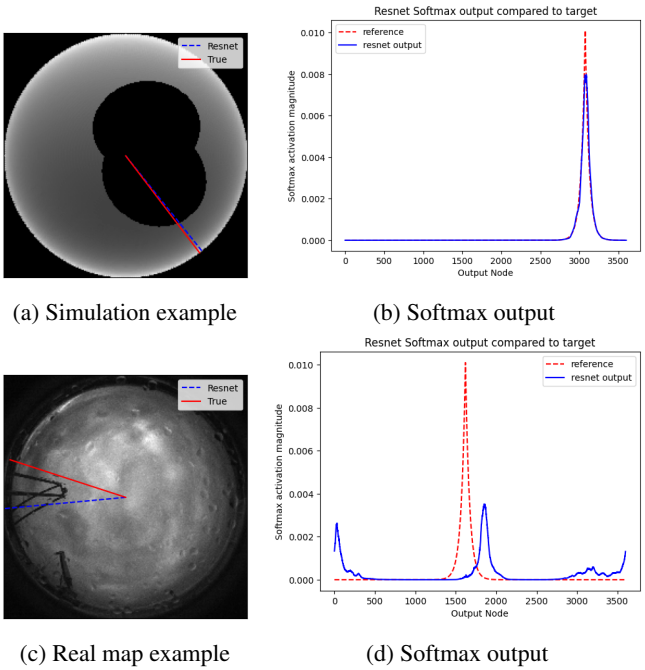
Learning rate	Exponential encoding decay, m			
	0.97	0.98	0.99	0.995
1e-4	29.86	22.72	26.29	26.83
5e-4	25.3	**21.81**	23.75	30.91
1e-3	30.69	24.60	27.68	24.02

Table 1: Average error (degrees) of different hyperparameters during gridsearch with a 65/20/15 train-val-test split after training for 50 epochs

Simulation

An augmented dataset of 2500 images is used as a basis for training the Resnet model using transfer learning [18]. The distribution of errors is visible in figure 9, with a final median absolute error of 1.00 degrees and a mean absolute error of 3.28 degrees.

To show the result of the model training, a simple visualization is shown in figure 12a. The Resnet model is able to mimic the exponential target output. This means the azimuth angle is accurately estimated, while being robust to the cloud mask.



(a) Simulation example (b) Softmax output

(c) Real map example (d) Softmax output

Figure 12: Examples of maps with their corresponding Softmax outputs. The top figures are simulated examples (a, b), and the bottom figures are real examples (c, d).

Real Data

Using the same train-test split of 85/15 that was used in the hyperparameter tuning, the model is trained on the dataset of 2618 real images. Figure 11 shows the learning curves, averaged over 5 training runs of 100 epochs each. The shaded area signifies a single standard deviation. The best performing model achieves a median absolute error of 4.30 degrees, and an mean absolute error of 12.10 degrees. The error distribution is visible in figure 15b. Figure 1 shows two example predictions of the trained ResNet model under different weather conditions.

We compare the robustness of the model by using a brightest spot predictor. This brightest spot model assumes the brightest spot in the image to be the sun as shown in figure 10. The best median error performance was found by pre-processing the images with a 33x33 Gaussian blur kernel. The resulting distribution of errors in shown in figure 15a. Figure 10 shows example predictions of the brightest spot model.

The brightest spot estimator performs worse than ResNet, with a median absolute error of 10.9 degrees and a mean absolute error of 27.24 degrees. The scatter plot in figure 14b and 14a shows the influence of cloud cover on both model performances. In general, a higher degree of cloudiness leads to both models estimating the solar meridian poorly (the brightest spot algorithm more so than the ResNet model). This reinforces the hypothesis of clouds introducing noise in polar-

ization information.

Looking at one such cloudy picture (figure 12c), the influence of the ResNet model becomes clear. The peak is lower when compared to 12a, which means the model is less certain about its prediction. Note that a second, lower peak is located 180 degrees from the predicted angle. Although the model correctly predicts the angle, this indicates the model is certain about the solar meridian orientation, but uncertain about the direction of the sun.

5 DISCUSSION

The accuracy of the model with simulated clouds is much higher compared to using real images. The median simulation accuracy is 1.0 degrees, whereas with real images it is 4.3 degrees. This highlights that the simulated clouds do not obscure polarization information in the same way that real clouds do. Another explanation for this difference in performance is the fact that the simulation produces maps in the global reference frame, where AOP symmetry is easily found.

From the plot in figure 14b, it is clear that from cloudiness levels above 0.7, the model accuracy is impacted. This demonstrates that our model is able to determine heading to some extent even with patchy skies, but it becomes increasingly difficult when there is less direct sky visibility.

In terms of performance, our model shows a clear advantage over the brightest spot algorithm, especially for cloudier skies. This indicates that our model is able to effectively estimate the solar meridian even in challenging conditions.

The collected dataset has limitations that need to be considered. The images have obstacles that obstruct the clear view of the sky, which can impact the accuracy of our model. These obstacles could have been used by the model as landmarks to help determine the solar angle. However, training the model with the objects all masked out with a circular mask, as shown in Figure 13, doesn't seem to significantly affect model performance. When trained on these masked images, even with a narrower field of view, the median error is 5.4 degrees compared to the original median error of 4.3 degrees. This shows that the model is in fact determining the heading from the polarisation information and not the obstacles. The Appendix shows the distribution of the azimuthal angles in our dataset is imbalanced. This might introduce bias in the model's performance.

6 CONCLUSION

In this paper, a polarization dataset has been created of skylight hemisphere pictures. The dataset contains 2618 images with various weather conditions. Additionally, a ResNet-18 model has been trained to perform solar azimuth estimation based on Stokes' parameters in the instrumental plane. The model attains a median absolute error of 1.0 degrees on simulation data with cloud masks, and 4.3 degrees on the real data. To compare, a brightest spot algorithm that predicts the direction of the brightest intensity achieves a me-

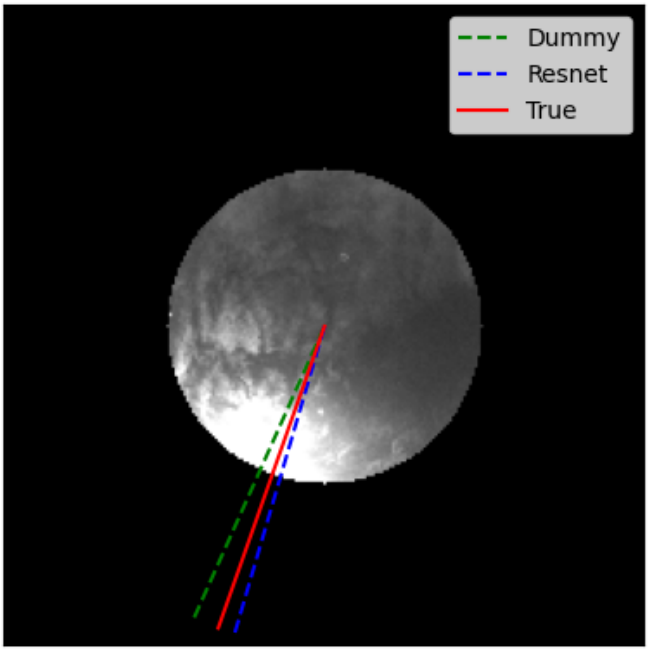
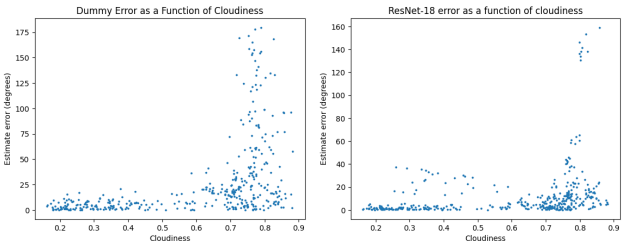
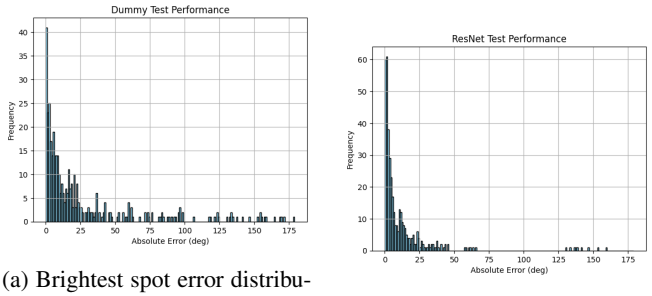


Figure 13: Masked image to determine effect of obstacles on model accuracy



(a) Brightest spot algorithm (b) ResNet

Figure 14: Influence of cloudiness on error



(a) Brightest spot error distribution (b) ResNet error distribution

Figure 15: Comparative error distributions

dian absolute error of 10.9 degrees on the real data. This research showcases that azimuthal angle estimation is possible with reasonable accuracy even under non-ideal weather conditions.

Future research is possible in expanding the polarization dataset to contain more images and improving the balance of the azimuth distribution. Also, a more refined metric for cloudiness could be developed to better quantify the influence of weather conditions on model performance. More advanced or custom machine learning models for azimuth estimation can be developed and validated on the dataset.

ACKNOWLEDGEMENTS

Setting up the physical setup would not have been possible without Stein Stroobants, Jesse Hagenaars, Erik van der Horst, Martin Søndergaard, and Bart Root.

DATA AVAILABILITY

The code and real dataset are both available upon reasonable request.

REFERENCES

[1] Kimon P Valavanis. Advances in unmanned aerial vehicles: state of the art and the road to autonomy. 2008.

[2] G Balamurugan, J Valarmathi, and V P S Naidu. Survey on uav navigation in gps denied environments. 2016 International conference on Signal Processing, Communication, Power and Embedded System (SCOPES), IEEE, 10 2016.

[3] D. Gebre-Egziabher and B. Taylor. Impact and mitigation of gps-unavailability on small uav navigation, guidance and control white paper. 10 2012.

[4] H. Liang, H. Bai, N. Liu, and X. Sui. Polarization navigation simulation system and skylight compass method design based upon moment of inertia. *Mathematical Problems in Engineering*, 204, 2020.

[5] H. Liang, H. Bai, K. Hu, and X. Lv. Polarized light sun position determination artificial neural network. *Applied Optics*, 61(6):1456, February 2022.

[6] G. Qian, J. Tang, C. Shen, and J. Liu. A highly robust polarization orientation method based on antisymmetry of skylight polarization pattern. *Measurement*, 2022.

[7] H. Zhao, W. Xu, Y. Zhang, X. Li, H. Zhang, J. Xuan, and B. Jia. Polarization patterns under different sky conditions and a navigation method based on the symmetry of the aop map of skylight. *Optics Express*, 26(22):28589–28603, 2018.

[8] L. M. Eshelman and J. A. Shaw. Visualization of all-sky polarization images referenced in the instrument, scattering, and solar principal planes. *Optical Engineering*, 2019.

[9] L. Guan, S. Liu, J. Chu, R. Zhang, Y. Chen, S. Li, L. Zhai, Y. Li, and H. Xie. A novel algorithm for estimating the relative rotation angle of solar azimuth through single-pixel rings from polar coordinate transformation for imaging polarization navigation sensors. *Optik*, 178:868–878, 2019.

[10] T. Kornland-Martinet, L. Poughon, M. Pasquinelli, D. Duché, J. R. Serres, and S. Viollet. Skypole—a method for locating the north celestial pole from skylight polarization patterns. *PNAS*, 120(30), 2023.

[11] X. Wu, C. Shen, D. Zhao, C. Wang, H. Cao, J. Tang, and J. Liu. Robust orientation method based on atmospheric polarization model for complex weather. *IEEE Internet of Things Journal*, 10(6):5268 – 5279, 2022.

[12] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition, 2015.

[13] D. M. Stam, J. F. De Haan, J. W. Hovenier, and P. Stammes. Degree of linear polarization of light emerging from the cloudless atmosphere in the oxygen a band. *Journal of Geophysical Research*, 104(D14):16843 – 16858, 1999.

[14] H. Liang and H. Bai. Polarized skylight navigation simulation (psns) dataset. 2021.

[15] Aravis Project. Aravis.

[16] L. Shunzi, K. Fang, X. Han, G. Xiaohan, L. Haozhe, R. Yaohuang, C. Shouhu, and G. Yinjing. Biomimetic polarized light navigation sensor: A review. *Sensors*, 23(13), 2023.

[17] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. 2017.

[18] J.Plested and T.Gedeon. Deep transfer learning for image classification: a survey. May 2022.

APPENDIX A: DATASET DETAILS

The cloudiness values of the collected images are bimodal as the collection period was characterized by many mainly cloudy weeks followed by mainly clear skies as shown in figure 16.

Figure 17 shows that the azimuthal angles distribution in the dataset is very unbalanced. This will most likely introduce significant bias in trained models, so this is something to be aware of.

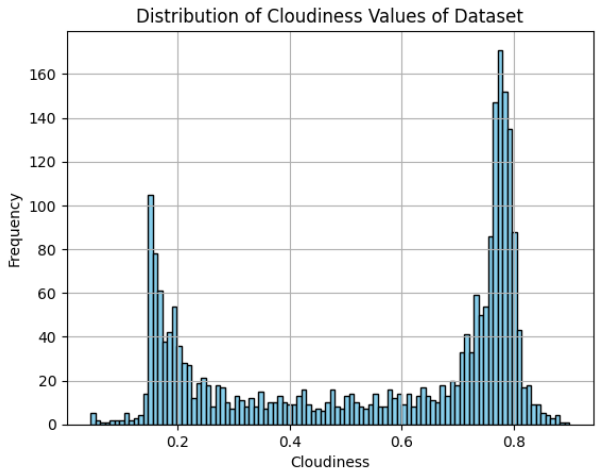


Figure 16: Distribution of cloudiness values in the dataset

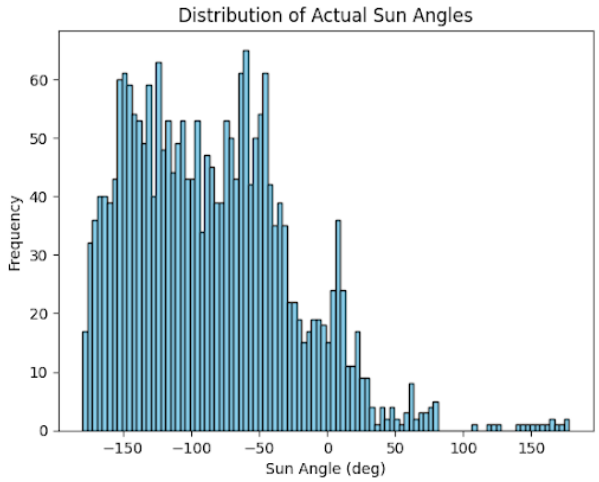


Figure 17: Distribution of azimuthal angles in dataset