Information Engineering Express
International Institute of Applied Informatics
2025, Vol. 11, No. 1, IEE866

Improving a Night-Time Weather Estimation Method with Low-Cost Devices using a Support Vector Machine

Shota Kusabiraki * , Shin Suzuki †,
Nobuhiro Inuzuka * , Kazunori Iwata ‡

Abstract

Because of the increasing density of meteorological observation networks and technological advances in data processing, weather forecast accuracy has improved in the past decades and is projected to increase further, requiring measurements at the local scale. However, their biggest problem is the equipment cost; the ultrasensitive atmospheric light cameras cost approximately ten million yen. This paper introduces a weather observation system designed exclusively with low-cost commercial products. The combined cost of the fabricated device was approximately 100,000 yen. An algorithm was also established to estimate night-time weather from the number of visible stars in all-sky images by assuming that more visible stars imply fewer clouds and more precise weather (the star count method). The resulting weather estimation accuracy was about 80 %. In addition, to improve the accuracy, this paper adopts a supervised machine learning method, Support Vector Machine (SVM), and compares the classification accuracy with the star count method.

Keywords: image processing, low-cost devices, support vector machine, weather estimation

1 Introduction

Benefiting from the densification of observation networks and improvements in data processing, the accuracy of daily weather forecasts has been continuously increasing, but requires more measurements of higher quality at the local scale. Network densification is mainly achieved with simpler, self-sufficient observation equipment such as the Automated Meteorological Data Acquisition System (AMeDAS). Improvement in data processing currently depends primarily on machine-learning techniques. Combining network densification and improved data processing technologies yields weather forecasts with higher accuracy.

Although local meteorological observations are essential, their biggest problem is the equipment cost. For example, observations of atmospheric airglow, an emission phenomenon

* Graduate School of Engineering, Nagoya Institute of Technology, Nagoya, Japan
† Faculty of Regional Policy, Aichi University, Toyohashi, Japan
‡ Faculty of Business Administration, Aichi University, Nagoya, Japan

that occurs in the upper atmosphere, have traditionally been acquired with ultrasensitive atmospheric light cameras costing approximately ten million yen [1, 2, 3]. More recently, airglow has been observed with inexpensive CCD video cameras [4, 5]. The images are taken using a CCD camera and optical filter, which limits the angle of view, but they can be used for atmospheric light observation and analysis. Atmospheric light observations must avoid cloud cover at night, and in fact [6], the analysis was done at night when there was no cloud cover for at least one hour.

In [7], thus, using a low-cost nighttime weather observation system that combines commercial products and image processing techniques, the proposed method could estimate the cloud cover based on the number of stars and classify the weather as Clear, Sunny, or Cloudy. In that paper, the weather was estimated with about 80 % accuracy, but the edges of clouds were mistakenly recognized as stars. Thus, even when most of the sky was covered with clouds, the weather was sometimes estimated as "clear" or "sunny" in the all-sky. In addition, the number of images from sunset to dawn in these studies is enormous because the images are taken every 5 seconds. Then, such a large amount of image data requires much time and work to find suitable images to analyze atmospheric light observations.

Therefore, this study aims to develop a better weather estimation model using supervised machine learning and to compare the classification accuracy with the star count method.

This paper is an extension of work originally presented in 14th International Congress on Advanced Applied Informatics [7].

2 Data and Analysis Methods

2.1 Image Data and Weather Definition

In this study, all-sky nighttime images acquired with an airglow observation system. The system exclusively includes widely-available commercial equipment. Although the field-of-view is limited, the spatial resolution is markedly higher than that of geostationary meteorological satellites. A commercial CCD camera is equipped with an optical filter that transmits only at the wavelengths of the hydroxyl radical airglow (600–900 nm). Therefore, atmospheric light collected by the camera originates at the mesopause, at an altitude of approximately 85 km. This image was taken in 2019 at in Kühlungsborn, Germany (54°N, 11°E), size is 640×480 pixels, exposure time is 4 seconds, temporal resolution is 5 seconds, and sampling interval is 1 minute. To minimize the influence of ground-level artificial lights and to account for image distortion by the fisheye lens, the edge of the difference image was masked to limit the analysis to the image portion within zenith angles of 60°–90°.

Figure 1 shows a typical image used for this analysis. It is an all-sky nighttime image acquired at 23:13 on July 1st, 2019 (JST: Japan Standard Time). The edge of the circular image corresponds to the horizon and the center is the zenith. The image is oriented with north at the top and east on the right.

The all-sky images were divided into "Clear," "Sunny," "Cloudy," and "Full Clouds," and there are 1,000 images for each weather condition. Table 1 is a summary of the weather categories and their details. This determination is based on Japan Meterological Agency (JMA) weather summaries [8] with blind estimation and exclusion of ambiguous weather.

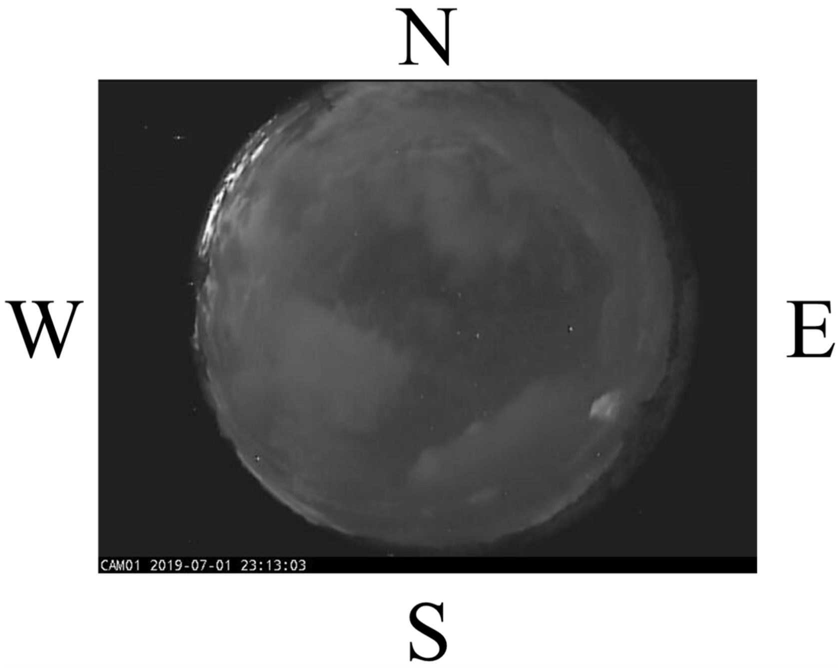

Figure 1: All-Sky Nighttime Circular Image

Table 1: Weather Type Definition as a Function of Cloud Cover

Weather Type	Amount of Clouds
Clear	Clouds cover from 0 % or more to under 20 %
Sunny	Clouds cover from 20 % or more to under 50 %
Cloudy	Clouds cover from 50 % or more to under 80 %
Full Clouds	Clouds cover from 80 % or more to 100 % or less

2.2 Star Count Weather Estimation

The following image processing algorithm applies to each nighttime image to count the number of stars:

1. A median filter (kernel size: 11×11 pixels) was applied to create a smoothed image.
2. A difference image was created by subtracting the smoothed image from the original image.
3. Pixels in the difference image with a value exceeding a preset brightness threshold were assumed to represent the positions of stars.

The brightness threshold value (T) was defined in Equation (1) as the product of the sum of the mean pixel value (m) and corresponding standard deviation (s) from the difference image and of an arbitrary coefficient ($2m/s$):

$$T = \frac{2m}{s}(m+s) \tag{1}$$

Moreover, stars present in each image were spread over several pixels. Therefore, a minimum discrimination distance of 8 pixels was set, within which distinct brightness peaks, possibly representing different stars, were counted as one object.

Amount of cloud (A) with Star Count was estimated in each image as:

$$A = \frac{C - P}{C}, \tag{2}$$

where P is the number of identified stars and C is the maximum number of stars in clear-sky conditions. Clear-sky conditions correspond to situations with no obstruction of stars by clouds, when star count is the largest.

2.3 SVM Weather Estimation

This study uses the supervised machine learning method, Support Vector Machine (SVM) [9], to improve weather estimation accuracy. The SVM model estimates the weather in an image.

2.3.1 Model Creation and Model Accuracy

One thousand images of each weather condition, described in subsection 2.1, are used for training an SVM. To confirm the accuracy of the created model, 5-fold cross-validation was performed. Table 2 shows the results of the cross-validation and Table 3 illustrates precision and recall for each class. The accuracy is 0.915, the macro averaging precision is 0.924 and the macro averaging recall is 0.915 [10]. These results indicate that the SVM model has enough accuracy to predict weather from an image.

Table 2: Confusion Matorix for the Classification

		Correct Weather Type			
		Clear	Sunny	Cloudy	Full Clouds
Weather Estimation	Clear	989	0	0	0
	Sunny	0	725	55	0
	Cloudy	11	275	945	0
	Full Clouds	0	0	0	1000

Table 3: Precision and Recall by Class

	Clear	Sunny	Cloudy	Full Clouds
Precision	1.000	0.929	0.768	1.000
Recall	0.989	0.725	0.945	1.000

3 Weather Type Estimation with Each Method and Comparison

The target image to estimate weather is on the date that the star count method fails to estimate some weather (from July 3rd to July 4th, 2019) [7], and this image is not included in the SVM training data. The results of the estimations by each model and the actual weather are shown in Table 5. The estimations by the SVM model are labeled as “Weather Estimations with SVM.” In contrast, the estimations based on the stars count method are labeled as “Weather Estimations with Star Count.” In addition, “Actual weather” is the correct weather at the time. The color-coding rules for Table 5 are outlined in Table 4.

Table 4: Color-Coding Rules in Table 5

Color	Rules
White	Both SVM and Star count are consistent with Actual weather.
Bule	Only SVM agrees with Actual weather.
Yellow	Only Star count agrees with Actual weather.
Gray	Neither SVM nor Star count is consistent with Actual weather.

4 Consideration

This section discusses the advantages and disadvantages of the SVM model according to the results shown in Table 5. Some figures in this section are used to explain their details. The zone inside the red circle corresponds to zenith angles of 60°–90°, and each green circle in the figures indicates the estimated position of a star.

4.1 Pros

4.1.1 Time at 23:57

Figure 2 (a) and (b) show the original image at 23:57 on July 3rd and the processed image for the tar count method, respectively. The star count method judged the weather to be "Cloudy," despite the actual weather being "Full Clouds," the cloud edge was incorrectly identified as a star in the yellow line area. On the other hand, the SVM estimation was "Full Clouds," which is consistent with the actual weather.

4.1.2 Time at 0:09

The original and the processed images at 0:09 on July 4th are shown in Figure 3 (a) and (b), respectively. In this case, the clouds are sparsely distributed in the image; hence, star detection is difficult. The estimation based on the star count method results in "Sunny." However, some areas in the yellow lines in Figure (b) include cloud edges rather than stars. Therefore, the SVM estimation, which predicts "Cloudy", is accurate.

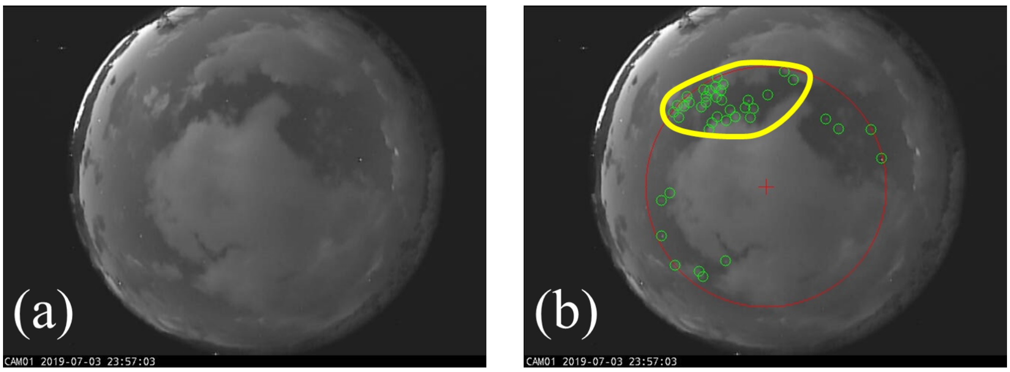

Figure 2:
(a) July 3rd, 23:57 Original image (b) July 3rd, 23:57 Processed image

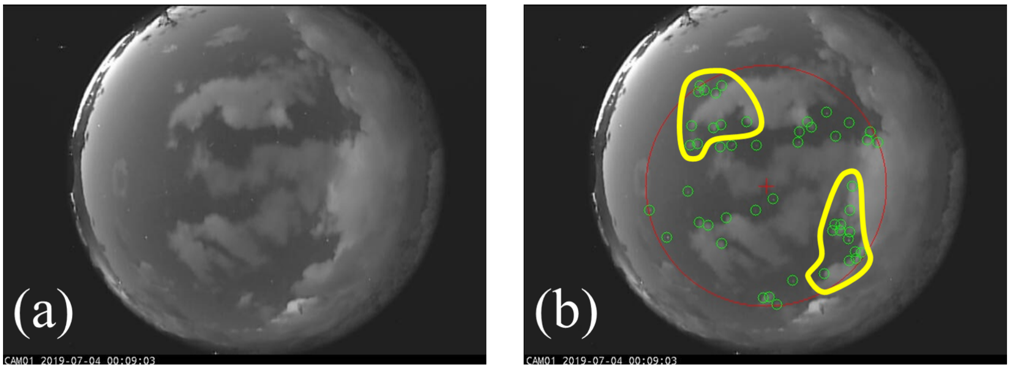

Figure 3:
(a) July 4th, 0:09 Original image (b) July 4th, 0:09 Processed image

4.2 Cons.

4.2.1 Time at 23:46

Figure 4 (a) and (b) are the original image and the processed image at 23:46 on July 3rd, respectively. The SVM estimation indicates “Full Clouds,” but the star count method estimates “Cloudy.”. The actual weather is “Cloudy” by considering the gaps among the clouds, so the star count method is more appropriate in this case.

4.2.2 Time at 0:16

During this period, bright clouds exist in the east and southeast. Consequently, weather estimation based on the star count misidentifies the edges of these clouds as stars, resulting in an excessively high star count and leading to the weather “Clear.” Additionally, the SVM estimation does not yield “Clear” or “Sunny,” and categorizes the weather as “Cloudy.” Therefore, during such periods, neither method can estimate the weather correctly.

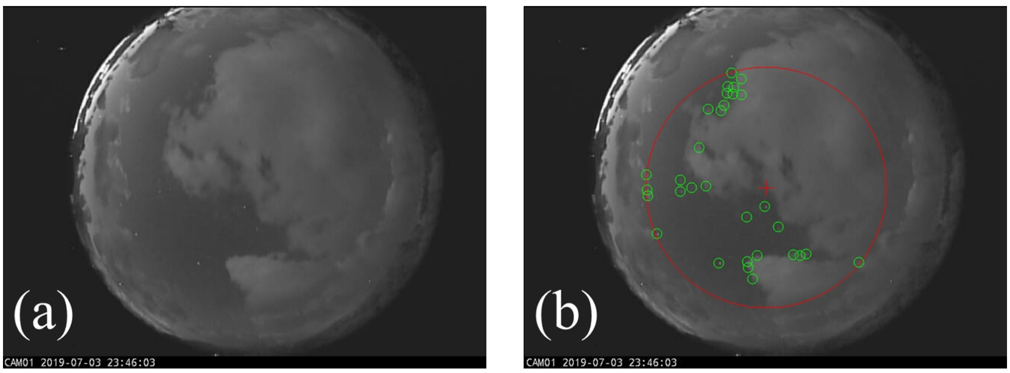

Figure 4:
(a) July 3rd, 23:46 Original image (b) July 3rd, 23:46 Processed image

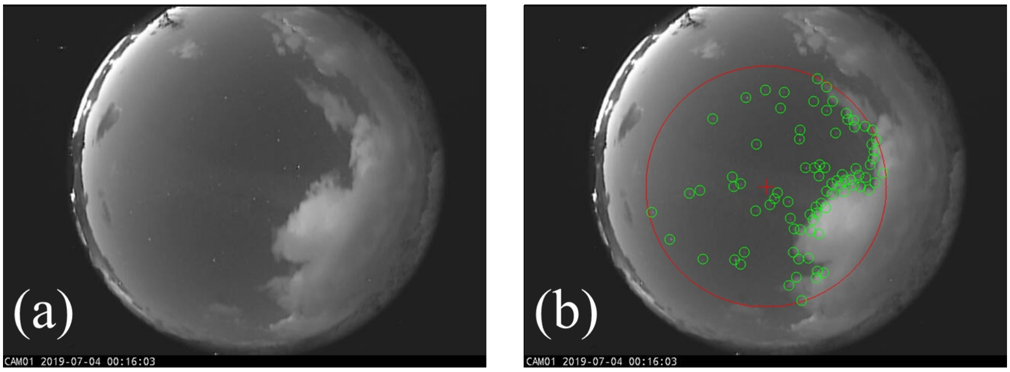

Figure 5:
(a) July 4th, 0:16 Original image (b) July 4th, 0:16 Processed image

4.3 Summary of Consideration

Table 5 reveals that the SVM model and the star count method have different strengths and weaknesses in estimating weather. To improve the accuracy of nighttime weather estimation, both methods must be combined and complement each other. In addition, it is also necessary to create more accurate models to distinguish between “Sunny” and “Cloudy”, and between “Cloudy” and “Full Clouds.”

5 Conclusion and Future Works

In this study, to improve the accuracy of nighttime weather estimation, the supervised machine learning method, SVM, is used. One thousand images of each weather condition, ”Clear,” ”Sunny,” ”Cloudy,” and ”Full Cloud,” were used to train an SVM, and the accuracy of the SVM model was confirmed using the macro averaging precision and the macro averaging recall. In addition, this paper compares the classification accuracy of the SVM model with that of the method of weather estimation based on the number of stars. The results show that the SVM model is superior to the star count method in some cases but only sometimes, and reveal that the SVM model and the star count method have different strengths and weaknesses in estimating weather.

Due to the differences in the types of images suitable for estimation between the SVM model and the star count method, developing a combined system that integrates both estimations is expected to improve nighttime weather estimation. Specifically, when the star count method fails to estimate correctly, the system automatically utilizes the SVM model for weather estimation. Furthermore, the goal is to create more accurate models to distinguish between “Sunny” and “Cloudy” and between “Cloudy” and “Full Clouds ” and to incorporate the proposed models in this paper into a comprehensive system.

References

[1] Shin Suzuki. Development of a multi-station airglow imaging system using commercial cameras (in Japanese). In *MGI34-P08, Japan Geoscience Union Meeting*, 2021.

[2] Shiokawa, Kazuo et al. Ground-based instruments of the PWING project to investigate dynamics of the inner magnetosphere at subauroral latitudes as a part of the ERG-

ground coordinated observation network, url = https://doi.org/10.1186/s40623-017-0745-9. *Earth, Planets and Space*, 69(1):160, 2017.

[3] Kazuo Shiokawa and Shin Suzuki. *Wonder of the Airglow (in Japanese)*. Gentosha, 2018.

[4] Shin Suzuki and Keiichiro Fukazawa. Development of Airglow Camera System for Multi-Point Airglow Imaging Measurements: Detection of Wave Structures (in Japanese). *Aichi University Journal of Regional Policy Studies*, 8(1·2):3–8, 2019.

[5] Yasunobu Ogawa et al. Development of low-cost multi-wavelength imager system for studies of aurora and airglow. *Polar Science*, 23:100501, 12 2019.

[6] Shin Suzuki and Kazuo Shiokawa and Alan Z. Liu and Yuichi Otsuka and Tadahiko Ogawa and Takuji Nakamura. Characteristics of equatorial gravity waves derived from mesospheric airglow imaging observations. *Annales Geophysicae*, 27(4):1625–1629, 2009.

[7] Shota Kusabiraki and Shin Suzuki and Kazunori Iwata. Development of a Night-time Weather Estimation Algorithm Using Image Processing with Low-Cost Devices. In *2023 14th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, pages 271–276, Los Alamitos, CA, USA, Jul. 2023. IEEE Computer Society.

[8] Japan Meteorological Agency. Weather overview terminology. https://www.data.jma.go.jp/obd/stats/data/mdrr/man/gaikyo.html (Accessed: 2023/12/4).

[9] Ioannis Tsochantaridis, Thorsten Joachims, Thomas Hofmann, and Yasemin Altun. Large margin methods for structured and interdependent output variables. *Journal of Machine Learning Research*, 6(50):1453–1484, 2005.

[10] Evidently AI. Accuracy, precision, and recall in multi-class classification. https: // www.evidentlyai.com/classification-metrics/multi-class-metrics (Accessed: 2024/1/28).

Table 5: Estimation Results with Two Methods and Actual Weather Between 23:20 on July 3rd, 2019 and 00:19 on July 4th, 2019

Time	**Weather Estimation with SVM**	**Weather Estimation with Star Count**	**Actual Weather**
23:20	Full Clouds	Full Clouds	Full Clouds
23:21	Full Clouds	Full Clouds	Full Clouds
23:22	Full Clouds	Full Clouds	Full Clouds
23:23	Full Clouds	Full Clouds	Full Clouds
23:24	Full Clouds	Full Clouds	Full Clouds
23:25	Full Clouds	Full Clouds	Full Clouds
23:26	Full Clouds	Full Clouds	Full Clouds
23:27	Full Clouds	Full Clouds	Full Clouds
23:28	Full Clouds	Full Clouds	Full Clouds
23:29	Full Clouds	Full Clouds	Full Clouds
23:30	Full Clouds	Full Clouds	Full Clouds
23:31	Full Clouds	Full Clouds	Full Clouds
23:32	Full Clouds	Full Clouds	Full Clouds
23:33	Full Clouds	Full Clouds	Full Clouds
23:34	Full Clouds	Full Clouds	Full Clouds
23:35	Full Clouds	Cloudy	Full Clouds
23:36	Full Clouds	Cloudy	Full Clouds
23:37	Cloudy	Sunny	Full Clouds
23:38	Cloudy	Sunny	Full Clouds
23:39	Sunny	Sunny	Sunny
23:40	Sunny	Sunny	Full Clouds
23:41	Sunny	Sunny	Full Clouds
23:42	Sunny	Cloudy	Full Clouds
23:43	Sunny	Sunny	Full Clouds
23:44	Sunny	Sunny	Full Clouds
23:45	Sunny	Sunny	Full Clouds
23:46	Full Clouds	Cloudy	Full Clouds
23:47	Full Clouds	Cloudy	Full Clouds
23:48	Full Clouds	Full Clouds	Full Clouds
23:49	Full Clouds	Full Clouds	Full Clouds
23:50	Cloudy	Cloudy	Full Clouds
23:51	Sunny	Cloudy	Cloudy
23:52	Sunny	Sunny	Sunny
23:53	Sunny	Cloudy	Cloudy
23:54	Full Clouds	Cloudy	Cloudy
23:55	Full Clouds	Cloudy	Full Clouds
23:56	Full Clouds	Full Clouds	Full Clouds
23:57	Full Clouds	Cloudy	Full Clouds
23:58	Full Clouds	Cloudy	Cloudy
23:59	Full Clouds	Cloudy	Cloudy
0:00	Full Clouds	Cloudy	Cloudy
0:01	Cloudy	Sunny	Sunny
0:02	Cloudy	Sunny	Sunny
0:03	Sunny	Sunny	Sunny
0:04	Sunny	Clear	Sunny
0:05	Sunny	Sunny	Sunny
0:06	Sunny	Sunny	Sunny
0:07	Sunny	Sunny	Sunny
0:08	Sunny	Sunny	Sunny
0:09	Cloudy	Sunny	Cloudy
0:10	Cloudy	Sunny	Cloudy
0:11	Cloudy	Sunny	Cloudy
0:12	Sunny	Sunny	Sunny
0:13	Sunny	Sunny	Sunny
0:14	Sunny	Sunny	Sunny
0:15	Cloudy	Sunny	Sunny
0:16	Cloudy	Clear	Sunny
0:17	Cloudy	Sunny	Sunny
0:18	Sunny	Sunny	Sunny
0:19	Sunny	Sunny	Sunny