



HAL
open science

StrikeNet: A Deep Neural Network to predict pixel-sized lightning location

Melanie Bosc, Adrien Chan-Hon-Tong, Aurélie Bouchard, Dominique Béréziat

► **To cite this version:**

Melanie Bosc, Adrien Chan-Hon-Tong, Aurélie Bouchard, Dominique Béréziat. StrikeNet: A Deep Neural Network to predict pixel-sized lightning location. 2024. hal-04593481

HAL Id: hal-04593481

<https://hal.science/hal-04593481v1>

Preprint submitted on 29 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

StrikeNet: A Deep Neural Network to predict pixel-sized lightning location

1st BOSCH Mélanie 2nd CHAN-HON-TONG Adrien 3rd BOUCHARD Aurélie 4th BEREZIAT Dominique
ONERA ONERA ONERA LIP6 - CNRS
Université Paris-Saclay Université Paris-Saclay Université Paris-Saclay Sorbonne Université
Palaiseau, France Palaiseau, France Palaiseau, France Paris, France
melanie.bosc@onera.fr adrien.chan-hon-tong@onera.fr aurelie.bouchard@onera.fr dominique.bereziat@lip6.fr

Abstract—Forecasting the location of electrical activity at a very short time range remains one of the most challenging predictions to make, primarily attributable to the chaotic nature of thunderstorms. Additionally, the punctual nature of lightning further complicates the establishment of reliable forecasts. This article introduces StrikeNet, a specialized Convolutional Neural Network (CNN) model designed for very short-term forecasts of electrical activity locations, utilizing sequences of temporal images as input and only two data types. Employing soft Non-Maximum Suppression (NMS) techniques, incorporating morphological features within residual blocks, and implementing dropout regularization, StrikeNet is specifically designed for detecting and predicting pixel-sized objects in images. This design seamlessly aligns with the task of forecasting imminent electrical activity giving great scores of 0.42 for precision, 0.78 for detection, and an F1-Score about 0.54.

Index Terms—Deep learning, convolutional neural networks, thunderstorm risk, very short-term forecasting

I. INTRODUCTION

Atmospheric events such as violent winds, rain, hail, and lightning can significantly impact aviation. These dangerous phenomena can be produced by cumulonimbus clouds, the Earth’s main lightning generators. These clouds need special atmospheric conditions to emerge with the presence of humidity, instability, and a mechanism allowing the development of convection as the convergence of cold and hot air masses. The lightning produced by these systems poses a high risk to aviation, striking aircraft once per year on average. Such incidents can lead to flight safety issues and necessitate mandatory maintenance operations. Therefore, the need for increasingly precise forecasts becomes crucial as thunderstorm systems continue to affect aviation safety.

The forecasting methods currently in use for short and long-term time range predictions are Numerical Weather Prediction (NWP) models, which involve resolving meteorological equations to forecast the state of the atmosphere. In addition, short-range lightning forecast is possible using belief functions or weighting functions as in [1]. To forecast at a very short time horizon (<1 h), a combination of observation-based approaches and algorithmic methods can be employed like in [9]. In recent years, studies have shown that the use

of Deep Learning could lead to better forecasting results for many atmospheric parameters such as precipitation rate as shown in [2] and thunderstorms in [12]. It has also shown improvement in lightning location forecasting by using CNN as in [14] and [7] yet also incorporating radar data in addition to satellite ones. We aim to implement a very short-term forecasting approach utilizing neural networks and only two kinds of satellite data, seeking to improve the overall prediction accuracy for pixel-sized phenomena, specifically lightning flashes within thunderstorm systems.

From a machine learning point of view, this article focuses on the precise prediction of small objects in images through the application of CNN. To our current knowledge, accurately forecasting such specific phenomena remains poorly understood and challenging, prompting our investigation to address this gap in understanding within the confines of this article.

In Section II, we will outline the data we used, followed by an explanation of the developed model in Section III, and ultimately, we will present the obtained results for the forecast of lightning at a very short time range in Section IV before concluding and giving some perspectives in Section V.

II. DATA

A. Data overview

The data employed in this study are collected from two sensors embedded on the Geostationary Operational Environmental Satellite (GOES-R/GOES-16), operated by the National Oceanic and Atmospheric Administration (NOAA) covering the Pacific Ocean, the Americas, and the Atlantic Ocean. The focus is on data captured by the Advanced Baseline Imager (ABI) and the Geostationary Lightning Mapper (GLM) sensors. These two sensors provide complementary information: the first one can identify cloudy areas, while the second can provide information about electrical activity.

B. Sensor’s data

The ABI sensor is a radiometer that offers a resolution of 0.5 km in the visible spectrum and 2 km in the infrared [11]. It has the capability to capture data across 16 wavelength bands. However, for this study, we specifically utilize the 13th band at 10.3 μm due to its heightened sensitivity for cloud classification. The ABI produces images every 5 minutes, offering

brightness temperature data that is particularly effective for detecting low temperatures on top of clouds. This capability enables us to determine if the top of a cloud is high, indicating the likelihood of it being a cumulonimbus and therefore the presence of lightning.

In contrast, the GLM sensor functions as a camera with a nadir spatial resolution of 8 km. Primarily designed for lightning detection, it boasts a detection rate between 70-90%. Operational day and night, the sensor exhibits superior performance during nighttime owing to improved contrast. It captures images every 20 seconds. We have selected flash information from the GLM L2 products, which are the closest to conventional lightning flashes. These flash positions are derived from optical pulse locations within a 330 ms timeframe and a 16.5 km area and are provided at a 1.1 km resolution. Among the various observable areas, data from these two sensors are retrieved over the CONUS (Continental United States) area, which covers North America to the Caribbean.

C. Dataset creation

We collect a dataset consisting of two distinct types of images, to represent near-future lightning strikes at a very short time range. Firstly, radiance images originating from band 13 of the ABI sensor are transformed into brightness temperature images as shown in Fig. 1. Secondly, the flashes captured by the GLM sensor are represented by white pixels set against a black background, as it can be seen in Fig. 2. In contrast to the CONUS images, which measure 1168×835 , all the images have been resized to 512×512 over the Gulf of Mexico. Spatial transformations such as down-sampling were employed for consistency in both data types. The lowest temporal resolution is selected, retaining brightness temperature every 5 minutes, and aggregating flashes for GLM sensor data in 5 minutes time steps. Furthermore, the final spatial resolution of the images is 4.5 km, which is a fine balance between a great resolution and an acceptable image size. The dataset encompasses data retrieved from 00:00 to 05:00 UTC, selected for specific days in January, February, and December, spanning the years 2020 to 2023. The overall database comprises 170 retrieved days, resulting in 10,200 pairs of ABI/GLM images and a total of 20,400 images. We choose to separate the dataset with 70% for the training phase, and 30% for the testing phase, separated by days. It means that each different date cannot be at the same time in both phases dataset.

III. METHOD

A. Model's sequence input

To train the model, a temporal sequence is used. Each brightness temperature image is paired with a corresponding flashes position image, constituting one input-label pair. To maintain temporal dependencies between images while ensuring efficient computational resource utilization, StrikeNet's input is designed as a sequence comprising 9 such pairs of images. In essence, it processes a 45 minutes sequence of images as input to forecast flashes occurring 5 minutes later as shown in Fig. 3. By learning to predict a future instance,

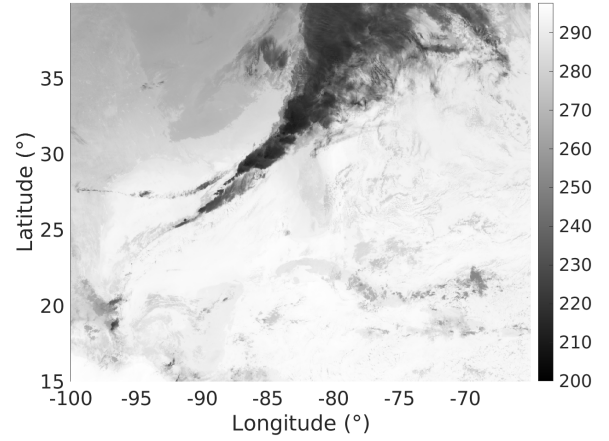


Figure 1. One sample of brightness temperature map acquired on 2023/13/1 at 02:01 UTC by the ABI sensor. Darker pixels have lower brightness temperatures and belong to higher top clouds. The colorbar is the brightness temperature in Kelvin.



Figure 2. Flashes location map acquired by the GLM sensor at the same date than in Fig. 1. White pixels identify the presence of flashes, and the background is represented with black pixels.

the algorithm must consider multiple moments from the past to preserve this temporal connection.

B. Model description

StrikeNet is a CNN designed for predicting the spatial distribution of future electrical activity. It draws inspiration from the encoder-decoder structure of U-Net [10], a neural network predominantly utilized for semantic segmentation, where each pixel in an output mask is associated with a specific label or class. The architecture of the StrikeNet model is illustrated in Fig. 4.

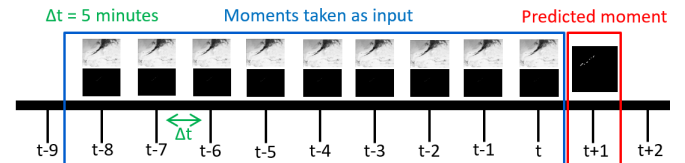


Figure 3. Sequential functioning of StrikeNet.

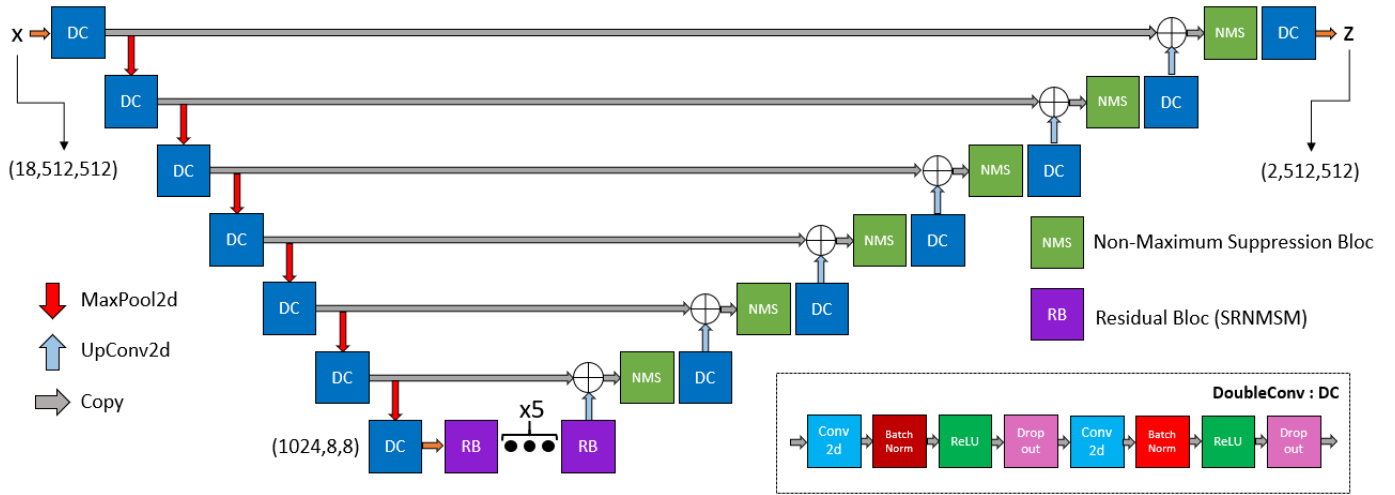


Figure 4. Architecture of StrikeNet.

On the one hand, the input sequence of images passes through an encoder. This consists of a sequence of two 3×3 convolutional layers, batch normalization layer, ReLU activation function, and then a 0.2 dropout. All of these define the DoubleConv (DC) block used in Fig. 4 and are followed by a max-pooling 3×3 layer. The process iterates six times, halving the image’s spatial resolution and doubling the number of channels with each step. Through it, images in the input sequence are reduced to a size up to 8×8 pixels thanks to the maxpooling layers, and the number of different channels grows up to 1024 features maximum in order to catch spatial and temporal dependencies over the input sequence.

Before going through the decoder, five residual blocks inspired by the Super Resolution Network using Multi-scale Spatial and Morphological features (SRNMSM) [5] are incorporated into the architecture. Compared to the original version, our block utilizes all its layers sequentially instead of in parallel to force morphological operations to be used rather than just offering them. The five blocks, placed between the encoder and decoder, enhance the network’s capacity to capture even finer image resolutions. By considering both morphological and spatial dependencies in the images, these blocks contribute significantly to the model’s ability to detect and predict lightning flashes, often represented by several single pixels in the image.

On the other hand, the decoder is composed of DC blocks followed by an upconvolution layer and soft NMS layer which is detailed in Subsection III-C. This is also repeated six times, and at the end, a 1×1 convolution layer is applied to produce the final mask representing the flashes’ location 5 minutes later. This mask consists entirely of black pixels, representing the background, while the white pixels delineate areas where lightning is predicted. These areas correspond to the two classes extracted in the final convolution.

C. Soft non-maximum suppression layers

These layers are placed in the decoder part and are inspired by the NMS technique explained in [6] and are used to obtain isolated pixels representing a lightning flash. The NMS layers retain local maxima by applying Eq. (1) in StrikeNet architecture:

$$x_{i,j} = 10x_{i,j} - 9 \max_{di,dj \in \{-1,0,1\}} x_{i+di,j+dj} \quad (1)$$

with $x_{i,j}$ representing the pixel we focus on, and $x_{i+di,j+dj}$ its 8 nearest neighbors. This allows us to keep the pixel value as the maximum value between him and its neighborhood. Using this type of layer enables the network to predict only the lightning when the probability is the highest in a 3×3 pixel area, thereby reducing the false alarms.

D. Model training

The training phase is launched using the Adam optimizer with a 10^{-4} learning rate. The standard Cross-Entropy loss function is chosen, but we also added a DiceLoss function with a coefficient of 0.1 to it in order to give greater importance to finding lightning than background. The algorithm adjusts the weights of the model to minimize the loss function by backpropagation of the loss gradient in order to find the most accurate class for each pixel on the output image to create an accurate prediction map of the flash positions.

IV. RESULTS

A. Evaluation metrics

StrikeNet has been tested on the flashes dataset, using different evaluation metrics. The first one is the confusion matrix which includes true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Using these quantities allows us to calculate the following metrics: Precision, Probability Of Detection (POD), F1-Score, False Alarm Rate (FAR), and Bias which are explained in Table I. The primary goal of this study is to make accurate predictions. Minimizing

missed lightning occurrences is critical, and keeping the FAR low is imperative to prevent unnecessary alerts.

In the testing phase, we establish an area of 9×9 pixels around each real lightning flash. For each predicted pixel, if the forecast is included in these areas, it is not counted as a false alarm. This means that the metrics are calculated within about 20km tolerance around actual flashes for all models which is something acceptable seeing the FAA standard recommendation to avoid thunderstorm areas by 37km [13].

B. Comparison with other methods

In Table II, we compared StrikeNet with various models, highlighting its significant outperformance over other models. All metrics were computed across all the test dataset. Each model underwent five training sessions, and assessments were conducted on these distinct trained models. The results were then averaged for tabulation.

Here, U-Net+dropout is the second-best model and it achieved an F1-Score of only 0.23, compared to 0.54 for StrikeNet. In addition, ED-DRAP [3], an encoder-decoder model that takes a sequence as input and uses spatial and temporal attention, also shows great performances with an F1-Score of 0.21. Furthermore, the remaining tested CNNs showed even lower scores between 0.10 and 0.17.

These results stem from the challenge of segmenting point-wise objects, where CNNs tend to produce blobs around strikes correlated with the corresponding cloud, deviating from the goal of segmenting the strike itself. For reference, the simple persistence model achieved an F1-Score of 0.43, surpassing all CNNs except for StrikeNet. While having the lowest POD, this model’s correct prediction of the punctual structure led to the lowest FAR, likely attributed to the relatively stable positions of flashes within 5-minute intervals.

Efficient prediction requires both improved POD and a low FAR, a feat challenging for deep networks, as demonstrated in Table II but effectively achieved by StrikeNet. Further discussions on the underlying mechanisms of StrikeNet’s success are provided in Subsection IV-D following graphical illustrations.

C. Graphical results

For better visualization of the results, we overlaid coasts, brightness temperature images, truth yellow dots, and red area predictions on a map. In Fig. 5 and Fig. 6, we compared graphical results obtained at the same moment and day, using a simple U-Net model and StrikeNet.

Fig. 5 illustrates that forecasted areas are significantly larger than the actual locations of lightning flashes when using classical U-Net. While this leads to a high POD, it also results in an excessively large FAR.

Fig. 6 shows graphical outcomes using StrikeNet. Lightning forecast areas are smaller and align more closely with actual flash locations. This results in a 1.5 times reduction in FAR and a threefold increase in the F1-Score. While this aligns with expectations, it is important to note that the POD has decreased as forecasted areas became less extensive.

D. Discussion

Our focus was on disentangling the components of StrikeNet, and our findings from Table II indicate that the model’s remarkable efficiency stems from the combination of morphological blocks, dropout, and soft NMS layers rather than from these elements taken individually.

Remarkably, within the limited dataset size, all deep networks (except DeepLab [4], a model using dilated convolutions and atrous spatial pyramid pooling) accurately predict isolated strikes on training data. However, this delicate balance, prone to limited generalization as documented in [8], is fortified by introducing dropout and morphological operations. For DeepLab, the challenge deepens due to low-resolution map architectures, making them almost incapable of predicting isolated strikes after rescaling.

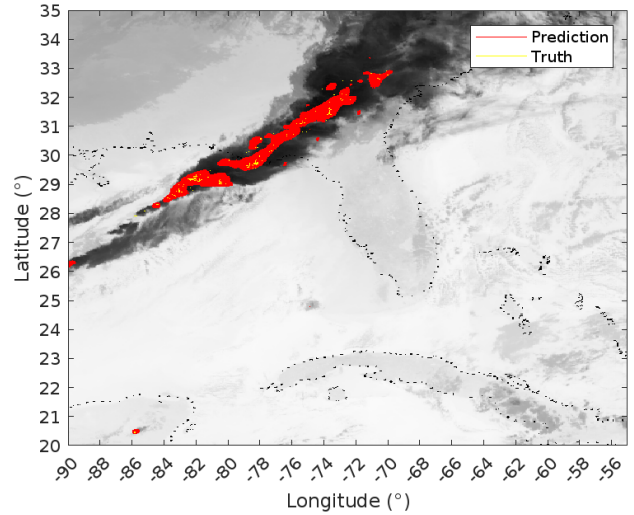


Figure 5. Graphical results of the forecast (red areas) compared to the truth (yellow pixels) with U-Net model for the 2023/13/1 at 02:06 UTC.

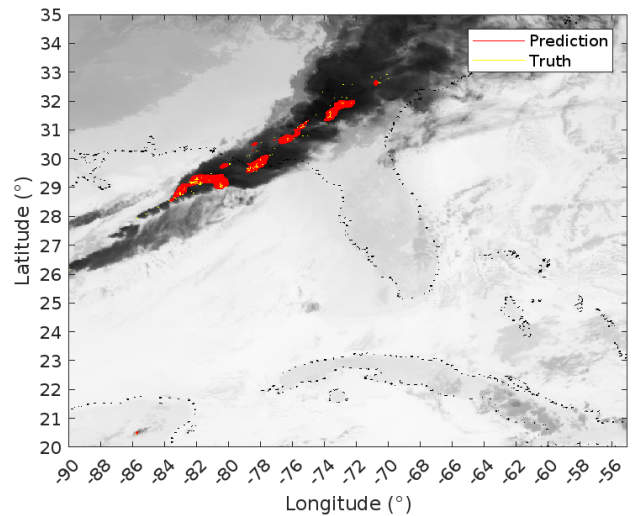


Figure 6. Graphical results of the forecast (red) compared to the truth (yellow) with StrikeNet model at the same date as Fig. 5.

Table I
DETAILED EVALUATION METRICS

Metrics	Equation	Explanations
Precision	$Precision = \frac{TP}{TP+FP}$	Ratio of the well-predicted flashes to the number of predicted flashes. Higher as possible.
POD	$POD = \frac{TP}{TP+FN}$	Ratio of well-predicted flashes to the real number of flashes. Higher as possible.
F1-Score	$F1 - Score = 2 \frac{POD * precision}{POD + precision}$	Harmonic mean between precision and POD. Higher as possible.
FAR	$FAR = \frac{FP}{FP+TP}$	Ratio between the number of false flashes to the number of predicted flashes. Lower as possible.
Bias	$Bias = \frac{TP+FP}{TP+FN}$	Ratio to determine the under or overestimation of the flashes by the network. Nearer to 1 as possible. If $bias < 1$, underestimates, if $bias > 1$, overestimates.

Table II
PERFORMANCES COMPARISON BETWEEN DIFFERENT MODELS

Models	Evaluation metrics				
	Precision	POD	F1-Score	FAR	Bias
Persistence	0.83	0.29	0.43	0.17	0.36
DeepLab [4]	0.06	0.82	0.11	0.94	12.55
U-Net [10]	0.095	0.95	0.17	0.91	16.65
U-Net+dropout	0.13	0.90	0.23	0.87	7.02
U-Net+soft NMS	0.053	0.96	0.10	0.95	23.65
U-Net+SRNMSM	0.05	0.96	0.10	0.95	24.60
U-Net+ 5xS-RNMSM	0.09	0.94	0.16	0.91	14.11
EDDRAP [3]	0.12	0.94	0.21	0.88	10.02
StrikeNet	0.42	0.78	0.54	0.58	1.96

V. CONCLUSION

This paper presents the StrikeNet neural network model, specifically designed for generating maps predicting the short-term location of electrical activity. The model incorporates two types of data, namely brightness temperature maps and flash position maps captured simultaneously, and takes these data in a temporal sequence as input.

The study showcases the adaptability of neural networks for predicting punctual objects in images, leveraging NMS techniques, residual blocks utilizing morphological features of the images, and their integration with dropout layers. StrikeNet yielded compelling outcomes, achieving a lower FAR of approximately 58%, a POD of 78%, and a precision close to 40%, resulting in an F1-Score of 54%. These findings, in comparison to the utilization of more traditional semantic segmentation models, mark a significant advancement for this category of meteorological data.

Future objectives encompass the production of extended forecasts, spanning up to an hour rather than the next 5 minutes. Additionally, we plan to generate forecasts incorporating risk percentage zones to enhance precision.

ACKNOWLEDGMENT

For the GOES-R data, we thank the NOAA National Geophysical Data Center.

REFERENCES

- [1] Aurélie Bouchard, Magalie Buguet, Adrien Chan-Hon-Tong, Jean Dezert, and Philippe Lalonde. Comparison of different forecasting tools for short-range lightning strike risk assessment. *Natural Hazards*, 115:1011–1047, 2022.
- [2] Vincent Bouget, Dominique Béréziat, Julien Brajard, Anastase Charantonis, and Arthur Filoche. Fusion of rain radar images and wind forecasts in a deep learning model applied to rain nowcasting. *Remote Sensing*, 13(2), 2021.
- [3] Hongshu Che, Dan Niu, Zengliang Zang, Yichao Cao, and Xisong Chen. ED-DRAP: Encoder–decoder deep residual attention prediction network for radar echoes. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022.
- [4] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *CoRR*, abs/1706.05587, 2017.
- [5] Alireza Esmailzahi, M. Omair Ahmad, and M. N. S. Swamy. Srmsm: A deep light-weight image super resolution network using multi-scale spatial and morphological feature generating residual blocks. *IEEE Transactions on Broadcasting*, 68(1):58–68, 2022.
- [6] Jan Hosang, Rodrigo Benenson, and Bernt Schiele. Learning non-maximum suppression. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [7] Jussi Leinonen, Ulrich Hamann, and Urs Germann. Seamless lightning nowcasting with recurrent-convolutional deep learning. *Artificial Intelligence for the Earth Systems*, 1(4):e220043, 2022.
- [8] Yuanzhi Li, Colin Wei, and Tengyu Ma. Towards explaining the regularization effect of initial large learning rate in training neural networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [9] Stéphane Pédebois, Paul Barneoud, and Claude Berthet. First results on severe storms prediction based on the french national lightning locating system. *International Lightning Detection Conference, San Diego, California, USA*, 2016.
- [10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.
- [11] Timothy J. Schmit, Paul Griffith, Mathew M. Gunshor, Jaime M. Daniels, Steven J. Goodman, and William J. Lebar. A closer look at the abi on the goes-r series. *Bulletin of the American Meteorological Society*, 98(4):681 – 698, 2017.
- [12] Peter Ukkonen, Agostino Manzato, and Antti Mäkelä. Evaluation of thunderstorm predictors for finland using reanalyses and neural networks. *Journal of Applied Meteorology and Climatology*, 56(8):2335 – 2352, 2017.
- [13] Federal Aviation Administration (FAA) U.S. Department of Transportation. Advisory circular (ac) no 00-24c. 2/19/13.
- [14] Kanghui Zhou, Yongguang Zheng, Wansheng Dong, and Tingbo Wang. A deep learning network for cloud-to-ground lightning nowcasting with multisource data. *Journal of Atmospheric and Oceanic Technology*, 37(5):927 – 942, 2020.