Predictive Analysis of Cloud Motion with Convolutional and Recurrent Neural Networks

Preeti Venkataraman Hegde
Responsible for Front End Development, maintained documentation and wrote section 1

Siva Sai Kumar Reddy Kachana
Responsible for Motion Estimation Model Development and wrote sections 2, 4 and References

Tharun Kumar Korine Palli
Responsible for Segmentation Model Development and wrote Abstract, section 3 and Acknowledgements

ABSTRACT

This research paper introduces a classical approach for segmenting and predicting cloud patterns, addressing a significant gap in traditional cloud forecasting methods which often suffer from limited accuracy and temporal resolution. We employ the usage of cloud images from fish-eye camera, addressing the challenge of large-scale cloud behavior analysis. The raw images, characterized by their irregularity, are processed using a Convolutional Neural Network (CNN) for precise cloud segmentation. This segmentation is crucial for the accurate extraction and identification of cloud regions from the vast sky imagery. After Segmenting the cloud images, We proposed a deep learning model, integrating a time-distributed architecture combining CNN with Long Short-Term Memory (LSTM) layers. This model is trained on an extensive dataset of cloud images, taken at one-minute intervals, enabling it to learn and predict cloud behavior patterns over time. The model's unique ability to process sequences of images allows it to predict future cloud formations, a step forward in dynamic cloud pattern forecasting. The paper outlines the process of Manual data labeling, Segmentation model training, cloud motion estimation, and evaluation. Our comprehensive approach demonstrates significant improvements in cloud segmentation and prediction accuracy, making help in the fields of renewable energy generation and meteorological research.

KEYWORDS

Cloud forecasting, Deep Learning, Convolutional Neural Network, Long short-term memory, Segmentation, Fish-Eye camera.

1 INTRODUCTION

The research presented in this paper primarily focuses on the development of a methodology for cloud image segmentation and forecasting. The catalyst for this study is the significant impact of cloud movements in various scientific and practical domains, including weather forecasting and solar energy estimation. Traditional methods in cloud forecasting have been hindered by limitations in accuracy and temporal resolution, leading to the pursuit of more sophisticated and reliable techniques.

The problem addressed in this study is identifying clouds and clear-sky regions due to distortion present in cloud images captured by fisheye camera, chosen for their wide field of view, enabling the observation of cloud formations over extensive geographic regions. The images captured by this camera encompass distortion, rendering them unsuitable for immediate use in Predictive analysis. To address this challenge, our approach incorporates the use of a convolutional neural network (CNN) [1] for cloud image segmentation. This technique is pivotal in accurately identifying and extracting cloud formations from the raw, wide-view imagery.

The primary objective of our study is to develop a robust AI model that can accurately segment clouds from images. This involves a meticulous process beginning with the collection of a vast dataset of cloud images from fisheye cameras [10]. Each image is carefully labeled manually, forming the foundation of our training dataset for the CNN. This labeling is critical in ensuring that the CNN can effectively differentiate between clouds and clear skies, a crucial step in achieving precise cloud segmentation.

Once the CNN model demonstrates high accuracy in cloud segmentation, our secondary objective is to predict cloud motion. We integrate it into a more comprehensive predictive model that utilizes a time-distributed architecture, combining the spatial analytical power of CNNs with the temporal processing capabilities of LSTM layers [9]. The model is trained on sequences of images, allowing it to understand and anticipate cloud movements based on past patterns. This aspect of time series analysis is integral to our predictive model, as it provides the ability to forecast cloud positions and movements with greater precision than previous methods.

Overall, this paper details a comprehensive methodology for cloud image segmentation and forecasting, utilizing advanced AI modeling techniques. Our approach is designed to address the existing limitations in traditional cloud segmentation techniques and provides an approach for predicting cloud patterns. By integrating sophisticated convolutional neural networks and time-distributed architectures, this approach helps in enhanced capabilities for weather forecasting and solar energy system optimization.

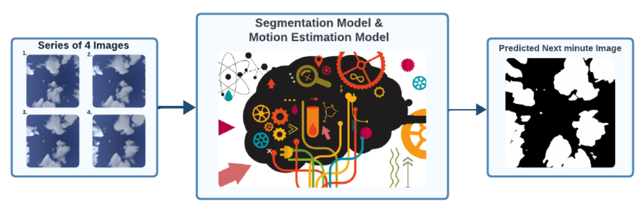

Figure 1: Schematic of the Predictive Analysis Pipeline Using Segmentation and Motion Estimation Models.

2 METHODOLOGY

This section details our methodological approach, encompassing the data sourcing, image segmentation model development, and cloud motion estimation. We begin by describing the datasets characteristics and preparation, crucial for the subsequent analysis. Following this, the construction and functionality of both the segmentation and cloud motion estimation models are elaborated, highlighting their roles in this research.

2.1 Data

The Dataset employed for this study consists of two categories : ASI (Atmospheric State Interference) [8] images and LES (Large Eddy Simulation) [4]. The LES run images are a sequential set of images that are created by 6h run with 1 min difference. ASI acquired images are used for training the cloud segmentation model to identify the cloud regions in the images. The actual ASI photos, required for the training of a deep learning model optimized towards cloud segmentation, were rigorously obtained from the foundational study undertaken by Philip Gregor [11]. These real world ASI images are used for cloud segmentation model because the images in the LES run is less and training such model might overfit the data. The real ASI images are of size 1024 x 1024 pixels but for training a deep learning model we have reduced the dimension to 256 x 256 pixels.

The LES dataset comprises 360 images taken at 60-second intervals during a 6-hour simulation using UCLALES. The data is publicly available here [5]. The images start with roughly 0% cloud fraction in the start of the simulation to 100% cloud fraction at the end of 6h run. This dataset allows for detailed analysis of the cloud motion estimation and allows for comparison with other related works in the field. In this LES dataset of images , the position of the Sun is fixed at a zenith angle of 30° to the south. The simulated images are projected onto a horizontal ground plane for seamless integration. During projection, images are resized to 280 x 280 pixels for efficient computation and reduced dimensionality, potentially resulting in some feature blurring. These datasets form the foundation of our study for estimating cloud motion from t-1 min to t min.

2.2 Cloud Segmentation

The first step in our study is to create an image segmentation model for our images. This model will be used to segment clouds, which will serve as a foundation for estimating future cloud movements at time t+n. The primary and significant aspect required for the cloud motion model is the ability to differentiate between clouds and clear sky regions within the images. Effective identification of these areas is crucial for providing accurate results by cloud motion model. The Real ASI images are first manually segmented using a specific tool built especially to label these images by using the Superpixels [5] technique. This tool helps the individual in labeling these images as either cloudy , clearsky or undecided. The individual is able to classify the images into different regions by either increasing or decreasing the superpixels for effective segmentation, even for areas with small clouds. By using this tool we have created a database of 1167 images with their corresponding segmentation masks. This usage of superpixels allows faster segmentation process based on the neighboring pixels. The labeling tool also has an option to select the number of superpixels in the images thus, making it easier to label very small cloud image precisely. The Segmentation of the clouds is needed in order to effectively estimate the cloud motion. The Precise segmentation of images into their clear and cloud regions allows for tracking individual cloud formations and characterizing their direction, speed and other motion characteristics in the sequence of images.

We utilize Convolutional Neural Networks for generating cloud masks in the cloud motion model images, as CNNs [2] [3] are widely employed for segmentation tasks. The CNN model is made to train on these ASI images ignoring the undecided class firstly, due to the fact that the images of LES run contains only 2 class either clear sky or cloudy, and excluding this class would help the model to understand better and secondly, since even humans sometimes find it challenging to classify certain areas as either cloudy or clear sky, especially when clouds are scattered throughout the sky, creating the impression of both cloudy and clear sky. The data is split into training and validation sets of 817 and 350 images respectively to train the network. The CNN network is trained with 50 epochs which uses the DeepLabV3+ [1] architecture which consists of the Encoder and Decoder network as described in the U-Net architecture model. The Encoder used for this study is resnet34 [14] pre-trained on the imagenet [15] dataset. The Adam [7] Optimizer has been used in this study along with the batch size of 6 images.

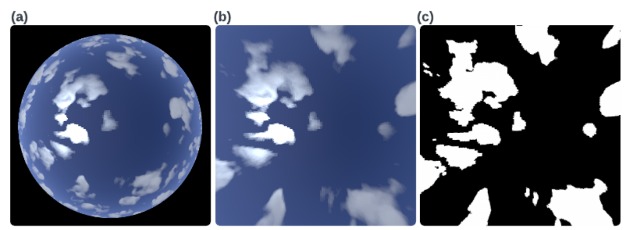

Figure 2: (a) LES cloud output field from fish-eye camera,(b) Projected Image of a, (c) CNN-segmented image of b

2.3 Cloud Motion Estimation

After training the model for segmentation and obtaining segmented masks, the subsequent phase involves developing a model that is capable of utilizing segmented images to forecast cloud motion. Cloud motion prediction refers to the process of forecasting the movement of clouds over a period. This forecasting primarily depends on sequential images of clouds, allowing algorithms to analyze the temporal evolution of cloud patterns and accurately track their movement over time. By harnessing the sequential information present in consecutive images, the algorithm differentiates subtle alterations in cloud positions and shapes, enabling precise estimation of cloud motion dynamics.

By leveraging the sequential information encoded in consecutive images, the algorithm learns subtle changes in cloud positions and shapes, enabling precise estimation of cloud motion dynamics. Various techniques, such as Optical Flow and Feature Tracking, are available for cloud motion estimation. These methods have limitations that render them unsuitable for real-world applications and may compromise accuracy in estimation. Optical flow calculates motion vectors between consecutive images based on edge detection. However, it assumes consistent pixel intensities between frames, which isn't always valid in real-world scenarios. Feature tracking, on the other hand, identifies distinctive features and tracks their movement over time, yet struggles with rapid or nonlinear cloud motion, leading to unreliable estimates. Consequently, these traditional methods, including Optical Flow, may not be ideal for prolonged motion estimation, requiring modern techniques for improved accuracy and reliability.

These days deep learning approaches [12] [13] showed significant improvement over the standard algorithms to estimate cloud motion. Their ability to automatically learn high level features in the data makes them very robust to be used in the cloud images as clouds show very high variability in seconds of time, also including changes in the lighting conditions and the structure of the clouds.

In this Study , we are training a deep learning model consisting of CNN and LSTM. The utilization of CNNs renders them exceptionally well-suited for extracting intricate cloud features from images, making them an ideal choice for this particular task. The LSTM network helps in understanding the temporal difference between the frames and thus helps in the motion estimation process.

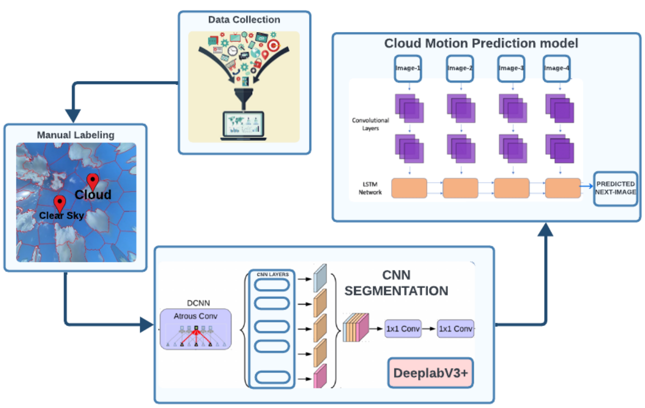

Figure 3: Workflow of the Cloud Motion Prediction System Incorporating Manual Labeling, CNN Segmentation, and Time Distributed Prediction Model

The model we have constructed leverages four consecutive images from a sequence of cloud data to predict the fifth image in the sequence. Our proposed architecture for cloud motion estimation meticulously integrates spatial and temporal information inherent in sequential cloud imagery. This design integration is made possible through the combined utilization of CNN and LSTM networks. By employing this hybrid model, we effectively capture spatial features from individual frames while preserving memory of temporal dependencies across sequential cloud images. These spatial features, meticulously extracted by the CNN, are then seamlessly integrated into LSTM layers. The LSTM network, augmented with TimeDistributed Layers, processes each frame of sequential cloud images individually while preserving the temporal sequence of information. During training, our model is fed sequences of four images as input, with the subsequent image serving as the target prediction. Training occurs over 50 epochs, optimizing the process with Mean Squared Error (MSE) loss. Through experimentation, we found that input sequences of size four yielded the highest accuracy, thus adopting this configuration for our model.

3 EVALUATION

The evaluation of our methodology focused on two main aspects: the accuracy and efficiencies of cloud image segmentation and cloud motion prediction. These evaluations are crucial to demonstrate the viability of our approach in practical applications.

3.1 Cloud Segmentation

Our methodology employs Cross Entropy Loss for training the segmentation model. This section details the metrics used for evaluating the segmentation accuracy.

Loss Function: We used Cross Entropy loss during the training cloud segmentation model. It measures the performance of the classification of cloud and clear-sky whose output is a probability value between 0 and 1. The cross entropy loss increases as the predicted probability diverges from the actual label. The equation for cross entropy loss is:

$$H(y, \hat{y}) = -\sum_i y_i \log(\hat{y}_i) \tag{1}$$

Where,

- $H(y, \hat{y})$ represents the cross-entropy loss.
- y_i is the true label for class i.
- $\hat{y}_i$ is the predicted probability for class i.

Metrics: We used 2 metrics named Intersection over Union (IoU) and Dice Coefficient to evaluate the accuracy of our segmented images.

Intersection over Union (IoU): IoU is a common metric for the evaluation of image segmentation models. It calculates the overlap between the predicted segmentation and the ground truth. A higher IoU score indicates better segmentation performance. The formula for IoU is:

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{2}$$

Where,

- Area of Overlap is the area of overlap between the segmented cloud image and the ground truth of the same cloud image.
- Area of Union is the area of the union of the segmented cloud image and the ground truth of the same cloud image.

Dice Coefficient: Also known as the Sørensen–Dice coefficient, this is a statistical tool that measures the similarity between two Images. For binary classification, it is especially useful. The Dice Coefficient is calculated as:

$$\text{Dice Coefficient} = \frac{2 \times |I_{\text{seg}} \cap I_{\text{gt}}|}{|I_{\text{seg}}| + |I_{\text{gt}}|} \tag{3}$$

Where,

- I_{seg} represents the flattened segmented cloud image.
- I_{gt} represents the flattened ground truth image of the same cloud.
- $|I_{\text{seg}} \cap I_{\text{gt}}|$ is the area of overlap between the flattened segmented image and the flattened ground truth.
- $|I_{\text{seg}}|$ and $|I_{\text{gt}}|$ are the total areas of the flattened segmented and flattened ground truth images, respectively.

3.2 Cloud Motion Estimation

In the evaluation of our cloud motion prediction model, we used a time-distributed model trained with Mean Square Error (MSE)

Loss, and we assessed its performance using both the Structural Similarity Index (SSIM) and the Dice Coefficient.

Loss Function: MSE is used as the loss function while training time distributed model. It measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value. The equation for MSE is:

$$\text{MSE Loss} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{4}$$

Where,

- y_i is the true value for the i-th observation.
- $\hat{y}_i$ is the predicted value for the i-th observation.
- N is the total number of observations.

Metrics: We calculated mean SSIM [12] and mean Dice Coefficient for 40 set of sequential test images.

Structural Similarity Index (SSIM): This metric is used to assess the similarity between the actual image and the predicted image. SSIM considers changes in texture, luminance, and contrast, making it ideal for evaluating the quality of images.

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{5}$$

Where,

- x is the actual cloud image, and y is the predicted cloud image.
- μ_x, μ_y are the average pixel values of images x and y respectively.
- σ_x^2, σ_y^2 are the variance of images x and y respectively.
- σ_{xy} is the covariance of images x and y.
- C_1 and C_2 are constants to stabilize the division with weak denominator.

Dice Coefficient: The Dice Coefficient, typically used in segmentation task, we also employed in our study to evaluate the cloud motion prediction model. It provides a measure of similarity between the predicted cloud motion images and the actual subsequent images.

4 RESULTS AND CONCLUSION

This section describes the outcomes of our CNN-based approaches for image segmentation and cloud motion estimation, setting the stage for subsequent conclusions and avenues for future research.

CNN Segmentation (Mean)		Cloud Motion Estimation (Mean)	
IoU	Dice Coefficient	SSIM	Dice Coefficient
0.88	0.93	0.836	0.94

Table 1: Summary of Performance Scores for CNN-Based Segmentation and Motion Estimation Techniques.

In our study, the CNN model for image segmentation achieved an Intersection over Union (IoU) mean score of 0.88, suggesting a high overlap between the model's predictions and the ground truth. The Dice Coefficient further confirmed the model's accuracy with a score of 0.93 as mentioned in Table 1. For cloud motion estimation, the model's performance was evaluated using the Structural Similarity Index Measure (SSIM) with a mean score of 0.836, indicating a strong resemblance between predicted and actual cloud patterns, complemented by a Dice Coefficient of 0.94, affirming the model's efficiency in capturing motion patterns.

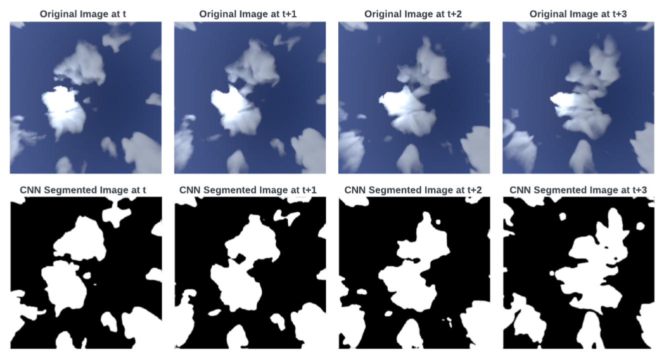

Figure 4: Series of 4 Original images and their CNN segmentated images

Figure 4 showcases a CNN segmentation model's adeptness in precisely isolating cloud formations from a series of cloud Images, showing a strong alignment with the actual cloud contours across time.

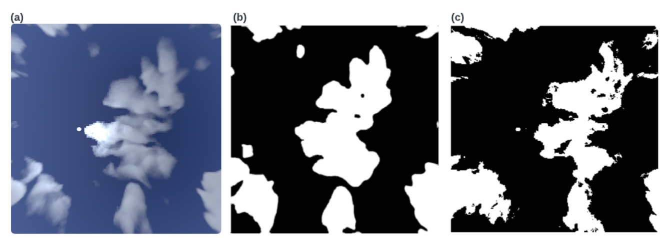

Figure 5: (a) original Projected Image at t+4, (b) CNN segmented Image at t+4, (c) Cloud Motion prediction at t+4

Figure 5 represents the output from a cloud prediction model, demonstrating strong predictive accuracy as **(c)** exhibits a forecasted cloud distribution that maintains consistency with **(b)** segmented cloud structures.

In conclusion, we defined a novel technique to automatically segment the cloud images and estimate the cloud motion within a 5 min time period. A CNN was trained using cloud images as inputs, and a deep learning based CNN-LSTM for forecasting future cloud motion was performed. By precisely segmenting cloud images and employing a hybrid CNN-LSTM model, we were able to capture both spatial and temporal details in cloud dynamics. This architecture accurately analyzes and predicts cloud movements by understanding complex properties and time patterns. The Limitation for this study is the dataset from the LES run which is just 360 images for 1 min apart and the constant zenith angle of the sun. By acquiring more images in sequence we can train our estimation model better and having an extensive validation set would enable us to provide more accurate estimations on our model further. Next steps further in our study would include the usage of the real world images to estimate the cloud motion with different positions of the sun and dynamics of the real world cloud images.

REFERENCES

[1] Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H.: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11211 LNCS, 833–851, https://doi.org/10.48550/arXiv.1802.02611, 2018.

[2] Ye, L., Cao, Z., and Xiao, Y.: DeepCloud: Ground-Based Cloud Image Categorization Using Deep Convolutional Features, IEEE Trans. Geosci. Remote Sens., 55, 5729–5740, https://doi.org/10.1109/TGRS.2017.2712809, 2017.

[3] Xie, W., Liu, D., Yang, M., Chen, S., Wang, B., Wang, Z., Xia, Y., Liu, Y., Wang, Y., and Zhang, C.: SegCloud: a novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation, Atmos. Meas. Tech., 13, 1953–1961, https://doi.org/10.5194/amt-13-1953-2020, 2020.

[4] Stevens, B., Moeng, C.-H., Ackerman, A. S., Bretherton, C. S., Chlond, A., de Roode, S., Edwards, J., Golaz, J.-C., Jiang, H., Khairoutdinov, M., Kirkpatrick, M. P., Lewellen, D. C., Lock, A., Müller, F., Stevens, D. E., Whelan, E., and Zhu, P.: Evaluation of Large-Eddy Simulations via Observations of Nocturnal Marine Stratocumulus, Mon.Weather Rev., 133, 1443–1462, https://doi.org/10.1175/MWR2930.1, 2005.

[5] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S.: SLIC Superpixels Compared to State-of-the-Art Superpixel Methods, IEEE T. Pattern. Anal., 34, 2274–2282, https://doi.org/10.1109/TPAMI.2012.120, 2012.

[6] Frydrych, J., & Pohl, G. (2022). UCLA-LES shallow cumulus dataset with 3D cloud output data [Data set]. LMU Munich, Faculty of Physics. https://doi.org/10.57970/5d0k9-q2n86

[7] Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, 3rd International Conference on Learning Representations, ICLR 2015 – Conference Track Proceedings, 1–15, http://arxiv.org/abs/1412.6980 (last access: 22 June 2023), 2014.

[8] Schmidt, T., Kalisch, J., Lorenz, E., and Heinemann, D.: Evaluating the spatio-temporal performance of sky-imager-based solar irradiance analysis and forecasts, Atmos. Chem. Phys., 16, 3399–3412, https://doi.org/10.5194/acp-16-3399-2016, 2016.

[9] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. Neural Comput. 9, 8 (November 15, 1997), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[10] Beekmans, C., Schneider, J., Läbe, T., Lennefer, M., Stachniss, C., and Simmer, C.: Cloud photogrammetry with dense stereo for fisheye cameras, Atmos. Chem. Phys., 16, 14231–14248, https://doi.org/10.5194/acp-16-14231-2016, 2016.

[11] Gregor, P., Zinner, T., Jakub, F., and Mayer, B.: Validation of a camera-based intra-hour irradiance nowcasting model using synthetic cloud data, Atmos. Meas. Tech., 16, 3257–3271, https://doi.org/10.5194/amt-16-3257-2023, 2023

[12] Su, X.; Li, T.; An, C.; Wang, G. Prediction of Short-Time Cloud Motion Using a Deep-Learning Model. *Atmosphere* **2020**, *11*, 1151. https://doi.org/10.3390/atmos11111151,2020

[13] A. H. Eşlik, E. Akarslan and F. O. Hocaoğlu, "Cloud Motion Estimation with ANN for Solar Radiation Forecasting," *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, Taiz, Yemen, 2021, pp. 1-5, https://doi.org/10.1109/ICOTEN52080.2021.9493523

[14] He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2016-December, 770–778, https://doi.org/10.1109/CVPR.2016.90,2015.

[15] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. and Berg, A.C., 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision*, *115*, pp.211-252. ,http://arxiv.org/abs/1409.0575, 2014.

[16] GitHub: A comprehensive repository containing code and resources for Predictive Analysis of Cloud Motion with Convolutional and Recurrent Neural Networks, https://github.com/tharun-kumar-22/Cloud-Segmentation-and-Forecasting

ACKNOWLEDGEMENTS

- We would like to express our deepest gratitude to **Prof. Magda Gregorova** for her invaluable guidance, supervision, and unwavering support throughout the duration of this project. Her expertise and insights have been instrumental in steering our work towards its successful completion.
- Our sincere thanks go to **Philip Gregor** for providing us with the critical dataset and segmentation tool. His immense patience, unconditional cooperation, and timely assistance have greatly contributed to our research efforts.
- We are also grateful to **David Benavente Ríos** for his constructive suggestions and expertise.
- Our project documentation could not have achieved its clarity and thoroughness without the support of **Confluence.** We appreciate its role in facilitating our project organization and documentation processes.
- For the efficient scheduling, creation of Gantt chart, and overall project management, **Monday.com** has been an essential tool. We acknowledge its contribution to keeping our project on track and well-organized.
- We extend our thanks to **Lucidchart** for enabling us to visualize our processes through flowcharts and process diagrams. Its intuitive design tools have been vital in communicating complex concepts clearly and effectively.
- Our gratitude also goes to **Canva** for its user-friendly platform that assisted us in designing impactful and visually appealing posters. Canva's tools and templates have greatly enhanced the presentation of our research findings.
- We also acknowledge the usage of **ChatGPT 3.5** for text optimization and **QuillBot** for grammatical checking, both of which were crucial in improving the quality and readability of our paper
- Lastly, we are thankful for **Overleaf** for streamlining the process of writing and formatting our paper in LaTeX. Its collaborative features and comprehensive LaTeX environment have been indispensable in producing a professional and well-structured document.

CONTACT INFORMATION

(1) **Siva Sai Kumar Reddy Kachana**
MAI Student at THWS
`sivasaikumarreddy.kachana@study.thws.de`

(2) **Tharun Kumar Korine Palli**
MAI Student at THWS
`tharunkumar.korinepalli@study.thws.de`

(3) **Preeti Venkataraman Hegde**
MAI Student at THWS
`preetivenkataraman.hegde@study.thws.de`