A study on using generative deep learning architectures for classification of ground-based cloud images

Student: Ștefan Alexandrescu

June 6, 2025

Outline

1. Introduction
2. Motivation
3. Related work
4. Research contribution
   - Proposal
   - Diffusion model
   - Vision Transformer
   - Case study on real data
5. Discussion
6. Conclusions and future work

Introduction

Context and importance

- Weather and climate change play a crucial role in society.
- Extreme weather events, predominantly caused by climate change, affect all everyday activities.
- Weather forecasting is essential for preventing large-scale disasters.
- Identification and prediction of severe weather phenomena is difficult.

Introduction (2)

Clouds classification

- There are various classifications for the clouds.
- The standard set is considered to be the WMO's [WMO]:
  - High clouds: **Cirrus** (Ci), **Cirrocumulus** (Cc), **Cirrostratus** (Cs);
  - Middle clouds: **Altocumulus** (Ac), **Altostratus** (As), **Nimbostratus** (Ns);
  - Low clouds: **Stratocumulus** (Sc), **Stratus** (St), **Cumulus** (Cu), **Cumulonimbus** (Cb).

Introduction (3)

Figure: Cumulus clouds [WMO]

Motivation

- Clouds display many characteristics (e.g., colour, shape) which are not easily quantifiable.
- **Cloud identification** represents one of the main activities of the weather stations personnel.
- Much data collected by satellites and instruments can be processed automatically.
- A relatively new topic in AI with much potential. *Machine Learning* proves to be a promising domain.

Motivation (2)

Figure: A tropical cyclone [WMO]

Related work

Details

- There is a progression from traditional image processing techniques toward more complex neural models.
- The approaches vary:
    - **supervised**
    - **unsupervised**
    - **self-supervised**
- **Public datasets involved:** Cirrus Cumulus Stratus Nimbus (CCSN) [Liu19], Ground-based Cloud Dataset (GCD) [LDZ+22], SWIMCAT, MODIS, LSCIDMR.

Related work (2)

- **Zhang et al. [ZLZS18]** introduced **CloudNet**, a CNN inspired by AlexNet, trained on the CCSN dataset with 11 cloud classes. Achieved 88% accuracy (CCSN) and 98.6% (SWIMCAT).
- **Mihuleț et al. [MC25]** proposed **X-Cloud** and **M-Cloud**, lightweight CNNs based on Xception, trained on CCSN and GCD datasets. Achieved 75% average accuracy with high efficiency.
- **Luo et al. [LPS+24]** developed a YOLOv8-based pipeline with haze-reduction preprocessing and image segmentation. Trained on 4000-image Tibet dataset, achieving near-perfect training accuracy.

Related work (3)

- **Yousaf et al. [YRK+23]** used over 100k satellite images (LSCIDMR), proposing a CNN-ResNet model inspired by VGGNet. Achieved 97.25% accuracy using GAP and softmax classification.
- **Togaçar et al. [TE22a]** implemented a mobile-friendly ShuffleNet with super-resolution, semantic segmentation, and SFO feature selection. Best accuracy: 98.56% using LDA.
- **Dev et al. [DLW15]** proposed a texton-based model using S-filters and k-means clustering, tested on SWIMCAT. Achieved near-perfect classification, struggling slightly with veil clouds.

Related work (4)

- **Vasylieva et al. [VM22]** trained a CNN on NOAA-20 satellite imagery for 4-class cloud classification. Achieved 95% training and 85% validation accuracy.
- **Geiss et al. [GCV+24]** proposed a self-supervised SNN with Barlow Twins loss. Trained on MODIS and ABI datasets, outperforming supervised models with accuracy surpassing 80%.
- **Kurihana et al. [KKF+19]** used a deep CAE with advanced loss metrics and HAC clustering on MODIS data. Showed good unsupervised class separation and high AMI stability scores.

GCD dataset peek

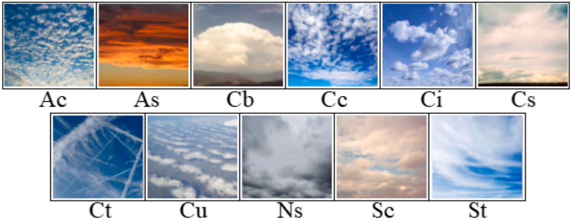


Figure: Sample images extracted from CCSN dataset [TE22b]

Proposal

- Proposed ideas reside at the intersection of generative models and transformer-based architectures.
- Enhancing the performance of cloud classification from images using a **Diffusion (Classifier)** and a **Vision Transformer** (ViT), both of which provide notable benefits:
  - contextual reasoning using mechanisms such as *attention*;
  - generative capabilities that enhance *generalization* and analyse small details.

Diffusion classifier (1)

- Learn a bijection between original domain and pure noise space.

- Consists of two steps: forward and **reverse diffusion** processes for altering an image.
  - the formulation of the forward process function:

$$q(x_t|x_{t-1}) = N(x_t, \sqrt{1-\beta_t}x_{t-1}, \beta_t I) \quad (1)$$

- x_{t-1} and x_t are the input and output at step t
- N is the normal distribution
- β_t is the noise degree scheduler.

Diffusion classifier (2)

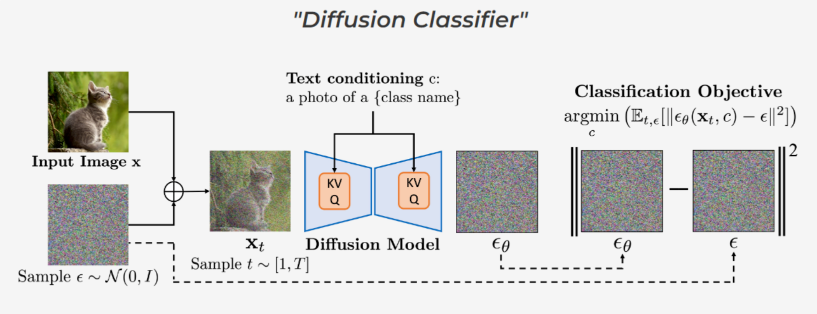
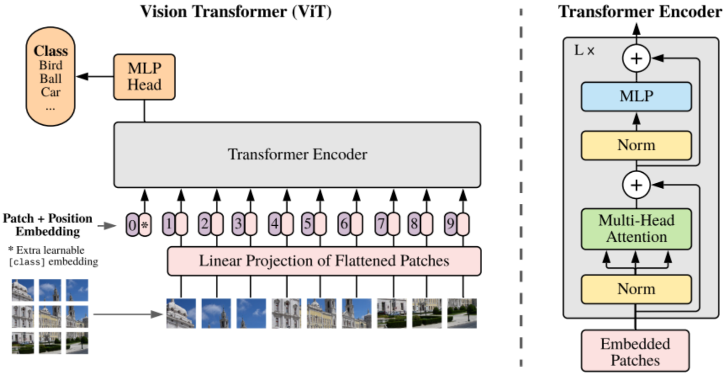


Figure: An overview of the diffusion classifier model [LPD+23]

Vision Transformer (1)

- Complex type of Deep Neural Networks (DNN) that rely on the mechanism of attention.
- Adjusted for image processing, as opposed to regular transformers.
- Techniques involved:
  - Patch Embeddings
  - Positional Embedding
  - Multi-Head Self-Attention

Vision Transformer (2)



Case study on real data

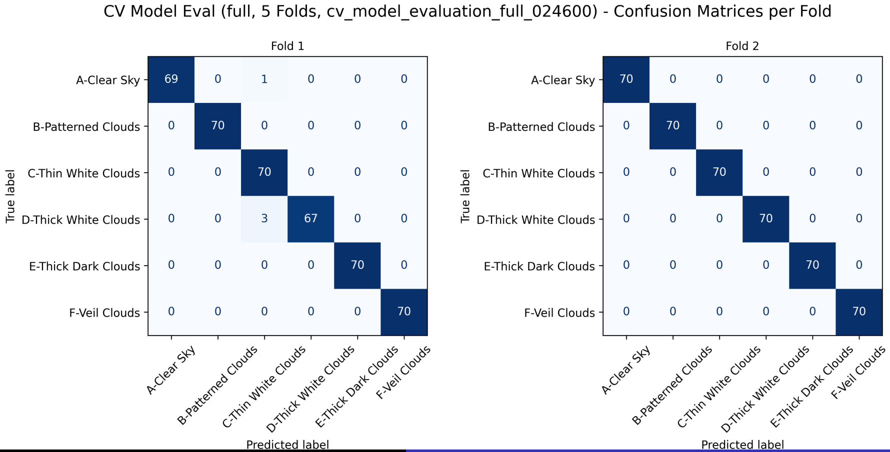
- Further testing is currently in progress, we evaluate multiple models (simple ViT model, hybrid ViT+efficient CNN feature extractor pre-trained on Image-Net).
- At this state, accuracy reached considerable values for some datasets (over 99% on Swimcat-extend, as well as all the other measured metrics.
- This shows the model's well performance and robustness.

Case study results

	Metric	Mean	Std Dev	SEM	CI Margin	CI Lower	CI Upper
1	Accuracy	0.9929	0.0045	0.0020	0.0055	0.9873	0.9984
2	F1 Macro	0.9929	0.0045	0.0020	0.0055	0.9873	0.9984
3	Precision Macro	0.9931	0.0042	0.0019	0.0053	0.9878	0.9984
4	Recall Macro	0.9929	0.0045	0.0020	0.0055	0.9873	0.9984
5	Specificity Macro	0.9986	0.0009	0.0004	0.0011	0.9975	0.9997
6	Roc Auc Macro	0.9997	0.0002	0.0001	0.0003	0.9994	1.0000
7	Pr Auc Macro	0.9930	0.0010	0.0004	0.0012	0.9917	0.9942

Figure: Classification performance of the ViT model on Swimcat-extend (5-Fold CV)

Case study results



Discussion

- CNN-based models (e.g., CloudNet [ZLZS18], ResNet-VGG [YRK+23]) achieve high accuracy (∼98.6%) on balanced datasets using strong augmentation.
- Lightweight architectures (e.g., X-Cloud, M-Cloud [MC25]) prioritize efficiency, trading off accuracy (∼75%) for deployment flexibility.
- Preprocessing-heavy pipelines [LPS+24] improve image quality (e.g., haze correction), but often suffer from limited generalizability due to geographic bias.
- Our ViT model achieved 99% accuracy, validating the value of attention-based mechanisms in capturing complex cloud patterns, while also indicating strong potential for scaling.

Discussion (2)

Approach	Dataset(s)	Classes	Accuracy (%)
CloudNet	CCSN, SWIMCAT	11	88.0 / 98.6
ResNet-VGG	LSCIDMR	11	97.25
X-Cloud / M-Cloud	CCSN, GCD	11, 7	75.0
YOLOv8 + Segmentation	Custom (Tibet)	4	~100 (train), low on Cu
ShuffleNet + SFO	CCSN, SWIMCAT-Ext	11, 6	98.56
CNN	NOAA-20 VIIRS	4	95 (train), 85 (val)
Texton + S-Filters	SWIMCAT	5	~ 100
SNN	MODIS, ABI	Self-supervised	>80
CAE + HAC	MODIS	Unsupervised	N/A
ViT (Ours)	Swimcat	6	**99.0**

Table: Summary of cloud classification models, datasets used, number of classes, and reported accuracies.

Conclusion

- Deep learning has transformed cloud classification, with CNNs and ViTs excelling in capturing visual patterns and contextual dependencies.
- This ViT and Diffusion design addresses key challenges such as class imbalance, image noise, and the difficulty of labelling large datasets.
- These innovations contribute toward automating weather phenomenon analysis, an increasingly vital task in climate-aware societies.

Future Work

- Implement a Diffusion Classifier with similar goals.
- Integrate ViT and Diffusion models into hybrid systems, leveraging both contextual understanding and generative capabilities.
- Investigate multimodal fusion of satellite, radar, and ground-based imagery for richer data representations.
- Utilize generative learning for other tasks such as weather event simulation and prediction.

References I

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby, *An image is worth 16x16 words: Transformers for image recognition at scale*, 2021.

Soumyabrata Dev, Yee Hui Lee, and Stefan Winkler, *Categorization of cloud image patches using an improved texton-based approach*, 2015 IEEE international conference on image processing (ICIP), IEEE, 2015, pp. 422–426.

References II

Andrew Geiss, Matthew W Christensen, Adam C Varble, Tianle Yuan, and Hua Song, *Self-supervised cloud classification*, Artificial Intelligence for the Earth Systems **3** (2024), no. 1, e230036.

Takuya Kurihana, Takuya Kurihana, Ian Foster, Rebecca Willett, and E. J. Moyer, *Cloud Classification With Unsupervised Deep Learning*, Proceedings of the 9th International Workshop on Climate Informatics (CI 2019), 2019, pp. 229–233.

References III

- Shuang Liu, Linlin Duan, Zhong Zhang, Xiaozhong Cao, and Tariq S. Durrani, *Ground-based Cloud Dataset*, https://github.com/shuangliutjnu/TJNU-Ground-based-Cloud-Dataset, 2022, Online; accessed 12 April 2025.
- Pu Liu, *Cirrus Cumulus Stratus Nimbus (CCSN) Database*, 2019.
- Alexander C. Li, Mihir Prabhudesai, Shivam Duggal, Ellis Brown, and Deepak Pathak, *Your diffusion model is secretly a zero-shot classifier*, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), October 2023, pp. 2206–2217.

References IV

Jingxuan Luo, Yubing Pan, Debin Su, Jinhua Zhong, Lingxiao Wu, Wei Zhao, Xiaoru Hu, Zhengchao Qi, Daren Lu, and Yinan Wang, *Innovative cloud quantification: deep learning classification and finite-sector clustering for ground-based all-sky imaging*, Atmospheric Measurement Techniques **17** (2024), 3765–3781.

Eugen Mihuleț and Gabriela Czibula, *Using deep learning for enhancing the performance of ground-based cloud images classification*, SiC **34(1)** (2025), 770–778.

References V

Mesut Toğaçar and Burhan Ergen, *Classification of cloud images by using super resolution, semantic segmentation approaches and binary sailfish optimization method with deep learning model*, Computers and Electronics in Agriculture **193** (2022), 106724.

Mesut Toğaçar and Burhan Ergen, *Classification of cloud images by using super resolution, semantic segmentation approaches and binary sailfish optimization method with deep learning model*, Computers and Electronics in Agriculture **193** (2022), 106724.

References VI

Anastasiia Vasylieva and Roman Melnyk, *Classification of cloud types on satellite images using deep learning*, Manazerska Informatica **3** (2022), 1–13.

World Metheorological Organization, https://wmo.int/, note = Online; accessed 12 April 2025.

Rukhsar Yousaf, Hafiz Zia Ur Rehman, Khurram Khan, Zeashan Hameed Khan, Adnan Fazil, Zahid Mahmood, Saeed Mian Qaisar, and Abdul Jabbar Siddiqui, *Satellite Imagery-based cloud classification using deep learning*, Remote Sensing **15** (2023), no. 23, 5597.

References VII

Jinglin Zhang, Pu Liu, Feng Zhang, and Qianqian Song, *CloudNet: Ground-Based Cloud Classification With Deep Convolutional Neural Network*, Geophysical Research Letters **45** (2018), no. 16, 8665–8672.