Image2image Tropical Cyclone wind field diagnosis with Pix2Pix generative adversarial networks (GANs)

Sarah J. Ollier, PhD[1], Eduardo Siman[1,2], David S. Nolan[1,2], Brian D. McNoldy[1,2]

[1]Worldsphere.ai, [2]University of Miami, sjo@worldsphere.ai, es@worldsphere.ai

1. Introduction

Challenge: Tropical cyclone (TC) wind fields are rarely observed directly with sufficient spatial detail. Faster and more accurate TC information could help operational meteorologists and emergency managers prepare better.

Operational methods:

- Subjective [1] or objective [2] satellite image analysis [3].
- Parametric models for wind radii for NE, SE, SW, NW sections around storm center for 34, 50, 64 knots (kt), (R34, R50, R64) [4,5] .
- **Low temporo-spatial resolution:** 6-hourly analyses [3].

Recent AI advances: TC intensity diagnosis [6], size [7] and inner core wind field estimation [8].

Solution: Computer Vision models [9] to transform single-channel infrared (IR) cloud top brightness temperature satellite imagery into detailed, spatially-resolved surface wind fields.

Advantages:

- **High spatial resolution:** 6°×6° complete wind fields.
- **Increased temporal resolution:** 10-15 minutes
- **Complete structure:** Inner core + outer rainbands.

2. Methodology

Training Data: 2 complementary datasets of **input** satellite TC cloud brightness temperature IR images, paired with **ground-truth** TC wind fields:

- **Simulations:** 1844 hourly synthetic IR/wind pairs from 14 high-resolution Weather Research and Forecasting (WRF) hurricane simulations [10, 11] (Fig. 1a).
- **Observations:** 1984 6-hourly IR satellite images [13, 14] from 49 storms paired with parametric wind fields (Fig. 1b). Parametric model includes translation and size asymmetries following [12].
- All data pre-processed as 6°×6° 256×256 pixels images then combined (Fig. 1).
- **Dataset splits:** to prevent data leakage between sets individual storms were not split between sets, resulting in train: validation: test split percentage ratios of 63:22:15 for wrf2wrf and 70:15:15 for mir2para.

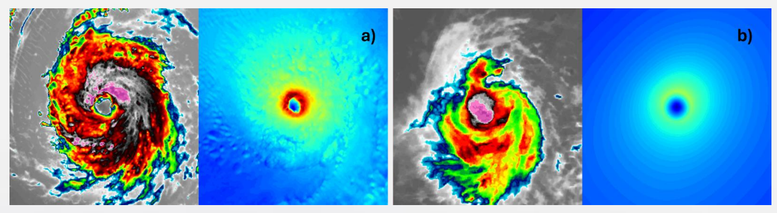

Figure 1: Example input data images a) WRF IR and wind field, b) MergeIR and parametric wind field. Combined input images are 512 x 256 pixels.

Model Architecture:

- Pix2pix GAN Framework [9] (Fig. 2.).
- Trained on NVIDIA L4 GPUs for computational and energy efficiency.

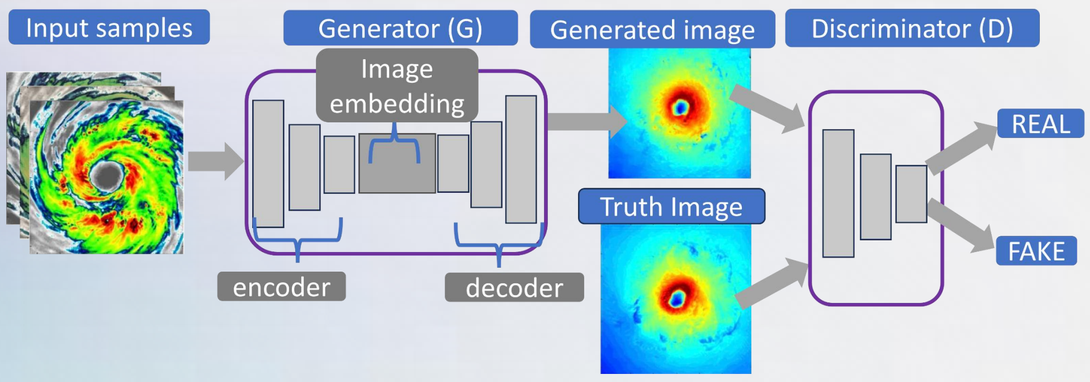

Figure 2: pix2pix GAN [9] for wrf2wrf. The generator (G) aims to create convincing copies of the truth image, and the discriminator (D) tries to determine if these are real or fake, each training iteration this improves the skill of both the G and the D.

Two Models Developed:

- **wrf2wrf**: Trained on WRF simulation data (170 epochs).
- **mir2para**: Trained on real satellite imagery (150 epochs).

Evaluation: Generated image pixel-wise wind speed root mean square error (RMSE) compared against ground truth, with operational parameter extraction (Radius of Maximum Wind (RMW), wind radii (R34, R50, R64)) for validation against IBTrACS (International best-track for climate stewardship) operational data [15,16].

3. Results and Discussion

Pixel-wise wind speed RMSE:

- wrf2wrf: 7.3 kt (min 3.6 kt) across 298 test images, mir2para: 6.7 kt (min. 1.2 kt) across 276 test images (Fig 3.).
- Stronger, more organised storms generally better predicted by both models, consistent with [8].

Cross-domain performance: wrf2wrf trained on synthetic data achieves 10.4 kt RMSE when tested on 73 real IR images without additional training compared to 7.5 kt for mir2para. This provides encouraging evidence that both models will perform well on operational data.

Comparison with operational IBTrACS best track data:

- Mean Absolute Error (MAE) between the prediction wind radii and IBTrACS as low as 8.9 NM (bias -3.3 NM) in NE quadrants (8.9 – 12.0 NM for all quadrants)(Fig. 4, Fig. 5).
- Spatial resolution: whole wind field rather than point estimates and fixed radii [3, 4, 5].

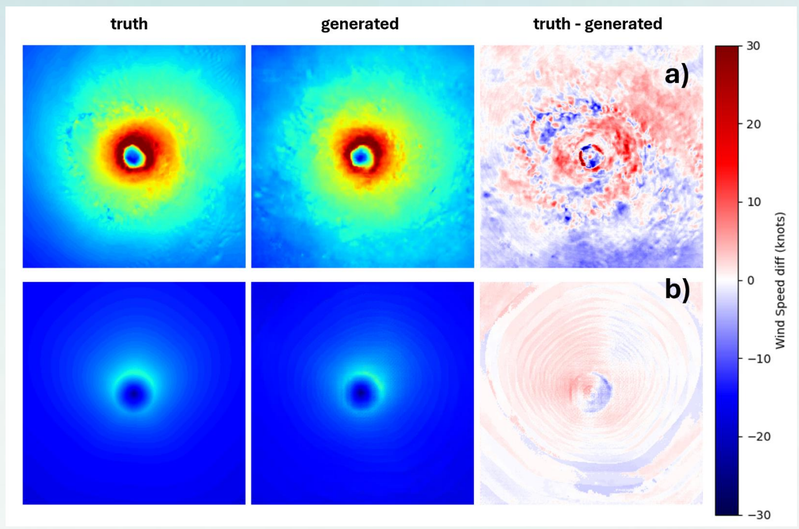

Figure 3: Test set validation images for a) wrf2wrf Hurricane Florence 2018. b) mir2para Hurricane Larry 2021. From left to right truth image, generated image, difference truth - generated. Legend shows wind speed difference in knots.

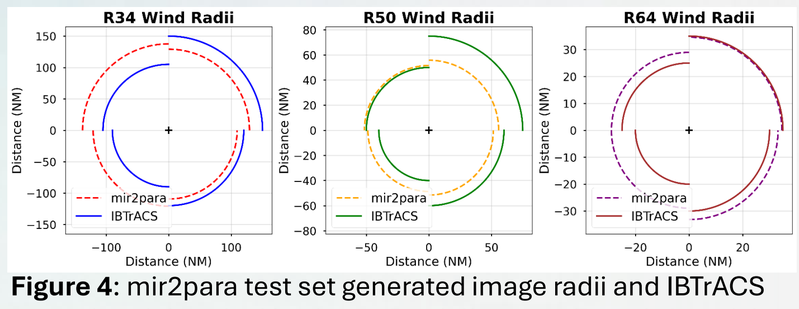

Figure 4: mir2para test set generated image radii and IBTrACS operational data for Hurricane Danielle 2010. From left to right truth R34, R50, R64. Dash lines - mir2para, IBTrACS - solid line.

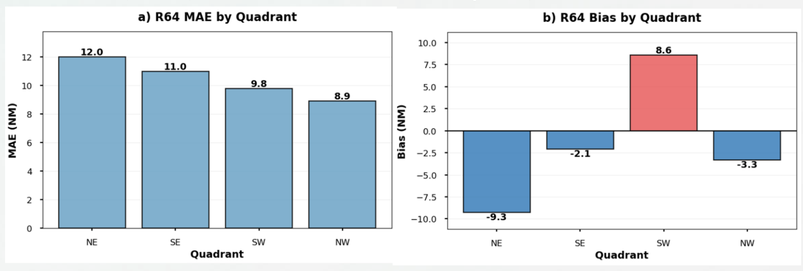

Figure 5: R64 wind radii comparison with IBTrACS for 278 mir2para test set images. Per quadrant: a) Mean Absolute Error b) bias.

Comparison with AI methods:

- RMW RMSE of 18.1 NM (~34 km), less than the 53 km RMSE reported between predictions and operational flight measurements in [8].
- Greater spatial coverage: 6°×6° generated images resolve inner-core and outer rainbands, compared to inner-core ~3.6°×3.6° with 1.25° usable radius in [8].

4. Conclusions

Our pix2pix GAN models successfully generate realistic TC wind fields from satellite imagery with:

- Low computational training cost.
- Performance comparable to operational best track data.
- Superior spatial and temporal resolution vs. existing methods.
- High potential for operational deployment.

Climate mitigation impact: Earlier and faster high-resolution TC wind information for Operational meteorologists, Emergency managers, Disaster preparedness and Climate resilience.

Climate mitigation impact:
Earlier and faster high-resolution TC wind information for Operational meteorologists, Emergency managers, Disaster preparedness and Climate resilience.

Future Work:

- **Category-specific tuning:** Develop classification-based routing to category-specific models for improved performance across storm intensities
- **Physics-conditioned learning:** Incorporate storm characteristics that influence wind field structure into model architecture and inference.

[1] Velden et al. 2006: Bull. Amer. Meteor. Soc., **87**, 1195–1210, doi:10.1175/BAMS-87-9-1195.
[2] Olander and Velden 2019: Wea. Forecasting, **34**, 905–922, doi:10.1175/WAF-D-19-0007.1.
[3] Knaff et al. 2011: J. Appl. Meteor. Climatol., 50, 2149–2166, doi:10.1175/JAMC-D-11-09.1
[4] Knaff et al. 2017: Wea. Forecasting, **32**, 629–644, doi:10.1175/WAF-D-16-0168.1.
[5] DeMaria et al. 2022: Wea. Forecasting, **37**, 2141–2159, doi:10.1175/WAF-D-22-0039.1.
[6] Chen et al. 2019: Wea. Forecasting, **34**, 447–465, doi:10.1175/WAF-D-18-0136.1.
[7] Wang and Li 2023: Mon. Wea. Rev., **151**, 403–417, doi:10.1175/MWR-D-22-0106.1.
[8] Wimmers et al. 2024: Artif. Intell. Earth Syst., **3**, e230084, doi:10.1175/AIES-D-23-0084.1
[9] Isola et al. 2017: Proc. IEEE CVPR, 1125–1134, doi:10.1109/CVPR.2017.632.
[10] Nolan et al. 2013: J. Adv. Model. Earth Syst., **5**, 382–405, doi:10.1002/jame.20031.
[11] Klotz & Nolan 2019: Mon. Wea. Rev., **147**, 247–268, doi:10.1175/MWR-D-18-0296.1.
[12] Knaff et al. 2007: Wea. Forecasting, 22, 781–791, doi:10.1175/WAF1026.1.
[13] Janowiak et al. 2017: NCEP/CPC L3 4km merged IR V1, GES DISC, doi:10.5067/P4HZB9N27EKU.
[14] NASA Earthdata: MERGIR dataset, https://search.earthdata.nasa.gov/search?q=MERGIR
[15] Knapp et al. 2010: Bull. Amer. Meteor. Soc., 91, 363–376, doi:10.1175/2009BAMS2755.1.
[16] Gahtan et al. 2024: NOAA National Centers for Environmental Information, doi:10.25921/82ty-9e16.