Application of Machine Learning for Meteorological Fronts Detection and Classification

Edson C. Tchivango Segunda[1], **Adriano A. Pereira**[2], **Alan James P. Calheiros**[3]

[1]National Institute for Space Research (INPE)
Avenida dos Astronautas, 1.758 - Jd. Granja - CEP 12227-010.
São José dos Campos - SP - Brasil

edson.segunda@inpe.br, adriano.almeida@inpe.br, alan.calheiros@inpe.br

Abstract. *One of the main drivers of weather change are fronts, which are transition zones between air masses of different temperatures and densities. They can cause rain, thunderstorms, wind shifts, and temperature changes. This article describes using a Convolutional Neural Network (CNN) to detect and classify fronts by learning patterns from weather variables such as temperature, wind magnitude, mean sea level pressure, and dew point. These patterns are linked to pixel classes that indicate front type and location.*

Resumo. *Um dos principais fatores que influenciam as alterações climáticas são as frentes, zonas de transição entre massas de ar com diferentes temperaturas e densidades. Podem provocar chuva, tempestades, mudanças na direção do vento e alterações de temperatura. Este artigo descreve a utilização de uma Rede Neural Convolucional (CNN) para detetar e classificar frentes, aprendendo padrões a partir de variáveis meteorológicas como a temperatura, intensidade do vento, pressão média ao nível do mar e ponto de orvalho. Estes padrões estão associados a classes de pixéis que indicam o tipo e a localização da frente.*

1. Introduction

Since the temperature distribution is the most important regulator of atmospheric density, a front almost invariably separates air masses of different temperature. Along with the basic density criterion and the common temperature criterion, many other features may distinguish a front, such as a pressure trough, a change in wind direction, a moisture discontinuity, and certain characteristic cloud and precipitation forms [Society 2024]. The study area was based on the ground truth data (synoptic map) done by the Center for Weather Forecasting and Climate Studies (CPTEC) which covers an area of latitude 15°N, 65°S and longitude 100°W, 20°W. Before diving into the type of fronts, let's understand air masses. [Office 2025], air masses are large bodies of air with similar temperature and moisture characteristics.

A cold front forms a boundary between warm and cold masses of air when cold air forces the warmer air rising. [Bergeron 1937]. On synoptic chart or weather maps cold fronts are illustrated using a blue line with triangles. - Warm fronts are represented by a line with red semicircles, resembling half suns. Continuing with, [Bergeron 1937] and [Office 2025], define, warm front as when warm air is advancing and rising over

cooler air. - Occluded fronts are more complex and are shown as a line with both semicircles and triangles, usually colored purple. An occlusion happens when a cold front catches up with a warm front. [Office 2025].

Exist different equations, ways or technique to detect fronts and here are two equations or technique: The slope of isentropes constituting a frontal zone by [Delden 1999] and Thermal Front Parameter equation by [Renard and Clarke 1965]. [EUMETRAIN 2015] defines Thermal Front Parameter (TFP) as a meteorological value calculated from the second derivative of temperature, which indicates the location of weather fronts by highlighting areas of significant temperature change.

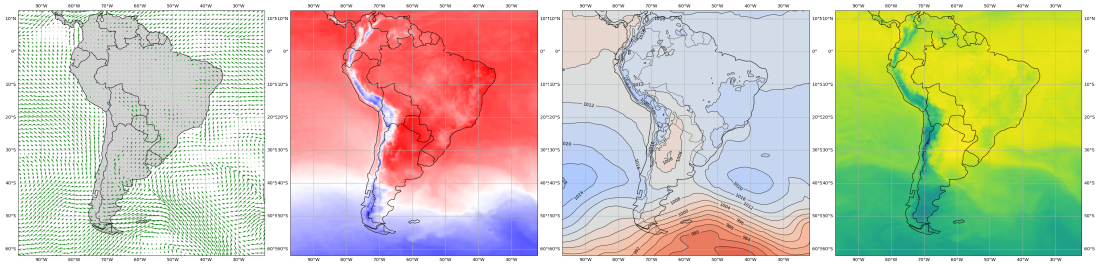
$$TFP = -\nabla|\nabla T| \cdot \frac{\nabla T}{|\nabla T|}$$

2. Methodology

For this study, a classification architecture was used to detect front classes directly, identifying which type of front each pixel belongs to rather than just its approximate color. This resulted in a model better suited for interpretation and quantitative validation and deviations (e.g., [0,0,240] vs. [0,0,255]) even when both represented a cold front.

2.1. Input/Features

For input data, atmospheric variables obtained from the ERA5 reanalysis were used [ECMWF], with a temporal resolution of 6 hours and a horizontal resolution of 0.25° x 0.25°. The selected variables were: Mean sea level pressure (msl); Air temperature at 2 meters (t2m); Zonal and meridional components of the wind at 10 meters (u10 and v10) combined into vector magnitudes and dew point temperature at 2 meters (d2m).



(a) Wind (b) Temperatur (c) Pressure (d) Dewpoint

Figura 1. Era5 Input Data

2.2. Target - Weather Front Maps (RGB Masks)

As the dependent variable (Y), were used georeferenced images in GeoTIFF format derived from synoptic analyses maps delivered every 6 hours from Center for Weather Forecasting and Climate Studies (CPTEC).

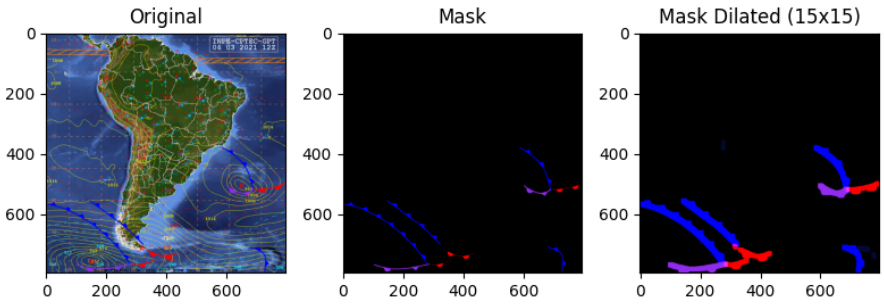


Figura 2. Mask Generation Phases - Original Image source: (CPTEC)

Dataset Division

The table below contains information about the dataset division for the training. Initial set of images covering 6-hour intervals. **Train** 2018/01/01 - 2020/10/28. Validation 2020/10/29 - 2020/12/31.Test 2021/03/01 - 2021/03/30

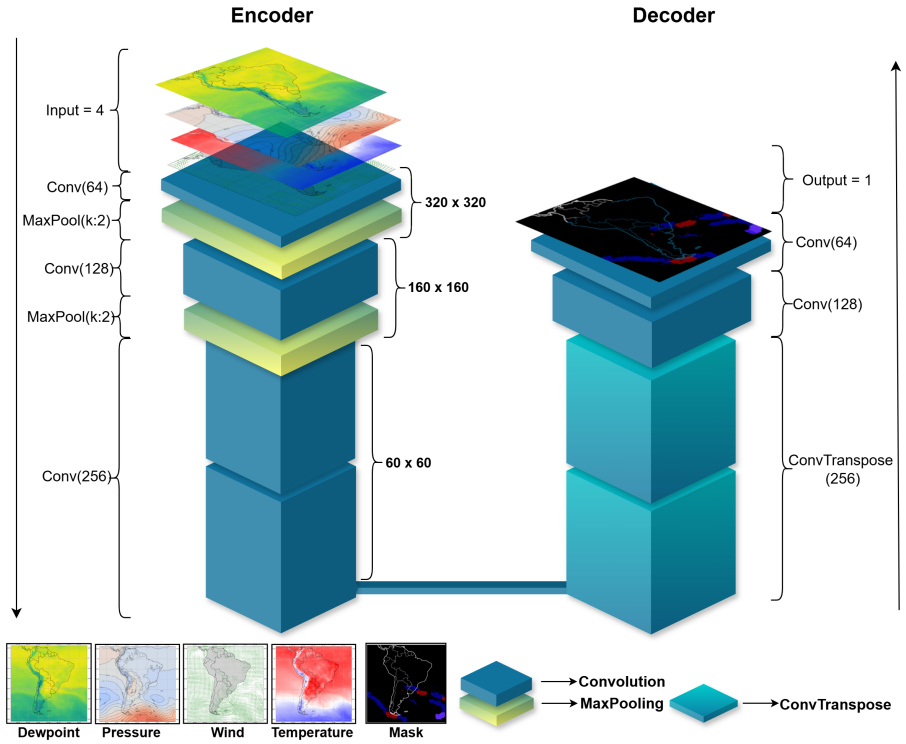


Figura 3. Model Architecture

The architecture follows an Encoder-Decoder framework (U-Net) designed for pixel-wise dense prediction tasks.

3. Results

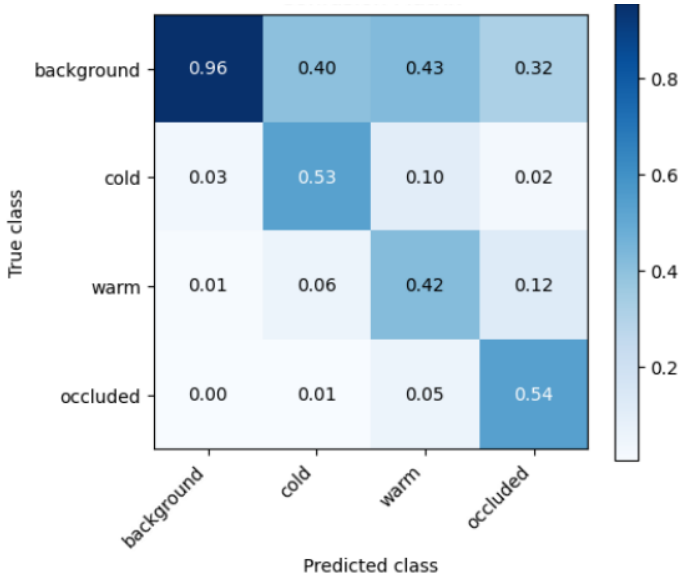


Figura 4. Confusion Matrix

Tabela 1. Meteorological Front Classification Performance Metrics

Class	Precision	Recall	F1-Score	CSI (Threat Score)
Cold Front	0.530	0.779	0.631	0.461
Warm Front	0.420	0.689	0.522	0.353
Occluded Front	0.540	0.900	0.675	0.510

Values along the main diagonal represent correct classifications (True Positives) per class: background 0.96, cold 0.53, warm 0.42, and occluded 0.54. Most misclassifications of fronts fall into the background class, which is expected due to dilation and the unbalanced data distribution. The confusion matrix also provides aggregated metrics for clearer comparison: accuracy (overall correct predictions), precision and recall (reported per class for imbalanced data), and F1 score (harmonic mean of precision and recall). Recall equals the diagonal value, while precision is true positives divided by all instances predicted as that class.

4. Conclusion

This study showed that deep learning can automatically detect and classify meteorological fronts (cold, warm, and occluded). Using ERA5 variables with curated RGB frontal masks, we built a consistent dataset for supervised training and a reproducible workflow for converting masks into one-hot labels. Morphological pre-processing (dilation) proved essential for learning fine frontal structures. Limited dataset size and natural class imbalance constrained performance, and predictions showed imperfections like irregular boundaries, noise, and positional inaccuracies. Overall, this research enables data-driven frontal detection with potential applications in synoptic analysis, climate monitoring, and real-time weather forecasting.

5. Reference

Bergeron, T. (1937). On the physics of fronts. *Bulletin of the American Meteorological Society*, 18(9):1–3.

Delden, A. V. (1999). The slope of isentropes constituting a frontal zone. *Tellus A: Dynamic Meteorology and Oceanography*, 51(5):603–611.

ECMWF. Climate data store.

EUMETRAIN (2015). Numerical parameters for synoptic to mesoscale cloud systems - thermal front parameter. page 1. EUTMETRAIN.

Office, M. O. (2025). What are weather fronts. page 1. Met Office.

Renard, R. J. and Clarke, L. C. (1965). Experiments in numerical objective frontal analysis. *Monthly Weather Review*, 93(9):547–556.

Society, A. M. (2024). Front. In *Front*, page 1. AMS.