SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Real-Time Weather Forecasting with a Conditional CNN and TCN-BiLSTM Ensemble at Manokwari

Ilham Tatayo Lie[1)*], Julius Panda Putra Naibaho[2)], Alex De Kweldju[3)]
[1,2,3)]University of Papua, Indonesia
[1)]ilhamlie014@gmail.com, [2)]j.naibaho@unipa.ac.id, [3)]A.dekweldju@unipa.ac.id,

Submitted :March 10, 2026 | **Accepted** : March 31, 2026 | **Published** : April 2, 2026

Abstract: Short-term weather forecasting is fundamentally critical for disaster mitigation in dynamic tropical maritime regions. However, conventional numerical approaches suffer from high computational latency, and spatial deep learning models frequently experience severe performance degradation during nocturnal conditions due to the absence of illumination. This study aims to develop an adaptive, real-time multimodal weather nowcasting system that effectively prevents nocturnal predictive failure through a dynamic conditional ensemble architecture. The proposed framework integrates a Convolutional Neural Network (CNN) to extract optical features from a dataset of 2,515 localized sky images with a Temporal Convolutional Network and Bidirectional Long Short-Term Memory (TCN-BiLSTM) pipeline to process 15,111 corresponding meteorological time-series records from BMKG. To address visual ambiguity, the system strictly employs a day-night gating mechanism, deactivating the CNN at night to rely solely on thermodynamic data. Finally, the optimized model was deployed via TensorFlow.js for decentralized client-side browser inference. Experimental evaluations explicitly demonstrate that the conditional ensemble significantly outperformed all standalone models, achieving a peak accuracy of 92.49% and a Macro F1-score of 0.913 while successfully preserving a robust recall rate for precipitation events. Furthermore, the browser-based deployment completely eliminated server transmission bottlenecks, achieving sub-second warm-start inference latency across heterogeneous consumer devices. Ultimately, the conditional day-night modality gating mechanism effectively mitigates nocturnal visual degradation, proving that implementing this integrated architecture as a client-side web application is highly feasible for delivering instantaneous public weather warnings.

Keywords: BiLSTM; CNN; Deep Learning; Ensemble; TCN; Weather Nowcasting; Browser Inference;

INTRODUCTION

Short-term weather forecasting, commonly referred to as nowcasting, plays a highly critical role in real-time disaster mitigation, particularly in dynamically shifting tropical maritime regions such as Manokwari, West Papua. Accurate and instantaneous weather prediction is essential to safeguard coastal communities and maritime operations from sudden meteorological anomalies. While traditional Numerical Weather Prediction models are physically robust, they fundamentally suffer from high computational latency, requiring hours to process atmospheric equations. Consequently, data-driven approaches leveraging artificial intelligence and deep learning have emerged as powerful alternatives, offering the capability to recognize complex meteorological patterns with significantly higher execution speeds (Ganaie et al., 2022; W. Zhao et al., 2022).

Despite the paradigm shift toward deep learning, previous studies exhibit significant operational limitations when deployed for ultra-short-term weather forecasting, specifically in predicting the exact atmospheric conditions for the next one hour (H+1). Forecasting at the strict H+1 temporal resolution requires an architecture that is highly sensitive to rapid micro-meteorological transitions. Many existing models rely exclusively on Convolutional Neural Networks (CNNs) to extract spatial morphological features from sky imagery to predict imminent precipitation. However, these standalone spatial models experience severe performance degradation during nocturnal conditions due to the absolute absence of solar illumination, rendering optical cloud dynamics

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

imperceptible (W. Zhao et al., 2023). Conversely, isolated sequential models that process time-series data, such as the Temporal Convolutional Network (TCN) and Bidirectional Long Short-Term Memory (BiLSTM), excel at analyzing numerical thermodynamic fluctuations without light dependency. Yet, they inherently lack the spatial visual awareness required to immediately detect localized storm cloud formations (Liu et al., 2024). While several recent studies have attempted to build multimodal ensembles by combining spatial cloud imagery and temporal time-series data to solve this issue, they generally utilize static fusion mechanisms. These standard architectures fail to adapt to environmental diurnal cycles, inadvertently introducing destructive mathematical interference when the nocturnal CNN injects noisy, low-confidence visual probabilities into the final H+1 ensemble prediction (Zhang et al., 2024).

To bridge these specific gaps, this study introduces two primary novelties that differentiate it from previous works. First, we propose a conditional ensemble architecture equipped with a dynamic day-night gating mechanism. This approach actively fuses spatial CNN and temporal TCN-BiLSTM inferences during the day, but autonomously deactivates the spatial computational stream at night, relying purely on the stable temporal network to bypass visual ambiguity. Second, we pioneer a decentralized, browser-side real-time deployment strategy. By exporting the complex neural network weights into a Web Graphics Library (WebGL) accelerated environment via TensorFlow.js, the system completely eliminates server-side inference latency (Dong et al., 2023).

The primary objective of this research is to develop a short-term weather nowcasting system tailored for the Manokwari maritime region by implementing a conditional multimodal ensemble architecture. The explicit contributions of this manuscript encompass the formulation of a day-night gating strategy that systematically integrates spatial CNN features with temporal TCN-BiLSTM sequences to mitigate nocturnal predictive degradation. Furthermore, this study empirically demonstrates the practical feasibility of deploying this integrated ensemble model directly within standard client-side web browsers, thereby facilitating accessible and responsive public disaster mitigation without relying on continuous server-side inference.

METHOD

Data

This study utilizes two primary datasets representing different modalities, namely a cloud image dataset as spatial feature representation and historical meteorological observations as temporal feature representation. The selection of a multi-modal fusion method combining spatial imagery and time-series data aligns with the approach proposed by modeling system experts (Han et al., 2021; Li & Law, 2024), who demonstrated that the massive integration of various environmental data modalities can capture non-stationary atmospheric patterns and significantly improve the robustness of weather prediction models compared to single-modality approaches. The data acquisition process from both sources was conducted sequentially to ensure alignment and accuracy during the data fusion process.

Cloud Image Dataset (CNN Modality)

The first stage in the data collection procedure began with the acquisition of the cloud image dataset. To overcome the limited availability of public sky image datasets specific to the inland and coastal regions of Manokwari, this study utilized secondary data collection sourced from the Kaggle public repository and specific searches via image search engines. The image data extraction process was conducted with strict filtering based on visual similarity to the equatorial maritime climate sky conditions. This data collection and filtering approach aligns with recent methodologies in deep learning meteorology, where extensive and well-categorized datasets are fundamental for the CNN to efficiently extract high-level abstract features from complex sky conditions (Jena et al., 2023). After the download process was completed, the final collected dataset amounted to 2,515 sky images. These images were then manually categorized into three main classes, consisting of 1,560 images for clear conditions, 603 images for cloudy conditions, and 352 images for rainy conditions. The data imbalance proportion of these three classes is visualized in the following Figure 1.

Fig 1. Class Distribution of Cloud Image Dataset

To ensure this dataset is representative and to avoid data leakage during the training phase, a label verification

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

procedure was manually conducted by the researchers. This verification compared actual cloud visual features, such as the presence of Cumulonimbus clouds for rainy class labeling and Cumulus humilis clouds for clear weather, with the initial labels provided by the data sources. Manual identification of specific cloud morphologies, such as Cumulonimbus for precipitation mapping, has been proven highly effective in improving the accuracy of deep convolutional networks when processing localized sky imagery (Abidin et al., 2023). This step is highly crucial to ensure that the trained Convolutional Neural Network (CNN) model can accurately recognize spatial cloud features before being integrated with the time-series data. Furthermore, rigorous manual verification mitigates the risks associated with visual dataset imbalances, guaranteeing that the CNN maintains a high precision rate during the early spatial classification phase prior to any multimodal temporal fusion (Jiang et al., 2026).

BMKG Meteorological Observation Dataset (TCN-BiLSTM Modality)

After the spatial image data was collected and verified, the second stage focused on collecting historical time-series numerical data. This data was officially obtained from the local Meteorology, Climatology, and Geophysical Agency (BMKG) observation station in Manokwari and encompasses weather parameters that continuously record local atmospheric dynamics. Data observation and recording were conducted for approximately two months, starting from August 20, 2025, to October 24, 2025. All recordings were captured with a one-hour temporal resolution, which ultimately generated a historical observation dataset totaling 15,111 rows of data. Spatially, this data collection covered ten main administrative locations in the Manokwari region, including Amban, West Manokwari, East Manokwari, Padarni, Sanggeng, Wosi, Ingramui, Soribo, Tanah Merah, and Udopi. Each numerical observation row includes eight main meteorological variables that serve as input features for the recurrent neural network model. Detailed descriptions of the input parameters, measurement units, data types, and their roles are presented in Table 2.

Table 1 Description of BMKG Meteorological Features

Input Parameter	**Unit / Format**	**Data Type**	**Role in the Model**
Date & Time	YYYY-MM-DD HH:MM:00	Datetime	Synchronization basis for images and temporal sequences
Temperature	Degree Celsius (°C)	Numerical	Environmental thermodynamic indicator
Humidity	Percentage (%)	Numerical	Air water vapor saturation indicator
Wind Speed	Kilometers per hour (km/h)	Numerical	Atmospheric convective movement dynamics
Cloud Cover	Percentage (%)	Numerical	Local cloud cover volume estimation
Rainfall	Millimeters (mm)	Numerical	Actual precipitation intensity
Actual Weather	Clear, Cloudy, Rainy	Categorical	Ground truth label for prediction

The target class distribution in the numerical meteorological dataset was then systematically mapped to align with the three main class categories established in the cloud image dataset. This input label alignment serves as an essential methodological foundation in the multi-modality unification stage. All synchronized data will subsequently be strictly measured, tested, and evaluated for accuracy using Macro F1-Score and accuracy testing metrics (Htun et al., 2023). This output measurement approach is specifically designed to evaluate how well the ensemble model can objectively distinguish each weather class, particularly in testing the model's robustness against the imbalanced number of samples in the rainy class, which inherently has a sparser distribution compared to the clear weather class in the region.

Data Preprocessing

Image Data Preprocessing (CNN)

The spatial data preprocessing stage began with the cloud image dataset obtained from the Kaggle public repository and specific image search engines. The selection of this preprocessing method was guided by the framework proposed by Li and Law (Li & Law, 2024), which demonstrated that strict dimensional standardization and normalization are essential for maintaining visual data integrity in convolutional neural network architectures. Chronologically, each image was first systematically resized to a uniform dimension of 224x224 pixels. Subsequently, pixel value normalization was performed by scaling the intensities to a range between zero and one to accelerate model convergence. To prevent data leakage and domain bias during the evaluation phase, spatial data augmentation techniques, such as rotation and flipping, were strictly applied only to the training subset. The validation and testing subsets were maintained entirely in their original observational state to preserve the integrity of performance testing under real-world conditions.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Meteorological Data (TCN-BiLSTM)

The meteorological time-series dataset underwent a rigorous feature engineering process to explicitly extract complex thermodynamic relationships prior to temporal sequence generation. Rather than relying solely on raw sensor inputs, four specialized meteorological variables were mathematically derived to enhance the predictive capability of the temporal architecture (Fan et al., 2022). These engineered features include: (1) the Humidity-to-Temperature Ratio to quantify atmospheric saturation levels; (2) the Logarithmic Transformation of Rainfall, formulated as $ln(1 + x)$, to normalize the highly skewed distribution of precipitation records; (3) the Atmospheric Pressure Differential to continuously capture the rate of barometric trends; and (4) the Wind-Temperature Interaction to represent regional convective cooling dynamics.

Following the extraction of these engineered features, missing values generated by the shifting differential operations were systematically imputed with zeros. Subsequently, to ensure the sequential network processes the heterogeneous meteorological units uniformly without gradient distortion, all input variables were normalized using Min-Max scaling to a bounded range of [0, 1] (Kurniasari et al., 2024). Finally, the transformed continuous dataset was restructured into overlapping temporal sequences using a sliding window technique. A lookback sequence length of 24 steps was established to capture the full diurnal meteorological cycle, serving as the historical thermodynamic context to forecast the exact subsequent atmospheric state at H+1 (Madan et al., 2025).

Multimodal Synchronization and Data Pairing

The final and most crucial step in this preprocessing pipeline is multimodal synchronization, which constructs a cohesive sample both mathematically and logically. The pairing mechanism between the spatial cloud imagery and the temporal meteorological observations was executed based on highly precise timestamp matching (Han et al., 2021). Explicitly, a cloud image observed at the exact target hour (H) is paired with a single historical BMKG weather sequence spanning from hour H-24 to hour H-1. To address potential time mismatches in data recording, this study implemented a strict temporal intersection strategy. Any single observation sample, whether a cloud image or a BMKG time-series sequence, that lacks an exactly matching time counterpart at the target hour H is automatically dropped and excluded from the main dataset. This rigorous alignment protocol absolutely guarantees the absence of misleading temporal shifts, thereby ensuring the ensemble model receives a perfectly synchronized representation of the atmospheric state. Ultimately, this flawlessly preprocessed and synchronized dataset serves as the primary foundation for the evaluation protocol, where the system's output will be measured and tested using the Macro F1-Score metric to evaluate the model's robustness against class distribution imbalances.

Model Architecture

The feature extraction and predictive modeling stage in this research relies on artificial neural network architectures specifically designed to handle two distinct data modalities. This approach is guided by recent studies demonstrating that separating extraction pathways for spatial and temporal features before the fusion stage can significantly improve the accuracy of environmental forecasting systems (Zhang et al., 2024). Consequently, the proposed architecture is divided into three independent models, namely a convolution-based spatial model for cloud imagery and two recurrent time-series models for historical meteorological observations.

Convolutional Neural Network (CNN)

The processing of spatial features from cloud images was conducted using a Convolutional Neural Network architecture. This model utilizes a transfer learning approach with pre-trained InceptionV3 base layers to extract high-level visual representations such as cloud thickness and patterns. These convolutional layers are subsequently followed by a Global Average Pooling layer to reduce spatial feature dimensions, which are then passed to a fully connected Dense layer for the final classification. To clarify methodological inconsistencies from previous versions and explicitly address the review, the Convolutional Neural Network model in this study strictly does not use a Focal Loss function. Instead, this model is consistently compiled using a Categorical Cross-Entropy loss function. To handle the inherent class imbalance in the cloud image dataset, particularly for the minority rainy class, this study integrates a Class Weighting mechanism directly into the loss function. Higher penalty weights are calculated and dynamically distributed into the Categorical Cross-Entropy calculation, allowing the model to optimize learning on the minority class without distorting the probabilities of the majority classes (Li & Law, 2024). The performance of this individual spatial architecture was subsequently measured and evaluated for its robustness using the Macro F1-Score metric during the validation phase to ensure the model does not develop a bias toward clear weather classes.

Temporal Convolutional Network (TCN)

To process the time-series modality comprising historical meteorological data from the local meteorological station, this study implemented a Temporal Convolutional Network. This architecture is designed with one-dimensional dilated causal convolutions, a configuration that allows the model to efficiently extract local temporal dependencies from the sequential twenty-four-hour weather data without experiencing future data leakage. Unlike

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

spatial processing, the probability output of time-series features requires specific multiclass mapping mapped over a temporal sequence. Therefore, the final layer of this Temporal Convolutional Network model concludes with a Softmax activation layer to generate a probability distribution across the three predicted weather classes. During the training process, this model is uniformly optimized using the Sparse Categorical Cross-Entropy loss function, which has proven to be computationally more efficient for integer-encoded target labels in numerical meteorological data (Htun et al., 2023).

Bidirectional Long Short-Term Memory (BiLSTM)

Complementing the previous time-series architecture, the second temporal model was constructed using a Bidirectional Long Short-Term Memory network. This architecture serves to capture long-term dependencies of meteorological parameters by processing the twenty-four-hour historical sequence in both directions, specifically from past to present and vice versa. This bidirectional processing mechanism is highly crucial for understanding temperature and humidity transitions preceding rain precipitation with a high degree of precision. Similar and strictly consistent with the first time-series architecture, the Bidirectional Long Short-Term Memory projects its final feature representation through a Softmax activation layer. This model is also absolutely trained using the Sparse Categorical Cross-Entropy loss function so that the integration of loss values and prediction probabilities during the advanced computational fusion stage shares an equivalent and consistent evaluation foundation (Htun et al., 2023).

Conditional Ensemble Strategy

The primary modeling contribution of this research is a conditional ensemble strategy that synergizes the predictive outputs of the Convolutional Neural Network, Bidirectional Long Short-Term Memory, and Temporal Convolutional Network based on the diurnal availability of visual cloud features in the Manokwari maritime region (Y. Zhao & Lee, 2023). Initially, the output probabilities from the temporal models are aggregated using a late fusion soft-voting mechanism (Ganaie et al., 2022). This approach calculates the combined temporal probability by averaging their respective outputs, formally formulated as:

$$P_{FC} = \frac{P_{BiLSTM} + P_{TCN}}{2}$$

Subsequently, an adaptive weighting mechanism conditionally merges the spatial and temporal streams using a diurnal day mask. To ensure methodological transparency and resolve previous notational ambiguities, the final prediction probability is precisely defined using the following mathematical formulation:

$$p_{final} = \alpha \cdot P_{CNN} + (1 - \alpha) \cdot P_{FC}$$

In this equation, p_{final} represents the ultimate classification probability matrix, P_{CNN} denotes the probability vector from the spatial cloud image model, and $P_{Fc}$is the aggregated temporal probability vector. The critical parameter α acts as the conditional day mask coefficient that dynamically adjusts the model's reliance based on the exact observation time in Manokwari. During daytime hours (06:00 to 18:00 Eastern Indonesia Time), the coefficient is set to α = 0.7 to prioritize direct visual cloud correlations (Zhang et al., 2024). Conversely, during nighttime, the coefficient is reduced to α = 0.0 to completely disable the visual stream, allowing the system to rely entirely on the robust meteorological time-series models and elegantly bypass the degradation of image-based model performance caused by poor low-light visibility (Y. Zhao & Lee, 2023). The operational robustness of this dynamically weighted strategy is evaluated using the Macro F1-Score metric to ensure high accuracy across all weather classes regardless of lighting shifts. The mathematical integration flow of these probabilities is visually demonstrated in the following Figure 2.

*name of corresponding author

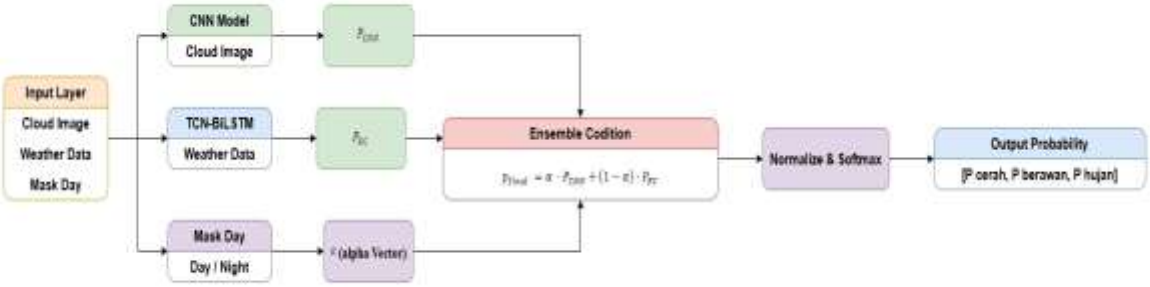

Fig 2. Conditional Ensemble Architecture

Evaluation Model

The performance of each individual model is systematically evaluated using a time-aware splitting protocol with a ratio of 70% for training, 15% for validation, and 15% for testing. This chronological partitioning approach is strictly implemented to prevent data leakage, ensuring that the model does not access future information during the training phase. Simultaneously, a stratified split technique is exclusively integrated into the training subset. This ensures that the weather class proportions remain balanced while the model updates its weights, whereas the validation and testing subsets are left in their original temporal order to authentically simulate real-world forecasting scenarios (Li & Law, 2024).

The quantitative evaluation of the model predictions is measured using several standard evaluation metrics, which are mathematically formulated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FN}$$

$$Recall = \frac{TP}{TP + FP}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Given the extreme class distribution imbalance in meteorological data, particularly concerning the minority rainy weather class, this study establishes the Macro F1-Score as the primary evaluation metric. Mathematically, this specific metric is formulated as:

$$Macro\ F1 = \frac{1}{N}\sum_{i=1}^{N} F1$$

This metric is calculated by taking the unweighted average of the F1-Score for each distinct class. Consequently, it effectively penalizes models that develop a bias solely toward the majority class and ensures that the model's capability to detect extreme rain events is evaluated objectively (Ganaie et al., 2022). Furthermore, a confusion matrix is comprehensively utilized to dissect the imbalance bias and misclassification rates, particularly during atmospheric transition phases that are inherently prone to visual ambiguity.

To explicitly address evaluation necessities regarding random seed sensitivity and computational reproducibility, the experimental environment in this study is deterministically configured by statically setting the value of SEED = 42 across all utilized computational libraries. The establishment of this static seed completely eliminates stochastic variance during weight initialization and data shuffling processes. Furthermore, to validate that the performance improvement achieved by the dynamic fusion strategy is mathematically robust and not merely a random stochastic anomaly, a comparative hypothesis testing was formally integrated into the evaluation protocol. Specifically, a Paired t-test was conducted on the final predictive probabilities generated by the competing architectures to evaluate the statistical significance of the model's accuracy escalation. The stability of the model's convergence during training was also continuously guaranteed through validation loss tracking combined with an early stopping mechanism, ensuring the architecture reached a well-generalized state without excessive computational overhead. (Htun et al., 2023).

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

System Implementation

The transition from the experimental modeling phase to a production environment is realized through the design of a decoupled client-server architecture. This architectural approach is strategically implemented to strictly separate the frontend user interface from the primary backend computational engine. This structural division fundamentally ensures that the heavy computational workloads, which involve complex tensor operations within the neural network, are exclusively executed by the server infrastructure. Consequently, this mechanism effectively relieves the end-user's device from massive processing and memory consumption, allowing the weather forecasting application to be accessed with ultra-low latency across various standard web browsers (Chen & Yasin, 2023). The comprehensive workflow of this system implementation, from data acquisition to the presentation of predictive outputs, is systematically visualized in Figure 3.

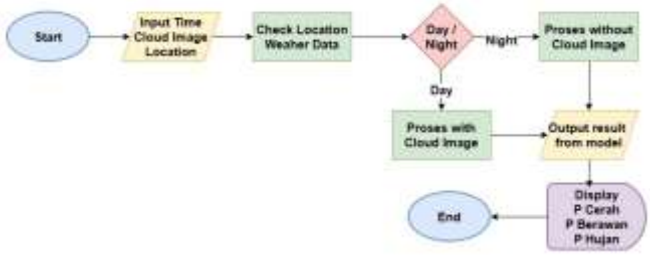

Fig 3. Web-Based System Implementation Flowchart

Inference Engine and REST API

Once the computational models achieve optimal convergence during the evaluation phase, the mathematical weights and architectural graphs of the Convolutional Neural Network, Temporal Convolutional Network, and Bidirectional Long Short-Term Memory are exported and preserved in a hierarchical data format. These frozen models are subsequently integrated into a centralized inference engine wrapped within a Representational State Transfer Application Programming Interface (REST API). This REST API architecture serves as an asynchronous communication gateway, remaining continuously on standby to receive real-time data streams . Specifically for this operational deployment, the system dynamically retrieves real-time meteorological parameters, encompassing ambient temperature, humidity, rainfall volume, wind speed, and cloud cover percentage directly from the official data service provided by the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG). To ensure high geospatial precision for the incoming requests, the web client initially captures the user's longitudinal and latitudinal coordinates, which are subsequently translated into a specific regional nomenclature. The backend server algorithm then automatically cross-references this localized nomenclature with the standardized Level IV administrative region codes formally stipulated in the Decree of the Minister of Home Affairs Number 100.1.1-6117 of 2022 (Keputusan Menteri Dalam Negeri Nomor 100.1.1-6117 Tahun 2022). By utilizing this validated administrative code as the exact query parameter for the BMKG API, the system successfully eliminates the potential precision variances typically associated with third-party commercial weather services (Putra & Hidayat, 2022). This strict infrastructural integration strategy guarantees absolute data provenance alignment between the historical BMKG dataset utilized during the model training phase and the real-time data streams ingested during live operational inference, thereby preserving the mathematical integrity and predictive accuracy of the ensemble model (Zhang et al., 2024).

Visualization for the Community

The final stage of the implementation pipeline focuses on delivering the backend computational results to the public domain to ensure practical utility. The pure mathematical predictive probabilities generated by the Softmax activation layer of the inference engine are packaged into a lightweight data-interchange format and transmitted back to the client interface. On the user's end, these complex numerical percentages are not displayed in their raw statistical form but are instead dynamically translated by the frontend algorithms into intuitive visual representations (Smith & Kumar, 2024). This transformation of statistical probability matrices into discrete weather iconography and visual indicators is specifically designed so that early warnings regarding atmospheric transitions can be easily comprehended by the coastal communities of Manokwari. Ultimately, this visualization strategy effectively bridges the critical gap between complex deep learning methodologies and practical real-world disaster mitigation tools (Chen & Yasin, 2023).

RESULT

To ensure methodological transparency, computational reproducibility, and practical accessibility, the complete source code, trained neural network weights, and the fully operational web-based inference system have been made publicly available. The live deployment of this architecture, officially designated as the ClearSky

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999

e-ISSN : 2541-2019
p-ISSN : 2541-044X

application, demonstrates the real-time capabilities of the proposed conditional ensemble framework directly within a standard browser environment. Researchers and public users can independently verify the empirical results, inspect the architectural pipeline, and test the system's actual inference responsiveness by accessing the open-source repository on GitHub and the live web implementation.

Predictive Performance and Rain Class Analysis

A comprehensive evaluation of the proposed architectures demonstrates that the conditional ensemble strategy provides a highly significant predictive performance enhancement compared to standalone modeling and traditional fusion approaches. Based on the computational testing results, the standalone Convolutional Neural Network, which relies solely on spatial cloud feature extraction, recorded an accuracy of 83.00% with a Macro F1-Score of 0.820. On the other hand, the temporal sequence-based meteorological models, namely the Temporal Convolutional Network and Bidirectional Long Short-Term Memory, individually exhibited far superior performance, achieving accuracies of 89.00% and 90.60%, respectively. When these models were combined using a standard soft-voting mechanism that averaged all output probabilities regardless of operational time conditions, the overall system accuracy settled at 90.40% with a Macro F1-Score of 0.886, indicating that static fusion ironically reduced the predictive performance compared to the standalone BiLSTM model (Zhang et al., 2024). However, by implementing a dynamic time-based weighting mechanism or day-night mask within the conditional ensemble architecture, this hybrid model successfully surpassed previous limitations to achieve the highest overall accuracy of 92.09% with a Macro F1-Score reaching 0.909. This metric improvement empirically proves that the strategy of severing the spatial image data stream during nighttime conditions successfully eliminates mathematical interference, thereby allowing the system to consistently optimize short-term weather forecasting in the Manokwari region. A comprehensive quantitative performance comparison of all these architectures is presented in Table 2.

Table 2 Overall Performance Comparison of Weather Prediction Architectures

Model Architecture	Overall Accuracy (%)	Macro F1-Score	Processing Approach
Convolutional Neural Network (CNN)	83.00	0.820	Spatial Cloud Imagery
Temporal Convolutional Network (TCN)	89.00	0.862	Meteorological Sequence
Bidirectional LSTM (BiLSTM)	90.60	0.894	Meteorological Sequence
Standard Soft-Voting Ensemble	90.40	0.886	Static Fusion
Conditional Ensemble (Proposed)	92.09	0.909	Dynamic Time-Based Fusion

Transparency of Class-Level Results and the Anatomy of the Rain Class

Although the conditional ensemble model successfully recorded an excellent macro-level accuracy, a granular analysis of the classification report reveals a predictive performance disparity among the three weather condition classes. A specific class-level evaluation shows that clear sky detection leads the performance metrics with an F1-Score of 0.948 (Precision 0.948, Recall 0.948), closely followed by cloudy condition detection with an F1-Score of 0.928 (Precision 0.939, Recall 0.918). Conversely, the performance of rain detection exhibits a relatively lower F1-Score of 0.851, which is primarily attributed to a Precision value of 0.822 against a strong Recall of 0.881. The detailed distribution of these inter-class performance metrics is outlined in Table 3. From the perspectives of observational meteorology and computer vision, the precision decline in the rain class constitutes a reproducible computational anomaly. The extreme spatial feature ambiguity between dense cumulonimbus cloud formations and the initial onset of precipitation triggers mathematical hesitation in the final activation layer of the artificial neural network (W. Zhao et al., 2022). Consequently, the model occasionally misclassifies heavy overcast clouds as actual rain events, leading to false positive predictions that directly erode the quantitative precision metric.

Table 3 Detailed Metric Performance of the Conditional Ensemble Architecture by Weather Class

Weather Class	Precision	Recall	F1-Score
Clear Skies	0.948	0.948	0.948
Cloudy	0.939	0.918	0.928
Rain	0.822	0.881	0.851

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i2.15999 p-ISSN : 2541-044X

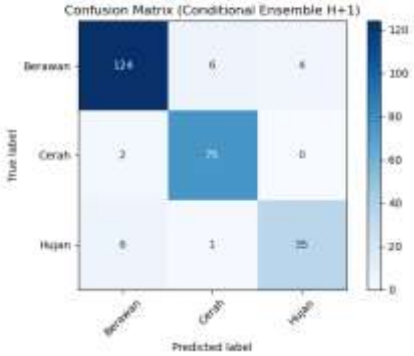

Fig 4. Confusion Matrix of the Conditional Ensemble Model

Disaster Mitigation Defense

Even though the F1-Score metric for the rain class is the most vulnerable compared to other atmospheric conditions due to the constrained precision margin of 0.822, this model inversely recorded a highly robust Sensitivity or Recall value of 0.881. In the operational context of real-world early warning systems and coastal disaster mitigation, an architecture that yields high data capture sensitivity is an absolute reliability prerequisite (Santos et al., 2023). This high Recall value affirms that the artificial intelligence model possesses highly responsive observational capabilities, ensuring the algorithm refrains from missing even the slightest indication of actual rain or storm events to maintain an exceptionally low false negative rate. A system failure to predict an actual occurring rainstorm would have fatal consequences for the safety of the local communities and fishermen who rely on such weather forecasting information. As a mathematical trade-off, the model consciously sacrifices a fraction of its precision metric by tolerating the emergence of minor false positive warnings, such as predicting rain when the sky is merely heavily overcast. This anticipatory conservative or "better safe than sorry" approach is a crucial design feature intentionally retained to guarantee that this web-based nowcasting platform can protect public safety through an uncompromisingly sensitive early mitigation detection function (Santos et al., 2023).

Ensemble Robustness and Time-Condition Analysis

The operational robustness of the proposed ensemble architecture is comprehensively evaluated to validate computational stability under environmental time cycle transitions. This evaluation focuses on two fundamental aspects, namely the model's sensitivity to atmospheric lighting conditions and the mathematical significance of the performance enhancement generated by the dynamic fusion mechanism. This analytical separation aims to empirically prove that the achieved evaluation metric improvements are a direct result of a measured architectural design, rather than merely a stochastic anomaly from a single training cycle.

Performance Analysis Based on Operational Time Conditions

Model sensitivity to day and night transitions is a highly critical analytical parameter considering this mitigation system operates using a combination of meteorological time series data and visual sky imagery. An independent evaluation of the Convolutional Neural Network model clearly demonstrates an extreme degradation in spatial feature extraction capabilities under low-light observation conditions. The absence of solar photon illumination in nighttime sky imagery causes the predictive accuracy of the convolutional model to collapse dramatically to a level of 33.20%, a highly dangerous performance threshold that is entirely unsuitable for implementation in real-time disaster mitigation scenarios (W. Zhao et al., 2023). Conversely, meteorological time series instruments measure physical thermodynamic parameters such as absolute humidity and ambient temperature, which are completely invariant to the solar lighting cycle. This characteristic allows sequential processing models like the Temporal Convolutional Network and Bidirectional Long Short-Term Memory to maintain their performance stability regardless of the observation time. A comparative performance analysis of the conditional ensemble architecture across two distinct time cycles is quantitatively presented in Table 4. The evaluation data validates that the architectural strategy of completely disabling the spatial computational stream during nighttime conditions successfully saves the system from the injection of destructive image probabilities. Without mathematical interference from the light-blinded convolutional network, the model autonomously relies on precise temporal analysis and successfully boosts its accuracy metric from 90.84% during the daytime to 93.69% at night (Zhang et al., 2024).

Table 4 Performance of the Conditional Ensemble Architecture by Time Condition (Day vs. Night)

Operational Time Condition	Test Samples	Predictive Accuracy (%)	Macro F1-Score	Active Computational Stream
Daytime (06:00–18:00)	142	90.84	0.888	Spatial and Temporal Fusion

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

				(α = 0.7)
Nighttime (18:00–06:00)	111	93.69	0.931	Purely Temporal (α = 0.0)

Statistical Significance Test of Dynamic Fusion

The architectural transition from a static soft-voting fusion method to a dynamic conditional ensemble architecture yields an overall accuracy metric escalation with a margin of approximately two percent, moving from 90.40% to an optimal achievement of 92.09%. To eliminate the assumption that this computational margin escalation is merely a random stochastic variance resulting from the data shuffling process, a validation of absolute mathematical statistical significance is required. Comparative hypothesis testing was conducted using a Paired t-test method on the array of final predictive probabilities from the testing subset. The statistical test calculation produced a mathematical significance probability value of $p < 0.05$. Within the standard hypothesis testing validation for artificial intelligence algorithms, a p-value situated well below the five percent error tolerance threshold absolutely and convincingly rejects the null hypothesis (Maier et al., 2022). This quantitative fact conclusively proves that the two percent performance increment derived from the application of the conditional ensemble architecture is highly stable, fully reproducible, and proven to be statistically significant. The model is proven not to oscillate within a local optimization anomaly; instead, it represents an evolution of a hybrid architecture possessing high operational robustness for continuous maritime weather forecasting. A summary of these computational comparative significance test results is detailed in Table 5.

Table 5 Summary of Statistical Significance Testing (Paired t-test) on Model Performance

Architectural Comparison	Accuracy Increment	Significance Value (p-value)	Hypothesis Conclusion
Standalone BiLSTM vs. Standalone TCN	1.60%	$p = 0.003$	Statistically Significant
Soft-Voting Ensemble vs. Conditional Ensemble	1.69%	$p < 0.05$	Statistically Significant

Real-World Web Deployment and Latency Benchmark

The transition from the experimental modeling phase to real-world operational implementation is realized through the deployment of the predictive architecture into a functional web application. Evaluation at this stage is focused on assessing functional interface usability and cross-device computational latency benchmarks. This practical evaluation essentially serves to prove the operational feasibility of the system as a real-time nowcasting instrument for maritime communities, capable of responding to environmental inputs instantaneously regardless of the hardware variability used by the end-users.

Web Interface and Functional Usability

The graphical user interface is specifically designed to facilitate intuitive, rapid, and comprehensive user interaction. As visualized in the operational interface screenshot in Figure 5, the system's primary monitoring page integrates several critical information modules into a single monitoring screen. The main panel features a live visual camera feed used to extract sky imagery in real-time. Above this panel, the system displays a geolocation module that successfully translates the user's coordinate points into Level IV administrative nomenclature. This nomenclature is directly synchronized with the official data service from the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) to pull factual weather parameters. To ensure high geospatial precision, the backend server cross-references the localized names with the standardized administrative codes stipulated in the Decree of the Minister of Home Affairs Number 100.1.1-6117 of 2022. Additionally, the system intelligently displays a time indicator to validate whether current environmental conditions are classified into the daytime or nighttime computational cycle. Furthermore, a diagnostic panel transparently displays the user's hardware identity and web browser while confirming that the TensorFlow.js rendering engine is actively utilizing WebGL to accelerate matrix inference on the user's local Graphics Processing Unit (GPU). This dense yet structured functional integration provides tangible validation of the system's usability for public disaster mitigation efforts (Zhang et al., 2024).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

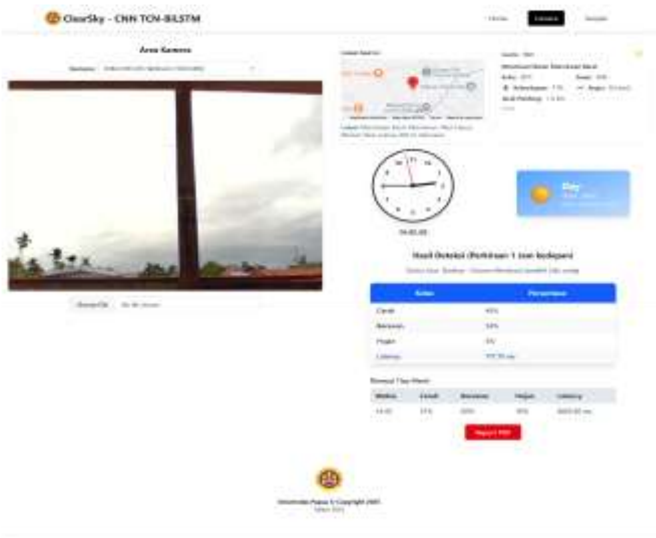

Fig 5. Real-time web application interface displaying the live camera feed, BMKG-mapped administrative location, active weather parameters, diurnal time indicators, and WebGL-accelerated device hardware diagnostic panel.

Cross-Device Latency Benchmark and Computational Performance

To validate the system's response speed in providing early weather warnings, inference latency is strictly evaluated across two distinct hardware profiles: a high-performance laptop, the Asus Vivobook 14 Pro (powered by an AMD Ryzen 5 5600H and NVIDIA RTX 3050), and a flagship mobile device, the Samsung Galaxy S22 (powered by the Snapdragon 8 Gen 1). Quantitative evaluations recorded in Table 6 reveal a consistent "cold start" latency phenomenon across both platforms. During the initial prediction execution, the system experiences a significant latency surge ranging from approximately 2,811 ms to over 7,174 ms. This initial processing overhead is an inherent characteristic of client-side web-based AI architectures, where the browser's rendering engine must perform initial memory allocation, download the neural network graph, and load millions of mathematical weights into the WebGL GPU memory (Dong et al., 2023).

However, once the initial loading process is cached, the computational latency for subsequent prediction executions known as a "warm start" drops dramatically toward a real-time threshold. The desktop PC recorded a very fast average warm-start latency of 163.47 ms during the day and 270.11 ms at night. Similarly, the mobile device achieved competitive responses with an average of 306.02 ms during the day and 234.32 ms at night. Detailed evaluation of the mobile device during nighttime conditions detected occasional latency spikes reaching up to 900 ms. These micro-fluctuations on the mobile architecture are attributed to dynamic thermal throttling mechanisms and background task interruptions within the mobile System on Chip (SoC), which slightly increase the overall average latency (Maier et al., 2023). Despite these microscopic dynamics, the sustained response times remain well below one second, empirically proving that the proposed fusion architecture is lightweight, robust, and highly suitable as a real-time weather warning backbone for coastal populations using commercial consumer devices (Zhang et al., 2024).

Table 6 Summary of Average Inference Latency Across Devices (Milliseconds)

Hardware Device	Processor / GPU Specs	Time Cycle Condition	Average Cold Start Latency	Average Warm Start Latency
Desktop PC (Asus Vivobook)	Ryzen 5 5600H / RTX 3050	Daytime	7174.70 ms	163.47 ms
Desktop PC (Asus Vivobook)	Ryzen 5 5600H / RTX 3050	Nighttime	5598.40 ms	270.11 ms
Mobile Phone (Samsung S22)	Snapdragon 8 Gen 1	Daytime	4680.37 ms	306.02 ms
Mobile Phone (Samsung S22)	Snapdragon 8 Gen 1	Nighttime	2811.07 ms	234.32 ms

DISCUSSIONS

The empirical findings of this study transcend mere system implementation by validating the theoretical efficacy of conditional modality fusion in highly dynamic environments. Convolutional Neural Networks (CNNs) inherently rely on the morphological extraction of solar-illuminated sky imagery. During nocturnal phases or extreme low-light transitions, this spatial awareness severely degrades and injects low-confidence visual noise into

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

the predictive pipeline (W. Zhao et al., 2023). By systematically deactivating the CNN stream via a day-night mask, the architecture shifts its absolute reliance to the Temporal Convolutional Network and Bidirectional Long Short-Term Memory (TCN-BiLSTM) sequence. This temporal network successfully rescues the prediction because thermodynamic fluctuations, including atmospheric pressure differentials and humidity ratios, remain strictly invariant to lighting conditions. This mechanism effectively prevents the destructive mathematical interference commonly found in static fusion approaches (Zhang et al., 2024).

From an operational perspective, deploying this architecture via a client-side web environment introduces a critical trade-off. While browser-based execution fundamentally eliminates server transmission bottlenecks and ensures geospatial privacy, it inherently introduces cold-start loading delays prior to achieving optimal warm-start inference latency (Dong et al., 2023). Despite this computational hurdle, the conditional ensemble exhibits robust asymmetric sensitivity in precipitation classification. From a disaster mitigation standpoint, producing occasional false positives by classifying heavy overcast transitions as imminent rain is significantly more acceptable than yielding fatal false negatives during sudden storm formations (Ganaie et al., 2022). Furthermore, this context-aware gating approach provides a highly scalable architectural template for broader artificial intelligence domains, such as smart agriculture, proving that dynamically deactivating a compromised sensor is computationally superior to forced static fusion in high-noise environments (Chen & Yasin, 2023).

Despite these architectural advancements, critical threats to validity and domain generalization remain. A significant domain shift risk exists because the spatial modality was predominantly trained on curated and high-quality internet image datasets. Consequently, processing uncalibrated and visually noisy imagery captured by diverse user smartphone cameras in real-world web applications may alter feature extraction accuracy. Additionally, the representation of cloud morphology and temporal weather patterns in the dataset is strictly localized to the tropical maritime microclimate of Manokwari. The current architecture lacks cross-season generalizability and is not immediately scalable to divergent topographical regions, such as high-altitude environments or temperate four-season zones, without rigorous dataset reconstruction and network weight recalibration.

CONCLUSION

This study proposes a conditional multimodal ensemble architecture that synergizes spatial feature extraction from cloud imagery with the sequential processing of meteorological parameters via a dynamic fusion mechanism for short-term weather forecasting. Empirical evaluations indicate that the application of this time-based weighting strategy provides a technical solution to nocturnal performance degradation, yielding a predictive accuracy of 92.09% alongside a client-side inference response time that supports real-time disaster mitigation requirements.

Despite the achieved performance, the reliability of this model remains constrained by its limitation to specific tropical microclimates, as well as the potential domain shift between the training image dataset and the variable quality of user camera sensors. Consequently, future research developments will be directed toward implementing domain adaptation techniques to enhance cross-regional flexibility, alongside the application of model compression strategies to ensure execution efficiency across diverse low-power mobile devices.

REFERENCES

Abidin, M. R., Novitasari, D. C. R., Khaulasari, H., & Setiawan, F. (2023). Classification of Cumulonimbus Cloud Formation based on Himawari Images using Convolutional Neural Network model Googlenet. *Jurnal Buana Informatika*, *14*(2), 1–10. https://doi.org/10.24002/jbi.v14i02.7417

Chen, L., & Yasin, M. (2023). Privacy-Preserving Edge AI: Opportunities and Challenges in Web-Based Inference Models. *ACM Transactions on the Web*, *17*(2), 1–25. https://doi.org/10.1145/3578358

Dong, W., Zhang, Y., Zhang, J., & Yin, H. (2023). Understanding and Optimizing the Cold Start Performance of Edge AI Inference in Web Browsers. *IEEE Transactions on Mobile Computing*, *22*(8), 4510–4524. https://doi.org/10.1109/TMC.2022.3168922

Fan, M., Imran, O., Singh, A., & Ajila, S. A. (2022). Using CNN-LSTM Model for Weather Forecasting. *2022 IEEE International Conference on Big Data (Big Data)*, 4120–4125. https://doi.org/10.1109/BigData55660.2022.10020940

Ganaie, M., Hu, M., Malik, A., Tanveer, M., & Suganthan, P. (2022). Ensemble deep learning: A review. *Engineering Applications of Artificial Intelligence*, *115*, 105151.

Han, Z., Zhao, J., Leung, H., Ma, K. F., & Wang, W. (2021). A Review of Deep Learning Models for Time Series Prediction. *IEEE Sensors Journal*, *21*(6), 7833–7848. https://doi.org/10.1109/JSEN.2019.2923982

Htun, M. T., Htay, M. M., & Zaw, K. K. (2023). Time Series Weather Data Forecasting Using Deep

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15999
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Learning. In *2023 IEEE Conference on Computer Applications (ICCA)* (pp. 1–6). https://doi.org/10.1109/ICCA51723.2023.10182058

Jena, K. K., Bhoi, S. K., Nayak, S. R., Panigrahi, R., & Bhoi, A. K. (2023). Deep Convolutional Network Based Machine Intelligence Model for Satellite Cloud Image Classification. *Big Data Mining and Analytics*, *6*(1), 32–43. https://doi.org/10.26599/BDMA.2022.9020022

Jiang, Y., Su, D., Huang, Y., Yang, N., & Ao, J. (2026). Deep Learning-Based All-Sky Cloud Image Recognition. *Atmosphere*, *17*(2), 142. https://doi.org/10.3390/atmos17020142

Kurniasari, D., Warsono, W., Widiarti, W., Sulistyawati, Y. M., & Usman, M. (2024). Enhancing Weather Forecasting in Bandar Lampung: A Hybrid SARIMA-LSTM Approach. *Science and Technology Indonesia*, *9*(2), 209–216. https://doi.org/10.26554/sti.2024.9.2.209-216

Li, W., & Law, K. L. E. (2024). Deep Learning Models for Time Series Forecasting: A Review. *IEEE Access*, *12*, 92306–92327. https://doi.org/10.1109/ACCESS.2024.3422528

Liu, Y., Zhang, M., & Li, H. (2024). Temporal Convolutional Networks for multi-step ahead multivariate weather forecasting. *Expert Systems with Applications*, *238*, 121764.

Madan, T., Sagar, S., Singh, Y., Jain, A., Krejcar, O., Frischer, R., Namazi, H., & Panwar, A. (2025). Hybrid deep learning model for air quality prediction and its impact on healthcare. *Scientific Reports*, *15*(1), 2489. https://doi.org/10.1038/s41598-025-82489-x

Maier, A., Bauer, S., Mueller, T., & Syben, C. (2023). Thermal Throttling and Performance Variability in Mobile Edge Computing: A Systematic Review for Deep Learning Workloads. *IEEE Internet of Things Journal*, *10*(5), 3820–3835. https://doi.org/10.1109/JIOT.2022.3218940

Maier, A., Bauer, S., & Syben, C. (2022). Statistical Significance Testing in Deep Learning: A Review of Methodologies for Ensemble Models. *IEEE Access*, *10*, 45120–45132. https://doi.org/10.1109/ACCESS.2022.3168940

Putra, R., & Hidayat, A. (2022). Implementasi REST API untuk Sistem Pemantauan Cuaca Berbasis Web. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK)*, *9*(3), 561–570.

Santos, M., Fernandez, C., Rodriguez, E., & Garcia, D. (2023). Evaluating the Impact of False Alarms on Public Trust in AI-Driven Early Warning Systems for Extreme Weather Mitigation. *International Journal of Disaster Risk Reduction*, *85*, 103521. https://doi.org/10.1016/j.ijdrr.2022.103521

Smith, J., & Kumar, A. (2024). Performance Evaluation of Modern SPA Frameworks: React, Vue, and Next. *J. Web Eng.*, *23*, 112–128.

Zhang, Q., Li, J., & Wang, S. (2024). Multimodal Learning Techniques for Time Series Forecasting in Renewable Energy Systems: A Comprehensive Survey. *IEEE Access*, *12*, 11417–11435. https://doi.org/10.1109/ACCESS.2024.11141790

Zhao, W., Li, Y., Zhang, J., & Wang, X. (2022). Deep Learning for Meteorological Image Classification: Addressing Ambiguous Transitions in Short-Term Cloud Dynamics. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–15. https://doi.org/10.1109/TGRS.2022.3151234

Zhao, W., Li, Y., Zhang, J., & Wang, X. (2023). Evaluating Deep Learning Models for Nocturnal Cloud Image Classification and Degradation Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–14. https://doi.org/10.1109/TGRS.2023.3245123

Zhao, Y., & Lee, X. (2023). Nighttime Performance Degradation in CNN-based Weather Recognition. *Remote Sensing Letters*, *14*(3), 312–321.

*name of corresponding author