

One-hour-ahead forecasting of ionogram morphology and spread-F signatures using a spatial group-wise enhanced ConvTransformer

Jinhui Cai¹ , Pengdong Gao² , Zheng Wang^{3,4,*}, Bo Wang², Guojun Wang^{3,4}, Quan Qi¹ , Chu Qiu¹, Yudong Ye⁵, Jiankui Shi^{3,4,6}, and Xiao Wang^{3,4}

¹ State Key Laboratory of Media Convergence and Communication, Communication University of China, Beijing, PR China

² Key Laboratory of Media Audio & Video (Communication University of China), Ministry of Education, Beijing, PR China

³ State Key Laboratory of Solar Activity and Space Weather, National Space Science Center, Chinese Academy of Sciences, Beijing, PR China

⁴ Hainan National Field Science Observation and Research Observatory for Space Weather, Danzhou, Hainan Province, PR China

⁵ Planetary Environmental and Astrobiological Research Laboratory (PEARL), School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, PR China

⁶ University of Chinese Academy of Sciences, Beijing, PR China

Received 12 August 2025 / Accepted 16 April 2026

Abstract—The Digisonde Portable Sounder (DPS) ionosonde at Hainan Station (19.5°N, 109.1°E; magnetic latitude: 11°N) has been monitoring ionospheric conditions since 2002, routinely recording ionospheric plasma profiles, sporadic E layers, and Spread-F structures through ionograms. A Spatial Group-wise Enhanced ConvTransformer (SGE-ConvTransformer) is proposed in this study for spatiotemporal ionospheric prediction at the Hainan station, with emphasis on Spread-F, enabling a one-hour lead time with a 15-minute sampling resolution. The SGE module optimizes semantic feature extraction from the global spatial context, dynamically recalibrating attention to prioritize information-rich regions, such as F-layer traces, over background noise. To further improve visual clarity, a super-resolution Enhanced Deep Super-Resolution (EDSR) module is integrated to sharpen the predicted ionograms. Leveraging DPS ionosonde data from 2002 to 2015, we constructed a large-scale ionogram sequence dataset comprising 36,240 Spread-F instances and 396,931 non-Spread-F instances, which were further categorized into five distinct classes. On the 2016 test set, our model achieved an average Spread-F classification accuracy of 90.05% and a correlation coefficient of 0.8115 for the predicted F-trace. Demonstrating superior robustness under disturbance conditions, the model maintained high performance during six representative geomagnetically disturbed intervals (2023–2024), achieving a classification accuracy of up to 95.69%. Furthermore, the model's generalizability was examined by applying pre-trained weights to data from low-latitude (Brazil, Peru), mid-latitude (Irkutsk), and high-latitude (Zhigansk) stations. Quantitative Spread-F Classification Accuracy (SFCA) metrics at low latitudes and qualitative visual assessments across all regions demonstrate the morphological transferability of our approach across diverse geospatial environments.

Keywords: SGE-ConvTransformer / Spatiotemporal feature / Forecasting / Ionogram morphology / Spread-F

1 Introduction

The ionosphere, a highly dynamic region of Earth's upper atmosphere extending from approximately 60 km to more than 1000 km, exerts a profound influence on radio wave propagation, Global Navigation Satellite System (GNSS) performance, over-the-horizon radar, and remote sensing applications. Its spatiotemporal variability is driven by complex solar–terrestrial interactions, including solar radiation, geomagnetic activity,

thermospheric circulation, and coupling with the lower atmosphere. These interactions give rise to phenomena spanning from regular diurnal and seasonal cycles to short-lived but operationally disruptive events, such as equatorial plasma bubbles, traveling ionospheric disturbances (TIDs), and storm-time ionospheric irregularities.

Over the decades, two principal classes of ionospheric models have been used for specification and forecasting. Numerical physics-based models, such as the National Center for Atmospheric Research thermosphere-ionosphere-electrodynamics general circulation model (NCAR/TIE-GCM; Richmond et al., 1992),

*Corresponding author: zwang@spaceweather.ac.cn

Coupled Thermosphere–Ionosphere–Plasmasphere Electrodynamics model (CTIPE; Millward et al., 2001), and SAMI3 (Huba et al., 2008), solve coupled equations for ion momentum, continuity, and energy balance. They can explicitly resolve multi-physics processes during both quiet and disturbed conditions, but require extensive geophysical inputs and are computationally intensive, limiting their real-time usability. By contrast, empirical and statistical models, exemplified by the International Reference Ionosphere (IRI; Bilitza et al., 2017) and NeQuick (Radicella & Leitinger, 2001), rely on decades of observations to produce climatological averages of ionospheric parameters such as foF2 and hmF2. They operate at low computational cost and were robust for long-term mean-state specification, but inherently smooth over short-term, nonlinear disturbances, especially during geomagnetically active periods, because they represented multi-day to monthly averages, rather than hour-scale dynamics (Pignalberi et al., 2021; Cherniak & Zakharenkova, 2016; Nava et al., 2008).

To bridge predictive skill gaps, data-driven approaches, particularly recurrent neural networks (RNNs) and their gated variants, have been widely adopted for ionospheric modeling and ionospheric parameter forecasting. For example, LSTM-based networks have been used to predict foF2 variations (Li et al., 2021; Yang et al., 2023), TEC dynamics (Ren et al., 2024), storm-time foF2 (Thammavongsy et al., 2023), ROTI for equatorial plasma bubble characterization (Zhao et al., 2025), and global TEC reconstruction in combination with IRI-2020 (Gao et al., 2024b). ConvGRU and ConvLSTM have also been applied to reconstruct two-dimensional ionospheric fields such as TEC maps (Gao et al., 2024b) and ionograms (Gao et al., 2024a). While LSTM and GRU mitigate vanishing-gradient issues and capture temporal dependencies, their iterative multi-step strategies tend to accumulate prediction errors, and their capacity to model highly nonlinear spatiotemporal structures, especially those embedded in both background diurnal cycles and transient storm-time perturbations, remains limited.

The digital ionosonde is among the most comprehensive ground-based diagnostic tools for monitoring both the background ionospheric condition and disturbed states. By transmitting vertically directed HF pulses and recording their echoes, ionosondes produce ionograms, frequency–virtual height plots from which full vertical electron density profiles for the E, F1, and F2 layers can be derived (Hargreaves, 1992). Parameters such as foF2 (F-layer critical frequency), hmF2 (F-layer peak height), h'F (bottomside virtual height), and TEC can be directly extracted (Reinisch & Huang, 1983; McNamara et al., 2008). Beyond background specification, ionograms reveal disturbed conditions such as Spread-F (SF), traces exhibiting vertical and/or horizontal spreading due to plasma irregularities, large-scale equatorial plasma bubbles, or small-scale turbulence. At low latitudes, SF is typically classified into frequency spread-F (FSF), range spread-F (RSF), strong range spread-F (SSF), and mixed spread-F (MSF) (Shi et al., 2011), each linked to distinct instability drivers, including Rayleigh–Taylor instability, gradient–drift instability, seed perturbations from atmospheric gravity waves, wind shear, and TIDs (Abdu et al., 2003; Kelley, 2009).

Accurately forecasting a full ionogram entails predicting not only scalar parameters (e.g., foF2 trends) but also the morphology of the entire ionospheric vertical structure under varying

geophysical conditions. Owing to this complexity, previous studies have typically decomposed the task into narrower sub-problems. For parameter prediction, prior efforts have focused on forecasting individual quantities such as foF2 (Li et al., 2021; Yang et al., 2023), the occurrence probability of range Spread-F (RSF) (Thammavongsy et al., 2020, 2023), total electron content (TEC) variations (Ren et al., 2024), or ROTI anomalies for equatorial plasma bubble characterization (Zhao et al., 2025). For event occurrence prediction, approaches have included equatorial Spread-F (ESF) and general SF detection using classifiers based on handcrafted features (Pillat et al., 2015; Lan et al., 2018), convolutional neural network (CNN)-based morphological recognition (Lan et al., 2020; Luwanga et al., 2022; Benchawattananon et al., 2024), and ensemble learning for regional ESF occurrence forecasting (Gao et al., 2025). More recently, generative approaches, notably GAN-based methods such as IonoGAN (Qiu et al., 2025), have reframed ionogram prediction explicitly as an image generation problem, aiming to preserve the spatial structures and morphological integrity of ionograms rather than reducing the task to a limited set of scalar outputs.

During geomagnetic storms and post-sunset instability growth phases, the ionosphere undergoes complex multi-physics coupling involving electrodynamics, plasma transport, and atmospheric wave forcing. For example, the initiation of ESF is influenced by a combination of factors, including background electron density gradients, unstable electrodynamic configurations driven by pre-reversal enhancement (PRE), penetration electric fields, and disturbance dynamo effects, as well as small-scale seed perturbations that are amplified through Rayleigh–Taylor and gradient–drift instabilities. Climatological models such as IRI and NeQuick, which average over multi-week to monthly timescales, are inherently unable to resolve the hour-scale nonlinear evolution associated with these processes, and thus can only provide probabilistic occurrence estimates rather than detailed morphological forecasts during disturbed conditions. Similarly, statistical regression and classical time-series models (e.g., ARIMA) fail to capture the coupled evolution of slow-varying background trends, such as diurnal and seasonal cycles, and rapid transient disturbances that occur on storm–substorm timescales.

In this work, we frame future ionospheric condition forecasting in terms of full ionogram sequence prediction, as ionograms provide a comprehensive representation of the ionosphere's vertical structure, including both quiet-time background parameters and disturbed features. This formulation allows the forecasting framework to encapsulate both the large-scale background state and small-scale transient irregularities, delivering a complete morphological prediction of the ionosphere. A generative sequence-to-sequence forecasting model directly operating on ionogram time series offers a unified framework to capture both the evolving background state and transient disturbed features, including the initiation and development of SF patterns. By treating ionograms as spatiotemporal fields, the model can encode physical morphology implicitly, preserving relationships between height- and frequency-dependent features, while learning from sequential dynamics over forecast horizons of several hours.

While recurrent neural network (RNN) variants such as long short-term memory (LSTM) and gated recurrent units (GRU)

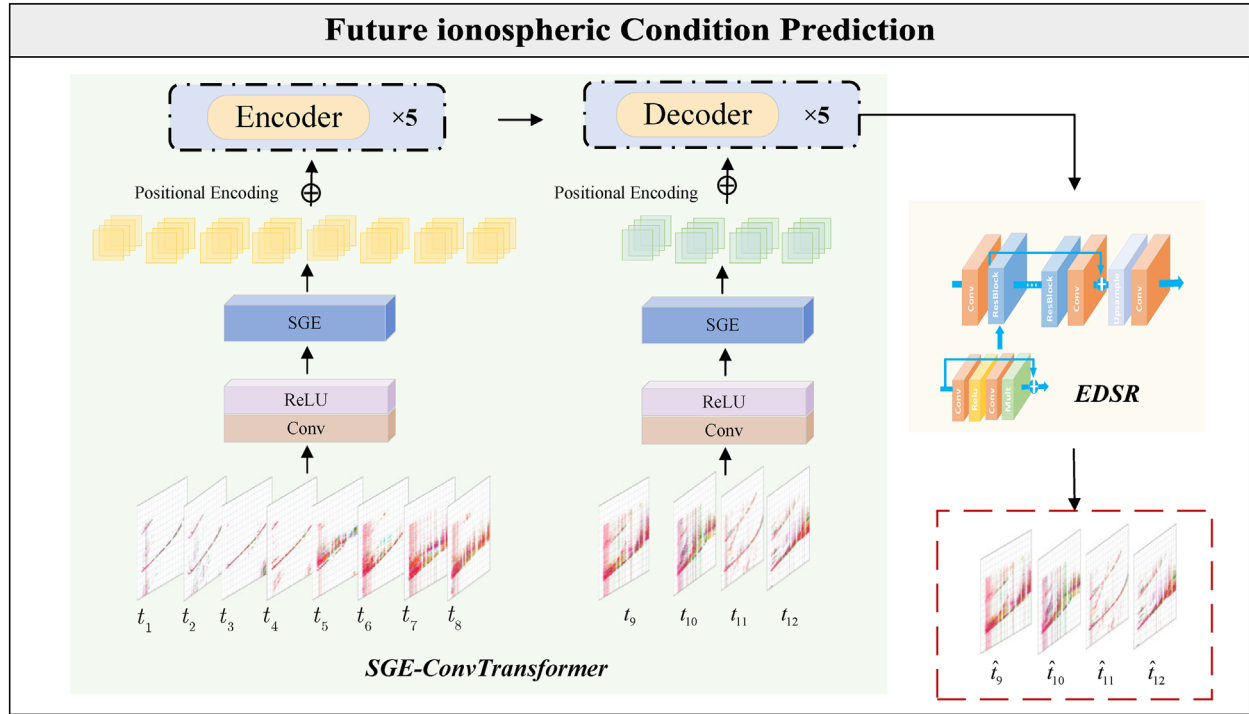


Figure 1. The whole process to predict 1-hour ahead ionograms from a 2-hour input sequence.

have been applied to ionospheric forecasting, they remain subject to inherent limitations. Iterative multi-step prediction schemes tend to accumulate errors; their inherently sequential processing hinders parallelization, and the fixed dimensionality of hidden states constrains their ability to capture the full complexity of spatiotemporal patterns. ConvLSTM/ConvGRU architectures offer improved spatial representation, yet still struggle with long-range dependency modeling. In contrast, Transformer-based architectures (Vaswani et al., 2017) leverage a self-attention mechanism to model dependencies across all time steps simultaneously, enabling efficient parallel computation and enhancing scalability, computational efficiency, and stability over extended forecasting horizons.

Therefore, we address the limitations of prior methods by proposing a Spatial Group-wise Enhanced ConvTransformer framework for ionogram sequence prediction. Unlike parameter-by-parameter LSTM forecasts, ConvTransformer (Liu et al., 2020) models ionogram sequences with multi-head convolutional attention, simultaneously capturing local high-frequency structures (important for SF morphology) and contextual evolution patterns over timescales from tens of minutes to an hour. This enables the model to detect and propagate transient perturbation signatures while maintaining background continuity. The Spatial Group-wise Enhancement (SGE) module (Li et al., 2022) selectively amplifies informative spatial groups, enhancing semantic extraction from the F-region layer, where most SF-related structures develop. To ensure predicted ionograms preserve fine-scale morphological fidelity, an Enhanced Deep Super-Resolution (EDSR) module is applied to enhance spatial clarity and mitigate blurring common in generative tasks. Crucially, comparative experiments demonstrate that the proposed model outperforms existing advanced methods, exhibiting superior predictive stability and minimizing per-

formance variance even under complex storm-time conditions. Furthermore, to verify the model's physical consistency beyond the training site, we conducted a zero-shot cross-site evaluation. Quantitative tests on two additional low-latitude stations (Fortaleza, Brazil, and Jicamarca, Peru) confirmed robust performance across distinct longitudinal sectors, while qualitative inspections on mid- and high-latitude stations (Russia) demonstrated the model's ability to reconstruct fundamental trace structures even in unseen climatological regions. By predicting the entire ionogram, the model provides a holistic forecast of the ionospheric vertical structure, offering both operational and scientific value in monitoring and understanding space weather dynamics.

2 Data and methods

At the Chinese Meridian Project low-latitude station in Hainan (Hainan Fuke, 19.5°N, 109.1°E, magnetic latitude 11°N), the ionograms have been produced since 2002 and typically have a temporal resolution of 15 minutes. Since the Spread-F events always last 1–4 hours, in this paper, we seek to accurately generate the next 4 ionograms (1 hour) to assess ionospheric variations and choose an ionogram sequence with 8 ionograms (2 hours) in the past as input. The entire workflow is illustrated in Figure 1.

2.1 Ionogram sequence dataset

To capture both short-term precursors and overall variation trends, each training sample was defined as a sequence of 12 consecutive ionograms, corresponding to a three-hour observation window. Following an approach similar to that of

Gao et al. (2024a), a total of 517,331 chronologically recorded ionograms, spanning from 2002 to 2015, have been manually labeled into five categories and processed into 75,930 sequences. The dataset comprises 10,505 frequency spread-F (FSF) sequences, 4,848 range spread-F (RSF) sequences, 14,111 mixed spread-F (MSF) sequences, 6,776 strong range spread-F (SSF) sequences, and 39,690 sequences without any spread-F signatures. These sequences were randomly divided into a training subset (80%) and a validation subset (20%). We retained an entirely independent test set, containing 47,278 ionogram sequences collected in 2016, which is completely disjoint from the training and validation sets. This test set was used to evaluate the performance of the proposed model, thereby ensuring an unbiased assessment of its generalization capability.

For the quantitative evaluation (specifically the Spread-F Classification Accuracy (SFCA) metric, defined in Section 2.4) on the 2016 test set, a unified validation strategy was adopted: the Spread-F labels for both the ground truth and predicted ionograms were determined by the same pre-trained automatic classifier (Wang et al., 2023). While this strategy provides an objective and scalable benchmark for relative model comparison, we explicitly acknowledge that it does not constitute an independent physical ground truth. Consequently, there is an inherent risk of classifier bias propagation, as the absolute evaluation performance is implicitly bounded by the auxiliary classifier's limitations. Therefore, the SFCA should be interpreted primarily as a measure of relative semantic consistency under a fixed automated standard, rather than a direct measurement of absolute physical forecasting skill. To mitigate this bias and ensure scientific rigor, our evaluation framework does not rely solely on SFCA; it is robustly complemented by the Absolute Value of the Correlation Coefficient for the F-trace (AVCC-F) metric (which directly quantifies fine-scale vertical structure correlation, also described in Section 2.4) and comprehensive morphological visual inspections.

2.2 Spatial group-wise enhanced ConvTransformer

The ConvTransformer model (Liu et al., 2021) integrates the representational power of Convolutional Neural Networks (CNNs) with the global dependency modeling capability of self-attention mechanisms, making it particularly well-suited for longer-range spatiotemporal prediction tasks. Its architecture comprises three primary components: an encoder, a decoder, and a positional encoding module. The encoder leverages a multi-head convolutional attention mechanism in conjunction with a two-dimensional convolutional feed-forward network to capture both local spatial structures and long-range temporal dependencies within the input image sequence, thereby producing a compact yet information-rich spatiotemporal representation. The decoder processes this latent representation through additional convolutional attention and 2D convolutional feed-forward layers, before generating the final predicted images via a convolutional layer with sigmoid activation to constrain pixel intensities to the valid range. To preserve temporal order information that is otherwise lost in pure convolutional processing, the positional encoding module employs sine and cosine functions of varying frequencies to embed temporal indices directly into the feature space, enhancing the model's ability

to learn sequential dependencies. This design enables ConvTransformer to simultaneously exploit fine-grained spatial detail and long-horizon temporal correlations, addressing limitations inherent in purely recurrent or convolutional architectures.

Spatial Group-wise Enhance (SGE) was introduced by Li et al. (2022). This method was designed to enhance feature learning in convolutional neural networks (CNNs) by grouping spatial features and applying group-wise normalization and attention mechanisms. The core idea of SGE is to divide the feature map into multiple groups along the channel dimension and process each group independently. This allows the model to focus on distinct spatial regions, capturing fine-grained details and improving the representation of local features. By combining group-wise processing with attention mechanisms, SGE enhances the model's ability to emphasize important spatial information while maintaining global context. This makes it particularly effective for tasks requiring precise spatial feature extraction, such as image segmentation, object detection, and, in this study, ionogram prediction.

In our research, particular emphasis is placed on the F-layer region of ionograms, with the objective of analyzing and characterizing the developmental trends of the ionospheric F-layer in the predicted ionogram sequences. This focus enables the detection of Spread-F phenomena and the identification of their specific morphological types. Consequently, the proposed network is explicitly designed to allocate greater representational capacity to the F-layer region. To this end, the Spatial Group-wise Enhancement (SGE) module is integrated into the ConvTransformer architecture as a natural extension. As illustrated in Figure 2, within the integrated framework, the SGE module selectively amplifies salient local spatial features, particularly in the F-layer, while the multi-head attention mechanism of the ConvTransformer captures long-range dependencies and temporal correlations across the sequence. This synergistic combination equips the framework with the capability to effectively model multi-hour spatiotemporal context, thereby enhancing its suitability for high-fidelity ionogram prediction tasks.

2.3 Super-resolution EDSR model

Although the integration of the SGE module with the ConvTransformer yields measurable improvements across multiple evaluation metrics, the visual clarity of the predicted ionograms remains insufficient for precise interpretation. To address this limitation, a post-processing stage based on the Enhanced Deep Super-Resolution (EDSR) network (Lim et al., 2017) is incorporated into our framework. EDSR employs a deep residual learning strategy to enhance spatial resolution and recover fine-grained structural details. In our implementation, the network is trained directly on Hainan ionograms, leveraging their relatively simple visual patterns to achieve convergence within a limited number of training epochs (Gao et al., 2024a). By refining the intermediate outputs of the SGE-ConvTransformer, the EDSR module produces high-clarity ionograms that preserve both large-scale morphology and subtle layer structures. This enhancement not only improves the perceptual quality of the forecasts but also provides a more reliable basis for subsequent analyses of ionospheric phenomena, including the identification and classification of Spread-F events.

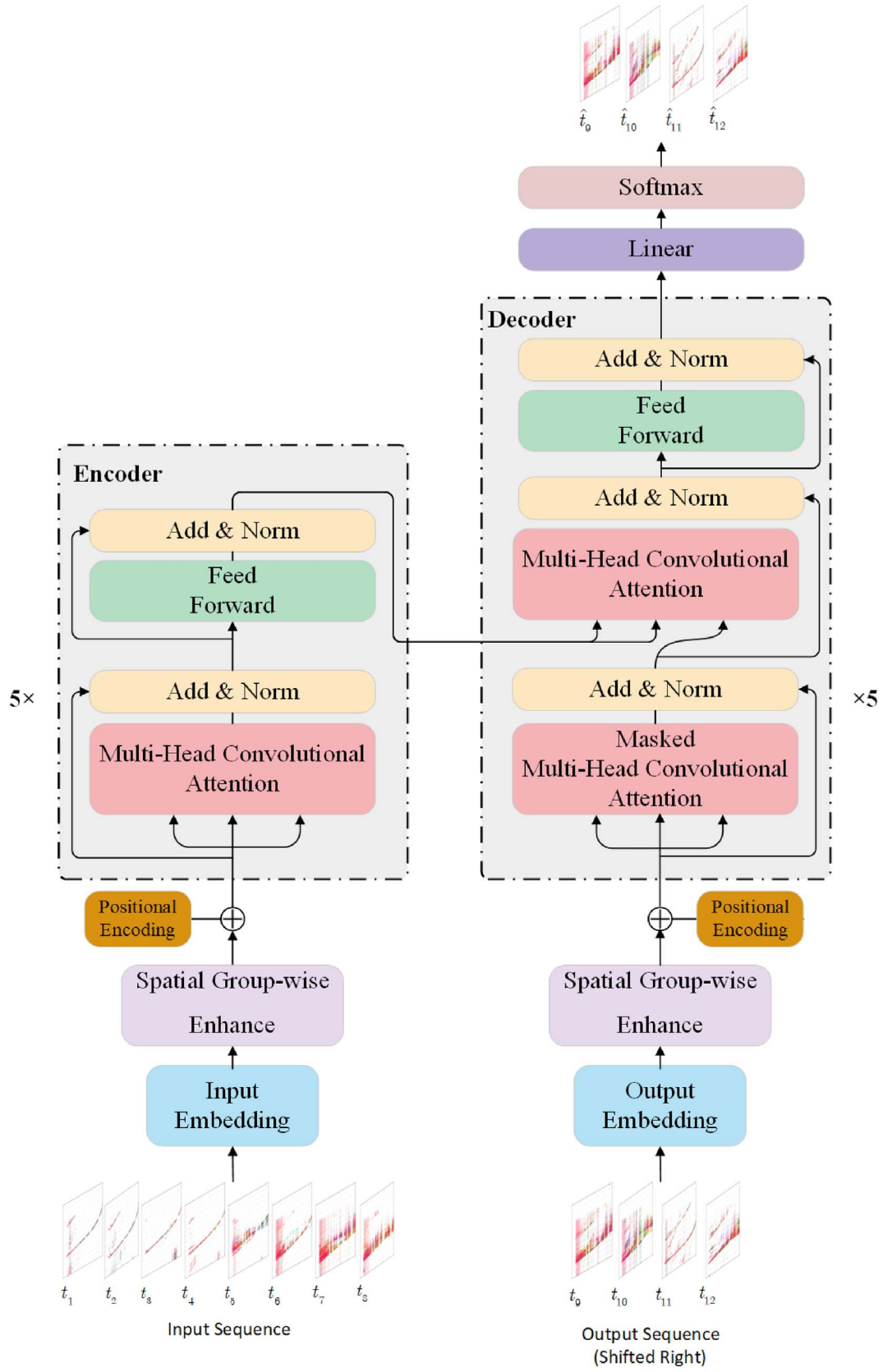


Figure 2. The architecture of SGE-ConvTransformer.

2.4 Quality evaluation metrics

In assessing the quality of the predicted ionograms, our evaluation framework goes beyond conventional image reconstruction metrics such as mean squared error (MSE), structural similarity index (SSIM), and peak signal-to-noise ratio (PSNR), which, while valuable for measuring generic visual fidelity, do not fully capture the geophysical relevance of ionospheric data. To bridge this gap, we introduce two ionospheric domain-specific metrics: Spread-F Classification Accuracy (SFCA) and the Absolute Value of the Correlation Coefficient for the F-trace (AVCC-F). These metrics are designed not only to quantify reconstruction accuracy but also to directly reflect the scientific utility of the predicted ionograms in space weather research and operational monitoring.

To evaluate the semantic consistency between the predicted and ground truth ionograms, we introduce a metric termed Spread-F Classification Accuracy (SFCA). This metric utilizes a pre-trained classification model (Wang et al., 2023), serving as an auxiliary discriminator, to determine the occurrence and specific type of Spread-F. Formally, we define SFCA using an indicator function $\mathbb{I}(\cdot)$. For a given sample, if the classification result of the predicted ionogram matches that of the ground truth, implying that the model correctly captured the Spread-F characteristics, the function returns 1; otherwise, it returns 0. Consequently, SFCA is defined as the proportion of total samples for which the auxiliary discriminator yields identical classification labels for both the predicted and ground truth ionograms. Mathematically, it is expressed as:

$$\text{SFCA} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(C(\hat{y}_i) = C(y_i)), \quad (1)$$

where N denotes the total number of test samples, y_i represents the ground truth ionogram, and $\hat{y}_i$ denotes the corresponding predicted ionogram. Additionally, $C(\cdot)$ refers to the pre-trained auxiliary discriminator, which determines the occurrence and specific type of Spread-F based on the class with the highest output probability score (argmax). And $\mathbb{I}(\cdot)$ is the indicator function that equals 1 if the classification results are consistent, and 0 otherwise. Furthermore, the final reported SFCA represents the average accuracy calculated frame-by-frame across the entire 4-step prediction window.

In contrast, the AVCC-F metric quantifies the correlation between the F-traces extracted from the predicted ionograms and the ground truth observations. Here, the F-trace is defined as the lower envelope of the ordinary wave (O-mode) echoes. Given that the frequency components (abscissa) are identical for both traces, the comparison is effectively reduced to an evaluation of their virtual height vectors. Let $\mathbf{h}^{GT}$ and $\mathbf{h}^{Pre}$ denote the virtual height vectors of the F-trace for the ground truth and predicted ionograms, respectively. Consequently, AVCC-F is defined as the absolute value of the Pearson correlation coefficient (ρ) between these two vectors:

$$\text{AVCC-F} = |\rho(\mathbf{h}^{GT}, \mathbf{h}^{Pre})| = \left| \frac{\text{Cov}(\mathbf{h}^{GT}, \mathbf{h}^{Pre})}{\sigma(\mathbf{h}^{GT}) \cdot \sigma(\mathbf{h}^{Pre})} \right|, \quad (2)$$

where $\text{Cov}(\cdot)$ denotes the covariance and $\sigma(\cdot)$ represents the standard deviation. We define AVCC-F using the absolute

Pearson correlation to standardize the range to $[0, 1]$, thereby quantifying the magnitude of structural similarity. This formulation prevents statistical cancellation during aggregation and ensures scale consistency with other metrics like SFCA. The AVCC-F metric specifically evaluates the reconstruction accuracy of the F-layer's fine-scale vertical structure. This ensures that the model not only captures large-scale features but also preserves high fidelity in subtle layer characteristics, particularly during ionospheric quiet periods.

Together, SFCA and AVCC-F extend evaluation beyond image-level similarity, enabling a scientifically grounded assessment that balances model performance in both disturbed and quiet ionospheric regimes. By incorporating these metrics into the testing process, the framework is guided toward producing outputs that are not only visually convincing but also scientifically meaningful, thereby ensuring the reliability of subsequent analyses and the credibility of conclusions drawn from the predicted ionograms.

3 Experimental results

In this section, we detail the experimental configuration and full training process for the Spatial Group-wise Enhanced ConvTransformer. The investigation commenced with a systematic exploration of the key hyperparameters in the baseline ConvTransformer, with the objective of maximizing predictive skill across both quiet and disturbed ionospheric conditions. Once the optimal configuration was established, the Spatial Group-wise Enhancement (SGE) module was incorporated into the architecture. Its contribution was rigorously evaluated using domain-relevant assessment strategies, including expert-driven qualitative inspection of ionogram morphology and quantitative metrics encompassing both general image reconstruction measures (e.g., MSE, SSIM, PSNR) and ionospheric-specific indicators (e.g., SFCA, AVCC-F), and diagnostic assessments of classification reliability (e.g., Confusion Matrix, F1-score). This integrated evaluation framework ensured that the observed performance improvements were not only statistically significant but also scientifically meaningful within the context of ionospheric forecasting.

All experiments were performed on a Dell workstation equipped with dual NVIDIA GeForce RTX 3090 Ti GPUs. The principal hyperparameters employed in the experiments are summarized in Table 1. The implementation was developed using the PyTorch framework, building upon the publicly available codebase.¹

Each ionogram in the input sequence is represented as a $224 \times 224 \times 3$ array, where 224×224 denotes the spatial resolution and the final dimension of “3” corresponds to the RGB color channels. The “Batch_Size” was set to 6, a value determined by the available GPU memory capacity, balancing computational efficiency and memory constraints. Within the ConvTransformer architecture, “Num_layers” specifies the number of stacked encoder and decoder layers, directly influencing the network’s capacity to model complex spatiotemporal dependencies. “D_layer” denotes the dimensionality of the hidden representations in each layer, which governs the expressiveness of the learned features. Both “Num_layers” and “D_layer”

¹ <https://github.com/iamrakesh28/Video-Prediction/>

Table 1. Hyperparameters for Our SGE-ConvTransformers.

Parameters	Value	Unit/Description
Input_Shape	224*224*3	pixels × pixels × channels
Batch_Size	6	number of ionogram sequences
Num_Layers	2/3/4/5	number of layers
D_layer	8/16/32	number of dimensions
Num_heads	4	number of heads
Filter_size	3*3	pixels × pixels
Epoch	50	epochs
Optimizer	Adam	—

Table 2. SFCA comparison of different hyperparameter combinations (%).

	Hyperparameters	Average	Pred 1	Pred 2	Pred 3	Pred 4
D_layer	2_8_4_3×3	88.64	88.75	88.64	88.60	88.57
	2_16_4_3×3	88.67	88.77	88.67	88.64	88.62
	2_32_4_3×3	88.78	88.88	88.85	88.71	88.67
Num_layers	3_32_4_3×3	88.84	89.03	88.80	88.77	88.76
	4_32_4_3×3	88.86	88.82	88.84	88.90	88.87
	5_32_4_3×3	88.99	88.95	89.02	88.98	89.02

are critical hyperparameters whose optimal values were identified through a series of comparative experiments. The parameter “*Num_heads*” refers to the number of attention heads in the multi-head convolutional attention mechanism. Increasing the number of heads allows the model to capture a broader range of spatial–temporal patterns in parallel, albeit with higher computational cost. The “*Filter_size*” was set to 3×3 , corresponding to the kernel size of the convolutional layers, a widely adopted choice that balances local feature extraction with computational efficiency. The networks were trained for 50 epochs using ionogram sequences from the Hainan station dataset. Gradient updates were conducted using the “*Adam*” optimization algorithm in combination with the chosen optimizer, facilitating efficient backpropagation and convergence.

As mentioned above, the ionogram sequences from 2016 were used as an independent test set to evaluate the performance of the proposed SGE-ConvTransformer in all subsequent experiments, ensuring a fair assessment of its generalization capability.

3.1 Two key hyperparameters

The number of encoder–decoder layers (*Num_layers*) and the hidden dimension of the ConvTransformer layers (*D_layer*) are two critical hyperparameters in the ConvTransformer framework (Liu et al., 2020). A series of experiments was first conducted to identify the optimal combination of these parameters. As shown in the “Hyperparameters” column of Table 2, the four entries correspond to *Num_layers*, *D_layer*, *Num_heads*, and *Filter_size*, respectively. The results indicate a clear upward trend in the SFCA metric as the hidden-layer dimension increases. The configuration “2_32_4_3 × 3” yielded the highest SFCA score of 88.78%. Furthermore, fixing the hidden-layer dimension at 32 reveals that increasing the number of encoder–decoder layers leads to additional improvements in SFCA, consistent with the widely recognized principle in deep learning that deeper architectures generally exhibit stronger feature-extraction capability. Based on these results,

the optimal hyperparameter configuration for the baseline ConvTransformer was determined to be five encoder–decoder layers, a hidden-layer dimension of 32, four attention heads in the multi-head attention mechanism, and 3×3 convolution kernels in the convolutional layers. Under this setting, the baseline ConvTransformer achieved an SFCA of 88.99%.

3.2 The SGE module

In the SGE method, the number of groups (*G*) plays a critical role in determining the granularity of distinct semantic sub-features. When the total number of channels is fixed, an excessively large *G* reduces the dimensionality of the sub-features within each group, resulting in weaker and less discriminative feature representations for individual semantic responses. On the other hand, an insufficient number of groups may constrain the diversity of semantic features, limiting the model’s ability to capture nuanced spatial patterns. Therefore, selecting an optimal *G* is essential to balance feature richness and representation strength, ensuring robust and diverse semantic feature extraction. In the following experiments, three configurations of the Spatial Group-wise Enhance module were tested: the first configuration excluded the SGE module entirely, the second divided the embedded image features into 4 groups, and the third divided them into 16 groups.

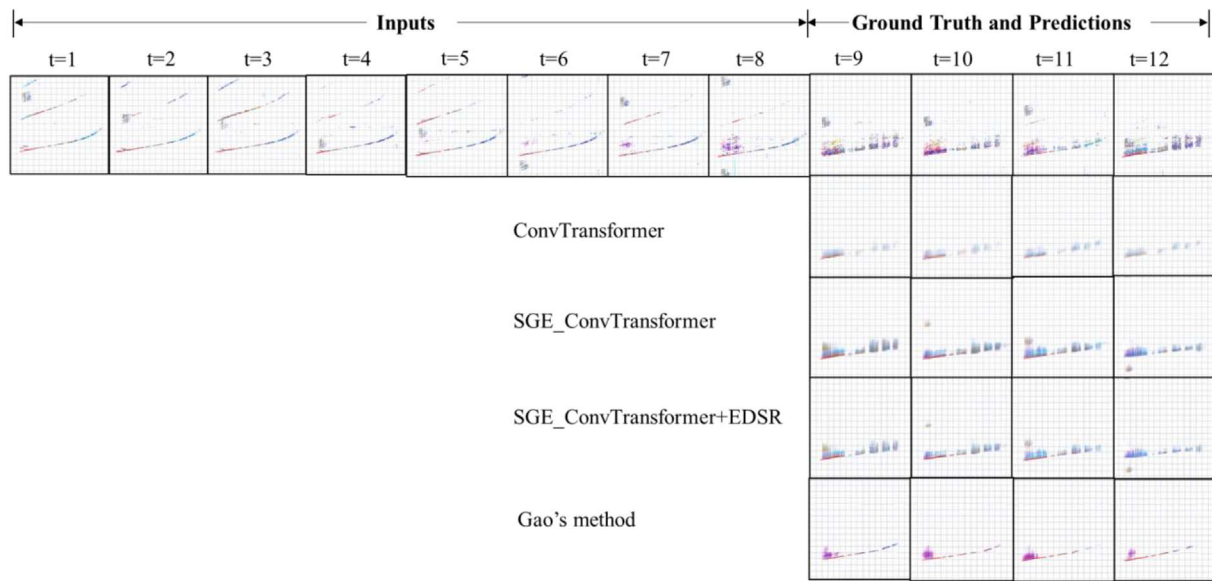
To rigorously assess the contribution of the SGE module to the ConvTransformer architecture, we employed both ionosphere-specific performance indicators and conventional image quality metrics. Table 3 presents the results for two key ionospheric metrics: SFCA and AVCC-F. Among all tested configurations, the 16-group SGE-ConvTransformer achieved the best overall performance. This configuration not only demonstrated an enhanced ability to predict the occurrence of Spread-F phenomena and accurately classify their morphological types, but also produced more realistic F-trace reconstructions within the F-layer region. The observed improvement in SFCA reflects the model’s strengthened capability to detect and differentiate fine-scale ionospheric irregularities, while the gain in AVCC-F

Table 3. Comparison of three ConvTransformer networks.

Hyperparameters	SFCA (%)					AVCC-F	
	Average	Pred 1	Pred 2	Pred 3	Pred 4	Mean	Variance
5_32_4_3×3	88.99	88.95	89.02	88.98	89.02	0.6704	0.0866
5_32_4_3×3_4G	89.03	89.05	89.03	89.04	89.00	0.6794	0.0860
5_32_4_3×3_16G	89.22	89.22	89.28	89.26	89.13	0.7013	0.0846

Table 4. Comparison of three ConvTransformer networks.

Hyperparameters	MSE↓	SSIM↑	PSNR↑	LPIPS↓	Mutual Information↑
5_32_4_3×3	0.0031	0.9100	25.7352	0.0953	0.9817
5_32_4_3×3_4G	0.0029	0.9180	26.0870	0.0931	1.0427
5_32_4_3×3_16G	0.0029	0.9384	26.1525	0.0836	1.0700

**Figure 3.** Visual comparison of predicted ionograms by convtransformer variants and Gao's method.

indicates improved fidelity in representing the background F-layer structure, both of which are critical for understanding the physical evolution of ionospheric disturbances. Notably, for the AVCC-F metric, critical for quantifying the fidelity of F-layer structure reproduction under both quiet and disturbed ionospheric conditions, the 16-group SGE-ConvTransformer improved the score from 0.67 (baseline) to 0.70, representing a gain of more than 0.03.

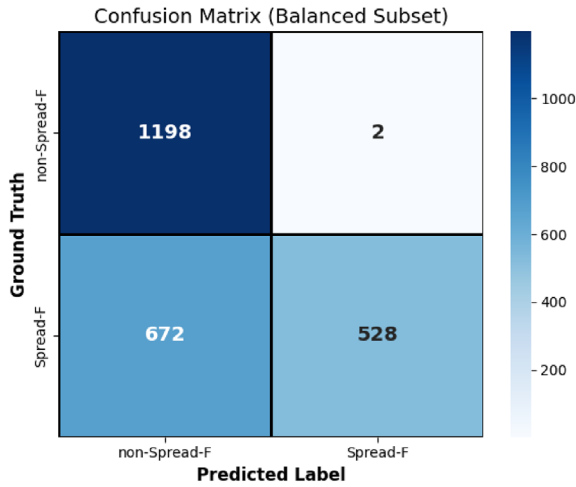
As presented in Table 4, a suite of conventional image quality metrics, including mean squared error (MSE), structural similarity index measure (SSIM), peak signal-to-noise ratio (PSNR), Learned Perceptual Image Patch Similarity (LPIPS), and mutual information, was employed to assess the clarity, structural fidelity, and overall visual quality of the predicted ionograms. The results clearly demonstrate that the 16-group SGE-ConvTransformer consistently outperforms all other configurations across every evaluated parameter, confirming that partitioning the embedded features into 16 groups yields more effective feature representation than using only 4 groups. Consequently, the 16-group configuration was selected as the optimal architecture for our ionogram prediction tasks.

3.3 Super-resolution

In terms of visual quality, the integration of the SGE module substantially enhances the performance of the ConvTransformer. As shown in Figure 3, the predicted ionograms in the third row on the right exhibit notably greater structural fidelity and clarity compared with those produced by the baseline model without SGE. Nevertheless, when compared with the corresponding ground truth (GT) ionograms, the F-layer region still presents suboptimal sharpness, which may limit the precision of subsequent scientific analyses. To address this limitation, a super-resolution post-processing step, based on a pre-trained EDSR network (Gao et al. 2024a), was applied to the SGE-ConvTransformer outputs. The resulting ionograms (the fourth row in Figure 3) exhibit markedly improved morphological fidelity within the F-layer region, along with enhanced overall image clarity. The bottom row of Figure 3 presents the results obtained by the Gao et al. (2024a) model. Since their ConvGRU-based approach is sequential and originally limited to predicting only the next 2 frames, we employed a recursive strategy to extend the prediction to the four frames.

Table 5. Comparison of three ConvTransformer networks after super-resolution.

Hyperparameters	SFCA (%)					AVCC-F	
	Average	Pred 1	Pred 2	Pred 3	Pred 4	Mean	Variance
5_32_4_3×3	89.04	89.14	89.07	88.99	88.96	0.7972	0.0844
5_32_4_3×3_4G	89.13	89.15	89.13	89.14	89.10	0.8060	0.0710
5_32_4_3×3_16G	90.05	90.17	90.17	90.06	89.80	0.8115	0.0781

**Figure 4.** Confusion matrix for the binary detection task (SF vs non-SF) evaluated on the class-balanced test subset.

The visual comparison indicates that our SGE-ConvTransformer can generate slightly richer echo content than the recursive baseline. Even at the initial forecasting steps ($t = 9$ and $t = 10$), our model preserves more morphological details compared to Gao’s model. This disparity becomes more distinct in the subsequent frames ($t = 11$ and $t = 12$). As the prediction time extends, Gao’s model shows signs of rapid echo content loss. This degradation is attributed to its recursive generation architecture, where prediction errors inevitably accumulate step-by-step. In comparison, while our model also experiences minor signal attenuation, the ConvTransformer’s parallel processing capability allows it to model global temporal dependencies simultaneously. This approach alleviates the error propagation inherent in recursive schemes, thereby preserving significantly more morphological details across the entire sequence.

In addition to the vision-based qualitative comparisons shown in Figure 3, quantitative evaluations were also conducted on the 2016 test set using two ionospheric-specific metrics, with results summarized in Table 5. The 16-Group SGE-ConvTransformer model maintained superior overall performance across all configurations. By comparing Table 3 with Table 5, the following conclusion is obvious. For the SFCA metric, while absolute improvements appeared modest, all three models with super-resolution outperformed their original counterparts. Specifically, the model with hyperparameters “5_32_4_3×3_16G” achieved a 0.83% SFCA enhancement, increasing from 89.22% to 90.05% after super-resolution implementation. More notably, the annual averaged AVCC-F exhibited substantial improvement, rising from 0.7013 (variance = 0.0846) to 0.8115 (variance = 0.0781) after super-resolution. This significant AVCC-F enhancement is physically meaningful, as sharper

and more structurally accurate ionograms enable more reliable F-trace extraction, particularly under complex disturbed conditions. Such improvements directly benefit space weather monitoring and ionospheric morphology studies, providing researchers with higher-quality data for both automated algorithms and manual interpretation.

Although the absolute numerical differences for the classification metric (SFCA) in Tables 2–3 appear modest, these gains are highly meaningful within the context of ionogram forecasting. Because the baseline ConvTransformer already achieves strong predictive performance (approaching 89% accuracy), the model encounters a “ceiling effect” where large absolute improvements are inherently difficult to extract from the morphologically ambiguous transitional Spread-F cases. More importantly, the proposed SGE enhancements demonstrate systematic advantages across multiple complementary evaluation criteria, including SSIM, PSNR, and LPIPS (Table 4), indicating a robust stabilization of the generative quality rather than incidental fluctuations. Furthermore, while generic image classification gains are incremental, the physics-specific structural metric reveals a much more profound impact. The incorporation of the super-resolution module yields a substantial leap in AVCC-F (from 0.7013 in Table 3 to 0.8115 in Table 5). This confirms that the proposed architecture effectively preserves the fine-scale morphological structures, particularly the F-region trace and Spread-F boundaries, that are critical for subsequent geophysical interpretation and space weather monitoring.

3.4 Classification performance diagnosis

While the SFCA and AVCC-F metrics confirm the model’s overall efficacy, the extreme class imbalance in the 2016 test set (where “Strong Range” samples constitute less than 1% of the data) can mask the true performance on minority classes. To address this and ensure a rigorous assessment that reflects both operational reality and diagnostic precision, we employed a diagnostic evaluation strategy using a class-balanced subset.

We constructed this balanced subset by randomly sampling 300 images for each Spread-F subtype and 1,200 “non-Spread-F” images, maintaining a strict 1:1 ratio between Spread-F (1,200 samples) and non-Spread-F (1,200 samples) events. This setup allows for a clearer visualization of the model’s discriminative capability across different morphological types without the bias of the majority class.

Figure 4 presents the Confusion Matrix for the binary detection task (Spread-F vs. non-Spread-F) evaluated on this balanced subset. A key finding is the model’s exceptional Precision for the “Spread-F” class, reaching 99.62% with only 2 false positives misclassified as Spread-F out of 1,200 quiet samples. In the context of space weather forecasting, minimizing false alarms is often critical to prevent “alarm fatigue” among human operators.

However, as highlighted by the confusion matrix, this high precision comes at the cost of limited sensitivity (Recall = 44.0%), where the number of false negatives (672) significantly exceeds the true positives (528). This clear precision-recall trade-off characterizes the SGE-ConvTransformer as a low-false-alarm but sensitivity-limited predictor. The elevated false-negative rate is fundamentally tied to both algorithmic characteristics and the physics of early-stage ionospheric irregularities. Driven by the MSE loss function, the generative model tends to smooth out faint, high-frequency textural details in favor of predicting the structural mean. Consequently, weak, onset-stage, or morphologically ambiguous Spread-F signatures are inherently difficult to reproduce faithfully one hour in advance and are often filtered out, leading to missed detections.

This inherent morphological ambiguity and transitional overlap between sub-classes also extend to the detailed 5-class classification, where the model yielded an overall F1-score of 0.62. While this metric might appear modest when compared to benchmark results in generic artificial intelligence tasks, it represents a highly competitive and physically meaningful achievement for Spread-F subtype classification and broader ionospheric studies. While the model performs robustly in identifying the “Frequency” type, distinguishing between specific transitional subtypes (e.g., “Range” vs. “Mixed”) remains a significant hurdle. This observation aligns with the conclusions drawn by Wang et al. (2023), who noted that “it is relatively easy to determine whether the ionograms show the presence of Spread-F, but is more challenging to classify the type of Spread-F based on the image features only”.

Taken together, these diagnostic results confirm that while capturing every marginal event and achieving perfect fine-grained typing remain open challenges, the SGE-ConvTransformer excels at the most critical operational task. When it generates a Spread-F alert, the prediction is highly credible, validating its practical utility for operational warning systems.

4 Discussion and analysis

While the experimental results and diagnostic analyses in Section 3 demonstrated the superior quantitative performance of the SGE-ConvTransformer on the Hainan dataset, establishing a model's operational viability requires a more rigorous validation beyond standard metrics. In this section, we extend the evaluation to assess the model's performance from multiple dimensions: morphological fidelity, robustness under extreme conditions, spatial generalization capability, and physical interpretability.

4.1 Fine-grained morphological fidelity and example analysis

Given that the EDSR module contributes significantly to the improvement in AVCC-F (from ~ 0.70 to 0.81), it is critical to verify that this gain stems from accurate signal restoration rather than the “hallucination” of non-existent structures. To rigorously validate this, we conducted both visual and quantitative fidelity checks.

Figure 5 presents a visual overlay of F-traces extracted from the Ground Truth (Green), Pre-EDSR (Blue), and Post-EDSR (Red) ionograms. As illustrated in Figures 5(a) and 5(b), the Post-EDSR traces demonstrate tighter adherence to the Ground Truth compared to the traces before super-resolution. A notable

improvement is observed at the high-frequency tail of the traces, where the Pre-EDSR predictions tend to be truncated due to low-resolution blurring. The EDSR module effectively recovers these faint signals, extending the trace to match the physical ground truth without generating spurious artifacts or outliers. Furthermore, we quantified this improvement using the Mean Absolute Error (MAE) and Dynamic Time Warping (DTW) distance across the entire 2016 test set. The application of EDSR reduced the average MAE from 23.77 km to 23.62 km and the DTW distance from 2749.15 to 2645.62. These results confirm that the super-resolution module enhances the physical fidelity of the ionograms by sharpening genuine features, thereby facilitating more accurate F-trace extraction for downstream analysis.

To provide an intuitive visualization of the model's capabilities and limitations, Figure 6 illustrates representative predictions across distinct ionospheric conditions, ranging from typical quiet periods to complex failure examples. Overall, visual inspection of Figure 6 reveals that the model excels in capturing the principal morphological dynamics of ionograms while filtering out transient artifacts, including multi-hop echoes and Sporadic E (Es) layers. Specifically, under non-Spread-F conditions, the model accurately reconstructed the complete F-layer echo traces in the ionograms, demonstrating high localization accuracy. In complex scenarios involving Frequency Spread-F (Fig. 6b) and Range Spread-F (Fig. 6c), the SGE-ConvTransformer effectively preserves the global semantic structure and distinct spread patterns. Although the generative process introduces a degree of texture smoothing due to the loss of high-frequency details, this trade-off significantly enhances visual clarity by eliminating background clutter, ensuring that the primary ionospheric features remain distinct and identifiable.

The bottom row, Figure 6(d), presents a challenging sample that highlights certain limitations of the model. In the prediction of this sequence, although the model successfully captures the dominant trends in the morphological evolution of the ionograms, yielding shapes that closely resemble the ground truth, there is a significant discrepancy between the predicted colors and those of the ground truth. Given that color in these ionograms encodes the directionality (Angle of Arrival) of received echoes, this discrepancy signifies a distortion of directional information. Such a distortion poses a risk of misleading human experts during the manual interpretation of signal sources.

The observed chromatic discrepancy can likely be attributed to two primary factors. First, the architectural design of the SGE module and convolutional operations is primarily aimed at extracting “semantic features” and the “global spatial content.” This design implicitly guides the model to prioritize structural information (morphology) over textural details (i.e., color information). Second, the use of the MSE loss function during training drives the model to minimize the aggregate prediction error. Consequently, in the face of uncertainty, the model tends to predict the statistical mean of the samples, resulting in a smoothing effect on the color values. Therefore, we explicitly advise that while the proposed SGE-ConvTransformer is highly effective for the morphology-oriented forecasting of Spread-F evolution, it is not yet suitable for operational applications or scientific analyses requiring reliable Angle-of-Arrival interpretation. Users must exercise caution and avoid relying on the predicted color mapping for directional physical studies.

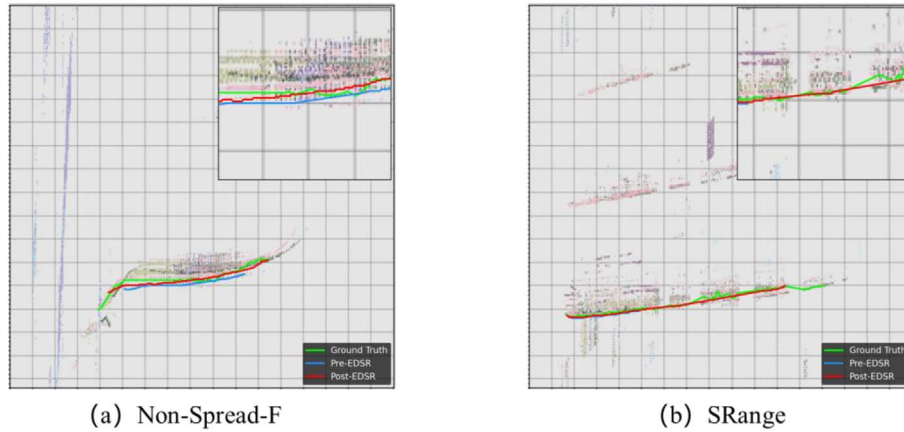


Figure 5. Detailed visual comparison of the extracted F-traces before and after the application of the EDSR super-resolution module for two representative ionospheric scenarios: (a) a quiet Non-Spread-F condition and (b) a disturbed Strong Range (SRRange) Spread-F condition. The background images display the corresponding ground-truth ionograms. Within each panel, the solid lines represent the ground-truth F-trace (green), the SGE-ConvTransformer prediction prior to EDSR refinement (blue), and the final high-fidelity prediction after EDSR post-processing (red). The insets provide a magnified view of the high-frequency tail regions, illustrating the model's enhanced capability to recover faint echo signals and extend the trace to match physical observations without generating spurious artifacts.

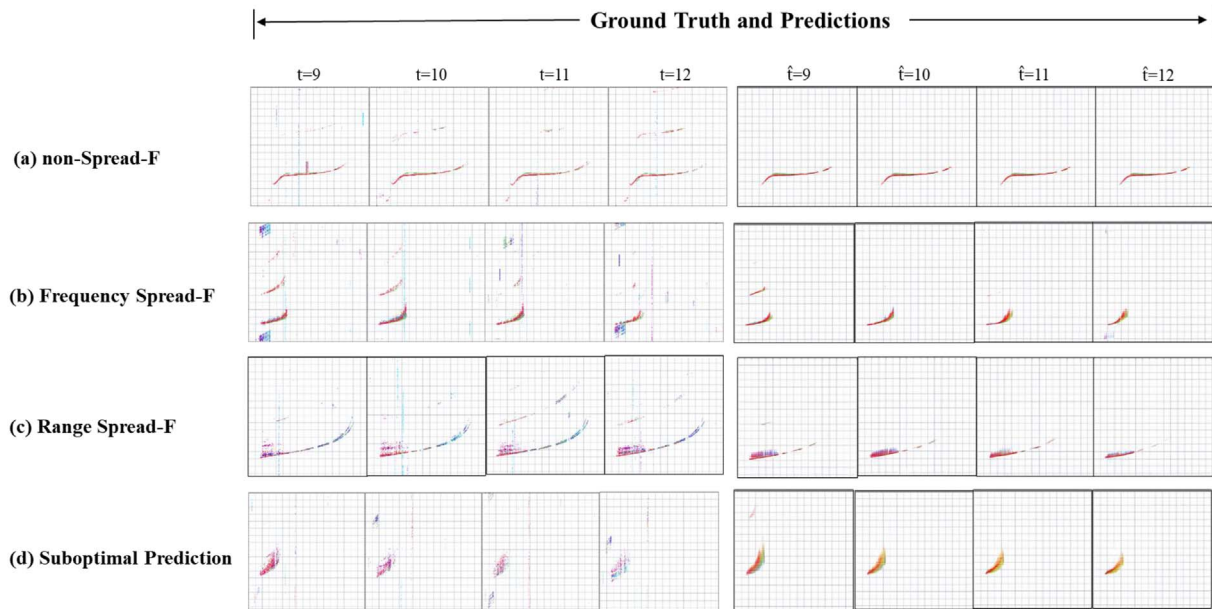


Figure 6. Visual comparison of representative four-step, one-hour-ahead ionogram forecasts generated by the SGE-ConvTransformer across four distinct ionospheric conditions: (a) Non-Spread-F, (b) Frequency Spread-F, (c) Range Spread-F, and (d) a challenging suboptimal prediction. For each condition (row), the four left panels display the ground-truth observational sequence at consecutive 15-minute intervals (from $t = 9$ to $t = 12$), while the four right panels (from $\hat{t} = 9$ to $\hat{t} = 12$) display the corresponding model predictions. While cases (a)–(c) demonstrate the model's proficiency in capturing principal morphological dynamics and global semantic structures, case (d) highlights a limitation. Although the structural shapes are largely preserved, noticeable chromatic discrepancies occur, indicating a potential distortion in the Angle of Arrival (directional) information.

4.2 Performance under extreme conditions

Previous studies have noted that LSTM-based ionospheric prediction models often suffer from performance degradation during geomagnetically disturbed periods (Kim et al., 2021). To evaluate the resilience of the proposed SGE-ConvTransformer under such conditions, we selected six representative geomagnetically disturbed intervals from 2023 and 2024. Each

interval contains at least one hour with Disturbance Storm Time (Dst) ≤ -50 nT based on the WDC for Geomagnetism, Kyoto provisional hourly Dst record. Among them, the 10–17 May 2024 interval was the most severe, reaching a minimum hourly Dst of -406 nT.

The comparative performance of our model versus (Gao et al., 2024a) is illustrated in Figure 7. The upper x-axis denotes the years, while the lower x-axis identifies the specific disturbed

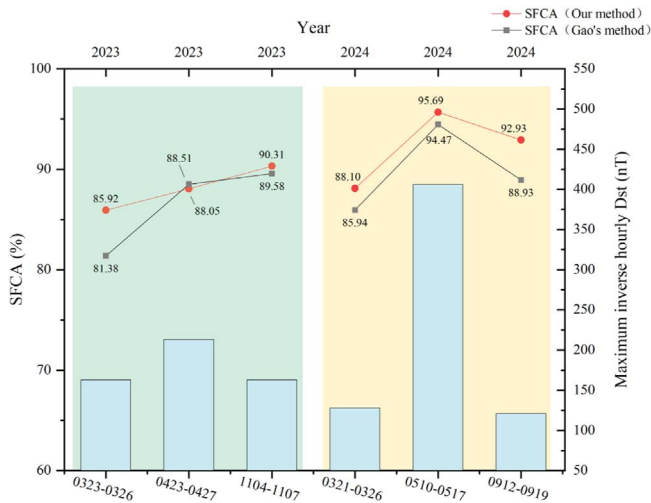


Figure 7. Comparison of mean SFCA during six selected geomagnetically disturbed intervals in 2023–2024. Red circles denote the performance of the proposed SGE-ConvTransformer, and gray squares denote the Gao et al. (2024a) baseline. The blue bars represent the maximum inverse hourly Dst within each labeled interval, calculated as $\max[-Dst(t), 0]$ using provisional hourly Dst data (UT) from the WDC for Geomagnetism, Kyoto. The upper x-axis indicates the year, the lower x-axis defines the specific interval, and larger bar heights correspond to stronger peak geomagnetic disturbances.

intervals. The left y-axis shows the SFCA values, and the right y-axis indicates the maximum inverse hourly Dst. In the figure, the bar heights correspond to the maximum inverse hourly Dst within each interval, calculated as $\max[-Dst(t), 0]$ from the WDC Kyoto provisional data (UT), with taller bars signifying greater geomagnetic disturbance intensity. For instances such as the March and September 2024 cases, which contain partially overlapping storms, the bar height represents the strongest hourly disturbance within that selected time window. The red circles represent the mean SFCA achieved by our SGE-ConvTransformer, while the gray squares denote the results from the Gao et al. (2024a) baseline model.

Notably, our model exhibits robust performance across intervals with different geomagnetic intensity. During the September and May 2024 disturbed intervals, our model achieved accuracies of 92.93% and 95.69%, respectively, outperforming the Gao model (88.93% and 94.47%). While both models capture disturbed-time features effectively, our approach demonstrates slightly better stability, minimizing performance fluctuation across different levels of geomagnetic disturbance.

This improvement can be attributed to the distinct ionospheric morphology induced by intense geomagnetic forcing. At the low-latitude Tucumán-Argentina station (26.9°S, 294.6°E; magnetic 15.5°S), González (2021) found that geomagnetic activity creates favorable conditions for the initiation of ionospheric irregularities, manifested by ionogram Spread-F and TEC fluctuation. Li et al. (2010) suggested that the simultaneous turning of IMF-Bz led to the rapid penetration of the eastward electric field into low latitudes, enhancing the normal ionospheric electric field. This process lifted the F-layer, creating favorable conditions for the development of spread-F. It was also confirmed by Rastogi et al. (2014) by checking

the spread-F at low-latitude stations around the world during a magnetic storm.

These physical processes result in stronger morphological signatures than those observed during quiet times. Unlike the faint, patchy, or diffusive irregularities characteristic of mild disturbances, storm-induced patterns typically manifest as large-scale, high-intensity Spread-F signatures, such as deep equatorial plasma bubbles with sharp gradients. From a computer vision perspective, these high-contrast features are more salient and easier for the Convolutional Neural Network to extract and classify, which helps explain the relatively high accuracy maintained during strongly disturbed intervals. Furthermore, a secondary factor contributing to this improvement is the alleviation of data imbalance. Statistical analysis reveals that the occurrence rate of Spread-F increases significantly during disturbed intervals. For instance, during the March 23–26, 2023 interval, the Spread-F occurrence rate surged to 19.53% (225/1,152), nearly triple the annual average of 7.12% (4,266/59,912). This increased density of positive samples locally mitigates the class imbalance inherent in ionospheric datasets, reducing the model's bias toward the negative class and further enhancing detection sensitivity.

4.3 Generalization capability: a zero-shot cross-site evaluation

A critical question remains: does the model's success rely on overfitting to the specific local trends of the Hainan station? To address this, we conducted a zero-shot evaluation to assess the generalization capability of the proposed SGE-ConvTransformer beyond its primary training site.

We first extended the evaluation to two additional low-latitude stations located in distinct longitudinal sectors: FZA0M (Fortaleza, Brazil, 3.9°S, 38.4°W) and JI91J (Jicamarca, Peru, 12.0°S, 76.8°W). Although these stations share similar geomagnetic characteristics with Hainan (19.5°N, 109.1°E), they operate in different electromagnetic environments. Crucially, the model, trained exclusively on Hainan data, was directly applied to these test sets without any fine-tuning.

The experimental results demonstrate robust cross-site performance. For the FZA0M station, utilizing a full-year dataset from 2016 (1,191 valid sequences), our model achieved a stable average SFCA of 87.55%. To further validate this resilience, we tested the model on another randomly sampled dataset of 1,500 sequences from the JI91J station, where it yielded an even higher average accuracy of 89.01%. This consistent performance across three geographically dispersed sites indicates that the SGE-ConvTransformer effectively captures the fundamental morphological features of Spread-F, such as characteristic diffuse traces and layer distortions, that are physically consistent across the equatorial ionosphere, rather than overfitting to site-specific artifacts or local background trends.

A key advantage of our proposed method lies in its reliance on pure visual data. Unlike physics-based models that often require rigorous site-specific parameter calibration, our model learns to forecast the morphological evolution of traces directly from image sequences. This characteristic allows the model to be efficiently deployed on various ionosondes capable of generating standard ionograms, significantly reducing the barrier for cross-site application.

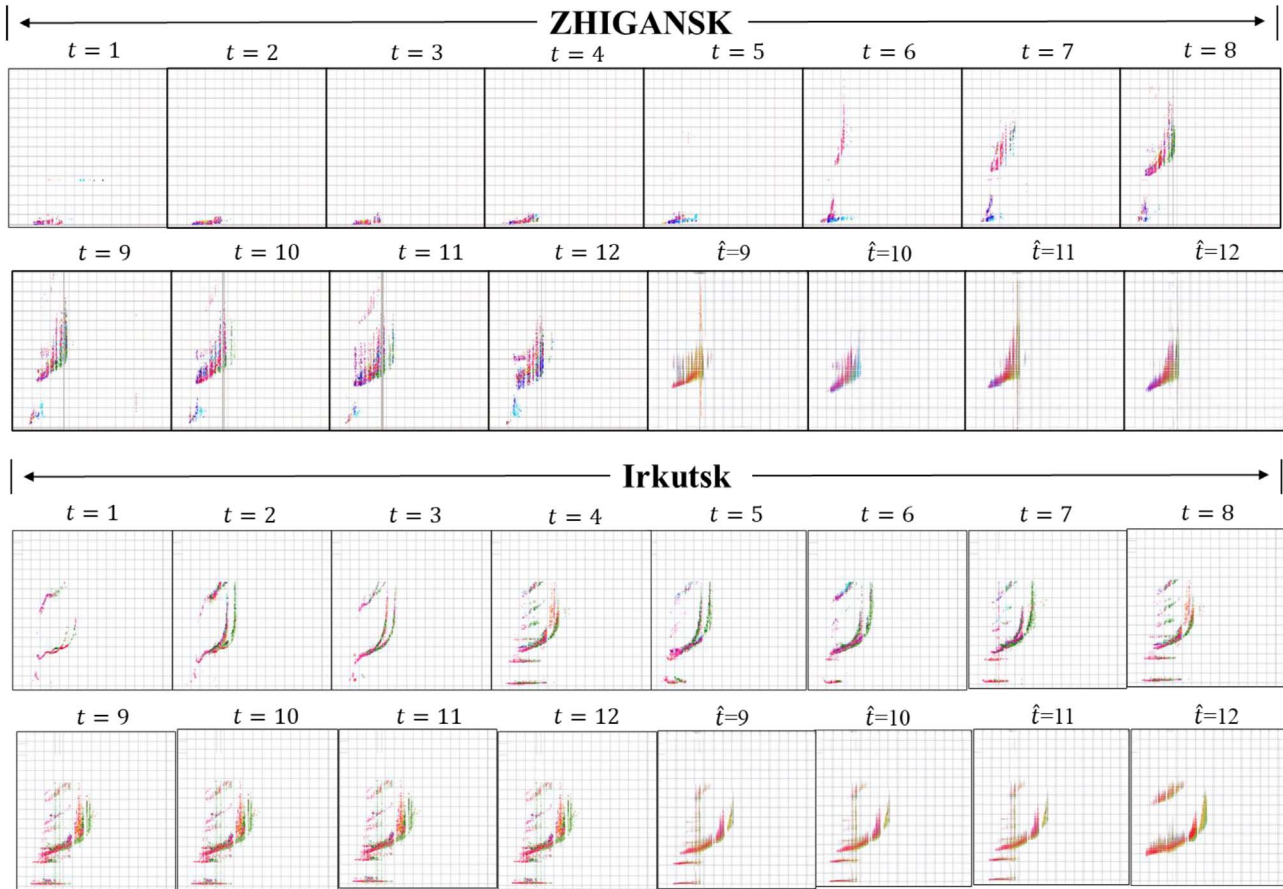


Figure 8. Zero-shot generalization results on unseen mid-latitude (Irkutsk) and high-latitude (Zhigansk) stations.

To further explore the limits of the model’s generalization capability, we extended the visual inspection to regions with distinct climatological characteristics. We randomly selected representative ionogram sequences from a mid-latitude station IR352 (Irkutsk, Russia, 52.3°N, 104.3°E, year 2011) and a high-latitude station ZH466 (Zhigansk, Russia, 66.8°N, 123.4°E, year 2014). As illustrated in Figure 8, even without any fine-tuning on data from these regions, the model demonstrates a capability to reconstruct the primary F-layer traces and background structures.

However, it is important to note that ionospheric dynamics at high latitudes are governed by different physical mechanisms compared to low latitudes. A specific limitation arises from the training data distribution: the Hainan dataset follows a specific 5-class taxonomy for low-latitude Spread-F. Consequently, the model has not learned specific spread types more common in other regions, such as “branch” or “slant” spread. When encountering these unseen morphologies, the model tends to approximate them as a generic range or frequency spread, leading to a potential loss of fine-grained structural details. Nevertheless, this zero-shot qualitative success suggests that the SGE-ConvTransformer has learned robust, generalized visual representations of ionograms that are not strictly bound to a single station’s data distribution. Therefore, we explicitly state that these zero-shot results should be interpreted strictly as evidence of morphological plausibility and structural transferability under out-of-distribution conditions. They do not constitute a validated measure of physical forecasting skill at mid- and high-latitude

stations, as establishing true predictive capability in these fundamentally different regimes would require region-specific quantitative benchmarks and station-level fine-tuning.

4.4 Predictive mechanism and physical interpretability

To rigorously verify whether the model suffers from a characteristic delay, specifically, whether it can anticipate an event onset without simply extrapolating the state of the immediate past, we conducted a statistical analysis on “Sudden Onset” cases within the 2016 test set. We focused on critical transition sequences where the ground truth remains in a Non-Spread-F state at the third prediction step ($t_{+45\text{min}}$) but abruptly transitions to a Spread-F event at the fourth prediction step ($t_{+60\text{min}}$). This scenario presents a rigorous test for the model: a naive predictor driven by temporal inertia would likely fail to detect the change, propagating the “quiet” status from the preceding frame. However, our results demonstrate that the SGE-ConvTransformer successfully predicted the Spread-F onset in 73.49% of these challenging cases.

This capability is well-supported by the specific timescales of ionospheric dynamics. Based on the studies mentioned above, it is evident that the ionosphere is highly susceptible to sudden space weather events such as solar flares and geomagnetic storms, which typically induce abrupt ionospheric changes on timescales ranging from minutes to a few hours. Although 2–3 hours of historical data may be insufficient to capture long-term periodicities (e.g., diurnal or seasonal cycles), such

short-term sequences can effectively characterize the initial features of sudden disturbances, including sharp enhancements in electron density and early gradient changes associated with the formation of equatorial plasma bubbles (EPBs).

In this context, the ConvTransformer architecture incorporating convolutional multi-head attention mechanisms is believed to better extract localized transient features (e.g., spiky pulses or abrupt gradient changes), outperforming traditional recurrent models such as LSTMs. Furthermore, the attention mechanism of ConvTransformer enables the model to learn and identify temporal dependencies before and after disturbance events, thereby substantially improving the prediction of ionospheric sudden events. This advantage over conventional LSTM models in capturing nonlinear transient phenomena has been demonstrated in the original study by Liu et al. (2021).

Another critical question arises regarding the role of external geophysical drivers (e.g., solar flux, geomagnetic indices), which are known to govern the generation of Spread-F. While our current SGE-ConvTransformer relies solely on visual inputs, it effectively internalizes these physical dependencies. We posit that the ionogram sequence serves as a high-fidelity, integrated observational manifestation of the underlying physical state. For instance, the rapid uplift of the F-layer, a known precursor to the Rayleigh-Taylor instability driven by pre-reversal enhancement electric fields, is captured as a distinguishable vertical displacement trend in the input image sequence. Consequently, the model is able to predict the onset of Spread-F by recognizing these morphological precursors inherent in the kinematic history of the plasma, rather than merely identifying existing spread traces.

This mechanism is supported by our empirical experiments, which indicated that directly fusing coarse-resolution scalar indices (e.g., 3-hour Kp) with high-resolution ionogram features yielded negligible performance gains for short-term (1-hour) forecasting. This suggests that for nowcasting tasks, the dense spatiotemporal information encoded in the historical ionograms provides a sufficient predictive signal, rendering explicit scalar inputs redundant in this specific architecture. Nevertheless, we acknowledge that explicitly incorporating these indices remains a high-value direction for future research, particularly for longer-term forecasting. As the prediction horizon extends beyond the immediate kinematic trajectory of the ionosphere, the "inertia" of the historical image sequence diminishes, making external forcing parameters increasingly critical. However, a central challenge in extending the forecasting horizon is that the model may begin to capture trend features similar to those of climatological statistical models like IRI or NeQuick, thereby diminishing its sensitivity to the instantaneous, transient variations that characterize Spread-F events.

5 Conclusions

In this work, we have developed an end-to-end framework for ionospheric nowcasting through full ionogram sequence prediction, advancing beyond traditional scalar-parameter forecasting. By formulating the problem as spatiotemporal field generation, the proposed Spatial Group-wise Enhanced ConvTransformer (SGE-ConvTransformer) model simultaneously captures background trends in electron density profiles,

morphological evolution of storm-induced irregularities (e.g., equatorial Spread-F), and precursors to plasma instability onset. Methodologically, our approach overcomes key limitations of prior data-driven models: Unlike sequential recursive models (e.g., ConvGRU used in Gao et al., 2024) which are prone to error accumulation over multi-step forecasts, the ConvTransformer architecture leverages multi-head convolutional attention to model long-range dependencies (60–120 min) and the initial features of sudden disturbances while avoiding such compounding errors; the Spatial Group-wise Enhancement (SGE) module dynamically amplifies discriminative features in the F-region structures where over 90% of Spread-F signatures originate; and EDSR-based super-resolution restores high-frequency details in predicted ionograms, achieving a 15.7% increase in F-trace correlation (AVCC-F) versus baseline generative models. Using ionosonde observations from Hainan Station, the model achieves up to 95.69% accuracy in predicting disturbed ionospheric conditions during geomagnetic storm periods, effectively capturing transient and small-scale variations that are often missed by statistical models such as IRI and NeQuick. Importantly, beyond the primary training site, the model's cross-site transferability was also examined. In a zero-shot cross-site evaluation on two additional low-latitude stations (Fortaleza, Brazil, and Jicamarca, Peru), the model yielded high average accuracies of 87.55% and 89.01%, respectively. Moreover, qualitative inspections on mid- and high-latitude stations (Russia) suggest that the SGE-ConvTransformer relies on intrinsic visual patterns rather than site-specific parameters, allowing it to reconstruct fundamental F-layer traces even in unseen climatological regions. These results indicate that the SGE-ConvTransformer captures fundamental, morphologically plausible Spread-F features across different longitudinal sectors, rather than overfitting to site-specific artifacts. The generation of four future ionograms (typically one hour ahead) based on eight preceding ionograms has been shown to offer a robust balance between sensitivity to rapid fluctuations and the stability necessary for operational use.

The experimental results confirm that advanced AI generative architectures can substantially enhance the predictive capability of ionospheric condition models, with implications for space weather services, HF communication planning, and GNSS performance management. Unlike purely statistical or physics-only models, our data-driven framework maintains responsiveness to short-term ionospheric disturbances while exhibiting promising transferability across varying spatial-temporal conditions.

Looking forward, two main research directions are envisaged. First, while our qualitative evaluation confirmed the model's basic adaptability to high-latitude regions, specific limitations remain. Since our current model was trained exclusively on low-latitude data with a specific 5-class taxonomy, it effectively approximates general spread features but struggles to resolve complex morphologies unique to high-latitude dynamics, such as "branch" or "slant" spread types. Therefore, extending the training dataset to include high-latitude observations will allow the model to learn these diverse scattering signatures, thereby improving generalization and fine-grained classification on a global scale. Second, regarding methodological evolution, a key objective is to bridge the gap between vision-based nowcasting and physics-driven long-term forecasting. While this study

suggests that visual features may be sufficient for short-term forecasting horizons, extending the prediction window beyond the immediate kinematic trajectory will likely require the integration of external geophysical drivers (e.g., solar and geomagnetic indices). Developing a hybrid architecture that fuses these physical constraints with data-driven visual representations offers a promising pathway to unify operational real-time warning with medium-term background specification, ultimately enhancing the resilience of space-based systems against adverse space weather.

Acknowledgments

The authors acknowledge the Public Computing Cloud of CUC for supporting the computational resources. We would also like to acknowledge the use of data from the Chinese Meridian Project. We express our sincere gratitude to the editor and the two anonymous reviewers for their constructive comments, which have significantly improved the quality and enriched the content of this paper. The editor thanks Marco Guerra and Hossein Ghadjari for their assistance in evaluating this paper.

Funding

This research was supported by the National Key R&D Program of China (2023YFB3905100); Youth Innovation Promotion Association CAS (2023000116); The Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA0470301); Project of Stable Support for Youth Team in Basic Research Field, CAS (YSBR-018); Specialized Research Fund for State Key Laboratories; Pandeng Program of National Space Science Center, Chinese Academy of Sciences; the High-quality and Cutting-edge Disciplines Construction Project for Universities in Beijing (Internet Information, Communication University of China); the Fundamental Research Funds for the Central Universities; and the Chinese Meridian Project.

Data availability statement

The ionogram data (.RSF/.PNG format) could be obtained from the Fuke Station (FKT) at Hainan, China, downloaded from the Data Center for Meridian Space Weather Monitoring Project /NSSC (<https://data2.meridianproject.ac.cn/>). Please choose the English version in the top right-hand corner, finish the sign-up (<https://data2.meridianproject.ac.cn/registration>), and sign in to the official website. If you are still unable to access the database and require the corresponding ionogram set, feel free to contact us directly (Email: pd_gao@cuc.edu.cn, zwang@swl.ac.cn).

References

Abdu MA, MacDougall JW, Batista IS, Sobral JHA, Jayachandran PT. 2003. Equatorial evening prereversal electric field enhancement and sporadic E layer disruption: A manifestation of E and F region coupling. *J Geophys Res Space Phys* **108**(A6): 1254. <https://doi.org/10.1029/2002JA009285>.

Benchawattananon P, Siritratiwat A, Supnithi P, Nishioka M, Perwitasari S, et al. 2024. Automatic classification of spread-F types in ionogram images using support vector machine and convolutional neural network. *Earth Planets Space* **76**: 56. <https://doi.org/10.1186/s40623-024-02002-x>.

Bilitza D, Altadill D, Truhlik V, Shubin V, Galkin I, et al. 2017. International Reference Ionosphere 2016: From ionospheric

climate to real-time weather predictions. *Space Weather* **15**: 418–429. <https://doi.org/10.1002/2016SW001593>.

Cherniak I, Zakharenkova I. 2016. First observations of super plasma bubbles in Europe. *Geophys Res Lett* **43**: 12169–12177. <https://doi.org/10.1002/2016GL071421>.

Gao P, Cai J, Wang Z, Qiu C, Wang G, et al. 2024a. Prediction of ionograms with/without spread-F at Hainan by a combined spatio-temporal neural network. *Space Weather* **22**: e2023SW003727. <https://doi.org/10.1029/2023SW003727>.

Gao S, Xiong C, Cai H, Pan Q, Zhan W, Zhang H, Zheng Y. 2025. Predicting equatorial spread F at Jicamarca sector via supervised machine learning. *Space Weather* **23**: e2024SW004214. <https://doi.org/10.1029/2024SW004214>.

Gao X, Yao Y, Wang Y. 2024b. Reconstruction of global ionospheric TEC maps from IRI-2020 model based on deep learning method. *J Geod* **98**: 10. <https://doi.org/10.1007/s00190-023-01818-x>.

González GL. 2021. Spread-F occurrence during geomagnetic storms near the southern crest of the EIA in Argentina. *Adv Space Res* **67** (3): 1058–1084. <https://doi.org/10.1016/j.asr.2020.10.051>.

Hargreaves JK. 1992. *The solar-terrestrial environment*. Cambridge Univ Press, Cambridge, UK. <https://doi.org/10.1017/CBO9780511628924>.

Huba JD, Joyce G, Krall J. 2008. Three-dimensional equatorial spread F modeling. *Geophys Res Lett* **35**: L10102. <https://doi.org/10.1029/2008GL033509>.

Kelley MC. 2009. *The Earth's ionosphere: plasma physics and electrodynamics* (2nd ed.). Academic Press/Elsevier, Amsterdam. ISBN 978-0-12-088425-4.

Kim J-H, Kwak Y-S, Kim Y-H, Moon S-I, Jeong S-H, et al. 2021. Potential of regional ionosphere prediction using a long short-term memory deep-learning algorithm specialized for geomagnetic storm period. *Space Weather* **19**: e2021SW002741. <https://doi.org/10.1029/2021SW002741>.

Lan T, Zhang Y, Jiang C, Yang G, Zhao Z. 2018. Automatic identification of spread-F using decision trees. *J Atmos Solar-Terr Phys* **179**: 389–395. <https://doi.org/10.1016/j.jastp.2018.09.007>.

Lan T, Hu H, Jiang C, Yang G, Zhao Z. 2020. A comparative study of decision tree, random forest, and convolutional neural network for spread-F identification. *Adv Space Res* **65**(8): 2052–2061. <https://doi.org/10.1016/j.asr.2020.01.036>.

Li G, Ning B, Hu L, Liu L, Yue X, et al. 2010. Longitudinal development of low-latitude ionospheric irregularities during the geomagnetic storms of July 2004. *J Geophys Res Space Phys* **115**: A04304. <https://doi.org/10.1029/2009JA014830>.

Li X, Zhou C, Tang Q, Zhao J, Zhang F, et al. 2021. Forecasting ionospheric foF2 based on deep learning method. *Remote Sens* **13** (19): 3849. <https://doi.org/10.3390/rs13193849>.

Li Y, Li X, Yang J. 2022. Spatial Group-wise Enhance: Enhancing semantic feature learning in CNN. *Lect Notes Comput Sci* **13843**: 687–702. (ACCV 2022). https://doi.org/10.1007/978-3-031-26348-4_19.

Lim B, Son S, Kim H, Nah S, Lee KM. 2017. Enhanced deep residual networks for single image super-resolution. *Proc IEEE/CVF Conf Comput Vis Pattern Recognit Workshops (CVPRW 2017)* **13843**: 1132–1140. <https://doi.org/10.1109/CVPRW.2017.151>.

Liu Z, Luo S, Li W, Lu J, Wu Y, et al. 2020. *ConvTransformer: A convolutional Transformer network for video frame synthesis*: arXiv:2011.10185.

Luwanga C, Fang T-W, Chandran A, Lee Y-J. 2022. Automatic spread-F detection using deep learning. *Radio Sci* **57**(5): e2021RS007419. <https://doi.org/10.1029/2021RS007419>.

McNamara LF, Baker CR, Decker DT. 2008. Accuracy of USU-GAIM specifications of foF2 and M(3000)F2 for a worldwide distribution of ionosonde locations. *Radio Sci* **43**: RS1011. <https://doi.org/10.1029/2007RS003754>.

- Millward GH, Müller-Wodarg ICF, Aylward AD, Fuller-Rowell TJ, Richmond AD, et al. 2001. An investigation into the influence of tidal forcing on F region equatorial vertical ion drift using a global ionosphere–thermosphere model with coupled electrodynamics. *J Geophys Res Space Phys* **106**(A11): 24733–24744. <https://doi.org/10.1029/2000JA000342>.
- Nava B, Coisson P, Radicella SM. 2008. A new version of the NeQuick ionosphere electron density model. *J Atmos Solar-Terr Phys* **70**: 1856–1862. <https://doi.org/10.1016/j.jastp.2008.01.015>.
- Pignalberi A, Pietrella M, Pezzopane M. 2021. Towards a real-time description of the ionosphere: A comparison between IRI and IRTAM. *Atmosphere* **12**(8): 1003. <https://doi.org/10.3390/atmos12081003>.
- Pillat VG, Fagundes PR, Guimarães LNF. 2015. Automatic identification of Equatorial Spread-F occurrence on ionograms. *J Atmos Solar-Terr Phys* **135**: 118–125. <https://doi.org/10.1016/j.jastp.2015.10.015>.
- Qiu C, Cai J, Wang Z, Gao P, Wang G, et al. 2025. IonoGAN: An enhanced model for forecasting quiet and disturbed ionospheric features from predicted ionograms. *Space Weather* **23**: e2025SW004463. <https://doi.org/10.1029/2025SW004463>.
- Radicella SM, Leiting R. 2001. The evolution of the DGR approach to model electron density profiles. *Adv Space Res* **27**(1): 35–40. [https://doi.org/10.1016/S0273-1177\(00\)00138-1](https://doi.org/10.1016/S0273-1177(00)00138-1).
- Rastogi RG, Chandra H, Janardhan P, Hoang TL, Condori L, et al. 2014. Spread-F during the magnetic storm of 22 January 2004 at low latitudes: Effect of IMF-Bz in relation to local sunset time. *J Earth Syst Sci* **123**(6): 1273–1285. <https://doi.org/10.1007/s12040-014-0467-3>.
- Reinisch BW, Huang X. 1983. Automatic calculation of electron density profiles from digital ionograms. *Radio Sci* **18**(3): 477–492. <https://doi.org/10.1029/RS018i003p00477>.
- Ren X, Zhao B, Ren Z, Xiong B. 2024. Ionospheric TEC prediction in China during storm periods based on deep learning: Mixed CNN–BiLSTM method. *Remote Sens* **16**(17): 3160. <https://doi.org/10.3390/rs16173160>.
- Richmond AD, Ridley EC, Roble RG. 1992. A thermosphere/ionosphere general circulation model with coupled electrodynamics. *Geophys Res Lett* **19**(6): 601–604. <https://doi.org/10.1029/92GL00401>.
- Shi JK, Wang GJ, Reinisch BW, Shang SP, Wang X, et al. 2011. Relationship between strong range spread F and ionospheric scintillations observed in Hainan from 2003 to 2007. *J Geophys Res Space Phys* **116**: A08306. <https://doi.org/10.1029/2011JA016806>.
- Thammavongsy P, Supnithi P, Phakphisut W, Hozumi K, Tsugawa T. 2020. Spread-F prediction model for the equatorial Chumphon station Thailand. *Adv Space Res* **65**(1): 152–162. <https://doi.org/10.1016/j.asr.2019.09.040>.
- Thammavongsy P, Supnithi P, Myint LMM, Hozumi K, Lakanchanh D. 2023. Equatorial spread-F forecasting model with local factors using the long short-term memory network. *Earth Planets Space* **75**: 118. <https://doi.org/10.1186/s40623-023-01868-7>.
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, et al. 2017. Attention Is All You Need. *Adv Neural Inf Process Syst* (NeurIPS 2017). arXiv:1706.03762.
- Wang Z, Zhan M, Gao P, Wang G, Qiu C, et al. 2023. Automatic detection and classification of spread-F from ionosonde at Hainan with image-based deep learning method. *Space Weather* **21**: e2023SW003498. <https://doi.org/10.1029/2023SW003498>.
- Yang Z, Qiao L, Su M, Hu Z, Teng X, et al. 2023. Ionospheric foF2 nowcast based on a machine learning GWO-ALSTM model. *Adv Space Res* **72**(11): 4896–4910. <https://doi.org/10.1016/j.asr.2023.09.028>.
- Zhao X, Li G, Xie H, Hu L, Sun W, et al. 2025. A novel short-term prediction model for regional equatorial plasma bubble irregularities in East and Southeast Asia. *Space Weather* **23**: e2024SW004224. <https://doi.org/10.1029/2024SW004224>.

Cite this article as: Cai J, Gao P, Wang Z, Wang B, Wang G, et al. 2026. One-hour-ahead forecasting of ionogram morphology and spread-F signatures using a spatial group-wise enhanced ConvTransformer. *J. Space Weather Space Clim.* **16**, 20. <https://doi.org/10.1051/swsc/2026015>.