

Article

ATCNet: A Novel Approach for Predicting Highway Visibility Using Attention-Enhanced Transformer–Capsule Networks

Wen Li ^{1,2,3}, Xuekun Yang ², Guowu Yuan ^{1,3}  and Dan Xu ^{1,*} 

¹ School of Information Science and Engineering, Yunnan University, Kunming 650504, China; liwen@stwp.com (W.L.); gwyuan@ynu.edu.cn (G.Y.)

² National Engineering Laboratory for Surface Transportation Weather Impacts Prevention, Broadvision Engineering Consultants Co., Ltd., Kunming 650299, China

³ Yunnan Key Laboratory of Digital Communications, Kunming 650103, China

* Correspondence: danxu@ynu.edu.cn; Tel.: +86-871-605033748

Abstract: Meteorological disasters on highways can significantly reduce road traffic efficiency. Low visibility caused by dense fog is a severe meteorological disaster that greatly increases the incidence of traffic accidents on highways. Accurately predicting highway visibility and taking timely countermeasures can mitigate the impact of meteorological disasters and enhance traffic safety. This paper introduces the ATCNet model for highway visibility prediction. In ATCNet, we integrate Transformer, Capsule Networks (CapsNet), and self-attention mechanisms to leverage their respective complementary strengths. The Transformer component effectively captures the temporal characteristics of the data, while the Capsule Network efficiently decodes the spatial correlations and hierarchical structures among multidimensional meteorological elements. The self-attention mechanism, serving as the final decision-refining step, ensures that all key temporal and spatial hierarchical information is fully considered, significantly enhancing the accuracy and reliability of the predictions. This integrated approach is crucial in understanding highway visibility prediction tasks influenced by temporal variations and spatial complexities. Additionally, this study provides a self-collected publicly available dataset, WD13VIS, for meteorological research related to highway traffic in high-altitude mountain areas. This study evaluates the model's performance in terms of Mean Squared Error (MSE) and Mean Absolute Error (MAE). Experimental results show that our ATCNet reduces the MSE and MAE by 1.21% and 3.7% on the WD13VIS dataset compared to the latest time series prediction model architecture. On the comparative dataset WDVigoVis, our ATCNet reduces the MSE and MAE by 2.05% and 5.4%, respectively. Our model's predictions are accurate and effective, and our model shows significant progress compared to competing models, demonstrating strong universality. This model has been integrated into practical systems and has achieved positive results.

Keywords: traffic; meteorological disaster; atmospheric visibility; forecasting; deep learning; transformer networks; capsule networks (CapsNet); attention mechanisms



Citation: Li, W.; Yang, X.; Yuan, G.; Xu, D. ATCNet: A Novel Approach for Predicting Highway Visibility Using Attention-Enhanced Transformer–Capsule Networks. *Electronics* **2024**, *13*, 920. <https://doi.org/10.3390/electronics13050920>

Academic Editors: Haihua Chen, Junhua Ding, Yunhe Feng and Tozammel Hossain

Received: 21 January 2024

Revised: 20 February 2024

Accepted: 26 February 2024

Published: 28 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the advancement of vehicles and the continuous development of transportation infrastructure, the need for road safety has become more urgent. Typical meteorological disasters along highways are a major cause of decreased road traffic efficiency [1]. In recent years, there has been growing concern about the impact of traffic-related meteorological disasters on road traffic. These disasters refer to traffic accidents and congestion caused by sudden or adverse weather events, severely affecting road traffic efficiency [2]. For example, severe meteorological disasters like dense fog and heavy snow can cause low visibility and slippery road surfaces, increasing the likelihood of various traffic accidents [3]. Extreme weather conditions such as typhoons and heavy rain can lead to road damage and traffic control, significantly impacting transportation efficiency [4]. Therefore, effective

warning and control of traffic meteorological disaster events can reduce the incidence of traffic accidents, thereby enhancing road traffic efficiency [5].

In traffic meteorological disasters, low visibility is a critical factor, especially when fog suddenly envelops highways, increasing the risk of multiple-vehicle collisions. For instance, according to data from the U.S. Federal Highway Administration, over 38,700 vehicle accidents occur annually in fog, resulting in more than 600 deaths and over 16,300 injuries [6]. Yunnan Province in China, characterized by its plateau and mountainous terrain, commonly faces low-visibility issues. The province has an extensive road network, including significant highways, which are particularly susceptible to risks associated with low visibility in high-altitude areas. Therefore, accurate visibility prediction is crucial for effective traffic control in these regions.

Figure 1 shows real-scene photographs collected from the Yunnan Province Plateau Mountain Area Traffic Meteorological Database, documenting the visibility changes at the same spatial location over 36 min.



Figure 1. Real-world scenario of highway visibility changes. (a) Visibility 3406 m (datetime: 24.12.2023 14:40:04). (b) Visibility 528 m (datetime: 24.12.2023 14:52:04). (c) Visibility 166 m (date: 24.12.2023 15:04:04). (d) Visibility 323 m (datetime: 24.12.2023 15:16:04). The image's timestamp in the upper left is captured by the video surveillance, displayed in Chinese datetime format. The highway's central electronic sign, in Mandarin, advises 'Drive cautiously when entering tunnel'.

Highway visibility prediction technology is based on analyzing meteorological and related data, utilizing computer models and algorithms to predict future visibility values under certain assessment criteria. Highway visibility prediction can be provided in real-time to traffic management departments or drivers through means such as implementing variable speed limits on roads, assisting them in taking appropriate control measures to prevent or reduce traffic accidents caused by adverse weather conditions.

Traditional road visibility prediction methods primarily use statistical and regression approaches. However, these methods have lower prediction accuracy, longer computation times, and cumbersome model structures, making them unsuitable for the demand of fast and accurate visibility prediction on highways.

In recent years, with the widespread application of deep learning technology in fields such as computer vision and natural language processing, progress has also been made in road visibility prediction. Road visibility prediction based on deep learning utilizes neural networks' adaptability and non-linear mapping capabilities. By learning

and feature extraction from sample data, it achieves road visibility prediction results that meet evaluation standards.

There are still challenges in applying deep learning methods in road visibility prediction. For instance, current deep learning-based visibility prediction techniques mainly rely on image data samples and require high-quality image samples for data collection. However, image collection and transmission can be easily affected by factors unrelated to meteorological conditions, such as network transmission issues, hardware limitations, and environmental interference, which may degrade data quality or interrupt data transmission. Additionally, some prediction models are overly complex, and the meteorological dimension of sample data is insufficient, affecting the accuracy and reliability of predictions.

Therefore, deep learning techniques for visibility prediction based on multidimensional non-image data are urgently needed to improve the accuracy and robustness of highway visibility prediction.

This paper constructs a dataset based on meteorological data collected from specialized meteorological equipment on highways in the plateau mountain areas of Yunnan Province, China, and proposes the ATCNet model. This model combines the Transformer, Capsule Networks, and self-attention mechanisms, fully leveraging their complementary advantages in capturing temporal features, understanding spatial relationships, and optimizing the decision-making process. These are used to improve the accuracy and robustness of highway visibility prediction under complex meteorological conditions. Comparative experiments and ablation studies with competing models demonstrate that each component of these three elements is indispensable. The model has achieved good results, effectively meeting the practical needs of highway management in road visibility prediction, and significantly reducing the probability of traffic accidents.

2. Related Work

At the outset of visibility prediction research, traditional methods primarily leveraged statistical and regression approaches to model visibility under varying weather conditions. These foundational techniques, while effective in their time, often faced limitations in handling the complex, multifaceted nature of visibility prediction, especially under diverse and dynamic meteorological conditions. Recent advancements, as explored in this section, have significantly extended beyond these traditional frameworks, employing numerical simulations, advanced machine learning models, and multimodality approaches to offer enhanced accuracy, adaptability, and comprehensive understanding of visibility factors. This evolution reflects the field's progression towards more sophisticated, data-driven methodologies capable of addressing the nuanced challenges of visibility prediction.

2.1. Numerical Simulation-Based Methods

Numerical simulation methods can model the occurrence, development, and dissipation of low-visibility weather, providing a theoretical basis for prediction. Fernández-González et al. [7] conducted numerical prediction research on radiation and CBL fog events in Iran using the WRF model, a globally used numerical weather forecasting model suitable for simulating low-visibility weather. Pahlavan et al. [8] focused on different model configurations and visibility predictions using the WRF model, exploring the impact of different configurations on visibility prediction accuracy, and offering new perspectives for enhancing accuracy. He et al. [9] combined data and video to quantitatively analyze the evolution of dense fog. Kim et al. [10] used ground observation data from the Automated Surface Observing System (ASOS) and air pollutant data from the ECMWF Copernicus Atmospheric Monitoring Service (CAMS) model to predict visibility in Korea using the Random Forest (RF) model (VISRF). Their method showed smaller biases below 2 km than other visibility parameterization schemes. Qian et al. [11] investigated the application of anomaly-based weather analysis for low-visibility prediction in coastal fog at Ningbo Zhoushan Port in East China, enhancing prediction accuracy with anomaly weather analysis.

2.2. Machine Learning-Based Approach

Machine learning methods based on single-modal data, such as deep learning and support vector machines, have significantly progressed in low-visibility prediction. Min et al. [12] proposed a deep learning framework with an attention mechanism for visibility prediction, achieving state-of-the-art accuracy (68.9%) in runway visual range prediction using a custom dataset collected from airport observation stations. Cornejo-Bueno et al. [13] developed polynomial regression and deep neural network (DNN) models for visibility prediction. While the polynomial regression model is simple to use, its prediction accuracy is limited. The DNN model, with stronger learning capabilities, requires extensive data for training. Peláez-Rodríguez et al. [14] focused on applying machine learning-based fusion models for visibility prediction, improving prediction accuracy by combining multiple machine learning models. Zang et al. [15] developed a recurrent neural network (RNN) prediction model named SwiftRNN, which outperformed ConvLSTM and PredRNN models in visibility prediction skill scores. The RNN model, suitable for processing time series data, is effective in predicting future visibility. Han et al. [16] explored using the Long Short-Term Memory (LSTM) model for visibility prediction. The LSTM, a special type of RNN, has memory capabilities, making it better suited for handling long-time series data. Peláez-Rodríguez et al. [17] proposed a hybrid model for visibility prediction, improving accuracy by fusing multiple machine learning models.

2.3. Multimodality-Based Approach

Unlike machine learning methods using single-modal data, multimodal-based approaches integrate multi-source data such as ground observations, satellite remote sensing, and numerical forecast data, enhancing visibility prediction accuracy. Gavahi et al. [18] proposed a deep learning model based on multi-source data for precipitation prediction, outperforming traditional machine learning models in various datasets. Bai et al. [19] introduced a multimodal fusion technique, integrating ground observation, satellite remote sensing, and numerical forecast data for weather visibility prediction. This method leverages the strengths of different data sources to improve prediction accuracy. Kim et al. [20] used data from an automatic visibility observation network for data assimilation to enhance visibility forecast accuracy. Data assimilation corrects numerical forecast results with observation data, improving prediction accuracy. Qin et al. [21] proposed a deep learning-based fog visibility prediction method using diverse data sources, including ground observations, satellite remote sensing, and numerical forecasts, showing superior performance over traditional machine learning-based methods.

Additionally, studies like Guijo-Rubio et al. [22] on the influence of meteorological factors on visibility duration and Zhang et al. [23] comparing different machine learning methods in visibility prediction are noteworthy.

Our analysis identified the following issues or shortcomings in the related work:

- Although many studies have utilized datasets from airports or ports, these often represent conditions specific to certain plain areas. There is a gap in visibility prediction for roads in different terrains, especially in mountainous areas with significant terrain differences. Moreover, the training sample datasets used in most studies are not self-collected, which may compromise the reliability of model performance due to uncertain quality.
- Drawbacks in the architectural design of related work, such as limited spatial-temporal dynamics understanding, inefficiency in handling long-term dependencies, and inadequate integration of features, significantly restrict these models' potential for further enhancing visibility prediction accuracy.
- The methods' universality is limited, with almost all studies not performing comparative performance tests on multiple datasets. They are typically optimized for specific regions or climatic conditions, limiting their applicability in different geographical areas.
- While we have datasets from structured environments like airports and ports, there is a significant gap in validating prediction models in existing highway systems. Ensuring

the effectiveness of models in the real world is crucial, requiring reliance on real data under diverse and dynamically changing traffic conditions, especially in complex terrain environments.

This paper proposes a comprehensive highway visibility prediction method that utilizes deep learning models and multidimensional meteorological data. Specifically, we introduce the Transformer–CapsNet network (ATCNet) with an attention mechanism, a leading prediction model designed to provide accurate and timely short-term visibility predictions for highways. Our method has the following advantages:

- We have proposed a validated complex mountainous area road scenario visibility prediction model by integrating various meteorological sample data (including visibility, wind speed, air temperature, humidity, precipitation, and road surface temperature). This is detailed in Section 4, Methodology.
- Our model demonstrates strong universality, with tests on our self-established and public datasets showing its general applicability to various visibility prediction scenarios. This is detailed in Section 5, Experiments.
- The accuracy of our predictions has been verified through precise application data. Over the past four years, our model has been integrated into an actual “Highway Traffic Meteorological Intelligent Monitoring and Active Control System” and has been validated for its accuracy by frontline users. Details are provided in Section 5.7.

3. Dataset

3.1. Data Collection

The National Engineering Laboratory for Surface Transportation Weather Impacts Prevention in China is a leading research institution in transportation meteorology. It holds a significant position in the study of highway traffic meteorology. Figure 2 displays the meteorological data collection equipment set up by the laboratory, which is crucial for the collection of traffic meteorological data and the creation of proprietary datasets in this study. These devices are equipped with a local data buffering feature, crucial for ensuring data quality. This functionality enables the devices to temporarily store data during network disruptions, safeguarding against data loss and ensuring the continuity and integrity of the dataset for our analysis.



Figure 2. Self-developed traffic meteorological data collection equipment.

The Zhao Hui, Dai Gong, and Qu Sheng highways are three typical mountainous highways in the northeastern region of Yunnan Province, China. These highways are

characterized by complex climatic conditions and dramatic visibility changes, making them highly suitable for our road visibility prediction research. According to the construction technical standards of the highway meteorological station network [24], by the end of 2023, our team had installed a total of 13 sets of multi-element traffic meteorological stations on these three highways. Each data record collected by these meteorological devices includes nearly 20 meteorological elements, such as air temperature, precipitation, visibility, relative humidity, road surface temperature, road conditions, wind speed, wind direction, air pressure, roadbed temperature, water film thickness, freezing point temperature, ice layer thickness, snow layer thickness, and slipperiness coefficient. Figure 3 displays the geographical locations of these 13 meteorological stations. Tables 1 and 2 show database snapshots of the original real-time and historical meteorological data collected.

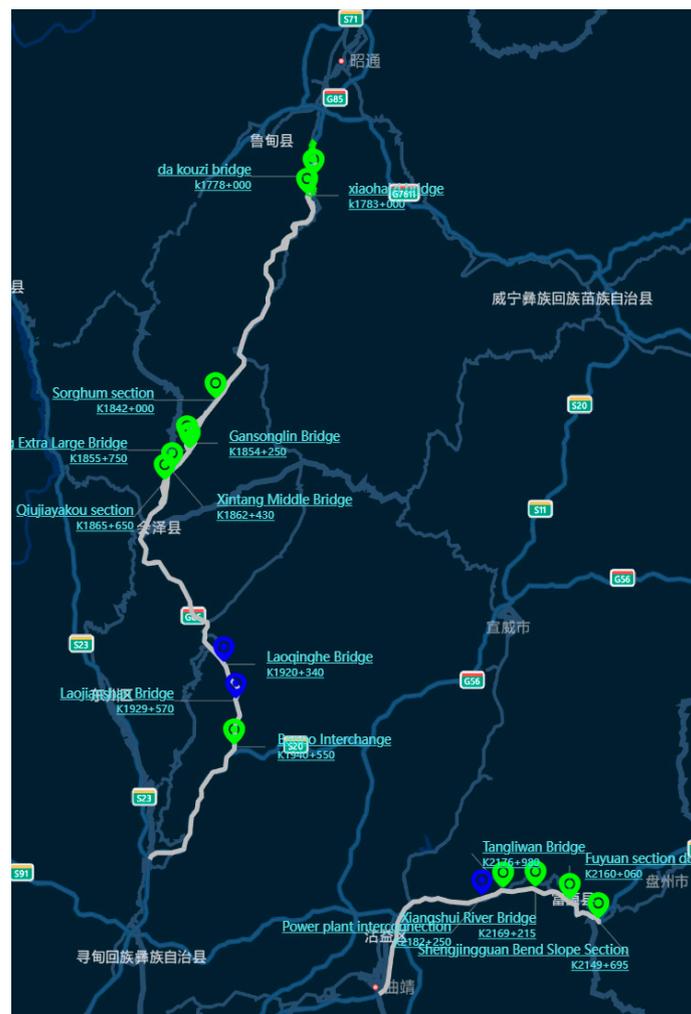


Figure 3. Geographical locations of the 13 meteorological stations shown on a map of China.

Based on the data from these 13 meteorological stations, we constructed a road meteorological element dataset named WD13VIS [25], which can be used for visibility prediction tasks. This dataset encompasses meteorological sample data collected from October 2023 to January 2024, characterized by its multidimensionality, high precision, high integrity, and high quality. Data are recorded every minute, resulting in a total of 563,715 data entries.

Table 1. Collected real-time raw meteorological data.

Id	DeviceId	Name	Description	Unit	Quality	Value	MinValue	MaxValue	Precision	Factor	TimeStamp
1	1	AMA	minute visibility	m	1	30,000	10	80,000	0	0	20240101231900000
2	1	AMB	ten minutes of visibility	m	1	30,000	10	80,000	0	0	20240101231900000
3	1	ABA	temperature	°C	1	53	−40	50	1	1	20240101231900000
4	1	ACA	humidity	%RH	1	697	0	100	1	1	20240101231900000
5	1	AEA	instantaneous wind speed	m/s	1	20	0	50	1	1	20240101231900000
6	1	ADA	instantaneous wind direction	°	1	76	0	359	0	0	20240101231900000
7	1	AEAEX	maximum wind speed in minute	m/s	1	20	0	60	1	1	20240101231900000
8	1	ADAEX	maximum wind direction in minute	°	1	76	0	359	0	0	20240101231900000
9	1	AEC	2 min average wind speed	m/s	1	10	0	60	1	1	20240101231900000
10	1	ADC	2 min mean wind direction	°	1	58	0	359	0	0	20240101231900000
11	1	AED	10 min average wind speed	m/s	1	16	0	60	1	1	20240101231900000
12	1	ADD	10 min mean wind direction	°	1	63	0	359	0	0	20240101231900000
13	1	AGA	atmospheric pressure	hPa	1	7637	500	1100	1	1	20240101231900000
14	1	AFA	minute of rainfall	mm	1	0	0	400	1	1	20240101231900000
15	1	AFB	hours of rainfall	mm	1	0	0	400	1	1	20240101231900000
16	1	APA	pavement temperature	°C	1	46	−50	80	1	1	20240101231900000
17	1	APH	pavement conditions		1	1	0	254	0	0	20240101231900000
18	1	APD	water film thickness	mm	1	0	0	25.4	1	1	20240101231900000
19	1	APE	ice thickness	mm	1	0	0	25.4	1	1	20240101231900000
20	1	APF	snow layer thickness	mm	1	0	0	25.4	1	1	20240101231900000
21	1	AQA	slip factor		1	82	0	1	2	2	20240101231900000

Table 2. Collected historical raw meteorological data.

Stationnum	Date_time	Visibility	Temperature	Humidity	Precipitation	Meteorological_Phenomenon	Dmwind speed	Dmwind Direction	INWINDSPEED_max	Inwind Direction_max	Temperaturea
G85_1920_00340	2023-12-16 18:22:00.000	153	0.1	97.6	0	0	0.6	232	1	226	3.4
G85_1920_00340	2023-12-16 18:23:00.000	170	0.2	97.6	0	0	0.4	235	0.5	233	3.4
G85_1920_00340	2023-12-16 18:24:00.000	145	0.2	97.6	0	0	0.3	252	0.8	237	3.5
G85_1920_00340	2023-12-16 18:25:00.000	158	0.2	97.6	0	0	0.5	244	1	243	3.5

Table 2. Cont.

Stationnum	Date_time	Visibility	Temperature	Humidity	Precipitation	Meteorological_Phenomenon	Dmwind speed	Dmwind Direction	INWINDSPEED_max	Inwind Direction_max	Temperaturea
G85_1920_00340	2023-12-16 18:26:00.000	184	0.2	97.7	0	0	0.3	229	0.6	274	3.5
G85_1920_00340	2023-12-16 18:27:00.000	226	0.2	98.1	0	0	0.2	137	0.9	25	3.4
G85_1920_00340	2023-12-16 18:28:00.000	256	0.2	98	0	0	0.3	84	0.9	5	3.4
G85_1920_00340	2023-12-16 18:29:00.000	223	0.2	98	0	0	0.4	165	1.4	147	3.4
G85_1920_00340	2023-12-16 18:30:00.000	148	0.2	98.1	0	0	0.6	201	1.1	222	3.3
G85_1920_00340	2023-12-16 18:31:00.000	149	0.1	97.9	0	0	0.6	240	1.2	237	3.4
G85_1920_00340	2023-12-16 18:32:00.000	251	0.2	97.9	0	0	0.5	214	1.1	128	3.4
G85_1920_00340	2023-12-16 18:33:00.000	185	0.2	98	0	0	0.4	132	0.9	156	3.3
G85_1920_00340	2023-12-16 18:34:00.000	246	0.2	98	0	0	0.4	123	1	123	3.4
G85_1920_00340	2023-12-16 18:35:00.000	335	0.2	98.1	0	0	0.3	60	0.6	144	3.4
G85_1920_00340	2023-12-16 18:36:00.000	170	0.2	98.1	0	0	0.2	132	0.5	122	3.4
G85_1920_00340	2023-12-16 18:37:00.000	179	0.2	98.1	0	0	0.3	184	0.7	230	3.4
G85_1920_00340	2023-12-16 18:38:00.000	189	0.2	98	0	0	0.2	164	0.7	109	3.4
G85_1920_00340	2023-12-16 18:39:00.000	207	0.2	98	0	0	0.2	120	0.8	114	3.4
G85_1920_00340	2023-12-16 18:40:00.000	286	0.2	98	0	0	0.2	127	0.7	4	3.4
G85_1920_00340	2023-12-16 18:41:00.000	257	0.2	98	0	0	0.4	221	0.8	233	3.4
G85_1920_00340	2023-12-16 18:42:00.000	382	0.1	98	0	0	0.4	255	0.7	230	3.4
G85_1920_00340	2023-12-16 18:43:00.000	790	0.2	98	0	0	0.4	353	1.2	85	3.5
G85_1920_00340	2023-12-16 18:44:00.000	582	0.3	98.1	0	0	0.5	60	1.3	98	3.4
G85_1920_00340	2023-12-16 18:45:00.000	481	0.3	98.1	0	0	0.5	81	1.5	82	3.6
G85_1920_00340	2023-12-16 18:46:00.000	443	0.3	98.1	0	0	0.5	353	1.1	96	3.4
G85_1920_00340	2023-12-16 18:47:00.000	279	0.3	98.1	0	0	0.6	301	0.9	294	3.4
G85_1920_00340	2023-12-16 18:48:00.000	393	0.3	98.1	0	0	0.5	259	0.6	279	3.4

It is important to note that while the selected features across the two datasets may vary due to the distinct characteristics and availability of meteorological elements at different collection locations with different collecting devices, we have ensured a consistent dimensionality of six for the multivariate meteorological data used in our analysis. This standardization is crucial for maintaining the integrity and comparability of our predictive model's performance across diverse data environments. The adaptability of our proposed model to efficiently handle variations in feature sets underscores its robustness and applicability to real-world visibility prediction scenarios.

4. Methods

Inspired by the successful experiences of sequence data processing techniques in fields such as speech and natural language processing, we proposed the ATCNet model, a method for highway visibility prediction characterized by the integrated application of Transformer, CapsNet, and attention mechanisms. ATCNet is primarily composed of a time-series Transformer module and a Capsule Network module, with an attention module integrated before the output. Its process mainly includes the following steps. Initially, a multivariate time series of meteorological elements with a fixed time window size is input into a time-series Transformer module consisting of multiple layers of encoders and decoders. The Transformer module is used to capture long-term dependencies in time series data. Next, the processed meteorological element features are input into a Capsule Network module. The Capsule module captures both local and global features of the matrix composed of meteorological elements and the dynamic features arising from the temporal changes in meteorological elements. It further extracts and utilizes features associated with visibility values. Finally, an attention mechanism module is used to automatically learn the weights of different outputs, adaptively adjusting the representation of features to obtain the final output features. These features are then passed through a fully connected layer to output short-term road visibility predictions. The overall architecture of this model is illustrated in Figure 4. Please notice that in the figure, colors represent different elements across three sections. In the 'T' area, colors correspond to various meteorological elements. In the 'Time Series Transformer', colors denote distinct functional modules. Lastly, in the 'Capsule module', colors illustrate different weather data features learned by the Transformer module.

In ATCNet, integrating Transformers, Capsule Networks, and self-attention leverages their complementary strengths. The Transformer efficiently captures the temporal characteristics of data, the Capsule Network effectively understands the spatial relationships and hierarchical structures among multidimensional meteorological elements, and self-attention, as the final decision refinement step, ensures that all relevant time-series and spatial hierarchical information is fully utilized. This enhances the accuracy and reliability of the final predictions. This triple approach is crucial for comprehensively understanding highway visibility prediction tasks influenced by temporal changes and complex spatial factors.

This integration aligns with our architectural design, ensuring robust and accurate predictions essential for traffic management and safety. The design of ATCNet accounts for the complexity of different weather conditions and environmental changes, enabling it to adapt and provide reliable visibility forecasts even in extreme weather scenarios like heavy rain and severe haze.

4.1. Problem Formulation

In our research, the time series data of meteorological elements collected by the meteorological devices on the highways are divided by a time window of length n . The data sequence within a time window is represented as $T = \{T_1, T_2, \dots, T_n\}$, where $T_i \in R^d$ ($i = 1, 2, \dots, n$) represents a multidimensional vector of meteorological elements corresponding to the sequence at the i^{th} moment. T_i includes values of several meteorological elements such as temperature, precipitation, wind speed, humidity, and road surface

conditions. The input to the Transformer is a data matrix of a time window $T \in N \times D$, where N is the size of the window and D is the dimension of each sequence in the window. F represents the feature representation of the multivariate time series learned through the Transformer module. This feature representation is then input into the Capsule module and further processed through an attention module to produce the output of the model. Given a fixed-length time series of meteorological elements T , the task of the model is to predict the value of road visibility for the next period.

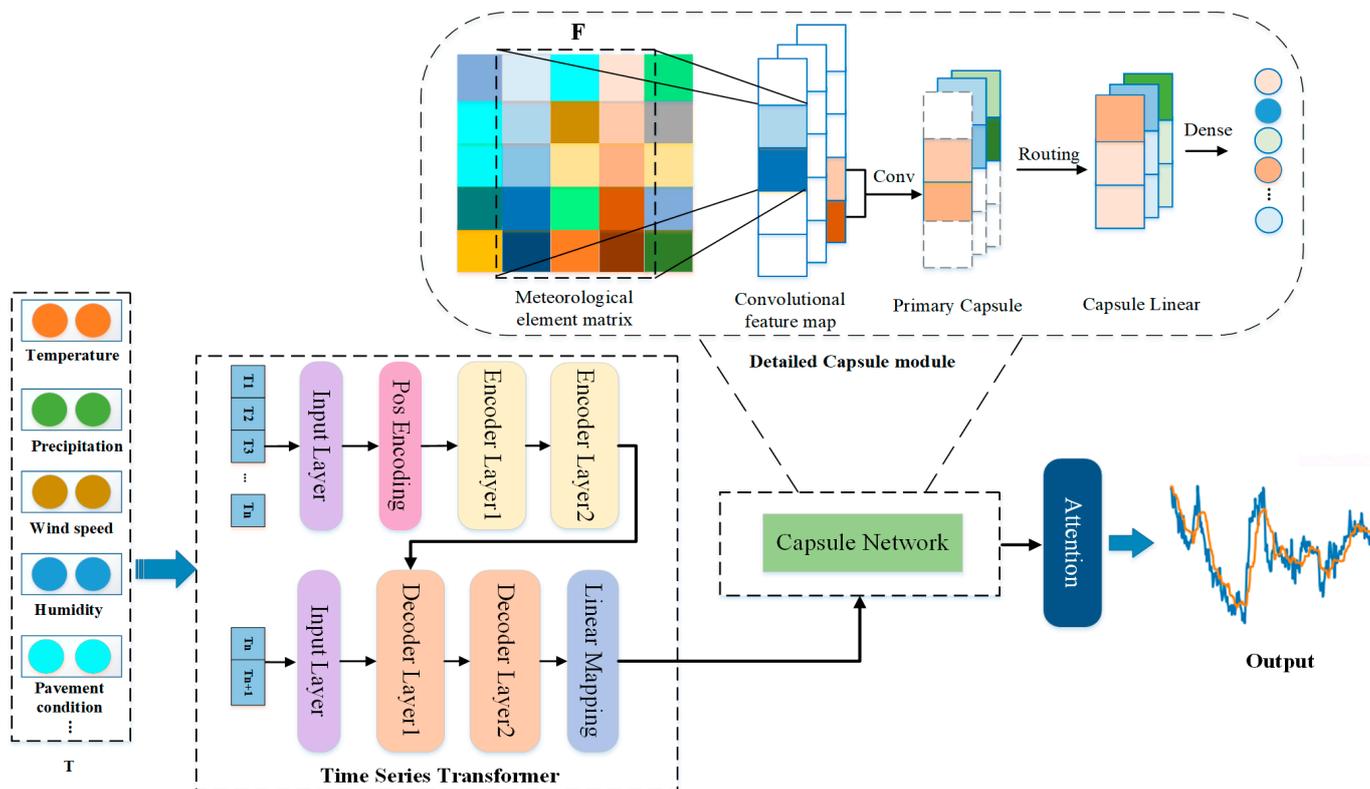


Figure 4. The overall structure of ATCNet.

4.2. Time Series Transformer

In this study, the core role of the time series Transformer is to capture the temporal dynamics of meteorological elements closely related to highway visibility. The input to the model is a multivariate time series dataset that includes temperature, precipitation, wind speed, humidity, and road surface conditions. Each meteorological element is encoded as a part of the time series data and processed through the Transformer to reveal how these elements change over time and how they interact with each other to influence visibility. The following Figure 5 illustrates the module composition of the Time Series Transformer part in the ATCNet architecture.

The Transformer network’s architecture is specifically designed to mitigate information loss during feature extraction, distinguished by its unique configuration of two encoders and two decoders. This design choice is pivotal in preserving critical information throughout the feature extraction process.

The dual-encoder and decoder setup allows for a more nuanced processing of the input data, ensuring that both global and local dependencies are captured and retained. The first encoder focuses on extracting broad, contextual information, while the second encoder refines this output, focusing on preserving detail-rich features. Similarly, the decoders work in tandem to reconstruct and predict the output, leveraging the comprehensive feature set preserved by the encoders.

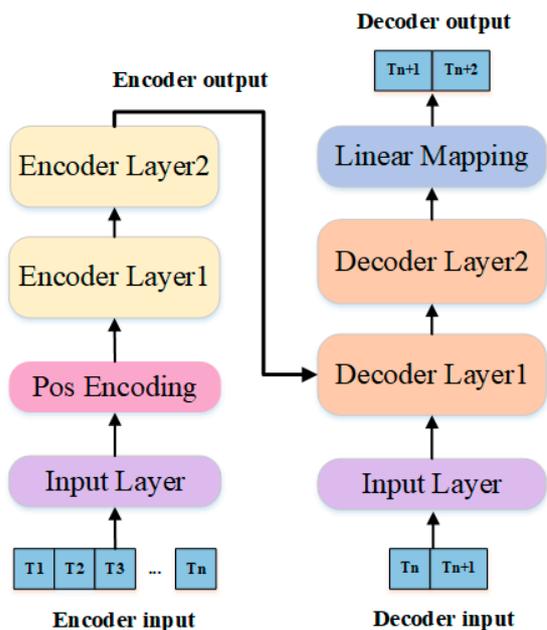


Figure 5. Diagram of Time Series Transformer module in ATCNet architecture.

In Time Series Transformer, the output of the encoder module is used as the input for the decoder module. In the encoder, each meteorological element’s time series is first transformed into a high-dimensional space to capture deep temporal features. Position encoding ensures that the model can understand the timing in the sequence, which is crucial for predicting future visibility. The multi-head self-attention mechanism allows the model to focus not only on the current meteorological condition at each time step but also consider conditions at other times, thus facilitating a comprehensive understanding of the factors influencing visibility.

The decoder part uses the deep temporal features provided by the encoder to generate predictions for future visibility. The introduction of masking ensures that predictions are based only on known historical information, preventing the model from “cheating” by seeing future data.

The encoder module takes historical traffic meteorological time series data as input. It consists of an input layer, a position encoding layer, and a stack of two identical encoder layers. The input layer maps the input time series data to vectors of a certain dimension through a fully connected network, crucial for the model to use multi-head attention mechanisms. Positional encoding with sine and cosine functions is used to encode the order information in the time series data by adding the positional encoding vectors to the elements of the input vectors.

Positional encoding (Position Encoding Layer): Firstly, the input layer projects the encoded feature representation into a matrix of hidden dimensions of length d . Positional encoding encodes this matrix to produce a vector of length d , uniquely representing an index in the token sequence. Positional encoding uses sine and cosine functions, as shown in Equation (2).

$$\begin{cases} PE_{pos, 2i} = \sin\left(pos / 10,000^{2i/d} \right) \\ PE_{pos, 2i+1} = \cos\left(pos / 10,000^{2i/d} \right) \end{cases} \quad (2)$$

In this formula, d represents the dimension of the input matrix, pos denotes the index of each meteorological element, and i indicates the index within the matrix. The term $2i$ corresponds to even positions (used in the sine function), and $2i + 1$ corresponds to odd positions (used in the cosine function).

The vectors generated by the positional encoding layer are fed into two consecutive encoder layers. Each encoder consists of a self-attention sublayer and a fully connected

feed-forward sublayer. Each sublayer is followed by a normalization layer. The encoder generates vectors of a certain dimension that are provided to the decoder.

The detailed structure of the encoder and decoder is depicted in Figure 6.

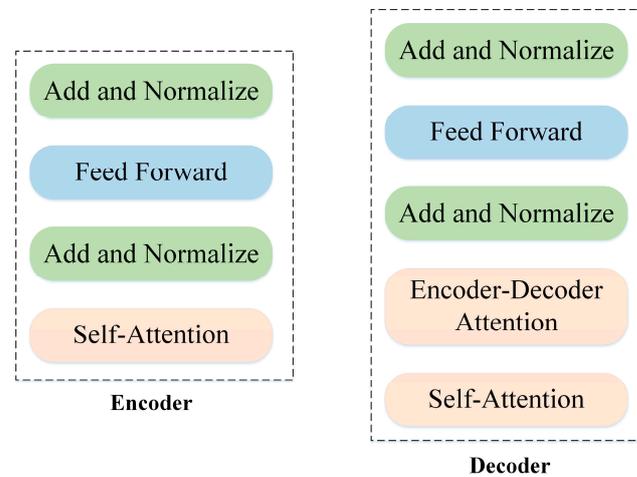


Figure 6. Diagram of the encoder and decoder in the Transformer module.

The decoder module predicts future values in an autoregressive manner, consisting of an input layer, a stack of two identical decoder layers, and an output layer. The decoder input starts from the last data point outputted by the encoder. The input layer maps the decoder input to vectors of a certain dimension. In addition to the two sublayers in each encoder, the decoder also inserts a third sublayer to apply self-attention mechanisms on the encoder output. Finally, the output layer maps the output of the last decoder layer to the target time series. The decoder can only rely on the previously generated parts when producing the output for each position. The Transformer uses attention masking to shield information from future positions to ensure the model does not utilize future information.

Add and Normalize (residual connection) is a critical module in both the encoder and decoder. After each sublayer, the Transformer uses residual connections and layer normalization to stabilize the training process. This helps to avoid the problem of gradient vanishing or exploding in deep neural networks. The network can focus only on the current differences by applying elements from the lower layer outputs to the higher layers. ‘Add’ involves adding the data before self-attention to the data after self-attention. ‘Normalize’ uses layer-normalization, which mainly includes two steps:

Step 1: Normalize each value (i.e., subtract the mean and divide by the standard deviation).

Step 2: Perform an affine transformation on the values obtained in the first step using two learned scalars, γ and β , as shown in Equation (3):

$$y_i = \gamma \hat{x}_i + \beta \quad (3)$$

In the formula, y_i represents the final normalized value, and \hat{x}_i is the value normalized in the first step.

The feed-forward layer in the Transformer architecture consists of two fully connected layers. The first layer uses the ReLU activation function, and the second layer does not use an activation function, as shown in Equation (4).

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (4)$$

In the formula, x is the output from the previous layer, W_1 and b_1 are the weight and bias of the first layer, and W_2 and b_2 are the weight and bias of the second layer, respectively. These parameters are typically initialized with random values and then adjusted through the backpropagation algorithm during model training.

The contributions of Transformer in the ATCNet model are as follows:

- **Powerful time series data processing capability:** The ability of the Transformer to process serialized data is crucial for highway visibility prediction. Thanks to its self-attention mechanism, the Transformer excels at understanding the temporal dependencies present in such data. For instance, the Transformer can analyze several days or even weeks of meteorological data to identify complex patterns that might lead to sudden changes in visibility. This is vital for accurately predicting highway visibility.
- **Fine feature extraction and multivariate data fusion:** The Transformer accurately processes multidimensional data. It can extract details from each meteorological factor and also fuse these factors to provide a more comprehensive visibility prediction. For example, the Transformer can provide a comprehensive visibility prediction analysis of the combined effects of temperature, humidity, and wind speed. This integration is necessary for a thorough understanding of the conditions that affect visibility.
- **Efficient data parallel processing capability:** The Transformer model can process the entire data sequence simultaneously, making it more efficient when dealing with large volumes of time series data, such as minute-level highway meteorological data. This is particularly important for real-time prediction capabilities, as it allows the model to quickly adapt to and respond to the latest changes in meteorological conditions.

4.3. Capsule Network

In ATCNet, CapsNet is introduced to further process the features extracted by the Time Series Transformer, extracting spatial features from the matrix composed of time series. This allows for a more detailed understanding of the spatial aspects of meteorological data. By incorporating a Capsule Network after the Transformer module, the model can learn the complex spatial hierarchies among these meteorological elements and utilize this information during the prediction process. The dynamic routing mechanism further strengthens the correct combination of features, making the predictions more accurate. Figure 7 illustrates the module composition of the Capsule Network part in the ATCNet architecture. The assorted colors reflect a spectrum of meteorological data characteristics that have been extracted by the Transformer module, highlighting the nuanced feature mapping capabilities of the capsule network.

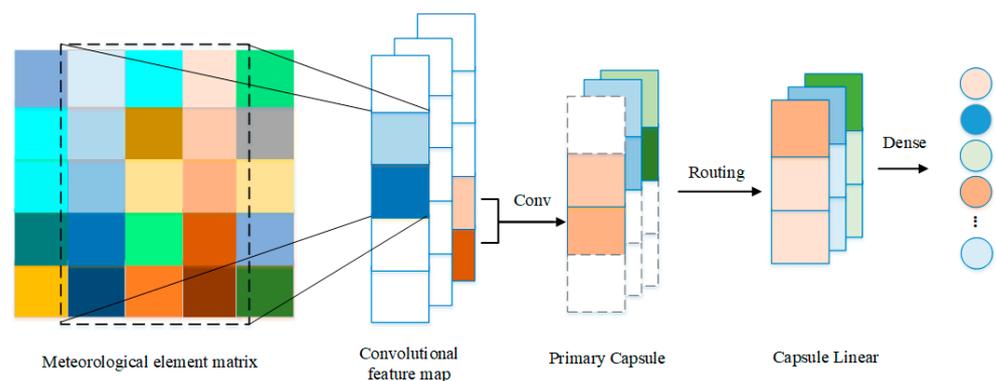


Figure 7. Diagram of Capsule Network module in ATCNet architecture.

The output of the Transformer module is then passed to the Capsule Network. In this module, a convolutional layer first transforms the output of the Transformer into a series of convolutional feature maps. These maps are passed to the Primary Capsule layer, which consists of a set of capsules, each capturing a local combination of a group of features.

The feature matrix outputted by the Time Series Transformer module is first considered as a single-channel image and processed by a convolutional kernel to produce an output tensor. This tensor serves as the input to the primary capsule layer. Each capsule in the primary capsule layer encodes specific features detected in the matrix and their instance parameters (such as position, size, orientation, etc.). Then, a routing algorithm is used to

compute the optimal feature combinations. These combinations, as the features ultimately extracted by the capsules, are processed by a Capsule Linear layer, which transforms the spatial relationship representations of these feature vectors.

The core idea of CapsNet is to use “capsules” to represent hierarchical features, which can better capture the spatial relationships and pose information of targets. Compared to traditional pooling layers and fully connected layers, capsule networks represent features in a vector form and introduce a dynamic routing mechanism, allowing the network to automatically learn the hierarchical relationships between different parts. Suppose we have a layer containing N capsules, each represented by a vector. The direction and length of the vector typically represent the probability of the existence of a specific type of feature and some of its attributes. Each capsule is represented by a d -dimensional vector, denoted as $v_j = (v_{j1}, v_{j2}, \dots, v_{jd})$, where v_j represents the output vector of the j th capsule.

The operation of a single capsule in a capsule network is illustrated in Figure 8:

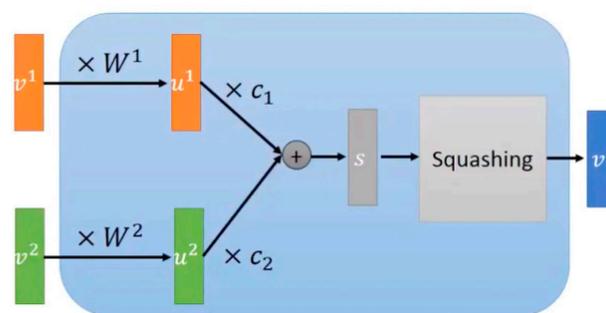


Figure 8. Manipulation of individual capsules.

The operation of a single capsule in a capsule network involves the following four steps:

1. Multiplication of input vectors: The input vectors v^1 and v^2 , which are outputs from the previous capsules, are each multiplied by their respective weights W^1 and W^2 within a single capsule. This results in new vectors u^1 and u^2 .
2. Scalar weighting of input vectors: The input vectors u^1 and u^2 are scalar weighted by multiplying them with routing coefficients c_1 and c_2 respectively. These routing coefficients are scalars and satisfy the condition $c_1 + c_2 = 1$.
3. Summation of vectors: The vectors obtained are summed to produce s , which is calculated as $s = c_1u^1 + c_2u^2$.
4. Vector-to-vector non-linearity: The resultant vector s is transformed to produce the vector v . This transformation involves compressing the vector such that its length lies between 0 and 1 while maintaining its direction. This non-linear transformation is represented by the following Equation (5):

$$v = \text{Squash}(s) = \frac{\|s\|^2}{1 + \|s\|^2} \frac{s}{\|s\|} \quad (5)$$

v as the output of this capsule can be used as input to the next capsule.

The outputs of these primary capsules are then passed to higher-level capsules through a dynamic routing process. The dynamic routing algorithm enables the model to learn the spatial hierarchies in the input data, achieved by iteratively adjusting the connection weights between capsules. The routing process can be described using the following algorithm:

1. For all primary capsule i and all secondary capsule j , initialize connection weight $b_{ij} = 0$;
2. For each i and j , calculate $c_{ij} = \text{softmax}(b_{ij})$ where c_{ij} is the contribution weight of capsule i to capsule j ;
3. The input for each secondary capsule s_j is calculated as $s_j = \sum_i c_{ij}u_{ij}$, where u_{ij} is the prediction vector of the primary capsule i ;

4. Apply a nonlinear activation function $v_j = \text{squash}(s_j)$, where v_j is the output vector of capsule j .
5. Update b_{ij} , passed $b_{ij} \leftarrow b_{ij} + u_{ij} \cdot v_j$.

The contributions of the Capsule Network in ATCNet are as follows:

- Spatial hierarchies of multiple meteorological elements: Thanks to its unique hierarchical structure, the Capsule Network is particularly effective in understanding the spatial relationships between different meteorological elements. It processes not only individual element data but also automatically recognizes the combinations and interactions of these elements. This capability is crucial for considering how different weather elements interact to affect visibility, such as how a combination of high humidity and low temperature might impact visibility.
- Contextual understanding through dynamic routing: The Capsule Network uses a dynamic routing mechanism to understand the context and significance of different features in multidimensional meteorological data. For visibility prediction, this means the model can prioritize the most relevant weather features under varying conditions, enhancing prediction accuracy.
- Robustness to changes: Equivariance in Capsule Networks, where subtle changes in the input lead to predictable changes in the output, is highly beneficial for highway visibility prediction tasks. Highways may experience various visibility conditions due to different weather situations. The equivariance property in Capsule Networks significantly enhances the model's robustness, enabling it to effectively process and predict visibility under diverse meteorological conditions.

4.4. Attention

After the Capsule Network, an attention layer is applied to further refine the feature representation. This attention mechanism focuses on the most informative features based on the requirements of the prediction task.

The role of the attention mechanism [29] in this model is crucial. When processing features outputted by the Transformer and Capsule Network, the attention layer identifies which features are most important for predicting visibility. Given that different meteorological conditions have varying impacts on visibility, the attention layer reflects these differences by allocating different weights to the features. For instance, in foggy weather conditions, humidity and road surface conditions might be more critical than temperature. The attention mechanism, by weighting these features, ensures that the prediction model focuses on the most critical factors under the current environmental conditions.

It is important to note that while the Transformer itself includes an attention mechanism, adding an additional attention mechanism before the output can still be valuable. This additional layer can provide a more refined, targeted feature weighting method specifically for visibility prediction tasks. It might enable the model to adaptively adjust the final feature representation of the output, thereby potentially increasing the accuracy and robustness of the predictions. The results of the ablation experiments in Section 5.5 demonstrate that this attention module is indispensable.

In implementation, we opted to use self-attention rather than standard attention. Standard attention is typically used in sequence transformation tasks, where one sequence's information is used to guide the generation of another sequence. Self-attention, by design, allows the model to weigh and prioritize different parts of the input meteorological data based on their relevance to the task at hand. This is particularly advantageous in tasks like visibility prediction, where not all input features contribute equally to the outcome. Theoretically, self-attention mechanisms offer a more dynamic and flexible approach to understanding data relationships compared to traditional methods. They enable the model to focus on the most informative features without being constrained by the sequential nature of the data. This adaptability results in a more nuanced and accurate representation of the data, leading to improved prediction outcomes.

The attention mechanism can score each dimension of the input data, then weigh the features based on these scores to highlight essential characteristics. The attention mechanism may influence downstream models or modules. The attention mechanism can be described by Equation (6).

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

In the formula, Q represents the query vector, K is the key vector, and V is the value vector. Q , K , and V are all weight matrices that are initially random and then optimized during the gradient descent process of training data. The use of the attention mechanism typically involves the following steps:

1. Mapping inputs: Each element of the input sequence (such as different meteorological elements) is mapped to a high-dimensional vector space, forming a matrix X , where each row represents a vector representation of an element.
2. Computing attention weights: For each element in the matrix X , calculate the similarity (or relevance) to all other elements using Q , K , and V . These similarities are used as weights for a weighted sum.
3. Weighted sum: Based on the calculated attention weights, perform a weighted sum of the vectors V to produce the final output.

The contribution of the Attention module in ATCNet is as follows:

- Feature Integration: In the final stage of the model, the Attention module is used to integrate all hierarchical features extracted by the Transformer and Capsule Network. This assists in further refining the model's global understanding of the input data before making the final prediction.
- Importance Weighting: Although earlier layers have already assessed the importance of features, self-attention evaluates these features' contributions to the final output at a higher level. This weighting is based on a deep understanding of the entire data flow, optimizing the model's decision-making process.
- Output Refinement: In the final stage of the model, Attention is responsible for mapping more abstract feature relationships, which might not have been as apparent in earlier processing stages. The placement of Attention just before the output indicates that it is used for refining and adjusting the final predictions, ensuring that all relevant features and their interactions are considered.

4.5. Output

The output structure of the ATCNet model is a predicted time series, succinctly expressed as $\hat{Y} = \{\hat{y}_{t+1}, \hat{y}_{t+2}, \dots, \hat{y}_{t+m}\}$. In this notation, the symbol \hat{y}_{t+k} represents the visibility prediction value at the future time point $t+k$, where m represents the time step from the current time point in t . This sequence format provides a clear and continuous representation of the forecast, extending from the immediate future to the specified forecast time step. Each element in the series reflects a discrete point in time, providing a continuous and detailed view of the predicted values over the forecast period.

The output stage of the model encapsulates the collaborative processing results of the entire ATCNet architecture. This includes the temporal dependencies captured by the Transformer module, the spatial features optimized by the CapsNet module, and the final feature representation outputted after adaptive weight adjustment by the Attention module. This output demonstrates the effective integration of advanced feature extraction with sequential analysis.

5. Experiments and Analysis

5.1. Dataset Division

In our experiments, we applied the N Fold technique to both the WD13VIS and WDVigoVis datasets to ensure a robust evaluation of our model's performance. Specifically,

we utilized a 5-Fold cross-validation approach, where each dataset was divided into five equal parts. In each iteration, four parts were used for training, and the remaining part was used for testing. This process was repeated five times, with each part serving as the test set once, ensuring that every data entry contributed to both the training and testing phases. This method allowed us to comprehensively assess the model's performance across different subsets of data, enhancing the reliability of our findings.

The experimental results presented in Sections 5.4–5.6 are based on the averages obtained from the N Fold technique.

5.2. Experimental Environment Configuration and Model Parameter Settings

The experiments were conducted on a server running Windows Server 2019, equipped with an Intel(R) Xeon(R) Silver 4210R CPU, 256 GB of memory, a 6-TB hard drive, and 6 NVIDIA GeForce GTX 3090 GPUs. Python 3.9 was used as the programming language, and TensorFlow 2.10.1 was employed as the machine learning software development library.

Below are the relevant parameters of the ATCNet model used in this paper (Table 5):

Table 5. Parameter settings for ATCNet model.

Parameter	Description	Value
Lr	learning rate	0.001
Bs	batch size	256
Activation	activate the function	Sigmoid
Ts	time step (min)	15/30/60/120
Tw	time Window (h)	0.25/0.5/1/2/3/4/5/6/10
Epoch	training iterations	100
Dropout	dropout rate	0.01
head_size	size of each head in multi-head attention	128
num_heads	number of heads in multi-head attention	4
ff_dim	dimensions of the feedforward network	4
num_transformer_blocks	number of Transformer codecs	4
total_params	total number of model parameters	2,113,809

5.3. Evaluation Metric

In this study, we utilized five evaluation metrics: MSE, MAE, MAPE, R-squared (R^2) and NMBD to comprehensively assess the model's performance. MSE reflects the model's performance under extreme conditions by assigning higher penalties to larger errors, crucial for predicting highway visibility in extreme weather, where significant prediction deviations can lead to severe safety issues. In contrast, MAE offers an intuitive understanding of the magnitude of prediction errors, representing a more stable error metric that helps evaluate the model's average performance under normal conditions. Meanwhile, MAPE measures the percentage of the model's prediction error relative to the actual values, allowing us to assess the accuracy of the model's predictions irrespective of the scale of the actual data. This is particularly important for highway visibility data of varying scales and ranges, as it ensures consistency in model evaluation. R-squared evaluates the proportion of variance in highway visibility data captured by our model, offering insight into the predictive accuracy and effectiveness in various weather conditions. Lastly, NMBD quantifies the model's systematic bias in visibility forecasts, ensuring balanced accuracy across different visibility levels.

Here are the definitions of these metrics:

- Mean Squared Error Loss (MSE): The average of the absolute squared error between the predicted value and the actual value. The smaller the MSE value, the better the predictive power of the model.
- Mean Absolute Error Loss (MAE): Measure the average absolute error between the predicted and actual values. It is used to evaluate the model's prediction accuracy. A smaller MAE value indicates that the model performs better.

- Mean Absolute Percentage Error (MAPE): A normalized version of MAE that is more sensitive to relative error and is not affected by the absolute value of the target variable. A smaller MAPE value indicates that the model has higher prediction accuracy.
- R-squared (R^2): Measures the proportion of variance in the dependent variable explained by the independent variables, indicating model fit quality. Higher R^2 signifies better predictive accuracy.
- Normalized Mean Bias Deviation (NMBD): Assesses model accuracy by calculating the mean prediction bias, normalized by observed values' mean. NMBD close to 0 means minimal bias.

These metrics are used to evaluate the model's performance in a time series forecasting task. The five indicators are defined in Equations (7)–(11) below:

$$MAE(y, \tilde{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} |y_i - \tilde{y}_i| \quad (7)$$

$$MSE(y, \tilde{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \tilde{y}_i)^2 \quad (8)$$

$$MAPE(y, \tilde{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} \frac{|y_i - \tilde{y}_i|}{|y_i|} \times 100\% \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n_{samples}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{samples}} (y_i - \bar{y})^2} \quad (10)$$

$$NMBD = \frac{\frac{1}{n_{samples}} \sum_{i=1}^{n_{samples}} (y_i - \hat{y}_i)}{\frac{1}{n_{samples}} \sum_{i=1}^{n_{samples}} y_i} \quad (11)$$

where $n_{samples}$ represents the number of samples, y_i is the actual value, \tilde{y}_i is the predicted value, and \bar{y} is the average value of the actual value.

5.4. Performance Comparison with Different Window Sizes and Time Steps

In time series prediction tasks, the size of the time window and the length of the prediction time step are key parameters. The time window size indicates the length of the historical data used, while the prediction time step determines the length of the forecast sequence output by the model.

To study the performance of the ATCNet model in short-term road visibility prediction under different combinations of time window sizes and prediction time step lengths, we set the following time window sizes: 15 min, 30 min, 1 h, 2 h, 3 h, 4 h, 5 h, 6 h, and 10 h. Additionally, we set the prediction time step lengths to 15 min, 30 min, 1 h, and 2 h.

We conducted experiments on the WD13VIS dataset and evaluated the model's performance under different configurations using the MSE and MAE metrics. The specific results are presented in Tables 6 and 7.

Single-step forecasting predicts the next observation in time series prediction tasks. The previous tables show that multi-step forecasting methods depend a lot on the time window size and the prediction time step length. A fixed window size leads to higher error as the prediction time step grows. A fixed prediction time step has an optimal window size that maximizes performance. Larger window sizes can have redundant features that lower performance. This optimal window size is wider for longer prediction time steps because they need more feature information from a longer input time series.

Table 6. Model’s MAE under different time window and forecast time step combinations. For a fixed prediction time step, the optimal window size that maximizes performance is indicated in bold.

Window Size	Forecast Time Step			
	15 min	30 min	1 h	2 h
15 min	0.411	1.021	2.112	3.212
30 min	0.323	0.733	1.122	1.617
1 h	0.113	0.419	0.944	1.323
2 h	0.035	0.065	0.325	0.902
3 h	0.021	0.033	0.314	0.811
4 h	0.024	0.027	0.265	0.732
5 h	0.025	0.034	0.263	0.680
6 h	0.026	0.045	0.268	0.665
10 h	0.033	0.062	0.272	0.673

Table 7. Model’s MSE under different time window and forecast time step combinations. For a fixed prediction time step, the optimal window size that maximizes performance is indicated in bold.

Window Size	Forecast Time Step			
	15 min	30 min	1 h	2 h
15 min	0.1051	0.2121	0.3465	0.6013
30 min	0.0851	0.1021	0.2173	0.3911
1 h	0.0502	0.0702	0.1039	0.3480
2 h	0.0120	0.0113	0.0796	0.1375
3 h	0.0027	0.0031	0.0128	0.0311
4 h	0.0029	0.0024	0.0095	0.0112
5 h	0.0032	0.0041	0.0088	0.0101
6 h	0.0039	0.0058	0.0069	0.0087
10 h	0.0044	0.0063	0.0074	0.0098

In practical applications of highway visibility forecasting, short prediction time steps, such as 15 min, or longer ones, like over 1 h, do not significantly benefit highway management. The most suitable prediction time step is 30 min. Therefore, combining practical needs and experimental results, this paper ultimately selected a 4 h time window size and a 30 min prediction time step as the optimal parameters for ATCNet.

5.5. Ablation Experiment

To validate the rationale and superiority of our proposed model architecture, and to better demonstrate the contributions of different modules in the ATCNet architecture, we conducted ablation experiments on the WD13VIS dataset. We removed the CapsNet and Attention modules from ATCNet and compared the changes in model performance. Based on the experimental results in Section 5.3, we set the time window to 4 h and the prediction time step to 30 min.

Table 8 shows the performance metrics MSE, MAE, MAPE, R^2 , and NMBD of each module in our model in the individual ablation experiments.

Table 8. Ablation experiments of the ATCNet model on WD13VIS datasets.

Model	MSE	MAE	MAPE	R^2	NMBD
Transformer	0.0207	0.060	0.2806	0.853	0.024
Transformer + CapsNet	0.0124	0.048	0.1528	0.948	0.015
Transform + Attention	0.0131	0.042	0.1445	0.942	0.019
Transformer + CapsNet + Attention(Ours)	0.0024	0.027	0.0414	0.987	0.005

The following Table 9 provides a detailed analysis of the results for each step of the ablation experiment:

Table 9. Analysis of the results for each step of the ablation experiment.

Model Combination	Analysis
Transformer Only	Using only the Transformer module, the model focuses on capturing long-term dependencies in time series data. The results at this stage indicate that although the Transformer can effectively process time series data, it has limitations in dealing with the complexity of visibility prediction. This reflects the inadequacy of a single module in capturing the spatial features and subtle changes in multidimensional meteorological data.
Transformer + CapsNet	The introduction of CapsNet significantly improved model performance. CapsNet helps capture the spatial relationships and hierarchical structures between meteorological elements, which is crucial for accurate visibility prediction. The performance improvement highlights the Capsule Network's strong ability to understand complex interactions in multidimensional data.
Transformer + Attention	The addition of the attention mechanism led to an improvement in model performance. The attention mechanism enables the model to better focus on the most critical parts of the time series data, thus enhancing prediction accuracy. This improvement demonstrates the importance of adaptively focusing on significant features when processing time series data.
Transform + CapsNet + Attention(ours)	The complete model exhibited the best performance on all metrics, proving the integration of Transformer, CapsNet, and attention mechanisms is crucial for enhancing the model's accuracy and robustness in predicting visibility. This combination effectively merges the strengths of each module, providing a more comprehensive understanding of input features, thus enabling the model to generalize better to new data and offer more accurate predictions.

Through the above analysis, it is clear how each component uniquely contributes to the model's performance and how they work together to enhance overall efficiency. The ablation study results not only reveal the importance of each module but also demonstrate how they complement each other in the entire prediction task. Particularly in the complex task of visibility prediction, the model needs to simultaneously understand the long-term dependencies of time series, the spatial relationships of multidimensional meteorological data, and the dynamic changes of important features. This integrated approach precisely meets these requirements.

The results validate the significant advantage of the ATCNet model, which combines these technologies to predict highway visibility accurately. It provides an effective solution to enhance traffic safety and efficiency, showcasing the potential of such sophisticated models in real-world applications, especially in critical areas like transportation.

5.6. Performance Comparison with Competitive Models

We conducted experiments to validate the performance of our proposed ATCNet model in highway visibility prediction tasks. We compared it with several of the most competitive time series prediction methods, using the WD13VIS and WDVigoVis datasets. Based on the results from Section 5.3, for the WD13VIS dataset, we set the time window to 4 h and the time step to 30 min. For the WDVigoVis dataset, both the time window and time step were also set to 4 h and 30 min.

We used several of the most competitive time series prediction methods for comparison with ATCNet, including the traditional ARIMA model, the machine learning decision tree ensemble algorithm XGBoost, various LSTM deep learning models, and the latest time series prediction model, Informer. We prepared corresponding datasets for each competitive model and performed customized preprocessing steps according to their specific requirements. For example, for the ARIMA model, we performed differencing operations to stabilize the mean of the time series. Each model (including ARIMA, XGBoost, LSTM variants, and Informer) underwent a rigorous configuration process, using grid search methods to adjust hyperparameters and determine the best settings.

Here is a brief description of these models:

- ARIMA [30]: A differential autoregressive moving average model for predicting non-stationary time series.

- XGBoost [31]: An efficient gradient-boosting decision tree algorithm that combines multiple weak learners into a single strong learner by forward addition.
- LSTM [32]: A variant of RNN [33] that is commonly used to deal with nonlinear features in time series.
- LSTM + CNN [34]: Combining CNN and LSTM networks to extract spatial and temporal features, respectively.
- GRU + Attention [35]: An LSTM-based model variant that merges the forgetting gate and the input gate into an update gate, and emphasizes the importance of the output of each hidden layer through the attention mechanism.
- Informer [36]: Employing a new attention mechanism that automatically adjusts the attention range according to the sequence length, effectively processing long sequences. It also employs a multi-scale time encoder/decoder structure that considers information at different time scales.

Tables 10 and 11 showcase the performance metrics MSE, MAE, MAPE, R^2 and NMBD for visibility prediction across the WD13VIS and WDVigoVis datasets, respectively. It is important to highlight that the observed differences in these metrics can be attributed to the distinct sampling frequencies of the datasets: WD13VIS with a 1 min granularity provides a more detailed temporal resolution compared to WDVigoVis's 30 min sampling rate, influencing the absolute values of the evaluation metrics. This discrepancy should not be interpreted as a bias towards the WD13VIS dataset; rather, it highlights the model's adaptability and consistent performance across varying data environments, affirming its general applicability.

Table 10. MSE, MAE, and MAPE for various models on WD13VIS datasets.

Model	MSE	MAE	MAPE	R2	NMBD
ARIMA [30]	0.0667	0.657	0.2726	0.612	0.102
XGBoost [31]	0.0371	0.343	0.2383	0.695	0.088
LSTM [32]	0.0254	0.132	0.1853	0.764	0.079
LSTM + CNN [34]	0.0223	0.086	0.1027	0.825	0.043
GRU + Attention [35]	0.0192	0.078	0.0895	0.874	0.024
Informer [36]	0.0145	0.064	0.0733	0.933	0.018
ATCNet (ours)	0.0024	0.027	0.0414	0.987	0.005

Table 11. MSE, MAE, and MAPE for various models on the WDVigoVis dataset.

Model	MSE	MAE	MAPE	R ²	NMBD
ARIMA [30]	0.1233	0.295	0.4100	0.644	0.135
XGBoost [31]	0.0971	0.262	0.3822	0.687	0.092
LSTM [32]	0.0874	0.246	0.3703	0.732	0.083
LSTM + CNN [34]	0.0822	0.227	0.3612	0.810	0.047
GRU + Attention [35]	0.0628	0.193	0.3146	0.858	0.030
Informer [36]	0.0517	0.174	0.2213	0.914	0.014
ATCNet (ours)	0.0312	0.120	0.1450	0.967	0.008

The above tables show that neural network models significantly outperform the machine learning model XGBoost and the traditional time series prediction model ARIMA in visibility prediction on these two datasets. The performance of LSTM and its variants is comparable.

The ATCNet model proposed in this paper, which combines Transformer, Attention, and CapsNet modules, can effectively extract multivariate time series and spatial feature information, and adaptively assign weights to each feature. It significantly outperforms other models.

On the WD13VIS dataset, compared to the latest time series prediction model architecture, Informer [36], ATCNet reduced the Mean Squared Error (MSE) by 1.21% and the Mean

Absolute Error (MAE) by 3.7%. Similarly, on the WDVigoVis dataset, ATCNet reduced MSE and MAE by 2.05% and 5.4%, respectively. This demonstrates ATCNet's outstanding performance in visibility prediction tasks in various scenarios, such as highways and airports, showcasing its strong universality.

Figure 9 clearly illustrates the performance advantage of the ATCNet model, providing an intuitive comparison of various model evaluation metrics.

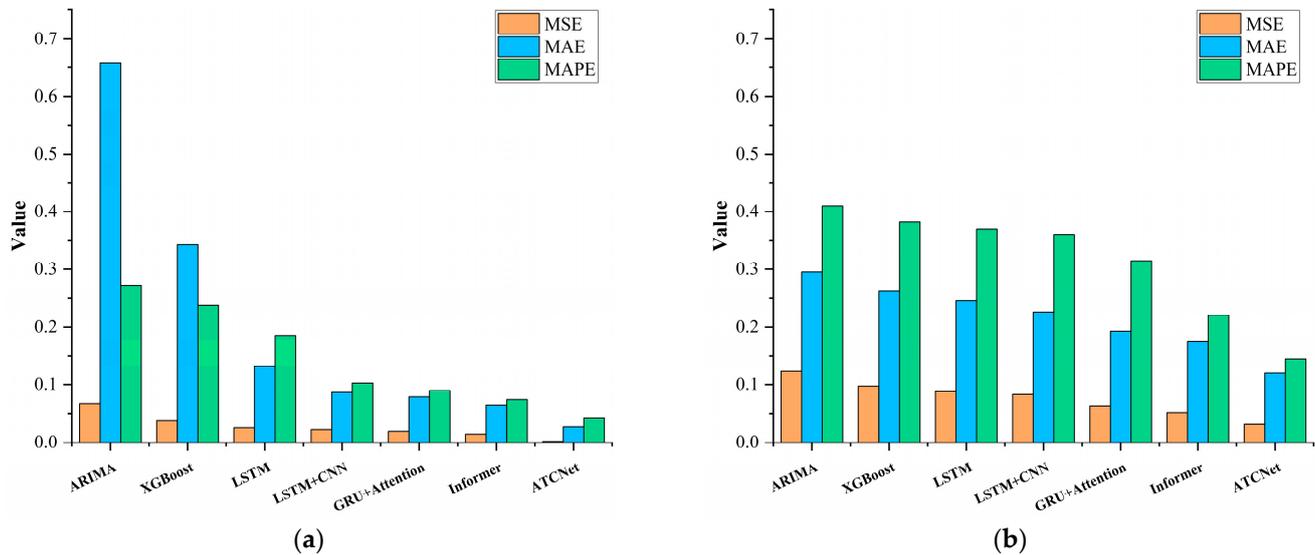


Figure 9. Model performance comparison of (a) WD13VIS dataset and (b) WDVigoVis dataset.

Experiments were conducted on the WD13VIS dataset, comparing actual and predicted visibility values. Figure 10 illustrates the comparison results across four different time intervals. Specifically, the y-axis represents the normalized predicted and actual visibility values, with each unit on the x-axis corresponding to a real-time minute. For instance, a significant increase in visibility values can be observed in the upper left subplot of Figure 10 during the 100 to 200 min interval, while the upper right subplot shows a marked decrease in visibility within the same interval. These results demonstrate that the ATCNet model can accurately predict sudden low-visibility events, with the trends in the predicted values consistent with actual changes, thereby confirming the model's effective performance.

In summary, through comparative analysis on the WD13VIS and WDVigoVis datasets, we have demonstrated that our proposed model exhibits robustness and adaptability in different data environments. The model shows good performance and predictive ability under various conditions, underscoring its potential for wide-ranging applications, particularly in scenarios where accurate visibility prediction is crucial. This versatility and effectiveness reinforce the model's suitability for practical deployment in diverse meteorological and traffic conditions.

5.7. Practical Application System Validation

Currently, the model has been successfully integrated into the "Highway Traffic Meteorological Intelligent Monitoring and Proactive Control System" [37] independently developed by our research team for the Yunnan Province Transportation Investment and Construction Group Co., Ltd which is located in Kunming City, China. This system has been deployed on three highways managed by the Qujing Management Office—ZhaoHui, DaiGong, and QuSheng—for the verification of accuracy in visibility prediction tasks. The practical verification results aligned with anticipated standards, allowing the system to aid local traffic management authorities in real traffic control.

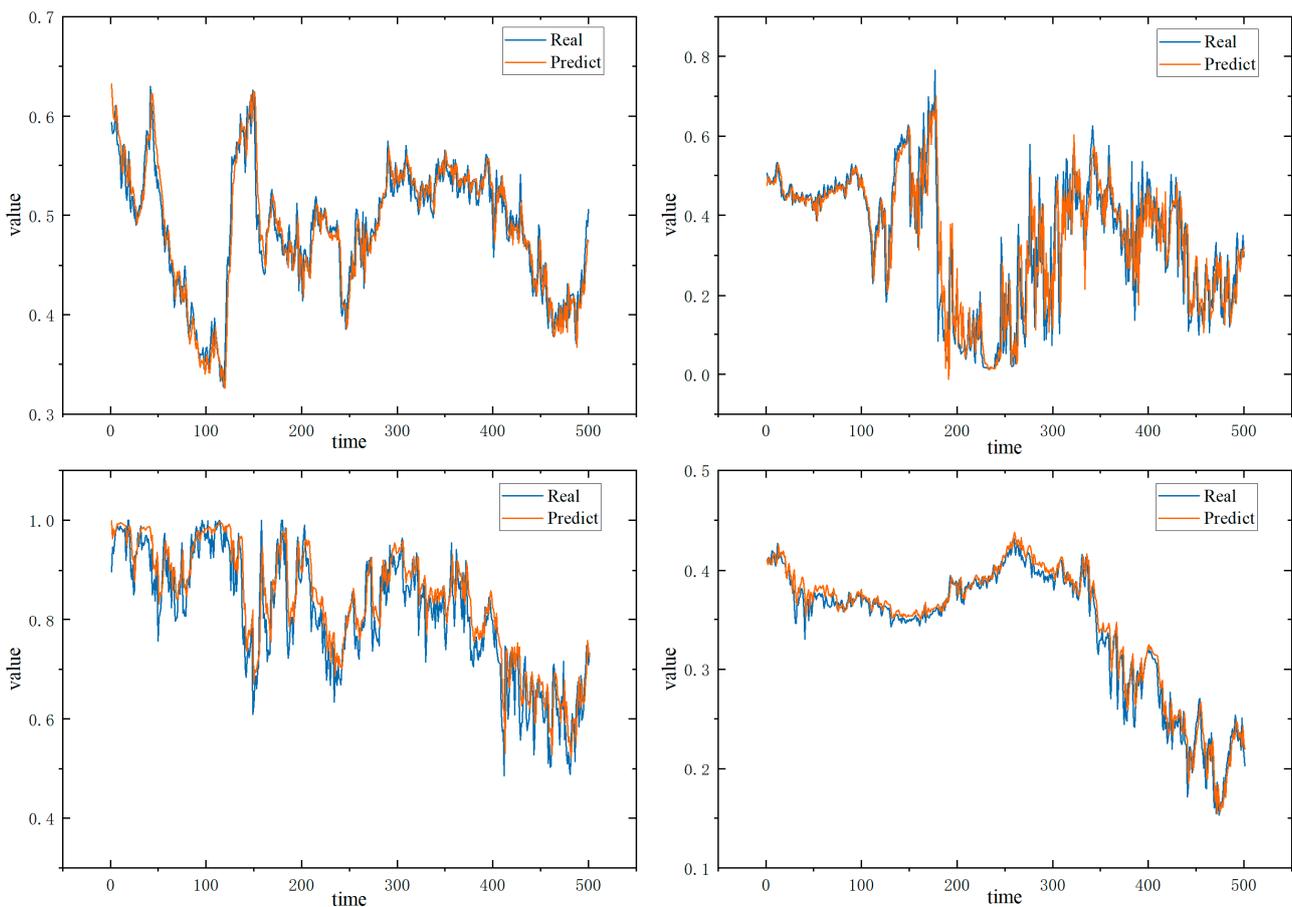


Figure 10. Comparison of predicted and actual values for four different time series intervals on the WD13VIS dataset.

We have included detailed screenshots that showcase the system’s interface in Figure 11, including the future visibility trend predictions over a 3 h window. These screenshots are intended to provide a clear view of how our model’s predictions are presented within the system. Furthermore, the system’s backend automatically compares the predicted visibility values against the actual ground truth values collected during the same intervals. This process, based on the algorithm depicted in Figure 11, assesses the model’s prediction accuracy. If significant discrepancies are noted, the system triggers alerts to indicate the potential need for recalibrating meteorological data collection devices or to verify that the visibility prediction program is functioning correctly.

Two challenges encountered and resolved in actual applications were: firstly, the real-time processing of incomplete or inaccurate meteorological data. We maintained the model’s accuracy and reliability in scenarios of data loss or sensor malfunction by employing data interpolation and anomaly detection techniques. Secondly, optimization of the algorithm’s computational efficiency and data processing workflow was undertaken to minimize prediction delays. This was achieved by utilizing parallel computing and data caching mechanisms, ensuring timely updates of predictions under rapidly changing weather conditions.

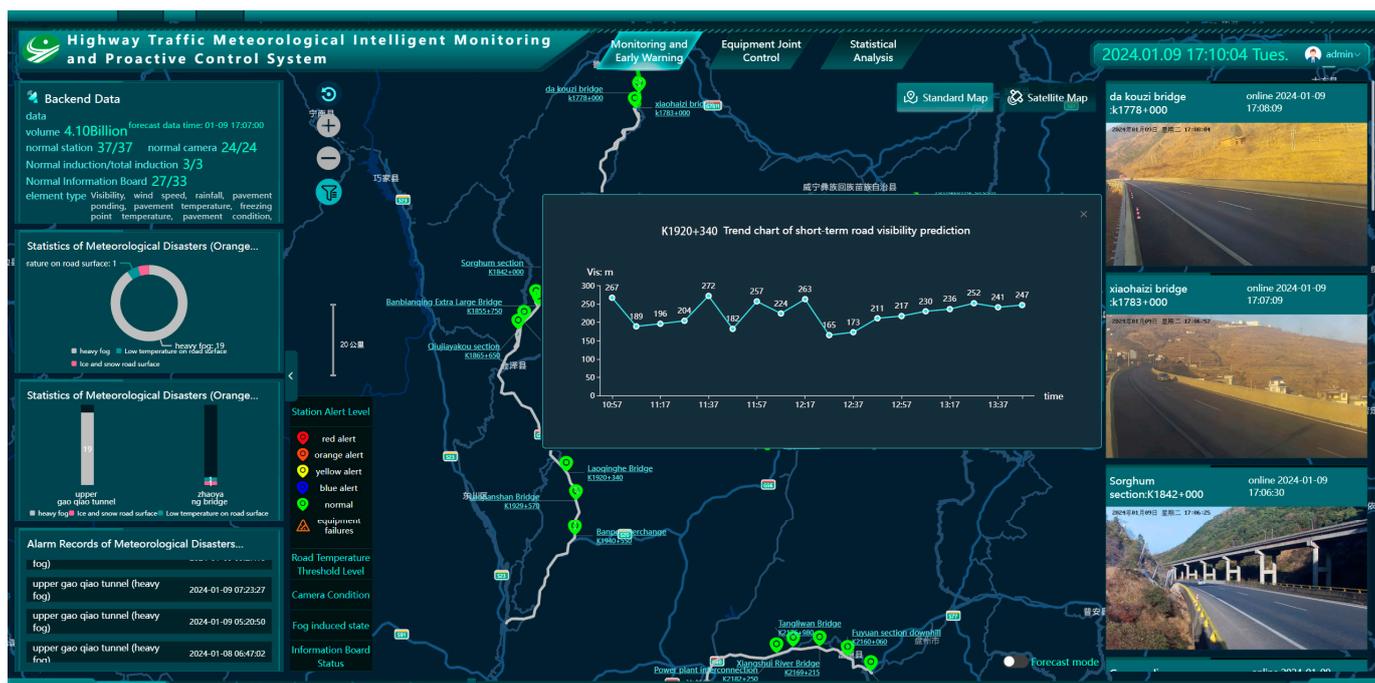


Figure 11. Demonstration of the visibility prediction performance of the model of highway traffic meteorological intelligent monitoring and early warning and active prevention and control system.

6. Conclusions

This study presents a novel highway visibility prediction model, ATCNet, which combines Transformer, Capsule Networks (CapsNet), and self-attention mechanisms. Our experimental results demonstrate that ATCNet excels in handling complex and varied meteorological data, significantly surpassing current advanced time series prediction methods. Through detailed ablation experiments, we have confirmed the contributions of Transformer, CapsNet, and self-attention mechanisms individually and in combination to enhance model performance. Particularly under extreme meteorological conditions, ATCNet shows excellent robustness and high accuracy.

We also explored the potential of ATCNet in practical application scenarios. The model has been successfully integrated into the Highway Traffic Meteorological Intelligent Monitoring and Proactive Control System and has been practically deployed on multiple highways in Yunnan Province. Preliminary application results indicate that ATCNet can effectively assist traffic management authorities in real-time prediction and decision-making, thereby improving traffic safety and efficiency. However, we also recognize the challenges in actual deployment, including the need for real-time data processing, data quality assurance, and continual model optimization.

Future work will focus on further enhancing the model's computational efficiency, optimizing its performance in a broader range of application scenarios, and expanding the model to handle more types of meteorological data. The success of ATCNet demonstrates the potential of deep learning in the field of traffic meteorology, paving new pathways for future research and applications.

Author Contributions: Conceptualization, W.L. and X.Y.; methodology, W.L. and X.Y.; software, X.Y.; validation, W.L., X.Y. and G.Y.; formal analysis, W.L. and X.Y.; investigation, W.L.; resources, W.L.; data curation, W.L.; writing—original draft preparation, W.L.; writing—review and editing, G.Y.; visualization, X.Y.; supervision, D.X.; project administration, D.X.; funding acquisition, W.L. and D.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Yunnan Key Laboratory of Digital Communications (Grant No. 202205AG070008); the National Natural Science Foundation of China (Grant No. 62162068); the

Yunnan Ten Thousand Talents Program and Yunling Scholars Special Project (Grant No. YNWR-YLXZ-2018-022); the Science and Technology Innovation Project of Yunnan Communications Investment & Construction Group Co., Ltd.—Research and Application of Meteorological Disaster Monitoring, Early Warning, and Joint Control Technology for Highway Traffic Prevention and Control Equipment (Grant No. YCIC-YF-2021-08); and the Science and Technology Innovation Project of Broadvision Engineering Consultants—Research and Application of Video Analysis Technology for Mountainous Highway Traffic Meteorological Disasters Based on Artificial Intelligence (Grant No. ZL-2021-04).

Data Availability Statement: The data presented in this study are openly available in “Mountainous Highway Weather Dataset”, at <https://www.kaggle.com/datasets/liphynix2003/mountainous-highway-weather-dataset> (accessed on 16 February 2024). The meteorological data used for the analysis in this paper were sourced from specialized meteorological equipment installed on highways in the Yunnan Province, China. We believe this is the first minute level mountainous highway meteorological public dataset. The meteorological elements in the dataset are as follows: Visibility (m), Temperature (°C), Humidity (%), Precipitation (mm), Wind speed (m/s), Pavement temperature (°C). The datetime is 2023 December’s next-to-last week. Due to privacy or ethical restrictions, some data or details are not publicly available but can be obtained from the corresponding author upon reasonable request.

Conflicts of Interest: Author Xuekun Yang was employed by the company Broadvision Engineering Consultants Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Faturechi, R.; Miller-Hooks, E. Measuring the performance of transportation infrastructure systems in disasters: A comprehensive review. *J. Infrastruct. Syst.* **2015**, *21*, 04014025. [CrossRef]
2. Zou, Y.; Zhang, Y.; Cheng, K. Exploring the impact of climate and extreme weather on fatal traffic accidents. *Sustainability* **2021**, *13*, 390. [CrossRef]
3. Petrova, E. Natural hazard impacts on transport infrastructure in Russia. *Nat. Hazards Earth Syst. Sci.* **2020**, *20*, 1969–1983. [CrossRef]
4. Lu, H.; Chen, M.; Kuang, W. The impacts of abnormal weather and natural disasters on transport and strategies for enhancing ability for disaster prevention and mitigation. *Transp. Policy* **2020**, *98*, 2–9. [CrossRef]
5. Zhu, S.; Yang, H.; Liu, D.; Wang, H.; Zhou, L.; Zhu, C.; Zhi, X. Observations and Forecasts of Urban Transportation Meteorology in China: A Review. *Atmosphere* **2022**, *13*, 1823. [CrossRef]
6. Federal Highway Administration. *Traffic Safety Facts 2022*; U.S. Department of Transportation: Washington, DC, USA, 2022.
7. Fernández-González, S.; Bolgiani, P.; Fernández-Villares, J.; González, P.; García-Gil, A.; Suárez, J.C.; Merino, A. Forecasting of poor visibility episodes in the vicinity of Tenerife Norte Airport. *Atmos. Res.* **2019**, *223*, 49–59. [CrossRef]
8. Pahlavan, R.; Moradi, M.; Tajbakhsh, S.; Azadi, M.; Rahnama, M. Numerical prediction of several radiation and CBL fog events over Iran using the WRF model for late December 2015. *J. Earth Space Phys.* **2020**, *46*, 561–582.
9. He, J.; Ren, X.; Wang, H.; Shi, Z.; Zhang, F.; Hu, L.; Jin, X. Analysis of the microphysical structure and evolution characteristics of a typical sea fog weather event in the eastern sea of china. *Remote Sens.* **2022**, *14*, 5604. [CrossRef]
10. Kim, M.; Lee, K.; Lee, Y. *Fog Prediction with Visibility Data Assimilation in South Korea*; American Geophysical Union: Washington, DC, USA, 2018; Volume 2018, p. A31J–3003.
11. Qian, W.; Leung, J.C.H.; Chen, Y.; Huang, S. Applying anomaly-based weather analysis to the prediction of low visibility associated with the coastal fog at Ningbo-Zhoushan Port in East China. *Adv. Atmos. Sci.* **2019**, *36*, 1060–1077. [CrossRef]
12. Min, R.; Wu, M.; Xu, M.; Zu, X. Attention based Long Short-Term Memory Network for Coastal Visibility Forecast. In Proceedings of the 2022 IEEE 8th International Conference on Cloud Computing and Intelligent Systems (CCIS), Chengdu, China, 26–28 November 2022; pp. 420–425.
13. Cornejo-Bueno, L.; Casanova-Mateo, C.; Sanz-Justo, J.; Cerro-Prada, E.; Salcedo-Sanz, S. Efficient prediction of low-visibility events at airports using machine-learning regression. *Bound.-Layer Meteorol.* **2017**, *165*, 349–370. [CrossRef]
14. Peláez-Rodríguez, C.; Pérez-Aracil, J.; de Lopez-Diz, A.; Casanova-Mateo, C.; Fister, D.; Jiménez-Fernández, S.; Salcedo-Sanz, S. Deep learning ensembles for accurate fog-related low-visibility events forecasting. *Neurocomputing* **2023**, *549*, 126435. [CrossRef]
15. Zang, Z.; Bao, X.; Li, Y.; Qu, Y.; Niu, D.; Liu, N.; Chen, X. A modified rnn-based deep learning method for prediction of atmospheric visibility. *Remote Sens.* **2023**, *15*, 553. [CrossRef]
16. Han, X.; Zhou, T.; He, Y.; Chen, Y.; Chen, R.; Zhou, W. LSTM-Based Visibility Detection for Airport Images in Time Series. In Proceedings of the 2021 40th Chinese Control Conference (CCC), Shanghai, China, 26–28 July 2021; pp. 7241–7246.
17. Peláez-Rodríguez, C.; Pérez-Aracil, J.; Casanova-Mateo, C.; Salcedo-Sanz, S. Efficient prediction of fog-related low-visibility events with Machine Learning and evolutionary algorithms. *Atmos. Res.* **2023**, *295*, 106991. [CrossRef]
18. Gavahi, K.; Foroumandi, E.; Moradkhani, H. A deep learning-based framework for multi-source precipitation fusion. *Remote Sens. Environ.* **2023**, *295*, 113723. [CrossRef]

19. Bai, C.; Zhao, D.; Zhang, M.; Zhang, J. Multimodal information fusion for weather systems and clouds identification from satellite images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 7333–7345. [[CrossRef](#)]
20. Kim, M.; Lee, K.; Lee, Y.H. Visibility data assimilation and prediction using an observation network in South Korea. *Pure Appl. Geophys.* **2020**, *177*, 1125–1141. [[CrossRef](#)]
21. Qin, H.; Gu, H.; Zhao, Y. Embedded weather forecast system based on multisource information fusion and perception. In Proceedings of the 2021 International Conference on Electronic Information Engineering and Computer Science (EIECS), Changchun, China, 23–26 September 2021; pp. 52–58.
22. Guijo-Rubio, D.; Casanova-Mateo, C.; Sanz-Justo, J.; Gutierrez, P.A.; Cornejo-Bueno, S.; Hervás, C.; Salcedo-Sanz, S. Ordinal regression algorithms for the analysis of convective situations over Madrid-Barajas airport. *Atmos. Res.* **2020**, *236*, 104798. [[CrossRef](#)]
23. Zhang, Y.; Wang, Y.; Zhu, Y.; Yang, L.; Ge, L.; Luo, C. Visibility prediction based on machine learning algorithms. *Atmosphere* **2022**, *13*, 1125. [[CrossRef](#)]
24. *DB53/T 1091-2022*; Technical Specifications for Construction of Highway Meteorological Station Networks. Yunnan Provincial Market Supervision Administration: Kunming, China, 2022.
25. Li, W. “Mountainous Highway Weather Dataset” National Engineering Laboratory for Surface Transportation Weather Impacts Prevention; Broadvision Engineering Consultants Co., Ltd.: Kunming, China, 2023. Available online: <https://www.kaggle.com/datasets/liphynix2003/mountainous-highway-weather-dataset> (accessed on 16 February 2024).
26. Deng, J.; Wang, T.; Jiang, Z.; Xie, M.; Zhang, R.; Huang, X.; Zhu, J. Characterization of visibility and its affecting factors over Nanjing, China. *Atmos. Res.* **2011**, *101*, 681–691. [[CrossRef](#)]
27. Cheng, M.T.; Tsai, Y.I. Characterization of visibility and atmospheric aerosols in urban, suburban, and remote areas. *Sci. Total Environ.* **2000**, *263*, 101–114. [[CrossRef](#)] [[PubMed](#)]
28. Robinat, J. “Weather Forecasting at Vigo Airport Using AI”, Raw Data from THREDDs Meteorological Raw Iowa State University Database. 2020. Available online: <https://www.kaggle.com/datasets/jorgerobinat/weather-forecasting-at-vigo-airport-using-ai/data> (accessed on 10 January 2024).
29. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
30. Kontopoulou, V.I.; Panagopoulos, A.D.; Kakkos, I.; Matsopoulos, G.K. A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. *Future Internet* **2023**, *15*, 255. [[CrossRef](#)]
31. Asselman, A.; Khaldi, M.; Aammou, S. Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. *Interact. Learn. Environ.* **2023**, *31*, 3360–3379. [[CrossRef](#)]
32. Huang, R.; Wei, C.; Wang, B.; Yang, J.; Xu, X.; Wu, S.; Huang, S. Well performance prediction based on Long Short-Term Memory (LSTM) neural network. *J. Pet. Sci. Eng.* **2022**, *208*, 109686. [[CrossRef](#)]
33. Wang, J.; Li, X.; Li, J.; Sun, Q.; Wang, H. NGCU: A new RNN model for time-series data prediction. *Big Data Res.* **2022**, *27*, 100296. [[CrossRef](#)]
34. Zha, W.; Liu, Y.; Wan, Y.; Luo, R.; Li, D.; Yang, S.; Xu, Y. Forecasting monthly gas field production based on the CNN-LSTM model. *Energy* **2022**, *260*, 124889. [[CrossRef](#)]
35. Munir, H.S.; Ren, S.; Mustafa, M.; Siddique, C.N.; Qayyum, S. Attention based GRU-LSTM for software defect prediction. *PLoS ONE* **2021**, *16*, e0247444. [[CrossRef](#)] [[PubMed](#)]
36. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 11106–11115. [[CrossRef](#)]
37. National Engineering Laboratory for Surface Transportation Weather Impacts Prevention & Broadvision Engineering Consultants Co., Ltd. Highway Traffic Meteorological Intelligent Monitoring and Proactive Control System. 2019. Available online: <https://220.163.107.220:11020/jtqx> (accessed on 30 December 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.