Precipitation Nowcasting with Graph Neural Network and Gated Recurrent Units

Soobin Park, Yoonhyuk Choi, Chong-Kwon Kim*
Department of Energy Engineering
Korea Institute of Energy Technology(KENTECH)
Naju-si, Jeollanam-do 58217, Republic of Korea
soobin, younhyuk.choi, ckim@kentech.ac.kr

***Abstract*—Precipitation is one of the most unpredictable events in meteorology. While long-term forecasting remains in the domain of conventional NWP (Numerical Weather Prediction), recent advances in deep learning-based forecasting techniques provide competitive performances in nowcasting, forecasting over zero to two hours into the future. This paper introduces a novel nowcasting method that combines Graph Convolutional Network and Gated Recurrent Unit techniques. The proposed method utilizes optimized edge weights that capture the non-linear relations between the end nodes via an MLP. The proposed model first applies a GCN (Graph Convolutional Network) to uncover latent features of weather stations. Then node feature vector is input to a Gated Recurrent Unit that captures temporal relations and predicts precipitation in the future. We performed extensive experiments to evaluate the performance of the proposed method using a real-world dataset obtained from KMA (Korea Meteorological Administration). The experiments show that the proposed method predicts precipitation nowcasting comparable to the NWP system.**

***Index Terms*—Graph Neural Network, Deep Neural Network, Graph Convolutional Network, Gated Recurrent Unit, Nowcasting.**

I. INTRODUCTION

GLobal warming causes climate change, which eventually poses frequent threats to human life through drought, flooding, and extreme temperature fluctuations[1]. Especially, localized heavy rainfall can be an important factor in emergency control and land/air/water transportation. Precipitation forecasting is driven by the comprehensive influence of internal dynamics, heat, and moisture, along with the interaction of external environmental conditions and large-scale weather systems[2]. It is characterized by obvious nonlinearity and is difficult to forecast, necessitating the development of a new precipitation forecasting methodology.

Machine learning-based weather forecasting has attracted tremendous research attention during the last several years[3][4]. Many prior methods utilize conventional machine learning techniques such as decision trees, random forests, and SVM (Support Vector Machine) [5][6]. Even though these methods can make predictions without radar-based weather images, their prediction accuracy usually is not better than that of NWP. More recently, several techniques that adopt advanced machine learning methods such as ANN, LSTM, and autoencoder have been proposed, and they show significant performance enhancements.

We propose a novel precipitation nowcasting technique that utilizes Graph Convolutional Networks(GCN)[7][8] and Gated Recurrent Units(GRU)[9][10]. A GCN, which aggregates features of homogeneous neighbors, has the capability to extract latent features from meteorological attributes by interacting with neighbors. Contrary to the conventional GCN where all edge weights are set to be equal, we optimize the edge weights by applying MLP (Muiti-Layer Perceptron) which uses various meteorological attributes of two nodes as input. Optimized edge weights control the degree of interaction in graph aggregation phases. Finally, a GRU extracts temporal features from time-series input and makes predictions of several hours in the future.

We carried out extensive experiments with a meteorological dataset obtained from the real world to evaluate the performance of the proposed method. The dataset used in this study is provided by the KMA (Korea Meteorological Administration). The dataset contains a plethora of meteorological information observed every hour from March 1 to May 31 in the year 2020 at about 100 weather stations spread in South Korea. Our experimental results show that the proposed method accurately predicts precipitation in one or two hours in the future; in terms of RMSE, the one hour and two hour prediction errors are 0.3350 and 0.4215, respectively.

II. RELATED WORK

This section describes ML techniques and prior weather forecasting proposals based on ML techniques.

A. Modeling Description

While conventional ML techniques such as decision trees and SVM, DL based precipitation forecasting models primarily use recurrent neural networks (RNNs). However, RNN suffers from two problems; one is oscillation or instability of weights and the other is gradient Vanishing and Exploding[11], [12] problem. Long Short-Term Memory(LSTM) greatly improves RNN performances by overcoming the error backpropagation issues. In particular, it compensates for the long-term dependence of RNNs by processing data with the concept of gates[13]. GRU[10] is another technique that efficiently preserves the important dependencies in past time series data. GRU is simpler than LSTM and reduces the number of learnable parameters.

Graph convolutional networks (GCNs) have been recently proposed and successfully applied in irregular data representation and analysis[14]. CNNs (Convolutional Neural Networks), effective in computer vision, fail to properly address problems with non-Euclidean data. To overcome this challenge, Graph Convolutional Networks (GCNs) represent non-Euclidean data as graphs, and borrowing concepts from CNNs, GCNs consider neigh nodes as adjacent pixels[15][16][17].

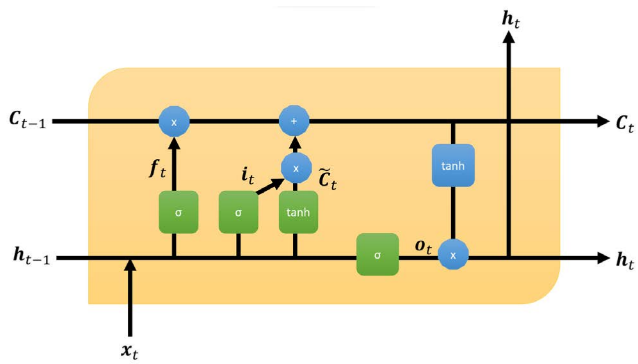

Fig. 1: LSTM Schematic

B. ML Based Forecasting Techniques

Accurate weather forecasting affects huge impacts on society and economy, and meteorologists have developed sophisticated mathematical models called NWP (Numerical Weather Prediction) from 20th century. NWP builds fluid and thermodynamic models based on past and current atmospheric observations and solves large and complex mathematical equations. The performance of NWP models have improved along with the advancements in computation. Advanced weather forecasting agencies are equipped with high performance computers that can solve large and detailed NWP models within forecasting periods. As a result, only large agencies such as national institutions can make accurate weather forecasting.

To get over the limitation of NWP, many researchers have attempted to devise new weather forecasting systems that can eliminate massive computational requirements during the last decade. Many notable proposals utilize various ML techniques that exploit the inference capability of ML. Early approaches[5][6] adopted conventional ML techniques including decision tree, random forest, regression and SVM. However, these methods fail to provide adequate performances.

Successes of DL (Deep Learning) techniques in computer vision and NLP (Natural Language Processing) fields attracted attentions from the ML based weather forecasting research community. In energy system, [18] utilized LSTM model to identify the correlation of output power from input characteristics[19][20][21]. A data decomposition method, applied in the pre-processing stage in order to enhance the forecasting accuracy of nonlinear and nonstationary time series, has been successfully adopted in several proposals[22][23][24].

More sophisticated ML based models accommodate recent advances in ML such as GNN and generative methods. A prior model[25] pioneers the adoption of GCN for weather forecasting. Even though it utilizes GCN and RNN techniques, it simply applies a plain GCN with equal edge weights. For day-ahead short-term forecasting, both historical data and future meteorological data can be used as input to further improve the prediction accuracy[26]. Noting that neighbor nodes have different effects depending on meteorological attributes, we propose to assign different edge weights.

III. METHODOLOGY

First, we explain a method to construct a graph from meteorological information. Then, we detail the proposed method that incorporates GCN with tunable edge weights and GRU.

A. Graph Construction

Weathers of a certain geographical region is affected by atmosphere of surrounding regions. Therefore, it is natural to allocate a node corresponding to a weather station. A weather station collects meteorological data which become the attributes of the corresponding node. The node attributes include average temperature, daily precipitation, average wind speed, prevailing wind direction, average dew point temperature, average relative humidity, average vapor pressure, average local air pressure, average sea surface pressure, altitude, latitude, longitude, and 15 other elements. After nodes are created, we then construct edges. Because the meteorological states of near regions have stronger effect than those of further regions, we construct edges based on the geographical distance between two nodes. In this study, we use the 2-Dimensional Euclidean distance ignoring the effect of altitude.

TABLE I: Meteorological attributes

Variable Name	Unit
Average Temperature	(°C)
Daily precipitation	(mm)
Average wind speed	(m/s)
Prevailing wind direction	(16 directions)
Average dew point temperature	(°C)
Average relative humidity	(%)
Average vapor pressure	(hPa)
Average local pressure	(hPa)
Average sea level pressure	(hPa)

The Graph Convolutional Network(GCN) has good performance in complex nonlinear data structures. The constructed graph represents the relationship of node and edge in non-euclidean space. For example, V and E denote the vertex and edge sets in an undirected graph $G = (V, E)$[27]. The adjacent matrix A can show the relationship in the graph, and it is computed by Equation(1):

$$A_{i,j} = exp\Big(- \frac{(x_i - x_j)^2}{\sigma^2} \Big) \tag{1}$$

In addition, the feature selection is used in this graph that manages feature dimensions[7]. The reduced dimension improves the pattern recognition in the graph. These selections retain the most influential information in the original feature set and reduce its redundancy to improve the performance and accuracy of the model[8].

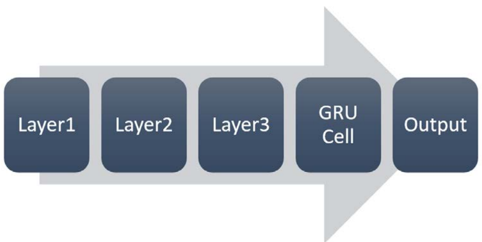

Fig. 2: Layer

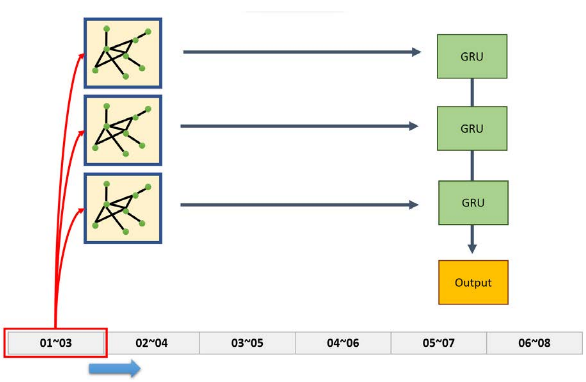

Fig. 3: Model Schematic

The model consists of five main layers; Three GCN layers, a GRU, and an output layer. The meteorological attributes are given as an input to each node and the three GCN layers transform the attributes to model the interactions between the nodes. The parameters between these layers are tuned through training. Each layer adjusts parameters to more accurately model the relationships between nodes[28]. This multi-layered structure allows graph attachment models to be good for learning important patterns and relationships in graph data. Therefore, model achievement is better results than other neural networks.

B. Modeling

This model consists of a Graph Convolutional Network(GCN) and a Gated Recurrent Unit(GRU).

1) Graph Convolutional Network(GCN): GCN is specialized to express graph topological structure with interdependencies features. For this reason, each node is set by region embeddings. This directly mapping node labeling represents a set of interdependent datasets in GCN[29]. The previous machine-learning methods ignored topological information in the graph because there was no way to efficient graph modeling. These convolutional neural networks only process data in Euclidean space. However, the realistic data exist in non-Euclidean structured data. To resolve this problem, the emergence of graph convolution fills the gap of neural networks to obtain topological graph-type features[30].

2) Gated Recurrent Unit(GRU): GRU is a variant of recurrent neural networks[9]. In recurrent neural networks, the previous input dataset can affect other datasets. This characteristic of recurrent neural networks, the exploding and vanishing gradient problems frequently occur in RNN modeling. The GRU modeling solves the previous layer gradient information learning problems to jointly participate[31]. In GRU, all layers participate in the next output calculation. This improvement modeling is specialized to manage time series datasets[10] because it predicts time series data and makes deep spatio-temporal features[32]. In addition, GRU determines the dataset remembered or forgotten by the gates. Moreover, the back-propagation updates the weight of gates[31].

IV. EXPERIMENTS

We performed extensive experiments with the dataset obtained from the real world to assess the performance of the proposed method. Beginning with the description of the dataset and testing environment, this section explains the experimental results.

A. Dataset

The data was collected from the Korea Meteorological Administration Weather Data Service Open Meteorological Data Portal. The Open Meteorological Data Portal is a site under the Korea Meteorological Administration. We collected 12 parameters from this site. A total of three years of data were collected from the year 2020 to the year 2022.

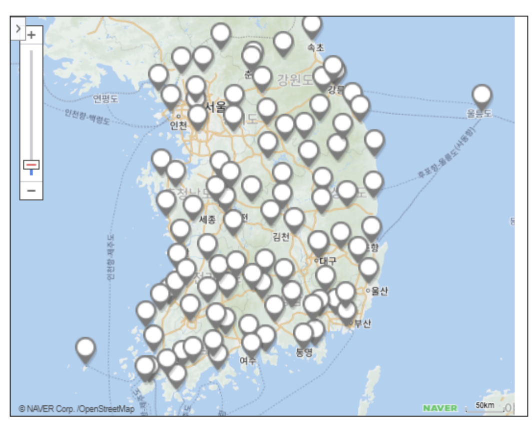

Fig. 4: Collected Data Points

TABLE II: Train, validation, and test data

Train	Validation	Test
May 1, 2000 - September 30, 2012	May 1, 2013 - September 30, 2016	May 1, 2020 - September 30, 2017

The collected points (nodes) are all in Region 103, which includes Ulleung Island and Jeju Island (Fig. 4). Since terrain is also important for weather forecasting, we also collected geographic data; The latitude, longitude, and altitude of every point were collected and added to the attributes. We rely on Google Earth to obtain the latitude, longitude, and altitude information, and they may contain small errors due to the location of the Korea Meteorological Administration or weather stations.

B. Training

We experiment with four different training conditions depending on the adoption of Multi-Layer Perceptron (MLP) for edge weight optimization and backpropagation methods. MLPs receive meteorological information of two adjacent nodes as input and determine the optimal weight for the edge between the nodes. It consists of an input layer, one or more hidden layers, and an output layer. Backpropagation finds the difference between the actual target value and the output calculated by the model and then propagates the error backward.

Note that this paper focuses on nowcasting, forecasting of zero to two hours after. For nowcasting, we use hourly meteorological information for both training and testing. We use 48 hour long data for training and make rainfall predictions one and two hours after. Table II describes the data used for the train, validation, and test of this experiment.

C. Performance Results

We use two performance metrics; RMSE (Rooted Mean Square Error) and MAE (Mean Absolute Error) as shown in (2) and (3). Because the most accurate results were obtained when MLPs are adopted for edge weight optimization, we show only the experimental results with MLP. Table III shows one hour and two hour prediction accuracy in terms of the two performance metrics. It indicates that MAE is 0.40 and 0.43 for one and two hour predictions, respectively. Generally, it is more difficult to make long-term forecasting and two hour prediction is about 7.5% worse than one hour forecasting.

$$MAE = \frac{1}{\tau}\sum_{t=1}^{\tau}\frac{1}{n}\sum_{i=1}^{n}|y_m - y_p| \quad (2)$$

$$RMSE = \frac{1}{\tau}\sum_{t=1}^{\tau}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_m - y_p)^2} \quad (3)$$

The image in Fig. 5 visualizes the result on June 6, 2020. Comparing the actual and experimental values shown in Fig. 5, we can observe that there are notable differences in prediction accuracy depending on regions. Generally, no rainfall regions experience small errors and heavy rainfall regions suffer from large errors. This phenomenon was observed in several

TABLE III: Experimental results

Type	1 hour prediction	2 hour prediction
Train loss	0.8281	0.8750
Test loss	0.7344	0.8381
MAE	0.3976	0.4312
RMSE	0.3350	0.4215

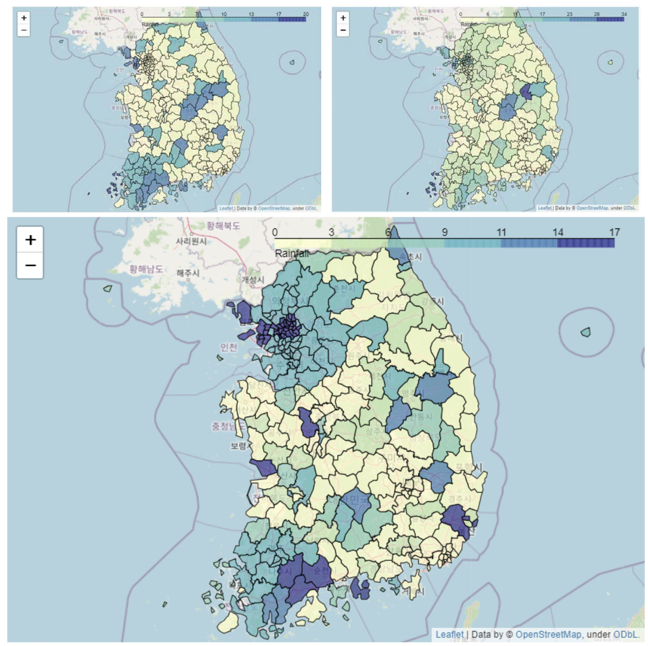

Fig. 5: The picture of (a) prediction, (b) real value, (c) difference between them

prior studies and it seems the problem remains as a further investigation in the future. In addition, we can observe that mountainous areas suffer from large errors; The mountainous areas receive more rain than predicted.

V. CONCLUSIONS

In this paper, GCN and GRU were used to forecast precipitation. The accuracy is overall good, and the overall error was about 20 millimeters. The slight error is due to localized heavy rainfall. The localized heavy rain is made by the updraft of water vapor and mountain ranges help it go up. This is why the reason for the mountainous terrain of Korea, altitude is a significant variable. This error is expected to be resolved by collecting additional altitude data. In this paper, data is

TABLE IV: Experimental results

Type	1 hour prediction	2 hour prediction
Train loss	0.8281	0.8750
Test loss	0.7344	0.8381
MAE	0.3976	0.4312
RMSE	0.3350	0.4215

collected and predicted at a daily interval, but if the data is increased to an hourly interval, good results can be obtained for localized heavy rain forecasting.

ACKNOWLEDGMENT

This work was supported by the KENTECH Research Grant(202200019A), IITP(No.2021-0-02068, 2023-RS-2022-00156287) grant funded by the Korean government(MSIT).

REFERENCES

[1] Abhishek Kumar, Shilpi Nagar, and Shalini Anand. Climate change and existential threats. In *Global climate change*, pages 1–31. Elsevier, 2021.

[2] Lin Kaiping, Chen Binglian, Dong Yan, and Huang Ying. A genetic neural network ensemble prediction model based on locally linear embedding for typhoon intensity. In *2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA)*, pages 137–142, 2013.

[3] Yu Feng, Weiping Hao, Haoru Li, Ningbo Cui, Daozhi Gong, and Lili Gao. Machine learning models to quantify and map daily global solar radiation and photovoltaic power. *Renewable and Sustainable Energy Reviews*, 118:109393, 2020.

[4] Yuchen Zhang, Mingsheng Long, Kaiyuan Chen, Lanxiang Xing, Ronghua Jin, Michael I Jordan, and Jianmin Wang. Skilful nowcasting of extreme precipitation with nowcastnet. *Nature*, 619(7970):526–532, 2023.

[5] R Meenal and A Immanuel Selvakumar. Assessment of svm, empirical and ann based solar radiation prediction models with most influencing input parameters. *Renewable Energy*, 121:324–343, 2018.

[6] Ümit Ağbulut, Ali Etem Gürel, and Yunus Biçen. Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*, 135:110114, 2021.

[7] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.

[8] Honglai Xu, Zhao Zhen, and Fei Wang. Nwp feature selection and gcn-based ultra-short-term wind farm cluster power forecasting method. In *2022 IEEE Industry Applications Society Annual Meeting (IAS)*, pages 1–22, 2022.

[9] Jun Wang, Junxing Cao, Shan Yuan, and Ming Cheng. Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the ceemdan-se-and the pso-als-optimized gru network. *Energy*, 233:121082, 2021.

[10] Cheng Fan, Jiayuan Wang, Wenjie Gang, and Shenghan Li. Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. *Applied energy*, 236:700–710, 2019.

[11] Malti Bansal, Apoorva Goyal, and Apoorva Choudhary. A comparative analysis of k-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics Journal*, 3:100071, 2022.

[12] Malti Bansal, Naman Oberoi, and Mohd Sameer. Iot in online banking. *Journal of Ubiquitous Computing and Communication Technologies (UCCT)*, 2(04):219–222, 2020.

[13] Sasan Karamizadeh, Shahidan M Abdullah, Mehran Halimi, Jafar Shayan, and Mohammad javad Rajabi. Advantage and drawback of support vector machine functionality. In *2014 international conference on computer, communications, and control technology (I4CT)*, pages 63–65. IEEE, 2014.

[14] Danfeng Hong, Lianru Gao, Jing Yao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. Graph convolutional networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7):5966–5978, 2021.

[15] Qimai Li, Zhichao Han, and Xiao-Ming Wu. Deeper insights into graph convolutional networks for semi-supervised learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[16] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24, 2020.

[17] Jie Zhou, Ganqu Cui, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. arxiv 2018. *arXiv preprint arXiv:1812.08434*, 1812.

[18] Didar Tukymbekov, Ahmet Saymbetov, Madiyar Nurgaliyev, Nurzhigit Kuttybay, Gulbakhar Dosymbetova, and Yeldos Svanbayev. Intelligent autonomous street lighting system based on weather forecast using lstm. *Energy*, 231:120902, 2021.

[19] Marcello Anderson FB Lima, Paulo CM Carvalho, Luis M Fernández-Ramírez, and Arthur PS Braga. Improving solar forecasting using deep learning and portfolio theory integration. *Energy*, 195:117016, 2020.

[20] Chaman Lal Dewangan, SN Singh, and S Chakrabarti. Combining forecasts of day-ahead solar power. *Energy*, 202:117743, 2020.

[21] Mingming Gao, Jianjing Li, Feng Hong, and Dongteng Long. Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using lstm. *Energy*, 187:115838, 2019.

[22] Hui Liu, Shi Yin, Chao Chen, and Zhu Duan. Data multi-scale decomposition strategies for air pollution forecasting: A comprehensive review. *Journal of Cleaner Production*, 277:124023, 2020.

[23] Lishu Wang, Yanhui Liu, Tianshu Li, Xinze Xie, and Chengming Chang. Short-term pv power prediction based on optimized vmd and lstm. *IEEE Access*, 8:165849–165862, 2020.

[24] Yongkang Zeng, Jingjing Chen, Ning Jin, Xiaoping Jin, and Yang Du. Air quality forecasting with hybrid lstm and extended stationary wavelet transform. *Building and Environment*, 213:108822, 2022.

[25] Mahdi Khodayar and Jianhui Wang. Spatio-temporal graph deep neural network for short-term wind speed forecasting. *IEEE Transactions on Sustainable Energy*, 10(2):670–681, 2018.

[26] Lilin Cheng, Haixiang Zang, Tao Ding, Zhinong Wei, and Guoqiang Sun. Multi-meteorological-factor-based graph modeling for photo-voltaic power forecasting. *IEEE Transactions on Sustainable Energy*, 12(3):1593–1603, 2021.

[27] Danfeng Hong, Lianru Gao, Jing Yao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. Graph convolutional networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7):5966–5978, 2021.

[28] Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 922–929, 2019.

[29] Zhao-Min Chen, Xiu-Shen Wei, Peng Wang, and Yanwen Guo. Multi-label image recognition with graph convolutional networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5172–5181, 2019.

[30] Yingya Guo, Yufei Peng, Run Hao, and Xiang Tang. Capturing spatial-temporal correlations with attention based graph convolutional network for network traffic prediction. *Journal of Network and Computer Applications*, page 103746, 2023.

[31] Gülsüm Yiğit and Mehmet Fatih Amasyali. Simple but effective gru variants. In *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, pages 1–6, 2021.

[32] Haixu Wang and Shiyong Shao. Prediction of pm2.5 in hefei based on a hybrid cnn-gru model. In *2022 5th International Conference on Data Science and Information Technology (DSIT)*, pages 1–6, 2022.