UFSP-Net: a neural network with spatio-temporal information fusion for urban fire situation prediction

Guangyin Jin¹, Cunchao Zhu¹, Xiaoxuan Chen¹, Hengyu Sha¹, Xingchen Hu¹, Jincai Huang¹
¹College of System Engineering, National University of Defense Technology, Changsha, China
Guangyin Jin: jinguangyin18@nudt.edu.cn

Abstract. Capturing the dynamics of urban fire situation is a basic but challenging task, which takes an indispensable role in the field of urban security and fire emergency decision. Traditional methods approach the urban fire prediction via stochastic process based on physics or statistics, which may be interpretable but less practical in real applications. Recently, some data-driven models, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Graph Convolutional Neural Network (GCN), seem to be fruitful in capturing spatio-temporal dynamics with massive high-dimension data. In this paper, we process some regional urban fire dataset of recent six years in the fire situation awareness images (FSAIs) and extract pixel-level latent representations with CNNs while GCNs are applied to process some spatial graph structure auxiliary information to obtain graph-level latent representations. And then Urban Fire Situation Prediction Neural Network (UFSP-Net) is formulated as a novel urban fire prediction model, integrating these two different kind of spatial latent representations and RNN structure. In comparison to other traditional algorithms, such as Conv-RNN, UFSP-Net demonstrates its superior prediction performance for multi type urban fire in spatio-temporal scale.

Key Words: Fire Situation Prediction, Graph Convolutional Neural Network, Convolutional Neural Network, Spatio-Temporal Analysis

1. Introduction
With the advancement of urbanization, the prevention and control of urban fires has become a tricky task in the field of urban security. According to the official fire statistical report of America in 2018 [1], a total of 237,000 fires were reported nationwide, 1,407 people were killed and 798 people were injured. The direct property loss has been calculated to exceed 500 million dollars. Among them, urban fires still account for the largest proportion and nearly 80% of the deaths are concentrated in urban residential region. Compared with wildfires, urban fires manifest more complex dynamics and lower predictability due to more impact factors. According to the types of urban fire impact factors, it can be divided into two categories: spatial impact factors and temporal impact factors. In the space perspective, a modern city can be divided into many different functional regions, for instance, residential regions, commercial regions, industrial regions, tourist regions, etc. Empirically speaking, the incentives of urban fires vary with spatial
regions. For instance, fire in residential regions are often caused by fires in boilers and appliances while fires in industrial regions are often caused by chemical and electrical failures. In addition to the functional region, the regional fire situation is also inseparable from the fire equipment investment and deployment in the particular region [2]. In the time perspective, fire incidents also have strong temporal correlation and we can argue from two aspects [3]. For one thing, the occurrence of fires varies with the seasonal climate. According to previous investigation researches, the hot and dry seasons are more prone to fire than usual. For another, the occurrence of fires varies with the time of one day. Empirically speaking, working hours in industrial regions during the day may have a higher probability of fire than non-working hours at night due to equipment’s operation or other reasons. In summary, space and time are the two most crucial dimensions of urban fire prediction tasks. However, the most serious challenge of urban fires prediction tasks is modeling spatio–temporal dynamics to achieve real-time prediction [4]. In some previous literatures, the theoretical model of urban fire prevention and control has been developed. In the past decade, as some data-driven approaches [5, 6] were gradually proposed, these models have also been transferred to urban fire prediction tasks. Although some statistical learning models have achieved state-of-art performance, these models still have difficulties to capture complex spatio–temporal dynamics in urban fire prediction. In recent year, development of deep learning is considered promising to capture complex spatial and temporal dynamics [7]. However, the application of deep learning in the field of fire prediction has not been fully developed.

In this paper, we proposed a deep learning approach, Urban Fire Situation Prediction Neural Network (UFSP-Net), which integrates CNN, GCN and RNN model to cope with urban fire prediction. We initially arrange spatio-temporal fire record data and region boundary data into Fire Situation Awareness Images (FSAs) and region graphs respectively. Then, convolutional neural networks (CNN) [8] and graph neural network (GCN) [9] are applied to embed the spatial information from the sequences of FSAs and region graphs into latent space to obtain joint spatial representation. Meanwhile, extraction of temporal information can be tackled by Gated Recurrent Unit Networks (GRU) [10], which can capture relevance of former and future time slots. Subsequently, the correlation of spatial and temporal dynamics has been established. To summarize, we make the following main contributions:

1. We first proposed the real-time deep learning model to capture the spatial temporal dynamics from joint latent representation of Urban Fire Situation Awareness Images and Region Boundary Distance Graph graphs.
2. We conduct experiments on real world dataset, and the proposed approach achieves remarkable performance in comparison to state-of-art baseline methods for urban fire prediction.

In the remainder of the paper, we begin with a review of the related work on the traditional urban fire prediction methods in section 2. Then we review background of CNN, GCN and GRU in section 3. The research data process and methodology adopted in this study is presented in section 4, followed by the experimental results and analysis in section 5. The final section concludes the achievements of this research and proposes future work.

2. Related work

There exists a large body of work on fire prediction and early methods for addressing this task were mainly based on some mathematical modeling and stochastic process simulation technology. Himoto et al investigated the physics-based computational model for urban fire spread [11], which proposed a mathematical physics model that interpreted urban fire as a collection of multiple building fires and simulated the spread of a fire by predicting the behavior of a single building fire under the thermal influence of a neighboring building fire. To better simulate the stochastic process of urban fire spread, the Cellular Automata (CA) model was introduced into the field of fire control and prevention in the work [12]. Based on the traditional cellular automata model, the work [13] presented a novel model integrating the Extreme Learning Machine (ELM) with the CA framework to improve simulation performance. However, these simulation methods are difficult to capture the spatio-temporal
3. Background

In this section, we briefly call the related methods of this research. In particular, the basic principles of Convolutional Neural Network (CNN), Graph Convolutional Neural Network (GCN) and Recurrent Neural Network (RNN).

3.1. Convolutional neural network

Convolutional Neural Network has achieved remarkable breakthroughs in image processing, speech recognition, etc. This model is designed to automatically and adaptively learn spatial hierarchies of features by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. Among them, convolution layer is the most crucial part of CNN and convolution layer operation is formulated as:

$$ Y^l(m, n) = X^k(m, n) * H^kl(m, n) = \sum_{k=0}^{K-1} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} X^k(m+i, n+j)H^kl(i, j) $$ (1)
It is assumed that the convolutional layer has \( L \) output channels and \( K \) input channels, so that KL convolution kernels are needed to realize the conversion of the number of channels. \( X^k \) is the two-dimensional feature map of the \( k \)-th input channel, \( Y^l \) is the two-dimensional feature map of the \( l \)-th output channel, and \( H^{kl} \) represents the \( k \)-th row and the \( l \)-th column of the two-dimensional convolution kernel. Assuming that the convolution kernel size is \( I \times J \) and the feature map size of each output channel is \( M \times N \).

The pooling layer can reduce the dimension of the feature map while maintaining the most important information. It has different types such as max pooling, average pooling, sum pooling and so on, but max pooling is default in this paper. The fully connected layer is a traditional multi-layer perceptron whose main function is to combine the features extracted by the previous convolutional layer and pooling layer into one-dimensional vectors.

3.2. Graph convolutional neural network
GCN is defined on graph structure data \( G = (V, E) \) where \( V \) is the set of graph nodes and \( E \) is edges between two different nodes. In a defined graph, the set of connected edges weights between different nodes can be represented by an adjacency matrix \( A \in \mathbb{R}^{|V| \times |V|} \) and a feature matrix \( F \in \mathbb{R}^{|V| \times |N|} \). The notation \( N \) represents the feature dimensions of each node. GCN is able to obtain local features with different reception fields from translation variant non-Euclidean structures [9]. Let \( L = I - D^{-1/2}AD^{-1/2} \) denotes the graph Laplacian matrix, where \( D \) is the degree matrix, a graph convolution operation is defined as:

\[
\begin{align}
\delta(F, A) &= \sigma(L^{(1)}FW^{(1)}), \quad l = 1 \\
\delta(H^{(l)}, A) &= \sigma(L^{(l)}H^{(l)}W^{(l)}), \quad l \geq 2
\end{align}
\]  

Where \( H^{(l)} \) denotes the hidden layer features in the \( l \)-th layer, \( W^{(l)} \) is the weight of each layer, \( L^{(l)} \) is the graph Laplacian matrix of \( l \)-th layer, \( \sigma \) represents an arbitrary activation function. And the input feature of the first layer equals the original feature matrix \( F \).

3.3. Recurrent neural network
RNN has been widely applied to capture the dynamics of sequence data in recent year. However, some disadvantages of RNN, for instance, gradient explosion, gradient disappears and long-term dependency problem, have hindered its promotion and application [23]. Hence, many variants of RNN were proposed such as long and short term memory network (LSTM) and gated recurrent units network (GRUN). GRUN is easier to train but can achieve performance comparable LSTM, so it is applied in our research to capture temporal dynamics. GRU is the basic component of GRUN and two gate operators are included in GRU as reset gate \( R_t \) and update gate \( U_t \):

\[
\begin{align}
R_t &= \sigma(w_{xr}x_t + w_{hr}h_{t-1}) \\
U_t &= \sigma(w_{xu}x_t + w_{hu}h_{t-1})
\end{align}
\]

Meanwhile, the candidate hidden information \( h_t^c \) and merged hidden information \( h_t \) can be calculated as:

\[
\begin{align}
h_t^c &= \tanh(w_{xh}x_t + w_{hh}(R_t \odot h_{t-1})) \\
h_t &= U_t \odot h_{t-1} + (1 - U_t) \odot h_t^c
\end{align}
\]

4. Data processing and methodology
In this section, we provide details of data processing approach and our proposed Urban Fire Situation Prediction Neural Network (UFSP-Net).
4.1. Data processing

The urban fire dataset in this research originates from a public safety database managed by San Francisco government. And the interest region is defined in San Francisco city, which cover the square region in the map with Latitude Interval [37.71, 37.8] and Longitude Interval [-122.50, 122.38]. Taking into account the integrity and time-effectiveness of the information, fire incidents in the period of last six years until June 2019 are processed, and detailed column attributes mainly include some temporal information (incident date, incident time, incident year), some spatial information (latitude, longitude, location description and region boundary), fire types (Potentiality Life-threatening, Non-Life-threatening and Alarm), battalion, neighborhoods and police district information. **Fig 1** illustrates the general situation of fire incidents data.

![Figure 1. Spots of Fires in Research Dataset. The orange dotted marked regions are some locations fire incidents happened.](image)

With this dataset, we focus on the spatio-temporal distribution of different types of fire, and the frequency of some fire types in some region lattice at some time slot is a direct statistics revealing the spatio-temporal situation of urban fire. Hence, we apply one channel fire situation awareness image to represent spatio-temporal distribution of each fire type.  

**Definition 1. Fire Situation Awareness Image of specific fire type** $X^{(t)}$. For the region lattice indexed (i, j), the time slot indexed t and one specific fire type, we define the element $x_{ij}^{(t)}$ as the occurrence count of the specific fire type in the particular time slot and region. And FSAIs of one specific fire type are composed of basic elements $X^{(t)} = \left[ x_{ij}^{(t)} \right]_{I*J}$, where $I*J$ represents geographical indexes.

Target area are uniformly partitioned into 10*10 lattices in this research and the FSAIs of three different fire types respectively as Potentiality Life-threatening (Fire Type 1), Non-Life-threatening (Fire Type 2) and Alarm (Fire Type 3).

**Definition 2. Region Boundary Distance Graph.** In urban fire prediction task, for two region boundaries near each other, they may share similar fire patterns. Following this idea, we use the distance to construct the region-based graphs. More specifically, reciprocal of the distances are used to mark the weight between two region boundaries so that closer regions will be linked with higher weights. The adjacency matrix of Region Boundary Distance Graph is defined as:

---

1 https://data.sfgov.org/Public-Safety/Fire-Department-Calls-for-Service/nuek-vuh3
In this dataset, region of San Francisco has been divided into 41 region boundaries and we can mark the latitude and longitude of each region boundary by GIS.

4.2. Urban fire situation prediction neural network

To address the challenge of capturing spatio-temporal dynamics from non-linear events, we propose a deep neural network framework, referred to as Urban Fire Situation Prediction Neural Network (UFSP-Net), to understand fire situation and perceiving potential dynamics. The overview of UFSP-Net is revealed in Fig 2.

As illustrated in Fig 2, UFSP-Net is mainly composed of three components, the spatial latent representation extraction model, temporal dynamics capture model and decoder model respectively. The first model consists of CNN, GCN and information fusion feedforward neural network (FNN), and attempts to map the high dimensional FSAIs and Region Boundary Distance Graphs into low dimensional joint latent representation. The second model is actually GRU model to capture temporal dynamics while the third model is decoder CNN model to decode FSAIs from predicted latent representation. The goal of UFSP-Net is to predict the next time slot FSAI by the last time slot FSAI and Region Boundary Distance Graph.

In UFSP-Net model, we applied two convolution layers and one fully connected layer in CNN. The parameters of the first convolution layer are [1, 16, 3, 1, 1], which represent input channel, output channel, kernel size, stride and padding value respectively. The parameters of the second convolution
layer are \([16, 32, 3, 1, 1]\), which represent the same definition as the first convolution layer. After each convolution layer, there will be a batch normalization layer and an activation function LeakyReLU. Following the second convolutional layer is a fully connected layer to map the FSAIs to one-dimensional latent variables with size \(1*10\). The process is formulated as:

\[
 z_{p_k}^{t} = \text{Linear}(\text{LeakyReLU} \left( \text{BN2} \left( \text{Conv2} \left( \text{LeakyReLU} \left( \text{BN1} \left( \text{Conv1}(x_k^t) \right) \right) \right) \right) \right)) \tag{9}
\]

Where \(\text{Conv}\) and \(\text{BN}\) represent convolution operation and batch normalize operation respectively. \(x_k^t\) is input FSAI and \(z_{p_k}^{t}\) is pixel-level latent representation of fire type \(k\).

In GCN model, we applied only one graph convolution layer and the parameters of it are \([1, 10]\), which represent the input dimension of feature matrixes and output dimension of hidden representation. In this paper, feature matrixes are one-dimensional vectors which represent number of fire incidents in some regions for particular time slot. Following the graph convolution layer is a fully connected layer to map the Region Boundary Distance Graphs to one-dimensional latent variables with size \(1*10\). The process is formulated as:

\[
 z_{g_k}^{t} = \text{Linear}(\text{LeakyReLU} \left( \text{GC}(A^t, f_k^t) \right)) \tag{10}
\]

Where \(A^t, f_k^t\) are adjacency matrix and feature matrix respectively. \(\text{GC}\) represents the graph convolution operation and \(z_{g_k}^{t}\) is graph-level latent representation of fire type \(k\).

The pixel-level latent representation and graph-level latent representation are extracted by CNN and GCN respectively while the FNN acts as a spatial information fusion network to obtain a joint latent representation with size \(1*10\). And the process is defined as:

\[
 z_{k}^{t} = \text{FNN}(z_{p_k}^{t} \oplus z_{g_k}^{t}) \tag{11}
\]

The operator \(\oplus\) represents concentration operation and \(z_{k}^{t}\) is the joint latent representation converged by \(z_{p_k}^{t}\) and \(z_{g_k}^{t}\).

Through the process of the spatial latent representation extraction model, we get a series of joint latent representation \([z_{k}^{1}, z_{k}^{2}, ..., z_{k}^{T-1}, z_{k}^{T}]\). For GRUN, the training process of the model is also a process that maximizes the conditional probability as:

\[
 \arg \max_{\theta} \prod_{k=1}^{3} \prod_{t=1}^{T} P_{\theta}(x_{k}^{t+1} | z_{k}^{t}) \tag{12}
\]

The mathematical form of GRUN is defined as:

\[
 z_{k}^{t+1} = \text{GRU}(z_{k}^{t}, h_{k}^{t+1}) \tag{13}
\]

Where \(h_{k}^{t+1}\) represents the hidden representation in the time slot \(t+1\).

For the last part of UFSP-Net, decoder CNN model consists of one fully connected layer and two deconvolution layers. The parameters of the first deconvolution layer are \([32, 16, 3, 1, 1]\), which represent input channel, output channel, kernel size, stride and padding value respectively. The parameters of the second deconvolution layer are \([16, 1, 3, 1, 1]\), which represent the same definition as the first deconvolution layer. The computational process of decode is defined as:

\[
 \hat{x}_k^t = \text{Deconv2} \left( \text{LeakyReLU} \left( \text{BN} \left( \text{Deconv1}(z_k^t) \right) \right) \right) \tag{14}
\]

The loss function of UFSP-Net model is formulated as:

\[
 L_2(x_k^t, \hat{x}_k^t) = ||x_k^t - \hat{x}_k^t||_2^2 \tag{15}
\]
\[ x_k^{(t)}, \hat{x}_k^{(t)} \] are respectively the original FSAIs and predicted FSAIs.

5. Experiments and Analysis
In this section, the training details are described, and some numerical analysis are presented to show the performance of our model.

5.1. Training details and comparison algorithms

**Training Details:** The urban fire dataset employed in this research ranges from Jun. 1st, 2014 to June 12th, 2019, and the time series can be roughly partitioned into 1810 continuous disjoint time slots with one day as unit to formulate FSAIs and Region Boundary Distance Graphs. Taking variations in three types of fire incidents in spatio-temporal patterns, we respectively train three CNNs and GCNs, then three series of joint latent representation at initial time slot \( t \) are obtained by information fusion as \([z_1^k, z_2^k, z_3^k]\). For three types of FSAIs and region-based graphs, 90% of dataset are for model training with the rest for the performance testing. For UFSP-Net model, the Adam optimizer is used in training process and learning rate is defined as 0.001.

**Comparison Algorithms:** In the experiments, four algorithms are included in comparison to Urban Fire Situation Prediction Neural Network (UFSP-Net), including Random Forest, XGBoost, Gated Recurrent Unit Neural Network (GRUN) and Convolutional Gated Recurrent Unit Neural Network (Conv-GRUN). And the prediction process for comparison algorithms is to automatically generate the next time slot FSAIs by the last time slot FSAIs.

5.2. Result analysis

We evaluate the effectiveness of proposed framework UFSP-Net in this section. No matter for CNN, GCN or GRUN training, the quality of prediction of FSAIs is the most concerned. Three metrics are applied to evaluate the prediction performance of each algorithm, respectively as the Root Mean Square Error (RMSE), Mean Square Percentage Error (MSPE) and Jessen-Shannon Divergence (JS), whose definitions are as follows:

\[
\text{RMSE}_k = \left( \frac{1}{T} \sum_{t=1}^{T} \left| x_k^{(t)} - \hat{x}_k^{(t)} \right|^2 \right)^{1/2}
\]

\[
\text{MAPE}_k = \frac{1}{T} \sum_{t=1}^{T} \frac{|x_k^{(t)} - \hat{x}_k^{(t)}|}{|x_k^{(t)}|}
\]

For the normalized spatial frequencies:

\[
p_{l,j,k}^{(t)} = \frac{x_{l,j,k}^{(t)}}{\sum_{i=1}^{M} \sum_{j=1}^{N} x_{i,j,k}^{(t)}}
\]

\[
p_{l,j,k}^{(t)} = \frac{\hat{x}_{l,j,k}^{(t)}}{\sum_{i=1}^{M} \sum_{j=1}^{N} \hat{x}_{i,j,k}^{(t)}}
\]

The JS-Divergence is computed as:

\[
D_{JS}(P, P^\sim) = \frac{1}{2} D_{KL}(P || 1/2(P + P^\sim)) + \frac{1}{2} D_{KL}(P^\sim || 1/2(P + P^\sim))
\]

Where \( x_k^{(t)}, \hat{x}_k^{(t)} \) are respectively the original FSAIs and predicted FSAIs. In this case, MSPE can be viewed as a squared individualized weighted RMSE, measuring the ratio of approximation regardless of scales. And JS divergence measures the spatial distribution similarity between two FSAIs.
In Table 1, the comparison results on Random Forest, XGBoost, GRUN, Conv-GRUN and UFSP-Net are displayed. We find UFSP-Net is significantly superior to others in testing dataset and performances fluctuate for three types of fire. On the datasets of these three fire types, the effect of these four algorithms are successively decreasing and this suggests that deep learning model could be harder to perform well on sparse datasets. We also find that combination of pixel-level latent representation and graph-level latent representation can improve the evaluation metrics on all of fire type dataset.

Table 1. Prediction Results on Five State-of-art Models. We conducted 10-fold cross validation and the results by the best performer in each column are in boldface. We assume that ten results are consistent with a normal distribution and statistical significance of pairwise differences of the structure UFSP-Net vs. the other state-of-art Models is determined by a t-test (▲/▼ for α = 0.1)

| Algorithms  | Metrics | Fire Type 1       | Fire Type 2       | Fire Type 3       |
|-------------|---------|-------------------|-------------------|-------------------|
| Random Forest | RMSE    | 2.6391±0.0312     | 1.8320±0.0291     | 1.0371±0.0268     |
|             | MAPE    | 0.1803±0.0035     | 0.2201±0.0042     | 0.4788±0.0139     |
|             | JS      | 0.1059±0.0031     | 0.1281±0.0029     | 0.2286±0.0038     |
| XGBoost     | RMSE    | 2.5478±0.0277     | 1.7679±0.0167     | 0.9901±0.0125     |
|             | MAPE    | 0.1712±0.0017     | 0.2107±0.0019     | 0.4611±0.0073     |
|             | JS      | 0.1048±0.0015     | 0.1275±0.0008     | 0.2264±0.0019     |
| GRUN        | RMSE    | 2.5762±0.0260     | 1.7710±0.0145     | 0.9939±0.0119     |
|             | MSPE    | 0.1719±0.0014     | 0.2115±0.0016     | 0.4656±0.0058     |
|             | JS      | 0.1047±0.0013     | 0.1273±0.0004     | 0.2265±0.0012     |
| Conv-GRUN   | RMSE    | 2.4612±0.0256     | 1.7012±0.0141     | 0.9815±0.0118     |
|             | MAPE    | 0.1683±0.0017     | 0.2078±0.0018     | 0.4589±0.0054     |
|             | JS      | 0.1045±0.0011     | 0.1272±0.0006     | 0.2223±0.0017     |
| UFSP-Net    | RMSE    | 2.2503±0.0229▲    | 1.6860±0.0132▲    | 0.9747±0.0148▲    |
|             | MAPE    | 0.1642±0.0012▲    | 0.2023±0.0024▲    | 0.4512±0.0057▲    |
|             | JS      | 0.1045±0.0008     | 0.1271±0.0009     | 0.2217±0.0011▲    |

As the instantiation, the predicted FSAIs at one randomly selected time slots from UFSP-Net in Fig 3 are satisfying to some extent. Notice that, to more easily outline the spatial distribution, the normalized FSAIs form are applied in this case. For Fire type1 and Fire type2 with more adequate spatial distribution statistics, details in images can be well predicted, accompanied with lower MSPE and JS divergence. And for Fire type3, rough hotspots of situation can be earlier perceived though regional details need further polished. The predicted situation images through UFSP-Net enable us to roughly master the spatial fire dynamics of different types.
6. Conclusion

In this paper, we make attempts to utilize deep learning approach for urban fire situation modelling and a novel deep learning framework UFSP-Net is introduced in this domain. Though it is still tough to precisely predict urban fire situation in both scales and spatio-temporal distribution, the experimental results of UFSP-Net is remarkable in comparison to other state-of-art traditional models in this domain and suggests that the combination of pixel-level information and graph structure information makes the awareness about urban fire situation more foreseeable.

In the future, it is potential to make further improvements on the structure of UFSP-Net by fusing more domain graph structure information that can be associated to real word statistical indexes and involving attention mechanism for capturing long-term dependence. Meanwhile, some possible extensions could be made to predict other social events for urban safety, for instance, urban drought and flood forecast and the urban air pollution monitoring with improved deep learning frameworks.

References

[1] "Fire department operations, management and safety," https://www.usfa.fema.gov/operations/.
[2] J. Corcoran, and G. Higgs, “Special issue on spatial analytical approaches in urban fire management,” Fire Safety Journal, no. 62, pp. 1-2, 2013.
[3] K. Wuschke, J. Clare, and L. Garis, “Temporal and geographic clustering of residential structure fires: A theoretical platform for targeted fire prevention,” Fire Safety Journal, vol. 62, pp. 3-12, 2013.
[4] M. Angelidou, A. Psaltoglou, N. Komninos, C. Kakderi, P. Tsalouroulos, and A. Panori, “Enhancing sustainable urban development through smart city applications,” *Journal of Science and Technology Policy Management*, vol. 9, no. 2, pp. 146-169, 2018.

[5] P. Cortez, and A. d. J. R. Morais, “A data mining approach to predict forest fires using meteorological data,” 2007.

[6] B. Singh Walia, Q. Hu, J. Chen, F. Chen, J. Lee, N. Kuo, P. Narang, J. Battis, G. Arnold, and M. Madaio, "A Dynamic Pipeline for Spatio-Temporal Fire Risk Prediction." pp. 764-773.

[7] R. M. Cichy, A. Khosla, D. Pantazis, A. Torralba, and A. Oliva, “Deep neural networks predict hierarchical spatio-temporal cortical dynamics of human visual object recognition,” *arXiv preprint arXiv:1601.02970*, 2016.

[8] Y. LeCun, and Y. Bengio, “Convolutional networks for images, speech, and time series,” *The handbook of brain theory and neural networks*, vol. 3361, no. 10, pp. 1995, 1995.

[9] T. N. Kipf, and M. Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.

[10] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur, "Recurrent neural network based language model."

[11] K. Himoto, and T. Tanaka, “Development and validation of a physics-based urban fire spread model,” *Fire Safety Journal*, vol. 43, no. 7, pp. 477-494, 2008.

[12] J. Quartieri, N. E. Mastorakis, G. Iannone, and C. Guarinaccia, “A cellular automata model for fire spreading prediction,” *Latest Trends on Urban Planning and Transportation*, pp. 173-178, 2010.

[13] Z. Zheng, W. Huang, S. Li, and Y. Zeng, “Forest fire spread simulating model using cellular automaton with extreme learning machine,” *Ecological Modelling*, vol. 348, pp. 33-43, 2017.

[14] R. Ardianto, and P. Chhetri, “Modeling Spatial–Temporal Dynamics of Urban Residential Fire Risk Using a Markov Chain Technique,” *International Journal of Disaster Risk Science*, vol. 10, no. 1, pp. 57-73, 2019.

[15] L. Wu, and A. Ren, “Urban fire risk clustering method based on fire statistics,” *Tsinghua Science and Technology*, vol. 13, no. S1, pp. 418-422, 2008.

[16] J. Clare, L. Garis, D. Plecas, and C. Jennings, “Reduced frequency and severity of residential fires following delivery of fire prevention education by on-duty fire fighters: Cluster randomized controlled study,” *Journal of safety research*, vol. 43, no. 2, pp. 123-128, 2012.

[17] H. K. Preisler, and A. L. Westerling, “Statistical model for forecasting monthly large wildfire events in western United States,” *Journal of Applied Meteorology and Climatology*, vol. 46, no. 7, pp. 1020-1030, 2007.

[18] T. Cheng, and J. Wang, “Integrated spatio-temporal data mining for forest fire prediction,” *Transactions in GIS*, vol. 12, no. 5, pp. 591-611, 2008.

[19] S. Lall, and B. Mathibela, "The application of artificial neural networks for wildfire risk prediction." pp. 1-6.

[20] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, “Efficient deep CNN-based fire detection and localization in video surveillance applications,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, no. 99, pp. 1-16, 2018.

[21] X. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." pp. 802-810.

[22] D. Chai, L. Wang, and Q. Yang, "Bike flow prediction with multi-graph convolutional networks." pp. 397-400.

[23] T. M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur, "Recurrent neural network based language model."