IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 19, 2020, accepted January 27, 2021, date of publication February 3, 2021, date of current version February 10, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3056695*

# Super Resolution Deduction: Inferring Fine-Grained Capacity for Urban Signal Station Deployment

**YONGJIAN YANG, JUFENG HOU, AND YUANBO XU**
College of Computer Science and Technology, Jilin University, Changchun 130012, China

Corresponding author: Yuanbo Xu (yuanbox@jlu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61772230 and Grant 61972450, in part by the Natural Science Foundation of China for Young Scholars under Grant 61702215 and Grant 62002132, in part by the China Postdoctoral Science Foundation under Grant 2017M611322 and Grant 2018T110247, and in part by the Changchun Science and Technology Development Project under Grant 18DY005.

**ABSTRACT** With the boosting of mobile devices, wireless sensor networks, and the internet of things, abundant multi-modal data, such as GPS signal, sensor data, are produced intentionally or unintentionally, which can represent the people's active patterns, vehicle's routes, and city's flows to develop a smart city. These multi-modal data are usually transmitted and received by signal stations deployed in the city. However, reasonably choosing the signal stations' locations is still an open issue for enhancing people's life quality in the smart city. To this end, we propose the Super Resolution Deduction (SRD) model for solving the signal station selection problem. SRD first initializes the city map as a coarse-grained heat map representing the capacity of the signal stations. Then an image-based super-resolution deduction model is proposed to obtain a fine-grained signal station capacity for deploying. To be specific, we employ Dense Block to capture the spatio-temporal correlations, C-Attention to selectively enhance useful feature maps, and S-Distribution to impose structural constraints. By sharing the GPS data load with the new deployment of signal stations, we ensure the smart city's efficiency and effectiveness. Extensive experimental results on real-world dataset Changchun City demonstrate that our proposed model achieves the superior performance among the state-of-the-art baselines.

**INDEX TERMS** Signal station capacity, signal station selection, C-Attention, spatio-temporal correlations.

## I. INTRODUCTION

Nowadays, the smartphone has become an indispensable part of people's daily life and its penetration rate in China has exceeded 112.23 percent. The signal strength of a smartphone often greatly affects the user experience. Specifically, due to the limited load capacity of signal stations, the high traffic flow may cause the overload of the signal stations and affects the signal of the smartphone. Whereas, the region with fewer people may face the problem of signal station deployment waste. Hence, how to choose reasonable locations to deploy new signal stations is particularly important. We need a reliable method to deploy new signal stations to share the GPS data load of the existing signal stations. So we propose the Super Resolution Deduction (SRD) model for solving the signal station selection problem. It can infer the fine-grained signal station capacity from the coarse-grained signal station capacity to get the accurate locations of new signal stations for deployment. Moreover, this method can make a great contribution to the deployment of new signal stations in urban development planning and the new city construction.

But traditional computer vision methods can't directly obtain accurate results under this complex background. Meanwhile, inferring the fine-grained signal station capacity faces the following challenges: 1) It is generally believed that the bustling commercial and residential areas are the most populous. However, the signal station capacity is not only related to the function of the area but also the temporal factors, i.e., the signal station capacity at different periods is significantly different. For example, on working days, people move from residential areas to office areas in the morning, and the number of people in entertainment venues increases at night. It can be seen that the signal station capacity in different places changes over time. As a result,

The associate editor coordinating the review of this manuscript and approving it for publication was Sathish Kumar.

how to incorporate both spatial and temporal factors in our problem is an important issue. 2) There are also some structural constraints [1] (i.e., spatial levels) in our problem. The sum of the signal station capacity volumes in subregions strictly equals that of the corresponding superregion as shown in Fig. 2. Furthermore, the signal station capacity in a region will be affected by the signal station capacity in the nearby regions, which will influence the fine-grained inference. Failure to accurately capture these features will result in a degraded performance, so we ought to consider the spatio-temporal correlations between different regions to improve the accuracy of our model. Fig. 1 shows the signal station capacity of residential regions near non-commercial areas and central commercial areas during the different times of working days and non-working days. It can be seen that even in the same functional area, the signal stations are under different pressures. 3) Since using the computer vision method for urban signal station deployment and each feature map carries different information, we need to think about how to selectively enhance useful feature maps to improve the accuracy of the model. 4) The distribution of the signal station capacity in a given region is affected by various external factors, especially special events. There are many typical special events, such as the beginning of the new term and convening of the meeting, which will put a lot of pressure on the signal stations in these areas. Therefore, the trend of crowd movement is also necessary to be considered.

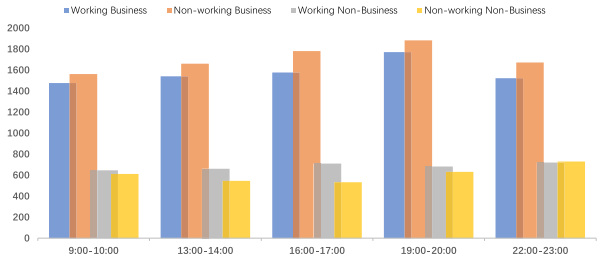



**FIGURE 1.** The pressure of signal stations in different areas.

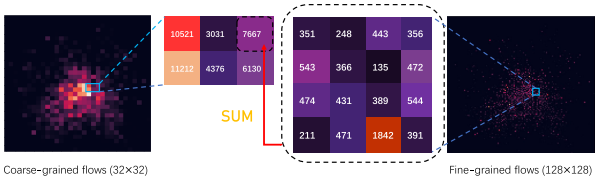



**FIGURE 2.** Signal station capacity in two granularities in Changchun.

Following the above comprehensive considerations, we propose a deep neural network model named SRD. The main contributions of this paper are as follows:

- By analyzing the particularity and challenge of the urban signal station deployment, we design a novel inference network called Super Resolution Deduction (SRD) to solve this spatio-temporal correlation problem, which employs Dense Block to address the influence of nearby regions and capture the trend of crowd movement.
- We employ the S-distribution module to impose structural constraints on the SRD, and also embed the C-Attention module to selectively enhance useful feature maps by adaptively adjusting the weight of each feature map. It turns out that using these modules will improve the inference performance of our method.
- We conduct extensive verification experiments on mobile signaling data collected from Changchun. Our verification experimental results demonstrate the significant advantages of SRD compared with the state-of-the-art baselines in effectiveness. Besides, we also verify the effectiveness of different modules in SRD. The experimental results prove that our method is a better and more effective method for urban signal station deployment.

## II. FORMULATION

*Definition 1 (Region):* As shown in Fig. 2, we divide the city into $a \times b$ grids based on latitude and longitude, where a grid represents an region [3]. By increasing $a$ and $b$ we can divide the city into smaller regions. This means we can get a smaller and more detailed mobile signaling data, which produces a more fine-grained signal station capacity.

*Definition 2 (Signal Station Capacity):* When a mobile phone is connected to a signal station to obtain service, it will passively generate a mobile signaling record. Let $\mathrm{L} \in \mathbb{R}_{+}^{A \times B}$ represents the signal station capacity at a specific time, where each $l_{a,b} \in \mathbb{R}_{+}$ represents the signal station capacity in the region $(a, b)$.

*Definition 3 (Structural Constraints):* As described in Fig. 2, given the scale factor $T$, each superregion is composed of $T \times T$ subregions. To a certain extent, it also indicates that there is a spatial structure relationship between regions that cannot be ignored. Fig. 2 illustrates an example as the scale factor $T = 4$: the superregion consists of $4 \times 4$ subregions. $l_{a,b}^{C}$ is the volume of the coarse-grained signal station capacity in a superregion, and $l_{a',b'}^{F}$ is the volume of its corresponding fine-grained signal station capacity. Given the scale factor $T$, they both obey the following equation:

$$l_{a,b}^{C} = \sum_{a',b'} l_{a',b'}^{F} \qquad \text{s.t.} \lfloor \frac{a'}{T} \rfloor = a, \lfloor \frac{b'}{T} \rfloor = b, \tag{1}$$

for simplicity, $a = 1, 2, \ldots, A$ and $b = 1, 2, \ldots, B$ in our formulation.

*Problem Statement:* Infer the fine-grained signal station capacity $L^{F} \in \mathbb{R}_{+}^{TA \times TB}$ by given a scale factor $T \in Z_{+}$ and coarse-grained signal station capacity $L^{C} \in \mathbb{R}_{+}^{A \times B}$.

## III. METHODOLOGY

Our framework follows the general procedure of super-resolution image processing. Fig. 3 depicts the framework of SRD which consists of two main components: Feature Extraction and Upsampling. The Feature Extraction module takes coarse-grained signal station capacity as its input, and then uses Dense Block [4] with bottleneck to extract

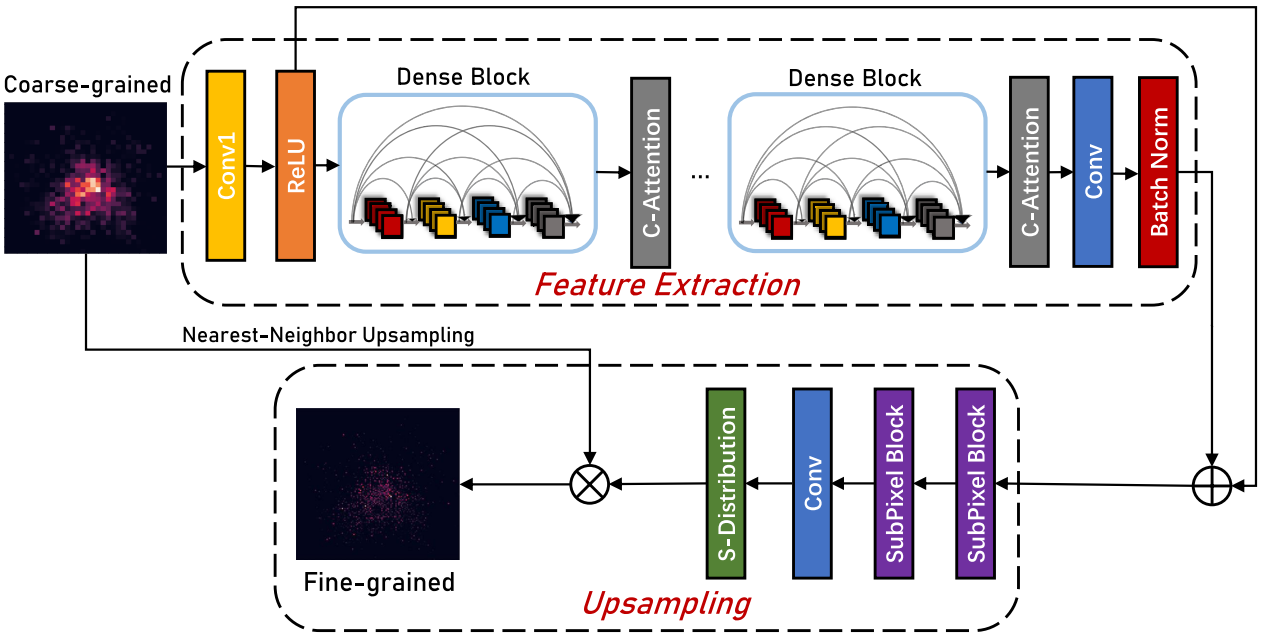



**FIGURE 3.** The SRD framework for 4× upscaling (Scale facter = 4). ⊕ denotes addition and ⊗ denotes Hadamard product.

spatio-temporal features of the city. After each Dense Block, we introduce the C-Attention layer to selectively enhance useful feature maps by the Squeeze and Excitation method [5]. By using a skip connection [6], we use identity mapping between low-level and high-level features to build an information highway [6] to skip Dense Blocks to achieve an effective gradient back-propagation. Taking weighted features as a priori knowledge, the upsampling outputs a signal station capacity distribution over subregions with respect to each superregion by employing the S-distribution [2] which imposes the structural constraints. Finally, the Hadamard product of the distribution inferred by S-Distribution with the upsampled coarse-grained signal station capacity gives the fine-grained signal station capacity as the output of SRD. We will elaborate on the detailed designs of Feature Extraction and Upsampling modules in this section.

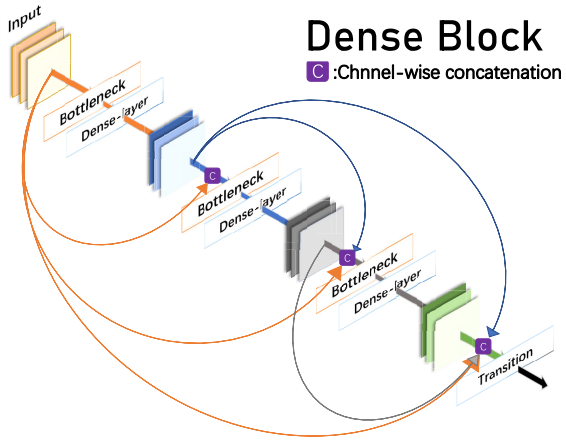



**FIGURE 4.** The Dense Block composed of dense layer and Bottleneck.

### A. FEATURE EXTRACTION

In the input stage, we use a convolutional layer [7] (F kernels with size of 9 × 9) to extract low-level features. Then M Dense Blocks with the same structure take the low-level feature map as input to construct the high-level feature map. We use C-Attention to selectively enhance useful feature maps by adaptively adjusting the weight of each feature map. As shown in Fig. 4 and Fig. 5, the Dense Block is composed of two parts: Bottleneck and Dense layer which both consist of BN, ReLU [8], and conv for reducing the amount of calculation and feature extraction respectively. Differently, the conv in Bottleneck is 1 × 1 and the one in dense layer is 3 × 3. There are M Dense Blcoks in our framework and they all contain 32 Dense layers and corresponding Bottlenecks. Since Dense Blocks can enhance feature reuse, we superimpose a depth framework to make the receiving field larger so that the features of the entire city can be fully captured. In other words, every pixel on the high-level feature map can capture the dependencies of distant or even the whole city. After feature extracting by M Dense Blocks, we use another convolutional layer (3 × 3) and Batch Normalization (BN) [9] to ensure the feature extraction. In the end, the output signal station capacity distribution will show regional dependence with the original region.

### B. C-ATTENTION

Essentially, convolution is the feature fusion of a local area, which includes feature fusion in space (H and W

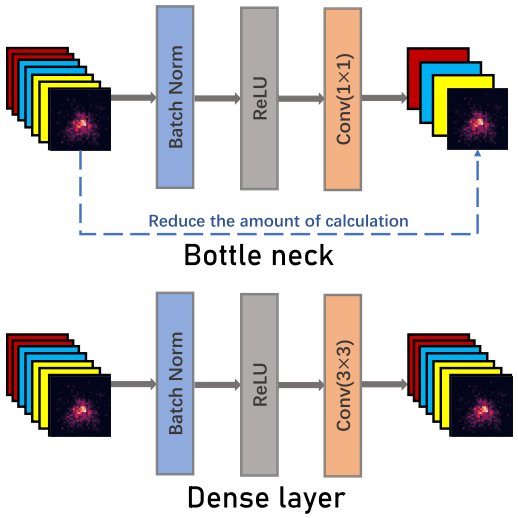



**FIGURE 5.** The structure of Bottleneck and Dense layer in Dense Block.

dimensions) and between channels (C dimension). For the channel dimension, the convolution operation fuses all the input channels by default. As shown in Fig. 6, we employ C-Attention to focus on the significance of the feature maps so that the model can adaptively adjust the weight of each feature map.

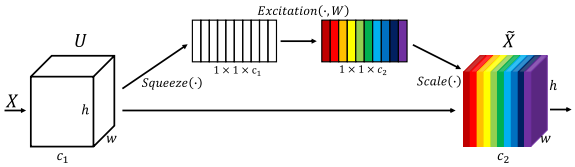



**FIGURE 6.** The structure of C-Attention module with Squeeze and Excitation.

The C-Attention mainly includes two operations, Squeeze and Excitation [5], which can be applied to any mapping $X \rightarrow U, X \in R^{H' \times W' \times C'}, U \in R^{H \times W \times C}$. Taking convolution as an example, the kernel is $k = [k_1, k_2, \cdots, k_c]$, where $k_c$ represents the $c^{th}$ kernel. Then output $U = [u_1, u_2, \cdots, u_c]$:

$$u_c = k_c \cdot X = \sum_{f=1}^{C'} k_c^s \cdot x^s, \tag{2}$$

where $\cdot$ represents the convolution operation, and $k_c^s$ represents a 3D kernel, which learns the spatial relationship of the features and sums the results of each channel, so the feature relationship of each channel is mixed with the space relationship learned by the kernel. The C-Attention is to separate this fusion in order that the model can directly learn the feature relationship between different channels.

**Squeeze (Sq).** Since convolution only takes effect in a local area, it is difficult for U to obtain enough information to extract the relationship between channels. We utilize the Squeeze in the C-Attention, using the global average pooling to encode the entire spatial features on a channel into a global feature:

$$z_c = \mathrm{Sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j). \tag{3}$$

**Excitation (Ex).** Due to Squeeze gets the global feature, we need another operation next to capture the significance of feature maps. And this operation ought to meet two criteria: first, it can learn the non-linear relationship between each feature map; second, the learned weight is not mutually exclusive. Based on this, the gating mechanism in the form of sigmoid is adopted:

$$s_c = \mathrm{Ex}(z_c, W) = \sigma(W_2 \mathrm{ReLU}(W_1 z_c)), \tag{4}$$

where $W_1 \in R^{\frac{C}{r} \times C}$, $W_2 \in R^{C \times \frac{C}{r}}$, $\sigma$ represents sigmoid.

To reduce the complexity of the model and improve the generalization ability, a bottleneck consists of two fully connected (FC) layers is adopted, and the first fully connected layer ($W_1$) is used for dimensionality reduction with the coefficient r (we set r to 16 in our experiment). And after ReLU activation, there is another FC layer ($W_2$) used to restore the original dimensions. Finally, we get the weight of each feature map:

$$\tilde{x_c} = \mathrm{Scale}(u_c, s_c) = s_c \cdot u_c. \tag{5}$$

### C. UPSAMLING

In the Upsampling module, the features extracted by Dense Blocks first pass through two SubPixel Blocks to perform an $T = 2^2$ amplification operation, thereby generating hidden features. As shown in Fig. 7, the SubPixel Block uses a convolutional ($3 \times 3$, $F \times 2^2$) layer, followed by Batch Normalization to extract features. Furthermore, it employs PixelShuffle [10] to rearrange and upsample the feature maps to double the size, and applies ReLU activation at the end. As shown in Fig. 8, the PixelShuffle layer reshapes the tensor from (H, W, $C \times r^2$) to ($H \times r$, $W \times r$, C). After each SubPixel Block, the size of the output feature maps will increase twice, while the number of channels remains unchanged.

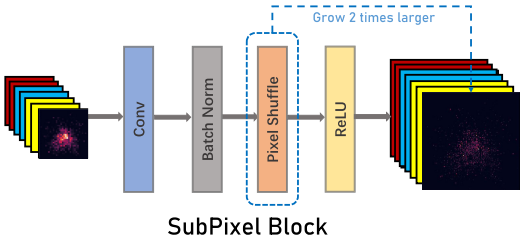



**FIGURE 7.** The overview of SubPixel Block.

Because it is different from ordinary image super-resolution, in order to infer the accurate signal station capacity, structural constraints must be fully considered. For this reason, we have introduced a structure [2] to output signal station capacity distribution.

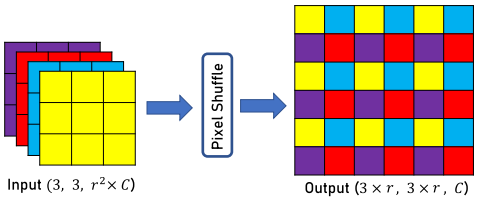



**FIGURE 8.** The PixelShuffle layer reshapes the tensor.

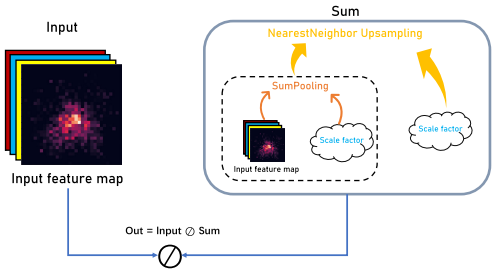



**FIGURE 9.** S-Distribution.

To be specific, we design S-Distribution as shown in Fig. 9. First, it performs sumpooling on input according to the required scale factor. Its result goes through Nearest-Neighbor Upsampling (NNUpsampling) with a scale factor to get Sum. In the end, we get the probability distribution: $Out = Input \oslash Sum$.

$$\begin{aligned} Sum_p &= \text{Sumpooling}(Input, Scalerfactor), \\ Sum &= \text{NNUpsampling}(Sum_p, Scalerfactor), \\ Out &= Input \oslash Sum, \end{aligned} \tag{6}$$

where $\oslash$ represents division.

Finally, we upscale $L^C$ to get $L^C_{up} \in R_+^{TA \times TB}$ by NNUpsampling with the Scale factor $T$ and then generate the fine-grained signal station capacity to get the final result: $\tilde{L^F} = L^C_{up} \odot Out$.

### D. OPTIMIZATION

For SRD, we can train the network through automatic backpropagation by providing training pairs $(L^C, L^F)$ and calculating the experience loss between $(L^F, \tilde{L^F})$, where $L^F$ is the true value and $\tilde{L^F}$ is the inference of our model. Concretely, we use Mean Absolute Error (MAE) for the loss function, as shown below:

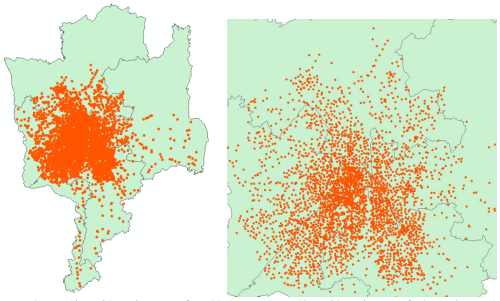
$$\text{MAE}(L^F, \tilde{L^F}) = \frac{1}{N} \sum_{i}^{N} \left| L_i^F - \tilde{L_i^F} \right|. \tag{7}$$

## IV. VERIFICATION EXPERIMENT

To verify the effectiveness of our method, we use Changchun mobile signaling data to conduct verification experiments. In order to better verify the effectiveness of our model, we select the signal stations in the main districts of Changchun. The latitude and longitude range we select is shown in Table 1 and Fig. 10.

**TABLE 1.** Latitude and longitude.

| North | South | West | East |
|---|---|---|---|
| 44.087914 | 43.729081 | 125.071501 | 125.537615 |



(a) The distribution of all signal stations. (b) The distribution of the signal stations we choose.

**FIGURE 10.** The distribution of signal stations in Changchun.

### A. SIGNALING DATASETS

Since lots of fine-grained mobile signaling data is available as ground truth, we obtain the coarse-grained signal station capacity by aggregating subregion signal station capacity from the fine-grained counterparts. Specifically as shown in Fig. 2, we divide the study area into $32 \times 32$ grids, count the signal station capacity information in a grid every hour, and infer the fine-grained signal station capacity of the four-fold resolution ($128 \times 128$). In our experiment, we divide all data into non-overlapping training data, test data, and verification data according to a ratio of about 3: 1: 1. The training datasets are shown in Table 2.

**TABLE 2.** Training dataset description.

| Dataset | Value |
|---|---|
| Time span | 9/1/2017–9/30/2017 |
| Time interval | 1 hour |
| Coarse-grained size | $32 \times 32$ |
| Fine-grained size | $128 \times 128$ |
| Scaler factor (T) | 4 |

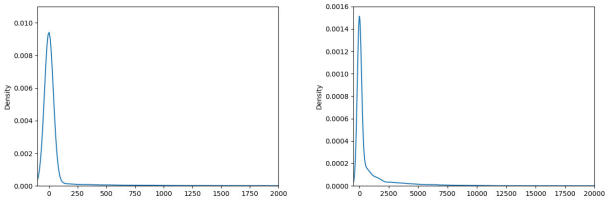
The sample of the mobile signaling data is shown in Table 3. The serial number is the unique identification of each signal station. And Connection time is the moment when the device connects to the signal station.

**TABLE 3.** The sample of the mobile signaling data.

| Longitude | Latitude | Serial number | Connection time |
|---|---|---|---|
| 125.32472 | 43.88861 | 17167 | 201709022202 |

### B. NORMALIZATION METHOD

To speed up the convergence of SRD, data normalization is adopted before the training. Since the input and output are both positive, we use Min-Max normalization to map



(a) The density of fine-grained signal station capacity. (b) The density of coarse-grained signal station capacity.

**FIGURE 11.** The distribution of signal station capacity.

the input and output to [0,1]. From the coarse-grained and fine-grained density curve in Fig. 11, we can see a long tail in both settings. The explanation is that there is sometimes a high signal station capacity in some regions, which may be due to the rush hours or traffic jams [1]. If we directly use the maximum signal station capacity as the max-scaler, it is inaccurate and most values will be much smaller than 1. Therefore, we choose 250 as the max-scaler for fine-grained data, and 5000 for coarse-grain data.

### C. EVALUATION METRICS

We use two evaluation metrics to assess the model from different aspects: Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The formulas are as follows:

$$\begin{aligned} \mathrm{MAE}(L^F, \tilde{L^F}) &= \frac{1}{N}\sum_{i}^{N}\left|L_i^F - \tilde{L_i^F}\right|, \\ \mathrm{MAPE}(L^F, \tilde{L^F}) &= \frac{100\%}{N}\sum_{i}^{N}\left|\frac{L_i^F - \tilde{L_i^F}}{L_i^F}\right|, \end{aligned} \tag{8}$$

where $N$ is the samples' total number, $\tilde{L_i^F}$ is $i^{th}$ the inferred value and $L_i^F$ is the corresponding true value.

### D. BASELINE

We compare the model with the following seven baselines in a variety of fields. The baselines are Mean Partition, Historical Average, SRCNN, ESPCN, VDSR, SRResNet, DeepSD. We detail them as follows:

**Mean Partition**: We evenly distribute the signal station capacity from each superregion to $4 \times 4$ subregions in the coarse-grained map.

**Historical Average**: Similar to distributional upsampling, Historical Average treats the value over each subregion a fraction of the value in the respective super region, where the faction is computed by averaging all training data.

**SRCNN** [11]: SRCNN is the first successful method to introduce convolutional neural networks (CNN) into image super-resolution problems.

**ESPCN** [10]: An efficient sub-pixel convolutional neural network (ESPCN) employs a sub-pixel convolutional layer, which aggregates feature maps in the LR space and constructs an SR image in one step.

**VDSR** [12]: Since SRCNN has a series of shortcomings such as slow convergence speed, VDSR adopts residual learning to improve the learning rate, accelerate the convergence speed, and deepen the network structure to improve accuracy.

**SRResNet** [13]: SRResNet enhances VDSR by using the residual architecture proposed by Hu *et al.* [6].

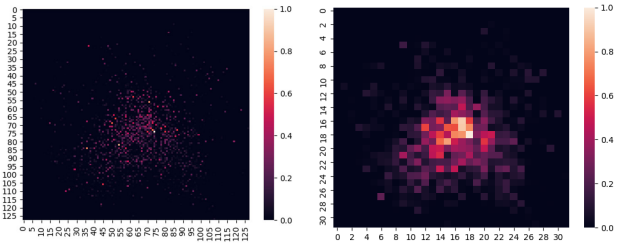
**DeepSD** [14]: DeepSD is the latest method for meteorological statistics, which is basically performed by simply stacking SRCNN. Compared with our method, the disadvantage is that it has too many parameters.

To evaluate each component of SRD, we also compare it with different variants of SRD:

**SRD-C**: In order to assess the effectiveness of the C-Attention part, we only remove this module from SRD and keep the other components unchanged.

**SRD-S**: Similarly, we remove the S-Distribution module from SRD to further verify its effectiveness.

**SRD-CS**: We remove the C-Attention and S-Distribution modules from the SRD and keep the other components unchanged.



(a) Fine-grained signal station capacity in Changchun($128 \times 128$). (b) Coarse-grained signal station capacity in Changchun($32 \times 32$).

**FIGURE 12.** Visualization of signal station capacity in Changchun.

### E. EXPERIMENTAL RESULT

The final result is shown in Fig. 12 (a). We summarize the experimental results and find that our model improve MAE and MAPE compared with other baselines, and other experiments on different M-F will be discussed later. In addition, we postpone the results of SRD-C and SRD-S to the next experiment for more detailed research. As shown in Table 4, we mark the best (★) and second best (∗) results in all experiments. And we observe that SRD is the best of all methods by 6.9% and 9% higher than the second for MAE and MAPE.

**TABLE 4.** Results comparisons. ★ and ∗ represents the best and second best result. SRD advances it by 6.9% and 9% for MAE and MAPE.

| Methods | MAE | MAPE |
|---|---|---|
| MEAN | 16.437 | 5.437 |
| HA | 5.377 | 0.671 |
| SRCNN | 4.165 | 0.411 |
| ESPCN | 4.051 | 0.501 |
| DeepSD | 3.997 | 0.408 |
| VDSR | 3.966 | 0.399∗ |
| SRResnet | 3.877∗ | 0.401 |
| **SRD** | **3.608**★(6.9%) | **0.363**★(9%) |

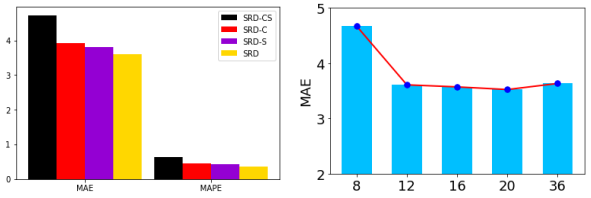
#### 1) STUDY ON S-DISTRIBUTION

To verify the effectiveness of the S-Distribution, we will compare SRD-S with other variants. From Fig. 13 (a), the S-Distribution module advances the SRD-CS in MAE by 17.08% and in MAPE by 28.37%, which verifies the superiority of the S-Distribution module.

#### 2) STUDY ON C-ATTENTION

The C-Attention incorporates two important operations: Squeeze and Excitation to describe the interdependence between feature maps and automatically obtains the weight of each feature map through learning. Hence, the C-Attention can better capture the significance of the different feature map. As shown in Fig. 13 (a), we compare SRD-C with other variants. The C-Attention module advances the SRD-CS in MAE by 19.3% and in MAPE by 21.3%, which verifies the superiority of the C-Attention module.

SRD is superior to SRD-C and SRD-S in MAE and MAPE, which proves that the combination of the two modules can further improve the model performance.




(a) The results of different variants of SRD. (b) The MAE of SRD in different Growth rate.

**FIGURE 13.** Performance comparison of various structural constraints.

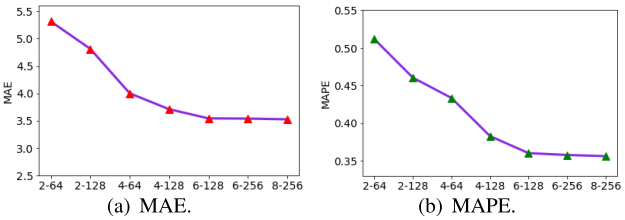
#### 3) STUDY ON PARAMETER SIZE

The setting of M-F in Fig. 14 presents the number of Dence Blocks (M) and the kernels (F) of Conv1. We can see that adding more Dense Blocks and more kernels can improve the performance of the model, but it also increases the training time and memory space. After considering the tradeoffs between cost and performance, we set the default M-F of SRD to 6-128.

Different from Resnet, each convolutional layer of the Dense Block outputs $k$ feature maps. Notably, due to the feature reuse of the Dense Block, as the network deepens, the Dense Block will produce more outputs, So the size of $k$ needs to be set appropriately. Specifically, $k$ is a hyperparameter called the growth rate, and then we conduct a series of comparison experiments on it. As shown in Fig. 13 (b), we find that setting a larger growth rate can improve our model's performance, but when the growth rate is 36, the error of our model increases. This can be explained as when the model pays too much attention to feature reuse, it will cause overfitting. After taking into account the tradeoffs of cost, performance and accuracy, we set the default $k$ of SRD to 12.

### F. INFERENCE ERROR ANALYSIS

Fig. 15 displays the inference error $|L_i^F - \tilde{L}_i^F|$ and the brighter the pixel, the greater the error. We find three areas with large error in the figure and locate the areas based on the latitude and longitude. Two of these regions are Jingyang Square and Rainbow Square, the other is Xinhai Fishing Park, which is an entertainment venue. According to the findings, we conclude that regions of the inference with large errors are the regions with high traffic flow or numerous traffic jams. Such areas are usually unstable and are easily affected by weather and external factors. Therefore, there will be a big deviation in the statistics of mobile signaling data of these regions. This indicates that compared with ordinary image super-resolution methods, we ought to focus on some constraints, which require special design.




(a) MAE. (b) MAPE.

**FIGURE 14.** The results of different M-F settings.

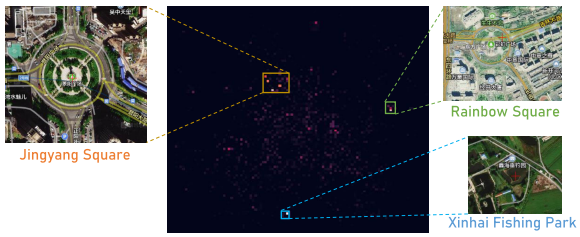



**FIGURE 15.** Visualization for inference errors in Jingyang Square, Rainbow Square and Xinhai Fishing Park.

## V. RELATED WORK

### A. IMAGE SUPER RESOLUTION METHOD

Research on super-resolution images, trying to recover from coarse-grained low-resolution images to high-resolution images, has received more and more focus in recent years. It has also been directly applied in many fields, such as face recognition [15], fine-grained crowdsourcing [16] and HDTV [17]. There are also a great number of image super-resolution algorithms developed in the computer vision field. To solve this problem, the early research pay attention to the interpolation methods, such as Lanczos resampling [18]. Furthermore, some researches introduce statistical image priors [19], [20] to obtain better performance. The advanced methods focus on learning nonlinear mapping between low-resolution images and high-resolution images, such as neighbor embedding [21] and sparse coding [22], [23]. However, these methods still have insufficiency, and they are still not enough to reconstruct realistic and fine-grained textures of images.In recent years, a series of approaches based on deep learning have made great progress in reconstructing high-resolution images. In the early days,

the CNN method is used to make low-resolution and high-resolution images through end-to-end mapping to form the image super-resolution architecture of CNN. Subsequently, an approach based on sub-pixel convolutional layer is proposed, which can reduce the cost of restoring high-resolution pictures. Inspired by VGGNet in image classification [24], a deep CNN [12] is applied to the method of recovering high-resolution images. Although the application of the residual module in [12] has shown that the depth has a positive effect on accuracy, as the network's depth increases, the efficiency and speed decrease significantly, so it is difficult to improve performance only by deepening the network depth. Later, in [13], a perceptual loss function is proposed to focus on the better reconstruction of high-frequency details. In [14], the performance of the model is improved by deleting the redundant modules. However, the above methods are only applicable to the pure image problem, and are not suitable for our actual problem in terms of efficiency and effectiveness, because mobile signaling data has a very specific spatio-temporal correlation compared with natural images. In the study of UrbanFM [2], we can fully understand how to improve the accuracy of the model through structural constraints. Nevertheless, in order to solve the deployment of signal station problem, we need to capture the trend of crowd movement, so we employ the Dense Block for feature reuse to capture this trend.

### B. SIGNAL STATION DEPLOYMENT ANALYSIS

The analysis and application of traffic data has become a hot research topic in recent years [1]. In the early work [25], [26], predictions of individual urban trajectories have already begun. Nowadays, the forecast of total traffic flow outside the trajectory [27] has gradually attracted people's attention. Inspired by deep learning techniques, more and more neural network models are applied in the transportation field. For example, in [28], the correlations between time and space are modeled and studied to achieve better prediction results. But in terms of the deployment of the signal stations, the existing methods cannot reasonably deal with this issue. Therefore, motivated by this, we propose the novel SRD framework to solve the problem with efficiency and effectiveness.

## VI. CONCLUSION

In this paper, we present a novel model called SRD based on deep convolutional neural networks, which improves the general process of image super-resolution methods to achieve fine-grained signal station capacity inference. SRD has addressed the challenges specific to the Urban Signal Station Deployment problem. Meanwhile, C-Attention and S-Distribution are employed to extract the feature importance and spatial structural constraints, and we employ Dense Block to further enhance feature reuse and capture the mobility trends. Experiments have shown that our method advances traditional methods by at least 6.9% and 9% in terms of MAE and MAPE. Moreover, we have made a new attempt on the deployment of urban signal stations and applied the SRD method in real life to deal with the issue of signal station deployment in cities. In the future, we will do more studies on improving the model structure and considering external factors, and pay more attention to reducing inference errors in some hard areas to solve the problem of signal station deployment.

## REFERENCES

[1] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing: Concepts, methodologies, and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, pp. 38:1–38:55, 2014, doi: 10.1145/2629592.

[2] Y. Liang, K. Ouyang, L. Jing, S. Ruan, Y. Liu, J. Zhang, D. S. Rosenblum, and Y. Zheng, "UrbanFM: Inferring fine-grained urban flows," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, A. Teredesai, V. Kumar, Y. Li, R. Rosales, E. Terzi, and G. Karypis, Eds., Jul. 2019, pp. 3132–3142, doi: 10.1145/3292500.3330646.

[3] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction," in *Proc. 31st AAAI Conf. Artif. Intell.*, S. P. Singh and S. Markovitch, Eds., San Francisco, CA, USA. Palo Alto, CA, USA: AAAI Press, Feb. 2017, pp. 1655–1661. [Online]. Available: http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14501

[4] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA. Washington, DC, USA: IEEE Computer Society, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.

[5] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA. Washington, DC, USA: IEEE Computer Society Jun. 2018, pp. 7132–7141. [Online]. Available: http://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Squeeze-and-Excitation_Networks_CVPR_2018_paper.html

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. 14th Eur. Conf., Comput. Vis. (ECCV)*, in Lecture Notes in Computer Science, vol. 9908, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Amsterdam, The Netherlands: Springer, Oct. 2016, pp. 630–645, doi: 10.1007/978-3-319-46493-0_38.

[7] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989, doi: 10.1162/neco.1989.1.4.541.

[8] R. H. R. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, and H. S. Seung, "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit," *Nature*, vol. 405, no. 6789, pp. 947–951, Jun. 2000.

[9] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn. Workshop Conf. (ICML)*, F. R. Bach and D. M. Blei, Eds., Lille, France, vol. 37, Jul. 2015, pp. 448–456. [Online]. Available: http://proceedings.mlr.press/v37/ioffe15.html

[10] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA. Washington, DC, USA: IEEE Computer Society, Jun. 2016, pp. 1874–1883, doi: 10.1109/CVPR.2016.207.

[11] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

[12] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA. Washington, DC, USA: IEEE Computer Society, Jun. 2016, pp. 1646–1654, doi: 10.1109/CVPR.2016.182.

[13] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 105–114, doi: 10.1109/CVPR.2017.19.

[14] T. Vandal, E. Kodra, S. Ganguly, A. Michaelis, R. Nemani, and A. R. Ganguly, "DeepSD: Generating high resolution climate change projections through single image super-resolution," in *Proc. 23rd Int. Conf. Knowl. Discovery Data Mining (ACM SIGKDD)*, Halifax, NS, Canada, Aug. 2017, pp. 1663–1672, doi: 10.1145/3097983.3098004.
[15] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, and R. M. Mersereau, "Eigenface-domain super-resolution for face recognition," *IEEE Trans. Image Process.*, vol. 12, no. 5, pp. 597–606, May 2003, doi: 10.1109/TIP.2003.811513.
[16] M. W. Thornton, P. M. Atkinson, and D. A. Holland, "Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping," *Int. J. Remote Sens.*, vol. 27, no. 3, pp. 473–491, Feb. 2006.
[17] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
[18] C. E. Duchon, "Lanczos filtering in one and two dimensions," *J. Appl. Meteorol.*, vol. 18, no. 8, pp. 1016–1022, Aug. 1979.
[19] Y.-W. Tai, S. Liu, M. S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA. Washington, DC, USA: IEEE Computer Society, Jun. 2010, pp. 2400–2407, doi: 10.1109/CVPR.2010.5539933.
[20] J. Sun, Z. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Anchorage, AK, USA. Washington, DC, USA: IEEE Computer Society Jun. 2008, pp. 24–26, doi: 10.1109/CVPR.2008.4587659.
[21] H. Chang, D. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*. Washington, DC, USA: IEEE Computer Society, Jun./Jul. 2004, pp. 275–282, doi: 10.1109/CVPR.2004.243.
[22] R. Timofte, V. D. Smet, and L. V. Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Proc. 12th Asian Conf. Comput. Vis. Comput. Vis. (ACCV)*, in Lecture Notes in Computer Science, vol. 9006, D. Cremers, I. D. Reid, H. Saito, and M. Yang, Eds. Singapore: Springer, Nov. 2014, pp. 111–126, doi: 10.1007/978-3-319-16817-3_8.
[23] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010, doi: 10.1109/TIP.2010.2050625.
[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, Y. Bengio and Y. LeCun, Eds., San Diego, CA, USA, 2015, pp. 1–14. [Online]. Available: http://arxiv.org/abs/1409.1556
[25] Z. Fan, X. Song, R. Shibasaki, and R. Adachi, "CityMomentum: An online approach for crowd behavior prediction at a citywide level," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (UbiComp)*, K. Mase, M. Langheinrich, D. Gatica-Perez, H. Gellersen, T. Choudhury, and K. Yatani, Eds., Osaka, Japan. New York, NY, USA: ACM, Sep. 2015, pp. 559–569, doi: 10.1145/2750858.2804277.
[26] X. Song, Q. Zhang, Y. Sekimoto, and R. Shibasaki, "Prediction of human emergency behavior and their mobility following large-scale disaster," in *Proc. 20th Int. Conf. Knowl. Discovery Data Mining (ACM SIGKDD)*, 2014, pp. 5–14.
[27] M. X. Hoang, Y. Zheng, and A. K. Singh, "FCCF: Forecasting citywide crowd flows based on big data," in *Proc. 24th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, S. Ravada, M. E. Ali, S. D. Newsam, M. Renz, and G. Trajcevski, Eds., Burlingame, CA, USA, Oct. 2016, pp. 6:1–6:10, doi: 10.1145/2996913.2996934.
[28] J. Zhang, Y. Zheng, D. Qi, R. Li, and X. Yi, "DNN-based prediction model for spatio-temporal data," in *Proc. 24th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, S. Ravada, M. E. Ali, S. D. Newsam, M. Renz, and G. Trajcevski, Eds., Burlingame, CA, USA, Oct. 2016, pp. 92:1–92:4, doi: 10.1145/2996913.2997016.

**YONGJIAN YANG** is currently a Professor and a Ph.D. Supervisor with Jilin University, the Vice Dean of the Software College of Jilin University, the Director of the Key Laboratory under the Ministry of Information Industry, the Standing Director of the Communication Academy, and a member of the Computer Science Academy of Jilin Province. He participated three projects of NSFC, 863 and funded by the National Education Ministry for Doctoral Base Foundation. He has authored 12 projects of NSFC, key projects of Ministry of Information Industry, Middle and Young Science and Technology Developing funds, Jilin Provincial programs, Shenzhen, Zhuhai, and Changchun. His research interests include network intelligence management, wireless mobile communication and services, research and exploiture for next generation services foundation, and key productions on wireless mobile communication.

**JUFENG HOU** received the B.E. degree in computer science and technology from Jilin University, Changchun, in 2018, where he is currently pursuing the M.E. degree with the College of Computer Science and Technology. His research interests include applications of data mining, urban computing, and mobile computing.

**YUANBO XU** received the B.Sc. and M.S. degrees from the College of Computer Science and Technology, Jilin University, Changchun, China, where he is currently pursuing the Ph.D. degree with the Key Laboratory of Symbol Computation and Knowledge Engineering, Ministry of Education. His research interests include applications of data mining, recommender systems, and mobile computing.