Research on Population Spatialization Method Based on AWA-ESPCN

Peng You¹,², Minghao Liu¹,²,*, Jiang Du¹,² and Lei Jing¹,²
¹College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
²Spatial Information Research Center, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

*Corresponding author: liumh@cqupt.edu.cn

Abstract. The spatialization of population is the process of rasterizing demographic data, which is of great significance for government decision-making and promoting the coordinated and sustainable development of social economy, resources and environment. In the continuous research of multiple linear regression, spatial interpolation, the use of multi-source heterogeneous data, random forest and other methods, a batch of highly representative global and national-scale population spatial databases have been born. Therefore, through the super-resolution of deep learning the method is more valuable for population low-scale sampling research and end-to-end mapping to achieve high-precision downscaling and enhance the robustness and generalization of the training model. This paper is based on the population statistics of the counties in Chongqing municipality in 2010, through the area-weighted rasterization coupled sub-pixel convolutional neural network (ESPCN) to learn the global and local features of the population to achieve 500m resolution population spatialization, and compare it with the other 2 compare plan. The results show that the training model using this method has the lowest mean square error, the smallest residual interval, and the best picture effect.

Keywords: area weighted average population rasterization (AWA), sub-pixel convolutional neural network (ESPCN), population spatialization

1. Introduction
Population spatialization is similar to the downscaling process of early climate data [1]. It is a way of converting the low spatial resolution statistical population at a certain point in the study area into a high resolution population distribution close to the real population distribution. Downscaling spatialization technology. As population spatialization data solves the spatial limitation of census data to a certain extent, it is widely used in disease management, urban planning, etc [2]. In summary, the current population spatialization modeling technology can be roughly divided into two major trends: global and local model construction.

The construction of the global model has gone through the calculation of the average population in the area [3] to meet the needs of low-resolution data, to the use of spatial interpolation methods to realize
the transformation of demographic data into fine spatial units; with the development of remote sensing and GIS technology. With development, multi-source heterogeneous data such as ETM image, elevation, slope, night light data, European distance between land and river, and land classification are gradually applied. Among them, night light data such as DMSP-OLP [4], NPP/VIIRS [5] and the difference between the two are frequently studied; and then the partition density mapping began to be combined with the tree model and the multiple regression model [10]. The random forest model is widely used in population spatialization research because of its high flexibility [6] and the measurability of the importance of variables [7, 8]. Compared with random forest, the multiple regression model [9] considers the correlation between different impact factors. In view of the fact that the global model is difficult to describe the heterogeneity of the population spatial distribution, local models such as geographically weighted regression and super-resolution convolutional networks [10] have been introduced into the research practice of population spatialization. Among them, the geographically weighted regression model reduces the simulation accuracy due to the influence of complex terrain. Based on the commonality of single image super-resolution and down-scaling technology, Zong Zefang et al. used the super-resolution convolutional neural network model (SRCNN) for the first time to introduce the whole-day population spatialization research in Shanghai, and obtained better results [11]. Thomas et al. tried to generate high-resolution climate change predictions through single-image super-resolution through SRCNN stacking elevations [12]. However, the SRCNN model has the problem of too many parameters in non-linear mapping steps and a large filter, which leads to certain information loss. Try to improve the convolution structure [13, 14] or use sub-pixel convolution [15] to better learn the local feature.

A single model can only extract global or local features separately, and cannot get rid of the influence of complex terrain. Therefore, this paper adopts two steps of area-weighted average population rasterization and sub-pixel convolutional neural network model to construct a population spatial coupling model to realize Chongqing's 2010 population spatialization with a resolution of 500m.

2. Research area and data processing

2.1. Research area

Chongqing, the city located in southwest China which is a transitional zone between the Qinghai-Tibet Plateau and the Middle and lower reaches of the Yangtze River between 105°11’~110°11’ east longitude and 28°10’~32°13’ north latitude. It is adjacent to Hubei and Hunan in the east, Guizhou in the south, Sichuan in the west and Shaanxi in the north. It covers an area of 82,400 square kilometers and has 40 districts, including 19 municipal districts, 17 counties and 4 autonomous counties. The spatial heterogeneity of population distribution in Chongqing is obvious due to the limitation of complex topography. Therefore, selecting Chongqing as the study area can better test the regional adaptability of spatial model.
2.2. Data sources and data processing

The demographic data sets used in this paper include demographic data and high-resolution raster demographic data. Demographic data Central district and county household registration population data from http://www.nature.com/sdata [16], High resolution raster population as model tag data from world Population Project (WorldPop) https://www.worldpop.org.

Driving factors in this paper includes natural factors (elevation, slope, land use type), social factors (night lighting, residential areas) and the distance factor (land and rivers Euclidean distance), a total of nine factors driving data, and the land use type of the natural factor can also been divided into four categories (water, rivers, farmland, forest land).

Elevation, slope and land use type data were obtained from global 30 m land cover data (GlobeLand30). The night light data is the 100m resolution calibration value (0-6300) obtained by multiplied the standard uncalibrated DMSP light source data (0-63) by 100. The residential data is the 100m resolution urban and rural settlements generated by Jeremiah J and other scholars using the random forest model. Vector data such as river and administrative boundary are derived from free street view data, so Euclidean distance between land and river can be calculated. The original land use type (cultivated land, forest land, shrub land, grassland, wetland, water, and land for construction purposes 7 class) be grouped into four broad categories, the wetlands, water, as uninhabited waters, incorporating forest and shrub land to forest land, arable land, grassland into farmland, construction land remains the same, extraction of each land use type at same time to form the 01 figure.

3. Research methods and basic processes

3.1. Research method

The low-resolution population data collected in this study are census data of districts and counties as statistical units. First of all, in the "De-Coarse" module, a area-weighted average rasterization model is adopted to realize global feature learning and complete the 500m resolution rasterization of district/county population census data. Although the area-weighted raster population is spatially low in accuracy, but ensures that the resolution of the raster is the same as that of the target raster, so it is
regarded as a process of "De-Coarse". In this paper, the results of "De-Coarse" are recorded as Low Resolution population (Low Resolution Population, LR population). Then, in "Take-Fine" module to construct a variety of grid model, take above rasterize population data and nine factors driving data as features, and the processing after the 500 m resolution raster data grid population as the label input to the super-resolution network model of blended learning to realize global and local features, get a higher accuracy in the spatial distribution of the population. Finally, numerical and spatial accuracy verification methods such as mean square error (MSE), root mean square error (RMSE) and spatial residual error are used to compare and verify the spatialization results of various coupling models.

In order to meet the requirements of the super-resolution model for the spatial position of the input data, all raster data were converted into ASCII matrix to maintain the spatial position relationship and input into the super-resolution model. In order to make the super-resolution model accurately learn the boundary characteristics of Chongqing, 127 districts and counties containing the matrix scope of Chongqing were processed in the process of "De-Coarse and Take-Fine", the 40 districts and counties of Chongqing were cut out in the model test part for accuracy verification.

3.2. Basic methods and processes of research

3.2.1. Area weighted average Population Gridding (AWA). The Area weighted average (Area weighted average, AWA) accounts for the total population of statistical units divided by the total Area of statistical units (Statistical unit area: hm2). By AWA method, the statistical population is evenly distributed in administrative units at district or county level (here the unit of population density is: person/hm2). The calculation formula is as follows:

\[ LR_{ki} = \frac{P_k}{S_k} (k : 1,2,3,\ldots,n) \]

(1)

Among them, \( LR_{ki} \) represents the population of grid \( i_{th} \) in the \( k_{th} \) administrative unit. \( S_k \) is the area of the \( k_{th} \) administrative unit (Units: hm\(^2\)).

3.2.2. Super-resolution convolutional neural network (SRCNN) model. First, the low-resolution image XLR is interpolated to the target resolution size by using the bicubic interpolation technology, remember to X. The task of SRCNN is to make the result \( F(X) \) obtained after image X passes through the three-layer convolutional neural network as similar as possible to the real target image Y. That is to optimize the mapping function \( F(X) \) by minimizing the objective function (2):

\[ \arg \min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \left\| F(X_i, \theta) - Y_i \right\|^2 \]

(2)

Where, \( \theta \) is the parameter of the convolutional neural network, and \( n \) represents the number of training samples. The mapping from X to F(X) consists of three steps, each corresponding to a convolution operation.

1) Extraction and representation. Slice the image from X at a certain step length (i.e. extract overlapping patches), and then represent each patch as a high-dimensional vector through a convolutional layer. The function is as follows:

\[ F_i(X) = \max(0, W_i * X + B_i) \]

(3)

Where, 33×33 is the patch size, \( W_i \) is 64 convolution kernels of 9×9×c, 9×9 is the size of each convolution kernel, and C is the number of channels. B1 represents a set of 64 dimensional biases. The
ReLU activation function \( \text{Max} \,(0, x) \) is used to make each 33\( \times \)33 patch have a higher-dimensional representation (64 dimensions).

(2) Nonlinear mapping. Each nonlinear vector in the 64-dimension vector in (3) is mapped to the 32-dimension vector, so the convolution kernel used here is 1×1. The function is as follows:

\[
F_2(X) = \max(0, W_2 \ast F_1(X) + B_2)
\]

(4)

Where, \( W_2 \) is 32 1×1×64 convolution kernels in the second layer, and \( B_2 \) is a group of 32-dimension bias.

(3) Reconstruction. This is equivalent to a deconvolution process. The task of this layer is to aggregate the 32-dimensional features in (4) to form a high-resolution image result, namely \( F(X) \).

\[
F(X) = \max(0, W_3 \ast F_2(X) + B_3)
\]

(5)

Where \( W_3 \) is a convolution kernel of size 1×5×5×32, and \( B_3 \) is a 1-dimensional bias.

3.2.3. Accelerate the super-resolution convolutional neural network (FSRCNN) model. FSRCNN can be decomposed into five parts: feature extraction, reduction, nonlinear mapping, extension and deconvolution.

(1) Feature extraction: FSRCNN performs feature extraction on the original LR image without interpolation. In order to differentiate from SRCNN, we represent the small LR image input as \( Y_s \). We refer to the selection of SRCNN parameters \( f_1, n_1 \) and \( c_1 \). In SRCNN, the filter size of the first layer is set to 9. These filters are executed on the enlarged image \( Y \). Since most of the pixels in \( Y \) are interpolated from \( Y_s \), the 5×5 patch \( S \) in \( Y_s \) can cover almost all the information of the 9×9 patch in \( Y \). Therefore, we can adopt a smaller filter size \( f_1 = 5 \), with little information loss. For the number of channels, we set \( c_1 = 1 \) according to SRCNN. Then we only need to determine the number of filters \( n_1 \). \( n_1 \) can be regarded as the number of LR image feature dimensions, denoted as \( d \), the first sensitive variable, so the first layer can be denoted as Conv (5, \( d \), 1).

(2) Contraction: In SRCNN, the mapping step usually follows the feature extraction step, and then the high-dimensional LR image features are directly mapped to the HR feature space. However, since LR image feature size \( d \) is usually very large, the computational complexity of mapping steps is very high. Therefore, the application of 1×1 layer is considered to save calculation cost, and a reduction layer is added after the feature extraction layer to reduce the LR image feature size \( d \). The filter size was fixed as \( f_2 = 1 \) and linear combination was carried out in LR image features. By using a smaller filter number \( n_2 = s << d \), the LR image feature size is reduced from \( d \) to \( s \). Here, \( s \) is the second sensitive variable that determines the degree of contraction. The second layer is denoted by Conv (1, \( s \), \( d \)), which greatly reduces the number of parameters.

(3) Nonlinear mapping: This step is the most important part that affects the performance of SR. We adopt medium filter size \( f_3 = 3 \). In order to maintain the same good performance as SRCNN, we use multiple 3×3 layers to replace a single network layer. The number of mapping layers (expressed as \( m \)) is another sensitive variable that determines the precision and complexity of the mapping. All mapping layers contain the same number of filters \( n_3 = s \) to be consistent. The nonlinear mapping part can be expressed as \( M \times \text{Conv} \,(3, s, s) \).

(4) Expansion: The effect of the expansion layer is similar to the reverse process of the contraction layer. In order to calculate the efficiency, the shrink operation reduces the LR image feature size. If HR images are generated directly from these low-dimensional features, the final recovery quality will be poor. Therefore, an extension layer is added after the mapping section to extend the HR image element dimension. In order to maintain consistency with the shrink layer, 1×1 filter is adopted, and the number...
is the same as the number of LR image feature extraction layers. Contrary to Conv (1, s, d), the network structure is Conv (1, d, s).

(5) Deconvolution: It uses a set of deconvolution filters to carry out up-sampling and aggregation of previous features, which is the reverse operation of convolution. The filter convolves the image with a step size k, and the output is \( 1 / k \) times the input. Similarly, if we swap the input and output positions, the output will be k times the input, and the final output will be the reconstructed HR image. The deconvolution layer is denoted by DeConv (9,1,d).

The above five parts form a complete FSRCNN network, such as Conv(5,d,1) -PRelu-Conv (1,s, d) - PRelu -mXConv (3,s,s) - PRelu -Conv (1,d,s) - PRelu -DeConv (1,d,s) - PRelu -DeConv (9,1,d). Three sensitive variables (image feature size D, number of shrink filters S and mapping depth m) control performance and speed.

3.2.4. Efficient subpixel convolutional neural network (ESPCN) model. The task of single image super resolution is to re-generate the IHR image from the IHR image obtained by scaling from a given original IHR image. First, gaussian filter convolution is used to simulate the point diffusion function of the camera, and image down-sampling is multiplied by r. In general, IHR images and IHR images have C color channels, so they are represented as true tensors of size H×W×C and rH×rW×C, respectively.

(1) Two-layer convolution is first used for feature extraction. Firstly, l layer convolutional neural network is applied directly to ILR image, and then subpixel convolutional layer is applied to ILR feature map to generate ISR. For a network consisting of L layer, The L-1 layer can be described as follows:

\[
f^l(I^{LR};W_1,b_1) = \phi(W_1 * I^{LR} + b_1) \tag{6}
\]

\[
f^l(I^{LR};W_{l+1},b_{l+1}) = \phi(W_{l+1} * f^{l-1}(I^{LR}) + b_{l+1}) \tag{7}
\]

Where \( W_1, b_1, 1 \in (1, L - 1) \) are respectively the weights and deviations of the learning network, \( W_1 \) is a 2D convolution tensor with the size of \( n_l \times n_l \times k_l \times k_l \), \( n_l \) is the number of features of l layer, \( n_0=C \), and \( k_l \) is the filter size of l layer. \( b_1 \) is a vector bias of length \( n_l \). Nonlinear function (or activate function) \( \Phi \) fixed. The last layer \( f^L \) converts LR feature map \( I^{LR} \) to HR image \( I^{SR} \).

(2) Sub-pixel convolutional layer was used to extract features: convolution was carried out with step size \( 1/r \) and filter size \( W_0 \) of \( Ks \). Weights that fall between pixels are not activated and do not need to be calculated. The number of activation modes is exactly \( R^2 \), and according to its position, it is activated at most \( \left[ \frac{Ks}{r} \right]^2 \), and the filter is periodically activated during the convolution of the image :mod(x, r), mod(y, r), where x and y are the output pixel coordinates in the HR image space. An effective way to do this is mod(kS, r) = 0:

\[
I^{SR} = f^L(I^{LR}) = \rho S(W_L * f^{L-1}(I^{LR}) + b_L) \tag{8}
\]

PS is a periodic mix-wash operator, which rearranges the elements of the H×W×C tensor into the shape tensor of rH×rW×C. This operation can be described in the following way

\[
PS(T)_{x,y,c} = T_{\left\lfloor \frac{x}{r} \right\rfloor, \left\lfloor \frac{y}{r} \right\rfloor, C - \text{mod}(y,r) + \text{mod}(x,r) + c} \tag{9}
\]

The convolution operator \( W_L \) has the shape \( n_{l-1} \times r \times C \times K_L \times K_L \), and do not apply nonlinearity to the convolution output of the last layer. When \( K_L = \frac{K_S}{r} \) and mod(kS, R) = 0, it is equivalent to using the
sub-pixel convolution of filter $W_s$ on the LR image space. Finally, a high-resolution image is generated directly from the low-resolution feature map by up-sampling.

(3) Mean square error as the objective function: Given the training set consisting of HR image examples $I_{HR}^n$, $n = 1...N$, we generate the corresponding $I_{LR}^n$, $n = 1...N$ and calculate the reconstructed pixel mean square error (MSE) as the objective function of the training network:

$$l(W_{1L}, b_{1L}) = \frac{1}{r^2HW} \sum_{x=1}^{H} \sum_{y=1}^{W} \left( I_{HR}^{x,y} - f_{x,y}^{L}(I_{LR}) \right)^2$$ (10)

Considering the influence of influencing factors on population distribution, the experiment took grid population and different influencing factors as the input of ESPCN in the form of channels, and the number of channels depended on the number of auxiliary data. Two convolutional layers and a sub-pixel convolutional layer used for feature map extraction are used to get the model output once. The task of population spatialization is to make the output of the model as similar as possible to the actual population data. Therefore, the standard back-propagation stochastic gradient descent method is adopted to minimize the mean square error between the output of the model and the actual population distribution, so as to obtain the high-resolution population distribution results.

3.2.5. Model test. In this paper, the output results of each coupling model are clipped to the scope of Chongqing, and the model errors are quantified from the numerical and spatial perspectives. Numerically, the paper uses the mean square error (MSE) and root mean square error (RMSE) to measure the total error of model output. Spatially, the residuals between labels and simulation results are used to visualize spatial errors. A negative residuals means that the simulation presents an underestimated result, and a regular value means that the simulation results present an overestimated result.

4. Experimental program design and result analysis

4.1. Experimental protocol

Option 1: AWA+SRCNN. The results of AWA rasterized population are evenly distributed according to the area of districts and counties, reflecting the overall characteristics of population distribution between districts and counties. SRCNN recognizes local features through 9×9 convolution, and the AWA+SRCNN model integrates district and county-level global features and local features.

Option 2: AWA+FSRCNN. In terms of local models, FSRCNN mainly improves the SRCNN neural network, formulates a compact hourglass CNN structure, uses a smaller filter size to reduce information loss, and achieves acceleration while maintaining its excellent performance.

Option 3: AWA+ESPCN. On the local model, the ESPCN sub-pixel layer has two convolutional layers for feature map extraction and a sub-pixel convolutional layer, using a smaller filter size to integrate the same information while maintaining the given context area, and has a larger receptive field, providing more contextual information; In addition, for the network with L layer, we learn the $n^{11}$ magnification filter of $n^{11}$ feature mapping instead of a filter that upgrades the input image, so the network can learn better and more complex end-to-end mapping from LR images to HR images, and can generate more realistic details.

4.2. Result analysis

The three schemes use the same impact factor, and through 200 iterations of training, the population spatialization results shown in Figure 4 are obtained. The results show that under the combined action of various natural and social factors, after various super-resolution models are 'extracted', the three schemes all show good simulation results. Because most of the southeast and northeast of Chongqing area are mountainous terrain, and the relatively low topography in the central and western regions, the
population distribution is affected by the elevation slope. Generally speaking, the population in the central and western regions is relatively dense, while the population distribution in the southeast and northeastern regions is sparse.

Further analysis of the forecast data: (1) Schemes 1 and 2 are clustered together in the nine districts of the main city, while scheme 3 is relatively scattered, not only close to the verification data, but also more in line with the actual population distribution; (2) scheme 3 the population distribution range of is closer to the verification data. The maximum value of Option 1 is much smaller than that of the verification data, and the maximum value of Option 2 is much larger. This is related to the lack of detailed local feature sampling. (3) Scheme 1 and Scheme 2 are not obvious in the population agglomeration distribution in southeastern Chongqing, while Scheme 3 can be well highlighted.

![Figure 2. Comparison of Population Spatialization Results with Verification Data.](image)

**4.3. Precision inspection**

In this paper, the mean square error (MSE) and root mean square error RMSE are used to test the model error numerically on the test set. It can be seen from Table 1 that the MSE of the AWA+ESPCN model is 3.1327, and the RMSE is 1.4603, which has the smallest error among all the schemes, followed by scheme 1 and scheme 2, MSE is 4.1733 and 3.5549 respectively. This shows that the spatialization result of scheme 3 has the highest credibility.

This paper uses the spatial residual distribution to test the spatial error of the model (Figure 5). In the residual distribution diagram, a positive residual value (red) represents the model's overestimation of the population, and a negative residual value (blue) represents the model's underestimation of the population. In general, the three types have good downscaling learning ability, and the residual value is mostly less than 1 person. Through further research on the residuals, it is found that: (1) The residual interval of Option 3 is (-90—100.0), the residual interval of Option 2 is (-95.3—104), and the residual interval of Option 1 is (-102.2—115.1), so the residual deviation of scheme 3 is smaller, indicating that the experimental results generated based on scheme 3 sub-pixel convolution learning will be more accurate; (2) the prediction value of the three schemes will be higher in the main urban area, which may be this it is related to the night light data, and there will be a small prediction value in the water...
distribution area. It may be related to the water data. The existence of driving factors can help us better simulate the spatialization of the population, but in some areas related to the land type it will also cause certain deviations to the forecast data. (3) Most of the population errors in Option 3 are distributed in -0.9—1, while in Option 1 and Option 2, except for some areas, the residual error is distributed in -0.9—1, and a considerable part of the regional errors are distributed in 1.1—10. Among them, the learning of local features in Scheme 3 is better than Scheme 1 and Scheme 2. (4) Compared with Option 1, there are more underestimated areas in the northeastern part of Chongqing with more water bodies. This may be related to the smaller filter and smaller convolution. Also in the southeast, there are fewer 1.1—10 interval, the picture generated in this area becomes more accurate. (5) The underestimated area of Option 3 will be much less than that of Option 1 and 2, which is mainly affected by the driving factors of the river, but is much less affected by the water body.

Table 1. Comparison of population spatialization model errors in different schemes.

| Scheme Number | Model Name     | MSE  | RMSE |
|---------------|----------------|------|------|
| Scheme 1      | AWA+SRCNN      | 4.1733 | 2.0429 |
| Scheme 2      | AWA+FSRCNN     | 3.5549 | 1.8854 |
| Scheme 3      | AWA+ESPCN      | 2.1327 | 1.4603 |

Figure 3. Comparison Chart of Spatial Residual Distribution of Population Spatialization Results of Each Scheme.

5. Conclusion and discussion

In this paper, the area-weighted population division rasterization "removing roughness" and the ESPCN "refining" are coupled to construct a population spatialization model to realize the integrated learning of the global and local characteristics of the population distribution in the study area. Three experimental schemes are formed by coupling the area-weighted partition rasterization method with different super-resolution models, and the accuracy verification and comparative analysis of the results of each scheme are carried out. The main conclusions are as follows:
(1) Under the combined effect of natural environmental factors and socio-economic factors, the AWA+ESPCN model coupling can compensate to a certain extent for the shortcomings of a single model in learning global features and generalizing local features. The ESPCN sub-pixel convolutional neural network has the best down-scaling learning ability, and the spatialized result is closer to the label data and closer to the true population distribution.

(2) The mean square error shows that the AWA+ESPCN coupling model has the highest simulation accuracy, and the space residual value interval error is the smallest. The residuals in most areas of Chongqing are distributed between -1 and 1, the results are more real and reliable.

(3) Compared with Option 1, Option 2 uses a compact hourglass CNN structure, with smaller convolution and fewer filters, so it can learn local features better, so there are fewer 1.1-10 in the southeast. In the interval, the pictures generated in this area become more accurate.

(4) The area-weighted average rasterized population learns the global population characteristics well. Although the driving factor as auxiliary data can improve the accuracy of the training model and reduce the residual size as a whole, but for the part of Yuzhong District where the density is particularly high, the lighting will As a result, the valuation is too large, and in some areas where water bodies are located, there will be too small valuations.

References
[1] Chen J, Xu C Y, Guo S L, et al. Research progress and challenges of statistical downscaling methods [J]. Water Resources Research, 2016, 5 (04): 299-313.
[2] Tatem A J, Noor A M, Von Hagen C, et al. High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa [J]. Plos One, 2007, 2 (12).
[3] Feng J, Zhou Y X. Population growth and distribution in Beijing metropolitan area in the past 20 years [J]. Acta Geographica Sinica, 2003 (06): 903-916.
[4] Tripathy B R, Tiwari V, Pandey V, et al. Estimation of Urban Population Dynamics Using DMSP-OLS Night-Time Lights Time Series Sensors Data [J]. Ieee Sensors Journal, 2017, 17 (4): 1013-1020.
[5] Liu Y J, Wang J Y, Wang Z W. Simulation of urban and rural permanent population pattern in Huanghuihai area based on NPP/VIIRS night light data [J]. Regional Research and Development, 2019, 38 (03): 176-180.
[6] Stevens F R, Gaughan A E, Linard C, et al. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data [J]. Plos One, 2015, 10 (2).
[7] Sinha P, Gaughan A E, Stevens F R, et al. Assessing the spatial sensitivity of a random forest model: Application in gridded population modeling [J]. Computers Environment and Urban Systems, 2019, 75: 132-145.
[8] Deng S Q. Research on Population Spatial Distribution Model Based on Random Forest Algorithm and Multi-source Data [D]. East China Normal University, 2018.
[9] Bakillah M, Liang S, Mobasher A, et al. Fine-resolution population mapping using OpenStreetMap points-of-interest [J]. International Journal of Geographical Information Science, 2014, 28 (9): 1940-1963.
[10] Wang Z, Chen J, Hoi S C H. Deep Learning for Image Super-resolution: A Survey [J]. IEEE transactions on pattern analysis and machine intelligence, 2020, PP.
[11] Zong Z, Feng J, Liu K, et al. DeepDPM: Dynamic Population Mapping via Deep Neural Network [M]. 2019: 1294-1301.
[12] Vandal T, Kodra E, Ganguly S, et al. DeepSD: Generating High Resolution Climate Change Projections through Single Image Super-Resolution [M]. 2017: 1663-1672.
[13] Chao D, Chen C L, He K, et al. Learning a Deep Convolutional Network for Image Super-Resolution [J], 2014.
[14] Dong C, Loy C C, Tang X: Accelerating the Super-Resolution Convolutional Neural Network, Leibe B, Matas J, Sebe N, Welling M, editor, Computer Vision - Eccv 2016, Pt Ii, 2016: 391-407.
[15] Shi W, Caballero J, Huszar F, et al.: Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network, 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1874-1883.
[16] Wang L, Chen L. Spatiotemporal dataset on Chinese population distribution and its driving factors from 1949 to 2013 [J]. Scientific Data, 2016, 3 (3): 160047.