

Article

Large-Scale Automatic Identification of Industrial Vacant Land

Yihao Sun ¹, Han Hu ², Yawen Han ¹ , Ziyang Wang ¹ and Xiaodi Zheng ^{1,3,*} 

¹ Department of Landscape Architecture, School of Architecture, Tsinghua University, Beijing 100084, China; sunyh20@mails.tsinghua.edu.cn (Y.S.); hanyw18@tsinghua.org.cn (Y.H.); ziyangw@mail.tsinghua.edu.cn (Z.W.)

² School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906, USA; huhan.opt@gmail.com

³ Key Laboratory of Eco Planning & Green Building (Tsinghua University), Ministry of Education, Beijing 100084, China

* Correspondence: xzheng@tsinghua.edu.cn

Abstract: Many cities worldwide have large amounts of industrial vacant land (IVL) due to development and transformation, posing a growing problem. However, the large-scale identification of IVL is hindered by obstacles such as high cost, high variability, and closed-source data. Moreover, it is difficult to distinguish industrial vacant land from operational industrial land based solely upon image features. To address these issues, we propose a method for the large-scale automatic identification of IVL. The framework uses deep learning to train remote-sensing images of potential industrial vacant land to generate a semantic segmentation model and further use population density and surface temperature data to filter model predictions. The feasibility of the proposed methodology was validated through a case study in Tangshan City, Hebei Province, China. The study indicates two major conclusions: (1) The proposed IVL identification framework can efficiently generate industrial vacant land mapping. (2) HRNet exhibits the highest accuracy and strongest robustness after training compared with other semantic segmentation backbone networks, ensuring high-quality performance and stability, as evidenced by a model accuracy of 97.84%. Based on the above advantages, the identification framework provides a reference method for various countries and regions to identify industrial vacant land on a large scale, which is of great significance for advancing the research and transformation of industrial vacant land worldwide.

Keywords: industrial vacant land; land use; satellite images; multi-source data; deep learning



Citation: Sun, Y.; Hu, H.; Han, Y.; Wang, Z.; Zheng, X. Large-Scale Automatic Identification of Industrial Vacant Land. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 409. <https://doi.org/10.3390/ijgi12100409>

Academic Editors: Wolfgang Kainz, Lili Jiang, Di Zhu and An Zhang

Received: 4 July 2023

Revised: 14 September 2023

Accepted: 28 September 2023

Published: 5 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The effective transformation of urban vacant land (UVL) has broad prospects and significant advantages for improving urban economy, ecological environment, transformational development, and urban planning [1–5]. The term UVL is broad and diverse but is usually defined as under-utilized lands including bare soil, derelict land, abandon buildings and structures, brownfields, greenfields, uncultivated land or marginal agricultural land, and recently razed land [6]. With the advent of the post-industrial era, some cities have begun to restructure their economy and transform their industrial areas. As a result, industrial areas in many cities have closed down, creating large amounts of industrial vacant land (IVL) [7–9]. Some examples include the “rust-belt” and Youngstown in the USA, and the Ruhr in Germany [10,11]. Urban development and expansion have often revolved around industrial land, particularly in cities that sprang up as a result of industrial production. Areas of IVL are widely distributed in cities, particularly in city centers, occupying prime urban locations, triggering environmental degradation, economic decline, and social exclusion; thus, many cities are confronted with the pressing need for revitalization of their older town centers [12], posing a serious challenge to urban planning [7–9]. These points demonstrate that IVL, as a category of UVL, is a valuable land resource and its reuse can effectively undermine urban sprawl and promote sustainable

urban development [13–17]. Many IVL areas are regenerated into new sites that integrate new functions into the urban system, such as real estate that provides economic value and urban green spaces that provide ecosystem services for inhabitants. For example, in the 1960s, the decadent Ruhr industrial district in Germany underwent a comprehensive renovation in terms of industrial restructuring and the optimization of the ecological environment. By changing the function of the original buildings, facilities, and sites, the history of the industrial area was effectively recreated while providing green spaces for culture and recreational life. There are many opportunities to redevelop vacant land in terms of ecological and social value [1]. Identifying IVL is, therefore, a prerequisite for its reuse to improve urban spatial patterns. However, IVL identification is challenged by the low efficiency and accuracy of previous IVL identification methods and other hindrances related to limited land-use data from some areas that are not yet open source. Certain researchers assume that vacant land exhibits certain patterns, such as the presence of buildings, vacant parking lots, and sparse vegetation. They employ object-based classification techniques using various data sources, including Ikonos, QuickBird satellite imagery [18], and 3D laser scanning data [19]. However, these studies face challenges with low accuracy in distinguishing between bare soil/land, built-up areas, and vacant land [20]. These studies primarily rely on image classification alone to identify vacant land. In reality, the term “vacant land” encompasses specific information about land use that is influenced by human activity. Image classification alone cannot provide such information; it needs to be complemented with additional data sources [20]. Furthermore, commercial data is used due to its high spatial resolution, such as QuickBird images with a resolution of 0.6 m. Despite the possible detection of specific types of vacant land, the cost of the data is not negligible. Consequently, the development of an effective and rigorous identification method is therefore crucial to strategizing IVL reuse and urban planning.

There are currently two main methods for identifying UVL and IVL. One is the traditional manual identification method, which includes field research, questionnaires, visual identification using satellite images, and marking areas using cadastral data. This approach requires a large amount of manpower and time. As a consequence, the literature mostly contains small-scale studies of urban areas. Li et al. (2018) used high-resolution satellite images of Shanghai from 2000, 2005, and 2010 to identify the UVL distribution during each time period using manual visual interpretation [21]. Similar methods were used in case studies in Guangzhou [22], Changchun [23], and Atibaia [24]. Wang et al. [25] used remote-sensing (RS) imagery combined with field research and questionnaires to identify UVL in Wuhu. Song et al. [26] combined manual field surveys with open-source data (e.g., maps, Baidu heat maps, business data, and Baidu open street maps) to quantify the UVL of the Changchun City center. Despite the significance of this work, most previous studies have faced limitations. For example, they are (1) expensive in terms of labor and time, (2) provide uncertain identification results due to subjective differences in the identification criteria between auditors, and (3) result in research that is difficult to expand to regional and national scales owing to limited data availability and labor time constraints.

Two methods have gained widespread attention in recent years that offer substantial savings in labor and time costs while identifying UVL and IVL on a large scale and have a high measurement accuracy. These methods include (1) machine-learning-based artificial intelligence recognition and (2) big-data-based data fusion recognition. Xu and Ehlers [20] established a rule-based data fusion framework that integrates RS images, geographical information system (GIS) layers, and citizen science data using datasets such as Urban Atlas (UA), impermeable subsurface data, land cover (LC), government data, OpenStreetMap, Wiki data, and social media data. They proposed that different datasets should be chosen for different UVL types to perform data fusion and identification, and they applied the method to identify UVL in 63 regions in Germany. Mao et al. [27] developed a large-scale automatic UVL identification framework using semantic segmentation of high-resolution remote-sensing (HRRS) images and city stratification, applied the framework in a case

study of 36 major Chinese cities, and obtained accurate and efficient UVL identification results. These studies have important implications for the field.

With the rapid development of deep-learning techniques in computer vision, semantic segmentation is becoming increasingly used in the field of the built environment and RS image recognition, including FCL [28,29], UNet [30], SegNet [29,31], and DeepLab [32,33], which are potential tools for achieving large-scale and automatic UVL identification. Most studies have focused on automatic UVL identification based on satellite images. However, little work has been conducted on automatic IVL identification because the RS images of IVL present a variety of visual characteristics, and their “vacant” or “idle” status cannot be judged by satellite images alone. Inconsistent identification and determination criteria have also been used. Thus, it has so far been difficult to establish an accurate and efficient automatic IVL identification framework.

To confront these problems, we aim to put forth a comprehensive methodology and framework for identifying large-scale IVL utilizing open-source data. This proposed methodology can be readily adapted and implemented in different countries and regions, offering a novel approach for IVL identification and analysis. Specifically, we propose a method to identify IVL based on satellite RS images and multi-source data. The main goal of this study is to achieve the automatic identification of potential IVL based on semantic segmentation of Sentinel-2A RS images, and to determine the true IVL characteristics by removing still-operational industrial sites based on population density and surface temperature inversion data. We compared three semantic segmentation frameworks—HRNet, ResNet-UNet, and VGG16-UNet—and then built a potential IVL semantic segmentation model based on the best-performing HRNet and a filtering mechanism using human density and surface temperature. We validated the whole identification workflow in Tangshan City, Hebei Province, China. The experiments show that the identification model based on HRNet is accurate, stable, efficient, and robust, and the filtering mechanism can cover most of the research objectives. This identification framework provides a reference method for various countries and regions to identify IVL on a large scale, which is of great significance for advancing IVL and UVL research and practice worldwide.

2. Study Area and Data

Three resource-based cities in China were investigated to study a large-scale automatic detection framework of IVL (Figure 1). Resource-based cities represent a common urban typology within China’s urban system and are characterized by a predominant reliance on the extraction and processing of nonrenewable resources (e.g., minerals, lumber, oil) within their respective regions. Such cities are often accompanied by a considerable amount of industrial infrastructure and industrial land. A total of 262 resource-based cities were identified in accordance with the National Sustainable Development Plan for Resource-Based Cities (2013–2020) published by the State Council in 2013, which were further categorized into four distinct types: growing (31), mature (141), declining (67), and regenerating (23). The industrial systems of resource-declining and resource-regenerating cities are often well-established and distributed in different provinces, with diverse core industrial types representing the industrial construction characteristics under different regional features and urban morphology backgrounds in China. For many resource-regenerating cities, some IVL may be reused and regenerated to reassume industrial or other functions such as shopping areas and industrial culture parks; however, they may retain specific industrial structures, making it difficult to distinguish operational industrial land from genuinely idle IVL simply by observing RS image characteristics, thereby increasing the complexity and challenges associated with automatic IVL detection.

We aim to make this automatic detection method broadly applicable and highly feasible. Therefore, we selected potential IVL from three declining cities (Xinyu, Wuhai, and Huangshi) and one regeneration city (Tangshan) as models for creating the semantic segmentation training dataset. The central urban area of Tangshan was selected as the verification case for the entire detection method. The four cities used for dataset creation are

located in three geographical regions in the west (Wuhai), middle (Xinyu and Huangshi), and east (Tangshan) of China. The diversity of surface features in the dataset was considered with respect to the various industrial systems established around iron mines, coal mines, and quarries, as well as the diversity of land-use characteristics in different cities. These characteristics are influenced by various factors such as different climatic conditions, urban culture, and planning and construction styles.

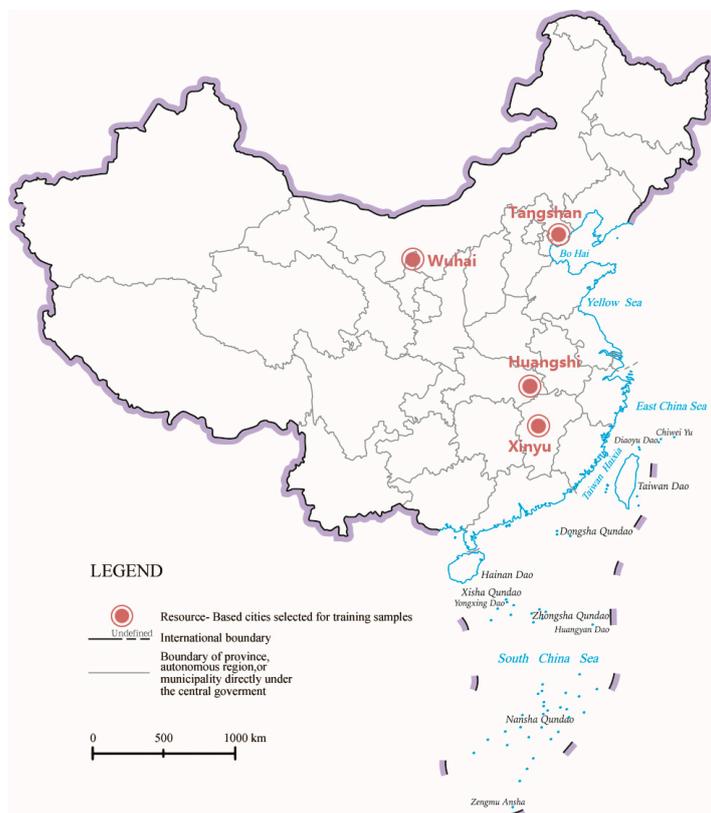


Figure 1. Locations of the four resource-based cities investigated here to study the large-scale automatic detection framework of IVL.

We collected Sentinel-2A RS images of the three cities from the European Space Agency (<https://scihub.copernicus.eu/dhus/#/home>, accessed on 11 November 2021). The images were synthesized from the red-green-blue channels to form false-color images with a pixel resolution of 10 m. Additionally, utilizing the bands 11 and 12 provided by the SWIR sensor aboard Sentinel-2A, we were able to ascertain the environmental temperature conditions at the target locations. This holds significant relevance in discerning the operational status of the land. These images are the highest-quality open-source and free RS data available. In contrast to lower-level L1C products that only undergo spatial correction, L2A-level images undergo atmospheric correction to obtain more accurate surface parameters. The data were collected from October 2020 to January 2021. The following criteria were followed for image selection to ensure the feasibility and data quality used in our method, including (1) low environmental temperature to ensure the exclusion of industrial sites that are still in operation using surface temperature data, and (2) cloud interference controlled at a very low level to allow clear analysis of land-use and land-cover features.

We also collected Landsat 8 RS images (<https://earthexplorer.usgs.gov/>, accessed on 23 December 2021) of the central urban area of Tangshan City, the validation target. The Band-10 data from the TIRS sensor were subjected to radiometric calibration. We computed the land surface temperature using the radiative transfer equation (RTE). Following this, we performed a merging process through the Environment for Visualizing Images (ENVI) platform's pansharpener technique. This involved integrating Band 2 (blue), Band 3

(green), and Band 4 (red) with a resolution of 30 m, along with Band 8 (pan) with a resolution of 15 m. This merging process provided a clear observation of the reasonableness of the inverted surface temperature raster data. Simultaneously, we obtained the Baidu heat map of the validation target during working hours from Baidu Map (<https://map.baidu.com/>, accessed on 12 April 2022) to calculate the population density in the city. The Baidu heat map can be converted into vector points representing specific population densities through vector calculations. The main calculation process involves using vectorization operations in ArcGIS to estimate the population distribution proportions provided by the Baidu Heat Map. This is combined with the total urban population to perform proportional sampling, thereby deriving the population density.

3. Methods

The entire automatic detection method includes two main stages: (1) predictive recognition of potential IVL through semantic segmentation models, followed by (2) IVL filtering using population density data and land surface temperature data. To ensure the accuracy of the semantic segmentation model and the feasibility of the subsequent filtering process, we developed a four-step framework (Figure 2) as follows: (1) data labeling, (2) model training, (3) potential IVL prediction, and (4) multi-source data filtering. The goal and contents of each step are described below.

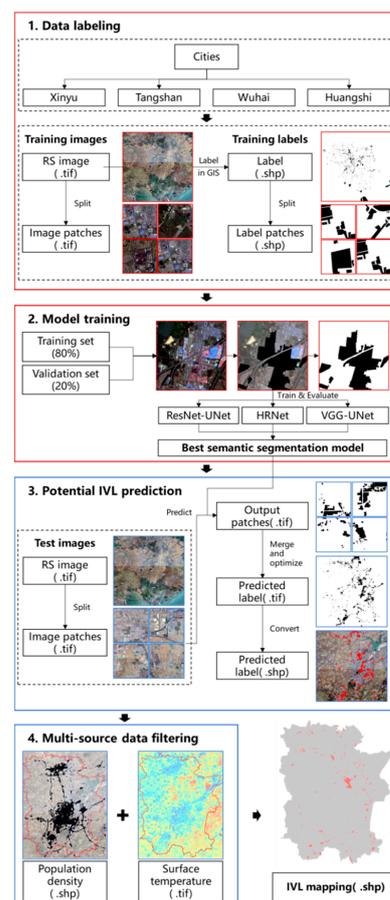


Figure 2. Large-scale automatic detection framework of IVL. The red boxes refer to model establishment and the blue boxes refer to prediction and filtering.

3.1. Data Labeling

The purpose of this step is to obtain high-quality training data for the semantic segmentation model of potential IVL. Labels were created for potential IVL with industrial surface features in the three investigated cities. To ensure effective training results, these

labels were required to meet the following three conditions: (1) cover a large number of potential IVL with different surface features, (2) ensure a wide range of industrial and non-industrial surface changes to enhance the robustness of the trained model, and (3) ensure a sufficient number of samples to control overfitting.

The dataset for potential IVL was accordingly created from RS images and potential IVL labeled by professional auditors. It should be noted that the surface features of IVL may vary owing to different industrial production tasks before the sites were abandoned (Table 1). The potential IVL in the four representative cities exhibits a wide range of these features, ensuring a selection principle of multiple types of training samples within the same category.

Table 1. Common and different aspects of the three categories of potential IVL: (a), (b), and (c) are the common surface features of the three categories; and (d) and (e) are the unique features presented by each category. We employed high-resolution imagery for effective visualization, and the images within the table were sourced from Google Earth Pro.

Category	Surface Features	Representative Remote-Sensing Image
All in common	<ul style="list-style-type: none"> (a) Boundaries are clear and generally demarcated by walls and roads. (b) Flat terrain. (c) Internal transportation system connects the entire venue. 	
Raw material processing and manufacturing sites	<ul style="list-style-type: none"> (d) Features buildings with distinct industrial characteristics, such as blast furnaces, chimneys, and pipelines at steel plants. (e) May have identifying colors, such as grayish-white for cement plants, brownish for steel plants, and jet-black for coking plants. 	
General manufacturing sites	<ul style="list-style-type: none"> (d) Has raw material and miscellaneous item storage but does not have large smelting. (e) Has large areas containing sheds with blue or red steel roofs or solar panels. 	
Infrastructure and warehouse	<ul style="list-style-type: none"> (d) May contain railway tracks, numerous containers, storage tanks, and a barge wharf on a riverbank. 	

After labeling the potential IVL of the three cities, we used label vector data as the reference extent. Subsequently, we employed the Split Raster tool within the GIS platform to segment the IVL label vector data and RS images in ArcGIS into 256×256 pixel rasters. Within each label raster, grid cells featuring industrial land surfaces were designated with a value of 1, while urban background grid cells received a value of 0. This approach not only ensured the effective utilization of data by accommodating partial overlaps at edges but also maintained a coherent correspondence between the label raster and RS image remote-sensing imagery. A total of 1246 patches were generated. The segmented dataset was randomly divided into a training set and a testing set with an 8:2 ratio, for the subsequent model training.

3.2. Model Training

Let us suppose that $L = \{(x_i, y_i)\}$ is a labeled training image, where x_i is the i th pixel in that image. Each sample pixel, $x_i \in \chi$, in the image is assigned with a ground truth label $y_i \in \{0, 1\}$, where $y_i = 0$ indicates that the pixel is the background and $y_i = 1$ indicates that the pixel belongs to the industrial surface class. The goal of our semantic segmentation task is to obtain a segmentation model $f_\theta : \chi \rightarrow \{0, 1\}$ from the labeled training dataset, where $\theta \in \Theta$ is the learned model parameter.

Many existing segmentation networks, such as UNet, were used to convert the entire image into a low-resolution visual representation before upsampling it back to the original high-resolution size to achieve pixel-level semantic segmentation. However, some positional information may become mismatched during the upsampling process in these networks. Our semantic segmentation task is highly sensitive to position. To address the issue of inaccurate segmentation in these networks, we used the HRNet [34] network, which preserves high-resolution visual representations during the forward propagation process, thus effectively retaining positional information.

Figure 3 provides an overview of the HRNet model, which starts with a high-resolution convolutional stream and progressively incorporates additional convolutional streams of decreasing resolution, connecting these multi-resolution streams in parallel. The segmentation result is generated by concatenating the representations from multi-resolution convolutional streams.

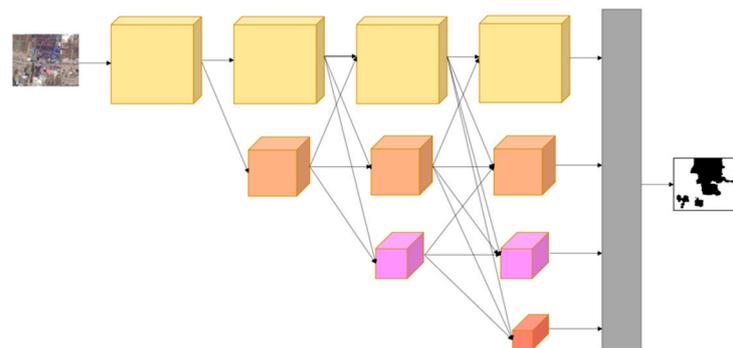


Figure 3. Overview of the HRNet model.

This HRNet model was trained from scratch without using any pretrained models. The loss function used to train the HRNet is binary cross entropy loss, which is commonly used in classification tasks. We used the Adam algorithm with an initial learning rate of 10^{-2} . The training epoch was set to 50.

3.3. Potential IVL Prediction

The purpose of this step is to obtain high-quality potential IVL prediction results in the form of vector data. We used the following methods to accomplish this goal and ensure the quality of the final results: (1) During model training, we installed the GDAL library for Python, enabling the use of TIFF data for both image input during model training

and the segmentation results of the model output. (2) We found fine points on both the edges and interior of the images. To reduce the amount of potential error when converting raster to vector files, we employed the ArcScan tool for raster refinement procedures. A morphological closing operation was performed on the predicted IVL raster. This operation effectively smooths out coarse boundaries within the raster while also appropriately filling small gaps between foreground objects, aiming to yield a raster with distinct boundaries and comprehensive internal filling. This practice is also advantageous for subsequent conversion into vector files, contributing to the creation of more precise potential IVL polygon vectors (Figure 4). Regarding the transformed outcomes, it was observed that, despite the morphology operations, certain independent and significantly small portions persisted within the vectorized polygons. Upon meticulous observation, statistical analysis, and considering the subsequent research's data requisites, independent polygons with an area of less than 400 m² were identified as noise and subsequently removed.

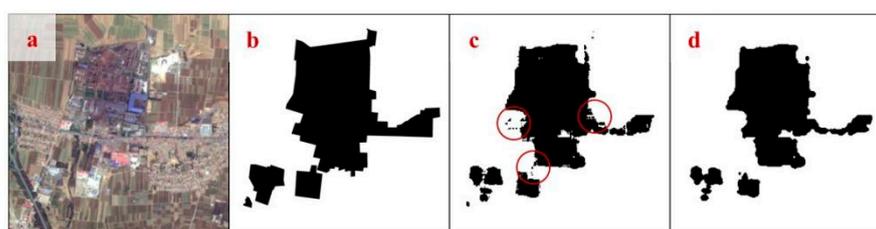


Figure 4. Comparison between optimized and non-optimized results: (a) original image; (b) ground truth; (c) non-optimized result with spots (red circles) on the edges; (d) optimized result with clear edges.

The process described of this section is shown in Figure 2: (1) Train a semantic segmentation dataset with the best-performing model architecture to obtain the prediction model and output the potential IVL prediction result raster tail. (2) Combine these tails into a complete raster. (3) Use image morphology operations, such as opening and closing in ArcScan, to optimize the raster. (4) Convert TIFF to SHP files. (5) Delete small-area polygons deemed as noise after area calculation.

3.4. Extraction of IVL

The potential IVL results obtained in the previous stage through semantic segmentation prediction consist of two parts: (1) industrial land that is still in use and (2) IVL. By analyzing the usage characteristics of these two types of land, we found that IVL is usually unoccupied or has very few people present during working hours. Additionally, the thermal radiation exhibited by buildings or structures within IVL sites is also often significantly lower than that of industrial land that is still in use, because heat is frequently generated during industrial processes. This phenomenon is more pronounced in northern China in winter owing to heating requirements in the cold climate. Two significant IVL characteristics are therefore sparse population density and low surface temperature. Considering that industrial production may require only a small number of on-site personnel, owing to a high degree of automation, the IVL identification criteria were set as follows. The population density of IVL is significantly lower than surrounding areas, and the maximum surface temperature within the land plot is also significantly lower than the average surface temperature of industrial areas that are currently operational.

At this stage, we utilized population density data and surface temperature data as the foundational elements for our analysis. Population density data were used for the initial filtering, followed by surface temperature data for further filtering. This systematic approach allowed us to separate operational industrial land from potential IVL, thus obtaining the recognition results for IVL. The filtering logic is shown in Figure 5.

The central urban area of Tangshan is taken as an example in this study to determine the target recognition area. The data used in this case study were from a statutory workday

(Friday, 18 December 2020). This specific day was chosen to ensure that the population distribution data were not affected by weekend or holiday features, which could lead to an underestimation of the true population distribution on active industrial land, thereby resulting in their misclassification as IVL. Conducting this study in winter also ensures that operating industrial sites can be distinguished from IVL based on their significantly higher surface temperatures. This distinction arises because heat radiation is not emitted from IVL owing to the absence of on-site production and heating.

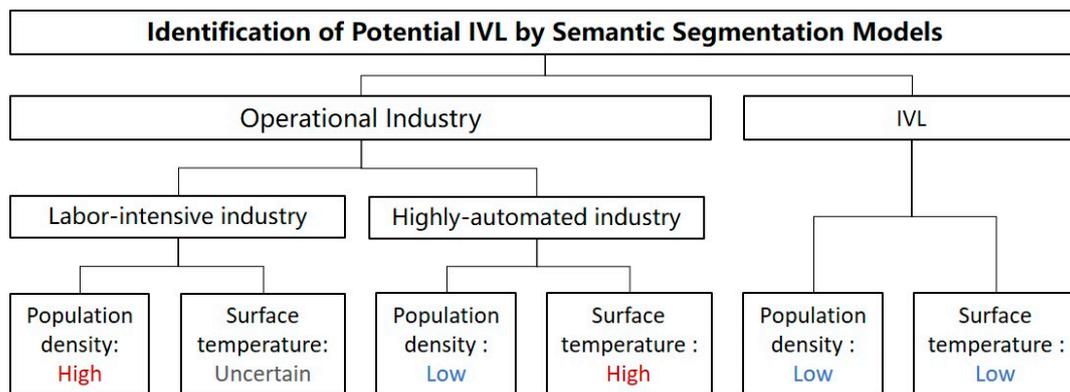


Figure 5. Logic of multi-source data filtering.

Due to the restricted accuracy of RS images, the semantic segmentation prediction results struggle to accurately separate large potential IVL in accordance with the city road configuration. Factory or enterprise areas are rarely separated by urban expressways or highways. To prevent scenarios where a large piece of land is mistakenly identified as IVL due to an abnormally high population density or surface temperature during the subsequent filtering process, this section uses vector data of Tangshan's major transportation network (including railways, highways, expressways, and main roads). This transportation network data are employed to separate the area identified as potential IVL (Figure 6).

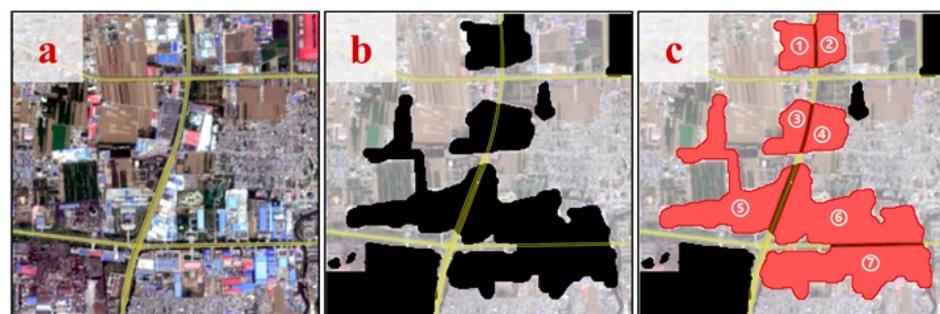


Figure 6. Potential IVL after road division: (a) original image; (b) prediction results before division; (c) optimized result based on a typical road network layout, the extensive potential IVL has been divided into 7 independent parcels.

The population density data obtained from vectorizing the Baidu heat map were imported into GIS (Figure 7). The vectorization process of Baidu heat maps involves the following procedural steps, including (1) reclassifying the thermal raster data acquired from the Baidu Maps web service to extract thermal values at varying levels; (2) converting the thermal raster image into pixel points and estimating population data by deducing RGB values corresponding to the point positions in the heat map; and (3) subdividing the research area into 100×100 m grid cells, aggregating pixel-based population data within their respective grid cells, and further expanding the vectorized grid-based population using the urban population figures from the Seventh National Census Data provided by

the National Bureau of Statistics (<http://www.stats.gov.cn/sj/pcsj/rkpc/d7c/>, accessed on 25 May 2023), thus yielding population density data measured in people per hectare (people/hm²). After vectorization, each vector point in the layer possessed a POP value representing a population density within a one-hectare square centered on the vector point. If the population density was less than 100 people/hm², the POP value was recorded as zero and excluded from the vectorization calculation results. Statistical analysis showed that the maximum population density within the central urban area of Tangshan at that time was 2277 people/hm², the minimum was 247 people/hm², and the average was 888.01 people/hm². In China, some control indicators of industrial land (e.g., plot ratio, building coefficient) must comply with the Industrial Project Construction Land Control Indicators (Ministry of Natural Resources of the People's Republic of China 2021, http://gk.mnr.gov.cn/zc/zxgfwj/202102/t20210226_2615419.html, accessed on 25 May 2023). Under these requirements, the population density of labor-intensive industries is usually set to 120–140 people/hm². We therefore assumed a conservative scenario where the population density within IVL was less than 120 people/hm², and used this as a criterion to filter out a set of IVL and highly-automated industry land.

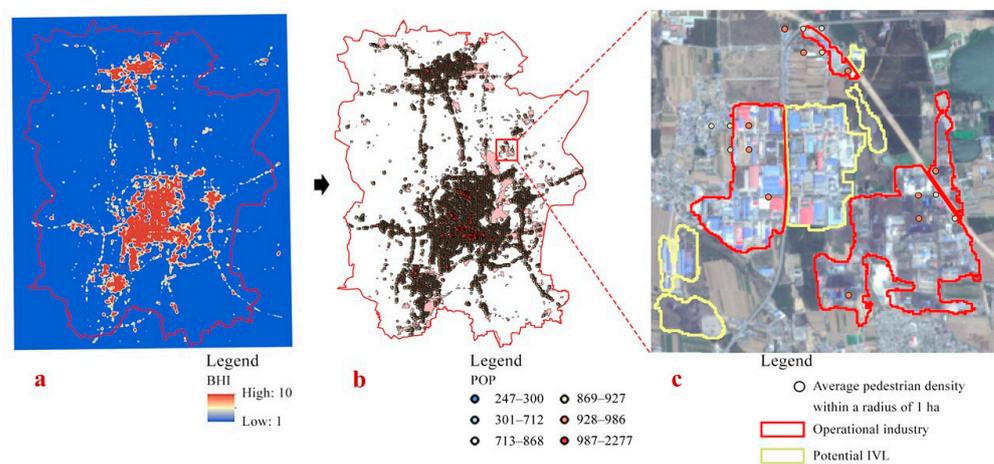


Figure 7. Filtering is first applied using population density data. (a) Performance on the client side of the Baidu heat map. (b) Population density vector point data after vectorization calculation. (c) Difference in population density between operational industry and potential IVL.

We utilized the spectral data from Landsat 8 images as the foundational basis, employing the radiative transfer equation (RTE) to perform the retrieval of regional land surface temperature [35]. This methodology primarily relies on the observation of surface thermal radiation Q by the satellite's infrared sensor (data encompassed by Band 10). Subsequently, atmospheric influences are subtracted, resulting in the determination of the intensity of surface thermal radiation. Finally, this intensity of thermal radiation undergoes conversion to ascertain the land surface temperature. The formulation of this method is presented as:

$$Q = [xD(t) + (1 - x)q_1]\alpha + q_2 \quad (1)$$

$$t = k_2 / \ln(k_1 / D(t) + 1) \quad (2)$$

In the given formulas, x represents the land surface emissivity, t denotes the land surface temperature, $D(t)$ signifies the blackbody radiance, and α signifies the atmospheric transmissivity in the thermal infrared domain. q_1 represents the atmospheric upwelling radiance, while q_2 denotes the atmospheric downwelling radiance. For distinct Landsat data, the values of k_1 and k_2 vary. For Band 10 of TIRS, k_1 is equal to $774.89 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1} \cdot \mu\text{m}^{-1}$, and k_2 equals 1321.08 K.

Following the aforementioned methodology, we obtain the surface temperature of the study area, which ranged from $-7.3 \text{ }^\circ\text{C}$ to $19.2 \text{ }^\circ\text{C}$. Statistical analysis showed that the lowest

surface temperature of operational industrial land with a human population density higher than 120 persons per hectare was 8 °C. Considering the winter research context, we inferred that sites with a surface temperature above 8 °C were indicative of thermal radiation generated by ongoing industrial production. We were consequently able to identify highly-automated industrial land that has a low human population density but is still engaged in industrial production (Figure 8). The contours of the buildings depicted in the figure were sourced from BIGEMAP (<http://www.bigemap.com/>, accessed on 28 May 2022), whereby the vector tools for buildings facilitate the provision of real-time vector data derived from the data source of Gaode Map (<https://www.amap.com/>, accessed on 28 May 2022). This serves as a means to examine and interpret extensive temperature anomalies, thereby aiding in the assessment of the logicity of such deviations, and additionally contributing to the assessment of the operational status of the sites where the buildings are situated.

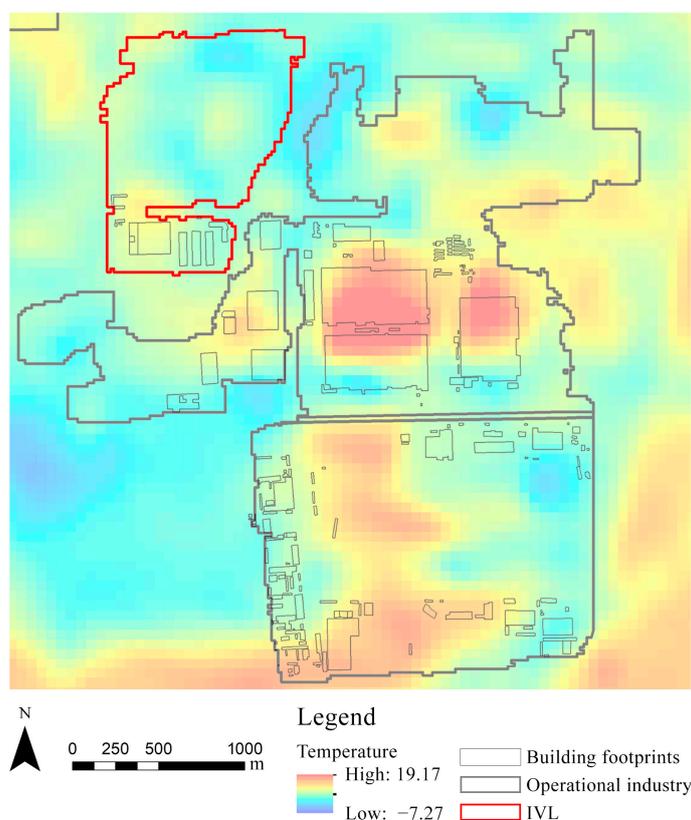


Figure 8. Difference in land surface temperature between land with operational industries and IVL.

4. Results

4.1. Overview of Extracted Sites

After completing the entire workflow, we mapped the industrial land still in operation and IVL in the central urban area of Tangshan City, as shown in Figure 9. The number, area, and geographic distribution of IVL, as well as the vacancy rate of urban industrial land were obtained from the results as expected. We identified 592 potential IVL areas in the central urban area of Tangshan. After screening with surface temperature and population distribution data, 431 IVL areas were extracted, covering an area of approximately 24.2 km². The results show that approximately 34% of the total industrial land in the central urban area of Tangshan is vacant.

Figure 10 shows a boxplot analysis of the area of the two land types. Although the number of IVL fields is significantly greater than that of operational industries, the maximum, minimum, upper quartile, lower quartile, and median areas of IVL are significantly lower (Figure 10). It can thus be concluded that as urban industrial transformation advances, small factories in Tangshan City are more likely to shut down than large factories,

resulting in a large amount of small IVL scattered throughout the city. The reasons for this phenomenon may include a single production business, lack of capital and assets leading to difficulties in capital turnover, poor product quality, and/or inadequate employee welfare [36,37].

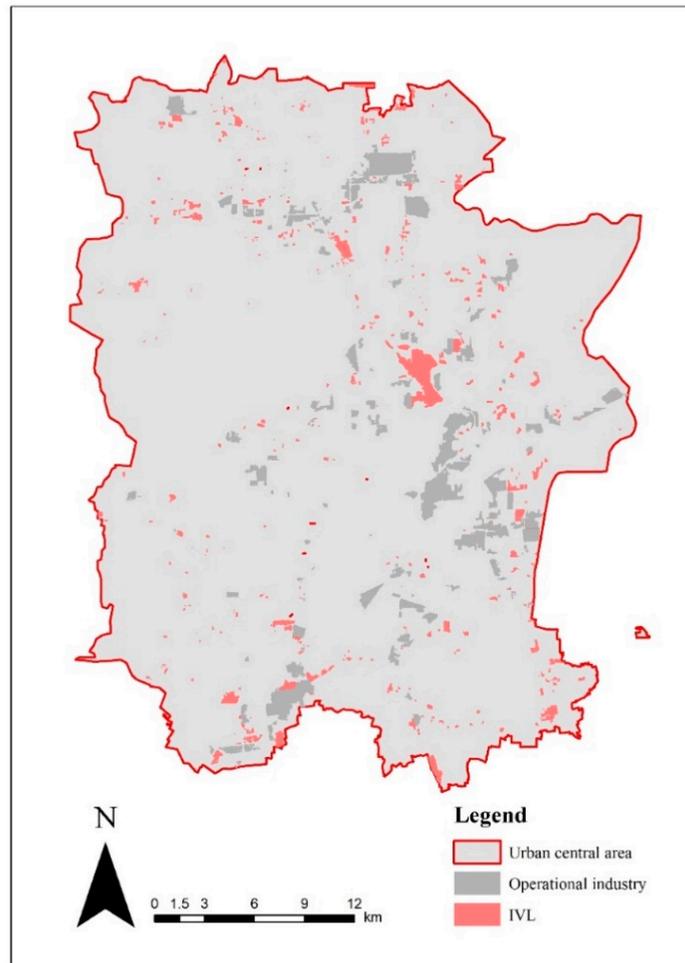


Figure 9. Mapping of IVL in the central urban area of Tangshan.

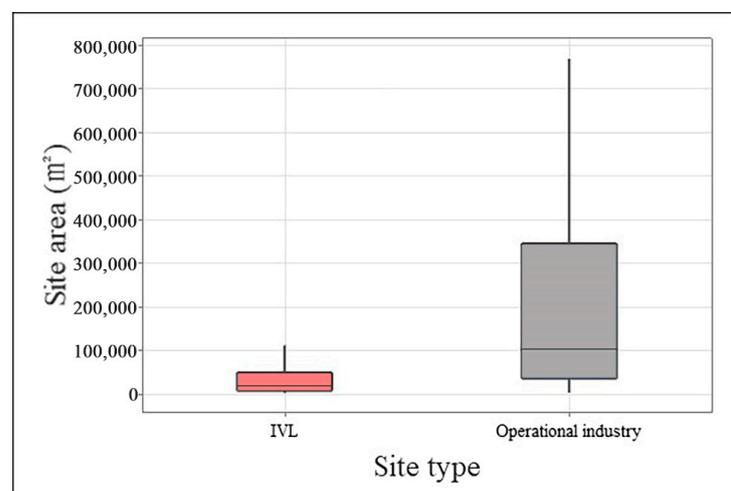


Figure 10. Site area differences between IVL and operational industries.

4.2. Model Evaluation

The ground-truth label for the test data is provided; thus, we employed three measures to evaluate the performance of the model: precision, recall, and F1-scores. The calculation formulas of the three metrics are as follows.

$$precision = \frac{TP}{TP + FP} \quad (3)$$

$$recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - score = \frac{2 \textit{precision} * \textit{recall}}{\textit{precision} + \textit{recall}} \quad (5)$$

where TP and FN represent the number of pixels belonging to the industrial surface that are correctly and incorrectly classified, respectively, and FP represents the number of pixels that belong to the background and are incorrectly classified. *Precision* represents the percentage of the correctly classified number of pixels in the predicted results, *recall* represents the percentage of the number of pixels classified in the ground truth, and *F1-score* refers to a tradeoff measure between the precision and recall. Therefore, one goal of a good classifier is to maximize the F1-score to a value of one. The HRNet evaluation results from the testing data are listed in Table 2.

Table 2. Performance evaluation of HRNet.

	Precision	Recall	F1-Score
HRNet	0.7782	0.7501	0.7640

4.3. Comparison Study

4.3.1. Comparison with Other Models

We compared the performance of HRNet to the model structure from two previous studies to examine the effect of the model structure. These methods include ResNet-UNet [38] and VGG16-UNet [39], which are commonly applied for semantic segmentation. The ResNet-UNet architecture uses the first four layers of ResNet50 for downsampling and replaces the transposed convolution using Pixel Shuffle in the upsampling. The VGG16-UNet architecture uses the UNet structure and uses the VGG-16 network as the encoder. The main difference between these two networks and our HRNet is that the previous methods first downsample the input image as a low-resolution representation, whereas our model maintains high-resolution representation throughout the entire forward process.

For a fair comparison, we trained each model from scratch for 50 epochs without pretraining. The initial learning rate was set to 10^{-2} and the Adam algorithm was used. The results are shown in Table 3, demonstrating that HRNet significantly outperformed the others in F1-score.

Table 3. Comparison of the three evaluation metrics in the three models.

	Precision	Recall	F1-Score
HRNet	0.7782	0.7501	0.7640
ResNet-UNet	1.0	0.4829	0.6513
VGG16-UNet	1.0	0.4829	0.6513

Figure 11 demonstrates that the accuracy of HRNet continues to increase and reaches nearly 1 at the end of training, whereas the accuracies of ResNet-UNet and VGG16-UNet increase over the first few epochs but then remain essentially constant at an accuracy of approximately 0.93. It is notable that the value of accuracy is much higher than precision and recall. This may be because the background pixels dominate the dataset, making it

easier for the model to learn the features of the background and predict most pixels as background, resulting in a higher accuracy rate.

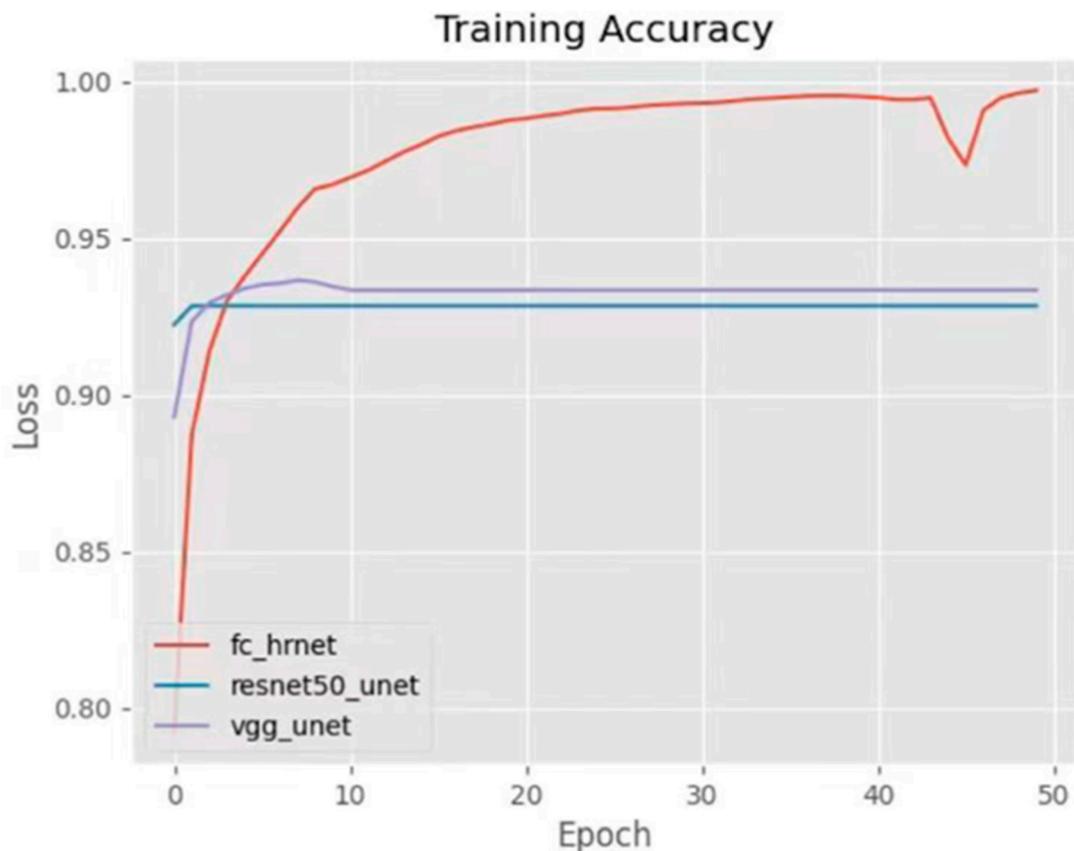


Figure 11. Comparison of accuracy curve of the three models.

Table 3 compares the performance of the three trained segmentation networks using the same training and testing datasets. When using ResNet-UNet and VGG16-UNet, a precision value of one and a low recall value indicate that these models classify almost all pixels as positive during segmentation, which suggests that these models are overfitting to the negative class. The F1-score provides a balanced evaluation of both precision and recall. HRNet shows the highest F1-score among all three models, which indicates that our proposed HRNet has an exceptional capacity to segment both positive and negative classes in a general sense.

4.3.2. Comparison with Traditional Identification Methods

Compared with traditional IVL identification methods, the innovation of this large-scale automatic detection framework lies primarily in the advantages demonstrated by the potential IVL semantic segmentation model. These include (1) the input data are highly accessible and can be used for more cities, covering a larger spatial range, (2) the method improves efficiency and reduces manpower, and (3) the conclusion demonstrates a significantly elevated accuracy comparable to manual recognition, while also maintaining a commendable level of stability and reliability and being unaffected by subjectivity owing to different auditors.

Traditional identification methods are heavily reliant on visual identification and resort to higher-resolution but non-open or commercially obtainable remote-sensing data, often entailing costs for annotators' convenience [27]. Moreover, outcomes from such visually driven conventional approaches typically pertain exclusively to specific regions for UVL identification. Consequently, traditional recognition methods impose stringent data prerequisites and exhibit limitations in terms of usage scenarios. Our innovation lies in

leveraging the superior performance of HRNet within the domain of image segmentation. By utilizing freely accessible and open-source Sentinel-2A data, we have achieved noteworthy image segmentation outcomes. Additionally, our approach demonstrates novelty through considerations of geographical diversity during the dataset construction phase, thereby rendering the model versatile across a wider array of usage scenarios.

In terms of efficiency comparison, we conducted an analysis by contrasting the total time required for the manual annotation of potential IVL across four cities with the combined time for model training and prediction. The outcomes demonstrate that the collective time for manually annotating potential IVL in these four cities amounts to approximately 50 h. Conversely, when leveraging the computational capabilities of one Nvidia RTX 3090 GPU within the automated framework for both training and prediction, the time required is approximately 4.5 h, representing an 11-fold enhancement in efficiency. Notably, if the computational power of the GPU employed for model training is augmented or if a multi-GPU configuration is employed, the efficiency of the automated framework would experience significant improvement. Moreover, when confronted with larger recognition areas, the efficiency advantages of the automated framework will become even more conspicuous.

We employed the intersection over union (IoU) value, which is the most widely adopted evaluation metric for the similarity between two image segmentation results, to contrast the differences in accuracy between the automated framework and traditional methods [40,41]. The equation for IoU value is as follows:

$$IoU(A, B) = \text{area}(A \cap B) / \text{area}(A \cup B) \quad (6)$$

During the dataset creation phase, the identification of potential IVL in the four cities was carried out by four professional auditors. In this phase of accuracy comparison and validation, we rotated the cities assigned to each of the four auditors. Subsequently, another round of visual identification for IVL was conducted across the four cities. The IoU value was calculated as a metric for benchmarking the accuracy of manual annotation, resulting in a value of 73.28%. Contrasting this outcome with the IoU achieved by the HRNet model after training, the IoU achieved by the automated framework reached 63.69%, achieving the human-derived metric by 86.91%. This substantiates that the automated recognition accuracy aligns with the proficiency of professional annotators engaged in visual identification. Simultaneously, the occurrence of variations in manual IoU values underscores the level of susceptibility to subjective influence from annotators.

5. Conclusions and Outlooks

In this study, we propose a framework for the large-scale automatic detection of IVL based on semantic segmentation of RS images and multi-source data filtering, including urban surface temperature and urban population density data. We selected the HRNet backbone network with the best performance after comparative experiments to train a semantic segmentation model, and verified its reliability and stability. We then validated the workflow using the central urban area of Tangshan City as an example and explained the distribution characteristics of IVL in the region.

In addition to the innovation of the semantic segmentation model, the introduction of multiple data sources ensures that the large-scale automatic IVL detection results are more reasonable and accurate. The segmentation of roads helps to extract IVL in a way that conforms to the basic logic of urban land use, which compensates for the resolution shortcomings of open and free RS images. The introduction of temperature and population density data enables operating industrial sites and IVL to be more easily distinguished. We must make clear, however, that a small number of special industrial types and industrial buildings with good insulation may result in an anomalously lower surface temperature, which could cause a small amount of error in the filtering of operating industrial land. However, for a large number of research targets, the overall framework can accurately map IVL in any city or region with sufficient open-source data.

Several problems remain to be solved to improve the model training results and IVL extraction, including (1) incorporating more bands of RS images to enhance the feature extraction of potential IVL; (2) generalizing the method to high-resolution RS imagery and more precise population density data for specific needs; and (3) increasing the sample size to include more cities to further enrich the training set and improve the model accuracy.

Author Contributions: Conceptualization, Yihao Sun and Xiaodi Zheng; data curation, Ziyang Wang and Yawen Han; formal analysis, Yihao Sun; funding acquisition, Xiaodi Zheng; methodology, Yihao Sun and Han Hu; project administration, Yihao Sun; software, Yihao Sun and Han Hu; supervision, Xiaodi Zheng; validation, Yihao Sun; visualization, Yihao Sun; writing—original draft, Yihao Sun and Yawen Han; writing—review and editing, Yihao Sun, Yawen Han, and Xiaodi Zheng. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Key Research and Development Program of China (No.2020YFC1807505).

Data Availability Statement: Supplementary data to this article are available online at <https://data.mendeley.com/datasets/jzvcgcxd9f/2> for details (accessed on 1 September 2023).

Acknowledgments: The authors would like to thank National Key Research and Development Program of China for helpful supports related to this work. The authors would like to thank the anonymous reviewers for their helpful remarks.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Kim, M.; Kim, G. Typology of urban left-Over space. In Proceedings of the Conference abstracts CELA Annual Meeting, Champaign, IL, USA, 28–31 March 2012.
- Kremer, P.; Hamstead, Z.A.; McPhearson, T. A social-ecological assessment of vacant lots in New York City. *Landsc. Urban Plan.* **2013**, *120*, 218–233. [\[CrossRef\]](#)
- Ardiwijaya, V.S.; Sumardi, T.P.; Suganda, E.; Temenggung, Y.A. Rejuvenating idle land to sustainable urban form: Case study of Bandung metropolitan area, Indonesia. *Procedia Environ. Sci.* **2015**, *28*, 176–184. [\[CrossRef\]](#)
- Martinat, S.; Dvorak, P.; Frantal, B.; Klusacek, P.; Kunc, J.; Navratil, J.; Osman, R.; Tureckova, K.; Reed, M. Sustainable urban development in a city affected by heavy industry and mining? Case study of brownfields in Karvina, Czech Republic. *J. Clean. Prod.* **2016**, *118*, 78–87. [\[CrossRef\]](#)
- Smith, J.P.; Li, X.; Turner, B.L., II. Lots for greening: Identification of metropolitan vacant land and its potential use for cooling and agriculture in Phoenix, AZ, USA. *Appl. Geogr.* **2017**, *85*, 139–151. [\[CrossRef\]](#)
- Pagano, M.A.; Bowman, A.O.M. *Vacant Land in Cities: An Urban Resource*; Brookings Institution, Center on Urban and Metropolitan Policy: Washington, DC, USA, 2000.
- Liebmann, H.; Kuder, T. Pathways and Strategies of Urban Regeneration-Deindustrialized Cities in Eastern Germany. *Eur. Plan. Stud.* **2012**, *20*, 1155–1172. [\[CrossRef\]](#)
- Nassauer, J.I.; Raskin, J. Urban vacancy and land use legacies: A frontier for urban ecological research, design, and planning. *Landsc. Urban Plan.* **2014**, *125*, 245–253. [\[CrossRef\]](#)
- Rieniets, T. Shrinking Cities: Causes and Effects of Urban Population Losses in the Twentieth Century. *Nat. Cult.* **2009**, *4*, 231–254. [\[CrossRef\]](#)
- Reckien, D.; Martinez-Fernandez, C. Why Do Cities Shrink? *Eur. Plan. Stud.* **2011**, *19*, 1375–1397. [\[CrossRef\]](#)
- Wiechmann, T.; Pallagst, K.M. Urban shrinkage in Germany and the USA: A Comparison of Transformation Patterns and Local Strategies. *Int. J. Urban Reg. Res.* **2012**, *36*, 261–280. [\[CrossRef\]](#)
- Sołtysik, M.; Mazur-Belzyt, K. City Space Recycling: The Example of Brownfield Redevelopment. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *960*, 042016. [\[CrossRef\]](#)
- Lange Deborah, A.; McNeil, S. Brownfield Development: Tools for Stewardship. *J. Urban Plan. Dev.* **2004**, *130*, 109–116. [\[CrossRef\]](#)
- Nagengast, A.; Hendrickson, C.; Lange, D. Commuting from US Brownfield and Greenfield Residential Development Neighborhoods. *J. Urban Plan. Dev. Asce* **2011**, *137*, 298–304. [\[CrossRef\]](#)
- De Sousa, C.A. Unearthing the benefits of brownfield to green space projects: An examination of project use and quality of life impacts. *Local Environ.* **2006**, *11*, 577–600. [\[CrossRef\]](#)
- Baing, A.S. Containing Urban Sprawl? Comparing Brownfield Reuse Policies in England and Germany. *Int. Plan. Stud.* **2010**, *15*, 25–35. [\[CrossRef\]](#)
- Potočník, J. Addressing soil contamination, land take and soil sealing: Good for the environment, good for the economy. In Proceedings of the Conference ‘Soil Remediation and Soil Sealing’, Brussels, Belgium, 10 May 2012.

18. Banzhaf, E.; Netzband, M. Detecting urban brownfields by means of high resolution satellite imagery. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *35*, 460–466.
19. Li, X. Use of LiDAR in Object-based Classification to Characterize Brownfields for Green Space Conversion in Toledo. Ph.D. Thesis, University of Toledo, Toledo, OH, USA, 2017.
20. Xu, S.J.; Ehlers, M. Automatic detection of urban vacant land: An open-source approach for sustainable cities. *Comput. Environ. Urban Syst.* **2022**, *91*, 12. [[CrossRef](#)]
21. Li, W.; Zhou, W.; Bai, Y.; Pickett, S.T.A.; Han, L. The smart growth of Chinese cities: Opportunities offered by vacant land. *Land Degrad. Dev.* **2018**, *29*, 3512–3520. [[CrossRef](#)]
22. Song, X.; Wen, M.; Shen, Y.; Feng, Q.; Xiang, J.; Zhang, W.; Zhao, G.; Wu, Z. Urban vacant land in growing urbanization: An international review. *J. Geogr. Sci.* **2020**, *30*, 669–687. [[CrossRef](#)]
23. Li, W.; Wang, D.; Li, H.; Wang, J.; Zhu, Y.; Yang, Y. Quantifying the spatial arrangement of underutilized land in a rapidly urbanized rust belt city: The case of Changchun City. *Land Use Policy* **2019**, *83*, 113–123. [[CrossRef](#)]
24. Sperandelli, D.I.; Dupas, F.A.; Dias Pons, N.A. Dynamics of Urban Sprawl, Vacant Land, and Green Spaces on the Metropolitan Fringe of Sao Paulo, Brazil. *J. Urban Plan. Dev.* **2013**, *139*, 274–279. [[CrossRef](#)]
25. Wang, Z.; Chen, X.; Huang, N.; Yang, Y.; Wang, L.; Wang, Y. Spatial Identification and Redevelopment Evaluation of Brownfields in the Perspective of Urban Complex Ecosystems: A Case of Wuhu City, China. *Int. J. Environ. Res. Public Health* **2022**, *19*, 478. [[CrossRef](#)]
26. Song, Y.; Lyu, Y.; Qian, S.; Zhang, X.; Lin, H.; Wang, S. Identifying urban candidate brownfield sites using multi-source data: The case of Changchun City, China. *Land Use Policy* **2022**, *117*, 106084. [[CrossRef](#)]
27. Mao, L.; Zheng, Z.; Meng, X.; Zhou, Y.; Zhao, P.; Yang, Z.; Long, Y. Large-scale automatic identification of urban vacant land using semantic segmentation of high-resolution remote sensing images. *Landsc. Urban Plan.* **2022**, *222*, 104384. [[CrossRef](#)]
28. Long, J.; Shelhamer, E.; Darrell, T.; IEEE. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 3431–3440. [[CrossRef](#)]
29. Liu, Y.; Duc Minh, N.; Deligiannis, N.; Ding, W.; Munteanu, A. Hourglass-ShapeNetwork Based Semantic Segmentation for High Resolution Aerial Imagery. *Remote Sens.* **2017**, *9*, 522. [[CrossRef](#)]
30. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 5–9 October 2015; pp. 234–241. [[CrossRef](#)]
31. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)]
32. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]
33. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In Proceedings of the 15th European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 833–851. [[CrossRef](#)]
34. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep High-Resolution Representation Learning for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 3349–3364. [[CrossRef](#)] [[PubMed](#)]
35. Mao, K.; Tang, H.; Zhou, Q.; Chen, Z.; Chen, Y.; Qin, Z. Retrieving land surface temperature from MODIS data by using radiance transfer equation. *J. Lanzhou Univ.* **2007**, *43*, 12.
36. Zhigan, Y. Most of China’s Small and Medium-sized Lead and Zinc Smelting Factories Have Ceased Production. *China Nonferrous Met. Mon.* **2009**, *1*, 2–4.
37. Binan Steel Factory Shut down after Explosions. (2013, 2013/08/09/). Philippine Daily Inquirer. Available online: <https://link.gale.com/apps/doc/A339193536/GBIB?u=tsinghua&sid=bookmark-GBIB&xid=bc6b6740> (accessed on 6 September 2023).
38. Pu, Y.; Yu, H. ResUnet: A Fully Convolutional Network for Speech Enhancement in Industrial Robots. In Proceedings of the 35th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE), Kitakyushu, Japan, 19–22 July 2022; pp. 42–50. [[CrossRef](#)]
39. Ye, H.-J.; Hu, H.; Zhan, D.-C.; Sha, F. Few-shot learning via embedding adaptation with set-to-set functions. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 8808–8817.
40. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017. [[CrossRef](#)]
41. Wang, T.C.; Liu, M.Y.; Zhu, J.Y.; Tao, A.; Kautz, J.; Catanzaro, B. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In Proceedings of the IEEE Conference on Computer Vision And Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018. [[CrossRef](#)]

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