This is the Accepted Manuscript version of an article published by Elsevier in the journal *Computers, Environment and Urban Systems* in 2025, which is available at: https://doi.org/10.1016/j.compenvurbsys.2025.102284
Cite as: Liu Y, Wang Z, Ren S, Chen R, Shen Y, Biljecki F (2025): Physical urban change and its socio-environmental impact: Insights from street view imagery. *Computers, Environment and Urban Systems*, 119: 102284.

Physical Urban Change and Its Socio-Environmental Impact: Insights from Street View Imagery

Yingjie Liu[a,1], Zeyu Wang[b,1], Siyi Ren[c], Runying Chen[c], Yixiang Shen[d], Filip Biljecki[e,f,*]

[a]*Department of Architecture, University of Washington, Seattle, 98195, WA, USA*
[b]*Department of Urban Design and Planning, University of Washington, Seattle, 98195, WA, USA*
[c]*The Paul G. Allen School of Computer Science & Engineering , University of Washington, Seattle, 98195, WA, USA*
[d]*Department of Electrical, Information and Physics Engineering, Tohoku University, Sendai, 9808579, Miyagi, Japan*
[e]*Department of Architecture, National University of Singapore, Singapore, 117566, Singapore*
[f]*Department of Real Estate, National University of Singapore, Singapore, 119245, Singapore*

Abstract

Urban transformation not only reshapes physical spaces but also impacts public perception, influencing how people experience their environments. This study utilizes Street View Imagery (SVI) as an emerging, human-level data source to assess urban changes, providing perspective beyond traditional datasets. Existing studies often focus on either urban physical changes or human perception changes, without bridging the two. This research integrates both aspects by combining a change detection model, trained on a self-labeled dataset, and a human perception model based on the crowdsourced Place Pulse 2.0 dataset with input from 81,630 online volunteers, to analyze urban transformation in New York City and Memphis from 2007 to 2023. Our findings reveal differences between the two cities: New York City exhibited small, isolated changes often driven by community needs, while Memphis transitioned from concentrated to more dispersed development patterns. This study provides insights into how physical changes in-

*Corresponding author
Email addresses: yj1996@uw.edu (Yingjie Liu), zeyuw1@uw.edu (Zeyu Wang), siyiren@uw.edu (Siyi Ren), rc76@uw.edu (Runying Chen), hirosawa.ishou.r6@dc.tohoku.ac.jp (Yixiang Shen), filip@nus.edu.sg (Filip Biljecki)
[1]These authors contributed equally.

fluence public perception within these two cities. It demonstrates how thoughtful, well-planned urban transformation can improve neighborhood's perception such as safety and livability, while also pointing out potential challenges like gentrification or social fragmentation. These findings provide policymakers with valuable insights into human perception, aiding in the creation of more inclusive, vibrant, and resilient urban transformation. This helps ensure that urban transformation efforts are based on community desires and align with long-term sustainability goals.

Keywords: Urban Renewal, Gentrification, Change Detection, Deep Learning

---

1. Introduction

Urban transformation, characterized by its variety in scale, duration, and spatial segregation, presents itself through physical changes that not only affect urban landscapes but also impact public perception, thereby having major impact on sustainable development (Koch et al., 2018; Maassen and Galvin, 2019). However, urban transformation occurs through formal and informal processes, posing issues that require close investigation in order to develop effective policy solutions that are adapted to the distinctive dynamics of each urban physical change, such as expansion of informal settlements and businesses, new development, and micro-scale urban renewal (Lara-Hernandez et al., 2020; Kamalipour and Dovey, 2019; Liu and Song, 2024).

To accurately capture and manage urban dynamics, existing studies use various data sources, including site photographs, building permits, construction records, and satellite imagery, to document and demonstrate physical urban changes at different scales (Wiatkowska et al., 2021; Zhang and Seto, 2013; Bennett and Smith, 2017; Venter et al., 2020). While these techniques effectively monitor physical changes in the city, they have limitations in directly capturing impacts at the human level. SVI provides a broad, historical, and detailed view of the urban environment from a human perspective, offering an objective and comprehensive record of physical changes (Biljecki and Ito, 2021; Wang et al., 2024). However, previous studies using SVI have primarily relied on single data epochs without considering multiple snapshots or historical versions to understand urban evolution and the various phenomena that occur as cities change.

Scholars have explored different methods to detect and monitor urban transformation (Ilic et al., 2019; Chen et al., 2021; Sakurada and Okatani, 2015; Huang et al., 2024). Traditional methods, such as questionnaires, interviews, or statistical analysis of socioeconomic data, provide valuable insights but lack detailed visual information on the urban landscape. Some scholars have employed machine learning models to compare streetscapes and detect changes, achieving high accuracy rates (Huang et al., 2024; Chen et al., 2021). However, studies focusing on change detection have primarily addressed physical changes, often neglecting to analyze their impact on human perception. Meanwhile, other scholars focus on changes in urban perception, often equating perceptual changes directly to the broad concept of changes in visual elements (Liang et al., 2023; He et al., 2023), without rigorously examining specific elements in the streetscape. This approach falls short in characterizing urban changes in a precise manner, and the machine learning models used cannot determine which type of urban change corresponds to the detected pixel changes and perception changes. Temporary in street elements, such as moving vehicles or swaying tree branches, may affect perceptual evaluations. This limitation restricts the depth and precision of analysis and the understanding of the detailed impacts of urban changes on human perception (Wang et al., 2024; Liu and Song, 2024).

Our research seeks to answer two questions: 1) Which parts in an urban area exhibit human level urban physical changes over time, and at what scale? 2) Do detected changes in specific locations lead to improvements or deterioration in human perception? The results of this study can help improve urban living standards, inform policy refinements, and the effective implementation and monitoring of these policies.

Our study informatively devised a workflow to quantify the distribution of urban physical changes over time and reveal their impacts through an analysis of perceived urban environmental quality scores, addressing both positive and negative consequences such as fragmentation or gentrification. Specifically, we used semantic segmentation to exclude temporary and irrelevant factors, such as vehicles, human activities, and vegetation changes, from SVIs. A simplified version of the change detection model (Liu and Song, 2024), as shown in Figure 1, was trained on a self-labeled dataset of 4,000 SVI pairs, which represents the largest change detection classification dataset for urban scenes within our studied cities. This model was then used to analyze time-series SVIs, aiming to identify whether physical change occurred in the city, particularly in buildings and streets. The change detection model employs two convolutional backbones derived from a pre-trained VGG16 model (Simonyan and Zisserman, 2014), generating two similarity scores to enhance feature extraction of the texture and detail of the input data. These similarity scores are combined and processed using a Support Vector Machine (SVM) classifier, which includes a fully connected layer network with ReLU activation and dropout regularization for classification. Our change detection results outperform the latest research results (Liu and Song, 2024) using similar size dataset of static SVIs.

Subsequently, we focus specifically on SVIs that exhibit physical changes. To assess whether these urban changes have positive or negative perceptions and impacts, we employ a pre-trained deep learning model of human perception. This model uses Place Pulse 2.0, a well-known comprehensive human perception dataset that is widely adopted (van Veghel et al., 2024), to predict perceptions of the urban environment before and after changes. It evaluates detected changes based on four key human perception metrics: 'beauty', 'wealth', 'safety', and 'liveliness'. Based on the positive and negative changes in these perception scores, we further analyzed the impact of physical changes, as shown in Figure 1. Through this dual approach of physical change detection and perception analysis, our research provides a more informative guidance for the policy maker to understand of how previous urban transformations impact both the physical landscape and people's perception and experience of these spaces in an inclusive and large-scale manner within the studied cities. Additionally, non-professionals may find it useful to understand how changes in their environment affect their daily experiences.

2. Literature Review

2.1. Theories related to urban change and their potential impacts

Urban physical change is a process driven by complex and interrelated factors, including physical, social, economic and ecological aspects of the urban environment (Shi et al., 2020). These factors not only shape the spatial layout and architectural character of cities, but also have a considerable impact on the structure of neighborhoods, the lifestyles of residents, and the state of the regional economy (Bratuškins et al., 2020). Urban renewal and redevelopment programs, while aimed at upgrading the city's infrastructure and living environment, may also trigger gentrification, forcing former residents out of their neighborhoods (Earley, 2023; Levine et al., 2022). At the same time, the growth and expansion of cities demonstrates their economic vitality, but such expansion may also be accompanied by urban contraction and land use change, exacerbating environmental pressures and social divisions.

The social ramifications of urban change encompass far-reaching and profound effects, not only altering the physical structure of neighborhoods but also significantly influencing the quality of life of their residents (Liang et al., 2023). Schelling and Grodzins' tipping point theory articulates how neighborhoods on

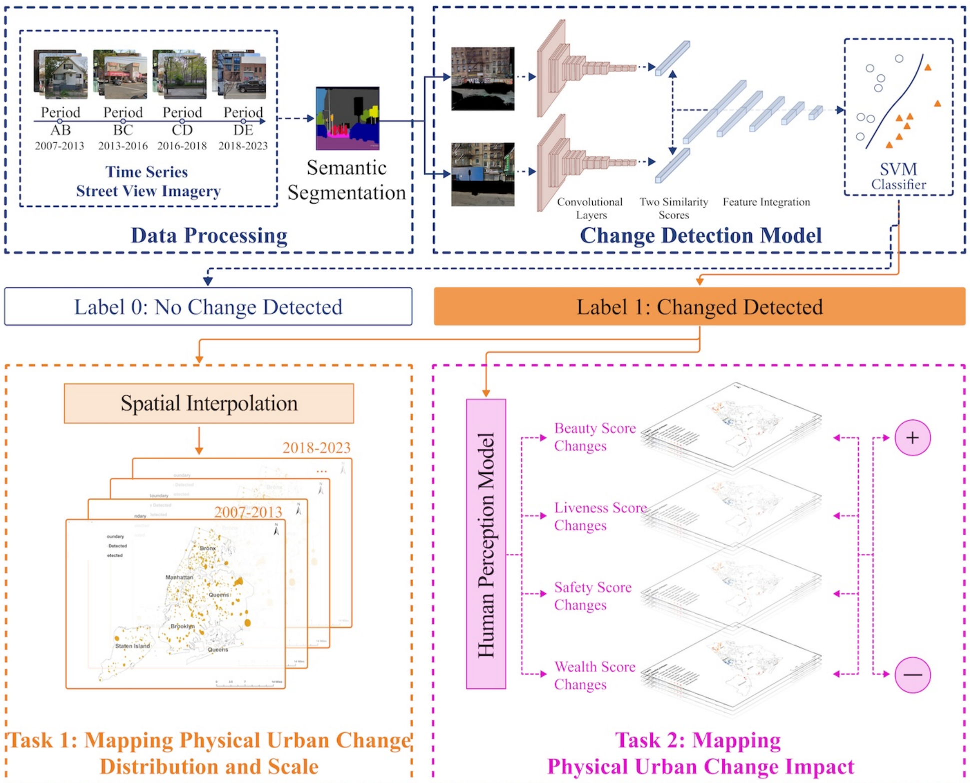

Figure 1: Overall workflow of the developed methodology.

the socio-economic brink can either decline further or improve, depending on various factors, including human capital development and educational achievements (Schelling, 1971; García, 2019). This concept emphasizes the critical role of fostering positive social progress to mitigate risks in poorer areas and enhance more affluent ones (Schelling, 1971). Concurrently, Burgess's intrusion-differentiation theory sheds light on the importance of geographic positioning and social networks in urban renewal efforts. It points out that urban improvement initiatives often cluster around central business districts and their adjacent desirable neighborhoods, indicating a targeted approach to urban enhancement (Schelling, 1971)

Building on these theories, the work of Naik et al. (2017) on urban environmental transformations reveals the cascading effects of changes within urban spaces, demonstrating the interconnectedness of modifications and their capacity to impact surrounding areas. This has been further illustrated in literature,

Figure 2: Selected pairs of SVIs showing different categories of changes. Data source: GSV.

e.g. through a seven-year study of Turkish neighborhoods, which showed that changes within communities disrupt residents' daily lives and social interactions, ultimately altering the social fabric and cultural identity of the area (Atay Kaya, 2021).

Given this complex backdrop, acknowledging the variability and breadth of urban physical changes becomes imperative. Figure 2 is provided for better understanding of the types urban changes. These changes, which vary in scale, duration, and spatial segregation, and can occur through both formal and informal processes, present challenges that require a comprehensive understanding. It is this variability and its consequences that underline the need for effective policy responses, tailored to address the specific dynamics of each urban transformation.

The influence of urban physical changes extends beyond tangible transformations to significantly impact human perception. Research demonstrates that modifications to the built environment, encompassing street layouts, building facades, and public spaces, fundamentally shape how residents experience and interpret their urban surroundings. This perceptual dimension adds another layer to the complex interplay of factors driving urban change, as discussed in theories like Schelling's tipping point and Burgess's intrusion-differentiation model. Recent technological advancements, particularly Street View Imagery (SVI) and machine learning models, have revolutionized our ability to study these perceptual impacts systematically (Ito et al., 2024). These tools enable researchers to conduct quantitative analyses of how the built environment and urban design influence subjective experiences of safety, vibrancy, and aesthetics (Zhang et al., 2021, 2018b; Wang et al., 2024; Kang et al., 2023; Biljecki and Ito, 2021).

Despite extensive research on urban environments and perception, a critical gap exists in our understanding of how physical urban changes affect human perception in controlled settings (Ito et al., 2024). While numerous studies examine static urban environments, there is a remarkable scarcity of research investigating how temporal changes at specific locations influence human perception. This knowledge gap is particularly significant given that urban transformation occurs as a continuous process rather than a static event, suggesting our understanding of the relationship between physical changes and their perceptual impacts remains largely unexplored in current literature.

2.2. Challenges and methodological advancements in monitoring urban change

The challenge of effectively monitoring urban change has captivated scholars for years. This ongoing phenomenon encompasses the transformation and development of various urban elements, including buildings, roads, green spaces, streetscapes, and neighborhoods (Ma et al., 2021). To grasp and manage these dynamics accurately, it necessitates sustained observation and thorough analysis. Traditional methods leaned on questionnaires and interviews (McGinn et al., 2007), visual surveys, or manual image assessments (Chen et al., 2009; Roth, 2006). These are typically hindered by low data yield and restricted spatial detail. Additionally, while socio-economic data offered valuable statistical insights, it fell short in detailing the visual information of urban landscapes (Liu and Song, 2024).

In recent years, the significance of visual data for urban change detection and monitoring has surged, propelled by advances in remote sensing technology (Wiatkowska et al., 2021; Wentz et al., 2014; Liang et al., 2023; Ito et al., 2024; Velasquez-Camacho et al., 2023). This data includes optical remote sensing imagery (ORSI) (Zitzlsberger et al., 2021), nighttime light imagery (NTL) (Bennett and Smith, 2017; Zhang and Seto, 2013), synthetic aperture radar (SAR) (Moulton et al., 2008; Cihlar et al., 1992), and light detection and ranging (LiDAR) (Venter et al., 2020). These methods facilitate large-scale, multi-level analysis of urban layout and land use changes at a fraction of the cost of traditional interviews and surveys, though each carries its own set of pros and cons. While these technologies offer substantial benefits in monitoring urban environments, challenges such as VHR's susceptibility to distortion, SAR's noise issues, and LiDAR's high costs and data complexity remain. Moreover, while these methods of data acquisition perform well in monitoring physical changes in cities, there are limitations in directly capturing urban physical changes' impacts on people, which is a particular concern according to existing literature (Liu and Song, 2024). Images acquired from satellite-based or airborne platforms, while effective in detecting topside changes, are less efficient in capturing subtle changes at the human level or street scale.

Over the last decade, the proliferation and utilization of SVI around the world has unveiled new datasets capturing urban changes, proving to be an invaluable asset for research (Biljecki and Ito, 2021). Offering a distinct vantage point significantly different from remote sensing techniques, SVI serves as an easily accessible data source that captures images from a human perspective (Zhang et al., 2018b; Biljecki and Ito, 2021). This unique feature enables SVI to convey the details of urban life and the tangible transformations within cityscapes in ways that surpass the capabilities of traditional statistical and satellite data (Li et al., 2015). It is adept not only at documenting broad urban development and renewal endeavors but also at detailing the minutiae of neighborhood beautification or aggravation efforts (Zhang et al., 2018a). In terms of time series, SVI tends to provide finer temporal coverage than ground-based LiDAR, which is more costly and often lacks optical data (i.e., color). Furthermore, when juxtaposed with social media imagery, SVI provides superior coverage (Wang et al., 2024; Zhang et al., 2018a; Liu et al., 2024; Hou et al., 2024). As a result, SVI maintains a good balance between efficiency, data completeness, and data collection costs, making it a reliable data source for large-scale mapping of physical urban change.

In recent years, scholars have increasingly employed deep learning techniques for urban visual element change detection research (Shi et al., 2020). Typically, evaluation metrics are utilized in supervised learning-based change detection research: the accuracy, which measures the ratio of correctly predicted observations to the total observations, is used to evaluate the performance of the image classification task. However, in change detection tasks, we often encounter imbalanced datasets. In such cases, the F1 score is introduced, as it indicates a balance between true positive rates and false positives. This helps ensure that the model correctly identifies both instances where changes have occurred and where they have not, making it a good choice for evaluating change detection model performance.

In 2015, Sakurada and Okatani (2015) proposed a convolutional neural network (CNN) combined with superpixel segmentation for change detection in panoramic images, achieving an F1 score of 0.639 on their self-collected and annotated Google Street View (GSV) dataset. Alcantarilla et al. (2018) used coarsely registered image pairs with an adaptation of Fully Convolutional Network (FCN) model for pixel-level change detection, achieving an F1 score of 0.614 on the same dataset. Chen et al. (2021) introduced the Dynamic Receptive Temporal Attention Module (DRTAM), achieving an impressive F1 score of 0.871 on the same dataset. Researchers have also explored self-supervised learning approaches. For instance, the Street2Vec method employs a ResNet-50 and multi-layer perceptron (MLP) model with the Barlow Twins self-supervised learning approach to extract embeddings from street-level images taken at the same location in different years (Stalder et al., 2024). These optimized embeddings are insensitive to irrelevant changes (e.g., lighting, seasons) but sensitive to structural urban changes, enabling urban change detection by comparing the cosine distance between embeddings. However, they did not demonstrate the explainability of their model, merely highlighting its performance surpassing generic pretrained embeddings. This raises concerns about the model's generalization capabilities in downstream tasks.

However, when tested on a real-world, randomly sampled, and annotated small-scale GSV dataset, where annotations were based on the presence or absence of urban physical changes (0/1 labels), Ilic et al. (2019) developed a Siamese network model using transfer learning, initialized with pre-trained VGG19 weights from ImageNet. Their SCNN-FC-8 model demonstrated high efficacy in the street view dataset, achieving an F1 score of 0.72. Huang et al. (2024) used fine-tuning of the Dino v2 model to achieve a change detection F1 score of 87.95%. Liu and Song (2024) employed a Markov SVM classifier integrated with ResNet as the encoder, achieving an accuracy of 0.76 after training on 2000 samples for 18 epochs. Given the similarities in labeling, training, and validation setups, their results serve as a benchmark for comparison with the findings of this study. These advancements effectively identified physical urban changes but left the exploration of impacts on human perception for future research.

In summary, prior studies have mainly leveraged diverse data sources for change detection, neglecting to evaluate its impacts, particularly regarding human perceptions. Our research seeks to address this gap by exploring in greater depth the actual impact of urban physical changes, especially from the perspective of human perception. We make use of these well-established change detection methods and take them forward by putting them in the context of impacts, ensuring that urban change detection not only maps physical alterations but also aligns better with the needs and well-being of urban residents.

3. Methodology

3.1. Study Area

In examining the dynamics of the physical environment, we focused on two considerably different American cities: New York City, a metropolis with a diverse demographic and economic composition (Savitch, 2010), and Memphis, historically known for its industrial base, currently undergoing significant political and economic transformation (Raciti et al., 2016). This selection enables comparing the differences in economic, social, and built environment development among large and medium-sized cities. Research in recent years has generally pointed to the fact that NYC has undergone significant socio-economic transformation, with the phenomenon of gentrification being particularly prominent (Chapple et al., 2021). While this process has contributed in part to the prosperity of its metropolitan area, it has also led to significant changes in community structure, with numerous neighborhoods facing different forms of displacement (Zukin, 1987). In particular, low-income groups have found themselves struggling to cope with rapidly rising housing prices, a trend that related wage increases and housing policies do not appear to have effectively mitigated (Zukin, 1987). In contrast, Memphis, although smaller in size, is a city with a well-known cargo airline hub, and its economic structure has shown changes and developments in recent years, which in turn may shape the urban physical environment. Thus, a comparative study of the two is of some interest, and will contribute towards understanding of the applicability of the developed method across a variety of urban settings.

3.2. Data preparation

To document the physical changes in NYC and Memphis from 2007 to 2023, we collected imagery from GSV at sample points with 300-meter intervals across the cities' street networks. Panoramic imagery spanning from 2007 to 2023 was requested and categorized by location for the sample points. However, it should be noted that not all sample points have imagery for the entire time span. Also, sample points on the road segments classified as highways and ramps were excluded due to their lower likelihood of exhibiting urban physical changes at the neighborhood level. The counts of sample points for NYC and Memphis are 26,175 and 20,003, respectively. The distribution of SVI over time was notably uneven. To mitigate the uneven temporal distribution of SVI, the dataset was segmented into four periods: period AB: 2007 to 2013; period BC: 2013 to 2016; period CD: 2016 to 2018, and period DE: 2018 to 2023. This segmentation process, involving 1278 potential configurations, was guided by the need to balance the number

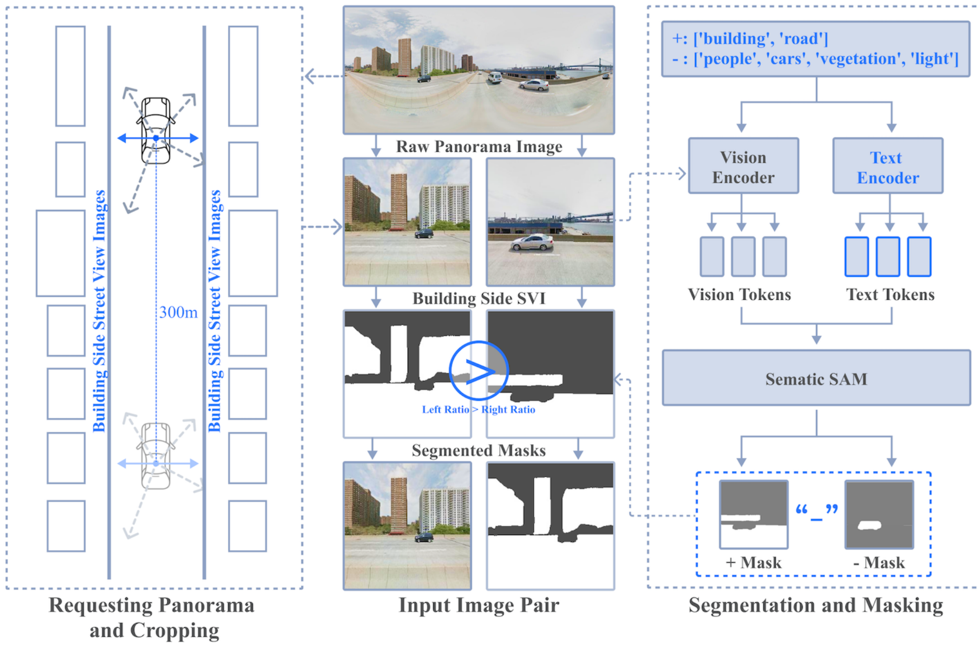

Figure 3: The sampling process for SVIs, including the steps for panorama cropping and masking.

of images across each time interval, with the final selection informed by a comprehensive assessment of the total number of observation points and the standard deviation of each time segment.

The subsequent processing of panoramic SVIs is shown in Figure 3. To crop panoramic images, adjusting the Field of View (FOV) and heading to produce 1024x1024 panoramic unfolded SVIs. An FOV of 60 degrees was chosen to minimize sky coverage, thereby focusing more on urban streetscapes. Images capturing both sides of the street were gathered, with adjustments made to the heading based on the heading information from panorama metadata, followed by a modification of ±90 degrees. The earliest heading recorded at each standpoint served as a baseline for aligning subsequent time-series SVIs from different time segments, ensuring consistency of heading for each standpoint. Furthermore, a pre-trained VGG model and SVM were employed to screen out SVI with sub-optimal alignment, enhancing the dataset's overall quality.

Our study adopts a broad definition of urban physical changes, specifically excluding temporary and irrelevant elements such as vehicles, human activities, and vegetation changes from the analysis. While vegetation can sometimes be important, its variability can introduce significant randomness, thereby affecting the overall performance of the model (Biljecki et al., 2023). By excluding these elements, we aim to reduce errors and minimize false positives in our analysis. Semantic-SAM (Li et al., 2023), a semantic segmentation model enable segmenting and recognizing streetscapes. The SAM model has been tested on various urban segmentation datasets(Zeng and Boehm, 2024), achieving accuracy and recall rates exceeding 80% for categories including 'building', 'road', 'sky', 'cars', 'bikes', and 'people' . The SAM model was used in this research to differentiate between categories such as 'building', 'road', 'people', 'cars', 'vegetation', creating corresponding masks for each SVI.

The categories 'building' and 'road' were highlighted in white in mask image, with the remaining categories set to black to exclude temporary elements from further analysis. The choice of representative images for each standpoint within a given time segment was determined by comparing the proportion of the sum of white pixels in the masks on both sides of the street, selecting the side with a higher white pixel ratio to represent that particular time segment and standpoint.

For the preparation of change detection dataset, 2000 images from each city were annotated for the training and validation sets, following the methodology of previous research (Liu and Song, 2024). The images were labeled as either "positive: detected change in the building and road categories" or "negative: no changes detected in the building and road categories" by PhD students with backgrounds in architecture and urban planning. These annotations were cross-validated by three additional students from computer science backgrounds. The benchmark dataset includes a variety of weather and lighting conditions, which helps improve the model's robustness across different urban scenes.

The annotated data were divided into training and validation sets at a ratio of 0.75 to 0.25. This approach was intended to ensure the model's accuracy in understanding and analyzing physical changes in urban street images, thereby supporting further research in urban planning and development.

3.3. Model Description

3.3.1. Change detection model

We address the challenge of urban imagery classification by implementing a simplified version of change detection model (Liu and Song, 2024), as shown in the Figure 1, which contains DPSM model (Ding et al., 2020) for feature extraction and SVM (Cortes and Vapnik, 1995) as classifier. The change detection model utilizes a Siamese network structure with two convolutional backbones derived from a pre-trained VGG16 model (Simonyan and Zisserman, 2014), whose parameters are kept constant during the training process. This setup combines the strong feature-extraction ability of VGG16 with specialized layers and processes to enhance the model's capability of evaluating similarities between pairs of images. SVM is well-suited for urban change detection tasks and is used in this research as the classifier because urban change detection is fundamentally a binary classification problem. SVM excels at handling such tasks by maximizing the margin between different classes, making it effective for distinguishing subtle differences in SVIs.

The model normalizes inputs to match the expected format for VGG16, using specified mean and standard deviation values. It then performs feature extraction using a five-stage pre-trained VGG16 model, where each stage consists of convolutional layers followed by pooling operations (Ding et al., 2020). This technique maintains spatial hierarchies while effectively minimizing dimensions and preserving essential data. The model applies masks at each stage of feature extraction to focus on relevant areas of the images, thereby sharpening the features' relevance for similarity evaluation. It generates two similarity scores: one reflects the average of the features, and the other focuses on variances and covariances to distinguish textures and details. These scores are combined and processed through SVM, a network of fully connected layers with ReLU activation and dropout regularization for classification.

The change detection model was trained over 19 epochs with a batch size of eight. To address the data imbalance, we adjusted weight using the inverse square root of the imbalance ratio (1.83) within our cross-entropy loss function. We used the Adam optimizer to refine the model, starting with a learning rate of 0.0001. A learning rate annealing strategy was employed at the 19th epoch, lowering the learning rate to 0.00001 to enhance the optimization process and ensure the model closely approximates the best solution (Nakamura et al., 2021). The repository of the model is available in Appendix A.

3.3.2. Human Perception model

Human perception plays a crucial role in how individuals evaluate and experience spaces, making it an essential metric for assessing the quality of physical environments (Liang et al., 2024). The ability to monitor urban transformations and determine whether they have positive or negative impacts is grounded in the analysis of these perceptions. The Place Pulse 2.0 dataset is a well-established and widely utilized dataset in urban perception studies, often regarded as a standard

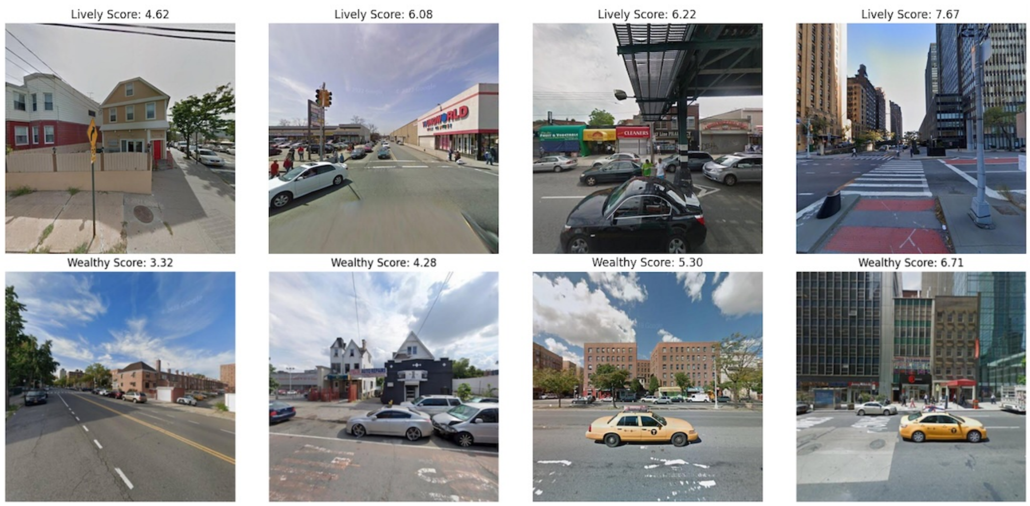

Figure 4: Examples of perception scores in New York City.

in the field. It has been extensively used in research to train deep learning models that predict human perceptions from imagery (Dubey et al., 2016; Wang et al., 2024; Zhang et al., 2018b; Wei et al., 2022; Salesses et al., 2013; Larkin et al., 2021; Zhou et al., 2021; Kruse et al., 2021; Zhang et al., 2020). This dataset contains over 100,000 Google SVIs rated by more than 80,000 volunteers, have been shown to reflect human perception scores in categories such as 'beauty,' 'wealth,' 'safety,' and 'liveliness' through SVI analysis. The crowdsourced nature of Place Pulse 2.0 ensures a broad and diverse spectrum of opinions, making it a valuable tool for categorizing urban scenes. Figure 4 illustrates an example of the 'wealth' and 'liveliness' scores in different areas of New York City, as these two indicators may be more abstract than others. "Wealthier" refers to visual signs of affluence, such as well-maintained buildings and clean streets. "Livelier" describes the perceived vibrancy or energy of an area, influenced by elements like colorful architecture, billboards, public spaces, or other features that suggest activity, beyond just the presence of people or vehicles.

Our research involves a comparative analysis of images from locations before and after they have experienced changes, utilizing the Deep Convolutional Neural Network (DCNN) models developed as part of the work by Hou et al. (2024). This process involves analyzing pairs of images, before and after the changes, with the pre-trained DCNN model (Hou et al., 2024) to generate respective human perception scores, with an accuracy of 76.9% for 'beauty', 72.9% for 'wealth', 76.7%

for 'safety' and 77.1% for 'liveliness'. In addition, we have evaluated the uncertainty of our human perception model predictions using Monte Carlo Dropout. In this approach, each test image is processed 30 times with dropout enabled, and the final prediction is computed as the mean of these outputs, while the uncertainty is quantified as the corresponding standard deviation. The result shows that the average and median uncertainty values across all four indicators (safer, livelier, wealthier, and beautiful) fall within an acceptable range, supporting the reliability of the model's predictions. For a detailed description of the methodology and statistical results, please refer to Appendix B. Finally, by comparing scores changes on the detected urban physical changing spot, we are able to determine the trajectory of urban physical changes in terms of human perception, thus evaluating and discuss the qualitative impact of physical modifications on urban environments.

For further analysis, we manually categorize the pairs of images that show changes. This inductive approach helps us illustrate and discuss the reasons and effects of these changes, enhancing our understanding of the different impact of urban transformations on human perception during the study period (2007-2023).

4. Result

4.1. Change Detection Model Performance

The change detection model classified the changed and unchanged SVI in four time periods. In the classification task, the label with the higher probability score was selected as the predicted label. After training, as shown in Figure 5, the change detection model achieved a validation accuracy of 76.87% in NYC's trainset after 19 epochs. For the Memphis dataset, training took 10 epochs, achieving a validation accuracy of 83.0%.

The analysis of NYC encompassed 16,343 pairs of SVIs from different time periods, while for Memphis the number is 14,988 pairs. The data points labeled as "changes detected" were 1,552 in New York City and 1,681 in Memphis. When compared with state-of-the-art models, using similar data sources, over similar training epochs (Liu and Song, 2024), our change detection model, which utilized a relatively simpler model without incorporating rule-based contextual knowledge, reached a comparable or higher accuracy. This outcome was especially evident in the mid-sized city of Memphis, whose benchmark dataset is unevenly distributed, as the urban changes are less commonly happening there when compared to larger cities like NYC. These results highlight the model's ability to maintain applicability without the need for supervision of building attributes data, ensuring

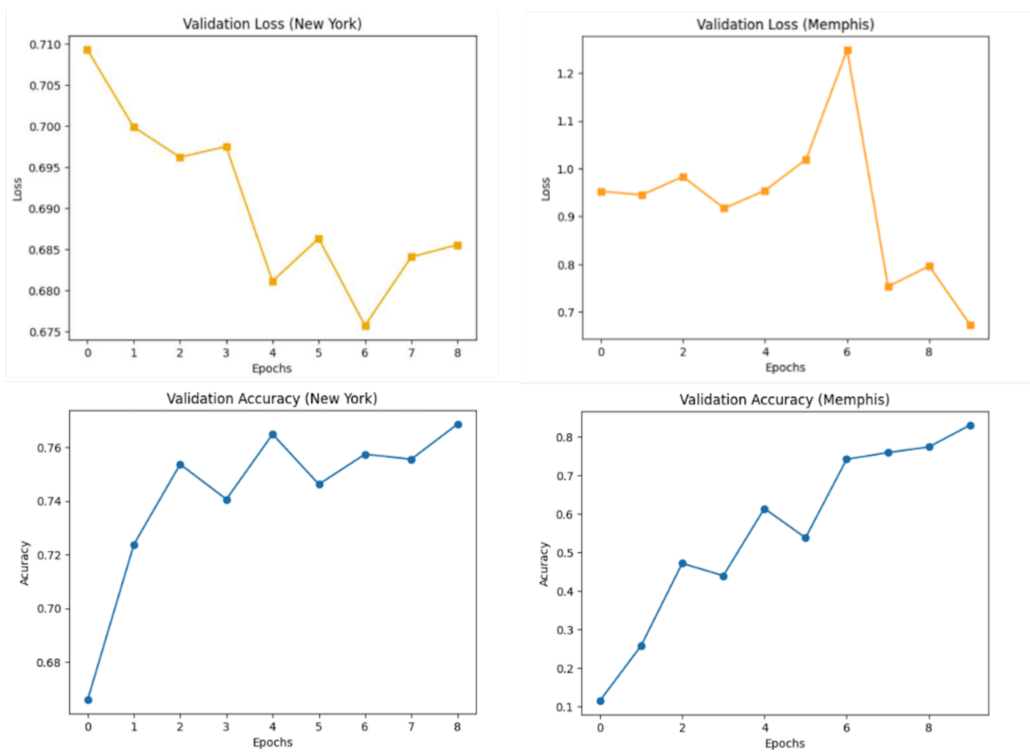

Figure 5: Change detection model results. The figure shows both training and validation loss and accuracy for the change detection model. It reports that the model achieves its best performance with the lowest loss and highest accuracy.

its broader usability, especially in those small cities that lack government official documents.

4.2. Change Detection Result for NYC and Memphis

The change detection results are presented in two steps. First, we analyze the spatial aggregation of all image pairs identified as positive, highlighting the locations and scales of detected changes over the relevant periods. Second, for each detected change, we evaluate whether it reflects a positive or negative human perception. The detailed data processing and visualization methods are outlined below.

The spatial distribution of the SVIs used for comparisons across different time intervals is inherently uneven, exhibiting a heterogeneous pattern. To address this, we employ the Inverse Distance Weighting (IDW) method, a widely used spatial interpolation technique. IDW estimates the value of an unknown location based on the known values of surrounding data points, with the influence of each point decreasing as distance increases. This approach enables us to capture the spatial patterns of urban changes more effectively by assigning greater weight to points closer to the location being estimated.

By applying IDW, we generate a continuous surface that reveals where and to what extent physical changes have been detected across the city. While the change detection model outputs binary values (0 or 1) indicating whether a change has occurred, IDW transforms these outputs into a gradient of values between 0 and 1. A value closer to 1 indicates a higher concentration of detected changes in a given area. For visualization purposes, areas with an estimated value of 0.5 and above were identified and marked with colors, indicating regions with significant urban change.

Focusing on the specifics, our examination of SVIs labeled as “1” revealed a range of urban physical changes over time. To assess the impact of these changes, it is essential to evaluate shifts in citizens’ perceptions. The perception model analyzes SVIs of the tested cities across four factors: beauty, wealth, safety, and liveliness. In this case, we calculated the difference in perception scores between two time periods and applied hotspot analysis to these differences. This method identifies regions where the aggregated positive or negative changes in perception scores deviate from the expected patterns. By mapping these areas, we can visually represent how perceptions of urban change have shifted over time, distinguishing areas where scores have significantly improved (hotspots) or declined (cold spots). This approach allows us to observe how urban changes influence perceptions by highlighting areas with noticeable positive or negative gaps between the start and end of each period.

To ensure the thoroughness and dependability of our analysis, we considered four levels of confidence: 80%, 90%, 95%, and 99%. These confidence levels help us determine how certain we are that the clustering of perception scores for beauty, wealth, safety, and liveliness is not due to random chance but instead reflects meaningful patterns. The higher confidence levels, such as 99% and 95%, represent areas where we are highly certain that the clustering is significant and reliable, providing robust insights into how urban changes are perceived. In contrast, the lower confidence levels, such as 90% and 80%, offer a more exploratory perspective, allowing us to capture broader regions where perception patterns might be emerging or less pronounced. By incorporating this range of confidence levels, we provide a comprehensive analysis that highlights both the most statistically significant areas and those where perception shifts may be developing. While all spots with a confidence level greater than 80% are important, the varying thresholds allow us to balance statistical rigor with exploratory insights. In the following subsection, we will show the results for NYC and Memphis respectively.

4.2.1. New York City

Spatial Agglomeration Analysis. The analysis depicted in Figure 6 illustrates urban physical changes in NYC from 2007 to 2023, identifying regions of aggregated transformation. Changes are marked in colors, indicating areas with an estimated change value of 0.5 and above, suggesting notable urban development or alteration.

Period AB (2007-2013) was marked by recovery and growth following the 2008 financial crisis, we have witnessed widespread urban changes across NYC. Notably, most areas in Bronx, Northern-Eastern Queens, lower Manhattan and Staten Island, experienced developments largely driven by housing expansions and community revitalization efforts. Specifically, starting from 2005, Bronx decided to provide 165,000 housing units on vacant land or redevelop existing estates (Chronopoulos, 2017). Manhattan's Hudson Yards project, initiated in 2012, became one of the largest private real estate endeavors in the U.S., transforming the West Side with its mix of commercial, residential, and public spaces.

Period BC (2013-2016) saw changes in the Northern part of Staten Island, particularly in Brighton Heights, as well as in Bay Terrace and the southeastern Bronx (Sound View). During this period, the city continued to focus on post-Sandy recovery, particularly in areas like Staten Island. This region has seen a mixture of residential development and environmental restoration projects, especially in the wake of Hurricane Sandy in 2012. In addition, Brookville in Queens near JFK

International Airport, and waterfront area Coney Island in Brooklyn likely saw development driven by new developments in the neighborhoods.

Period CD (2016-2018) highlighted substantial changes across Staten Island, including Mid Island, New Springville, and Richmond. A part of reason is driven by the demand for more affordable housing options in NYC, initially spurred interest and investment, which may have had ripple effects in adjacent areas like Mid Island and New Springville. For Queens, a spatial aggregation of detected changes was found in Eastern side, specifically in Jamaica center. Jamaica center is a highly strategic location with multiple rail lines connecting Manhattan, JFK airport and Long Island.

Period DE (2018-2023) indicated a shift towards smaller and more segregated changes, with no large-scale developments observed. However, mid-scale transformations in upper Manhattan, near Manhattanville and Rockaway Beach in Queens.

Throughout these periods, Queens, Staten Island, and the Bronx consistently showed a high concentration of changes, particularly near waterfronts and mass transit hubs. At the early stage, the scale of detected changes tends to be larger and the distribution is spread-ed city wide. However, as the time goes by, the scale of urban changes tends to be smaller and dis-aggregated. Developments often included mixed-use projects, affordable housing, flood defense projects and new public spaces. Despite these extensive changes, the broader impact of new developments on the urban fabric and community life remains to be fully assessed.

These results can be cross-validated with the research of Chapple et al. (2021), which employs traditional data analysis methods for evaluating gentrification and displacement risks, analyzing regional housing, income, and other demographic data.

Perceptual Impact Assessment. As illustrated in Figure7 and Figure 8, evaluating the impact on people's perceptions across distinct temporal period using four dimensions: 'beauty', 'wealth', 'safety', and 'liveliness'. The principal insight reveals that the study, anchored in an analysis of changes, underscores an alignment between the perceived urban physical changes impact and the outcomes from change detection efforts. This alignment especially emphasizes that the scale of changes in locales at the periphery of the city surpass those observed within the core city.

During the AB period, enhancements in 'beauty', 'wealth', 'safety', and 'liveliness' were noted in the southern Bronx and Staten Island. A decline in 'beauty', 'wealth', 'safety', and 'liveliness' was noted in central Manhattan, with'safety'

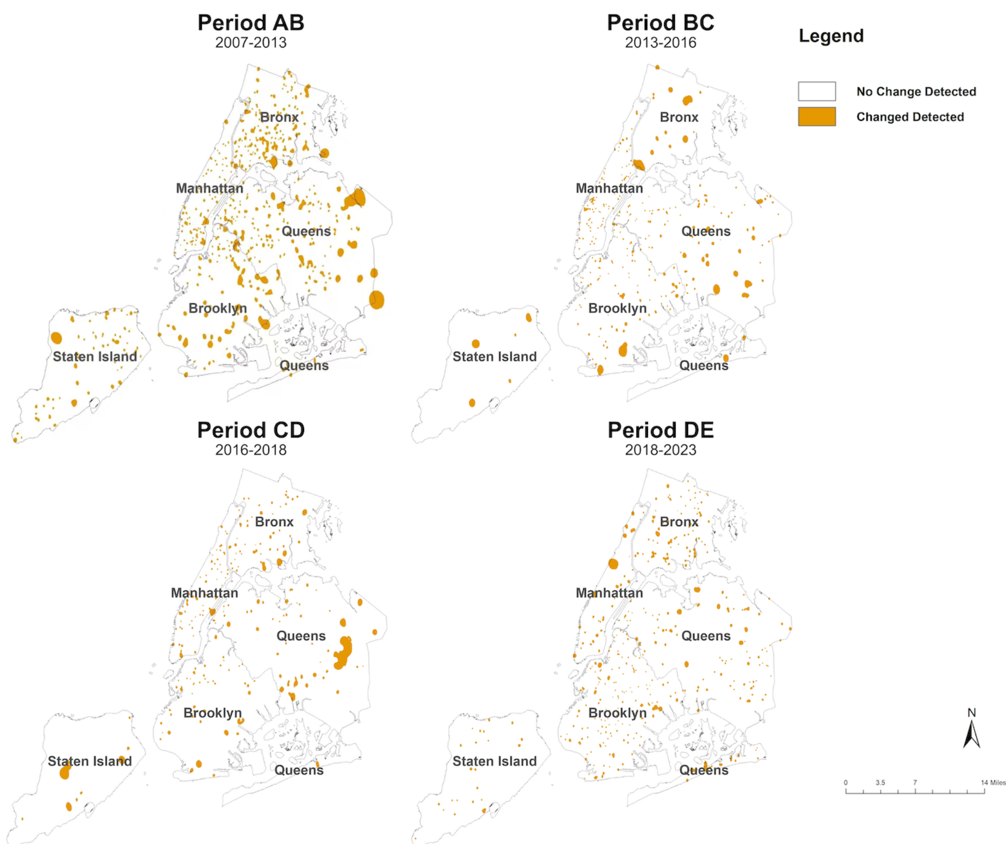

Figure 6: Detected urban physical changes in NYC. The maps show the locations of detected changes.

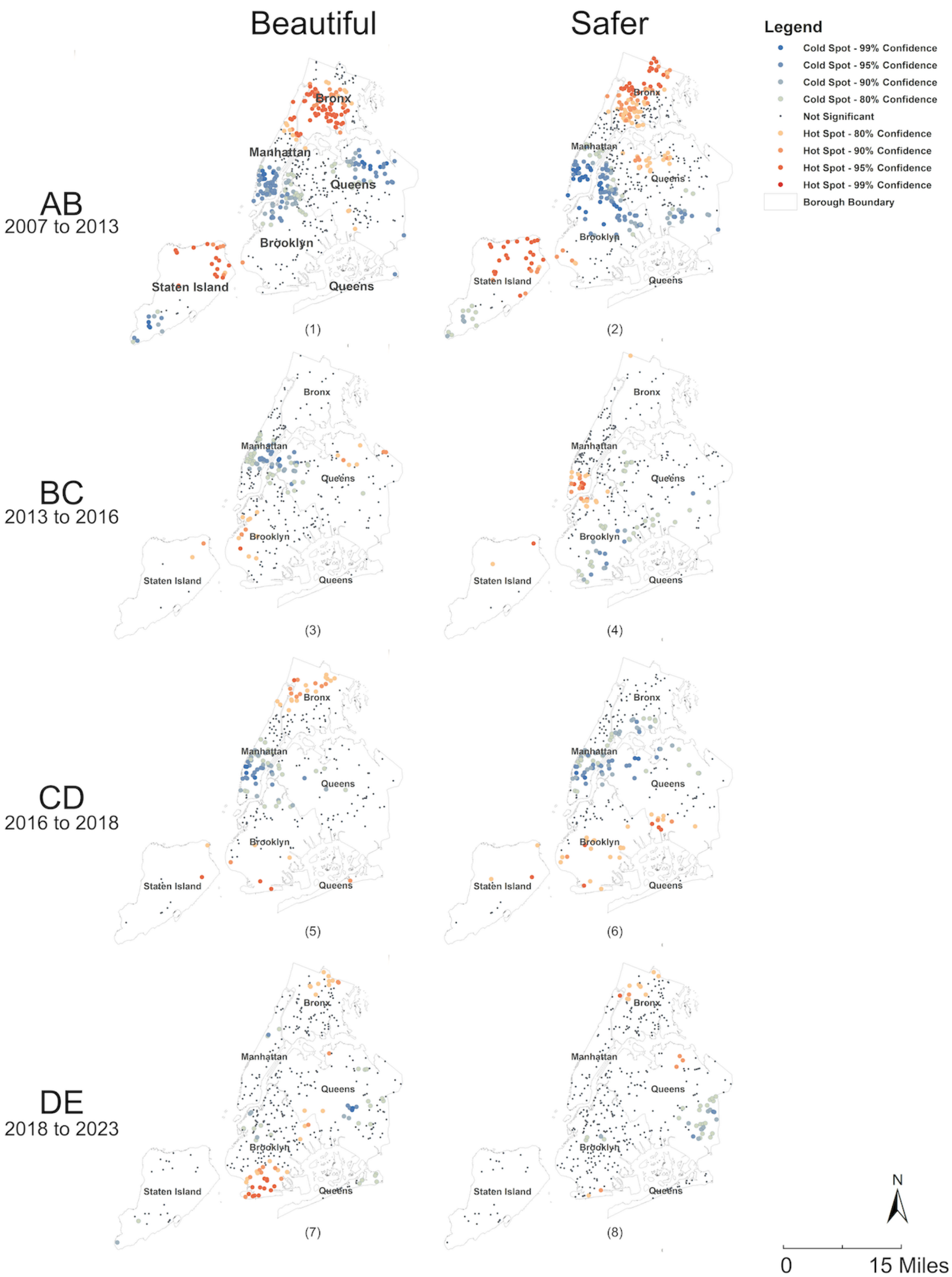

Figure 7: 'Beauty' and 'Safety' perception score changes in NYC.

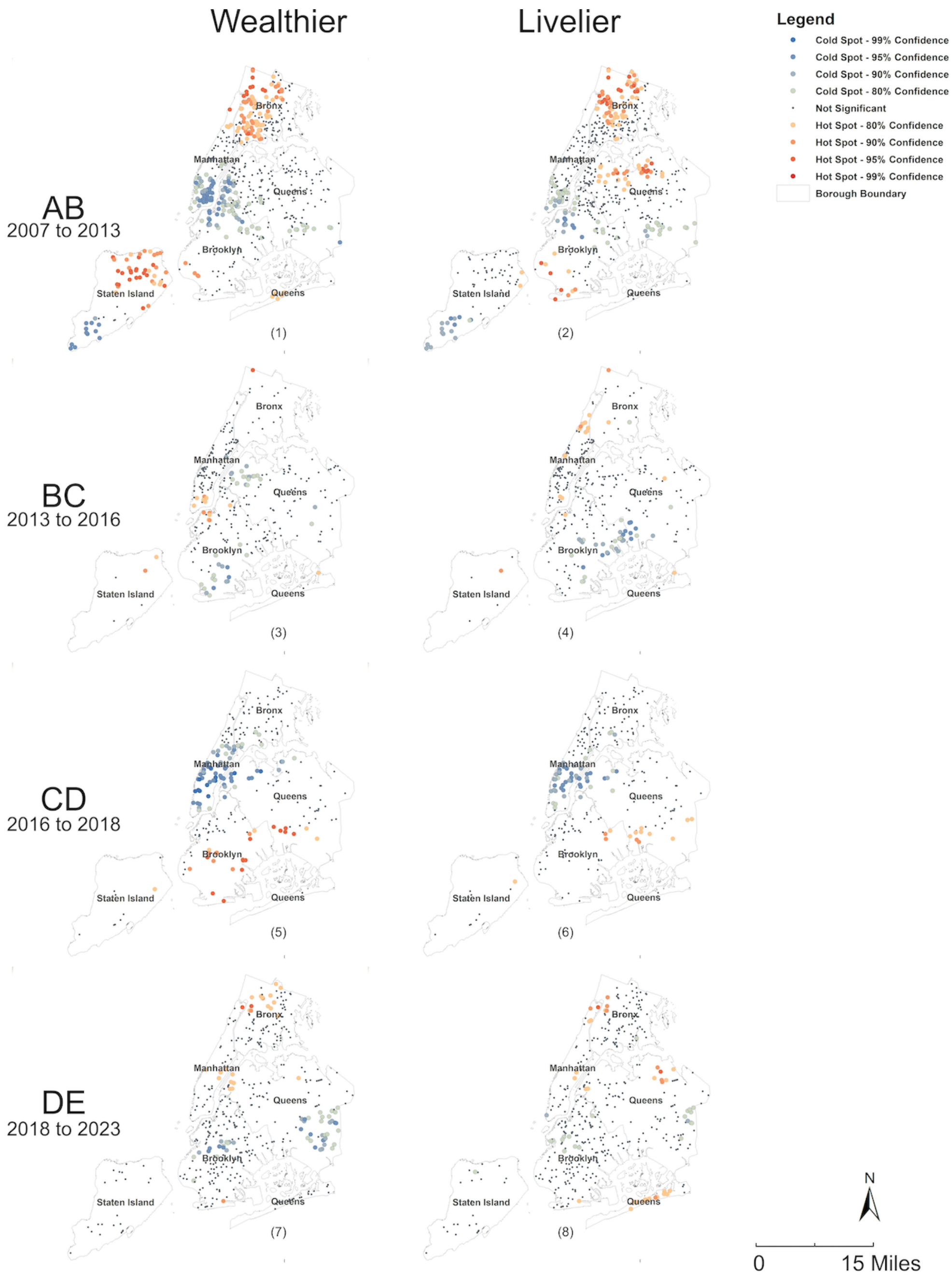

Figure 8: ‘Wealth’ and ‘Liveliness’ perception score changes in NYC.

specifically facing a downturn in the southwest of the island.

In the BC period, ‘beauty’ and ‘safety’ marked increases in northern Queens, with ‘safety’ and ‘wealth’ improving in the central to lower Manhattan areas. All four factors saw slight enhancements on Staten Island. Proximate to Midtown Manhattan and its neighboring areas to Queens, ‘beauty’, ‘wealth’, ‘safety’, and ‘liveliness’ experienced varying degrees of reduction. The nexus between northern Brooklyn and Queens recorded declines in ‘liveliness,’ ‘safety,’ and ‘wealth.’

The CD period showcased bigger scale of increases in all four factors compared to the BC period within the northern Bronx, southern Brooklyn, and their intersecting areas with Queens, alongside the southeastern edge of Staten Island. Comparably notable decreases were observed in central Manhattan, adjacent to the East River and Brooklyn, the central regions of Queens, and the southern boundary between the Bronx and Manhattan.

In the DE period , ‘beauty’, ‘wealth’, ‘safety’, and ‘liveliness’ scores increased in the southern Brooklyn and northern Bronx, with more modest uplifts in northern Queens. The eastern portion of Queens recorded relatively minor decreases across all dimensions, while central and northern Brooklyn witnessed comparable reductions on a small scale and threshold of change.

The analysis reveals that the detected changes in perception score align closely with the input detected physical changes data. Throughout the four phases examined—AB, BC, CD, and DE—the AB phase witnessed the most significant changes in perception scores due to detected changes, with the latter three phases showing comparably small quantities of changes. Notably, major perception alterations were concentrated along the East River and around TOD theory oriented key transportation centers, with changes in perception scores ranging from a minimal 0.01 to between 4.45. Variations in the ‘wealth’ factor may suggest a connection to gentrification (Glaeser et al., 2020).

Upon a thorough examination of the SVI pairs labeled as “1,” the detected changes were categorized into distinct groups: Road Signs, Street Furniture, Sheds, Building Facades, Renovations, and New Constructions, as depicted in the Figure 2. These modifications are dispersed across a variety of locations, emerging at different times and with varying frequencies. Such a pattern underscores the need for detailed analysis to grasp the full extent of their impacts comprehensively.

4.2.2. Memphis

Spatial Agglomeration Analysis. The analysis depicted in Figure.9 illustrates urban physical changes in Memphis from 2007 to 2023, identifying regions of aggregated transformation.

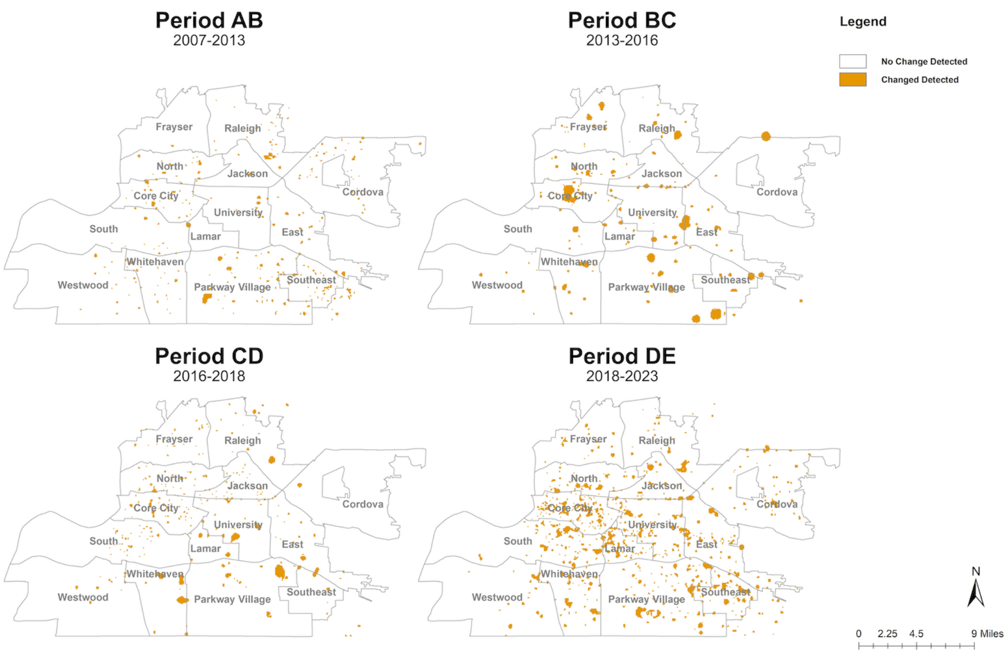

Figure 9: Detected urban physical changes in Memphis. The maps show the locations of detected changes.

Memphis's development shows progressive changes during the AB, BC, CD, and DE stages. Through each detected change is smaller in scale than in NY, but they are distributed more evenly across the city.

In the AB period (2007-2013), changes were mostly found in the downtown and southern parts of the city. Specifically, Parkway Village, representative of the southern districts, experienced multiple scattered but comparable large scale changes. Substantial private investments and worsening safety led to a middle-class exodus, ultimately causing economic collapse and the closure of many stores (Steimer and Steimer, 2024).

During the BC period (2013-2016), the Core City, University, Frayser and Raleigh districts experienced changes across several blocks. In the northern part, especially Raleigh, with the decline of the car dealership industry, the city government purchased the original community center's shopping mall in 2016 and began renovating it as a municipal center. The renewed building includes administrative and public library functions. The project continued until 2020, and changes were detected by the model in the CD/DE periods (Appeal, 2016; City of Memphesis, 2023). Additionally, the Core City and East districts' main commercial area experienced comparable large-scale changes (MLK50: Justice Through Journalism, 2022).

During the CD period (2016-2018), changes were detected in the southern part of the city, including Parkway Village, Whitehaven, the University area, and the Core City. These changes mainly occurred in residential areas, likely due to improvements in neighborhood streets and housing renovations or rebuilding. These areas, located on the city's periphery, have long suffered from high vacancy rates and community decline (MLK50: Justice Through Journalism, 2022). The movement of people led to renovation of building's facades, while vacant houses resulted in deteriorating landscapes. Both positive and negative changes were observed.

In the DE period (2018-2023), the city experienced more widespread, smaller-scale changes. In 2019, Memphis started to implement the 'Memphis 3.0' plan, which included over 100 projects focused on economic development and urban environmental improvements (Memphis et al., 2020). This plan involved multiple rounds of public participation for each community, analyzing local issues, especially in vulnerable city regions and communities (Wiloandco, 2024). The plan outlined key development projects and also small-scale projects, such as adding public spaces and enhancing landscapes, aimed at improving community vitality and overall quality of life. These small scaled projects are abundant in the quantity and scattered around the city, aligning with the change detection results for the DE period.

This series of urban developments is closely linked to Memphis's land policies. Before 'Memphis 3.0', the city had not updated its urban planning for nearly 40 years, resulting in a disconnect between planning and land use regulations (Saija et al., 2019; Barlow et al., 2015). The previous strategy focused on urbanizing undeveloped land as a way to drive economic growth. Memphis established the Center City Commission to attract private investment through incentives. However, due to a lack of effective constraints on private investors, this approach primarily served private interests, benefiting affluent communities the most. Meanwhile, communities most in need of resources struggled with ongoing disinvestment (Waters, 2022). Such developments repeatedly occurred as the economic situation remained stagnant, necessitating further land sales and the turnover of underutilized commercial buildings. The 2019 Memphis 3.0 plan shifted focus toward equitable economic policies to ensure that most Memphians could effectively benefit from limited investments, resulting in smaller but more practical urban changes (Opticos Design, 2024).

Perceptual Impact Assessment. For the perceptual impact of all detected changes, as shown in the Figure 10 and Figure 11, it is found that during the AB periods, overall, changes in 'beauty' score and 'liveliness' score were smaller, while changes in 'safety' score and 'wealth' score were larger. Safety improvements were concentrated in the central-northern part of the city, while declines were seen in the southeastern areas. The positive changes in 'safety' and 'wealth' scores in the central-north were due to public-private partnerships, allowing private investments to supplement public funding for major projects, such as developing historic districts, re-purposing buildings, and improving riverfront areas (Memphis et al., 2020). In contrast, deteriorating safety conditions in the southern city led to upscale establishments being replaced by lower level smaller businesses, which later resulted in vacant retail spaces and homes, contributing to the declines of the scores in the southeast.

During the BC and CD periods, overall changes were smaller in comparison with period AB. In the BC period, changes were evenly distributed across the city, with spotty urban renewal perception changes in areas like Westwood, Whitehaven, and Parkway Village in the south part of the city. 'Wealth' score changes were concentrated in Core City, University, South, and Whitehaven. In the north, Raleigh and Frayser saw decreases in 'wealth' and 'safety' scores.

The CD period witnessed declines in downtown for 'safety', 'liveliness', and 'wealth' scores. There were small increases in 'beauty', 'wealth', and 'safety'

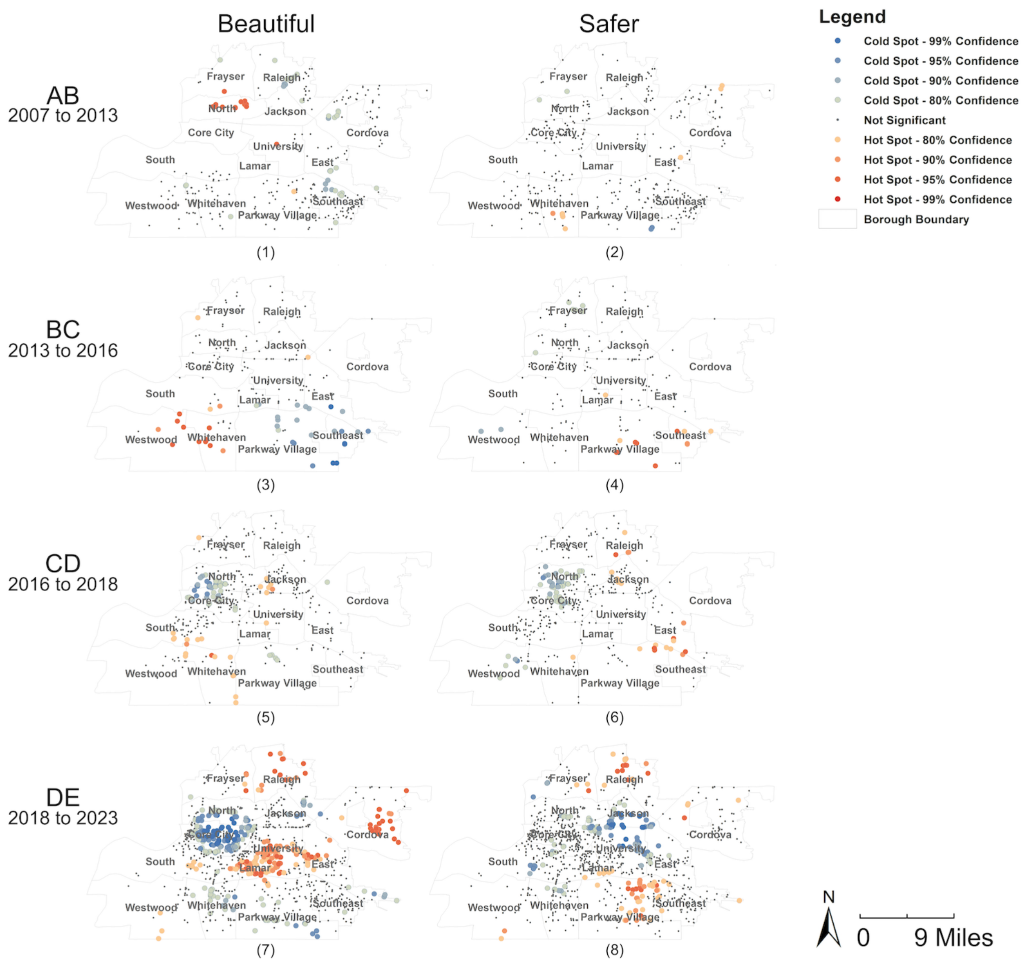

Figure 10: ‘Beauty’ and ‘Safety’ perception score changes in Memphis.

Figure 11: ‘Liveliness’ and ‘Wealth’ perception score changes in Memphis.

scores in the mid, south and north part of the city. ‘Beauty’ score improved in parts of Jackson, South, and Whitehaven, while ‘safety’ score improvements were concentrated in the western part of Jackson. ‘Wealth’ score increases were mainly in the northern part of Frayser. These areas, located on the periphery of the expanding city core, indicate that some previously vulnerable communities have received more attention and experienced positive urban changes.

In the DE period, downtown, midtown, and the southern city experienced declines in ‘beauty’, ‘wealth’, and ‘safety’ scores. In midtown, areas like University, Lamar, and East districts saw increases in ‘beauty’ and declines in ‘liveliness’ scores, with small improvements in ‘safety’ and ’wealth’ scores. There is a focus on mixed-use infill development along streets and the renovation of decaying buildings in these areas in the Memphis 3.0 plan, showing signs of gentrification. Additionally, in Raleigh and Parkway Village, positive perception changes were detected, contrasting with previous periods. The Memphis 3.0 plan specifically mentions new plans for mixed-use projects in these areas, providing more opportunities for small businesses (Memphis et al., 2020).

Overall, in the early stages of the study period, many positive changes occurred in the city center. However, over time, by the end of the study period, positive changes were more concentrated in relatively peripheral areas, with some previously seen as negative areas showing positive signs.

5. Discussion

In order to visualize the specific content of the changes, we extracted the eligible images from the our database of this study for further analysis. Specifically, images exhibiting a confidence level of 80% were arranged based on a sorted order of variations in perception scores, leading to the selection of pairs of time-series SVIs across four time periods for illustration. These image pairs were labeled with cluster IDs, while the time period was marked by varied stock colors, complemented by a corresponding map to pinpoint the locations of the physical changes detected. This categorization enabled a focused examination of the prevalent types of detected physical change impacts for subsequent in-depth analysis.

In NYC’s case, as shown in Figure 12, urban changes are detected in four different categories: furniture, sheds, facades, and new construction. For the detected changes classified as furniture, the DE time periods (2018-2023) witnessed changes due to the global spread of COVID-19, which profoundly affected street-level businesses(Meng et al., 2024). Initiatives such as the NYC Department of

Figure 12: Categories of perception score changes in New York. (Numbers below each pair of images are cluster IDs).

Transportation's Open Restaurant Program facilitated distinctive adaptations during this unique period (DOT, 2020). The detected urban physical changes labeled as cluster ID 9634 and 8188 led to an uplift in perception scores across various aspects, spanning from a minimal decrease of −0.02 to an increase of up to 1.87. Despite a marginal reduction in the perceived aspects of 'beauty' and 'wealth', these adaptations significantly bolstered the perception of 'safety', with increases surpassing 1.87. Such interventions, including the deployment of open wooden semi-open spaces for retail, outdoor dining, and social interactions (Gibson, 2020), contributed to fortifying communal bonds amongst citizens during and after the pandemic (Hu and Rosa, 2020).

The time periods AB, BC, and CD (2007–2018) highlighted the adverse impact of safety sheds installed along the East River due to deteriorating facades. While intended to enhance pedestrian safety, these semi-permanent structures led to a decline in resident perceptions, with reductions in safety indicators ranging from -0.86 to -4.42, especially in cluster ID 6920. Ironically, the sheds became potential hubs for nocturnal crime (Aneja, 2020). The absence of such sheds in the DE period reflects the success of removal efforts initiated in 2019 by the Adams administration and various NGOs (Siff, 2023). This underscores the importance of accelerated assembly, rigorous inspections, and strategic facade renovation planning to improve urban livability and pedestrian experiences (Harlem World Magazine, 2023).

Facade renovations detected in SVIs span several areas, including Lower Manhattan, Queens, the link between Downtown Brooklyn and Manhattan, and the southwest Bronx, across the AB, CD, and DE periods. These renovations display a diverse changes, encompassing commercial facade updates, color shifts in single family homes and community centers, the removal of declining businesses (Lee, 2023), and the emergence of graffiti. For instance, cluster id 9098 in Crown Heights, Brooklyn, noted for its graffiti, saw a detrimental impact on urban aesthetics due to low-quality graffiti between 2018-2023, leading to a perception decrease in safety and liveliness by -0.85 to -4.39. Conversely, cluster id 9907 in Downtown Brooklyn showed a notable improvement post-renovation, with a liveliness score increase of 3.08 and a 'wealth' increase of 2.23 during the AB period, highlighting potential gentrification concerns. Additionally, cluster id 6945 in Chinatown, Lower Manhattan, faced storefront vacancies spurred by rent hikes. A mere facade clean-up resulted in a modest perception improvement ranging from 0.59 to 1.54, with a 'wealth' increase of 1.25. This scenario subtly illustrates the gentrification challenge confronting local small enterprises in Chinatown post the 2008 East Village/Lower East Side Rezoning, where rental prices surged despite the community's median income stagnating around $40,000 (Xu, 2013).

During the BC, CD, and DE periods, new constructions led to varied changes in perception across four key indicators with changes ranging from -0.08 to a 5.72, notably in liveliness. Cluster id 7363 present a glass façade building inserted in the historical street, located in Park Slope, Brooklyn, saw 2 points 'wealth' score increased, suggesting potential links to gentrification. In fact, Park Slope has becoming a sought-after location for Manhattan residents. A 2011 survey highlighted a 19.3% surge in median housing prices over five years, positioning Park Slope's housing market 145.28% above Brooklyn's median sales price.

In the time period CD, DE ranging from 2016 to 2023, the Brooklyn area witnessed two instances of new construction that led to urban physical changes, yielding favorable perception score changes spanned from 0.19 to 1.69, with safety improvements being the most notable of 1.69. Specifically, for the site identified as cluster id 14249, the increase in perceived 'wealth' score after new construction was moderate at 0.88. However, improvements in safety and liveliness were exceeding 1.5. This balance suggests that the new constructions enhanced neighborhood comfort and safety without triggering the adverse effects typically associated with gentrification. The thoughtful approach to these constructions is praiseworthy for its sensitivity to the existing urban fabric and community well-being.

Between 2016-2023, two instances of new construction-related urban physical changes were detected in the Brooklyn area, bringing positive evaluation results. The score changes ranged from 0.19 to 1.69, with the most significant change being in safety, reaching 1.69. For cluster id 14249, the change in 'wealth' score after the new construction was only 0.88, while other data, including safety and liveliness, increased by more than 1.5. This ensured that the comfort of the neighborhood residents was maintained without risking gentrification caused by the physical changes of the construction. The design is commendable for its consideration.

In Memphis case, as shown in the Figure 13 urban changes are detected in four different categories, including facade, sign changes, renovation, and new construction.

For facade changes, we observed color changes and beautification efforts in various buildings during the BC, CD, and DE periods. Clusters 6486, 8049, and 10198 showed increases in all four factors from 0.1 to 3.55. Meanwhile, cluster 5804 showed an increase in 'beauty', with declines in 'liveliness', 'safety', and 'wealth' scores.

For sign changes, from AB to DE, all examples showed declines in 'beauty' and 'wealth' scores, ranging from -0.03 to -0.63. These changes occurred in the Frayser, Core City, South, and Whitehaven districts, covering retail sign changes and gas station signs. This reflects the industrial shifts possibly due to land policies in Memphis (MLK50: Justice Through Journalism, 2022).

During periods AB and BC, both northern and southern districts experienced renovation and new construction. Clusters 10195, 3649, and 1712 showed upward trends in most indicators, ranging from 0.1 to 3.19. In periods CD and DE, clusters 4018, 3719, 8578, and 2346 showed changes, including facade reconstruction, additions, renovations, and new single-family houses. All four indicators showed downward trends, ranging from -0.46 to -2.78.

Looking at the spatial aggregation of urban change in NYC and Memphis and the corresponding changes in human perception from a comparative perspective, we can draw some interesting findings. Urban change in Memphis shows a gradual shift from a pattern of few and relatively concentrated development to many and dispersed, while NYC has been characterized by small and dispersed. Combined with the corresponding changes in perceived scores, we get a glimpse of potential differences in the motivations and processes of urban change in the two cities.

Change in NYC has been driven in large part by economic and neighborhood needs. At various times, Queens, Staten Island, and the Bronx have consistently shown a high concentration of change, especially those area close to waterfront and mass transit hubs. These developments typically include mixed-use projects, affordable housing, flood control projects, and new public spaces. While the modifications brought about by the COVID-19 pandemic (e.g., outdoor dining spaces)

Figure 13: Categories of perception score changes in Memphis. (Numbers below each pair of images are cluster IDs).

have slightly reduced perceived ‘beauty’ and ‘wealth’, they have generally improved the sense of safety and community connection. While the broader impacts of these changes on the urban fabric and community life have yet to be fully assessed, they have dramatically changed the face of NYC and the way residents live.

In contrast, the changes in Memphis have generated different public perceptions in different areas. Some areas have progressed through facades beautification and signage updates, while others have declined under the influence of industry shifts and local policies. This pattern of growth has often ignored social and spatial inequalities, leading to increased fragmentation and suburbanization within the city. While the challenges of downtown Memphis remain, the public planning project referenced in “Memphis 3.0: Two Tales of the Same Plan” demonstrates that while the project was initiated by the public sector and managed by the Office of Public Planning, in practice it relied more on private funding and resources (Memphis et al., 2020). This reflects the passive role of municipalities in planning, which may lead to instability and lack of sustainability in long-term development.

NYC attempts to balance economic development with social justice and emphasizes public participation in its urban consolidation and planning strategies, while Memphis focuses more on economic growth and the promotion of private interests. These differences reflect the fundamental differences between the two cities in their approaches to urban development and socio-spatial inequality. The contrasts reveal the different paths Memphis and NYC have taken in adapting to and managing urban change and the socio-economic drivers behind them, and highlight the importance of planning strategies and public engagement in shaping the future of cities.

Of course, there is room for improvement in this study. The first issue concerns SVIs. Due to inherent limitations of the street view system, some areas exhibit angular deviations and uneven distribution. This problem was particularly evident during our test in Memphis, where SVIs were often poorly captured, less abundant, blurry, distorted, or severely tilted. We look forward to further refinement of this data source in small to medium-sized cities. The second issue is the interpretability of the model. Like many AI/ML models today, two of the models used in the paper remain black box to humans to a certain degree. Although we scrutinized the categories detected by the change detection model, it would be helpful to add an explainable component for the deep learning model to show where its attention is focused. This will provide further validation. Third, our study focuses on applying machine learning to detect urban changes in street view images and assess their impacts on human perception within the study area. However, the experimental design and analysis do not directly support specific or detailed policy recommendations. Additionally, while the model can be retrained without modifying its architecture, extending these findings to other cities would require further data collection, annotation, and analysis. Despite these constraints, we have made the model available for the community, enabling further experimentation and adaptation in different contexts. Future research could address these aspects across diverse urban contexts. Finally, while we have employed several spatial analysis techniques that are sufficient for the scope of this study, we acknowledge that future research could benefit from more formal spatial statistical approaches to further refine and deepen our understanding of spatial patterns.

6. Conclusion

In this work, we propose a framework to assess fine-grained urban change at scale with time series SVIs. We propose an end-to-end change detection pipeline to identify urban change points at scale. In conclusion, by comparing urban change and its impact on human perception in NYC and Memphis, we reveal the distinct paths taken by these cities in addressing urban development and socio-spatial inequality. In NYC, early large-scale changes transitioned to smaller, dispersed modifications, reflecting ongoing urban development and revitalization efforts. Perceptual assessments aligned with these changes, indicating significant impacts on 'beauty', 'wealth', 'safety', and 'liveliness', especially in peripheral areas. In Memphis, the model detected more evenly distributed but smaller-scale changes, consistent with the city's shift towards equitable economic policies and community-focused developments. This comparison emphasizes the crucial role of planning strategies and public participation in shaping the future of cities. Furthermore, our research highlights the importance of integrating human perception into the evaluation of urban changes. By bridging the gap between physical urban development and resident experience, it offers a comprehensive framework for creating more livable, inclusive, and resilient urban spaces.

We see this study as a step forward in utilizing SVI, an increasingly important dataset in urban studies (Hou et al., 2024), for large-scale urban change monitoring. By integrating a change detection model and a human perception model into the workflow, we bring an innovation in detecting the scale and magnitude of urban change and evaluating the impact of urban change on human perception. Based on this advancement, we interpret the model results by integrating urban development plans, enhancing the interpretability and policy relevance of the model.

Acknowledgements

YL and ZW acknowledge financial support from the University of Washington's Graduate School and College of Built Environments. The research team also extends their gratitude to the University of Washington for the facilities provided. This research is part of the project Large-scale 3D Geospatial Data for Urban Analytics, supported by the National University of Singapore under the Start Up Grant R-295-000-171-133. We gratefully acknowledge the data sources used in this study.

Appendix A. Model Repositories

The change detection model is available at:
https://github.com/Yevsquant/UrbanChangeDetection

The human perception model is available at:
https://github.com/strawmelon11/human-perception-place-pulse

Appendix B. Human Perception Model Uncertainty Analysis

In this section, we have illustrated how we have quantified the uncertainty of our model in human perception predictions by employing Monte Carlo Dropout during inference. In this framework, for each image in the test set, the model performs $N = 30$ stochastic forward passes with dropout enabled. The final predicted score for an image is computed as the arithmetic mean of these predictions,

$$\mu = \frac{1}{30}\sum_{i=1}^{30} s_i,$$

and the associated uncertainty is quantified by the standard deviation,

$$\sigma = \sqrt{\frac{1}{30}\sum_{i=1}^{30}(s_i - \mu)^2}.$$

Both μ and σ are scaled to the 0–10 range, which provides a direct measure of the prediction variability induced by the dropout mechanism.

To ensure that our evaluation comprehensively covers the full range of spatial perception, we applied a stratified sampling approach based on these scores in all the images we have used for the city of Memphis in the study. First, the continuous score range (0–10) is divided into three intervals: low (0–3), mid (3–7), and high (7–10). For each perception indicator, the number of images within each of these intervals is determined, and proportional sample sizes are calculated so that the final test set comprises 2,000 images in total.

Figure A.14 presents a combined scatter plot that visualizes the relationship between predicted scores and their corresponding uncertainty for all four perception indicators. In this plot, each point represents an image, with its predicted score on the x-axis and the computed uncertainty on the y-axis. The results show the majority of images for most indicators exhibit uncertainty values below 1.0, indicating high model confidence. In contrast, the *livelier* indicator tends to show a broader spread in uncertainty, which suggests that predictions for liveliness are more variable and possibly more subjective. The *beautiful* indicator, on the other hand, demonstrates the lowest overall uncertainty, implying more stable and consistent predictions, which aligns with the findings in the existing research (Rui and Cai, 2025).

Furthermore, an examination of the scatter plot indicates that images with extreme scores (either very low or very high) generally tend to have lower uncertainty, whereas those with mid-range scores display higher variability. This observation suggests that when the visual features are clear and distinctive, the model is more confident in its prediction, while ambiguous or borderline cases naturally result in greater prediction variability. Overall, the average and median

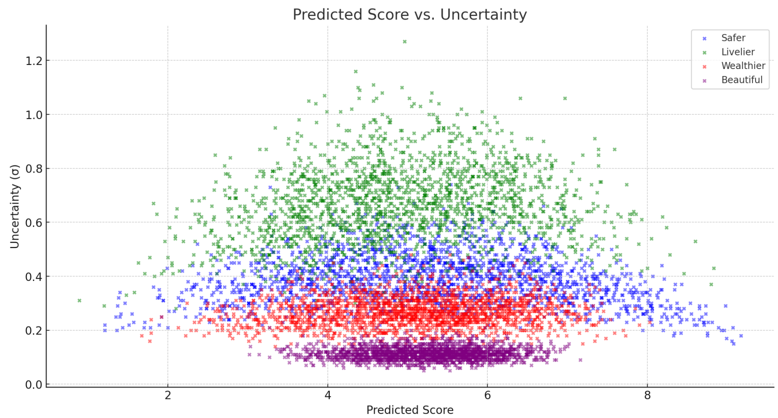

Figure B.14: Combined scatter plot showing the distribution of uncertainty values across the dataset.

uncertainty values across all indicators fall within an acceptable range, supporting the reliability of the model's predictions while simultaneously providing a useful measure to flag images that may require further review.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT in order to proofread the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Alcantarilla, P. F., Stent, S., Ros, G., Arroyo, R., and Gherardi, R. (2018). Street-view change detection with deconvolutional networks. *Autonomous Robots*, 42:1301–1322.

Aneja, S. (2020). *Alternative Shelters: Immortalizing the New York City Sidewalk Shed*. PhD thesis, Syracuse University.

Appeal, R. P. C. (2016). Memphis buys last raleigh mall property for development. *The Commercial Appeal*.

Atay Kaya, I. (2021). Changes in neighbourhoods near urban transformation areas: Izmir (turkey) example. *Uluslararası Sosyal Bilimler Akademi Dergisi*, pages 755–783.

Barlow, S., Pacello, T., and Whitehead, J. (2015). Regulatory created blight in a legacy city: What is it and what can we do about it. *U. Mem. L. Rev.*, 46:857.

Bennett, M. M. and Smith, L. C. (2017). Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics. *Remote Sensing of Environment*, 192:176–197.

Biljecki, F. and Ito, K. (2021). Street view imagery in urban analytics and GIS: A review. *Landscape and Urban Planning*, 215:104217. Publisher: Elsevier.

Biljecki, F., Zhao, T., Liang, X., and Hou, Y. (2023). Sensitivity of measuring the urban form and greenery using street-level imagery: A comparative study of approaches and visual perspectives. *International Journal of Applied Earth Observation and Geoinformation*, 122:103385.

Bratuškins, U., Zaleckis, K., Treija, S., Koroļova, A., and Kamičaitytė, J. (2020). Digital Information Tools for Urban Regeneration: Capital's Approach in Theory and Practice. *Sustainability*, 12(19):8082. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.

Chapple, K., Thomas, T., and Zuk, M. (2021). Urban displacement project website. *Berkeley, CA: Urban Displacement Project*.

Chen, B., Adimo, O. A., and Bao, Z. (2009). Assessment of aesthetic quality and multiple functions of urban green space from the users' perspective: The case of Hangzhou Flower Garden, China. *Landscape and Urban Planning*, 93(1):76–82.

Chen, S., Yang, K., and Stiefelhagen, R. (2021). Dr-tanet: Dynamic receptive temporal attention network for street scene change detection. In *2021 IEEE Intelligent Vehicles Symposium (IV)*, pages 502–509. IEEE.

Chronopoulos, T. (2017). The Rebuilding of the South Bronx after the Fiscal Crisis. *Journal of Urban History*, 43(6):932–959. Publisher: SAGE Publications Inc.

Cihlar, J., Pultz, T. J., and Gray, A. L. (1992). Change detection with synthetic aperture radar. *International Journal of Remote Sensing*, 13(3):401–414. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01431169208904045.

City of Memphesis (2023). Springs civic center - the city of memphis 2023.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Ding, K., Ma, K., Wang, S., and Simoncelli, E. P. (2020). Image quality assessment: Unifying structure and texture similarity. *IEEE transactions on pattern analysis and machine intelligence*, 44(5):2567–2581. Publisher: IEEE.

DOT, N. (2020). Nyc's temporary open restaurants program.

Dubey, A., Naik, N., Parikh, D., Raskar, R., and Hidalgo, C. A. (2016). Deep learning the city: Quantifying urban perception at a global scale. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pages 196–212. Springer.

Earley, A. (2023). Achieving urban regeneration without gentrification? Community enterprises and community assets in the UK. *Journal of Urban Affairs*, 0(0):1–24. Publisher: Routledge _eprint: https://doi.org/10.1080/07352166.2023.2229459.

García, I. (2019). Human Ecology and Its Influence in Urban Theory and Housing Policy in the United States. *Urban Science*, 3(2):56. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

Gibson, E. (2020). David rockwell designs kit-of-parts to build restaurants on streets following pandemic.

Glaeser, E. L., Luca, M., and Moszkowski, E. (2020). Gentrification and Neighborhood Change: Evidence From Yelp. *National Bureau of Economic Research*.

Harlem World Magazine (2023). Nyc mayor and dob commissioner unveil sidewalk shed removal plan from harlem to hollis.

He, J., Zhang, J., Yao, Y., and Li, X. (2023). Extracting human perceptions from street view images for better assessing urban renewal potential. *Cities*, 134:104189.

Hou, Y., Quintana, M., Khomiakov, M., Yap, W., Ouyang, J., Ito, K., Wang, Z., Zhao, T., and Biljecki, F. (2024). Global streetscapes – a comprehensive dataset of 10 million street-level images across 688 cities for urban science and analytics. *ISPRS Journal of Photogrammetry and Remote Sensing*, 215:216–238.

Hu, W. and Rosa, A. (2020). Outdoor dining in n.y.c. will become permanent, even in winter.

Huang, T., Wu, Z., Wu, J., Hwang, J., and Rajagopal, R. (2024). Citypulse: Fine-grained assessment of urban change with street view time series. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Ilic, L., Sawada, M., and Zarzelli, A. (2019). Deep mapping gentrification in a large Canadian city using deep learning and Google Street View. *PloS one*, 14(3):e0212814. Publisher: Public Library of Science San Francisco, CA USA.

Ito, K., Kang, Y., Zhang, Y., Zhang, F., and Biljecki, F. (2024). Understanding urban perception with visual data: A systematic review. *Cities*, 152:105169.

Kamalipour, H. and Dovey, K. (2019). Mapping the visibility of informal settlements. *Habitat International*, 85:63–75.

Kang, Y., Abraham, J., Ceccato, V., Duarte, F., Gao, S., Ljungqvist, L., Zhang, F., Näsman, P., and Ratti, C. (2023). Assessing differences in safety perceptions using GeoAI and survey across neighbourhoods in stockholm, sweden. *Landscape and Urban Planning*, 236:104768.

Koch, F., Kabisch, S., and Krellenberg, K. (2018). A Transformative Turn towards Sustainability in the Context of Urban-Related Studies? A Systematic Review from 1957 to 2016. *Sustainability*, 10(1):58. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

Kruse, J., Kang, Y., Liu, Y.-N., Zhang, F., and Gao, S. (2021). Places for play: Understanding human perception of playability in cities using street view images and deep learning. *Computers, Environment and Urban Systems*, 90:101693.

Lara-Hernandez, J. A., Coulter, C. M., and Melis, A. (2020). Temporary appropriation and urban informality: Exploring the subtle distinction. *Cities*, 99:102626.

Larkin, A., Gu, X., Chen, L., and Hystad, P. (2021). Predicting perceptions of the built environment using GIS, satellite and street view image approaches. *Landscape and Urban Planning*, 216:104257.

Lee, G. (2023). The state of storefronts: Alarming vacancy rates and rising rents during the pandemic.

Levine, D., Sussman, S., Yavo Ayalon, S., and Aharon-Gutman, M. (2022). Rethinking Gentrification and Displacement: Modeling the Demographic Impact of Urban Regeneration. *Planning Theory & Practice*, 23(4):578–597. Publisher: Routledge _eprint: https://doi.org/10.1080/14649357.2022.2117399.

Li, F., Zhang, H., Sun, P., Zou, X., Liu, S., Yang, J., Li, C., Zhang, L., and Gao, J. (2023). Semantic-SAM: Segment and Recognize Anything at Any Granularity. arXiv:2307.04767 [cs].

Li, X., Zhang, C., Li, W., Kuzovkina, Y. A., and Weiner, D. (2015). Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' socioeconomic conditions in Hartford, Connecticut, USA. *Urban Forestry & Urban Greening*, 14(4):751–759. Number: 4.

Liang, X., Chang, J. H., Gao, S., Zhao, T., and Biljecki, F. (2024). Evaluating human perception of building exteriors using street view imagery. *Building and Environment*, 263:111875.

Liang, X., Zhao, T., and Biljecki, F. (2023). Revealing spatio-temporal evolution of urban visual environments with street view imagery. *Landscape and Urban Planning*, 237:104802.

Liu, C. and Song, W. (2024). Mapping property redevelopment via GeoAI: Integrating computer vision and socioenvironmental patterns and processes. *Cities*, 144:104644. Publisher: Elsevier.

Liu, J., Xiao, L., and Wang, B. (2024). The varying effects of residential built environment on travel behavior of internal migrants and locals. *Travel Behaviour and Society*, 34:100692.

Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C., and Ren, F. (2021). Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities*, 110:103086.

Maassen, A. and Galvin, M. (2019). What Does Urban Transformation Look Like? Findings from a Global Prize Competition. *Sustainability*, 11(17):4653. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute.

McGinn, A. P., Evenson, K. R., Herring, A. H., Huston, S. L., and Rodriguez, D. A. (2007). Exploring Associations between Physical Activity and Perceived and Objective Measures of the Built Environment. *Journal of Urban Health*, 84(2):162–184.

Memphis, of Planning, S. C. D., and Department, D. C. P. (2020). Memphis 3.0. Technical report, Memphis and Shelby County Division of Planning and Development.

Meng, Y., Ho, H. C., and Wong, M. S. (2024). Changing associations of built environment with usage of urban space due to the covid-19 pandemic in the united states. *Cities*, 152:105205.

MLK50: Justice Through Journalism (2022). Feeling neglected, parkway village residents try to rebuild after white flight.

Moulton, J., Kassam, S., Ahmad, F., Amin, M., and Yemelyanov, K. (2008). Target and change detection in synthetic aperture radar sensing of urban structures. In *2008 IEEE Radar Conference*, pages 1–6.

Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L., and Hidalgo, C. A. (2017). Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences*, 114(29):7571–7576. Publisher: Proceedings of the National Academy of Sciences.

Nakamura, K., Derbel, B., Won, K.-J., and Hong, B.-W. (2021). Learning-Rate Annealing Methods for Deep Neural Networks. *Electronics*, 10(16):2029. Number: 16 Publisher: Multidisciplinary Digital Publishing Institute.

Opticos Design (2024). Memphis 3.0 comprehensive plan - opticos design.

Raciti, A., Lambert-Pennington, K. A., and Reardon, K. M. (2016). The struggle for the future of public housing in memphis, tennessee: Reflections on hud's choice neighborhoods planning program. *Cities*, 57:6–13.

Roth, M. (2006). Validating the use of Internet survey techniques in visual landscape assessment—An empirical study from Germany. *Landscape and Urban Planning*, 78(3):179–192.

Rui, J. and Cai, C. (2025). Plausible or misleading? evaluating the adaption of the place pulse 2.0 dataset for predicting subjective perception in chinese urban landscapes. *Habitat International*, 157:103333.

Saija, L., Santo, C. A., and Raciti, A. (2019). The deep roots of austere planning in memphis, tn: is the fox guarding the hen house? *International Planning Studies*, 25(1):38–51.

Sakurada, K. and Okatani, T. (2015). Change detection from a street image pair using cnn features and superpixel segmentation. In *Proc. Brit. Mach. Vis. Conf*, pages 61–1.

Salesses, P., Schechtner, K., and Hidalgo, C. A. (2013). The Collaborative Image of The City: Mapping the Inequality of Urban Perception. *PLoS ONE*, 8(7):e68400. Number: 7.

Savitch, H. V. (2010). What makes a great city great? an american perspective. *Cities*, 27(1):42–49.

Schelling, T. C. (1971). Dynamic models of segregation†. *The Journal of Mathematical Sociology*, 1(2):143–186. Publisher: Routledge _eprint: https://doi.org/10.1080/0022250X.1971.9989794.

Shi, W., Zhang, M., Zhang, R., Chen, S., and Zhan, Z. (2020). Change detection based on artificial intelligence: State-of-the-art and challenges. *Remote Sensing*, 12(10):1688.

Siff, A. (2023). Tired of scaffolding? nyc has new plan to get rid of long-standing sidewalk sheds.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Stalder, S., Volpi, M., Büttner, N., Law, S., Harttgen, K., and Suel, E. (2024). Self-supervised learning unveils urban change from street-level images. *Computers, Environment and Urban Systems*, 112:102156.

Steimer, J. and Steimer, J. (2024). Feeling neglected, parkway village residents try to rebuild after ‘white flight’.

van Veghel, J., Dane, G., Agugiaro, G., and Borgers, A. (2024). Human-centric computational urban design: optimizing high-density urban areas to enhance human subjective well-being. *Computational Urban Science*, 4(1).

Velasquez-Camacho, L., Etxegarai, M., and De-Miguel, S. (2023). Implementing deep learning algorithms for urban tree detection and geolocation with high-resolution aerial, satellite, and ground-level images. *Computers Environment and Urban Systems*, 105:102025.

Venter, Z. S., Barton, D. N., Gundersen, V., Figari, H., and Nowell, M. (2020). Urban nature in a time of crisis: recreational use of green space increases during the COVID-19 outbreak in Oslo, Norway. *Environmental Research Letters*, 15(10):104075.

Wang, Z., Ito, K., and Biljecki, F. (2024). Assessing the equity and evolution of urban visual perceptual quality with time series street view imagery. *Cities*, 145:104704.

Waters, D. (2022). The end of public housing in memphis, but not resident "relocations".

Wei, J., Yue, W., Li, M., and Gao, J. (2022). Mapping human perception of urban landscape from street-view images: A deep-learning approach. *International Journal of Applied Earth Observation and Geoinformation*, 112:102886.

Wentz, E. A., Anderson, S., Fragkias, M., Netzband, M., Mesev, V., Myint, S. W., Quattrochi, D., Rahman, A., and Seto, K. C. (2014). Supporting Global Environmental Change Research: A Review of Trends and Knowledge Gaps in Urban Remote Sensing. *Remote Sensing*, 6(5):3879–3905. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

Wiatkowska, B., Słodczyk, J., and Stokowska, A. (2021). Spatial-Temporal Land Use and Land Cover Changes in Urban Areas Using Remote Sensing Images and GIS Analysis: The Case Study of Opole, Poland. *Geosciences*, 11(8):312. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute.

Wiloandco (2024). Memphis 3.0 comprehensive plan five-year update to focus on zoning and new community investments - the city of memphis.

Xu, N. (2013). *Why Chinatown has Gentrified Later than Other Communities in Downtown Manhattan: A Planning History*. PhD thesis, Columbia University.

Zeng, Z. and Boehm, J. (2024). Exploration of an open vocabulary model on semantic segmentation for street scene imagery. *ISPRS International Journal of Geo-Information*, 13(5):153.

Zhang, F., Fan, Z., Kang, Y., Hu, Y., and Ratti, C. (2021). “Perception bias”: Deciphering a mismatch between urban crime and perception of safety. *Landscape and Urban Planning*, 207:104003.

Zhang, F., Zhang, D., Liu, Y., and Lin, H. (2018a). Representing place locales using scene elements. *Computers, Environment and Urban Systems*, 71:153–164.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., and Ratti, C. (2018b). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180:148–160.

Zhang, F., Zu, J., Hu, M., Zhu, D., Kang, Y., Gao, S., Zhang, Y., and Huang, Z. (2020). Uncovering inconspicuous places using social media check-ins and street view images. *Computers, Environment and Urban Systems*, 81:101478.

Zhang, Q. and Seto, K. C. (2013). Can Night-Time Light Data Identify Typologies of Urbanization? A Global Assessment of Successes and Failures. *Remote Sensing*, 5(7):3476–3494. Number: 7.

Zhou, H., Liu, L., Lan, M., Zhu, W., Song, G., Jing, F., Zhong, Y., Su, Z., and Gu, X. (2021). Using google street view imagery to capture micro built environment characteristics in drug places, compared with street robbery. *Computers, Environment and Urban Systems*, 88:101631.

Zitzlsberger, G., Podhorányi, M., Svatoň, V., Lazecký, M., and Martinovič, J. (2021). Neural Network-Based Urban Change Monitoring with Deep-Temporal Multispectral and SAR Remote Sensing Data. *Remote Sensing*, 13(15):3000. Number: 15 Publisher: Multidisciplinary Digital Publishing Institute.

Zukin, S. (1987). Gentrification: culture and capital in the urban core. *Annual review of sociology*, 13(1):129–147.