Essay on Urban Computing

Sumin Han
hsm6911@kaist.ac.kr

August 15, 2024

This essay encapsulates various reflections experienced during the completion of this doctoral program, which the author believes may hold even more significant value than the core argument of this dissertation—the existence and methodologies for processing Urban Embedding. While methodologies may evolve over time to better suit future advancements, the philosophical contemplations expressed in this preface represent the culmination of the author's lifelong efforts. The author hopes that someone seeking answers to similar challenges might reflect on the path taken and find some belief in the author's voice. Much like the philosopher Schopenhauer, the author aims to convey a message intended for future generations rather than the current era.

1 Urban Space Robot

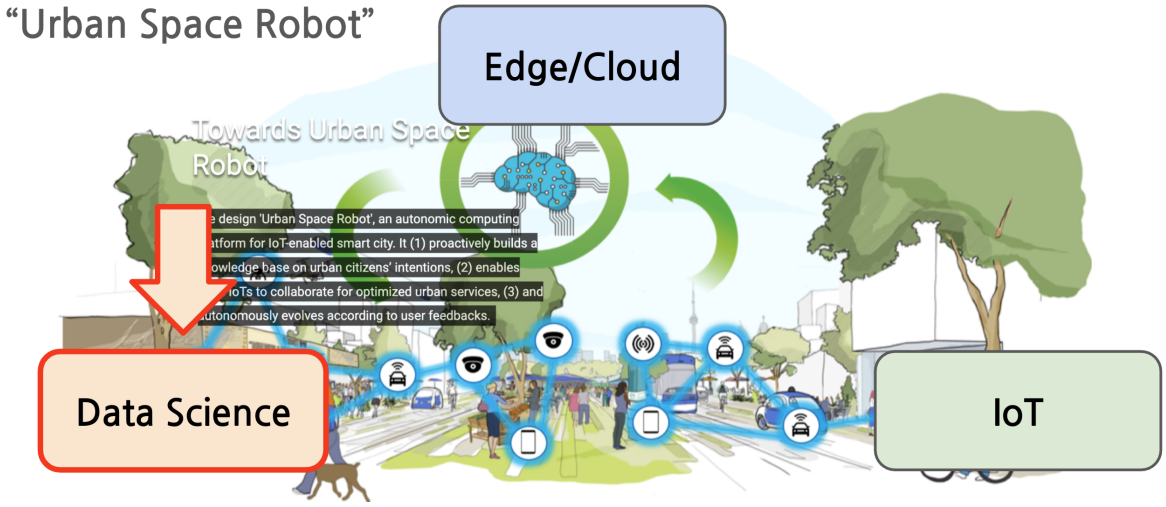

Figure 1: Urban Space Robot

What is an Urban Space Robot[1]? Unlike passive machines programmed to perform specific tasks, robots can autonomously provide appropriate solutions to resolve specific situations. An Urban Space Robot will prevent excessive urbanization, automatically construct necessary buildings, and continuously devise solutions to ensure the well-being and satisfaction of urban residents.

An Urban Space Robot is akin to a living, breathing animal in a balanced state, where factors like population density, traffic, and energy use must be managed efficiently.

In the near future, edge/cloud technologies will advance further, being widely applied across various sectors. These technologies will enable the use of IoT devices across multiple platforms, including smart traffic cameras, electric scooters, and electric vehicles. The vast amount of spatiotemporal urban data generated by these IoT devices will require solutions for understanding and processing it. This urban data will be learned through the Urban Embedding proposed in this dissertation and will be integrated into various deep learning-based AI applications.

[1]The Urban Space Robot is the core vision of the CDSN lab to which the author belongs – http://cds.kaist.ac.kr/

2 The City as Will and Representation

The author of this doctoral dissertation has been significantly influenced by Arthur Schopenhauer's (1788–1860) magnum opus "The World as Will and Representation (*German: Die Welt als Wille und Vorstellung*)".

> "The world is my representation. Man does not know the sun and the earth, but only an eye that sees the sun, a hand that feels the earth; this world that surrounds man is only a representation."
> – *The World as Will and Representation, Volume 1, Chapter 1*

Schopenhauer believed that the world could be understood through two concepts: representation and will. **Representation** encompasses everything we perceive through our senses. For example, when we see a number, what actually happens is that our eyes convert visual signals into neural electrical signals, transmitting them to the brain, which ultimately deciphers the value of the number. Schopenhauer regarded all these stages of perception as representations, and coincidentally, deep learning functions in a remarkably similar manner. In Figure 2, a model is proposed where a deep learning structure based on a Convolutional Neural Network (CNN) reads an image of a number and performs final classification, showing that at some latent vector stage, everything exists as an embedded representation in the form of a vector of real numbers.

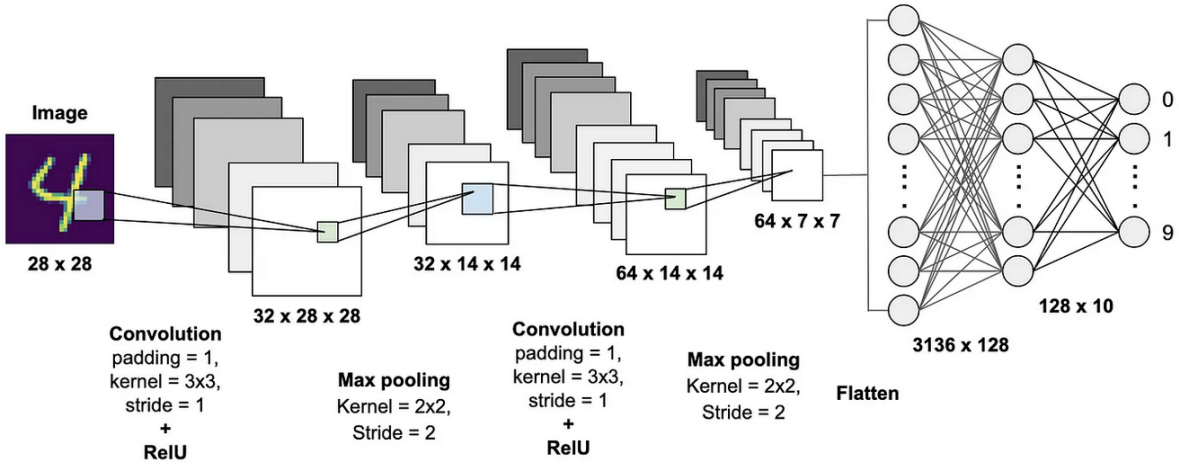

Figure 2: Convolutional Neural Network (CNN) and Embedded Representation

On the other hand, for Schopenhauer, the concept of **Will** extended beyond its general meaning to include notions such as 'desire', 'want', 'craving', 'pursuit', 'effort', and 'stubbornness'[2]. In Figure 3, we can sense the strong will of people wanting to visit the city. One of the peak areas is the Financial District of Boston, which could be seen as reflecting the material success desires of urban dwellers, as well as their values and will.

> "Just as the world is entirely representation on one hand, so it is entirely will on the other."
> – *The World as Will and Representation, Volume 1, Chapter 1*

When we consider the city divided into representation and will, many aspects can be understood more clearly. For instance, when comparing Asian cities and European cities, Asian cities (e.g., Korea, China) are often evaluated based on convenience, efficiency, and economy, all of which can be **represented** by material figures. Even personal values are often assessed quantitatively through education, performance, wealth, and other metrics. This could be because, in the past, the high population density in these regions provided fewer opportunities for education or jobs, leading to the establishment of quantitative systems that efficiently evaluate human value according to existing conventions, rather than qualitative evaluation criteria that consume a lot of mental energy based on a priori senses such as responsibility or mission.

[2]Refer to the footnote on page 44 of "The World as Will and Representation" (translated by Hong Sung-kwang)

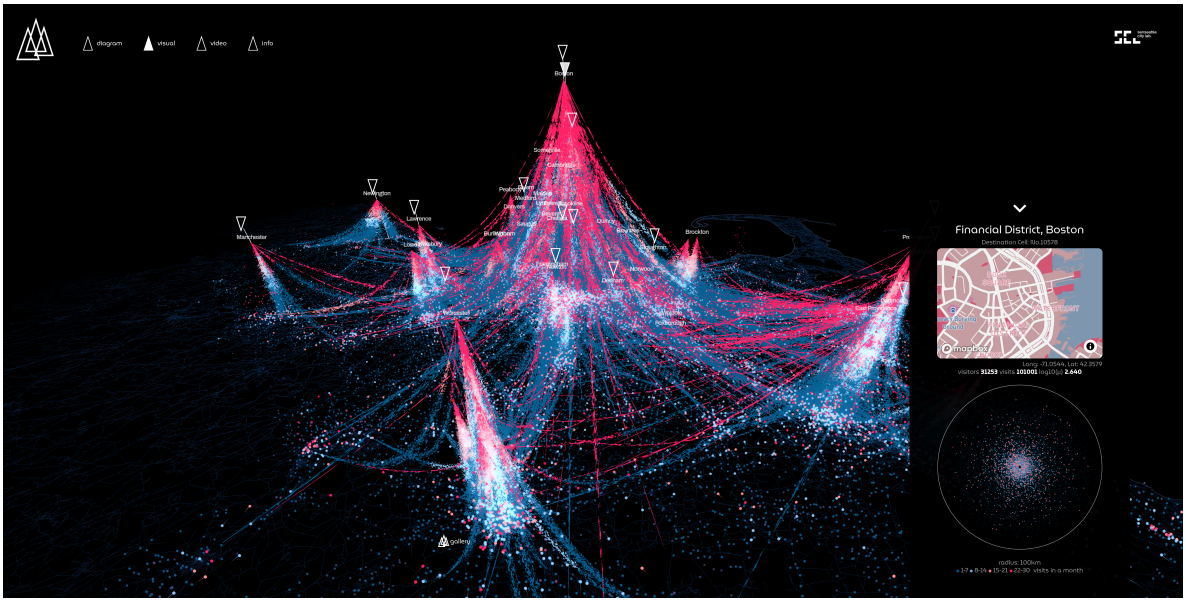

Figure 3: Human movement in WanderLust by MIT Senseable City Lab

In contrast, European cities are often associated with beauty, grandeur, religion, and historical value, with each city telling its own story. The citizens of these cities respect diversity, acknowledging that every individual has their own tastes and individuality, and they are cautious not to declare others' values as wrong. In other words, they respect the **will** that each individual pursues. Perhaps this writing was born out of a small world in Asia, specifically Korea, where individuality, will, and diversity have been stripped away, leaving only representation. It is written for scholars who have forgotten that pure will is essential in fields that contribute to humanity, such as science.

In conclusion, deep learning, the methodology utilized in this dissertation, can be applied only to issues within the realm of representation, not will. Its limitations are also clear. Just as artificial intelligence cannot perfectly recommend what clothes suit me best, it cannot judge why I find Baek Gun-woo's performance of the Appassionata more beautiful than Lang Lang's. The author would argue that the only intelligence that future generations may not be able to replace with AI is the primordial intelligence that enables one to decide for oneself which food makes one feel good, what clothes one wants to wear, what music one likes to listen to, and what needs to be done to be happy.

3 Contribution as a Computer Scientist

Research on cities is exceedingly complex and difficult because it involves studying virtually everything in the world, including politics, economics, culture, and race. It is challenging to accurately determine causality among the myriad variables that appear to be correlated, and there are too many factors to predict changes in a city. As mentioned earlier, cities often need to be understood not just through the **representation** of quantitative indicators but also through qualitative indicators that reflect **will**, such as the culture of the city and the values of individuals. For example, even in a country with a high GDP, it is difficult to predict future growth if there is a gloomy atmosphere where people do not communicate with each other and compete with each other over material possessions, such as luxury goods or real estate, which hold no real value. On the other hand, even in a country with a low GDP, there may be significant potential for future growth if the education level is high or if there are well-formed communities where people communicate actively and help each other more equally.

As a computer scientist, one can contribute to humanity in the realms of scientific art or engineering, as previously mentioned. A scientific artist can interpret qualitative meanings from quantitative indicators observed in the city, genuinely feel the social issues, and create works that allow people to empathize and communicate about these issues. MIT Senseable City Lab[3] has conducted numerous

[3] https://senseable.mit.edu/

studies on the visualization of such urban data. Meanwhile, an engineer can establish the foundations to efficiently address these issues using various solutions, including urban artificial intelligence. Urban spatiotemporal data, in particular, is often unrefined due to its complexity, and experiments conducted with unrefined data are likely to result in "garbage in, garbage out." For example, in the case of the traffic prediction model (Urban Activity-based Graph Convolutional Recurrent Neural Network: UAGCRN) published by the author[HPL+23], the author dares to comment that during the six years since the original Diffusion Convolutional Recurrent Neural Network (DCRNN)[LYSL18] was proposed, nearly 100 papers have been published, yet no researcher has scientifically justified the core graph connectivity weights of the originally proposed Graph Convolutional Recurrent Neural Network. Instead, there has been a focus on increasing the complexity of the deep learning model architecture to publish seemingly impressive papers.

Computer scientists must go beyond merely creating solutions that make transportation more convenient and continuously consider what urban residents truly want in a qualitative sense. For instance, while it is possible to train a city's traffic congestion prediction model using Mean Absolute Error (MAE), it is difficult to claim that the accuracy of the model genuinely enhances the actual satisfaction of urban residents with the service. In other words, 'Mean Absolute Error' is merely a representational tool for training the model. Although it can somewhat quantitatively measure the happiness of urban residents, it cannot definitively represent qualitative satisfaction. Therefore, urban space robots should be fundamentally designed based on the qualitative satisfaction of urban residents rather than quantitative indicators. A good city is not defined by any quantitative standard. Some people may want a place where they can enjoy tea and snacks while having a moment of peace, while others may desire a noisy place where they can go to a cocktail bar, play pool, and socialize with new people. A place where people live should not be a place where money is earned to buy real estate, but rather a happy society where one can communicate with the community, help those around them, and solve the problems of humanity together.

References

[HPL+23] Sumin Han, Youngjun Park, Minji Lee, Jisun An, and Dongman Lee. Enhancing spatiotemporal traffic prediction through urban human activity analysis. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 689–698, 2023.

[LYSL18] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018.