CSCI 1430 Final Project Report:
Seeing the Past: A Computer Vision Approach to University-led Gentrification

Perkins Pirates: Ronnie Shashoua, Seehanah Tang.
Brown University

Abstract

Inspired by the Brown Daily Herald article "'We didn't choose to be neighbors': A history of Brown's property impact on the East Side" by Katy Pickens and Finn Kirkpatrick [3], we created a model that takes inputs of the signs of certain Brown campus buildings. The model then outputs the old image of the building that stood there before Brown took over.

1. Introduction

Brown's campus expansion and university-led gentrification has drastically changed the landscape of College Hill by demolishing family homes, businesses, and historic buildings to make way for university housing and academic buildings, but as Brown continues to grow, there are few avenues for students to learn about the communities that preceded them. We want to show students the rich history of the area and impress upon them the magnitude of Brown's intrusion into communities.

We aimed to create a program that can take in an image of signage on a modern day building on Brown's campus and have it output images of that area prior to Brown's demolition of those historic buildings. First we needed to build a dataset of images of historical buildings, and a set of images of signage on those buildings. Then, we are utilizing traditional CV and optical character recognition (OCR) to identify the name spelled out on the signs in order to match it to the corresponding historical images.

Through our model, we hope that it demonstrates the drastic changes between the landscape of the current Brown campus and what used to be there. We attempt to highlight the disorientation and the sense of displacement that one might feel when trying to piece together the past and the present through images. This is an indirect call to the gentrification that has happened on College Hill that goes by un-noticed and not remembered.

2. Related Work

The inspiration for this project came from the Brown Daily Herald article "'We didn't choose to be neighbors': A history of Brown's property impact on the East Side" by Katy Pickens and Finn Kirkpatrick [3]. We also utilized the archives of the John Hay Library, including two theses that helped us understand historical context and find new sources to look for data for our datasets.

The first was "Competing Visions: Historic Preservation and Institutional Expansion on Providence's East Side, 1937-1966" by Peter Schermerhorn [1] which primarily helped us find information and sources regarding the 1951 expansion that led to the construction of Wriston Quadrangle. The other was "A Home for the Liberal Ideal: Brown University Housing Policy & the East Side of Providence, 1937-1997" by Nathaniel Pettit [2], which primarily helped us find information and sources regarding the 1957 expansion that led to the construction of Keeney Quadrangle.

Other sources include directories from the Rhode Island Historical Society, the Wriston Papers from the John Hay Library, miscellaneous archive boxes from the John Hay Library, and the digital photo collections of the Providence Public Library.

For the model, we used python for the software, which included certain packages such as pytesseract and cv2.

3. Method

A substantial part of this project was data collection, which included accessing Brown University library resources as well as online databases and Providence historical organizations.

John Hay Library. First, we pulled boxes of archival data to look through in the John Hay library, such as the Wriston Papers. This involved combing through boxes of images preserved by the university and searching for images which seemed to be located in the general area of today's campus. Many images were captioned using street names,

but some used addresses that no longer exist and some refer to houses by name (i.e. Wroth House). Therefore, to understand whether these photos were relevant to our work we had to research the exact locations of old images, which involved cross-referencing with other images of buildings and aerial photos, official university correspondences, and historic registries and ownership records.

Non-University sources. We also found images and reference sources from Rhode Island Historical Society directories, Providence Public Library digital photo collections, and the Brown University repository. The Rhode Island Historical Society directories contained information that we were able to use to match the sites of our historic images to the sites of current day buildings. They also had old photos organized by street in their archives, which was helpful in supplementing images found in John Hay. The Providence Public Library digital photo collections contained aerial shots that, along with aerial photos taken from the John Hay collection, were used to help pinpoint the location of the buildings in the old photos that we found.

Current day data. Once we knew the physical location where every one of our historical images were taken, we were able to create a list of current day buildings that occupy the spaces where those buildings once stood. Some images covered the physical space of multiple current day buildings, which was taken into account. Then we went to build up a dataset of images of the signage on these buildings that we knew corresponded to our historical data, which just involved going to every single one to take images of the signs with their names from several different angles.

For our training dataset, we took pictures of building signs from different angles. Our sample size was 92 images of 18 buildings.

Then, for developing our text extraction model, we built off of the python library pytesseract. The baseline accuracy without any preprocessing was low, below 50%. We tried different combinations of preprocessing, and the final model performs preprocessing by converting the image to black and white, applying median and Gaussian blur, and inverting the final image. To map words to the building, we created a dictionary of building to its characteristics (e.g. name, year it was built, location). If any of the characteristics are found in the text extracted, we return the building name, or key of the dictionary. The code below highlights this approach.

```
buildings_dict = {
        "Barbour":
            {"Barbour","Charlesfield",
               "1966"},
        # omitted #
        "Keeney":
            {"Keeney","Benevolent
                ","1957","Bronson"}
    }

for building in buildings_dict:
    for string in buildings_dict[
        building]:
        if string.lower() in text.lower
            ():
            return building

return "default"
```

4. Results

We were able to acquire images for the sites of Wriston Quadrangle, Keeney Quadrangle, Barbour Hall, Macmillan Hall, Nelson Fitness Center, New Pembroke, Perkins Hall, Rockefeller Library, and Wellness Center.

We wanted to emphasize the disorientation and the feeling of displacement we felt looking through images of a completely different landscape that once stood on College Hill.

Below are example images that show a side-by-side comparison of pre-1951 Wriston with the current site today.

Figure 1. A view down Benevolent Street from Thayer, a look down into a bustling commercial street that no longer exists after being demolished for the construction of Wriston Quadrangle.

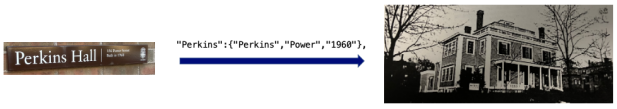
Figure 2. A view from further up Thayer Street, depicting the life that once existed in this neighborhood. In the current day, these businesses have been replaced by a brick wall and the Sharpe Refectory.

Our model achieved a final accuracy of around 62% (Table 1 In general, it performs well on direct, close-up photos of signs, especially signs that are the standard brown signs

with white letters. Below is an example of the model working accurately.

Building	Accuracy
Barbour	0.6
Buxton	0.25
Chapin	0
Diman	0.4
Goddard	1
Harkness	0.17
Hegeman	1
Keeney	0.86
Macmillan	1
Marcy	0.75
Nelson	0
New Pembroke	0.6
Olney	0.83
Perkins	0.8
Ratty	1
Rock	0.5
Sears	0.4
Wayland	0.4
Total	**0.62**

Table 1. Results. Sign recognition accuracy per building and cumulative. Total accuracy calculated by weighting accuracy for building with the number of images we have for that building



We found the pytesseract algorithm to be quite unreliable, depending on the angle, quality, and background of the image. Thus, next steps would be to train our own model (e.g. using CNNs). Unfortunately, since data collection was the most challenging and time-consuming part, we were unable to implement our own self-learning model.

4.1. Technical Discussion

Our text extraction model built off of tesseract, which is a popular optical character recognition (OCR) engine with open-source code. OCR uses artificial intelligence for text search and its recognition on images. Tesseract uses a two-step approach otherwise known as adaptive recognition. First, it reads in the image and recognizes characters, then in the second stage, puts any letters it wasn't secure in by letters based on context.

It is interesting that such a well-developed an popular python library only produced mediocre results. When only using tesseract, the produced text was usually incomprehensible.

Since a huge part of the project was data collection and research, we tried to use a simple and concise model. Thus, in doing so, we were able to get average results (62%). Using more complex models might take longer to train, but would probably lead to better results. By making our model more simple, it leads to better ability to replicate our results and reproduce our code. It is also more time efficient.

5. Conclusion

We created a model that shows someone from the present what someone in the past would see if they were standing in front of certain Brown University campus buildings and dorms just through a picture of the building's sign.

In the last century, through property expansion and student housing development, Brown has contributed to the displacement of low-income and non-white residents that get priced out by students and Brown's real estate developments and projects. Although a lot of the development Brown has created has encouraged economic development in the neighborhood, it has also spelled the end of many small, family businesses and torn down historic homes. By putting these two sets of data side by side, we hope to encourage users of our program to imagine themselves as a previous resident of the neighborhood whose home and community got bought up by Brown and demolished. We are attempting to emphasize the disorientation and the sense of displacement that those who call College Hill their home may experience when they remember what used to be before.

Our project is an attempt to bridge the gap between the before and the now. We hope we can continue to shed light on this issue, and encourage students to educate themselves on the communities that they are entering, both when they move into dorms and when they move into off-campus housing. Although many students can live their four years here in a bubble on campus, we aim to expand the knowledge of the general Brown student on the history of the space we now occupy. As decision makers in the city of Providence and future residents, those affiliated with Brown, both students and administration, should educate themselves on the historical context of the university and the space it takes up in this community.

Appendix

Team contributions

Person 1 This person took charge on dataset collection. They reached out to the BDH article authors, John Hay librarian Raymond Butti, the Rhode Island Historical Society (RIHS), Providence Preservation Society, and more. She also researched various image text extraction algorithms.

Person 2 This person took charge on coding. She implemented tesseract and tried various pre-processing techniques to improve accuracies. She also helped by going to John Hay and RIHS to help find old campus building data.

References

[1] Peter Schermerhorn. "Competing Visions: Historic Preservation and Institutional Expansion on Providence's East Side, 1937-1966". PhD thesis. Providence, RI: Brown University, May 2005.

[2] Nathaniel Pettit. "A Home for the Liberal Ideal: Brown University Housing Policy the East Side of Providence, 1937-1997". MA thesis. Providence, RI: Brown University, May 2020.

[3] Finn Kirkpatrick and Katy Pickens. "'We didn't choose to be neighbors': A history of Brown's property impact on the East Side". In: *Brown Daily Herald* (2022).