Categorising Buildings in Lausanne Based on Façade Material

Youssef Belghmi, Hamza Morchid
Machine Learning Project, Department of Data Science, EPFL, Switzerland

***Abstract*—This project is part of a collaboration with HERUS (Laboratory for Human-Environment Relations in Urban Systems) aimed at harnessing the capabilities of machine learning to solve a critical urban challenge. Our main goal is to classify buildings in Lausanne based on the composition of their facade materials, using advanced machine learning techniques. Our project aims to provide the HERUS laboratory with crucial data to inform their future initiatives in the expansion and transformation of Lausanne's built environment. This research represents an important step towards a more in-depth understanding of the urban structure of Lausanne, paving the way for informed decisions in urban planning and heritage conservation.**

I. Introduction

In the context of increasing urbanization and sustainable urban development, the analysis of the characteristics of buildings and their harmonious integration into the urban environment is of paramount importance. This report documents our collaborative research project with HERUS, which aims to solve a crucial challenge related to the architecture and composition of buildings in the city of Lausanne. The HERUS laboratory has clearly established its objectives in anticipating the planned expansion of the built environment in Lausanne, necessary to meet the needs of residents by 2050. This expansion involves the reuse of existing building components in new construction. To realize this vision, it is imperative to gain a thorough understanding of the materials currently present in the current stock of buildings.

It is in this context that our project takes off, by exploiting advanced machine learning techniques to accomplish this complex and essential task. In this report, we will explore our research methodologies, tools and results in detail.

II. Data Collection

Data collection in a machine learning project is crucial because it involves gathering essential information to train and test a model. In our project to classify Lausanne buildings based on their facade materials, we collect images of these facades to train our model to recognize them. The quality and diversity of training data are essential because they determine the accuracy and performance of classification. High-quality data ensures reliable results, while insufficient or biased data can lead to inaccurate predictions.

A. Dataset of Lausanne Buildings

As part of our project, we were responsible for collecting, on our own, images of the facades of buildings in Lausanne. However, the HERUS laboratory also offered two valuable resources for this task :

1) Swiss Real Estate Database [1] : This resource offers detailed information on buildings in Switzerland, taken from the official directory of building addresses.
2) Google Street View via the Google Maps API : Using this API allows us to download images of building facades at specific geographic locations in Lausanne.

By strategically leveraging these two resources, we managed to generate a dataset consisting of images captured from Google Street View, specifically of building facades in Lausanne. We accomplished this task using the geographic coordinates available in the official Swiss buildings database.

B. Data Cleaning and Labeling

Once we have collected our images, the next step is to conduct a brief analysis to better understand the nature and characteristics of the 1437 images we have collected.

One of the first observations made regarding the images obtained is the presence of blank images. These images appear blank because Google Street View did not capture imagery for the corresponding building. These images are not usable for our model, thus, we decided to remove them from our dataset. This ensures that only the remaining 918 relevant and usable images are used for our study.

However, taking into account the laboratory's request to classify building facades into three classes (brick, stone, mixed materials), it is clear that a data imbalance is present in our set. Of the 918 usable images, less than 10 belong to the "stone" class, the same for the "brick" and "mix" (facade with at least two distinct materials) classes. The majority of facades should be classified in a category other than the three initially proposed by the laboratory. Thus, we have made the decision to broaden the classification classes by adding a new class, "concrete", corresponding to buildings with a concrete facade.

This is not the only limitation of our dataset. Indeed, the analysis of images obtained using the Google Street View API reveals considerable variability in the quality and clarity of the images. Several factors contribute to this variability. Some images are blurry, possibly due to Google Street View camera movement or focus issues during capture. Many other images are partially or completely obstructed by elements such as trees, vehicles or fences, thereby largely or entirely obscuring the facade of buildings. In some cases, coordinates supposedly representing a building result in images where the building in question is not visible or is poorly positioned within the photo frame. Additionally, sometimes images show two neighboring buildings in a single frame, which can be confusing during analysis and classification. Finally, some images are taken from too far away or from angles which do not allow a clear and complete vision of the facades. These imperfections are likely to have a significant impact on the performance of our model.

To deal with this problem of images of facades that are difficult to identify, we have chosen to add an additional class, called ”other”, which will group together images not showing buildings, those showing facades of buildings with significant obstacles in front, or those with facade materials different from the other four classes, such as wood or glass.

All that remains is to complete the last essential step of our process, which is to manually label the buildings by classifying them into the class that corresponds to their respective facade construction material. This labeling task is crucial to creating an accurately and reliably annotated dataset, which will serve as the training base for our machine learning model. Thanks to this manual classification, each image will be assigned to one of the classes we have defined, whether ”brick”, ”stone”, ”mix”, ”concrete”, or ”other”.

To further illustrate, here is an example of the best photos taken from Google Street View for each of the five classes we will study :

Figure 1. Examples of facades for each class

C. Creating a New Dataset

This in-depth analysis of our dataset provided us with crucial insights, highlighting the diversity of the images, their varying quality, and their limited representation of the scope of our project. The great heterogeneity of the images is evident, with almost 90% of them requiring classification into the “concrete” and “other” classes, leaving a limited amount of data for the other classes. Poor image quality, as we explained previously, will inevitably have a direct impact on the accuracy and efficiency of our model.

To resolve these challenges and work with a more complete and balanced dataset, we made the decision to create a new dataset. This new dataset will be made up of the best images from our original collection, obtained via the Google Street View API, and will be supplemented by additional images. We will manually collect screenshots from Google Street View, allowing us to overcome the obstruction and perspective issues we encountered with the API. Additionally, we will incorporate images of building facades from datasets available online to further enrich our dataset. We notably used some of the images used for a Czech study [2], the dataset taken from the University of Bonn [3], and the few images found on a github project [4].

This methodical approach not only gives us the promise of a more balanced distribution of data within each category, but it also ensures better control over image quality, establishing a stronger foundation for training our model.

However, it is important to note that collecting images of building facades is complex. Despite our commitment and all our efforts, we only managed to gather around thirty images for each of the classes. This limitation highlights the scarcity of data available for our project, prompting us to work carefully to maximize the use of these valuable resources.

III. Data Preprocessing

Data preprocessing is a fundamental step in the data preparation process, it involves cleaning, transforming and organizing the collected raw data in order to make it suitable for the application of machine learning techniques. It is essential to ensure data quality, consistency and relevance, as poor quality raw data can lead to incorrect or biased results. In sum, data preprocessing creates a solid foundation for analysis and modeling, thereby improving the accuracy, generalization, and performance of data models.

Data augmentation is a data preprocessing technique commonly used in machine learning, particularly for computer vision tasks such as image classification. It consists of increasing the size of the dataset by creating new variations of existing data, while maintaining the same label or class. The main goal of data augmentation is to improve model generalization by exposing the model to a greater diversity of data during training, which can reduce overfitting.

A. Data Partitioning

Before moving on to data augmentation, we need to distribute the images of each class of the dataset into the training and testing folders according to certain predefined proportions. We will allocate 70% of the dataset for training and 30% for testing. This balanced distribution aims to provide the model with 70% of the data for training, thus allowing efficient acquisition of the characteristics of the different classes. Then, the remaining 30% is used to evaluate the model’s performance on data not seen during training. This helps estimate the model’s ability to generalize to new observations, while limiting the risk of the model overfitting to specific training data. This step of strategic data distribution is crucial to ensure an accurate assessment of the model’s ability to perform in real and unknown situations.

B. Data Augmentation

Now we can move on to the data augmentation phase, which involves applying transformations to the images before using them for training and testing the model. The data transformations we apply, to both the training data and the test data, are resizing (to ensure a uniform size, given that our images come from varied sources and have different dimensions) and normalization (to improve training stability and reduce bias introduced by differences in the scale of pixel values). To do this, we calculated the mean and standard deviation values from the training and testing data, which will be used to normalize the data, thus bringing it to a scale centered on the mean and with unit variance .

Due to the limited amount of images available in our dataset, applying various transformations to training images is essential to improve the robustness of the model and its ability to generalize. We implemented a variety of transformations, including horizontal and vertical flips, translations, random cropping, zooms, grayscale conversion, shearing, Gaussian blur, and brightness adjustments. For example, from a building facade image in the training set, we can randomly obtain a transformed image as shown below :

Figure 2. Examples of random transformations made on an image

In this way, data augmentation allows us to simulate different scenarios and enrich the diversity of images provided to the model, despite the limited amount of images in our training set. This diversity is essential so that the model can effectively learn to recognize patterns and generalize to new data.

IV. Summary

The aim of a scientific paper is to convey the idea or discovery of the researcher to the minds of the readers. The associated software package provides the relevant details, which are often only briefly explained in the paper, such that the research can be reproduced. To write good papers, identify your key idea, make your contributions explicit, and use examples and illustrations to describe the problems and solutions.

Acknowledgements

The author thanks Christian Sigg for his careful reading and helpful suggestions.

References

[1] Federal Office of Topography, swisstopo, "Official directory of building addresses," 2022.

[2] R. Tyleček and R. Šára, "Spatial pattern templates for recognition of objects with regular structure," 2013.

[3] F. Korč and W. Förstner, "eTRIMS Image Database for interpreting images of man-made scenes," 2009.

[4] R. Gadde, R. Marlet, and N. Paragios, "Learning grammars for architecture-specific facade parsing," 2016.