Using Remote Sensing to Track Sidewalk Accessibility Issues to Improve Urban Planning

Lynux Louie K. Ansay
De La Salle University
Manila, Philippines
lynux_ansay@dlsu.edu.ph

Nicolas Steven T. Sy
De La Salle University
Manila, Philippines
nicolas_sy@dlsu.edu.ph

Henric Ivan L. Tay
De La Salle University
Manila, Philippines
henric_tay@dlsu.edu.ph

Rafael Nicolas A. Topacio
De La Salle University
Manila, Philippines
rafael_topacio@dlsu.edu.ph

Unisse C. Chua
De La Salle University
Manila, Philippines
unisse.chua@dlsu.edu.ph

Briane Paul V. Samson
De La Salle University
Manila, Philippines
briane.samson@dlsu.edu.ph

ABSTRACT

The lack of open datasets regarding sidewalk accessibility has reduced the opportunity to assess and address inaccessible sidewalks. In this project, we assessed the features of Metro Manila sidewalks to determine its accessibility. The features are limited to sidewalk width, surface type, and common obstructions found on local sidewalks. By collecting Google Street View images of sidewalks in Makati and Manila city, we labeled common sidewalk obstructions and crowdsourced for accessibility score through a platform we developed. Additionally, we ran a regression model to identify the most influential features that affect sidewalk accessibility. Our results show that a rough surface and obstructions, such as construction materials, street vendor stands, and tricycles, negatively affect sidewalk accessibility. Oppositely, the presence of wide sidewalks and trees positively affect sidewalk accessibility. Lastly, we recommend using the platform to collect more data on sidewalk accessibility and to extend the scope to other cities in Metro Manila.

CCS CONCEPTS

• **Machine Learning**; • **Computer Vision** → *Object Detection, Semantic Segmentation*;

KEYWORDS

Urban Studies, Remote Sensing, Computer Vision, Crowd-sourcing.

ACM Reference Format:
Lynux Louie K. Ansay, Nicolas Steven T. Sy, Henric Ivan L. Tay, Rafael Nicolas A. Topacio, Unisse C. Chua, and Briane Paul V. Samson. 2021. Using Remote Sensing to Track Sidewalk Accessibility Issues to Improve Urban Planning. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/1122445.1122456

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

A city's urban design is responsible for shaping its character through the arrangement of infrastructure, transportation systems, and public areas [1]. Through urban design, a city becomes functional for its inhabitants through the network of streets, sidewalks, and roads that improve urban accessibility. However, majority of regions have been designed to favor vehicles rather than inhabitants. In the Philippines, the increase in paved roads from 22,468.71 km in 2009 to 32,087.08 in 2019 is proof of this bias [14]. These projects further incentivize the use of vehicles as the primary mode of transportation within a city at the cost of the well-being of its inhabitants [21]. Traffic congestion, public transportation, and pedestrian utilities have become a large factor in determining a city's accessibility. The impact of these factors on a city's accessibility has led urban designers to further look into urban accessibility.

2 RELATED WORK

We present the related studies that discuss urban studies, accessibility standards, remote sensing, crowdsourcing practices, semantic segmentation, and object detection.

2.1 Urban Studies

A significant part of urban accessibility has to do with sidewalk design and purpose. A well-designed sidewalk is one where a person with disability (PWD) can safely cross a street or make use of public transportation [2]. Urban planners are expected to design sidewalks in an accessible manner to prevent additional costs needed for reconstruction and adjustments. Architectural firm, Ayres Associates, mentions certain factors which go into the design of new sidewalks including placement, compliance with local laws, width, drainage, and driveway design. In our research, we aim to determine the level of accessibility of certain sidewalks in the Philippines, hence the importance we place on identifying the characteristics of a good sidewalk. With the data we acquire from our research, we hope to create a platform where data on accessibility can easily be obtained by urban planners capable of making a significant impact on accessibility in the Philippines.

2.2 Measuring Accessibility

Accurately measuring accessibility has long been an issue in the field of urban planning and infrastructure as a universal guideline for measuring accessibility has not been established. However, there are still many credible metrics from different organizations, one of them is the American Disabilities Act (ADA). The ADA 2010 Standards for Accessible Design, also known as the 2010 Standards, encompasses all the benefits and compromises required of public utilities to cater to disabled persons. The United States Department of Justice (2010) released the 2010 Standards as a metric for measuring accessibility in urban settings. It include provisions on how to properly construct facilities and public spaces to cater to all types of pedestrians. These include designing a network of travel by placing sidewalks in appropriate locations, improving the quality of surfaces to be slip-resistant in any weather, and the surface should be even and absent of irregularities to ensure the safety of the pedestrians walking on it. In this study, we aim to evaluate the sidewalks located in the Philippines through its design and usability for people with disabilities.

2.3 Remote Sensing

Remote sensing is the use of sensors to capture images on the Earth's surface to gather information of large geographical areas without the need to be physically present at the location [20]. Technologies such as drones, air crafts, and Earth orbit satellites are able to capture overhead images or top view level of the cities. However, the lack of ground-level details make it a difficult task to gather more information on the geographical area. Fortunately, the growing accessibility to different sources of geo-tagged data makes it possible to fuse remote sensing imagery with data of different modalities and observations. In this study, we utilized Google Street View (GSV) to gather street level images of Manila and Makati. It serves users with street-level panoramic imagery captured in thousands of cities worldwide, which makes it possible to observe street scenes in big cities and thus provides proximate sensing ability and ground-level details that overhead images lack [4].

2.4 Crowdsourcing

Crowdsourcing is the use of IT to outsource any organizational function to a strategically defined population of human and non-human actors in the form of an open call [11]. Crowdsourcing can be used in rapidly acquiring data with the use of volunteers and paid workers. In this paper, we explored how crowdsourcing influenced contemporary studies on location based data gathering especially with regards to accessibility [15]. With these studies, we looked into the three key factors that are crucial in building a successful crowdsourcing platform: *scalability*, *availability*, and *data quality*

2.5 Object Detection

In defining object detection, we must first define the concept of image classification and object localization. Image classification is the task of assigning a class label to an image. It takes an image as an input, and outputs a class label of the objects present in the image. On the other hand, object localization refers to identifying the location of one or more objects in an image by drawing a bounding box around them. Object detection combines the tasks of both image classification and object localization, classifying and localizing one or more objects in an image [3]. In this paper, we utilized the YOLO or You Only Look Once object detection algorithm introduced by Redmon, Divvala, Girshick, and Farhadi (2016). The most recent version of YOLO is the YOLOv5 introduced by Glenn Jocher [7]. It is built on a PyTorch framework, and is able to achieve 140 frames per second inference time on videos. The model is trained on the COCO 2017 train/val/test dataset which contains a list of 91 objects [18]. It achieves a score of 68.8 mAP@50 on its YOLOv5x (largest weights) model. In this study, we aim to use the YOLOv5 object detection model to identify objects situated on sidewalks. This is also to help reduce the tasks that our volunteers have to label when looking for objects or obstructions on the sidewalk.

3 RESEARCH METHODOLOGY

The development of the study follows a 7-stage pipeline namely image collection, preparing data for crowdsourcing, crowdsourcing platform development, crowdsourcing, sidewalk width collection, accessibility score modeling and validation, and improving the pipeline of Atlas. Each stage of the pipeline will be discussed thoroughly in the following sections.

3.1 Image Collection

An initial collection of Google Street View images was conducted in order to provide data to our research processes, namely running YOLOv5 on the GSV images and preparing it for deployment to the website. We collected unannotated GSV images representing streets in Makati and Manila City. For this task, we utilized PostgreSQL, PostGIS, OpenStreetMap (OSM), and Google Street View to implement a semi-automatic method of collecting images.

3.1.1 Retrieving the Road Network. Since Google Street View requires the longitude and the latitude of a specific location, we decided to utilize publicly available geospatial data to get the entire road network of Metro Manila. We were able to download the entire region of Metro Manila from Planet OSM[1] and loaded this into a PostgreSQL database using osm2pgsql[2]. The Planet OSM extract provides multiple tables, and we focused on extracting the data from `planet-osm-line` for road networks. This specific table provides all the `LINESTRING` geometries, which is composed of a series of longitude and latitude pairs that may be extracted from the database using the PostGIS functions. We rounded off the latitude and longitude pairs to the fourth decimal place and selected unique values based on the decimal degree precision values[3] to reduce redundant pairs. For the entire Metro Manila region, we were able to extract a total of 203,949 latitude and longitude pairs.

3.1.2 Google Street View API for Street-Level Imagery. After gathering list of coordinates, we were able to use the Google Street View API to get static images of roads. For every coordinate, we collected four (4) images representing the north, south, east, and west perspectives. We then pitched the images slightly downwards to mimic an eye-level perspective. This totals to 815,796 static road images. Keep in mind, with our limitations, some of these images are unusable due to some factors such as Google Street View's privacy policy with regards to private roads. We also limited our study to collecting street view images only in Makati and Manila.

[1] https://planet.openstreetmap.org/
[2] https://osm2pgsql.org/
[3] https://en.wikipedia.org/wiki/Decimal_degrees#Precision

With that, we were able to collect 109,832 street view images and stored them in the Azure Database. Figure 1 is a sample Google Street View image that is stored in our Azure Database.

Figure 1: Sample Google Street View image of the road with the sidewalk present in frame

3.2 Preparing Data for Crowdsourcing

Before the Google Street View images are uploaded to Atlas for crowdsourcing, we first needed to train our version of YOLOv5 to detect objects that are situated on Philippine sidewalks. This is to reduce the need for our crowdworkers to locate obstructions and draw a bounding box when using Atlas, our crowdsourcing platform.

3.2.1 Detecting Objects on Sidewalks. With the streetview images we collected, we used YOLOv5 to detect objects found on the images. Since YOLOv5 is trained after the Common Objects in Contect (COCO) 2017 train/val/test dataset which contains a list of 91 objects, we expect it to detect objects that do not affect the accessibility of the sidewalk. We utilized the existing COCO 2017 Dataset and filtered its classes to only detect objects located on sidewalks seen in Table 1. We also added to our list other objects that are commonly found on Philippine sidewalks. It is important to note that the objects we have added still need a number of instances for the YOLOv5 model to be able to detect them.

Objects from the COCO 2017 Dataset	Additional Objects Added
Bicycle	Tricycle
Car	Street Vendor Stand
Motorcycle	Street Sign
Traffic light	Cracked Pavement
Fire hydrant	Construction Materials
Stop sign	Utility Post
Parking meter	Lamp Post
Bench	Trees
	Curb Ramp

Table 1: List of classes for our YOLOv5 model to detect

3.2.2 Crowdsourced Manual Annotations to Train YOLOv5. In order for the YOLOv5 model to detect the additional objects we added we looked for volunteers who were willing to participate in the annotation process. A user guide was created to inform users how to use Makesense.ai, a free online image labeling tool which could export labels as usable files for training YOLOv5. Within the manual, we included tips on how to properly label the objects, such as not overlapping the bounding box too much around the object, so that irrelevant pixels would not be included in the label. In handling non-sidewalk images, we asked participants to note down these images and submit it together with their annotations. In total, we collected 12,732 annotated images.

3.3 Atlas Development

Atlas is an online crowdsourcing platform we developed to facilitate the crowdsourcing of accessibility scores by presenting pre-labeled sidewalk images to volunteers. Figure 2 shows the system architecture of the study. The data collection and processing scripts are separated from the main system, but all data collected were stored in a single database that Atlas was connected to.

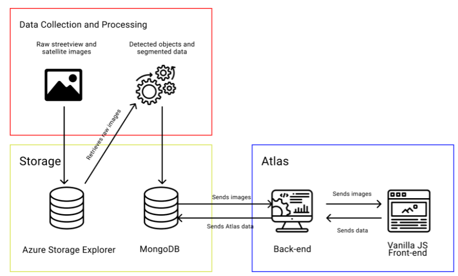

Figure 2: Diagram of Atlas System Architecture

In terms of storage, we used Azure Storage Explorer to store all raw images and blob data. MongoDB was used to store all data collected from Atlas, such as sidewalk accessibility scores, annotations, and participant information (non-personally identifiable information). MongoDB is also used to store the links to the pre-labeled images and their annotations from YOLOv5 and OCR (ResNet-101).

For the backend of the Atlas system, we use Next.js API routes for the API responsible for querying the data and saving the user inputs to the database. These API routes retrieve the pre-labeled images from MongoDB, authenticate users, register users, and save user annotation to the database.

For the frontend, we used Next.js and Tailwind CSS for our styling. Next.js is an open-source development framework built on top of Node.js enabling React based web applications functionalities such as server-side rendering and generating static websites. Tailwind CSS is a utility-first CSS framework that excels in rapid styling of HTML elements. We built the annotation tool by building on top of React Picture Annotation library and customizing it to our needs.

MongoDB was used as the database to store the data needed and obtained from Atlas. The database contains three tables, namely *User*,*Images*, and *Annotations*. The User table contains all relevant information in building a simple demographic of our crowdworkers and user credentials for authentication. The images table contains data on pre-processed images: this include the blob link url, longitude and latitude, direction, city, and the bounding boxes for the objects that our computer vision model detected. Lastly, the Annotations tables are where user annotations are stored when they contribute to our platform. Annotations are connected to the User table to connect an annotation to a specific user. In addition, an annotation also has two relationships to the Image table: here we get the imageID and the detected objects the user selects. The exact attributes of each document are presented in Figure 3.

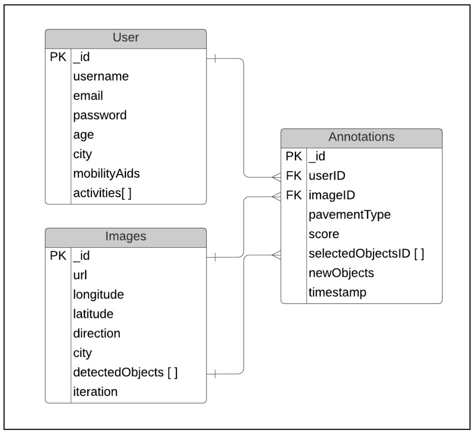

Figure 3: The Entity Relationship Diagram of the database schema used to store image and annotation data in Atlas

By developing Atlas, we were able to have a crowdsourcing platform wherein registered users could label, score, and further annotate sidewalks. We uploaded 12,372 preliminary streetview images during our data preparation to the crowdworkers. The whole development process took around five months and it was deployed and hosted on Vercel. [4]

3.4 Crowdsourcing through Atlas

3.4.1 Registration. For Atlas, users can access our online platform where they can register an account with their email. When creating an account, the user will be asked for the following participant information: *city of residence*, *age*, *use of mobility aid (wheelchairs, walkers, canes)*, and *frequency of using sidewalks in a week*. We collect this information in order to know the demographic of our users. Additionally, we want to explore the relationship of user demographic and the scores they provide. The account will be used by Atlas as a tracker on the images that the volunteer annotated. Additionally, the email collected from the users is strictly used for means of communication.

3.4.2 Identifying Objects. After creating an account, the volunteer can proceed to score images on Atlas. Annotators can select among 5, 10, 15, and 20 as the number of images they choose to annotate. This is to provide them with a sense of commitment and completion after they finish their annotation session.

Next, they will be shown an image containing objects on sidewalks with bounding boxes and tags annotated on them. From this point, the volunteer can do two things. They may click on the pre-labeled objects and identify if they are obstructions or not, and they may also choose to annotate more objects that they think contribute to the accessibility of the sidewalk by adding bounding boxes on objects that were not included in the pre-labeled images. They are then asked to select the class or label for these bounding boxes from a list of objects (refer to Section 3.2.1). If the object is not present in our list of objects, it means that the object is easily movable by a human being (e.g. tables) so we don't consider it as an obstruction. However, an *Others* option was also provided to catch potential items that may be considered by other as an obstruction.

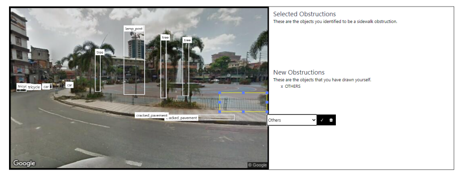

Figure 4: A screenshot of the Atlas prototype showing how to add a new annotation and label from a predetermined list of objects

3.4.3 Scoring Sidewalk Accessibility. Once participants finish annotating objects, they will then give the image a sidewalk accessibility score. Participants may select a score ranging from 1 to 10 based on their own perception of accessibility as seen in the given image, and end the annotation process by selecting the surface type of the sidewalk in the image as shown in Figure 5. To prevent any biased scores, did not provide any preliminary information or expectations of what a given score should like based on the street view image. We utilize a 10-point Likert scale as a means of capturing the widely varying levels of accessibility in urban streets. Similarly, [Naik et al.] measures the perceived safety of sidewalks in their work, Streetscore, using a 10-point scale. The reason being that the street level images in the dataset, when compared to each other, contain different elements which improve or reduce the perceived safety of the location in the image. These elements provide a more precise attribution to the indicated Streetscore when measured on a 10-point scale. Additionally, for this study, we are focusing more on the pedestrians' perception of accessibility and trying to figure out what contributes to an accessible sidewalk. We would like to note that the participants only rate the accessibility of the sidewalk in the image presented to them. We did not consider sidewalk continuity. Additionally, images that contain sidewalks on both sides of the road are given a single score; due to the fact that sidewalks that are near each other usually have a consistent design.

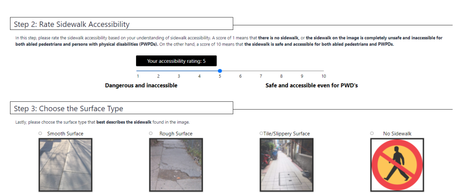

Figure 5: A screenshot of the Atlas platform showing the accessibility score slider as well as selecting the surface type

[4]https://we-the-streets-site-one.vercel.app/

After giving a sidewalk accessibility score, they can proceed to the next sidewalk image and repeat the scoring and annotation process. The annotation process is designed to be simple and convenient for crowdworkers to encourage them to score as many images as possible. The simple design also contributes to the speed of data collection so crowdworkers will not feel overworked from annotating multiple images.

The crowdsourcing platform was disseminated through online community forums, schools, and urban planning groups. The researchers aimed to incentivize crowdworkers by offering raffle draws and cash prizes to the users with the most number of annotations. We disseminated this competition through a graphic posted in the same social media groups mentioned above, as seen in appendix

3.5 Sidewalk Width Collection

In addition to collecting sidewalk accessibility scores, annotations, and surface types, we collected the widths of the annotated sidewalks due to its effect on sidewalk accessibility. As stated in Section 2.2, the width of the sidewalk is a feature that contributes to its overall accessibility. The collection of sidewalk widths was done through the use of the measure distance feature on Google Earth. As seen in Figure 6, the feature allows users to click and drag on a top-view satellite street view image and will output the width in meters. We used the latitude and longitude of the street view images that were annotated in order to locate it on Google Earth. By measuring the width of the sidewalks, we add another feature on sidewalk accessibility that will be used as a predictor for the model.

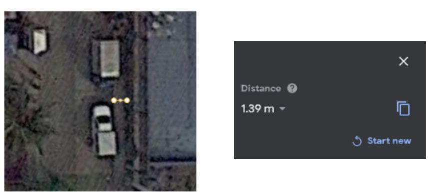

Figure 6: Measuring sidewalk width of a satellite image using the measure distance feature of Google Earth

3.6 Model Creation and Validation

After collecting accessibility scores, annotations, surface types, and sidewalk widths, we begin with modeling sidewalk accessibility. With the use of XGBoost, we created a regression model in order to see which features affect the accessibility of a sidewalk. The model was also used to predict the accessibility score of an entry given the features such as surface type, sidewalk width, and obstructions.

3.6.1 Pre-processing. Data pre-processing was done in order for us to be able to quantify our data for modeling and exploration. Initially, the annotation data we receive from our crowdworkers consists of the following columns: *id*, *date*, *username*, *imageID*, *accessibilityRating*, *pavementType*, *selectedObjectsID*, *newObjects*. In order to process this better, we needed to transform *pavementType*, *selectedObjectsID*, and *newObjects* in a more measurable data type. For the objects, we mapped the *selectedObjectsID* and *newObjects* column to their respective counts per class. As such, we generated the following columns: *objectType_total*, *objectType_obstruction*, *objecType_non-obstruction* for each of the 18 object types the crowdsourcing platform has. In addition, the *pavementType* column was converted to four numerical columns: *isRough*, *isSmooth*, *isSlippery*, *isNoSidewalk*. We would like to state that not all the user annotations were used due to our limited time and our manual process of acquiring the sidewalk width. Because of this, annotations without a corresponding sidewalk width were dropped. After this, we normalized the data by converting the counts of obstructions and non-obstructions into percentages by dividing them with the total counts of objects in an image. We also averaged the sidewalk widths of annotations on the same image in order to have a consistent width among images with multiple annotations.

3.7 Model Development

In developing our accessibility score model, there are two objectives: predicting the accessibility score of a sidewalk given several features and determining the features with the biggest impact on the prediction. To do this, we ran a regression analysis on the data through the use of XGBoost regressor. XGBoost is among the highest performing learning models due to its use of an collection of learning models to provide accurate results [5]. We identified 40 unique predictors for the model including the sidewalk width, three surface types, 18 sidewalk obstructions types, and 18 object types not tagged as obstructions.

In order to better understand the impact of each feature, we use SHAP values from the SHAP package to interpret them. SHAP values represent the contributions of each value to the output of the sidewalk accessibility model. Through SHAP values, we can see which features have a positive and a negative effect on the output of the model. In this case, the output of the model would be the predicted accessibility score.

3.7.1 Training and Validation. In order to train and validate the model, we split the data into train and test sets with a percentage of 80% and 20% respectively. We then look at modelling with a classifier (XGBClassifier) and a regressor (Linear Regression, XGBRegressor) and choose the model that performs best based of their scores on the dataset, 5-fold cross validation score, and other peformance measures such as the accuracy and root mean squared error for the classifier and regressor respectively. We then tune the parameters of the models to increase its performance. In order to determine the best parameters to be tuned, we use **GridSearchCV**, a function from sklearn's model selection package, to fit multiple configurations of parameters in our model. By running multiple configurations of parameters, GridSearchCV gets to see which configurations performs the best on the train set. The configurations of parameters that were tuned for the classifier and regressor can be seen in figures 2 and 3. These parameters were chosen due to their relevance on influencing the model's ability to generalize and reduce overfitting. We also use ridge as the regularization method for the models since we do not want to remove features when reducing the complexity of our model. After getting the best configurations of parameters for the XGBoost regressor and classifier, we ran the 5-fold cross validation with original default parameters and tuned parameters and compare their accuracy scores. Besides this, F1-score was used to measure the performance of the classifier models while the root mean squared error (RMSE) was used for the regression models.

Hyperparameter	Value
learning_rate	0.01
max_depth	3
min_child_weight	3
n_estimators	500

Table 2: Best hyperparameter values found for the XGBClassifier model

Hyperparameter	Value
colsample_bytree	0.7
learning_rate	0.01
max_depth	5
min_child_weight	1
n_estimators	500
subsamples	0.5

Table 3: Best hyperparameter values found for the XGBRegressor model

3.8 Improving the Pipeline of Atlas

We initially collected a total of 109,832 street view images from the Google Street View API. However, we encountered resource problems earlier on when training YOLOv5. With this, we used the 12,372 images we collected from manual annotations and combined it with the annotations from the YOLOv5 trained on the COCO 2017 annotations. We proceeded with our crowdsourcing on Atlas by only using the 12,372 street view images. However, training YOLOv5 would be helpful so that we can utilize all the 109,832 street view images that we initially collected. This section documents the process of how we were eventually able to train the YOLOv5 model. Afterwards, we uploaded these new images to Atlas.

3.8.1 Training of YOLOv5 with COCO 2017 Annotations and Manual Annotations. Training the YOLOv5 model would improve our study as we would be able to accurately measure the correctness of the annotations in the street view images. When training the model, we first shuffled the data to ensure that the all the different classes of the annotations get randomized among the train, test, and validation sets. We then partitioned the data by following a split ratio of 70% train, 15% valid, and 15% test. Our train data included 26,793 images, while our test and valid both contained 5,743 images. We also created a python script to compute for the total number of instances within the Train, Test, and Validation sets to ensure that the instances of the classes are more or less equally divided among the three sets according to their split ratio. Afterwards, we followed the guide instructions on YOLOv5's main GitHub repository on how to train on a custom dataset [8]. We trained on 300 epochs and a batch size of 64. It was recommended by the author of YOLOv5 [10] that the training should start on 300 epochs, and the batch size to be set to the largest that our hardware allows for. We also used YOLOv5x, the largest and most accurate model of YOLOv5 for images with 640 pixels in size. All our 109,832 street view images were saved in a resolution of 640x400, hence we used the YOLOv5x model. After the training was completed, we saved the best weights of our model, and used it to run inference on our dataset of 109,832 street view images. Figure 7 is a sample output of the inference using the trained YOLOv5 model.

Figure 7: Sample result of the inference using the trained YOLOv5 model

3.8.2 Attempt at Filtering Sidewalks Using Semantic Segmentation. We also attempted to improve the results of our crowdsourcing phase by filtering sidewalks from images using semantic segmentation. Semantic segmentation is defined as the process of classifying each pixel belonging to a particular label [12]. We used OCR (ResNet-101) semantic segmentation model through MMSegmentation, an open source semantic segmentation toolbox from OpenMMLab.

In total, we ran inference on 109,832 Google Streetview images which we previously collected. Unfortunately, the inaccuracy of the model out of the box led us to discontinue our plan to apply it on images during the crowdsourcing phase. Although the group considered morphological transformation methods such as dilation and erosion to complete the sidewalk area created by semantic segmentation, it would be difficult to handle the cases of mislabeled sidewalks where sidewalks are identified by the model even if there are not any sidewalks. Because of this, we did not push through with using these images for crowdsourcing since defining a generalized threshold would not guarantee accurate results for images with mislabeled sidewalks. In order to replicate the effect of semantic segmentation, we introduced a feature on Atlas where users can select a *"No sidewalk"* option on images with no sidewalks. This will help us exclude these images and annotations from our model.

4 RESULTS AND DISCUSSION

This section discusses the results from training of the YOLOv5 model on the COCO 2017 annotations and manual annotations, the crowdsourcing on Atlas, as well as the modeling of the obtained accessibility scores.

4.1 YOLOv5 Model Trained with COCO 2017 Annotations and Manual Annotations

Training the YOLOv5 model using YOLOv5x (the largest and most accurate model of YOLOv5) with 300 epochs and a batch size of 64 took 39.770 hours to complete. In Figure 8, we can see the graph of the precision, recall, and mean average precision training results. The values for precision, recall, mAP@0.5, and mAP@0.5:0.95

peaked at 0.6665, 0.5447, 0.5359, and 0.3300 respectively. It can be seen in the graphs that the precision continued to increase with the higher number of epochs, while the recall, mAP@0.5, and mAP@0.5:0.95 all peaked at around 60 to 70 epochs before decreasing as the number of epochs reached 300. The model was overfitting early and we could have reduced the number of epochs during training. However, we were no longer able to train the model again due to resource limitations.

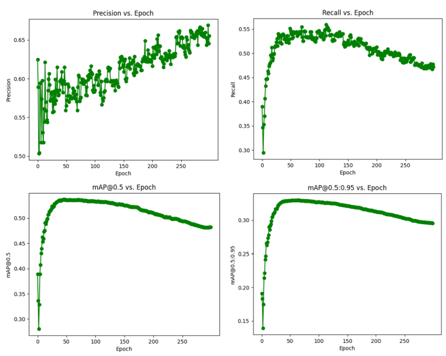

Figure 8: Precision, Recall, and Mean Average Precision (mAP) results of the YOLOv5 model

The main metric used to measure YOLOv5's performance is its mAP. After the training was completed, YOLOv5 generated the best training weights, and we tested its performance on our test set and validation set. For the test set, it scored 0.524 mean average precision (mAP)@0.5 on all classes, while it scored 0.534 mAP@0.5 on the validation set. The highest average precision (AP) scores were on objects such as *fire hydrant* and *stop sign*, with 0.882 and 0.865 respectively. This means that the model had lower false positives and lower false negatives for these two objects, resulting to more chances of correct predictions when seen on the streetview images. On the other hand, it scored the lowest on objects such as *cracked pavement*, *curb ramp*, and *construction materials* with 0.073, 0.106, and 0.187 respectively. These low scores would result to more mislabels of the three objects when seen on the streetview images. Comparing it to the pretrained checkpoints of YOLOv5x on their main GitHub repository [9], its benchmark mAP score for the mAP test 0.5:0.95, mAP val 0.5:0.95, and mAP val 0.5 are: 50.4, 50.4, and 68.8 respectively. The performance of our trained model only had scores of: 0.324, 0.329, and 0.534 respectively. The low average precision scores of our trained model on some of the objects of could be the reason as to why our model had lower scores than the benchmark mAP scores.

Upon analyzing the images from our train set for *cracked pavement* and *curb ramp*, we noticed certain inconsistencies on the annotations that might affect how the trained YOLOv5 model views the objects. For the *cracked pavement* object, the annotation could either include small pieces of cement, or vertical line cracks on the surface of the sidewalk. For *curb ramp*, some curb ramps that are not steep and not painted may easily blend in as part of the entire sidewalk pavement. The presence of shadows and harsh lighting conditions in some our streetview images could also make it difficult for the YOLOv5 model to detect the objects on the image. Lastly, another factor would be difference in the image quality of the streetview images and the COCO 2017 Dataset. The street view images were somewhat pixelated compared to the detailed images from the COCO 2017 Dataset. Also, the images from the COCO 2017 Dataset were focused on the specific object, whereas the street view images were focused on the street itself and not the object in particular.

4.2 Crowdsourcing with Atlas

As of August 13, 2021, we were able to collect a total of 5,047 annotations and accessibility scores from 48 unique users. As shown in Figure 5 of Chapter 3.4.3, users were able to select *"No sidewalk"* as an option for surface type, and due to that filtering we were able to reduce that number to 4,096 total annotations. To further substantiate the data for our accessibility model, we manually collected sidewalk widths for images annotated by users as described in section 3.5. This led to a final data set containing 2,476 annotations with corresponding sidewalk widths.

For accessibility scores, we can observe in Figure 9 that the most common scores given to sidewalks in Manila and Makati were 4, 7, and 8. The perceived mean accessibility score is 5.643 with a standard deviation of 2.673. Some annotations, however, were done on the same image by multiple users. If we instead observe the accessibility scores per unique image then we obtain a mean accessibility score of 5.733 with a lower standard deviation of 1.316. The higher standard deviation despite having repeating images led us to believe that users may have had their own biases or standards when scoring sidewalks. For this, we decided to look at user demographics and activity to find a possible explanation.

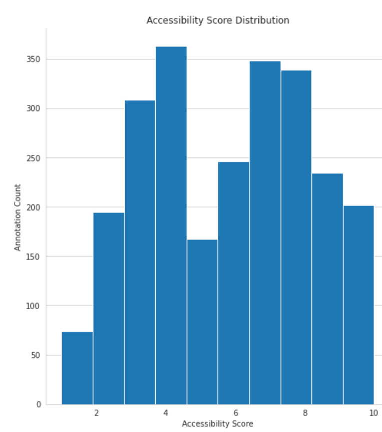

Figure 9: Perceived accessibility score distribution

4.3 User Demographics and Activity

We first look at the age of our users seen in Figure 10, where we see that most of our users are in their early 20's. It is common for people of this age to be fairly computer literate as well as have an understanding of commuting and pedestrian infrastructure, given that those in their early 20's in the Philippines usually take public transportation to school or work. We also have a few users above the age of 40 who have experienced more change in sidewalks from previous years compared to now. This may contribute to a better understanding of scoring accessibility. We also looked at the city of residence of our users as seen in Figure 11 where we see that the majority of our users come from Manila, one of the cities we selected for our crowdsourcing phase; we did not have any registered users though from Makati, the other city we had collected street view images from. Unfortunately, we were not able to gather users that make use of mobility aids who would have been ideal users to score accessibility.

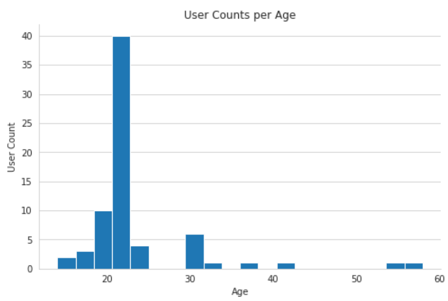

Figure 10: Registered user ages

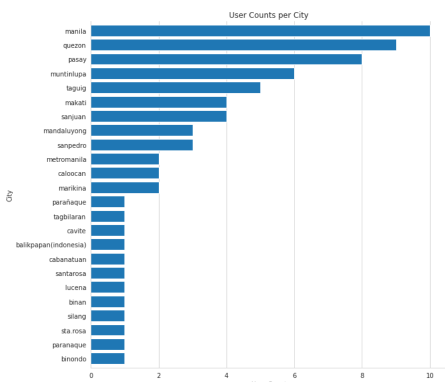

Figure 11: Registered user residences

Regarding user annotation activity, we observed a large distribution from the mean number of images annotated by users. With an average of 54 images annotated per user, we have a high standard deviation of 131.404, with our highest user annotating a total of 729 images. Figure 12 shows that 69% of our annotations come from the top 5 annotators, with the top annotator contributing 28% of all annotations just by themselves. The most likely reason for the top annotators having such high activity was the giveaway competition we conducted to gather more volunteers as mentioned in section 3.4.

Additionally, there is no distinguishable bias from the top annotators, or from any annotators for that matter, in terms of their given accessibility scores, seen in Figure 13. The users were shown different images and therefore we did not expect their accessibility scores to be similar.

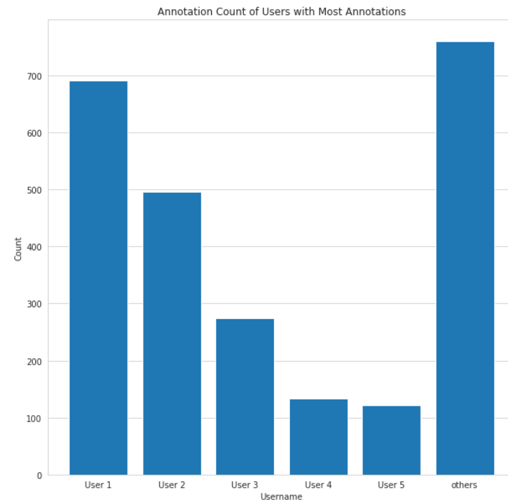

Figure 12: Annotations of top scorers

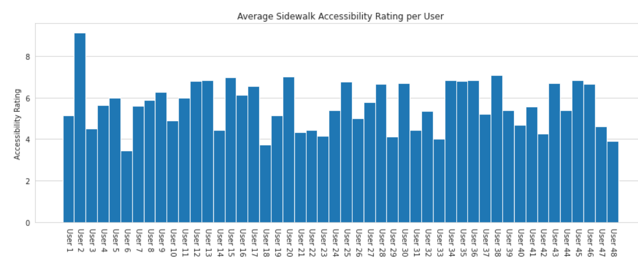

Figure 13: Perceived average accessibility scores of by each user

4.4 Sidewalk Obstructions

In total, we detected 21,590 objects from all the 12,372 images used on Atlas, with only 6,485 (29.9%) of objects being considered as obstructions by users. The images uploaded to Atlas were the combination of the manual annotations from our volunteers and the annotations of the YOLOv5 trained on the COCO 2017 annotations to allow users to quickly provide an accessibility score for the sidewalk. Users were also able to label new obstructions in case they were not pre-labeled by YOLOv5, and 1,943 new objects were annotated by users, equating to 8.9% of all objects. We should also consider that we included an *Others* class from the list of objects that users could select in case they found an obstruction that didn't fall under any class. In total, we found that there were 544 objects selected as *Others* which we manually labeled and grouped into the following classes and categories:

- **Aesthetics** - Plants, Greenery, Sidewalk Design, Grass Wall
- **Utilities** - Barriers, Infrastructure, Tower, Water Meter, Shed
- **Commercial/Livelihood**: Sign, Tables, Chairs, Bottle Case, Refrigerator, Basketball
- **Clutter**: Rocks, Debris, Concrete Box

The counts of these objects from each category can be seen in Figure 14.

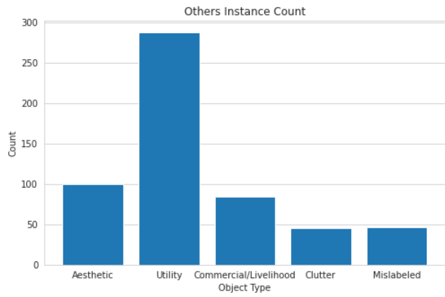

Figure 14: Counts of others objects

Accessibility Score	Surface Type			
	Rough	Smooth	Slippery	% of Smooth Surface
1	38	34	1	46%
2	83	106	6	54%
3	110	179	19	58%
4	106	236	21	65%
5	47	115	3	69%
6	54	166	25	67%
7	42	250	56	71%
8	29	261	49	76%
9	7	200	26	85%
10	5	185	12	91%

Table 4: Perceived accessibility scores and surface types table of comparison

4.5 What Constitutes a Good Sidewalk

Based on all the data we've collected, we can come up with some basic assumptions for what makes a good sidewalk and what makes a bad one using accessibility scores as the dependent variable. Firstly, there is an observable pattern that sidewalk width is directly proportional with accessibility scores as seen in Figure 15. To further validate this claim, we used a Spearman Rank-Order Correlation statistical test. The spearman's rho correlation coefficient are as follows: r_s=0.3306, p=2.9395 x 10^{-64} indicating a positive correlation between sidewalk width and accessibility score.

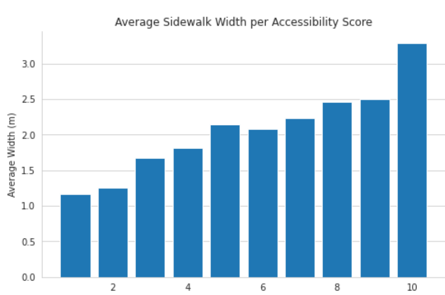

Figure 15: Comparison of sidewalk widths and perceived accessibility scores

Next, we take into account the surface types of sidewalks which we asked users to select among *Rough*, *Smooth*, and *Slippery*. Similarly to sidewalk width, the count of smooth sidewalks increases from accessibility scores 1 to 10 as seen in Table 4. The count of rough surfaces also seem to follow a downward trend, but slippery surfaces do not seem to affect the perception of accessibility. In accessibility score 10, we even see that 91% of surfaces were considered smooth.

Accessibility Score	Mean Total Objects	Mean Objects
1	7.652	1.859
2	7.865	3.200
3	8.099	2.701
4	9.082	3.068
5	8.827	2.751
6	8.785	2.931
7	8.890	2.789
8	8.637	2.345
9	7.811	1.859
10	6.797	0.856

Table 5: Perceived accessibility score and objects table of comparison

Finally, we take a look into obstructions which negatively affect accessibility score, while exploring those objects that may actually provide a positive effect on sidewalks. Table 5 displays key figures which allow us to analyze and observe patterns between objects and accessibility. We observed that there is a downward trend of the ratio between total objects found in the image to obstructions as accessibility score rises; however, this correlation is not as significant. Spearman's rho correlation coefficient was used to asses the relationship on total objects found in the image and accessibility score. There was no significant correlation between the two, r_s=-0.036, p=0.0717. On the other hand, we observed a clearer correlation with the number of obstructions found in each image. Again, the spearman's rho correlation was used to assess the relationship between accessibility score and the obstruction count. There was a negative correlation between the two: r_s=0.110, p=1.945 x 10^{-8}.

Seeing that these numbers are found at the extremities of the accessibility score scale, this led us to the inference that some objects contribute to a positive score just as much as obstructions contribute negatively to the score. We dove deeper into this by comparing objects commonly present in images with scores of 1 and 10, which can be seen in tables 6 and 7. In this figure we also show the obstruction status of these objects as determined by our users. The figure shows that trees are more present in 10-scored images and that most trees were not considered as obstructions. We can observe the same thing with benches, so we can interpret that locations with non-obstructing trees and benches tend to seem more accessible such as outdoor parks and public transportation spots. On the other hand, motorcycles, tricycles, and construction materials are more prevalent in 1-scored images than in 10-scored ones.

4.6 Accessibility Score Model

As mentioned in section 4.2, 2,476 annotations with accessibility scores were left after pre-processing. The annotations are split with a ratio of 80% and 20% respectively. This gives us 1,980 entries for the train set and 496 entries for the test set. In choosing the right model,

Perceived Accessibility Score = 1			
Objects	**Obstruction**	**Non-Obstruction**	**Total**
bench	1	4	5
bicycle	7	11	18
car	47	362	409
construction materials	35	13	48
curb ramp	1	7	8
cracked pavement	15	31	46
fire hydrant	1	3	4
lamp post	15	57	72
motorcycle	23	87	110
parking meter	0	0	0
stop sign	1	0	1
street sign	7	29	36
street vendor stand	19	12	31
traffic light	2	34	36
tree	20	193	213
tricycle	27	23	50
utility post	50	84	134
others	34	0	34

Table 6: Relation of Obstructions and Non Obstructions in Perceived Accessibility Score = 1

Perceived Accessibility Score = 10			
Objects	**Obstruction**	**Non-Obstruction**	**Total**
bench	6	19	25
bicycle	0	8	8
car	3	402	405
construction materials	2	13	15
curb ramp	1	18	19
cracked pavement	5	13	18
fire hydrant	1	3	4
lamp post	29	115	144
motorcycle	7	26	33
parking meter	0	2	2
stop sign	0	2	2
street sign	22	39	61
street vendor stand	1	2	3
traffic light	1	2	3
tree	51	490	541
tricycle	0	1	1
utility post	20	32	52
others	23	0	23

Table 7: Relation of Obstructions and Non Obstructions in Perceived Accessibility Score = 10

we ran different models and also perform hyperparameter tuning by getting the most optimal parameters through GridSearchCV. We also used ridge as the regularization for both the linear regression and XGBoost models. We used ridge regularization because we did not want to disregard features when performing regularization. From this, we select the model that performed the best in terms of cross-validation score through 5-fold cross validation. The cross-validation score represents the accuracy of the model. We also look at other performance measures such as the root mean squared error (RMSE) for the regression models while we looked at F1-score for classification models.

4.6.1 Accessibility Score Model Performance. First, we ran a linear regression model with ridge regularization from `scikit-learn`. After getting a accuracy of 29.14% and an RMSE of 2.13, we decided to look at XGBoost in order to see if the model performance would be better. Similar to the linear regression model, we used ridge as the reguralization parameter for the XGBoost models. Hyperparameter tuning is also made easier in XGBoost because of its popularity in literature. Initially, we opted to use XGBClassifier since the response variable for our model is a discrete value. However, when we ran the XGBClassifier model, we got a accuracy of 23.02% which was marginally worse than the linear regression model. Additionally, we looked at the weighted F1-score of the classifier and found that it was 20.21% meaning that it's precision and recall were low. After this poor result, we tuned the hyperparameters of the XGBClassifier with use of GridSearchCV. GridSearchCV runs several configurations of hyperparameters and helps in selecting the best configuration of hyperparameters given the training set.

Model	XGBClassifier (Base)	XGBClassifier (Optimized)
Train Score	45.80%	36.52%
Test Score	22.18%	21.77%
5-Fold Cross Validation Score (Accuracy)	23.02%	22.70%
Weighted F1-score	20.21%	17.81%

Model	Linear Regression	XGBRegressor (Base)	XGBRegressor (Tuned)
Train Score	32.60%	44.70%	50.08%
Test Score	29.33%	31.66%	33.66%
5-Fold Cross Validation Score (Accuracy)	29.14%	31.95%	33.43%
RMSE	2.13	2.10	2.13

Table 8: Classification and Regression Model Performances

With the tuned XGBClassifier, we got a accuracy of 22.70% and a weighted F1-score of 17.81%. The poor performance of the optimized XGBClassifier led us into using the XGBRegressor model to model our data. Without any hyperparameter tuning, the XGBRegressor got a 5-fold accuracy of 31.95% and an RMSE or 2.10. While the performance of the XGBRegressor was the best that we had so far, we tried to tune it and compare its performance.

With the tuned XGBRegressor model, we got a accuracy of 33.43% with an RMSE of 2.13. While the RMSE of the tuned XGBRegressor model was slightly higher than the base XGBRegressor model, it had a better accuracy than the base XGBRegressor model. In the end, we opted for the tuned XGBRegressor due its better accuracy based on its 5-fold cross-validation score.

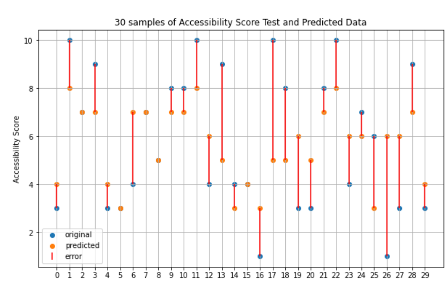

Figure 16: A sample of perceived vs the predicted values of the accessibility score model

As seen in figure 16, we can see multiple predicted entries with a noticeable margin of error. From this, we realize that the accuracy of the predictions are considerably low. 33.43% accuracy of the model and the RMSE of 2.13. Given that the accuracy of the tuned XGBRegressor is low, we can look into improving the accuracy of the model by providing a robust dataset. The inconsistencies in crowd sourced accessibility scores and labels could also have contributed to the accuracy of the model. Re-training the model with a dataset filled with images with multiple accessibility scores can also benefit the predictions of the model.

4.6.2 Feature Importance of Predicted Accessibility Scores. Another important finding that we get from running a regression model is the calculation of feature importance used in the model. Quantifying the influence of each feature in predicting the accessibility score allows us to analyze the features that are desirable and undesirable in assessing sidewalk accessibility. By using SHAP values, we are able to see the features with the most influence on the accessibility score.

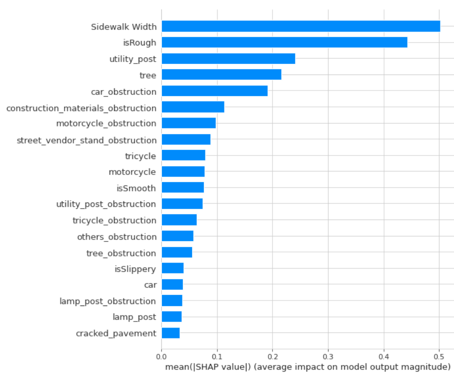

Figure 17: Mean absolute SHAP value of top 20 influential features on predicted sidewalk accessibility score

In Figure 17, the influence of a feature is measured in a mean absolute SHAP value. The mean absolute SHAP value does not look at the positive or negative influence but the total influence a feature has on a predicted sidewalk accessibility score. Based on Figure 17, the features with the considerable influence on a predicted sidewalk accessibility score are sidewalk width, rough sidewalk surface type, utility post non-obstruction, tree non-obstruction, car obstruction, construction material obstruction, motorcycle obstruction, steet vendor obstruction, tricycle non-obstruction, motorcycle non-obstruction, smooth sidewalk surface type, utility post obstruction, tricycle obstruction, tree obstruction, slippery sidewalk surface type, car non-obstruction, lamp post obstruction, lamp post, and lastly cracked pavements.

Most of the top 20 features listed in Figure 17 have a commonality within them that makes them influential to the predicted sidewalk accessibility score. Each of them has a high number of occurrence in the annotations. For sidewalk width and surface types, every annotation must have at least one sidewalk width and one surface type. Because of this, their mean absolute SHAP value is the highest among the features. In the top 20 features, we could also see the same objects with the obstruction and non-obstruction pairing. Objects such as utility post, trees, cars, motorcycle, tricycle and lamp post are the features that both have their obstruction and non-obstruction form in the top 20 features. In looking at the objects

Objects	Obstruction	Non-Obstruction	Total
bench	33	64	97
bicycle	82	150	232
car	666	6720	7386
construction materials	214	127	341
curb ramp	108	274	382
cracked pavement	267	360	627
fire hydrant	23	23	46
lamp post	667	951	1618
motorcycle	307	827	1134
parking meter	0	4	4
stop sign	11	17	28
street sign	462	598	1060
street vendor stand	171	72	243
traffic light	116	460	576
tree	1068	2796	3864
tricycle	213	289	502
utility post	1533	1373	2906

Table 9: Object Occurences

with the most counts on table 9, we confirm that these objects have the most count among the other objects we included in our list of obstructions. The occurrences of features has a direct relationship with the overall influence it has on the SHAP value.

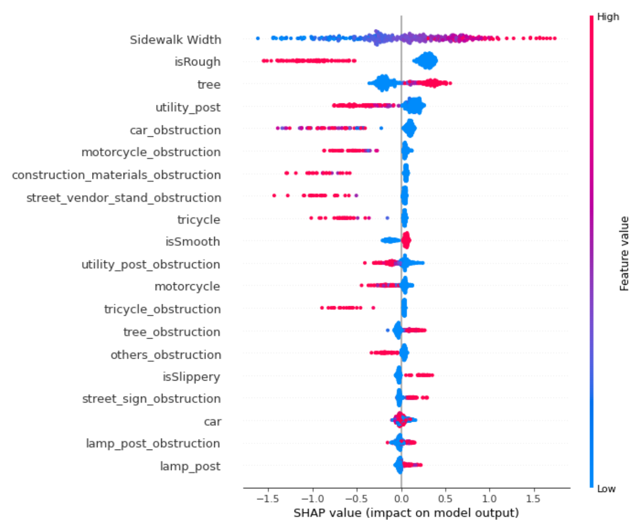

Figure 18: Feature influence based on their values

Now that we know the overall influence of a feature, let's look at another summary plot to see how which values of these features make them have a positive or negative effect on the predicted sidewalk accessibility score. In Figure 18, we can see the distribution of high and low feature values and their respective influence on the model. For sidewalk width, wider sidewalks tend to have a positive influence on the predicted accessibility score, with most red data points having 0.5 - 1.5 SHAP values. For rough sidewalk surfaces, a high value will lead to a negative influence on the score. For most objects, high values have a negative influence on the predicted sidewalk accessibility score. It does not matter whether that object is an obstruction or not, as long as it has a high value, it will have a negative influence on the predicted sidewalk accessibility score. However, the tree object is an exception to this trend. Its presence, whether as an obstruction or not, has a positive influence on the predicted sidewalk accessibility score. It's also good to point out that cars, utility posts, motorcycles, construction materials, street vendor stands, and tricycle are obstructions that have the largest negative influence on sidewalk accessibility. This is because these objects are notorious in being found in sidewalks. Some common examples of this are when tricycles, motorcycle, and street-vendors park on sidewalks. Besides this, construction materials such as cement powder disrupting the surface of a sidewalk are often the pain of pedestrians that walk in Metro Manila. Looking at the mean absolute SHAP values and feature influence, we can see that sidewalk width and rough surface type are the most influential features that affect the predicted sidewalk accessibility. To find out what constitutes a high value or low value for these two features, we can look at individual breakdowns of three levels of accessibility scores. The three levels of the predicted accessibility scores are low, mid, and high. The grouping of these levels were based on the average or base value of the predicted scores. Looking at the base value, the average score that was predicted by our model was 5.762. Now we look at an instance of a low accessibility score prediction.

Predicted Accessibility Score = 3		
Features	**Feature Value**	**SHAP value**
Sidewalk Width	0.8m	-0.97
Rough	TRUE	-0.85
motorcycle obstruction	1	-0.55
lamp post	1	0.17
car obstruction	0	0.08

Table 10: Feature and SHAP value of sidewalk width and rough surface in predicted low score

In this case it is 3.28. Looking at the features, we see that a sidewalk width of 0.8, a rough surface type, and a motorcycle obstruction push the predicted sidewalk accessibility score down. We also see that a lack of car obstructions and the presence of lamp posts have a little positive effect on the predicted sidewalk accessibility score.

Predicted Accessibility Score = 5		
Features	Feature Value	SHAP value
Sidewalk Width	1.5m	-0.092
Rough	FALSE	0.28
utility post	1	-0.57
trees	0	-0.22

Table 11: Feature and SHAP value of sidewalk width and rough surface in predicted mid score

For the mid accessibility score prediction, the presence of utility posts, even as a non-obstruction, negatively influenced the accessibility score. The lack of trees also negatively influences the accessibility score. A sidewalk width of 1.5 meters also has a slight negative effect on the prediction. However, we can see that when the surface type of the sidewalk is not rough, it has a large positive effect on the accessibility score.

Predicted Accessibility Score = 7		
Features	Feature Value	SHAP value
Sidewalk Width	2.0m	0.14
Rough	FALSE	0.34
utility post	0	0.25
tree	1	0.38
car obstruction	0	0.13

Table 12: Feature and SHAP value of sidewalk width and rough surface in predicted high score

Lastly for the high accessibility score prediction, the presence of trees, lack of utility posts, and car obstructions have a positive influence on the accessibility score. The non-rough surface also has a large positive effect on the predicted high accessibility score. The sidewalk width of 2 meters also produces a positive effect on the accessibility score.

From the 3 observations, we can see that the surface type of a sidewalk heavily influences the predicted sidewalk accessibility score due to the fact that the rough surface feature is among the features with the biggest portion of the breakdown. We can also see that the sidewalk width's influence slowly turns from negative to positive as its value increases. When the width of the sidewalk is 0.8 meters, it has a huge negative influence on the score, with 1.5 meters, it still has a negative influence but it is noticeably less. With 2 meters, we can see that it has a slightly positive influence on the predicted sidewalk accessibility score. From this we can say that the point of change from negative to positive influence could be in the region between 1.5 meters and 2 meters. Lastly, the presence of trees also provide a positive influence. This could be because of the protection that trees provide from the road. Also the trunk of a tree also makes the sidewalk width increase resulting in wider sidewalks. Lastly, the presence of utility posts have a negative influence on the predicted sidewalk accessibility score.

5 CONCLUSION

5.1 Training the YOLOv5 Model

The results of the YOLOv5 model we trained with COCO 2017 annotations and manual annotations did not perform as well as the ideal benchmark scores. For mAP test 0.5:0.95, mAP val0.5:0.95, and mAP val 0.5, our trained model only had scores of: 0.324, 0.329, and 0.534 respectively. Many factors could have been the reason as to why our model was not able to have good performance scores. One would be the difference in the image quality between the streetview images and the COCO 2017 Dataset. The COCO 2017 Dataset were more detailed compared to that of the streetview images we collected from the Google Street View API. Aside from image quality, images from the COCO 2017 Dataset were also focused on the specific object, whereas the streetview images were focused on the street itself and not on the object in particular. The model also performed poorly on objects that might appear ambiguous such as *cracked pavement* and *curb ramp*. The time of day, harsh lighting conditions, and shadows on the streetview images also affect the clarity of the object being detected. With these, it can be said that when training the YOLOv5 model, not only do the number of images per class or number of instances per class matter in achieving higher performance scores, but also the image quality and consistency of how the object appears to look on each image of the dataset. We would also like to highlight that despite the low mean average precision scores of our model, it was still successful in detecting Philippine specific objects on the sidewalk.

To further improve the performance of the YOLOv5 model when training on a custom dataset, we recommend future researchers to train the model more than once, given the luxury of time. If the model overfits early, the training could be improved by reducing the number of epochs. Hyperparameter tuning can also be performed to reduce and delay overfitting, resulting to higher final mAP scores. The consistency and accuracy of the labels and the quality of the images may also improve the YOLOv5 model to properly learn the objects during training.

5.2 Crowdsourcing

Previous works such as Project Sidewalk have heavily relied on crowdsourced data to support research purposes. Similarly, we have contributed to this method by creating our own crowdsourcing platform which we have introduced to dozens of new users, as well as to multiple academic institutions through email campaigns. The results we have obtained contribute to our understanding of crowdworkers' quality of work and motivation. One of the objectives of Project Sidewalk was to compare the quality of work between volunteers and paid workers from Amazon Mechanical Turk [16]. The number of users and annotations drastically increased when we announced the mechanics for the annotation competition, but it did not have a clear effect on the quality of work.

Our online crowdsourcing platform distributed through social media was able to garner 70 unique users, but only 48 of them actually contributed to our research. Through these users we were able to collect 5047 annotations, these were eventually reduced to only 2,476 annotations since those were the only ones we could provide sidewalk widths for in the given time span. The crowdsourcing platform which performed reliably throughout the crowdsourcing phase is one of our major contributions to the field of sidewalk accessibility and crowdsourcing. Atlas served as a consistent platform despite only being developed to its minimum capacity using bare bones features. Developing a custom crowdsourcing platform can be challenging given the obscurity of research projects which require specific features and functions, so we recommend a stricter time frame when building platforms such as these. We also recommend future researchers to build on the features of our crowdsourcing platform to encourage more people to register as users. It would be very beneficial to gather users who have experienced the challenges of using mobility aids to travel. Gamification and streetview traversal can be added to create a more immersive experience which would also promote more accurate accessibility scores for whole sidewalks rather than just sidewalk segments seen in static images. Finally, we recommend having multiple iterations of crowdsourcing to allow for the improvement of the crowdsourcing platform in between iterations.

5.3 Modeling Sidewalk Accessibility

In modelling sidewalk accessibility, we affirm previous literature from the likes of [AYRES], [Santos], and the [United States Department of Justice] that the perception of sidewalk accessibility is positively impacted by a more considerate sidewalk width, the presence of non-slip surfaces, and the minimization of obstructions and irregularities on the sidewalk. We have contributed to study of sidewalk accessibility by including positively impacting objects in creating our model. Previous studies have covered positively impacting objects such as pedestrian signs and crosswalks, but since we did not try to influence user accessibility scores while crowdsourcing we have captured an unbiased pedestrian perspective on these positive objects. In general, more developed areas usually have higher sidewalk accessibility scores. Underdeveloped areas contain factors which hinder accessibility such as obstructions caused from unfinished construction, irregularly parked vehicles due to lack of parking space, and poorly planned infrastructure which get in the way of the pedestrian path. The data we gathered on sidewalk widths also provide key insights on the state of Filipino sidewalks. The Department of Public Works and Highways (DPWH) issued in Order 37, Series of 2009, that busy sidewalks should have a minimum dimension of 1.3 meters for the sake of wheelchair users [6]. Our findings indicate many sidewalks far smaller than that number; urban planners should consider this regulation for renovation purposes and for the creation of new sidewalks the next time around.

For future technical studies regarding sidewalk accessibility, we recommend looking into the continuity of sidewalk networks within urban areas. We believe that determining the accessibility of a single sidewalk segment does not capture the level of accessibility in the general area, and that the high quality of the entire sidewalk should be upheld from landmark to landmark. Given the impact of sidewalk width to its accessibility, we recommend utilizing more sophisticated means of gathering width data to improve the efficiency of the research pipeline as well as the quality of the results. Surface types are almost just as significant in this project, future researchers may consider breaking surface types down into surface materials instead to specify the structure of the sidewalk in the image. Additionally, future studies could use the perceived sidewalk accessibility scores and predicted sidewalk accessibility through a city map. This can provide insights on the varying levels of sidewalk accessibility in the Philippines. The visualization could also be used to evaluate the accuracy of the sidewalk accessibility model. Lastly, the crowdsourcing platform could be used to collect accessibility scores on other cities of Metro Manila. This can be extended into a comparative study on the sidewalk accessibility of different cities in Metro Manila.

REFERENCES

[1] 2012. European Urban Knowledge Network. https://www.eukn.eu/eu-presidencies/danish-eu-presidency/urban-design/urban-design-how-important-is-it-for-cities/#:~:text=Urban%20design%20is%20the%20art,whole%20neighbourhoods%2C%20and%20the%20city.

[2] AYRES. 2017. Carefully Designed Sidewalks Take Center Stage. https://www.ayresassociates.com/carefully-designed-sidewalks-take-center-stage/

[3] Jason Brownlee. 2019. A Gentle Introduction to Object Recognition With Deep Learning. https://machinelearningmastery.com/object-recognition-with-deep-learning/

[4] Rui Cao, Jiasong Zhu, Wei Tu, Qingquan Li, Jinzhou Cao, Bozhi Liu, Qian Zhang, and Guoping Qiu. 2018. Integrating Aerial and Street View Images for Urban Land Use Classification. *Remote Sensing* 10 (Sept. 2018), 1553. https://doi.org/10.3390/rs10101553

[5] Tianqi Chen and Carlos Guestrin. 2016. XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Aug 2016). https://doi.org/10.1145/2939672.2939785

[6] Hermogenes Ebdane Jr. 2009. DPWH: Enforcement of the Accessibility Law (BP344) along National Roads. https://www.dpwh.gov.ph/dpwh/sites/default/files/issuances/DO_037_S2009.pdf Accessed August 27, 2021.

[7] Glenn Jocher. 2020. yolov5. https://github.com/ultralytics/yolov5

[8] Glenn Jocher. 2021. Train Custom Data. https://github.com/ultralytics/yolov5/wiki/Train-Custom-Data Accessed August 23, 2021.

[9] Glenn Jocher. 2021. Ultralytics/yolov5. https://github.com/ultralytics/yolov5 Accessed August 27, 2021.

[10] Glenn Jocher. 2021. YOLOv5 Tips for Best Training Results. https://github.com/ultralytics/yolov5/wiki/Tips-for-Best-Training-Results Accessed August 22, 2021.

[11] Jan Kietzmann. 2016. Crowdsourcing: A revised definition and introduction to new research. *Business Horizons* 60 (11 2016). https://doi.org/10.1016/j.bushor.2016.10.001

[12] Anil Matcha. 2020. A 2020 guide to Semantic Segmentation. (Aug. 2020). https://nanonets.com/blog/semantic-image-segmentation-2020/

[13] Nikhil Naik, Jade Philipoom, Ramesh Raskar, and Cesar Hidalgo. 2014. Streetscore-predicting the perceived safety of one million streetscapes. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (05 2014), 779–785.

[14] Department of Public Works and Highway. 2021. https://www.dpwh.gov.ph/dpwh/2019%20DPWH%20ATLAS/Table%205%20-%20Roads.pdf

[15] D. Quercia, R. Schifanella, and K. McLean. 2015. Smelly Maps: The Digital Life of Urban Smellscapes. https://www.semanticscholar.org/paper/Smelly-Maps%3A-The-Digital-Life-of-Urban-Smellscapes-Quercia-Schifanella/f50082a51ad8a975c2e01cc734addc24db16e48a

[16] Manaswi Saha, Michael Saugstad, Hanuma Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, and Jon Froehlich. 2019. Project Sidewalk: A Web-based Crowdsourcing Tool for Collecting Sidewalk Accessibility Data at Scale. https://doi.org/10.1145/3290605.3300292

[17] Paula Manoela dos Santos. 2018. The 8 Principles of the Sidewalk: Building More Active Cities. https://thecityfix.com/blog/the-eight-principles-of-the-sidewalk-building-more-active-cities-paula-santos/

[18] Mikael Simalango. 2018. What Object Categories / Labels Are In COCO Dataset? https://tech.amikelive.com/node-718/what-object-categories-labels-are-in-coco-dataset/

[19] United States Department of Justice. 2010. *2010 ADA standards for accessible design.* Department of Justice. https://www.ada.gov/regs2010/2010ADAStandards/2010ADAstandards.htm#curbramps

[20] USGS. 2020. When Remote Sensing Meets Artificial Intelligence & GIS. https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news%5Fscience%5Fproducts=0#qt-news%5Fscience%5Fproducts

[21] Liang Wen, Jeff Kenworthy, Xiumei Guo, and Dora Marinova. 2019. Solving Traffic Congestion through Street Renaissance: A Perspective from Dense Asian Cities. *Urban Science* 3, 1 (2019). https://www.mdpi.com/2413-8851/3/1/18