ASSESSING TRANSPORTATION INFRASTRUCTURE SEGMENTS FOR BIKE SUITABILITY

Final Report

by

Dimitra Michalaka
Associate Professor
Civil, Environmental, and Construction Engineering
The Citadel
Email: Dimitra.Michalaka@citadel.edu
Tel: 843-953-7676

Kweku Brown[1]
Chun-Hsing (Jun) Ho[2]
Kewei Ren[2]
Nathan Hunyh[2]
Yuche Chen[3]
William J. Davis[1]

1. The Citadel
2. University of Nebraska - Lincoln
3. University of South Carolina

September 2024

Center for Connected Multimodal Mobility (C^2M^2)

CLEMSON UNIVERSITY

SC State University | The Citadel, The Military College of South Carolina | University of South Carolina | Benedict College

200 Lowry Hall
Clemson, SC 29634

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by the Center for Connected Multimodal Mobility (C^2M^2) (Tier 1 University Transportation Center) Grant, which is headquartered at Clemson University, Clemson, South Carolina, USA, from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

ACKNOWLEDGMENT

This work was supported by the Center for Connected Multimodal Mobility (C^2M^2), a USDOT Tier 1 University Transportation Center. The research team extends sincere thanks to the participants, local city officials, and university collaborators in Charleston, SC, Columbia, SC, and Lincoln, NE, whose support and cooperation were instrumental to the successful completion of this study.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
Assessing Bike Suitability of Transportation Infrastructure Segments	September 2024
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Dimitra Michalaka, Ph.D., ORCID: 0000-0001-7001-0579 Kweku Brown, Ph.D., ORCID: 0000-0001-6497-8479 Chun-Hsing (Jun) Ho, Ph.D., ORCID: 0000-0002-6690-4403 Kewei Ren, Ph.D., ORCID: Ph.D., ORCID: 0000-0003-2577-2448 Nathan Hunyh, Ph.D., ORCID: 0000-0002-4605-5651 Yuche Chen, Ph.D., ORCID: 0000-0003-2577-2448 William J. Davis, Ph.D., ORCID: 0000-0002-3812-8654	

9. Performing Organization Name and Address	10. Work Unit No.
The Citadel Dept. of Civil & Environmental Engineering 171 Moultrie St. Charleston, SC 29409	11. Contract or Grant No.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Center for Connected Multimodal Mobility (C2M2) Clemson University 200 Lowry Hall, Clemson Clemson, SC 29634	Final Report (May, 2023 – September, 2024)
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Evaluating the built environment to determine pedalcyclist riding quality is essential for encouraging travelers to use bicycles as a means of transportation and ensuring rider safety. This research assesses the bike suitability of transportation infrastructure segments using motion and vibration sensors. Data collection was conducted using both iPhone and Android smartphone applications, recording triaxial acceleration data and GPS coordinates along various routes in three locations: Charleston, SC, Columbia, SC, and Lincoln, NE. Smartphones were mounted at the front of the bicycle to ensure an accurate representation of the bike's intrinsic vibrations. Euler's formula was used to reorient the 3-axis acceleration data to the bike's coordinate system, standardizing the data. Data were analyzed, and pavement anomalies such as cracks and potholes were detected and geocoded. The results were visualized using GIS software, with minor discrepancies observed between Android and iPhone anomaly detection, likely due to variations in GPS accuracy and processing. The comprehensive methodology utilized in this study ensures a thorough evaluation of cycling path quality, contributing to improved urban mobility and infrastructure planning. The study aims to provide valuable insights for policymakers and stakeholders to enhance urban mobility and infrastructure planning, ultimately improving road quality and safety for all users.

17. Keywords	18. Distribution Statement
Bike Infrastructure suitability, Pavement anomalies, Urban mobility	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	32	NA

Table of Contents

List of Tables

List of Figures

EXECUTIVE SUMMARY

This study investigates the suitability of transportation infrastructure for cyclists through data collection and analysis methods using motion and vibration sensors mounted on bicycles. Prompted by increasing bicycle crash rates, such as a 75% increase in South Carolina between 2015 and 2019, this research builds on earlier efforts to evaluate how built environments influence safety, mobility, and equity for cyclists.

The project utilized smartphones as low-cost, sensor-rich data collection tools to assess pavement quality along selected bike routes in Charleston, SC, Columbia, SC, and Lincoln, NE. A hybrid model combining Vector Quantized Variational Autoencoders (VQ-VAE) with Long Short-Term Memory (LSTM) networks was employed to detect anomalies in road conditions such as potholes and cracks. This approach enabled high-accuracy identification of surface irregularities, with validation accuracy exceeding 90% in Charleston and over 80% in Columbia.

By mapping defect locations using ArcGIS and Google Earth, researchers visualized and compared infrastructure quality across the three cities, offering targeted insights for improvement. In Charleston, for example, 23 segments totaling nearly 15 miles were analyzed, while in Columbia and Lincoln, specific high-traffic and residential routes were examined for surface quality and ride comfort.

The results show that urban routes with high traffic volumes pose challenges for less experienced cyclists, while multi-use paths and roads with dedicated infrastructure offer safer, more accessible alternatives. The algorithm's consistent performance across both Android and iPhone platforms confirms its robustness for wider application.

In conclusion, this research provides a replicable, data-driven framework for evaluating cycling infrastructure, offering communities, planners, and policymakers insights to enhance safety, equity, and mobility in urban transportation networks.

CHAPTER 1

Introduction

Based on data from the National Highway Traffic Safety Administration, there were 938 pedalcyclist fatalities in 2020, a 9% increase from the previous year (*1*). In that same year, South Carolina's pedalcyclist fatality rate per 100,000 population was 0.27, slightly below the national average of 0.28 (*1*). However, South Carolina experienced a 75% increase in bicycle crashes statewide from 2015 to 2019 (*2*). **Figure 1** shows the distribution of bicycle crashes by severity in South Carolina from 2015 to 2019, while **Figure 2** illustrates the fatal and serious injury crash rates by county (*2*).

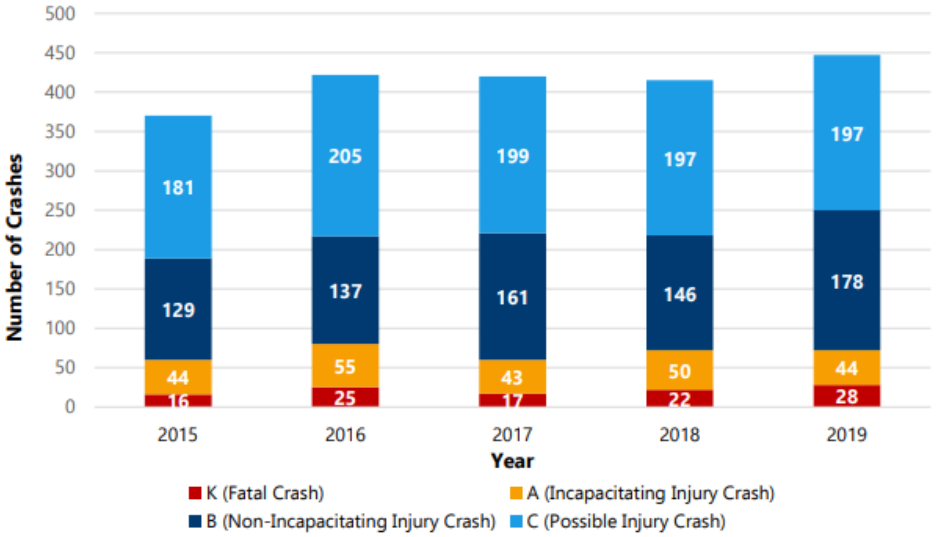


Figure 1: SC Statewide Bicycle Crashes by Severity (2015-2019) (2)

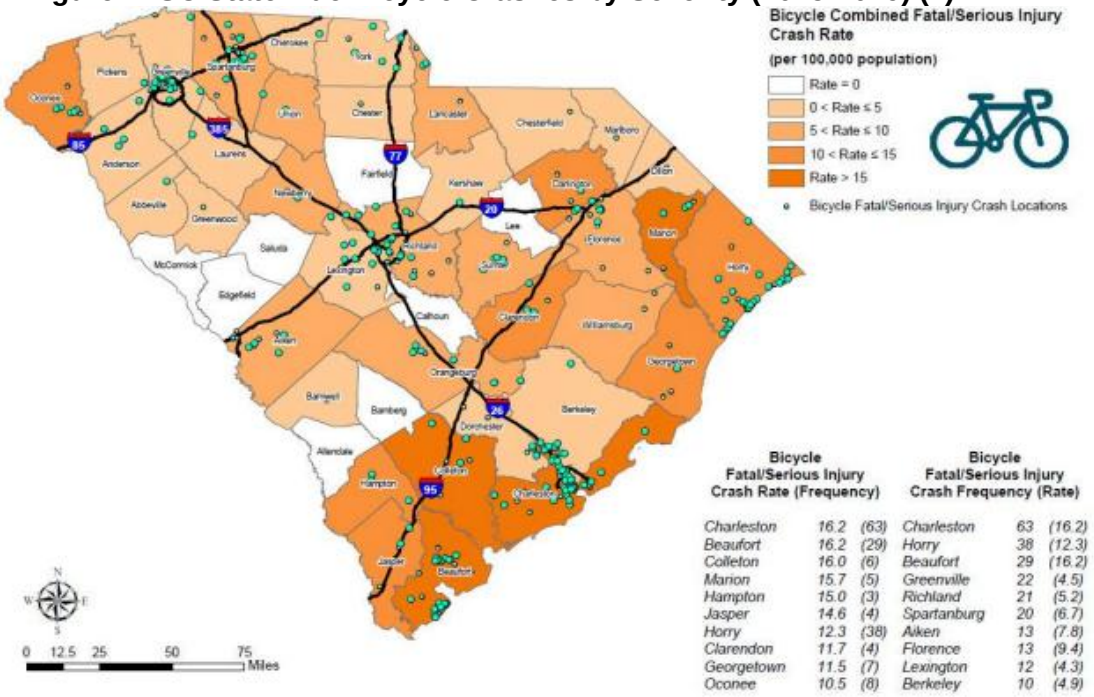


Figure 2: Bicycle Fatal and Serious Injury Crashes/Rates by County (2)

According to the Bureau of Transportation Statistics, in 2021, 52% of all driving, rail, and transit trips were less than three miles, with 28% being less than a mile (*3*). Utilizing bicycles and other alternative active transportation methods for short trips in urban areas can reduce congestion, improve mobility, encourage physical activity, and promote the use of alternative transportation modes for better overall wellness. Collecting bicycling infrastructure quality data can contribute to the understanding of barriers to active transportation.

There is a need for research and empirical measurements to assess how built transportation infrastructure accommodates bike trips in urbanized communities across the US. This project focuses on assessing transportation infrastructure segments for bike suitability using motion and vibration sensors. This effort supplements the previously completed Center for Connected Multimodal Mobility (C^2M^2) project, "Assessing Potential of Bike Share Networks and Active Transportation to Improve Urban Mobility, Physical Activity and Public Health Outcomes in South Carolina." Data are collected in Charleston, SC, Columbia, SC, and Lincoln, NE, providing case study locations for exploring insightful relationships that will inform communities. The data are analyzed to investigate route conditions, helping to better understand how the built environments in the respective cities meet users' needs.

Collecting motion and vibration data using sensor applications allows for an examination of bike suitability on commonly used routes. Research findings are directly useful to users, policymakers, and stakeholders in Charleston, SC, Columbia, SC, Lincoln, NE, and beyond, with results anticipated to be transferable to other communities. The intention is to create a blueprint for agencies to inform travelers and implement measures to improve road quality, ultimately making transportation infrastructure safer for all users.

CHAPTER 2

Literature Review

A central focus of transportation planning is to provide equitable access to all users and no user group should bear a disproportionate burden of the negative impacts of the built environment. Kawachi et al.'s work pointed out that neighborhood environments are one of the determinants of public health (*4*). For example, access to multi-use paths and/or bike lanes encourages residents to be physically active (*5*), and achieving recommended levels of physical activity is linked to lower rates of non-communicable diseases and overall mortality rates (*6*). Researchers have found evidence that suggests infrastructure development may propagate race-related health disparities. Melton-Fant (*7*) and Shertzer (*8*) found that Black/African American communities are consistently disadvantaged from a historical and current disinvestment standpoint regarding community design which contributes to higher levels of chronic disease among other race-related health disparities. Experts called for systemic and systematic change through new policies and implementation of existing policies as well as enhanced community inclusion in decision-making through ownership of policy and built environment change (*9*). This study contributes to this call and body of work by providing built-environment data for respective communities to make informed decisions regarding built-environment changes.

In addition to public health and traffic safety (discussed in the Introduction), a poorly built environment also has a negative impact on economic benefits for both individuals and societies as a whole. Mora et al. (10) found that high-income groups in Santiago, Chile have greater access to bike lanes than low-income groups. The authors concluded that fragmentation, inequality, and weak governance play crucial roles in those disparities. Similarly, Anaya-Boig et al. (11) found that the wealthier the population in Barcelona, Spain, the more they have access to cycling infrastructure, especially bike-sharing stations. Lindsay et al. (*12*) found empirical support for cycling advocates' claim that low-income and minority communities across the U.S. have disproportionately low access to bike lanes. They suggested the consideration of social equity in bicycle planning and advocacy. To our knowledge, no study has examined differences in the ride quality of the road infrastructure in or near a college/university in different U.S. cities. Understanding their similarities or differences will provide insights into the differences in cycling activities of respective cities.

The idea of using a bike as a probe vehicle to determine road roughness was investigated by Rizelioglu et al. (*13*). In their study which sought to measure nonmotorized road roughness, they used polyvinylidene fluoride sensors, attached to the front wheel of a mountain bike to capture road roughness through tire-road interaction. Smartphones mounted on bicycles have also been used as sensors to measure road roughness. Kay et al. (*14*) assessed the surface quality of bicycle lanes and paths using crowd data from smartphones carried on bicycles. Power Spectral Density was used to classify the roughness, texture, and structure of road surfaces. To make the sharing of road quality data easier for cyclists, Luedemann and Nascimento (*15*) developed BikeVibes1, an app that cyclists can use to log data regarding the smoothness of their rides. Recognizing that the ride quality on a bicycle path or street can vary from person to person, Ahmed et al. (*16*) investigated the ride quality along different bicycle paths, considering the perception of cyclists. They selected thirteen bicycle paths and streets around the city of Hasselt, Belgium, to measure ride quality. Similar to this study, Jarry et al. (*17*) evaluated cycling comfort using GPS and accelerometer-equipped bicycles in Montreal, Laval, and Longueuil, Canada. They found the participant ID and cyclist speed to be significant predictors of comfort. This study has a similar goal to the work of Ahmed et al. (*16*) and Jarry et al. (17) in quantifying the quality

of bike infrastructure at different locations. Our work collects data in Charleston, SC, Columbia, SC, and Lincoln, NE using a classification method developed by Ho et al. (*18*).

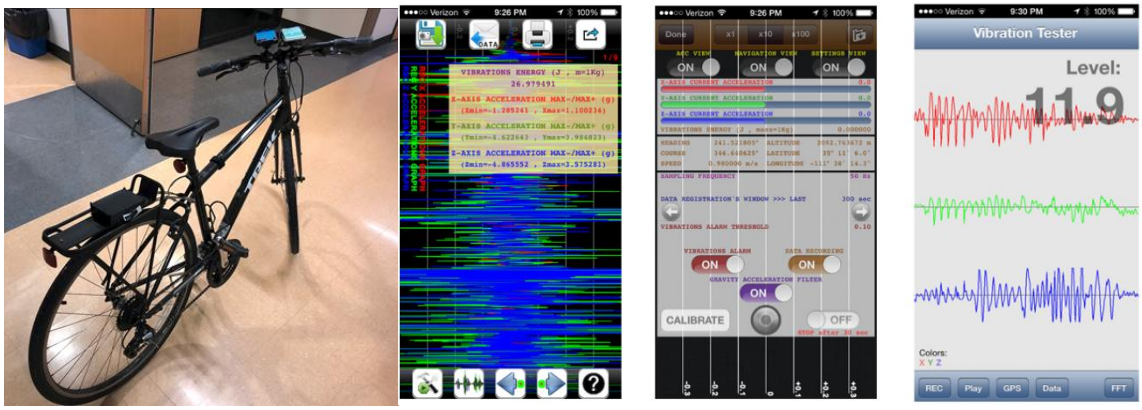
CHAPTER 3

Methods and Data Collection

3.1 Method

This research focuses on evaluating built environment by collecting data to determine pedalcyclist riding quality. Data collection was conducted using both iPhone and Android smartphone applications, recording triaxial acceleration data and GPS coordinates along various routes in three locations: Charleston, SC, Columbia, SC, and Lincoln, NE. Smartphones were mounted at the front of the bicycle as seen in **Figure 3 (A)**. **Figure 3 (B)** shows sample smartphone data collection screens. To ensure an accurate representation of the bike's intrinsic vibrations, Euler's formula was used to reorient the 3-axis acceleration data to the bike's coordinate system, standardizing the data. Data were analyzed and pavement anomalies, such as cracks and potholes, were detected and geocoded. The results were visualized using GIS software, with minor discrepancies observed between Android and iPhone anomaly detection, likely due to variations in GPS accuracy and processing. The comprehensive methodology utilized in this study, shown in **Figure 4**, ensured a thorough evaluation of cycling path quality, contributing to improved urban mobility and infrastructure planning.

As shown in **Figure 4**, the methodology consists of several components, starting with data processing, reorientation, and normalization. It is followed by the model development process which uses an encoder-decoder structure. The model training process uses a loss function and optimizer. Then experiments are performed and validated, showcasing the effectiveness of our model to detect anomalies (i.e., potholes, large and deep cracks). The data collection process, data analysis, autoencoder technique, and training process are elaborated in the following sections.



(A) **(B)**

Figure 3: (A) Instrumented Bicycle Setup; (B) Examples of Data Collection Screens

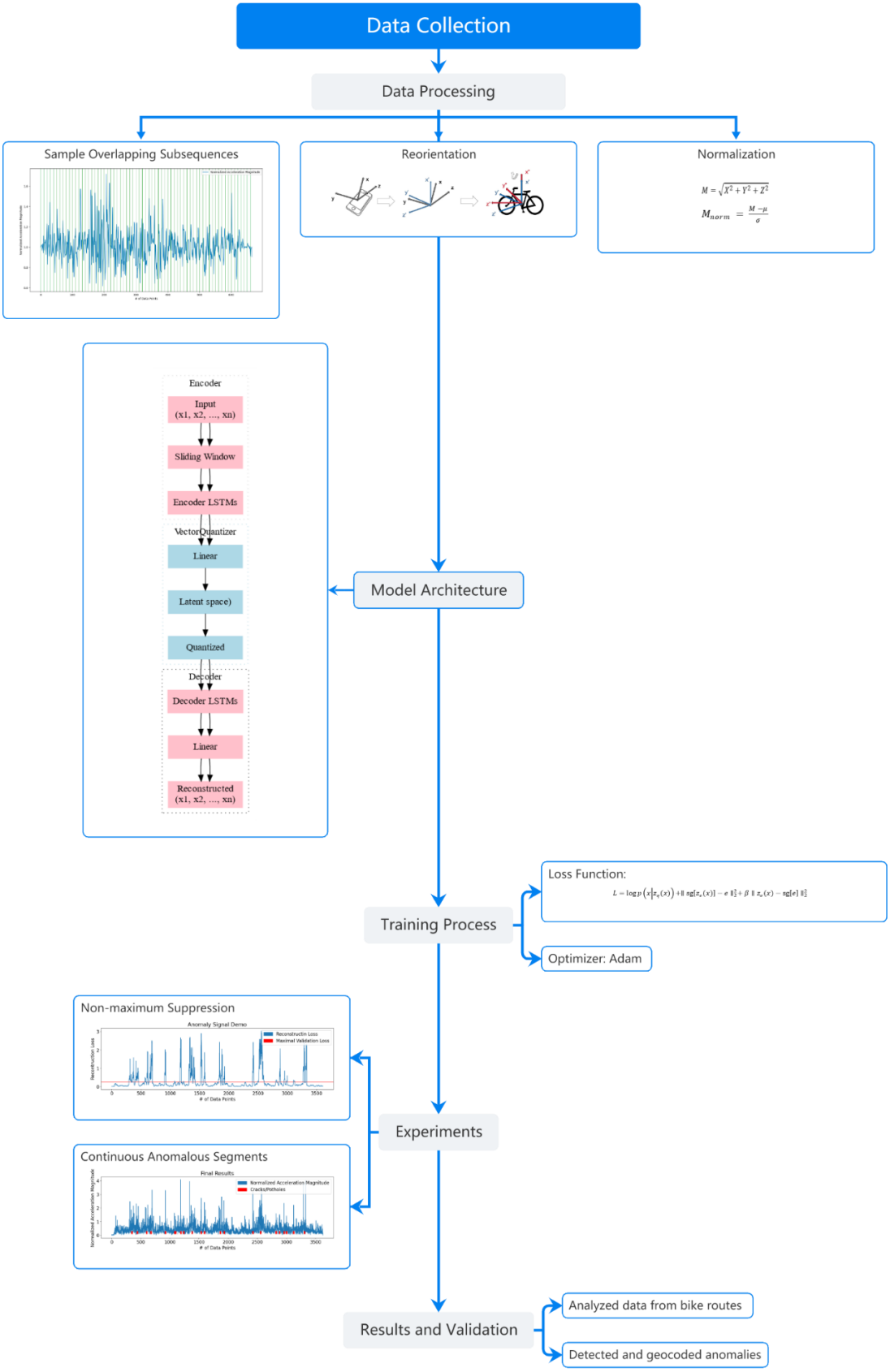


Figure 4: Methodology Flowchart

3.2 Data Collection

Smartphones, with their widespread availability and low cost, are ideal tools for real-time data collection and geospatial analysis, particularly in road condition assessments (*19, 20, 21*). These devices are equipped with precise built-in sensors, such as accelerometers and GPS, enabling detailed environmental monitoring. In this study, we utilized applications named Gauges (*22*) and Sensor Logger (*23*) to collect data at a high resolution of 100 points per second, encompassing metrics such as vibration across X, Y, and Z axes (in g), heading, course (both in degrees), speed (m/s), altitude (meters), and geocoordinates (latitude and longitude in degrees). Additionally, the integration of GPS with the mobile app allows for accurate tracking of bike paths. The collected data were compiled into a CSV file, facilitating easy import into visualization software such as ArcGIS or Google Earth for subsequent analysis. This format ensures compatibility and simplifies the process of integrating spatial data with geospatial analysis tools. To ensure the study's findings are universally applicable, data collection was performed concurrently on two major smartphone platforms, Android and iPhone, thus covering a broad spectrum of operating system environments.

This experiment differentiates between asphalt and concrete pavements to account for material-specific variations in vibration data, acknowledging that the algorithm's performance may vary across different surfaces. Additionally, to assess the robustness of the developed algorithm under varying conditions, data were also collected from bike paths exhibiting different levels of damage—categorized as severe and fair.

CHAPTER 4

Data Analysis

4.1 Data Processing

To evaluate cycling infrastructure across diverse urban environments, data were collected in three cities: Charleston, SC; Columbia, SC; and Lincoln, NE. Each location provided context in terms of road types, traffic volumes, and bicycle infrastructure. The following subsections describe the specific road segments assessed in each city, the methodology used for data collection, and key characteristics of the evaluated routes.

4.1.1 Reorientation

During data collection, smartphones were affixed in holders at the bike's front end (**Figure 4**). Variations in the placement and angle of these devices across multiple sessions introduced discrepancies in reflecting the bike's intrinsic vibrations. To address this, we applied Euler's formula (*24*) to reorient the 3-axis acceleration data from the smartphones to the bike's coordinate system. This standardization ensures that the data accurately represent the impact of road conditions on the bike, both during the training and testing phases of our model.

$$\begin{bmatrix} a_x'' \\ a_y'' \\ a_z'' \end{bmatrix} = R_z(\gamma)R_y(\beta)R_x(\alpha)\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \begin{bmatrix} cos\gamma & -sin\gamma & 0 \\ sin\gamma & cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} cos\beta & 0 & sin\beta \\ 0 & 1 & 0 \\ -sin\beta & 0 & cos\beta \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\alpha & -sin\alpha \\ 0 & sin\alpha & cos\alpha \end{bmatrix}\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$$

Where $\begin{bmatrix} a_x'' \\ a_y'' \\ a_z'' \end{bmatrix}$ is the transformed vector of accelerometer sensor data value along the three axes in the bike coordinate system and the second line of the equation is the calculation of rotation matrices about the X, Y, and Z axes, respectively. $\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$ is the collected three-axis acceleration in the smartphone coordinate system. $\alpha, \beta,$ and γ are the Euler angles about X, Y, and Z axes.

4.1.2 Normalization

Upon reorienting the accelerometer data from the device-specific coordinate system to the bike's frame of reference, the current dataset comprises reoriented values across three axes (x, y, z) along with corresponding geographical coordinates (longitude, latitude). Notably, accelerations in all three axes exhibit significant variations when encountering surface irregularities such as cracks or potholes. Consequently, the combined magnitude of acceleration, calculated using **Equation 1**

$$M = \sqrt{X^2 + Y^2 + Z^2} \quad (1)$$

where X, Y, and Z represent the acceleration values along the x, y, and z axes, respectively, was selected as the input feature. To facilitate consistent model input during both training and testing phases, the magnitude is normalized as shown in **Equation 2**,

$$M_{norm} = \frac{M - \mu}{\sigma} \quad (2)$$

where μ and σ denote the mean and standard deviation of the magnitude, respectively. This standardization, a common preprocessing step in neural network training, enhances the numerical stability of the model and promotes faster convergence.

4.1.3 Sample Overlapping Subsequences.

To ensure comprehensive and continuous data coverage in signal analysis, our sampling strategy incorporated overlapping subsequences. We utilized a sliding window technique, where each subsequence overlapped all elements of the preceding subsequence except the first and last. Given the smartphone's acquisition frequency of 50 Hz, an average bike riding speed of 10.2 mph, and a bike length of approximately 2.0 meters, our sliding window was designed to capture the time series data spanning at least the entire length of the bike. Consequently, all data were resampled into overlapping windows of size 21. This data augmentation strategy enhances the detection of localized surface irregularities, thereby enabling a more detailed and reliable analysis of bike lane conditions.

4.2 Model Architecture

Detecting anomalies in bike paths from acceleration data presents significant challenges due to the complexities of machine learning with unlabeled data. While triaxial acceleration variations encode anomaly patterns, constructing an effective model is complicated by the absence of direct, labeled observations of bike lane conditions during the learning process. A commonly employed solution is the use of autoencoders, which consist of an encoder and a decoder (*25*). The encoder maps input data to a latent space, and the decoder reconstructs the data from this latent representation, aiming to minimize the reconstruction error, as outlined in **Equation 3**.

$$\varphi, \Phi = \ arg\ min_{\varphi,\Phi} ||\chi - (\varphi \circ \Phi)\chi\ ||^2 \tag{3}$$

where χ denotes input data space, and Γ represents latent space, via a function $\varphi: \chi \rightarrow \Gamma$.

Despite its efficiency, the traditional autoencoder often fails to recognize unseen data features, potentially leaving vast regions of its latent space underutilized, which constrains its ability to generate novel data instances. To address these limitations, the variational autoencoder (VAE) introduces probability distribution learning during the encoding process, which enhances the model's ability to generalize by producing not just fixed coding vectors but also parameters that describe the input data distribution. This enables the generation of a richer set of data instances. However, VAEs typically struggle to capture smooth, anomaly-free sequences as effectively as standard autoencoders. This challenge was addressed by Niu et al. (*26*), who segmented the dataset into a training set devoid of anomalies and a test set that includes them.

In our study, we employed a Vector Quantized-Variational Autoencoder (VQ-VAE) to extract data features. VQ-VAE maps each input sample to the nearest discrete vector in a predefined codebook, a process that naturally filters out irrelevant details and noise by focusing on the most significant features for reconstruction (*27*). This selective compression is advantageous for identifying anomalies, such as potholes or other defects, during the inference phase if the reconstruction loss exceeds the established threshold from the training data.

Furthermore, capturing the temporal dependencies among data points is crucial for accurately assessing pavement conditions. Our approach uses Long Short-Term Memory (LSTM) networks, which are superior to conventional Recurrent Neural Networks (RNNs) for modeling long-term dependencies in time-series data (*28*). By integrating LSTM units into the VQ-VAE architecture, our model learns sequence representations in an unsupervised manner, enhancing its predictive accuracy. The architecture of this integrated network is depicted in **Figure 5**.

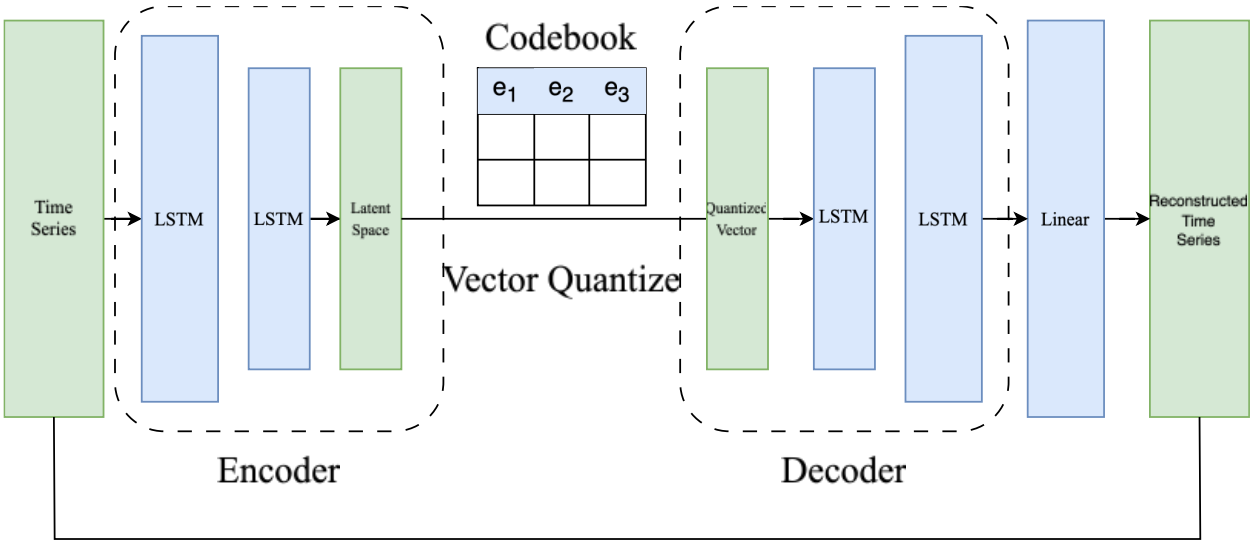


Figure 5: Proposed Deep Learning Neural Network Structure

The network architecture comprises an encoder, a codebook for vector quantization, and a decoder. The encoder utilizes multiple LSTM layers to process time series input, capturing temporal dependencies and distilling this data into a latent space representation that encapsulates key features of the input. This latent representation is then quantized using a pre-defined codebook containing a fixed set of vectors (e.g., e1, e2, e3), which are optimized during training to efficiently represent and compress the data. The quantized data serves as the input for the decoder, which also consists of multiple LSTM layers followed by a linear layer, tasked with reconstructing the original time series. This design ensures both the efficient compression of data for storage or transmission and the faithful reconstruction of the original inputs, critical for applications requiring accurate data recovery.

4.3 Training Process

A multi-stage training procedure was meticulously devised to enhance the performance of LSTM and VQ-VAE models in detecting pavement anomalies on bike routes. Initially, accelerometer vibration readings and GPS position data are normalized and reoriented, standardizing the feature scales to promote model convergence during training. Our loss function, adapted from Oord as outlined in **Equation 4** (*27*), incorporates three components: reconstruction loss, the minimal distance between the embedding vector and encoder outputs, and a commitment loss. Although the original authors reported minimal impact from the commitment loss, we assigned it a constant value of 0.25 to maintain consistency.

$$L = \log p\left(x \middle| z_q(x)\right) + \| \text{sg}[z_e(x)] - e \|_2^2 + \beta \| z_e(x) - \text{sg}[e] \|_2^2 \qquad (4)$$

Furthermore, we pre-initialized the codebook using k-means clustering to accelerate convergence and facilitate the learning of non-anomalous data features, essential for the model's classification tasks. For optimization, the Adam optimizer was selected for its adaptive learning rate capability, which effectively adjusts to different data conditions, enhancing learning efficiency throughout the training phases. Through the entire training process, the reconstruction loss is close to 0 as shown in **Figure 6**.

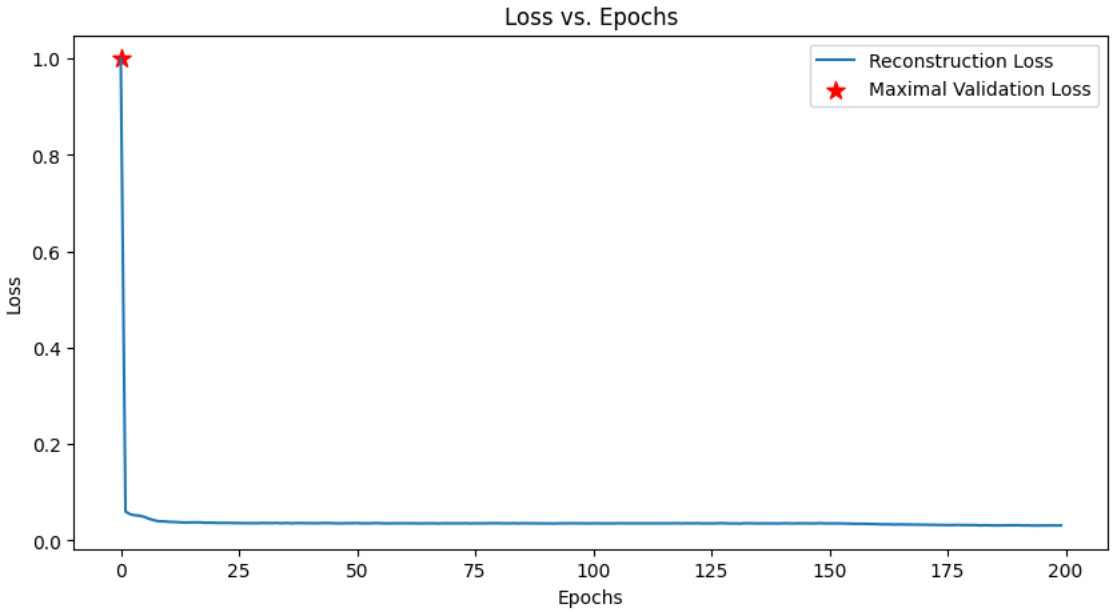


Figure 6: Reconstruction loss vs epochs

CHAPTER 5

Experiments and Algorithm Validation

5.1 Experiments

In this study, we deployed a pre-trained LSTM-VQVAE model for anomaly detection in bike path conditions, building on the foundational window interpolation method previously described (*25*). The procedure involves inputting acceleration magnitudes from a single sliding window into the model to compute the loss. An anomaly is identified if the output loss exceeds the maximum validation loss observed during training, classifying the input sequence as a potential anomaly.

To accurately localize these anomalies, we adopted two distinct strategies, recognizing the complex nature of bike paths, which include features like manholes and expansion joints. The first approach employs non-maximum suppression on the anomaly candidates identified within the sliding windows. By filtering out isolated anomalies who do not have consecutively neighboring anomaly candidates, we pinpoint the anomaly's midpoint coordinates—specifically, the 11th point in the sequence.

In the second approach, we address prolonged anomalous conditions by considering a sequence of 10 consecutive windows (each window consisting of 21 data points and substantial overlap) as a continuous anomalous segment. The segment’s start and end points are defined by the coordinates of the first and last points of the respective windows, effectively capturing the extended nature of road surface anomalies. The result is a comprehensive segment-based output that provides a detailed assessment of the road conditions.

5.2 Algorithm Validation

Upon completing the LSTM-based sliding window algorithm, we conducted several validation tests to assess the method's efficacy in collecting vibration signals and identifying potential road anomalies such as bumps, uneven surfaces, potholes, and cracks. For this purpose, several bike routes in three areas were selected. Following the cycling trials, all collected data were analyzed using a hybrid model that integrates VQ-VAE with the LSTM-based sliding window algorithm. This analysis facilitated the classification of anomalies, with the patterns of potential anomalies for each route being displayed and identified in **Figure 7** (vibration data from Charleston, SC).

Reconstruction losses exceeding the maximum validation loss, as denoted by the horizontal red line in **Figure 7**, are identified as anomalies. These anomalies typically align with the crack and pothole patterns observed in the original signal. The final summarized boxes containing detected candidate anomalies are correlated with these patterns. Selected anomalies were extracted from the raw dataset, geocoded, and visually represented using GIS software; **Figure 8** illustrates one of the selected routes as an example. Given that road surface anomalies usually extend over a continuous stretch rather than being isolated incidents—particularly on cycle paths where the road surface condition is poor—the research team consolidated all adjacent anomalies into "zones" where consistent detection was achieved across both Android and iPhone devices. Furthermore, for isolated anomalies such as cracks, the proposed algorithm demonstrated consistent detection across both platforms in terms of the vibration data collected.

Although our proposed hybrid model, which combines LSTM and VQ-VAE, has demonstrated superior performance in predicting road anomalies, **Figure 8** reveals a minor discrepancy in anomaly detection between Android and iPhone devices. In the figure, blue stars represent anomalies detected using iPhone, while red dots indicate anomalies detected using Android. Specifically, a small percentage of anomalies identified by Android phones were not detected by

iPhones, and vice versa. Upon further analysis, it became apparent that the anomaly patterns detected by both types of smartphones are remarkably similar once discrepancies due to operating systems are accounted for. This observation suggests that our algorithm effectively achieves its anomaly detection objectives. The differences likely stem from variations in GPS accuracy and the disparate processing of GPS data by different smartphone operating systems.

In the following section, details regarding the data collection areas are presented.

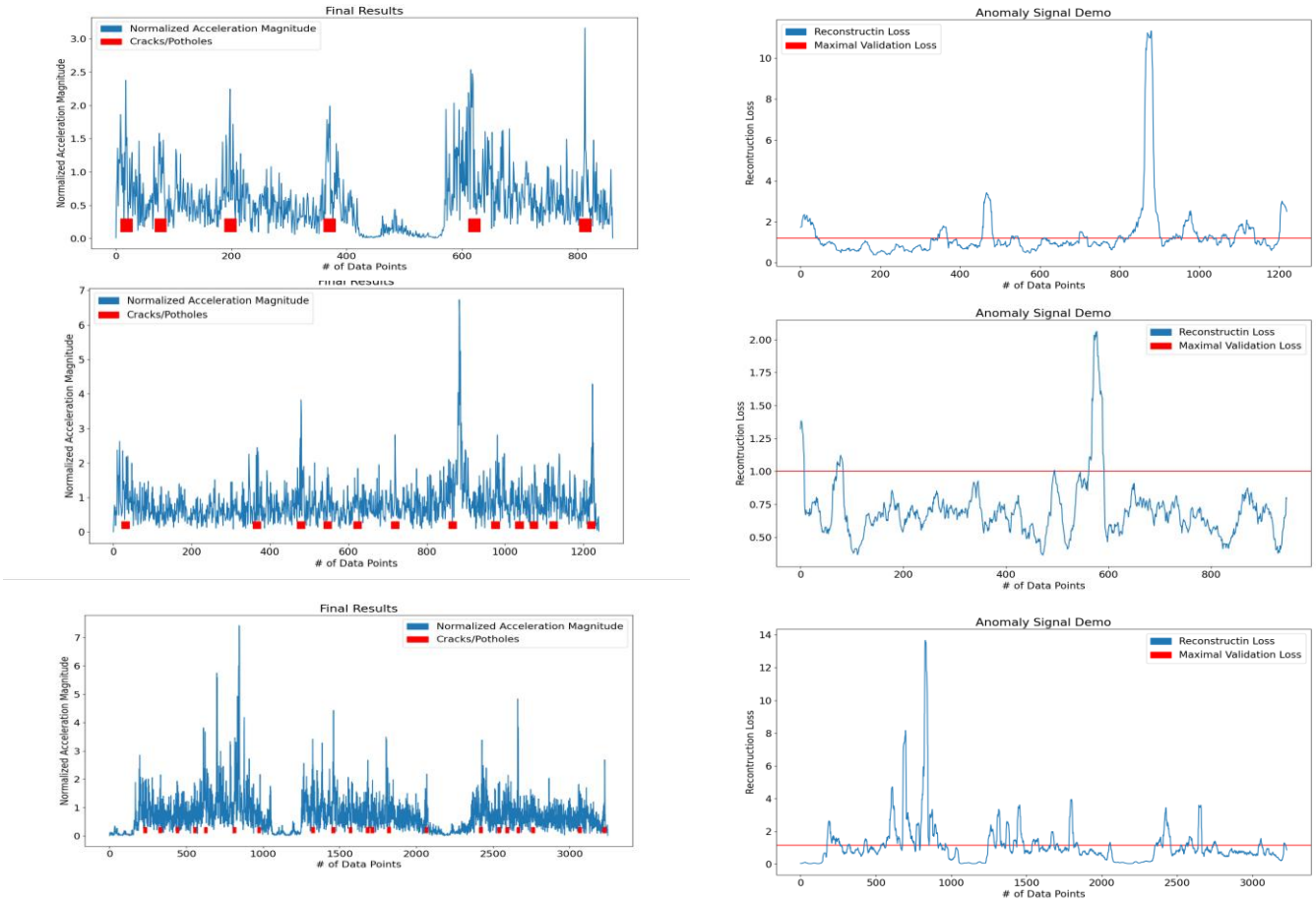


Figure 7: Idenfied Anomaly Pattern (left) and Corresponding Locations (right)

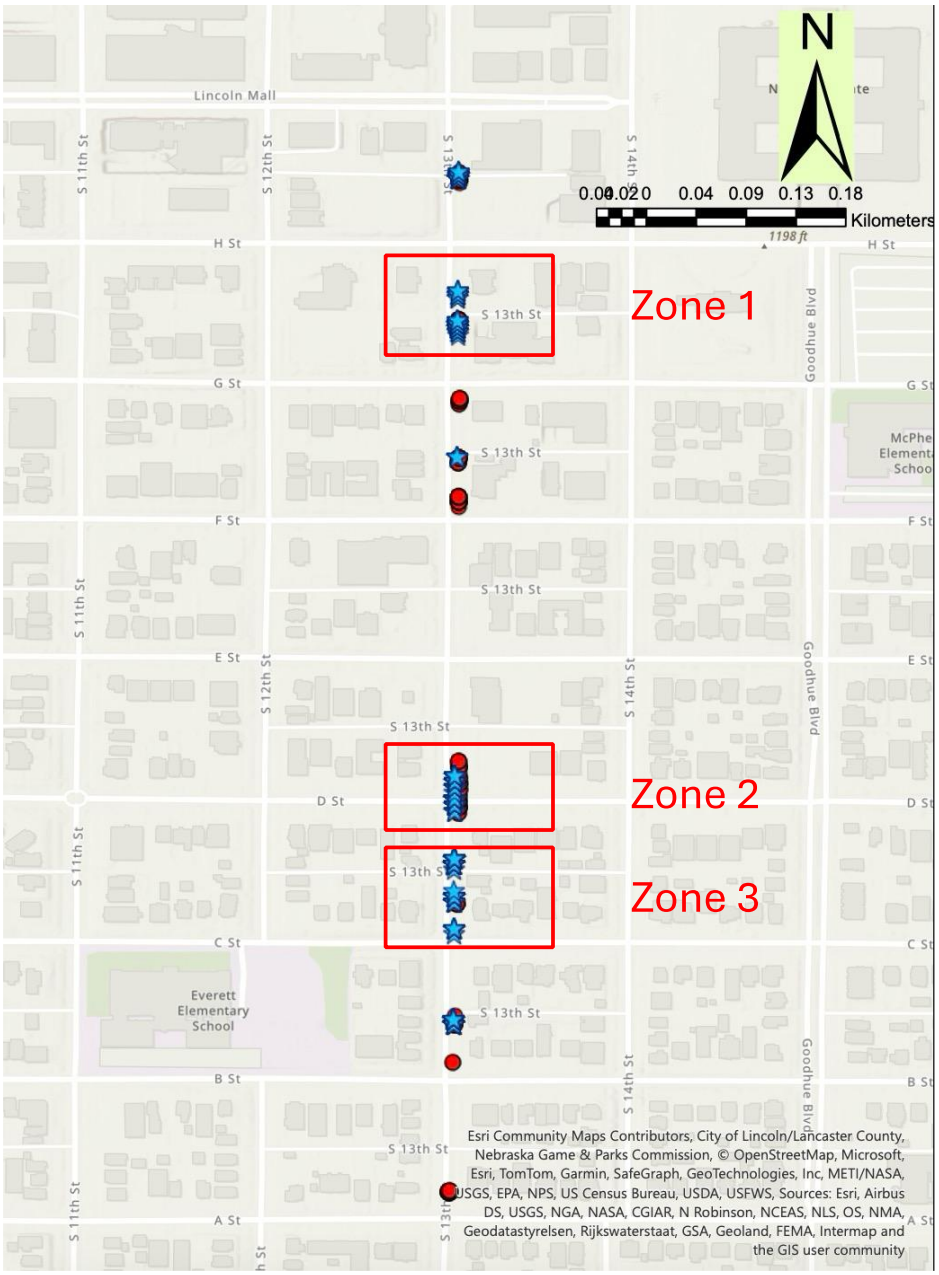


Figure 8: Sample ArcGIS output

CHAPTER 6

Data Collection Areas

This project involved data collection in three distinct urban locations: downtown Charleston, SC, Columbia, SC, and Lincoln, NE. These locations were selected due to their proximity to the researchers' campuses and their diverse urban cycling environments. Data collection encompassed various road segments, including multiuse paths, mixed-traffic areas, and roads with dedicated bike lanes.

In each city, specific segments were chosen to provide a comprehensive overview of the cycling infrastructure. These segments were selected based on their usage patterns and the variety of conditions they presented to cyclists. Detailed descriptions of the segments, along with the analysis are detailed in the following sections, highlighting key findings and infrastructure recommendations.

6.1 Charleston, SC

Two researchers rode their instrumented bikes on 23 roadway segments in Charleston to evaluate their bike suitability. The road segments are presented in **Table 1** along with their length and AADT (Annual Average Daily Traffic). Each road segment was ridden 4 to 6 times. One researcher had an Android phone and the other an iPhone. The data collection took place over several days during May 2023, December 2023, and January 2024. **Figure 9** shows the routes that covered the 23 segments. In total, a total of 14.72 miles were evaluated. There is little elevation change along the routes, except for the segment involving the Ravenel Bridge. The Charleston side of the bridge has an average grade of 2.5%, which is a manageable grade for most bicyclists.

Table 1: Data Collection Segments in Charleston, SC

Segment No	Type	Description/ Name	Segment Limits	Length (Miles)	AADT
1	Urban Principal Arterial	King St.	From Huger St. to Spring St.	0.64	10400
2	Urban Principal Arterial	King St.	From Spring St. to Hutson St.	0.45	11100
3	Urban Principal Arterial	King St.	From Hutson St. to Calhoun St.	0.12	15000
4	Urban Principal Arterial	King St.	From Calhoun St. to Hasell St.	0.32	3600
5	Urban Major Collector	King St.	From Hasell St. to Fulton St.	0.14	3600
6	Urban Major Collector	King St.	From Fulton St. to Board St.	0.23	3600
7	Urban Major Collector	King St.	From Broad St. to Murray Blvd.	0.47	3600
8	Urban Local	Murray Blvd.	From Lenwood Blvd. to E. Battery St.	0.38	2800
9	Urban Local	S. Battery St.	From Tradd St. to Lenwood Blvd.	0.43	2400
10	Multi-use path	Lockwood Blvd. multiuse path	From Broad St. to Fishburne St.	1.6	N/A

Segment No	Type	Description/ Name	Segment Limits	Length (Miles)	AADT
11	Urban Principal Arterial	Calhoun St.	From Courtenay Dr. to Smith St.	0.48	22000
12	Urban Principal Arterial	Calhoun St.	From Smith St. to E. Bay St.	0.83	21800
13	Urban Principal Arterial	Broad St.	From Lockwood Dr. to E. Bay St.	1.00	11400
14	Urban Principal Arterial	Meeting St.	From Broad St. to S. Battery St.	0.43	18800
15	Urban Major Collector	E. Battery/E. Bay St.	From Murray Blvd. to Broad St.	0.54	4500
16	Urban Principal Arterial	E. Bay St.	From Broad St. to N. Market St.	0.31	21100
17	Urban Principal Arterial	E. Bay St.	From N. Market St. to Calhoun St.	0.54	21100
18	Urban Minor Arterial	E. Bay St.	From Calhoun St. to South St.	0.4	21100
19	Urban Principal Arterial	E. Bay St. multiuse path	From South St. to Cooper St.	0.42	21100
20	Multiuse path	Cooper R. Bridge multiuse path	From Cooper St. to mid-span Cooper River shipping channel	2.5	N/A
21	Urban Minor Arterial	Morrison Dr. partial bike lane/sharrows	From Grace Bridge St. to Huger St. (adjacent major commercial dev. projects have changed NB bike lane)	0.41	13500
22	Urban Local	Huger St.	From Morrison Dr. to President St.	0.98	6400
23	Multi-use Path	Hampton Park Multi-Use Path	Circular Path	1.1	N/A

Figure 9: Routes qualitatively assessed for their bike suitability

6.2 Columbia, SC

The road segments were data was collected are presented in **Table 2.**

Table 2: Data Collection Segments in Columbia, SC

Segment No	Type	Description/ Name	Segment Limits	Length (Miles)	AADT
1	Urban Principal Arterial	Greene St.	From Huger St. to Pulaski St.	0.1875	3900
2	Urban Principal Arterial	Greene St.	From Pulaski St. to Lincoln St.	0.1875	3600
3	Urban Principal Arterial	Greene St.	From Lincoln St. to Assembly St.	0.1875	4400
4	Urban Principal Arterial	Greene St.	From Assembly St. to Main St.	0.1875	3900
5	Urban Principal Arterial	Greene St.	From Main St. to Sumter St.	0.1875	700
6	Urban Principal Arterial	Greene St.	From Sumter St. to Pickens St.	0.1875	0

Segment No	Type	Description/ Name	Segment Limits	Length (Miles)	AADT
7	Urban Principal Arterial	Greene St.	From Pickens St. to Barnwell St.	0.1875	3200
8	Urban Principal Arterial	Greene St.	From Barnwell St. to Harden St.	0.1875	3200
9	Urban Local	Airport Blvd.	From Jim Hamilton Blvd to Commerce Dr.	0.25	1400
10	Urban Local	Commerce Dr.	From Airport Blvd. to Howe St.	0.25	200
11	Urban Local	S Edisto Ave.	From Howe St. to Mitchell St.	0.25	200
12	Urban Local	Mitchell St.	From S Edisto Ave. to Superior St.	0.25	200
13	Urban Major Collector	Superior St.	From Mitchell St. to S Edisto Ave.	0.25	1675
14	Urban Major Collector	Superior St.	From S Edisto Ave. to Holt Dr.	0.25	1675
15	Urban Major Collector	Superior St.	From Holt Dr. to Graymont Ave.	0.25	225
16	Urban Local	Superior St.	From Graymont Ave. to S Maple St.	0.25	225
17	Urban Local	Superior St.	From S Maple St. to S Holly St.	0.25	225
18	Urban Major Collector	S Holly St.	From Superior St. to Montgomery Ave.	0.25	700
19	Urban Local	Montgomery Ave.	From S Holly St. to S Ott Rd.	0.25	300
20	Urban Local	S Ott Rd.	From Montgomery Ave. to S Bonham Rd.	0.25	300
21	Urban Local	S Ott Rd.	From S Bonham Rd. to Jim Hamilton Blvd	0.25	300
22	Urban Principal Arterial	Jim Hamilton Blvd	From S Ott Rd. to Airport Blvd.	0.25	1400

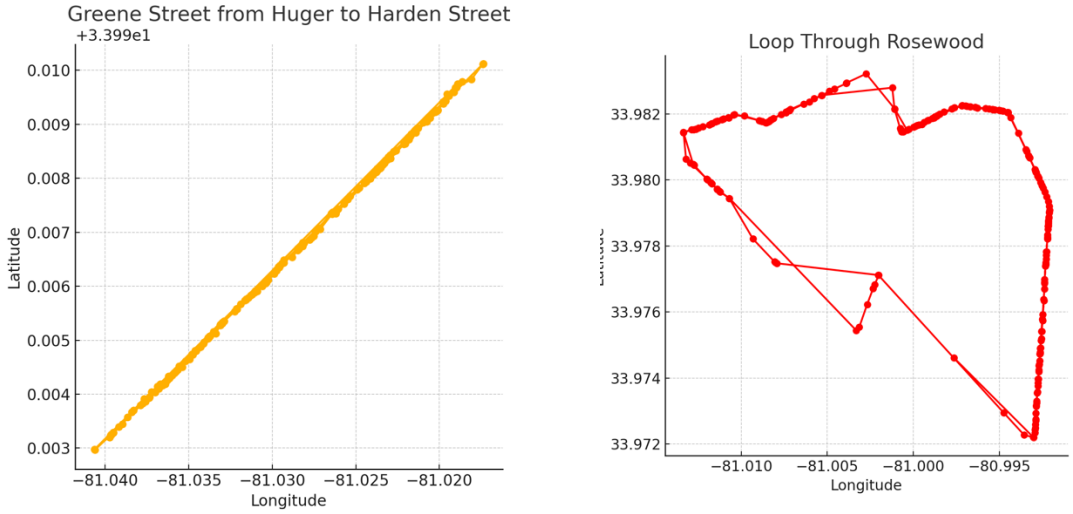
Greene Street from Huger to Harden Street

This 1.5-mile route primarily passes through urban areas with multiple turns and intersections. It experiences heavy traffic, with AADT data ranging from 700 to 4400, necessitating vigilance and adherence to traffic signals by cyclists. This route is suitable for experienced cyclists who can handle the complexities of urban traffic. Data collection for this segment was conducted using

bicycles equipped with smartphones running mobile applications. The data was collected between 04:05:42 PM and 04:53:30 PM on December 8, 2023. The latitude and longitude path map is shown in **Figure 10 (A)**.

Loop Through Rosewood

This 3.5-mile loop offers a relatively straight path with few sharp turns, making it accessible for all types of cyclists. The route mainly traverses flat terrain, ensuring accessibility for most riders. It passes through residential and commercial areas, making it ideal for both leisure rides and commuting. Traffic volume on different segments, according to AADT data, ranges from 200 to 1675. Data for this route was collected between 04:19:22 PM and 04:52:28 PM on December 12, 2023. The latitude and longitude path map is shown in **Figure 10 (B)**.



(A) **(B)**

Figure 10: Latitude and Longitude Path Maps (A) Greene St from Huger to Harden St. and (B) Loop through Rosewood

6.3 Lincoln, NE

Details of the data collection segments in Lincoln, NE are presented in **Table 3**.

Table 3: Data Collection Segments in Lincoln, NE

Segment No	Type	Description/ Name	Segment Limits	Length (Miles)	AADT
1	Arterial	16th St.	From X St. to W St.	0.1	3170
2	Arterial	16th St.	From W St. to Vine St.	0.1	3170
3	Arterial	16th St.	From Vine St. to U St.	0.1	5580
4	Arterial	16th St.	From U St. to S St.	0.1	5580
5	Arterial	16th St.	From S St. to R St.	0.1	5580
6	Local street	11th St.	From LINCOLN mall to H St.	0.1	2830
7	Local street	11th St.	From H St. to G St.	0.1	2830
8	Local street	11th St.	From G St. to F St.	0.1	2830
9	Local street	11th St.	From F St. to E St.	0.1	2830
10	Arterial	13th St.	From LINCOLN mall to H St.	0.1	5760

Segment No	Type	Description/ Name	Segment Limits	Length (Miles)	AADT
11	Arterial	13th St.	From H St. to G St.	0.1	5760
12	Arterial	13th St.	From G St. to F St.	0.1	6510
13	Arterial	13th St.	From F St. to E St.	0.1	6510
14	Designated bike only lanes	N St	From 9th to 10th St.	0.1	4380
15	Designated bike only lanes	N St	From 9th to 10th St.	0.1	3820
16	Designated bike only lanes	N St	From 10th to 11th St.	0.1	4320
17	Designated bike only lanes	N St	From 11th to 12th St.	0.1	4040
18	Designated bike only lanes	N St	From 12th to 13th St.	0.1	3060
19	Designated bike only lanes	N St	From 13th to 14th St.	0.1	4400
20	Designated bike only lanes	N St	From 14th to 15th St.	0.1	2890
21	Designated bike only lanes	N St	From 15th to 16th St.	0.1	2730
22	Designated bike only lanes	N St	From 16th to 17th St.	0.1	2830
23	Designated bike only lanes	N St	From 17th to 18th St.	0.1	2210
24	Designated bike only lanes	N St	From 18th to 19th St.	0.1	2620
25	Designated bike only lanes	N St	From 19th to 20th St.	0.1	3140
26	Designated bike only lanes	N St	From 20th to 21st St.	0.1	3140
27	Designated bike only lanes/paths	Billy Wolff Trail	N/A	0.5	N/A

Figure 11 presents a detailed map of bike routes in Lincoln, Nebraska, highlighting the distribution and condition of bike infrastructure. Local streets are indicated in red, while designated bike-only lanes are differentiated by surface material—concrete in red and asphalt in green. Arterial roads are classified based on their conditions: a bike route in fair condition is shown in dark blue, and one that is in poor condition are depicted in light blue. This mapping is essential for assessing the spatial distribution and quality of cycling paths, identifying critical areas for infrastructure improvement. It also supports our comprehensive analysis of how pavement performance varies under different conditions, aligning with the core objectives of our study.

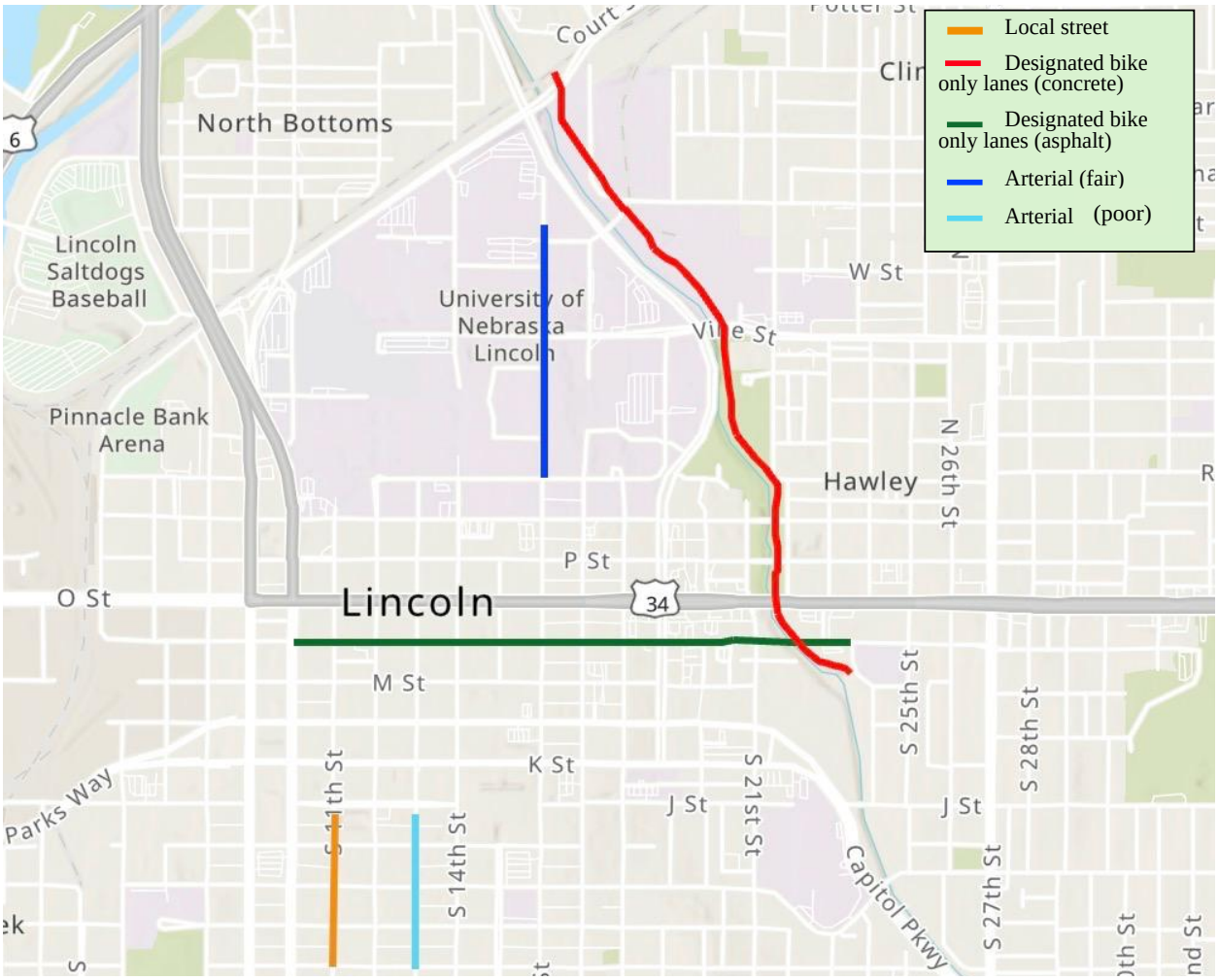


Figure 11: Routes qualitatively assessed for bike suitability around UNL

CHAPTER 7

Results, Discussion, and Conclusions

All data collected were analyzed using the proposed algorithm, and results were mapped using ArcGIS and Google Earth Pro. **Figure 12** shows sections of how the pavement defects were visualized in the three different collection areas. Blue stars represent results from analyzing data collected via iPhone, while the red circles show results from data collected using Android.

In Charleston, the individuals collecting data each had one phone mounted on the bicycle—one with an iPhone and one with an Android. The rider with the iPhone was more adventurous, preferring to ride on the rightmost travel lane, while the other rider was more conservative, riding on the sidewalk when possible. Therefore, the Charleston data points represent different locations on the pavement, which is why the points on the map do not overlap.

To ensure the finalized algorithm performed well, algorithm validation was conducted. In Charleston selected segments were verified by individuals who revisited the field to confirm that the identified points indeed had some kind of pavement defect. Over several miles of streets, the accuracy of the algorithm was over 90%. In Columbia, SC, a more detailed algorithm evaluation was conducted using Python. This method identified relevant timestamps in cycling videos, extracting 430 frames, of which 351 contained cracks, resulting in an accuracy of 81.63%. Overall, the algorithm was found to perform at an acceptable level.

The data from all areas highlight the varying conditions and challenges faced by cyclists. Urban routes with heavy traffic seem to be not suitable for inexperienced cyclists. In contrast, multi-use paths and roads with some type of cycling infrastructure offer better conditions, making them accessible for all cyclists and ideal for both leisure rides and commuting. These findings underscore the importance of targeted infrastructure improvements to enhance safety and accessibility for cyclists in different urban environments.

Additionally, the study revealed that discrepancies in GPS accuracy and processing between iPhone and Android devices did not significantly affect the overall performance of the algorithm. This validation confirms that the proposed methodology is robust and reliable for detecting pavement defects, contributing to improved urban mobility and infrastructure planning.

(A) Charleston, SC - Calhoun and Rutledge

(B) Columbia SC – Greene St. (In front of Preston Residential College)

(C) Lincoln, NE – N Street (Between Centennial Mall St and S 16 St)

Figure 12: Results Representation using ArcGIS and Google Earth Pro

The study successfully utilized motion and vibration sensors mounted on bicycles to assess the bike suitability of various transportation infrastructure segments in Charleston, SC, Columbia, SC, and Lincoln, NE. The results demonstrated that our proposed algorithm performs well. The findings indicated that most urban routes with heavy traffic require infrastructure improvements to enhance safety and accessibility for cyclists. The high accuracy of pavement distress prediction using image analysis suggests that such technologies can be effectively integrated into future infrastructure monitoring efforts. These findings underscore the value of targeted infrastructure improvements to enhance the safety and convenience of cycling routes based on specific conditions and challenges. This research provides valuable insights for policymakers and stakeholders to implement targeted improvements, ultimately making transportation infrastructure safer and more accessible for all users.

REFERENCES

1) National Highway Traffic Safety Administration (NHTSA). Traffic Safety Facts 2020 Data: Bicyclists and Other Cyclists. June 2022. Available from: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813322
2) South Carolina Department of Transportation. South Carolina Pedestrian and Bicycle Safety Action Plan: Final Report. May 2022. Available from: https://www.scdot.org/projects/pdf/SC%20Pedestrian%20and%20Bicycle%20Safety%20Action%20Plan.pdf
3) Vehicle Technologies Office. FOTW #1230, March 21, 2022: More than Half of all Daily Trips Were Less than Three Miles in 2021. Available from: https://www.energy.gov/eere/vehicles/articles/fotw-1230-march-21-2022-more-half-all-daily-trips-were-less-three-miles-2021
4) Kawachi I, Berkman LF, editors. Neighborhoods and health. Oxford University Press; 2003.
5) Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. *Epidemiol Rev*. 2009;31(1):7-20.
6) Gullón P, Bilal U, Cebrecos A, Badland HM, Galán I, Franco M. Intersection of neighborhood dynamics and socioeconomic status in small-area walkability: the Heart Healthy Hoods project. *Int J Health Geogr*. 2017;16(1):21.
7) Melton-Fant C. Relationship between state preemption of inclusionary zoning policies and health outcomes: Is there disparate impact among people of color? Housing Policy Debate. 2020;30(6):1056-65.
8) Shertzer A, Twinam T, Walsh RP. Zoning and segregation in urban economic history. *Reg Sci Urban Econ.* 2022;94:103652.
9) Prochnow T, Valdez D, Curran LS, Brown CT, Sammons Hackett D, Auld ME. Multifaceted scoping review of Black/African American transportation and land use expert recommendations on activity-friendly routes to everyday destinations. *Health Promot Pract.* 2024;25(2):293-308.
10) Mora R, Truffello R, Oyarzún G. Equity and accessibility of cycling infrastructure: An analysis of Santiago de Chile. *J Transp Geogr.* 2021;91:102964.
11) Anaya-Boig E, Cebollada À, Castelló Bueno M. Measuring spatial inequalities in the access to station-based bike-sharing in Barcelona using an Adapted Affordability Index. *J Transp Geogr.* 2022;98:103267.
12) Braun LM, Rodriguez DA, Gordon-Larsen P. Social (in)equity in access to cycling infrastructure: Cross-sectional associations between bike lanes and area-level sociodemographic characteristics in 22 large US cities. *J Transp Geogr.* 2019;80:102544.
13) Rizelioğlu M, Arslan T, Yigit E, Yazıcı M. Using a Bike as a Probe Vehicle: Experimental Study to Determine Road Roughness with Piezoelectric Sensors. *J Infrastruct Syst.* 2024;30(3):04024018.
14) Massow K, Maiwald F, Thiele M, Heimendahl J, Protzmann R, Radusch I. Crowd-Based Road Surface Assessment Using Smartphones on Bicycles. In: 2024 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA). *IEEE*; 2024. p. 1-8.
15) Luedemann K, Nascimento MA. Bikevibes: An app for crowdsourcing open road quality data from a cyclist perspective. In: Proceedings of the *15th ACM SIGSPATIAL International Workshop on Computational Transportation Science*. 2022. p. 1-4.

16) Ahmed T, Pirdavani A, Wets G, Janssens D. Evaluating cyclist ride quality on different bicycle streets. *Transp Res Procedia*. 2024;78:586-93.
17) Jarry V, Apparicio P, Gelb J. Bumpy rides: an extensive accelerometer-based cycling infrastructure survey. *Transp Res Rec*. 2023;2677(3):1217-29.
18) Ho C-H, Qiu P, Zhang Y, Ren K. A Generic Deep Learning–Based Computing Algorithm in Support of the Development of Instrumented Bikes. *ASCE OPEN Multidiscip J Civ Eng*. 2024;2(1):04024003.
19) Salau HB, Onumanyi AJ, Aibinu AM, Onwuka EN, Dukiya JJ, Ohize H. A Survey of Accelerometer-Based Techniques for Road Anomalies Detection and Characterization. *Int J Electr Syst Autom.* 2019;3(1):1.
20) RoADS: A Road Pavement Monitoring System for Anomaly Detection Using Smart Phones. In: International Conference on Advances in Computing and Communication. *Springer*; 2016. p. 306-15.
21) Lee T, Chun C, Ryu S-K. Detection of Road-Surface Anomalies Using a Smartphone Camera and Accelerometer. Sensors. 2021;21(2):536.
22) Gauges. App Store. Available from: https://apps.apple.com/us/app/gauges/id1056649790
23) Choi K. Sensor Logger [Google Play]. 2023. Available from: https://play.google.com/store/apps/details?id=com.kelvin.sensorapp&hl=en
24) Slabaugh GG. Computing Euler angles from a rotation matrix. Acta Mech. 2000;2000:39–63.
25) A Generic Deep Learning–Based Computing Algorithm in Support of the Development of Instrumented Bikes. ASCE OPEN: *Multidisciplinary Journal of Civil Engineering*. 2024 [cited 2024 Jul 06];2(1). Available from: https://ascelibrary.org/doi/full/10.1061/AOMJAH.AOENG-0025
26) Niu Z, Yu K, Wu X. LSTM-Based VAE-GAN for Time-Series Anomaly Detection. Sensors. 2020;20(13):3738.
27) Van Den Oord A, Vinyals O, Kavukcuoglu K. Neural discrete representation learning [Internet]. arXiv preprint arXiv:1711.00937. 2017 [cited 2024 Jul 06]. Available from: https://arxiv.org/abs/1711.00937
28) Bayer JS. Learning Sequence Representations [dissertation on the Internet]. Technische Universität München; 2015 [cited 2024 Jul 06]. Available from: https://mediatum.ub.tum.de/1256381