AUTOMATING CONSTRUCTION SUPPLY CHAIN USING ML: REVOLUTIONZNING PAKISTAN'S CONSTRUCTION SUPPLY CHAIN WITH AI AND ML

FINAL YEAR PROJECT UG – 2020

By

ABDULLAH BIN TANVEER (G.L)	341338
M. TALHA SATTAR	292750
AYAN NADEEM BHATTI	342762
UMAR FAROOQ	341183

NUST Institute of Civil Engineering (NICE)

School of Civil and Environmental Engineering (SCEE)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

2024

CERTIFICATION

This is to certify that the

Final Year Project Titled

AUTOMATING CONSTRUCTION SUPPLY CHAIN USING ML: REVOLUTIONZNING PAKISTAN'S CONSTRUCTION SUPPLY CHAIN WITH AI AND ML

Submitted By

ABDULLAH BIN TANVEER	341338
M. TALHA SATTAR	292750
AYAN NADEEM BHATTI	342762
UMAR FAROOQ	341183

has been accepted towards the requirements. for the undergraduate degree

in

BACHELOR OF ENGINEERING IN CIVIL ENGINEERING

Dr. Muhammad Usman Hassan

Assistant Professor

Head of Department Construction Engineering and Management

NUST Institute of Civil Engineering (NICE)

School of Civil and Environmental Engineering (SCEE)

National University of Sciences and Technology (NUST), Islamabad, Pakistan

ACKNOWLEDGEMENTS

We express our gratitude to Allah Almighty for enabling us to successfully conclude this research endeavor, granting us the strength, fortitude, and persistence needed. This accomplishment would not have been possible without the unwavering guidance and support of our Project Advisor, Dr. Muhammad Usman Hassan, and the esteemed faculty of the CEM department, especially Sir Abdullah.

We are also grateful to our industry contacts who supported us throughout the project and provided valuable feedback, helping us refine our work to better align with industry requirements.

ABSTRACT

Supply chain management in the construction industry is very critical phenomenon. However, challenges in the supply chain often led to delays, cost overruns, and quality issues in projects. This research focuses on addressing the significant challenges faced by the construction industry in Pakistan's Supply Chain Management, particularly in the realm of material estimation, procurement inefficiencies, and resource allocation. The study adopted a quantitative iterative experimental approach by utilizing Machine Learning AI, namely YOLO V8 and python coding to automate the process of Cost Estimation in construction supply chain. The study aims to serve as a proof for concept that "Detection using ML algorithms" can be used to automate Cost Estimation. The YOLOv8 model was trained on 5 blueprint components (walls, floor, columns, windows, and doors). Python codes were used to extract bounding box coordinates which were further used for cost estimation. The research findings show data that yielded high accuracy percentages, establishing that with better ML training sets, 100% accuracy of cost estimations is achievable.

Key Words: Cost estimation, AI, Machine Learning, Construction Supply Chain

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1. Introduction:

Worldwide, the construction industry plays a vital role in infrastructure development. According to a study by (GlobalData, 2023), in Pakistan alone, the construction industry market size was USD 17.4 Billion in 2022 and is estimated to have an Average Annual Growth Rate of 5% from 2022-2027. However, even in this modern era of constant development, the construction supply change industry faces significant challenges in material estimation, procurement, and resource allocation. A survey conducted by (Ineight, 2023) suggested that 70% of construction projects experience delays and disturbances in supply chain. An ineffective supply chain can cause delays. cost overruns and quality issues in a project (Badea et al., 2014). One prevalent issue could be the inadequate support and minimal active participation which leads to a deficiency in communication and information exchange among stakeholders (Kim & Nguyen, 2022). Moreover, the fragmented nature of the construction industry has been associated with a lack of coordination, trust and emphasis on the client objectives (SRM Riazi et al., 2020). These challenges underscore the urgent need for innovative solutions to revolutionize the construction supply chain. To address these issues, the integration of Artificial Intelligence (AI) and Machine Learning (ML) technologies presents a promising solution. Research by (Kudirat Ayinla et al., 2023) has used ML and ANN for cost estimation and showed how a vast dataset is required and that early cost estimation using AI may offer some advantages. Similarly, another research project by (Jain et al., 2023) showed how using COCOMO model, cost estimation may be achieved. A gap was observed that such projects aren't implemented in Pakistan. Also, the research do not use AI and Machine Learning to detect components directly off of blueprints. This research focuses on the implementation of YOLO v8 ML AI by Ultralytics, with the aim to improve cost estimation techniques in Construction Supply Chain of Pakistan. By training YOLO v8 on construction blueprints, it becomes capable of identifying and classifying components such as walls, floors, columns, windows, and doors with remarkable precision. This includes not only identifying structural components but also estimating the required number of bricks, cement bags, and associated costs. Our team used 3000 images as dataset and remained restricted to 5 blueprint components only.

1.1. What is Artificial Intelligence:

Artificial Intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think and learn like humans. It involves the development of algorithms that can perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation.

1.2. Types of Artificial Intelligence:

1.2.1. Narrow AI (Weak AI):

Narrow AI, also known as weak AI, is designed to perform a narrow task or a specific set of tasks. These AI systems are focused on a particular problem and excel at it. Examples include:

- Virtual Personal Assistants: Siri, Alexa, and Google Assistant are examples of narrow AI that assist users with tasks such as scheduling appointments, setting reminders, and answering queries.

- Recommendation Systems: Platforms like Netflix and Amazon use narrow AI algorithms to analyze user preferences and behavior to recommend movies, products, and services.
- Image Recognition Software: AI-powered image recognition software, such as Google Photos or facial recognition systems, can identify objects, people, and scenes in images with high accuracy.

1.2.2. General AI (Strong AI):

General AI, also referred to as strong AI, is a hypothetical form of AI that possesses the ability to understand, learn, and apply knowledge across different domains, much like a human being. Although true general AI has not yet been achieved, researchers continue to explore the possibilities of creating such systems.

1.3. Different Training Methods of Artificial Intelligence:

1.3.1. Supervised Learning:

Supervised learning involves training a model on a labeled dataset, where each input is associated with the correct output. Examples include:

- Image Classification: Convolutional Neural Networks (CNNs) trained on labeled image datasets, such as ImageNet, can classify images into predefined categories with high accuracy.
- Speech Recognition: Models trained on transcribed speech data, such as the Deep Speech model developed by Baidu, can accurately transcribe spoken language into text.

1.3.2. Unsupervised Learning:

Unsupervised learning involves training on unlabeled data, where the algorithm discovers patterns and structures on its own. Examples include:

- Clustering: K-means clustering algorithms can group similar data points together without prior knowledge of class labels.
- Dimensionality Reduction: Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE) are unsupervised learning techniques used to reduce the dimensionality of data while preserving important relationships among data points.

1.3.3. Reinforced Learning:

Reinforcement learning involves training an agent to interact with an environment and learn from feedback in the form of rewards or penalties. Examples include:

- Game Playing: DeepMind's AlphaGo, trained using reinforcement learning, defeated world champion Go players by learning optimal strategies through gameplay.
- Robotics: Robots trained with reinforcement learning algorithms can learn to perform complex tasks such as grasping objects, navigating environments, and manipulating tools.

1.4. History of Artificial Intelligence and Its Applications:

Artificial intelligence (AI) has a rich history dating back to the mid-20th century, marked by significant milestones and breakthroughs that have shaped its development. From early conceptualizations to practical applications, AI has evolved into a transformative technology with diverse uses across various industries.

1.4.1. Early Developments in AI:

The origins of AI can be traced back to the Dartmouth Conference in 1956, where the term "artificial intelligence" was first coined by John McCarthy and his colleagues. The conference laid the foundation for AI as an interdisciplinary field, bringing together researchers from computer science, mathematics, psychology, and other disciplines to explore the possibilities of creating intelligent machines. One of the earliest AI programs developed was the Logic Theorist by Allen Newell and Herbert A. Simon in 1956. The Logic Theorist could prove mathematical theorems using symbolic reasoning, demonstrating the potential for computers to perform tasks traditionally associated with human intelligence.

1.4.2. Evolution of AI Technologies:

Over the decades, AI technologies have advanced rapidly, driven by innovations in algorithms, computing power, and data availability. Key developments include:

- Expert Systems: In the 1970s and 1980s, expert systems emerged as a prominent AI technology, enabling computers to mimic the decision-making processes of human experts in specific domains. Examples include MYCIN, a diagnostic system for infectious diseases, and DENDRAL, a program for chemical analysis.
- Neural Networks: Neural networks, inspired by the structure and function of the human brain, gained popularity in the 1980s. Backpropagation, a method for training neural networks, was introduced by Rumelhart, Hinton, and Williams in 1986, paving the way for breakthroughs in pattern recognition, speech recognition, and other tasks.
- Machine Learning: Machine learning, a subfield of AI focused on algorithms that can learn from data, became increasingly prominent in the

late 20th century. Support vector machines (SVMs), decision trees, and random forests are examples of machine learning algorithms used in classification, regression, and clustering tasks.

1.4.3. Applications of AI Across Industries:

AI technologies have found applications across a wide range of industries, transforming business processes, enhancing productivity, and driving innovation. Examples include:

- **Healthcare:** AI is revolutionizing healthcare with applications such as medical imaging analysis, drug discovery, personalized medicine, and virtual health assistants. For instance, deep learning algorithms have shown promise in detecting diseases from medical images with accuracy comparable to human experts (Esteva et al., 2017).
- Finance: In the finance industry, AI is used for fraud detection, algorithmic trading, credit scoring, and customer service. Natural language processing (NLP) algorithms are employed to analyze news sentiment and social media data for predicting market trends and making investment decisions (Poria et al., 2017)
- **Transportation:** AI is driving innovation in transportation with the development of autonomous vehicles, route optimization systems, and predictive maintenance solutions. Companies like Tesla, Waymo, and Uber are investing heavily in AI technologies to develop self-driving cars that can navigate roads safely and efficiently. Also introducing applications of traffic simulations with different Systems. (Krajzewicz et al., 2012).

- **Manufacturing:** In manufacturing, AI is employed for process optimization, predictive maintenance, quality control, and supply chain management. For example, predictive maintenance algorithms use sensor data to forecast equipment failures and schedule maintenance tasks, reducing downtime and operational costs (Wang et al., 2016).
- **Retail:** AI is reshaping the retail industry with personalized recommendations, demand forecasting, inventory management, and customer service automation. E-commerce platforms like Amazon leverage AI algorithms to analyse customer behaviour and preferences, driving sales through targeted product recommendations. AI can be used to solve issues related to datamining and real-world problems (Liu et al., 2017).

1.4.4. AI in the Construction Industry:

In the construction industry, AI is being used to improve efficiency, safety, and cost-effectiveness across various stages of the project lifecycle. Examples include:

- Building Information Modeling (BIM): BIM software incorporates AI technologies for 3D modeling, clash detection, and project coordination, enabling stakeholders to visualize and simulate construction projects before they are built.
- Construction Robotics: AI-powered robots are deployed for tasks such as bricklaying, concrete pouring, and site inspection, enhancing productivity and safety on construction sites.
- Predictive Maintenance: AI algorithms analyze sensor data from equipment and machinery to predict maintenance needs and prevent costly breakdowns, optimizing asset performance and uptime.

1.5. Problem Statement:

- 70% of Construction projects experience delays and disturbances in Supply Chain (Ineight, 2023)
- Construction Projects suffer failure due to various factors including Cost Estimation Errors (Sunday Odediran & Abimbola Olukemi Windapo, 2014)
- Human Cost Estimators are prone to biased and inaccurate calculations. (Galorath, 2015)

"Human Cost Estimation being prone to error may lead to various delays in a construction project. These delays can lead to huge financial losses and may decrease the efficiency of the construction site".

1.6. Aims and Objectives:

The primary objectives for the project can be described as:

- Automating Cost Estimation through Machine Learning
- Developing a Proof of concept for the industry that Machine Learning is an efficient way for cost estimation.

1.7. Targeted SDG's:

A direct alignment of the research is with the UN Sustainable Development Goal (SDG) 9: Industry, Innovation, and Infrastructure, which is achieved through the application of Artificial Intelligence (AI) and Machine Learning to automate cost estimates in the construction industry. The project intends to transform conventional building methods by utilizing technologies like YOLOv8 and Python coding, minimizing issues like delays and cost overruns in material estimating and procurement. To promote economic growth and human well-being, SDG 9 emphasizes the significance of inclusive and sustainable industrialization, encouraging innovation, and modernizing infrastructure—all of which this research addresses. Incorporating artificial intelligence (AI) into cost assessment improves efficiency and accuracy while optimizing resource allocation, which is essential for the development of sustainable infrastructure.

Additionally, by promoting technological developments that help developing nations like Pakistan, where the construction sector is essential to economic progress, this research directly advances SDG 9. Setting an example for contemporary methods to enhance infrastructure development, project management, and environmental impact reduction is the successful application of AI and ML technologies in the construction industry. The thesis addresses important industrial issues by automating cost estimation procedures and optimizing supply chain management, eventually assisting in the development of robust networks.

In conclusion, this research emphasizes how technology can revolutionize the construction industry by advancing infrastructure development and sustainable practices that are in line with SDG 9 targets. This study emphasizes the role that innovation plays in propelling sustainable industrialization and infrastructure development for the good of society through the use of AI and ML technologies.

CHAPTER 2

LITERATURE REVIEW

2. Existing Literature:

Artificial Intelligence (AI) stands as a transformative field within computer science, focusing on the development of intelligent systems capable of mimicking human cognitive functions (LeCun et al., 2015). At its core, AI involves the creation of algorithms and models that enable machines to perceive their environment, reason, learn from data, and make decisions autonomously. This multidisciplinary domain integrates concepts from computer science, mathematics, neuroscience, and cognitive psychology to create systems that can process information and adapt to new situations. The crux of AI lies in its capacity to process vast amounts of data, recognize patterns, and make predictions or recommendations, surpassing traditional computing capabilities (Russell & Norvig, 2021). AI can be used vastly in the human world. The applications of AI span diverse sectors, revolutionizing industries and augmenting human capabilities. In healthcare, AI is instrumental in medical image analysis, drug discovery, and personalized medicine (Esteva et al., 2019). Machine Learning and Artificial Intelligence is also used in intelligent traffic control signals to reduce waiting time at intersections (Zhang et al., 2021) . AI can do whatever we want it to do, nothing more & nothing less. AI has also been used thoroughly in the construction industry. Artificial Intelligence (AI) has emerged as a transformative force in the construction industry, offering innovative solutions to longstanding challenges. In the realm of construction supply chain management, AI technologies such as machine learning algorithms and computer vision have shown great potential. These technologies enable automatic analysis of construction blueprints, identifying various components with precision. For instance, YOLO AI by Ultralytics has gained prominence for its speed and accuracy in object detection (Joseph Redmon & Ali Farhadi, 2018). By training the AI model on construction blueprints in our case, it becomes possible to detect elements such as walls, floors, columns, windows, and doors giving width and height for each identified class bounding box. Such applications of AI not only improve efficiency but also contribute to cost reduction and project timeline optimization. AI-powered systems facilitate predictive analytics in construction, aiding in decision-making processes. Predictive maintenance, for example, uses AI algorithms (BDA) to forecast equipment failures based on data patterns, allowing for proactive maintenance, and minimizing downtime (Bilal et al., 2016). This proactive approach to maintenance enhances operational efficiency and reduces unexpected costs. Artificial Intelligence may also be used for risk management by recognizing historical data and patterns. These capabilities of AI hold immense promise for enhancing the construction industry's resilience and competitiveness in Pakistan. The construction industry encompasses a variety of sub fields including the supply chain. The construction supply chain constitutes a crucial aspect of the construction industry, encompassing the flow of materials, information, and finances from suppliers to contractors and ultimately to the project site. CSC has been defined as the steps required to procuring the tools, building material and any other necessity essential to the construction site (Kristen Frisa & Ben Ashburn, 2024). Efficient management of this supply chain is essential for timely project completion and cost-effective operations. Within this framework, cost estimation and quantity surveying play integral roles, tightly interlinked with the construction supply chain. Cost estimation in construction projects is a multifaceted process that involves forecasting the expenses associated with labor, materials, equipment, and overheads. Many construction projects suffer failure due to factors including cost estimation errors (Sunday Odediran & Abimbola Olukemi Windapo, 2014) . The effectiveness of cost estimation is directly tied to the efficiency of the supply chain in procuring materials and managing logistics. Inaccurate estimates can lead to budget overruns and project delays, highlighting the importance of a well-coordinated supply chain. Quantity surveying, on the other hand, is closely linked to both cost estimation and the construction supply chain. Quantity surveyors are responsible for quantifying and managing the materials required for construction projects. Quantity surveying also includes advising on the client implications of client's requirements, cost management and procurement in the Supply Chain (Olanrewaju & Anahve, 2015). They play a crucial role in verifying the quantities of materials delivered to the site, ensuring that they align with project requirements and specifications. A good supply chain ensures good quantity surveying as the costs are tallied faster from the market. Cost Estimation and Quantity Surveying comes with its own set of challenges. One significant challenge is the inherent subjectivity and variability in human judgments. (Galorath, 2015) highlights how even experienced estimators have a delusion of success, optimism bias and various other psychological factors, leading to inconsistencies in quantifying materials and estimating costs. This subjectivity introduces the potential for errors in the early stages of a project, which can have cascading effects on subsequent phases. Furthermore, it has been noted that estimators and quantity surveyors tend to use their intuition, which has been observed to be biased and prone to error according to cognitive studies (Budi

Hartono et al., 2014). Manual cost estimation and quantity surveying are labor-intensive processes prone to human errors. The current cost estimation method, due to its manual nature, is time-consuming and leaves room for mathematical mistakes (Fonbeyin Henry Abanda et al., 2015). Simple data entry errors, such as incorrect unit conversions or misplaced decimal points, can result in substantial discrepancies in final cost estimates. The discrepancy between historical data and real-time market conditions can also lead to inaccurate estimations, affecting project budgets and timelines. Hiring skilled cost estimators in the construction industry presents a significant financial investment due to the specialized expertise required. According to the RICS Head of Policy Jeremy Blackburn, it is easier to hire a ballet dancer than a quantity surveyor. Companies often find themselves competing for a limited pool of talent, leading to higher salary expectations and recruitment expenses. As a result, the process of hiring cost estimators can be a resource-intensive endeavor, involving extensive recruitment efforts and significant financial commitments. ML and ANN have been used for cost estimation and have shown promising results but factors such as trust in AI and limited availability of datasets (Kudirat Ayinla et al., 2023). This research solves this gap by creating a custom AI model that can be used for detection of Blueprint Components with astonishing accuracy as will be described in the findings.

CHAPTER 3

METHODOLOGY

3. Methodology:

3.1. Research:

For our Project and our Research, we opted for an Iterative Experimental Approach. This approach constitutes a development method that underscores repeated cycles of prototyping, testing, and refinement; unlike traditional linear methods, it fosters continuous improvement and adaptation—each iteration reinforcing the last by weaving in fresh insights and feedback for progressive system enhancement. The dynamic process excels in managing complex projects: for instance, our project benefits from its flexibility—adjustments are continual as the path to completion is rarely clear-cut.

Furthermore, Iterative processes in development significantly benefit risk management: by segmenting the project into smaller, manageable iterations—we were able to quickly pinpoint and address potential issues early on; this proactive approach not only conserved time and resources but also kept the project trajectory firmly on course. Iterative development actively cultivates a culture of continuous learning and improvement: each cycle yields insightful lessons—informing subsequent iterations—which is essential in the fluid realm of cutting-edge technologies, such as deep learning.

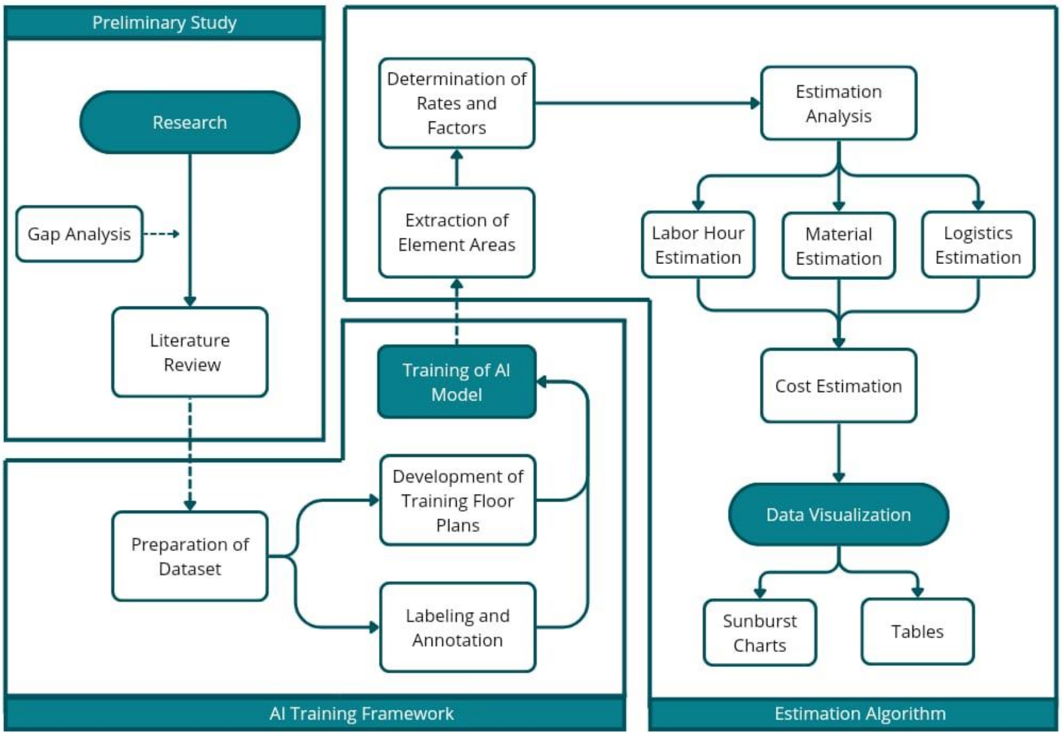

Figure 1: Flowchart of Methodology

Main points of investigation involved understanding the basic usage of convolutional neural networks (CNNs) and algorithms for detecting objects (Galvez et al., 2018). Care was taken to recognize the difficulties unique to the detection and analysis of architectural sketches, like differences in scale, viewpoints, overlapping parts and ways of depicting them.

Additionally, we also narrowed down an appropriate candidate for image object detection named YOLO (Redmon et al., 2016). YOLO (You Only Look Once) object detection system, specifically YOLOv8, was selected for identification of elements in a building.

3.2. Environment Setup for YOLOv8 Program:

Development of the program required the setup of the environment which entails the usage of specific software and hardware for the YOLOv8 model.

Installation of software packages such as CUDA drivers and Python 3.10 was done for smooth working of object detection (Raschka et al., 2020). The hardware specifications, especially the type of GPU for model training and inference were carefully thought about.

We used consumer NVIDIA GPUs consisting of a large amount of CUDA cores (such as RTX 4070 and RTX 3070; see Table 1). Since they were optimal for working with deep learning systems (Ilievski et al., 2018) and effective in computing. Moreover, we made the process of preparing the environment simpler by using code editing tools like PyCharm (Sanika Kendhe et al., 2023). This was to maintain uniformity and make sure that the results could be reproduced on various computer systems.

Table 1: Utilized GPU against number of CUDA cores.

GPU	CUDA cores
RTX 4070	4608
RTX 3070	5120

3.3. Creation of Custom Dataset:

Making a special custom dataset was an important part in creating a versatile model (Jain et al., 2020) for detection which would be able detect architectural elements in drawings. To do this, different types of architectural drawings were used for training, from drawings of residential to commercial buildings.

Drawings used for training were carefully selected to ensure that we included a variety (SUG, 2018) of architecture styles, types of buildings, sizes and how complex they were.

We chose to focus on 5 prominent features of a building drawing. Adding additional elements is just a matter of using more resources; the methodology for it remains the same. The elements that we considered for our project are the following:

Table 2: Component against Hatch Color

Component	**Solid Hatch Color**	**Color**
Column	ACI-253	Grey
Walls	ACI-1	Red
Doors	ACI-42	Brown
Windows	ACI-5	Blue
Floor	ACI-84	Green

Shades of above-mentioned colors were used to represent different ceiling heights where needed.

We edited the dataset to create a standardized form, which allows the YOLOv8 model to make accurate inferences (Redmon et al., 2016). This standardized dataset consisted of color coding (shown in Table 2) for every element type and hatch patterns for them in AutoCAD. Shades of these colors were used to represent varying ceiling heights.

Furthermore, data augmentation techniques were employed to enhance dataset diversity and improve model quality (Ding et al., 2019), (Terrance DeVries & Graham W. Taylor, 2017). Augmentation operations, such as rotation, translation, scaling, flipping, and color jittering, were applied to generate additional training samples while preserving spatial characteristics of architectural elements (Mishra et al., 2021)

3.4. Labelling and Annotation of Custom Dataset:

Every image in our custom dataset was carefully tagged and annotated for the training of the YOLO model (Russo et al., 2021). This task included manual drawing of bounding boxes around separate building elements, like walls, floors, window openings, doors, and columns.

Annotation was done by utilizing Roboflow (see figure 2), which allowed us to perform supervised learning for the model create accurate labels with consistency (James F. Mullen Jr. et al., 2019). Accuracy of the annotation process was further enhanced by labelling reviewing, where the same labels / annotations were reviewed by multiple team members. Complex spatial geometries such as overlapping elements, merging elements, and hidden elements were manually catered for during the annotation process.

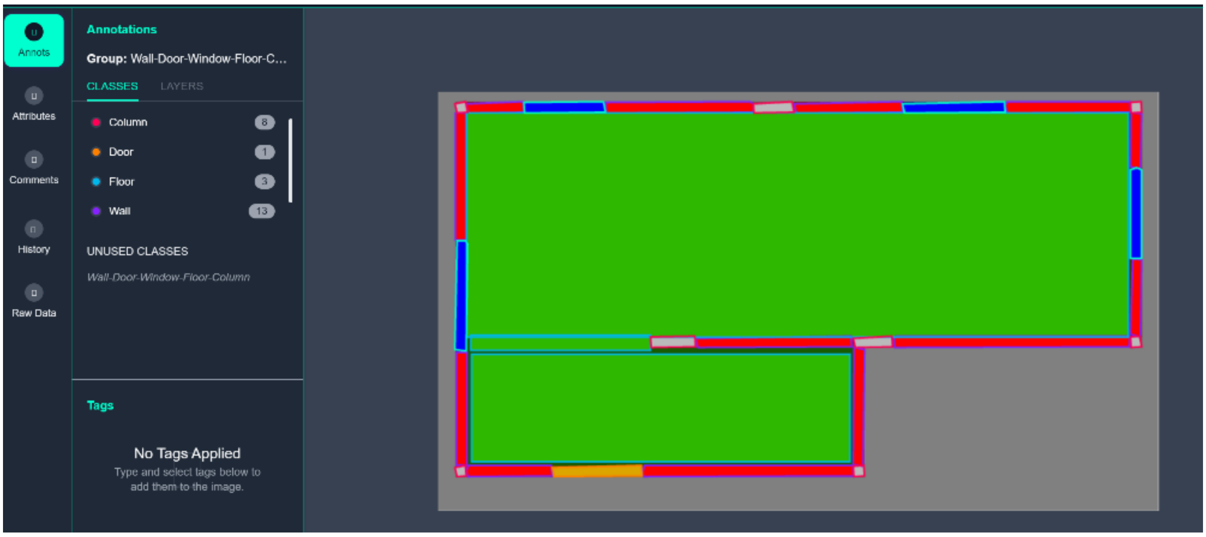

Figure 2: Annotation Process

3.5. Training of Custom Model:

The training of a custom YOLOv8 model for detecting architectural elements involved an iterative process of model initialization, optimization, and fine-tuning using the annotated dataset.

Model training was performed using deep learning, which was done by using high-performance computing hardware equipped with high-end NVIDIA GPUs

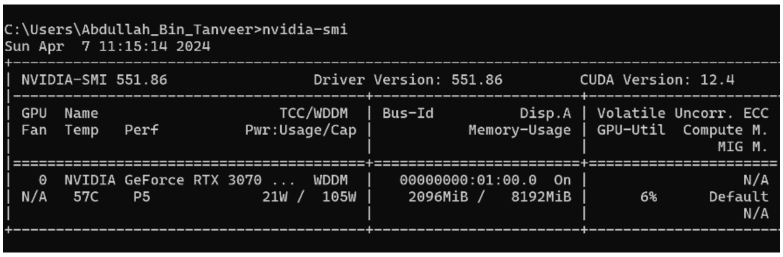

Figure 3: NVIDIA RTX 3070 used for training.

for accelerated training (Bhagirath et al., 2019). 35 epochs (iterations) were run to give optimal training vs. time (see Figure 3).

3.6. Validation of Model:

After the training was finished, we aimed to test the model to observe its performance on new data and the quality of its inference (Erickson & Kitamura, 2021).

The dataset was split into three parts for training, validation, and testing, which are essential for the creation of the model. Evaluation metrics, mainly precision and accuracy, were used to assess the model's robustness. Feedback from validation results (example: figure 4) were used to refine and optimize the model.

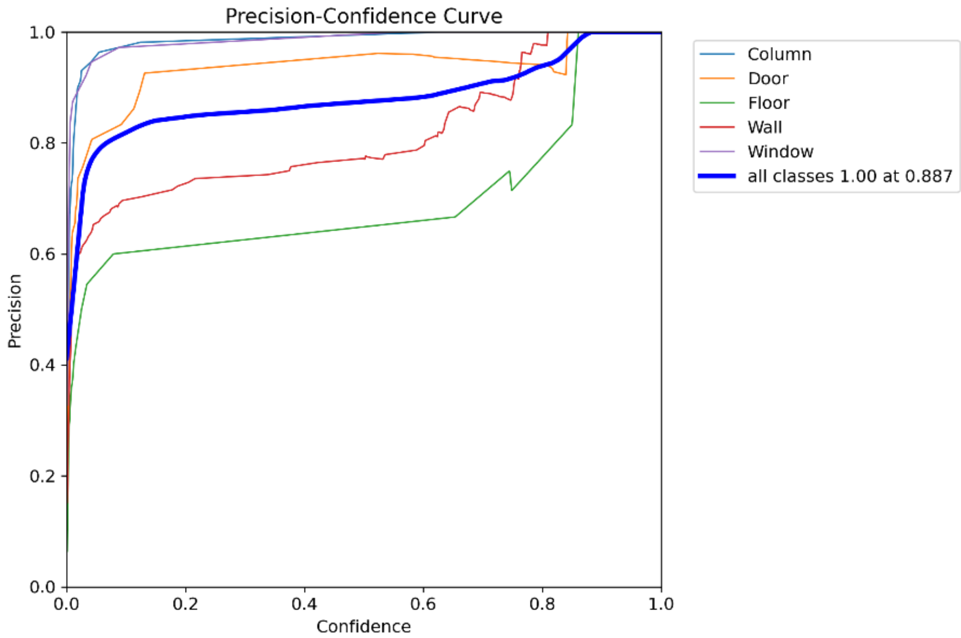

Figure 4: Precision Confidence curve of AI model trained.

3.7. Storing / Saving Model Inference Results:

Following successful validation and preparation of the model, it was necessary to store / save the results of any inferences that were done by the model (see figure 5).

Simple detection did not provide us with the necessary logistical information; hence we used special commands in YOLOv8 during inference. This allowed us to obtain a txt format file that contained a list of all detected elements along with important information of these elements.

Additional information obtained was the class id of each element, center x coordinates, center y coordinates, width of the element, and height of the element (see Figure 6).

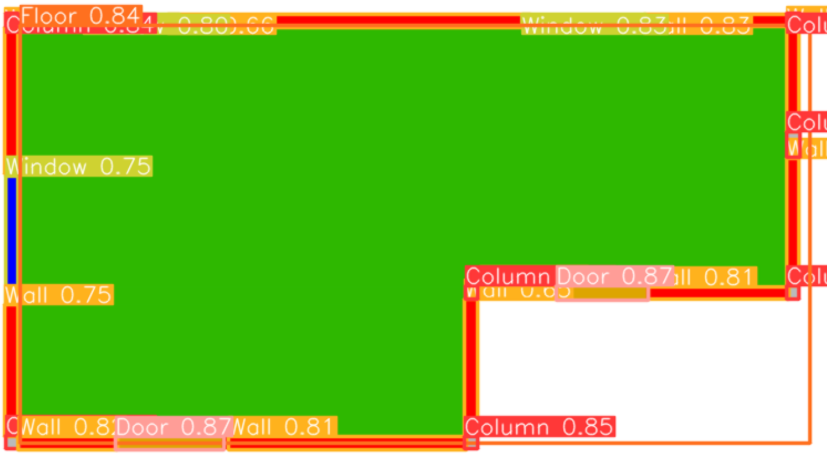

Figure 5: Example of bounding boxes.

```
File    Edit    View

1 0.218051 0.890059 0.128676 0.0261466
1 0.731782 0.59243 0.109163 0.0266351
0 0.958215 0.0504054 0.0160809 0.0263187
0 0.575917 0.889473 0.015842 0.0250639
0 0.95807 0.297232 0.0160767 0.0507748
2 0.509735 0.476678 0.938235 0.830657
0 0.0299843 0.0503877 0.0150746 0.0261012
0 0.576186 0.592101 0.0155059 0.0262688
4 0.709692 0.0506128 0.14759 0.0246546
3 0.867148 0.0503846 0.171723 0.0257123
3 0.957935 0.167958 0.0154156 0.209508
3 0.631623 0.591981 0.0921579 0.0258495
3 0.0968801 0.890192 0.118159 0.025977
0 0.957878 0.591404 0.0154307 0.0249889
3 0.869895 0.5922 0.167925 0.0256867
3 0.42888 0.890493 0.281989 0.0262282
4 0.171799 0.0504397 0.11461 0.0248345
0 0.0304761 0.889189 0.0162121 0.0256396
3 0.958045 0.451894 0.0151311 0.254237
3 0.0301679 0.746938 0.0157204 0.263561
4 0.0300669 0.487232 0.0146291 0.253227
3 0.0768539 0.0501307 0.0771293 0.0263529
3 0.0294463 0.210137 0.0148927 0.292368
3 0.430694 0.051228 0.431965 0.0275035
3 0.576053 0.740878 0.0160598 0.270715
```

Figure 6: Coordinates of Bounding Boxes

3.8. Extraction of Dimensional Data using Python:

Post-detection processing involved the usage of custom Python scripts that were developed to parse bounding box coordinates and dimensions from the .txt file that was previously generated by the YOLOv8 model (see Figure 5).

3.9. Derivation of Usable Logistical Information using Python:

The extracted dimensional data created the foundation for deriving useful logistical information that is critical for construction planning, management, and resource allocation.

Additional custom Python scripts were coded to process the dimensional data and compute various logistical parameters, such as material quantities required, estimated construction costs, labor requirements, and project timelines.

The logistical information derived provides valuable insights for stakeholders, which enables them to make informed decisions and allocate resources accordingly throughout the project. Moreover, interactive visualization tools and dashboards were developed to express data effectively, which leads to better analysis capabilities and interpretation by stakeholders.

Data-driven insights generated from analyzed logistical information allows for the creation of guided strategic plans and informed risk management strategies, which in the end improve project results and minimize cost overruns.

CHAPTER 4

RESULTS AND ANALYSIS

4. Results and Discussions:

4.1. Findings:

A total of 2 trials were conducted to ensure maximum efficiency and coherence in the detection by YOLOv8. Both trials will be discussed with their findings.

4.1.1. Trial 1:

In the first trial, only 4 components were labelled by the group i.e. walls, windows, doors, and floors. A total of 1200 blueprints (all made from scratch by the group) were manually annotated and labelled leading to a sufficient dataset for a project that serves as a proof of concept. 35 iterations were run. The findings for trial 1 are as follows:

- The background clutter was more than the subject in images.
- Walls were over-represented whereas windows were under-represented as shown in figure 7.
- The Detection could have been better as the area under PR curve (figure 8) was comparatively lower than what was expected.
- The training set used was therefore discarded and deemed unfit for further processing.

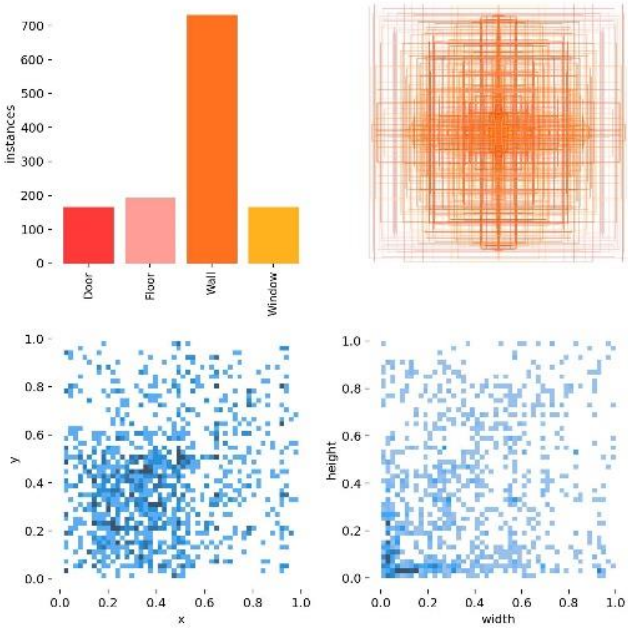

Figure 7: Labels vs. Instances for trial 1

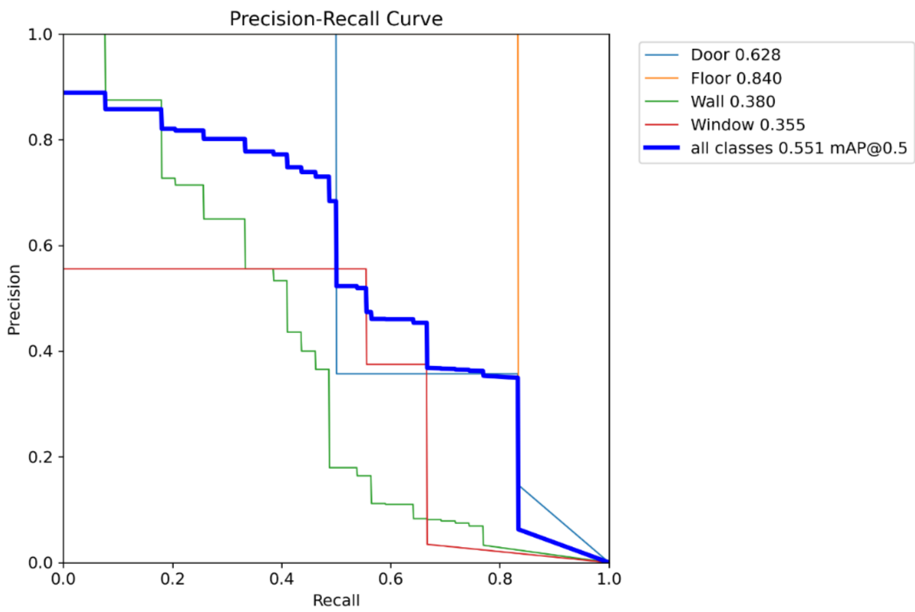

Figure 8: PR Curve for trial 1.

4.1.2. Trial 2:

Our second trial focused on 5 components of an engineering drawing. The dataset consisted of around 3000 images of blueprints of various sizes. 40 iterations were run. This trial gave us a better functioning AI ML model.

- Lesser background clutter was observed.
- Walls, even though still over-represented (figure 9), did not overshadow the rest of the components.
- A better area under PR curve (figure 10) was observed that satisfied our standards.
- This AI model was then used for cost estimation by using it in our python code.

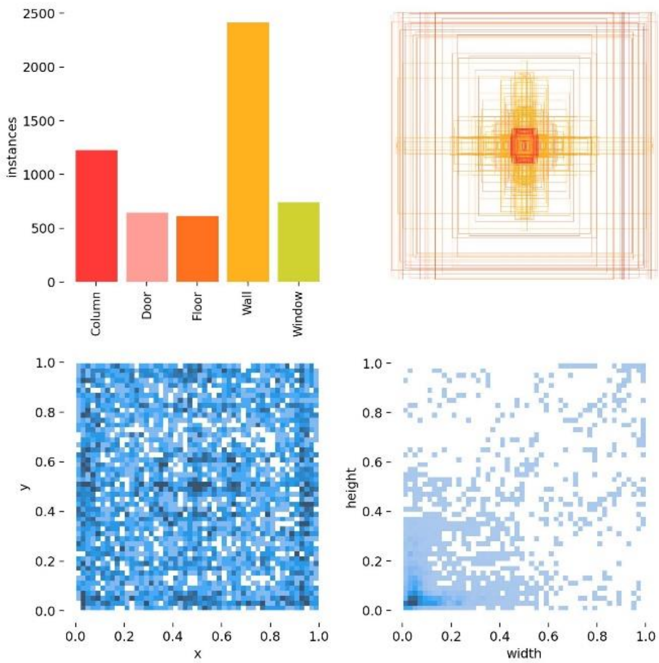

Figure 9: Labels vs Instances for trial 2

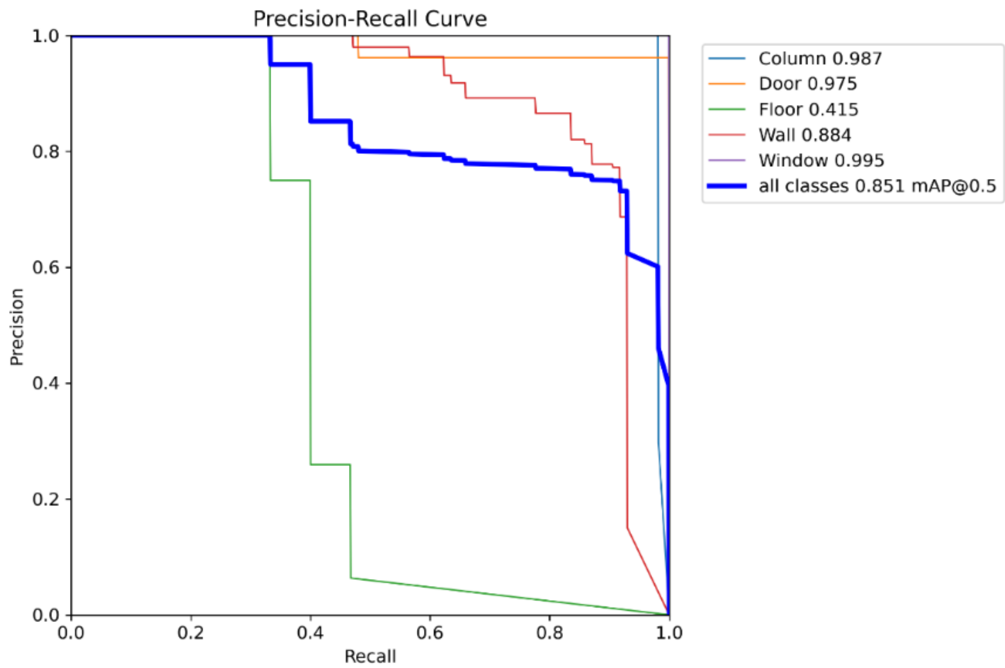

Figure 10: PR curve for trial 2

The following table shows accuracy of each component that was estimated using this model.

Table 3: Estimation of Sample 1

1 - FINAL ESTIMATION – BUILDING			
Quantities	**Actual Area (sqft)**	**Predicted Area (sqft)**	**Error**
Floor	2525	2421.84	4.25%
Walls	289	310.65	6.9%
Doors	29.77	30.92	4%
Windows	61	65.70	7%
Columns	51	57.29	11%

The average error for Building Sample 1 was calculated to be around 6%.

Figure 11: Detection in Sample 1

Table 4: Estimation of sample 2

2 - FINAL ESTIMATION – PARKING LOT			
Quantities	**AREA (sqft)**	**Predicted Area (sqft)**	**Error**
Floor	3360.125	3378.12	0.6%
Walls	126.38	135.68	6.9%
Columns	13.5	15.08	10.5%

The average error for Building Sample 2 was calculated to be around 6%.

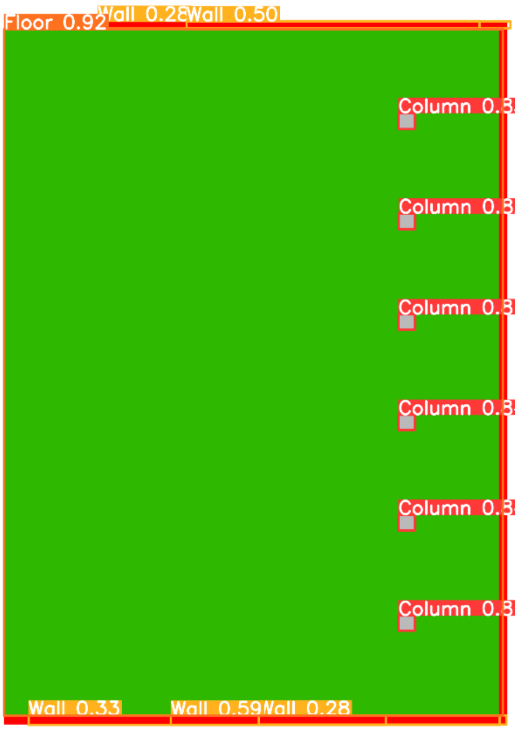

Figure 12: Detection for sample 2

4.2. Relevance to the Field:

In the Construction Industry, accurate cost estimation is essential for the successful completion of a project. The AI under discussion requires no knowledge of computing and a very basic understanding of how applications work, thus can be utilized by anyone, anytime, anywhere. As discussed in the literature review, human cost estimation is prone to a plethora of errors, causing problems such as insufficient estimates, poorly handled logistics and delays in construction projects. Use of Artificial Intelligence negates these issues in the construction industry. Many projects face delays due to inaccurate cost estimation and quantity surveying. While Human cost estimation may take several hours if not days, cost estimation through this AI will take seconds at most thus saving precious resources.

4.3. Discussions:

The AI model as well as the Integration of Python Code will now be able to Automate Cost Estimation of the Superstructure to a Degree. The 5 main quantities which are Floor, Walls, Doors, Windows, and Columns are now being detected by YoloV8 with more than an Average Accuracy of 80%, meaning that the quantities found from the usage of this AI has a Deviation of at most 20% from the cost estimation performed manually. Further Information is also derived from these quantities such as the amount of Labour Hours used to work on the project, the amount of steel used, the average number of trips which will be used. These Quantities are well divided between different components such as The Number of Labor hours used on Concreting, the number of Labor used on brick laying etc. The area estimation error in Building Sample 2 is considered to be a limitation of the project. Even though a dataset of 3000 images was used, it is still not enough for extremely accurate estimations. Roughly 10,000-20,000 images will give high end accuracy and will mitigate the error in instances of classes.

These Key Findings from our project will be extremely beneficial to the construction industry. As cost estimation is one of the vital steps involved in every project related to construction, with the help of the AI not only a person will be able to save on time but save on resources as well. These resources can be diverted on different components of the project which will result in Higher Quality as well as repel any chances of Cost Overruns.

4.4. Practical Implications:

Our research raises pertinent questions about the integration of advanced technologies in traditional industries. While our study highlights the potential benefits of AI in enhancing project efficiency, management, and cost accuracy, there are practical implications that warrant thorough consideration.

One aspect of concern is the readiness of the construction industry in Pakistan to adopt AI technologies. The transition from conventional methods to AI-driven cost estimation processes may face resistance due to factors such as workforce training, infrastructure requirements, and initial investment costs. Additionally, the cultural acceptance of AI in decision-making processes within the industry needs to be addressed to ensure successful implementation.

Moreover, the reliance on Machine Learning algorithms for cost estimation raises questions about the quality and reliability of the data inputs. The accuracy of cost predictions generated by AI systems heavily depends on the quality of historical data, which may be limited or inconsistent in certain contexts. Ensuring the integrity and relevance of data sources becomes crucial in mitigating the risks associated with inaccurate estimations.

Furthermore, while automation through AI has the potential to reduce project costs, there is a need to assess the broader socio-economic impact of displacing traditional roles with technology. The shift towards AI-driven cost estimation may lead to changes in job requirements, skill sets, and labor dynamics within the construction sector, necessitating proactive measures to address potential workforce disruptions.

In conclusion, while the integration of AI in cost estimation presents promising opportunities for enhancing project outcomes, the practical implications underscore the importance of addressing challenges related to industry readiness, data quality, cultural adaptation, and workforce implications. A comprehensive strategy that considers these factors is essential to facilitate a successful and sustainable transition towards AI-driven practices in the construction industry.

CHAPTER 5

CONCNLUSIONS

5. Conclusions:

5.1. Recommendations:

The AI Cost Estimator developed for the construction industry has been a significant undertaking for the final year project. This tool aims to revolutionize the estimation process by utilizing artificial intelligence to predict and analyze costs associated with various construction elements.

Upon thorough evaluation of the project, it is evident that the AI Cost Estimator has met the established standards. However, there are areas where enhancements could further elevate its performance. The analysis reveals that walls emerge as the primary quantity leading to deviations, overshadowing other elements during the iterative processes.

Based on the findings, it is imperative to implement strategies to address the challenges posed by the dominance of wall quantities. To mitigate this issue effectively, the following recommendations are proposed:

- Utilize Complex Drawings: Incorporating more intricate drawings into the training dataset can provide the AI with a comprehensive understanding of diverse wall structures. This exposure to varying complexities will enhance the Estimator's ability to accurately identify and quantify walls in different scenarios.
- Increase Data Samples: Augmenting the dataset with a larger volume of data samples will expose the AI Cost Estimator to a wider range of construction

scenarios. This expanded dataset will enable the AI model to learn from diverse examples, improving its capacity to recognize and estimate quantities such as floors, doors, windows, and columns with greater precision.

- Enhance Training Procedures: Enhancing the training procedures by incorporating advanced algorithms and methodologies can refine the Estimator's predictive capabilities. Implementing state-of-the-art techniques in data preprocessing, feature engineering, and model optimization can elevate the accuracy and efficiency of cost estimations.
- Continuous Monitoring and Calibration: Regular monitoring of the AI Cost Estimator's performance and conducting periodic recalibrations are essential to ensure its sustained accuracy. By evaluating its predictions against actual costs and making necessary adjustments, the Estimator can adapt to evolving construction dynamics and deliver reliable estimations consistently.

By implementing these recommendations, the AI Cost Estimator in the construction industry is poised to achieve several notable advantages:

- Improved Efficiency: The Estimator's enhanced ability to accurately identify and quantify construction elements will streamline the cost estimation process, saving time and resources.
- Enhanced Accuracy: With a refined training dataset and advanced algorithms, the Estimator will exhibit higher accuracy in predicting costs, minimizing deviations and errors in estimations.
- Optimized Performance: Continuous monitoring and calibration will ensure that the Estimator remains aligned with industry standards and adapts to changing construction trends, maintaining its relevance and reliability.

In conclusion, the proposed recommendations aim to elevate the AI Cost Estimator's capabilities, enabling it to deliver more accurate and efficient cost estimations in the construction industry. By addressing the challenges identified and implementing strategic enhancements, the Estimator is poised to become a valuable tool for cost estimation in construction projects.

5.2. Future Research:

As advancements in technology continue to revolutionize the industry, the integration of Artificial Intelligence (AI) in cost estimation processes has shown great promise. Building upon the foundation laid by our project, there are several avenues for future research that can further enhance the accuracy and effectiveness of cost estimation in the construction industry.

5.2.1. Expansion to Include MEP Plans

One key area for future research involves expanding the scope of the AI Cost Estimator to incorporate Mechanical, Electrical, and Plumbing (MEP) plans. While the initial project focused on the Plan Layouts of the structure, the inclusion of MEP drawings can provide a more comprehensive understanding of the cost components involved in a construction project. By considering elements such as wires, bulbs, plumbing fixtures, and other MEP-related items, the AI Cost Estimator can offer a more detailed and accurate estimation of project costs.

5.2.2. Utilization of Enhanced Datasets

Enhancing the quality and depth of datasets used in cost estimation is another crucial aspect of future research. By leveraging more extensive and refined datasets, the AI Cost Estimator can generate more precise cost values for different construction activities. Access to a broader range of data points, such as historical cost information, market trends, and material prices, can significantly improve the reliability and accuracy of cost estimates provided by the AI system.

5.2.3. Integration of Building Information Modeling (BIM)

The integration of Building Information Modeling (BIM) technology presents a promising avenue for enhancing the capabilities of the AI Cost Estimator. By incorporating BIM data into the cost estimation process, the system can leverage the detailed 3D models and information available in BIM to derive more accurate cost projections. BIM offers a holistic view of the project, enabling the AI Cost Estimator to account for spatial relationships, material quantities, and construction sequencing, leading to more precise cost estimations.

5.2.4. Optimization of Machine Learning Algorithms

Further research can focus on optimizing the machine learning algorithms used in the AI Cost Estimator to improve its predictive capabilities. By fine-tuning the algorithms based on feedback and continuous learning from cost estimation outcomes, the system can enhance its ability to analyze complex cost relationships and generate more reliable estimates. Continuous refinement of the AI models through algorithm optimization can lead to increased accuracy and efficiency in cost estimation processes.

5.2.5. Collaboration with Industry Experts

Collaboration with industry experts and professionals in the construction sector can provide valuable insights and feedback for refining the AI Cost Estimator. Engaging with practitioners who have firsthand experience in cost estimation and project management can help validate the accuracy of the system's estimates and identify areas for improvement. By incorporating real-world expertise and domain knowledge, the AI Cost Estimator can be tailored to meet the specific needs and challenges faced by construction industry professionals.

In conclusion, the future research directions outlined above present exciting opportunities to enhance the capabilities of the AI Cost Estimator in the construction industry. By expanding the scope of the system, improving dataset quality, integrating BIM technology, optimizing machine learning algorithms, and collaborating with industry experts, the AI Cost Estimator can evolve into a sophisticated tool for accurate and efficient cost estimation in construction projects. Embracing these avenues for further research will contribute to advancing the field of cost estimation and driving innovation in the construction industry.

5.3. Potential Topics:

- *Integrating Structural Drawings for Accurate Cost Estimation in Building Projects using AI*
- *Efficiency in Construction Cost Estimation: Leveraging Diverse Drawing Types for Precision.*
- *An Automated Cost Estimator Using Artificial Neural Networks.*
- *A Comparison between a Cost Estimator trained on Machine Learning and Artificial Neural Network.*

5.4. Conclusions:

Serving as a proof of concept, the project is an important milestone in the construction industry, more importantly, the construction supply chain. Using Artificial Intelligence to cost estimate and control an important decision-making step in construction projects is still something new and will require refinement along the way, but it holds great promises. It is necessary that civil engineers use the up and rising field of AI and ML to their great benefit, and in return, serve the community in an efficient, better way.

CHAPTER 6

6. REFERENCES

Badea, A., Prostean, G., Goncalves, G., & Allaoui, H. (2014). Assessing Risk Factors in Collaborative Supply Chain with the Analytic Hierarchy Process (AHP). *Procedia - Social and Behavioral Sciences*, *124*, 114–123. https://doi.org/10.1016/j.sbspro.2014.02.467

Bhagirath, Mittal, N., & Kumar, S. (2019). Machine Learning Computation on Multiple GPU's using CUDA and Message Passing Interface. *2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC)*, 18–22. https://doi.org/10.1109/PEEIC47157.2019.8976714

Bilal, M., Oyedele, L. O., Qadir, J., Munir, K., Ajayi, S. O., Akinade, O. O., Owolabi, H. A., Alaka, H. A., & Pasha, M. (2016). Big Data in the construction industry: A review of present status, opportunities, and future trends. *Advanced Engineering Informatics*, *30*(3), 500–521. https://doi.org/10.1016/j.aei.2016.07.001

Budi Hartono, Fandy I. Nugroho, & Bhery Aditya Saputra. (2014). *Biases in Project Estimation: Experimental Evidence*.

Ding, J., Li, X., Kang, X., & Gudivada, V. N. (2019). A Case Study of the Augmentation and Evaluation of Training Data for Deep Learning. *Journal of Data and Information Quality*, *11*(4), 1–22. https://doi.org/10.1145/3317573

Erickson, B. J., & Kitamura, F. (2021). Magician's Corner: 9. Performance Metrics for Machine Learning Models. *Radiology: Artificial Intelligence*, *3*(3), e200126. https://doi.org/10.1148/ryai.2021200126

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Erratum: Corrigendum: Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *546*(7660), 686–686. https://doi.org/10.1038/nature22985

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Fonbeyin Henry Abanda, Bernard Kamsu-Foguem, & Joseph H. M. Tah. (2015). Towards an Intelligent Ontology Construction Cost Estimation System: Using BIM and New Rules of Measurement Techniques. *World Academy of Science, Engineering and Technology International Journal of Industrial and Manufacturing Engineering*, *9*(1), 294–299.

Galorath, D. (2015, August 31). Why Are Estimates Always Wrong: Estimation Bias and Strategic Misestimation. *AIAA SPACE 2015 Conference and Exposition*. https://doi.org/10.2514/6.2015-4591

Galvez, R. L., Bandala, A. A., Dadios, E. P., Vicerra, R. R. P., & Maningo, J. M. Z. (2018). Object Detection Using Convolutional Neural Networks. *TENCON 2018 - 2018 IEEE Region 10 Conference*, 2023–2027. https://doi.org/10.1109/TENCON.2018.8650517

GlobalData. (2023). *Pakistan Construction Market Size, Trend Analysis by Sector, Competitive Landscape and Forecast to 2027*. https://www.globaldata.com/store/report/pakistan-construction-market-analysis/#:~:text=The%20Pakistan%20construction%20market%20size,5%25%20during%202024-2027.

Ilievski, A., Zdraveski, V., & Gusev, M. (2018). How CUDA Powers the Machine Learning Revolution. *2018 26th Telecommunications Forum (TELFOR)*, 420–425. https://doi.org/10.1109/TELFOR.2018.8611982

Ineight. (2023). *Global Capital Projects Outlook* (1st ed., Vol. 1). https://ineight.com/annual-global-capital-projects-report/?content=true&utm_medium=Earned_Media&utm_campaign=230614-PR-GCPO&utm_content=global-outlook-report

Jain, A., Patel, H., Nagalapatti, L., Gupta, N., Mehta, S., Guttula, S., Mujumdar, S., Afzal, S., Sharma Mittal, R., & Munigala, V. (2020). Overview and Importance of Data Quality for Machine Learning Tasks. *Proceedings of the 26th ACM*

SIGKDD International Conference on Knowledge Discovery & Data Mining, 3561–3562. https://doi.org/10.1145/3394486.3406477

Jain, A., Somwanshi, D., Talapula, D., & Joshi, K. (2023). *Artificial intelligence based software cost estimation*. 020006. https://doi.org/10.1063/5.0152957

James F. Mullen Jr., Franklin R. Tanner, & Phil A. Sallee. (2019). Comparing the Effects of Annotation Type on Machine Learning Detection Performance. *CVPR 2019*.

Joseph Redmon, & Ali Farhadi. (2018). *YOLOv3: An Incremental Improvement*.

Kim, S.-Y., & Nguyen, V. T. (2022). Supply chain management in construction: critical study of barriers to implementation. *International Journal of Construction Management*, *22*(16), 3148–3157. https://doi.org/10.1080/15623599.2020.1843768

Krajzewicz, D., Bieker, L., Härri, J., & Blokpoel, R. (2012). Simulation of V2X Applications with the iTETRIS System. *Procedia - Social and Behavioral Sciences*, *48*, 1482–1492. https://doi.org/10.1016/j.sbspro.2012.06.1124

Kristen Frisa, & Ben Ashburn. (2024, April 3). *A Contractor's Guide to Supply Chain Management*. Procore.

Kudirat Ayinla, Abdullahi Saka, Rafiu Dimeji Seidu, & Upeksha Madanayake. (2023, September). *The Impact of Artificial Intelligence on Construction Costing Practice*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Liu, S., Wang, X., Liu, M., & Zhu, J. (2017). Towards better analysis of machine learning models: A visual analytics perspective. *Visual Informatics*, *1*(1), 48–56. https://doi.org/10.1016/j.visinf.2017.01.006

Mishra, S., Hashmi, K. A., Pagani, A., Liwicki, M., Stricker, D., & Afzal, M. Z. (2021). Towards Robust Object Detection in Floor Plan Images: A Data Augmentation Approach. *Applied Sciences*, *11*(23), 11174. https://doi.org/10.3390/app112311174

Olanrewaju, A., & Anahve, P. J. (2015). Duties and Responsibilities of Quantity Surveyors in the Procurement of Building Services Engineering. *Procedia Engineering*, *123*, 352–360. https://doi.org/10.1016/j.proeng.2015.10.046

Poria, S., Peng, H., Hussain, A., Howard, N., & Cambria, E. (2017). Ensemble application of convolutional neural networks and multiple kernel learning for multimodal sentiment analysis. *Neurocomputing*, *261*, 217–230. https://doi.org/10.1016/j.neucom.2016.09.117

Raschka, S., Patterson, J., & Nolet, C. (2020). Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence. *Information*, *11*(4), 193. https://doi.org/10.3390/info11040193

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed., Vol. 1).

Russo, S., Besmer, M. D., Blumensaat, F., Bouffard, D., Disch, A., Hammes, F., Hess, A., Lürig, M., Matthews, B., Minaudo, C., Morgenroth, E., Tran-Khac, V., & Villez, K. (2021). The value of human data annotation for machine learning based anomaly detection in environmental systems. *Water Research*, *206*, 117695. https://doi.org/10.1016/j.watres.2021.117695

Sanika Kendhe, Abhishek Nishad, Diksha Labhade, & Vikas Magar. (2023). Comparative Analysis of Different Python Editors. *Vidhyayana ISSN 2454-8596*, *8*(7).

SRM Riazi, MF Zainuddin, MNM Nawi, S Musa, & A Lee. (2020). A critical review of fragmentation issues in the construction industry. *Journal of Talent Development and Excellence*, *12*(2), 1510–1521.

SUG, H. (2018). Performance of Machine Learning Algorithms and Diversity in Data. *MATEC Web of Conferences*, *210*, 04019. https://doi.org/10.1051/matecconf/201821004019

Sunday Odediran, & Abimbola Olukemi Windapo. (2014, February). Systematic review of factors influencing the cost performance of building projects. *Laryea, S. and Ibem, E. (Eds) Proceedings 8th Construction Industry Development Board (Cidb) Postgraduate ConferenceAt: University of the Witwatersrand, Johannesburg, South AfricaVolume: VIII*.

Terrance DeVries, & Graham W. Taylor. (2017). *Dataset Augmentation in Feature Space*.

Wang, S., Wan, J., Li, D., & Zhang, C. (2016). Implementing Smart Factory of Industrie 4.0: An Outlook. *International Journal of Distributed Sensor Networks*, *12*(1), 3159805. https://doi.org/10.1155/2016/3159805

Zhang, R., Ishikawa, A., Wang, W., Striner, B., & Tonguz, O. K. (2021). Using Reinforcement Learning With Partial Vehicle Detection for Intelligent Traffic Signal Control. *IEEE Transactions on Intelligent Transportation Systems*, *22*(1), 404–415. https://doi.org/10.1109/TITS.2019.2958859