AI-Driven Detection and Mitigation of Urban Heat Island Effects Using Image Analysis and IoT

R25-002

Project Proposal Report

Kumara B.D.A.N – IT21256264

B.Sc. (Hons) Degree in Information Technology Specialized in Software Engineering

Department of Computer Science and Software Engineering Faculty of Computing

Sri Lanka Institute of Information Technology Sri Lanka

January 2025

AI-Driven Detection and Mitigation of Urban Heat Island Effects Using Image Analysis and IoT

R25-002

Project Proposal Report

Kumara B.D.A.N – IT21256264

Supervisor: Mr. Vishan Jayasinghearchchi
Co-Supervisor: Ms. Kaushalya Rajapakse
External Supervisor: Dr. Rajitha De Silva

B.Sc. (Hons) Degree in Information Technology Specialized in Software Engineering

Department of Computer Science and Software Engineering Faculty of Computing

Sri Lanka Institute of Information Technology Sri Lanka

January 2025

Declaration

I declare that this is my own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature	Date
Kumara B.D.A.N	IT21256264		1/23/2025

Supervisor

Mr. Vishan Jayasinghearchchi

Date: 23/01/2025

Co-Supervisor

KR

Ms. Kaushalya Rajapakse

Date: 23/01/2025

Abstract

Urban thermal analysis is vital for sustainable city planning and environmental management. However, accurate segmentation and material identification of urban structures remain significant challenges. This research introduces an innovative framework that combines YOLOv8 for object detection with the Segment Anything Model (SAM) for mask generation, followed by advanced material classification algorithms and surface area calculations. The proposed approach begins with YOLOv8 identifying key urban components such as buildings, pavements, and rooftops in images. SAM then generates precise masks for these components, which are subsequently analyzed to determine material types, including concrete, glass, and metal. Finally, image processing techniques calculate the surface areas of these segments based on their material type, providing crucial metrics for thermal analysis. By leveraging the speed and accuracy of YOLOv8 for object detection and the generalization capabilities of SAM for segmentation, this framework addresses existing gaps in urban thermal analysis. The results have the potential to significantly enhance urban planning, resource management, and environmental assessments.

List of Abbreviations

AI	Artificial Intelligence
CNN	Convolutional Neural Network
CRF	Conditional Random Field
GIS	Geographic Information System
IoT	Internet of Things
LLM	Large Language Model
SAM	Segment Anything Model
UHI	Urban Heat Island
U-Net	Convolutional Neural Network Model for Semantic Segmentation
VLM	Vision Language Model
YOLO	You Only Look Once

Table of Contents

Table of Figures

List of tables

Introduction

Urbanization has transformed cities into complex landscapes, introducing challenges such as increased energy demands, urban heat islands, and resource inefficiencies. The Urban Heat Island (UHI) effect, characterized by elevated temperatures in urban areas compared to surrounding rural regions, exacerbates energy consumption, intensifies greenhouse gas emissions, and poses public health risks. Addressing these challenges requires a precise understanding of urban structures, their material properties, and their contributions to heat retention. Urban thermal analysis is an essential tool for achieving this understanding, enabling insights into heat distribution, energy optimization, and sustainable planning. However, traditional methods such as manual surveys and static imaging are time-consuming, lack scalability, and fail to provide the detailed segmentation and material-specific data required for effective UHI analysis [1].

The emergence of advanced computer vision and machine learning techniques has opened new possibilities for automating urban image analysis. Semantic segmentation models, such as U-Net and DeepLabv3+, have been widely used to segment urban images into meaningful regions like buildings, pavements, and rooftops [2][3]. Similarly, object detection models, including the YOLO family, are known for their real-time performance and high accuracy in identifying urban components [4]. Datasets like Cityscapes and ADE20K provide annotated urban images that facilitate the training and evaluation of these models [5]. While these innovations have significantly enhanced urban image analysis, current methodologies often focus on segmentation or detection in isolation, without integrating material classification and surface area calculations [6].

Recent advancements in tools like YOLOv8 and the Segment Anything Model (SAM) represent significant progress in object detection and general-purpose segmentation. These models enable efficient and precise analysis of urban scenes, even with minimal labeled data. Despite their capabilities, the integration of such tools into a unified framework that combines segmentation, material classification, and quantitative analysis for UHI studies remains an underexplored area [7].

Previous research has largely addressed specific components of UHI analysis, such as pixel-wise segmentation, material classification, or surface area calculations. For example, semantic segmentation models provide detailed delineations of urban features, but they often require extensive labeled datasets and lack integration with downstream material-specific analyses [2][3]. Material classification methods, while effective, typically depend on spectral imaging or manual annotations, making them less adaptable to diverse urban environments [6]. Similarly, surface area calculations are frequently performed as standalone tasks, disconnected from the segmentation pipeline [8].

To address these limitations, this research proposes an integrated framework that combines YOLOv8 for object detection with SAM for segmentation. The segmented components will then be analyzed to classify their material types using adaptive algorithms. Advanced image processing techniques will be employed to calculate the surface areas of these components. By addressing existing limitations, this approach aims to contribute to precise and efficient urban thermal analysis, supporting sustainable city planning and resource management

Background and Literature Survey

The Urban Heat Island (UHI) effect is a significant challenge faced by urbanized areas, characterized by elevated temperatures in cities compared to surrounding rural regions. This phenomenon arises from the extensive use of heat-retaining materials like concrete, asphalt, and glass, along with the reduction of vegetation and green spaces. The resulting elevated temperatures lead to increased energy consumption, heightened reliance on cooling technologies, and intensified greenhouse gas emissions. Addressing these challenges requires a comprehensive understanding of urban structures, their material properties, and their contributions to heat retention. Traditional approaches to studying UHI effects, such as manual surveys and static thermal imaging, have provided valuable insights but are often time-intensive, lack scalability, and fail to integrate material-specific data with spatial analysis [1]. Recent advancements in artificial intelligence (AI) and image processing techniques have paved the way for automating urban analysis, offering precise, efficient, and scalable solutions [7].

Semantic segmentation has become a foundational technique in urban image analysis, enabling the division of complex urban scenes into meaningful regions such as buildings, pavements, and rooftops. Models like U-Net and DeepLabv3+ have set benchmarks in segmentation tasks, achieving pixel-level accuracy through the use of convolutional neural networks (CNNs). These models have proven effective in providing detailed delineations of urban components [2][3]. However, much of the research in this area focuses solely on segmentation, often neglecting the integration of results with downstream analyses such as material classification or quantitative assessments like surface area calculations.

Material classification techniques are crucial for understanding the thermal properties of urban structures. CNN-based approaches have been employed to classify materials such as concrete, glass, and metal based on visual data. These methods have demonstrated their effectiveness in controlled settings but frequently depend on large labeled datasets, which limits their adaptability in diverse urban contexts. Furthermore, material classification is rarely integrated with segmentation workflows, leaving a gap in methodologies for comprehensive urban heat analysis [6].

Surface area calculation is another essential component of thermal analysis, providing quantitative data about the size and extent of heat-retaining structures. Traditional methods for surface area estimation often rely on pixel-based scaling or geometric modeling. While these approaches are computationally efficient, their accuracy depends heavily on the quality of segmentation masks. Moreover, these calculations are typically conducted as standalone tasks, disconnected from segmentation or material classification, which restricts their applicability in integrated urban heat studies [8].

Recent advancements in object detection and segmentation models, such as YOLOv8 and the Segment Anything Model (SAM), have further enhanced the capabilities of urban image analysis. YOLOv8 is renowned for its real-time detection performance, while SAM offers task-agnostic segmentation with minimal dependence on labeled datasets. These tools enable efficient and precise analysis of urban environments but are rarely utilized in unified frameworks that combine segmentation, material classification, and surface area calculations [4][7]. The lack of integration among these advanced tools remains a critical gap in the field.

This study aims to address these challenges by developing a cohesive pipeline that combines semantic segmentation, material classification, and surface area calculation. By leveraging state-of-the-art models like YOLOv8 and SAM, this research seeks to advance urban analysis methodologies, enabling more precise and scalable thermal assessments of urban areas

Research Gap

Despite advancements in urban image analysis techniques, significant gaps remain in achieving an integrated and scalable framework for Urban Heat Island (UHI) analysis. This section highlights the limitations in current methodologies and the areas that this research aims to address.

One of the critical challenges is the lack of integration among semantic segmentation, material classification, and surface area calculation. Models such as U-Net and DeepLabv3+ provide high-quality segmentation results, but their outputs are often not extended to downstream tasks such as material identification and quantitative thermal assessments [2][3]. Similarly, object detection models like YOLOv8 excel at real-time detection of urban components but are rarely combined with segmentation and material classification workflows [4]. This fragmentation results in isolated analyses that fail to provide holistic insights into the thermal impacts of urban structures.

Another key limitation is the dependency on extensive task-specific labeled datasets. Many segmentation and material classification models require large annotated datasets, such as Cityscapes and ADE20K, to achieve acceptable performance [5]. This reliance on labeled data limits the scalability of these methods, particularly in diverse urban environments where dataset availability and variety are limited. Emerging models like the Segment Anything Model (SAM) address some of these challenges by enabling task-agnostic segmentation, but their application in integrated UHI analysis frameworks remains unexplored [7].

Generalization across varied urban landscapes is another pressing issue. Urban environments differ significantly in terms of structure, lighting, and materials. Existing segmentation and classification methods often struggle to maintain accuracy across these variations, leading to inconsistent results [6]. This inconsistency hinders their effectiveness in real-world UHI analysis, where robust performance across diverse conditions is essential.

Furthermore, surface area calculation methods are often treated as standalone tasks, disconnected from segmentation and material classification. While geometric modeling and pixel-based scaling techniques are effective, their accuracy is highly dependent on the quality of segmentation masks. Additionally, these methods fail to account for the material-specific properties of urban components, which are critical for accurate thermal analysis [8].

Finally, there is a gap in the integration of real-time thermal data with image-based analysis. IoT devices that measure temperature and heat distribution could provide valuable insights when combined with segmentation and classification outputs. However, current frameworks lack the ability to merge these data sources into a unified pipeline for comprehensive UHI analysis and mitigation [9].

This research addresses these gaps by proposing an integrated framework that combines YOLOv8 for object detection, SAM for semantic segmentation, and algorithms for material classification and surface area calculation. The proposed solution aims to enhance scalability, generalization, and integration, paving the way for precise and efficient urban thermal analysis.

Research Problem

The Urban Heat Island (UHI) effect, a phenomenon where urban areas experience significantly higher temperatures than surrounding rural regions, has become a critical environmental and public health challenge. This issue arises from the extensive use of heat-retaining materials such as concrete, glass, and asphalt, coupled with reduced vegetation and green spaces. Mitigating UHI effects requires precise thermal analysis, which relies on accurate segmentation of urban structures, material identification, and surface area estimation. However, current methodologies for addressing these needs remain fragmented, inefficient, and inadequate for large-scale urban environments.

Existing segmentation techniques, such as U-Net and DeepLabv3+, have demonstrated effectiveness in dividing urban images into meaningful components. However, these methods often lack integration with material classification and surface area calculations, which are essential for understanding thermal properties. Similarly, object detection models like YOLOv8 excel in real-time detection but are rarely extended to support downstream analyses for UHI mitigation [2][3][4].

A major limitation of current approaches is their reliance on extensive task-specific labeled datasets, such as Cityscapes and ADE20K, which restrict their scalability to diverse urban environments. Emerging models like the Segment Anything Model (SAM) reduce the dependency on labeled data, offering task-agnostic segmentation capabilities. However, their potential remains underexplored in frameworks that combine segmentation with material classification and surface area analysis [5][7].

Another challenge is the generalization of existing methods across diverse urban settings. Urban environments vary significantly in terms of lighting conditions, weather patterns, and structural compositions, which negatively impact the performance of segmentation and material classification models. Additionally, surface area calculations are often conducted as isolated tasks, disconnected from segmentation and material-specific workflows, limiting their accuracy and real-world applicability [6][8].

Lastly, the integration of real-time temperature data from IoT devices with segmentation and classification workflows is an underdeveloped area. While IoT technologies provide valuable thermal insights, their potential to complement image-based analyses for UHI mitigation has yet to be fully realized [9].

This research addresses these challenges by proposing a unified framework that integrates YOLOv8 for object detection, SAM for segmentation, and algorithms for material classification and surface area calculation. By bridging these gaps, the study aims to develop a scalable, accurate, and efficient solution for urban thermal analysis, paving the way for actionable UHI mitigation strategies.

Objectives

The objectives of this research are structured to address the challenges of urban thermal analysis by developing an integrated framework for segmentation, material classification, and surface area calculation of urban components. These objectives are divided into **main** and **specific** goals to ensure clarity and focus.

Main Objective

Develop an end-to-end system for urban image analysis that combines object detection, semantic segmentation, material classification, and surface area calculation, enabling precise and scalable urban thermal analysis.

Specific Objectives

1. Data Collection and Preparation:
    a. Gather and preprocess urban imagery datasets, such as Cityscapes and ADE20K, for use in training and testing models.
2. Object Detection:
    a. Implement and fine-tune YOLOv8 to detect key urban components, including buildings, pavements, and rooftops, in diverse environments.
3. Semantic Segmentation:
    a. Use the Segment Anything Model (SAM) to generate precise segmentation masks for detected objects.
4. Material Classification: (Pending Finalization)
    a. Explore machine learning-based approaches, such as convolutional neural networks (CNNs), to classify material types (e.g., concrete, glass, metal) for each segmented component.
5. Surface Area Calculation:(Pending Finalization)
    a. Develop and implement pixel-based or geometric modeling techniques to calculate the surface area of each segmented component accurately.
6. System Integration:
    a. Combine all components into a unified pipeline, ensuring seamless interaction between object detection, segmentation, material classification, and surface area calculation.
7. Evaluation and Validation:
    a. Test the integrated system on diverse datasets to evaluate its accuracy, scalability, and reliability.
    b. Refine the system based on evaluation metrics such as Intersection over Union (IoU), precision, recall, and overall efficiency.

Methodology

Overall System Description

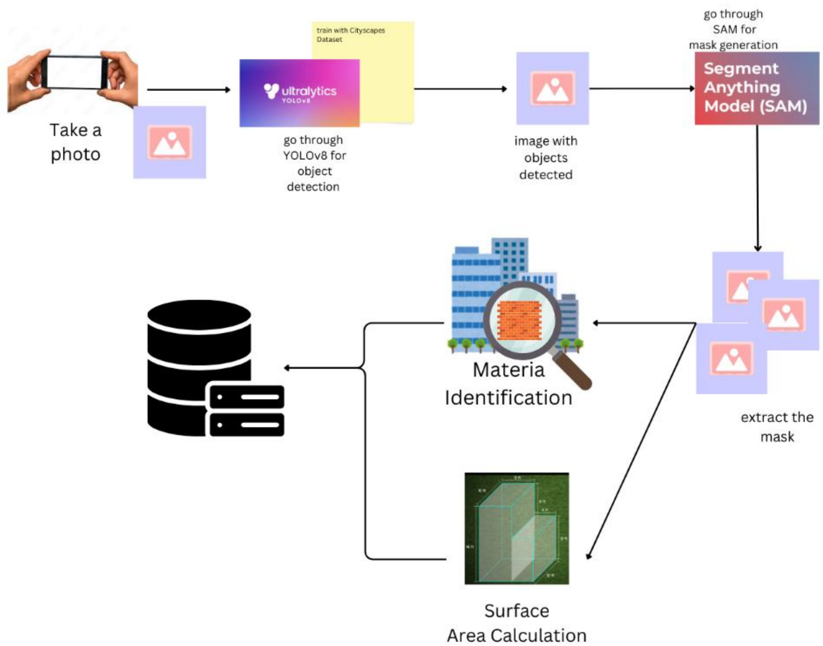

Figure 0-1 system diagram

This research proposes a framework that integrates object detection, semantic segmentation, material identification, and surface area calculation for analyzing urban images. The process begins with YOLOv8 for detecting key urban components, such as buildings, pavements, and rooftops, followed by precise mask generation using the Segment Anything Model (SAM). The segmented components will then be analyzed for material classification and surface area computation using flexible algorithms and image processing techniques.

Tasks and Sub-Tasks

1. Data Collection and Preprocessing
    a. Collect annotated urban images from datasets such as Cityscapes and ADE20K.
    b. Perform data augmentation (rotation, scaling, flipping) to enhance model generalizability.
2. Object Detection
    a. Utilize YOLOv8 for detecting urban components.
    b. Fine-tune YOLOv8 using collected datasets for optimal performance on urban features.
3. Semantic Segmentation

    a. Use SAM for generating segmentation masks for YOLOv8-detected objects.
    b. Refine segmentation masks using post-processing techniques such as morphological operations.
4. Material Identification (Method to be finalized)
    a. Evaluate machine learning-based classifiers (e.g., CNNs) for material identification.
    b. Train a model to classify materials such as concrete, glass, and metal using labeled data.
5. Surface Area Calculation (Method to be finalized)
    a. Apply pixel-based or geometric modeling methods to calculate the surface area of segmented components.
    b. Validate results by comparing calculated areas with ground truth or manually measured data.
6. Integration and Testing
    a. Develop a unified pipeline combining all components.
    b. Test the pipeline on diverse urban datasets to evaluate accuracy and scalability.

Materials and Resources Needed

1. Hardware
    a. High-performance GPUs for model training and inference.
2. Software and Libraries
    a. PyTorch/TensorFlow for implementing YOLOv8 and SAM.
    b. OpenCV for image processing tasks.
3. Datasets
    a. Cityscapes and ADE20K for urban imagery.
    b. Additional labeled datasets for material classification, if needed.

Data Collection and Analysis Methods

1. Data Requirements
    a. Urban images annotated with object categories and material properties.
2. Data Collection
    a. Use publicly available datasets.
    b. Create additional annotations for materials if required.
3. Analysis
    a. Evaluate the accuracy of object detection, segmentation, and material classification models.

Anticipated Timeline

Figure 0-2 Gantt chart

Work breakdown

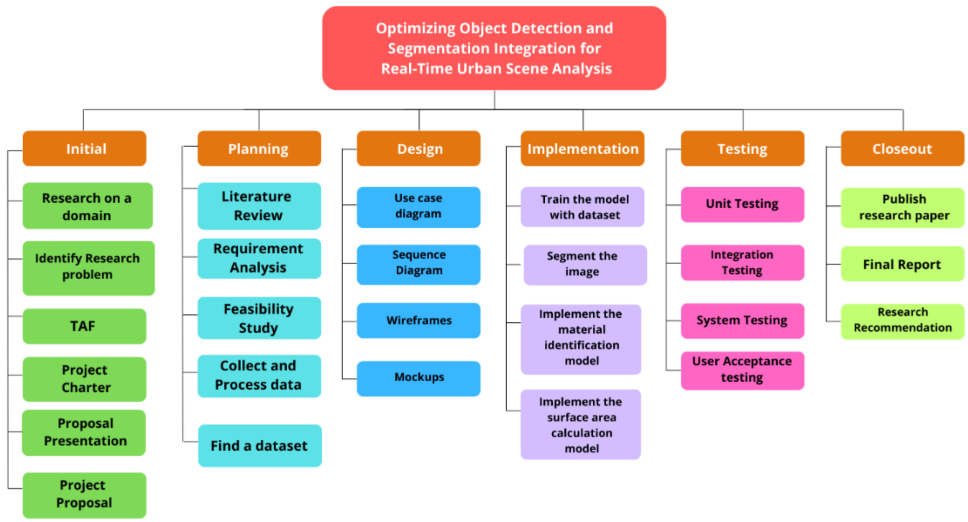

Figure 0-3 work breakdown

Expected results and applications

Segmentation Accuracy	High-quality masks generated for urban components.
Material Classification	Reliable identification of material types (e.g., concrete, glass).
Surface Area Calculation	Accurate area estimation for each segment.
Real-World Application	Enhanced tools for urban thermal analysis and planning.

Project Requirements

The project requirements encompass the functional, non-functional, user, and system needs to ensure the successful implementation of the proposed framework. These requirements provide a comprehensive overview of what is necessary to develop, deploy, and evaluate the system.

Functional Requirements

1. Detect urban components (e.g., buildings, pavements, rooftops) in images using YOLOv8.
2. Generate precise segmentation masks for detected objects using the Segment Anything Model (SAM).
3. Classify material types (e.g., concrete, glass, metal) for each segmented component.
4. Calculate surface areas of segmented components accurately using image processing techniques.
5. Integrate all components into a unified pipeline for urban image analysis.
6. Provide outputs in a format suitable for thermal analysis and urban planning.

Non-Functional Requirements

1. Accuracy:
    a. High detection and segmentation accuracy with minimal false positives/negatives.
    b. Reliable material classification and surface area calculation.
2. Efficiency:
    a. Real-time or near real-time performance for object detection and segmentation.
3. Scalability:
    a. The system should handle large datasets with diverse urban images.
4. Robustness:
    a. The framework should perform consistently across varying lighting, weather, and environmental conditions.
5. Maintainability:
    a. The system should be modular, allowing for future updates or integration of additional features.

User Requirements

1. The system should be easy to use with minimal manual intervention.
2. Provide clear visual outputs (e.g., segmented images with labeled components and calculated surface areas).
3. Offer compatibility with other tools for urban analysis and planning, such as GIS software.

System Requirements

1. Hardware:
    a. High-performance GPUs (e.g., NVIDIA A100) for training and inference.
    b. Sufficient storage for datasets and model outputs.
2. Software:
    a. Programming frameworks such as PyTorch or TensorFlow.
    b. Libraries including OpenCV for image processing and SAM/YOLOv8 implementations.
3. Data:
    a. Annotated datasets like Cityscapes and ADE20K for model training and testing.
    b. Supplementary datasets for material classification.

Budget Plan

Table 1 budget plan

Requirements	Price(Rs.)
Cloud Computing Service	10,000
Domain Registration	4,000
Hardware Products	20,000
Total estimated cost	34,000

Commercialization

Target Audience

- Urban Planning and Development Firms
- Environmental Management Agencies
- Real Estate and Construction Companies
- IoT and Smart City Solutions Providers

Market space

- Urban Planning and Development
- Environmental Sustainability and Research
- Smart City Technologies

References

[1] T. R. Oke, “The energetic basis of the urban heat island,” Quarterly Journal of the Royal Meteorological Society, vol. 108, pp. 1–24, 1982.

[2] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” arXiv preprint arXiv:1505.04597, 2015. [Online]. Available: https://arxiv.org/abs/1505.04597.

[3] L.-C. Chen et al., “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” arXiv preprint arXiv:1606.00915, 2016. [Online]. Available: https://arxiv.org/abs/1606.00915.

[4] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” arXiv preprint arXiv:1804.02767, 2018. [Online]. Available: https://arxiv.org/abs/1804.02767.

[5] M. Cordts et al., “The Cityscapes Dataset for Semantic Urban Scene Understanding,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 3213–3223. [Online]. Available: https://www.cityscapes-dataset.com/.

[6] S. Bell, P. Upchurch, N. Snavely, and K. Bala, “Material Recognition in the Wild with the Materials in Context Database,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015, pp. 3475–3483. [Online]. Available: https://arxiv.org/abs/1412.0623.

[7] A. Kirillov et al., “Segment Anything,” arXiv preprint arXiv:2304.02643, 2023. [Online]. Available: https://arxiv.org/abs/2304.02643.

[8] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3rd ed., Pearson, 2008.