Urban Slum Detection and Mapping: Semantic Segmentation on VHR Satellite Imagery ROBO043

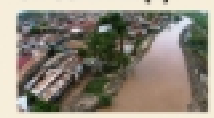
Meenakshi Nair, Mission San Jose High School, Fremont, California, USA

All figures and images are generated by student researcher unless otherwise cite
I certify the item is created by myself without any AI.

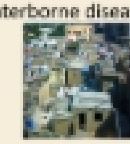
Q1- Research Question & Goal

- Worldwide, **~1.1B** people live in slums with an additional **2B people** expected to live in such settlements in the next 30 years
- Slums face **high temperatures and floods** endangering dwellers
- Adequate housing is a **human right**, and the absence of it negatively affects urban equity, health and, livelihood opportunities
- Slum rehabilitation efforts rely heavily **on slum mapping and monitoring**
- **Current mapping efforts are time consuming and labor intensive**
- Accurate mapping and monitoring of informal settlements requires **manual digitization of satellite imagery**

UN GOAL #11 [1]



Slums in Tanzania exposed to waterborne diseases[2]



Crowded Karachi slums[3]

Goal: Build a robust segmentation & mapping tool using AI to accurately identify structures that aid officials to understand the specific aspects that make slum areas inhabitable

Rationale: Slum mapping technology acts as a crucial tool for
- Informed decision-making
- Assessing vulnerability levels to disasters
- Prioritizing resources and for targeted interventions.

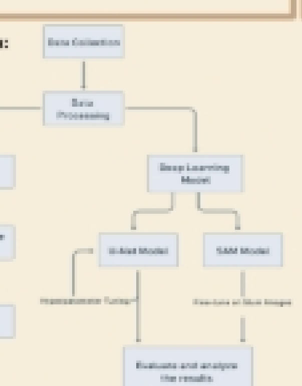
Q2- Framework & Methodology

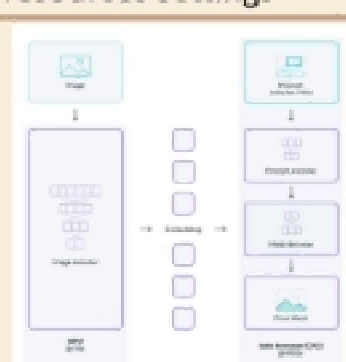
Interpretable AI Model
- Investigating heuristics to unravel factors in slum identification, offering actionable insights for rehabilitation.
- Trained Support Vector Machine (**SVM**) for patch-level predictions

GeoPrompt SAM AI Model
- Pre-trained weights for SAM
- Fine tune the SAM Model by retraining the mask decoder on slum images
- Computationally efficient alternative for less resources settings

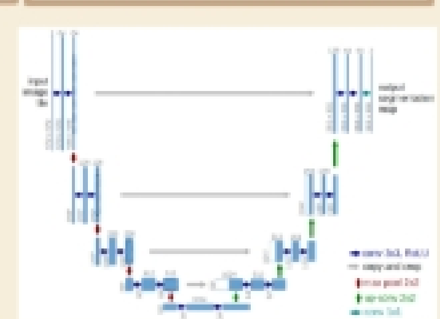
GeoSat-UNet AI Model
- U-Net custom model
- Skip connections preserve spatial information
- Sigmoid activation to transform the model into a binary task: slum or not slum.

Design Diagram:





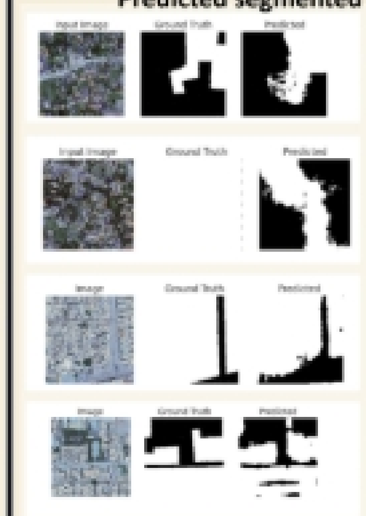
Source - https://arxiv.org/abs/2304.02643



Source: https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5

Q3- Findings

GeoSAT-Unet Model Results: Actual and Predicted segmented mapping



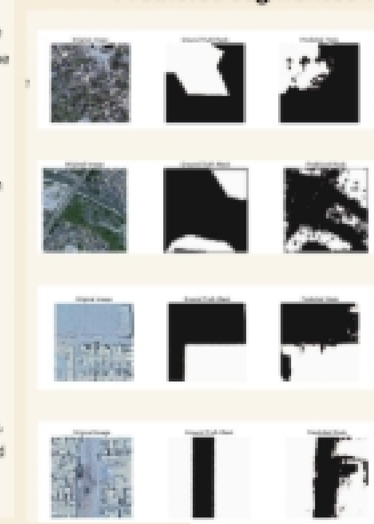
Struggled to capture the intricate details along the edges of slum regions

Predicted the whole image was a slum, when some areas were urban areas

Correctly captured the edges of the road, but some areas of slum were not identified

Correctly identified the upper part of the image, but incorrectly predicted a small strip of urban area in a slum area

GeoPrompt-SAM Model Results: Actual and Predicted segmented mapping



Finding general slum area, but limitations in segmenting edges sharply as well as some random white points scattered

Struggling with precise segmentation with multiple roads and vegetation areas, but capturing the general areas well

Correctly identifying the general area, but struggling with completing some areas

Correctly segmenting the road, but incorrectly identifying some buildings as urban

Model	IoU	Dice Coefficient	Pixel Accuracy
Baseline SAM Model			
SAM - Karachi	25.4%	38.5%	59.4%
SAM - Tanzania	45.9%	60.7%	51.2%
GeoPrompt-SAM & GeoSAT-UNet			
GeoPrompt-SAM - Karachi	78.3%	87.1%	90.7%
GeoPrompt-SAM - Tanzania	68.1%	80.4%	87.1%
GeoSAT-UNet - Karachi	74.7%	85.5%	85.8%
GeoSAT-UNet - Tanzania	61.7%	76.3%	74.9%

Table 1 : Table of results for all segmentation models showing the IoU, dice coefficient, and pixel accuracy. The baseline model was the SAM model before fine tuning on slums

Best Model: GeoPromptSAM in Karachi
- Accuracy: **90.7% (+31.3% from baseline)**
- IoU: **78.3%**
- Dice Coefficient: **87.1%**

Best Model: GeoSAT-UNet in Karachi
- Accuracy: **85.8% (+34.6% from baseline)**
- IoU: **74.7%**
- Dice Coefficient: **85.5%**

Q4- Interpretation & Conclusions

Research addresses the **critical need** for efficient slum monitoring by creating **novel, robust, fast, and accurate** deep learning models for mapping with **accuracy over 90%**

- The interpretable model provides valuable insights into feature importance, aiding **refining slum identification** processes
- GeoSat-UNet, with further computational resources and training time has potential for further improved performance.
- GeoPrompt-SAM presents a promising avenue for future research in leveraging advanced pre-trained models for specific urban mapping applications
- Both models show **comparable performance** in terms of segmentation IoU and accuracy to existing models

Future work
- Beyond slum mapping, broader applications in Earth and environmental remote sensing data
- Explore partnerships with governmental agencies and international organizations, such as the United Nations, to integrate the models into **practical applications** for slum mapping and urban development initiatives

Image source: [1] https://sdgs.un.org/goals/goal11 [2]https://floodlist.com/africa/tanzania-drones-help-communities-map-flood-risk-dar-es-salaam-slums