Green Sense

Dor , Adi and Noam

[1] Affiliation 1; e-mail@e-mail.com
[2] Affiliation 2; e-mail@e-mail.com
* Correspondence: e-mail@e-mail.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials)

Abstract

Monitoring the quality of urban green spaces is critical for urban resilience, resident wellbeing and efficient maintenance planning. Green Sense is a computer vision pipeline designed to assess green space quality automatically from street level imagery, distinguishing between three conditions of vegetation: Healthy, Dried and Contaminated. A labelled synthetic dataset has been constructed using a visual generative AI model (FIBO) configured to mimic typical street camera views. Methodologically, the work follows a staged design. First, a baseline classifier that replicates the approach of the study Green Space Quality Analysis Using Machine Learning Approaches is trained on the synthetic dataset to provide a comparable reference point. Second, a structured model improvement pipeline is applied, combining hyperparameter tuning, exploration of alternative architectures and refined data preparation, with all configurations evaluated on both the internal test set and additional generated images that simulate street camera conditions. The best performing configuration is designated as the Green Sense model. In the next stage, the selected model will be evaluated on authentic street camera images from a municipal authority, in order to compare synthetic and real world performance and to assess readiness for practical deployment in urban green space monitoring.

Keywords: Urban green spaces, street level imagery, synthetic dataset, generative AI, image classification, deep learning, transfer learning, hyperparameter tuning, environmental monitoring, municipal maintenance

1. Introduction

Aaaאבסטרכט בסוף ואינטרודקשן אפשר לכתוב

2. Literature Review

Urban green spaces are a critical component of modern cities, shaping residents' quality of life, social cohesion and public health. As these areas age and intensify in use, municipalities face growing pressure to systematically monitor their cleanliness, vegetation quality and overall maintenance level at scale. In parallel, advances in computer vision and spatial analysis have enabled new approaches to automatically classify and evaluate green spaces in civil environments, moving beyond traditional manual surveys. This literature review will first show the societal importance of clean and well maintained urban greenery and existing monitoring practices, and then examine the main methodological frameworks and models used to classify green spaces from visual and spatial data.

According to Bedimo Rung, Mowen and Cohen, urban green spaces and parks provide accessible settings for regular physical activity, which is associated with reduced risks of obesity, cardiovascular disease, diabetes and other chronic conditions. They also report psychological, social and environmental benefits of green spaces, including improved mood, lower stress and fewer symptoms of depression and anxiety, stronger neighborhood ties and social capital, and contributions to cleaner air and mitigation of urban heat [1].

Kaplan’s evaluation of a “vest pocket” park in downtown Ann Arbor shows that even a very small green space in a dense business district functions as a valued refuge that offers quiet, rest and visual relief. Users reported that simply having this planted, shaded spot nearby improved their daily experience of the city, highlighting the strong restorative benefits of accessible greenery in urban cores [2].

According to Madureira et al., the most valued green space characteristics are cleanliness and good maintenance, richness in plant species, the presence of water bodies, sufficient benches and a generally tranquil atmosphere. Cleanliness and maintenance stand out as the single most important attribute across the three cities, more important than park size, parking or high visitor numbers [3].

Greenspaces are increasingly recognized as critical urban infrastructure rather than a “nice to have” amenity, because they support physical and mental health, reduce health inequalities and help cities adapt to challenges such as air pollution, heat and flooding. In response, municipalities and their partners now treat parks and wider green infrastructure as natural capital: they use spatial planning tools, inclusive design, community programs and formal valuation methods to demonstrate benefits and to prioritize investment. Funding typically comes from a mix of local authority budgets, national health and environment strategies, and dedicated mechanisms such as the Community Infrastructure Levy (a planning charge on new developments used to fund local infrastructure, including parks and green spaces) and Section 106 developer contributions (site specific legal agreements where developers fund or provide infrastructure to mitigate the impacts of their projects), as well as external grants from bodies like the National Lottery Heritage Fund and National Trust. While the review stresses that budgets are under pressure, it shows that targeted investment in high quality, accessible greenspace is seen as a relatively low cost way to deliver health, social and environmental priorities [4].

This paper [5] contrast conventional satellite based indicators of urban green cover, particularly the Normalized Difference Vegetation Index (NDVI), with a human scale measure of street greenery that better reflects residents’ daily visual exposure. Using Google Street View imagery and a SegNet convolutional network, they segment vegetation at pixel level and compute a Green View Index (GVI) for each street segment, which is then classified into low, medium and high greenery based on expert judgements. This GVI based metric is integrated with space syntax measures of street accessibility to derive indicators of “daily accessed greenery,” capturing where people are most likely to experience greenery during routine pedestrian and commuting trips. When compared with NDVI at planning area scale, the daily accessed greenery indicators show only partial correspondence, leading the authors to conclude that top down NDVI alone cannot represent the benefits actually enjoyed by city residents and should be complemented by street level, machine learning based assessments of visible greenery in planning practice.

In the study “Deep Green Diagnostics Urban Green Space Analysis Using Deep Learning and Drone Images,” [6] the authors present an operational framework for fine grained assessment of urban vegetation health and surface contamination using high resolution aerial imagery. The work addresses the limitation of satellite based indices such as NDVI that only capture overall vegetation quantity and cannot distinguish between healthy, dry or polluted green areas. To overcome this, the authors construct a dataset of geo referenced RGB images acquired by a DJI Phantom 4 drone flown over four distinct urban environments in Mexico. The raw images are tiled into patches of 200 by 200 pixels, from which 9 901 samples are manually labelled into eight classes that jointly encode vegetation condition healthy, dry, unhealthy or no vegetation and the presence or absence of visible solid waste contamination. On this dataset they train a convolutional neural network followed by a multilayer perceptron classifier that learns discriminative visual features directly from the image patches and assigns each tile to one of the eight semantic categories. The resulting system achieves around 72 percent accuracy on a held out test set, with particularly strong performance for healthy vegetation and no vegetation classes, and demonstrates reasonable transferability when retrained on small samples from a new area. Overall, this article shows that deep learning applied to drone imagery can provide spatially detailed, operational information on both vegetation condition and urban cleanliness, and offers an open source prototype that municipal or health authorities can adapt for monitoring and decision support.

The article "Green Space Quality Analysis Using Machine Learning Approaches" [7] develops a full pipeline for assessing the visual quality of urban green spaces using machine learning on eye level images. It defines quality through conditions such as cleanliness, maintenance and dryness, builds a labelled dataset with three classes healthy, dried and contaminated, and compares several transfer learning image classification models, selecting ResNet50 as the best performing model and deploying it as a practical web tool for green space monitoring. Methodologically, the study follows the CRISP DM lifecycle, collecting 944 smartphone photos of Kuala Lumpur green spaces in 2022, manually annotating them into the three quality classes, and then addressing class imbalance through data augmentation using flips, rotations, brightness and contrast changes and CLAHE to obtain a balanced dataset of 986 images. The augmented images are resized to 224 by 224 pixels and split into training, validation and test sets, and nine ImageNet pretrained convolutional networks ResNet50, ResNet101, DenseNet201, VGG16, VGG19, XCeption, MobileNet, InceptionResNetV2 and EfficientNetB7 are fine tuned using Adam optimisation, categorical cross entropy loss, early stopping and learning rate reduction, with performance evaluated via accuracy, precision, recall, F1 score, Cohen Kappa and ROC AUC. ResNet50 is chosen as the overall best model, reaching about 99 percent accuracy together with almost perfect precision, recall and ROC AUC, and therefore selected for deployment in a Streamlit web application that allows users to upload or capture a photo and receive an automatic quality label, while the authors conclude that such transfer learning pipelines on relatively small eye level datasets can already deliver operational tools for green space maintenance support and point to mobile deployment and segmentation based models as important future directions.

3. Methodology

This chapter outlines the methodological pipeline of the Green Sense project, from dataset construction through baseline replication, model enhancement and real world validation.

A labelled image dataset is constructed to simulate street camera views of urban green spaces. The dataset includes approximately 300 images for each of three classes: Healthy, Dried and Contaminated. Healthy denotes vital, well maintained vegetation; Dried denotes visible lack of irrigation or plant vitality; Contaminated denotes vegetation areas affected by litter, waste or similar pollution.

Images are generated using a visual generative AI model (FIBO) configured to mimic realistic street level perspectives with variation in lighting, distance, angle and surrounding built environment. All images are manually reviewed and labelled according to a written annotation guide with visual examples. The dataset is split into training, validation and internal test sets with preserved class proportions.

The model improvement pipeline is implemented as three concrete steps, evaluated on newly generated data that mimics street camera conditions in order to compare all configurations to the baseline. First, hyperparameters such as learning rate, batch size, number of epochs and regularisation strength are tuned on the validation set using grid or random search, and the best configuration is re evaluated on both the internal test set and the new synthetic street camera set. Second, alternative architectures, including different convolutional backbones and efficient modern networks, are trained under the tuned hyperparameters and compared on the same evaluation sets. Third, the data pipeline is refined through augmentations such as random crop, flip, rotation and colour jitter, alternative preprocessing schemes and, where needed, class balancing using class weighted loss or oversampling. For each stage, models are scored on the generated street camera like data and on the internal test set, and the configuration with the highest overall performance is designated as the Green Sense model.

In the final stage, external validity is assessed on authentic street camera images provided by a municipal authority. Frames containing public green spaces are labelled using the same protocol and evaluated with the selected Green Sense model, enabling direct comparison between synthetic and real world performance and indicating the readiness of the system for practical deployment.

4. Generative

This chapter describes the construction of a synthetic generative dataset that transforms the original baseline images into street camera view representations of urban green spaces. It then outlines the validation procedure used to ensure that the generated images remain realistic and that their inherited labels for Healthy, Dried, and Contaminated conditions remain consistent and reliable.

4.1 Synthetic generative Dataset

In order to create…

4.2 Validation of Dataset

Validating the perceptual quality of generated images is essential to ensure that the synthetic street view dataset remains visually realistic and that its assigned labels remain trustworthy, especially under the substantial domain shift from the original eye level close up imagery to a wider street camera viewpoint. In this work, validation is performed in two stages. A human screening step first verifies scene plausibility and label correctness. NIQE, the Natural Image Quality Evaluator, is then applied as an objective no reference measure of image naturalness, enabling consistent threshold based filtering.

NIQE is a no reference image quality metric that estimates perceptual naturalness from the image itself, without relying on human opinion scores. It is widely used to compare the visual quality of outputs in settings where a true reference image is not available or not meaningful. The method is based on the idea that real world photographs follow stable natural scene statistics, and that artifacts or unrealistic textures shift these statistics in measurable ways.

NIQE computes natural scene statistics features from local 96×96 luminance patches, fits these features with a multivariate Gaussian model, and measures the distance to a reference model learned from high quality natural images. The score is nonnegative and lower values indicate better perceptual quality. In practice, typical values for natural images are often reported around 3 to 20, while much higher scores indicate stronger deviations from natural image statistics.

Overall, the validation results indicate that the synthetic street view dataset satisfies the perceptual quality requirements for reliable model evaluation. The average NIQE score of the generated images was lower than that of the original eye level baseline set, suggesting higher statistical naturalness in the synthetic outputs. This trend is also evident in the Dry class, where the original images achieved an average NIQE of 10.70 while the synthetic images achieved 8.55. When combined with human inspection confirming visual realism and label consistency, these findings support that the generated dataset meets the defined quality standards and is appropriate for testing model performance and strengthening confidence in generalization toward future real world deployment.

להוסיף השוואה של בריא ומלוכלך ולשים בטבלה

לעשות תמונות דוגמאות ללפני ואחרי

לתת פרקי משנה בין התאוריה לתוצאות

5. Results

This Results chapter first reports the benchmark performance from the original study, "Green Space Quality Analysis Using Machine Learning Approaches," in order to establish a reference baseline for the green space quality classification task. It then presents results on the synthetic generative dataset of street camera view images derived from the original data in two stages. First, the baseline method is evaluated on the synthetic dataset to quantify performance under the shift from standard eye level imagery to street camera perspectives. Second, a new approach is evaluated on the same synthetic dataset and compared directly against the baseline results.

3.1 Baseline results

The baseline research used transfer learning with pre-trained Convolutional Neural Network (CNN) models to assess green space quality. Nine different ImageNet pre-trained models served as baselines: ResNet50, ResNet101, DenseNet201, VGG-16, VGG-19, XCeption, MobileNet, InceptionResNetV2, and EfficientNetB7.

All green space images were resized to 224×224 pixels before being fed into the pre-trained models. The dataset contained 986 total images after augmentation, split into 70% training, 20% testing , and 10% validation. A dense layer with SoftMax activation was built in the final layers to provide probability distributions across the three green space quality classes: Healthy, Dried, and Contaminated. All nine baseline models achieved strong performance, as shown in Table 1, with accuracy ranging from 92.13% to 99.75%.

Table 1. Performance metrics of benchmark's models

Model	Accuracy	Precision	Recall	F1-Score	Cohen's Kappa	ROC-AUC
EfficientNetB7	99.75%	99.75%	99.75%	99.75%	1.00	1.00
ResNet101	99.49%	99.49%	99.49%	99.49%	0.99	1.00
VGG-19	99.24%	99.24%	99.24%	99.24%	0.99	1.00
ResNet50	98.98%	99.00%	98.98%	98.98%	0.98	1.00
DenseNet201	98.22%	98.24%	98.22%	98.23%	0.97	1.00
MobileNet	95.69%	95.81%	95.69%	95.70%	0.94	0.99
VGG-16	95.43%	95.43%	95.43%	95.43%	0.93	0.99
InceptionResNetV2	94.92%	94.96%	94.92%	94.93%	0.92	1.00
XCeption	92.13%	92.13%	92.13%	92.12%	0.88	0.98

3.2 Results on original data

The first step was to replicate the benchmark model architecture in order to enable a rigorous comparison. The architecture diagram is shown in Figure 1. Because the original implementation was not publicly available, the initial task was to reconstruct a model that matches the benchmark as closely as possible and achieves comparable performance.

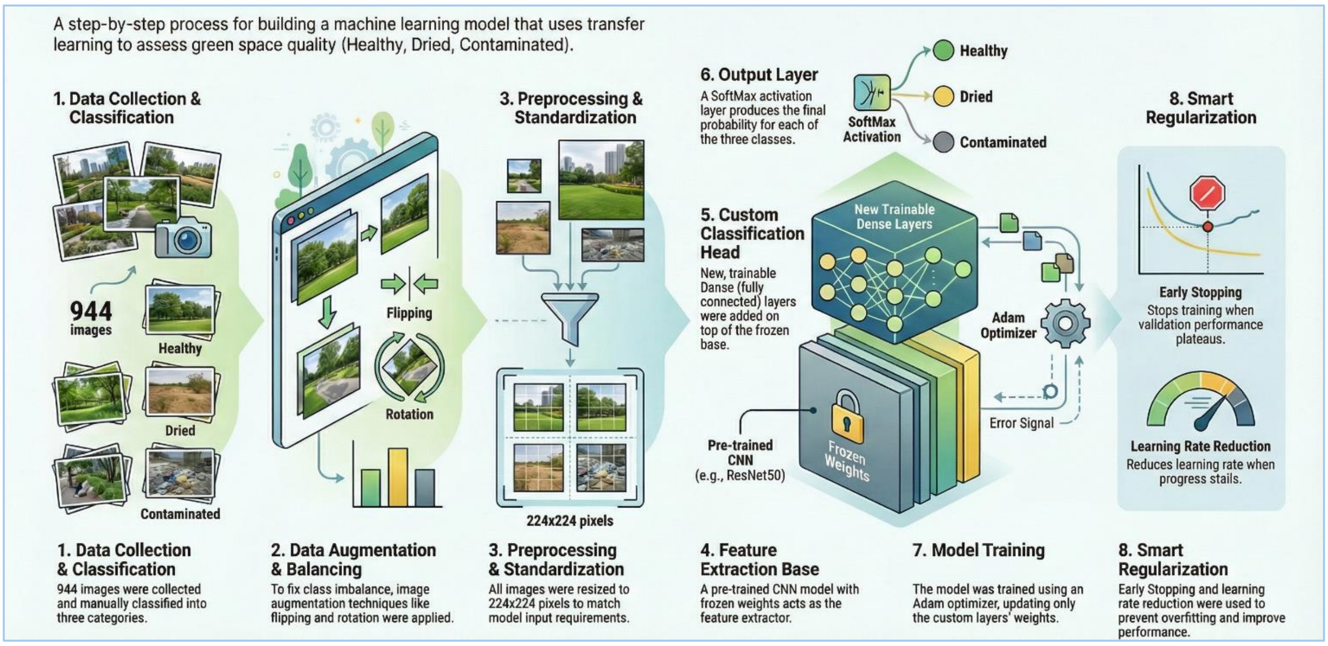

Figure 1. Architecture diagram for baseline comparison.

3.3 Results on synthetic data

References

[1] P. A. J. M. P. D. A. C. M. Ariane L. Bedimo-Rung, "The Significance of Parks to Physical Activity and Public Health".

[2] R. Kaplan, "EVALUATION OF AN URBAN VEST-POCKET PARK °," *USDA FOREST SERVICE.*

[3] F. N. ,. J. V. O. a. T. M. Helena Madureira, "Preferences for Urban Green Space Characteristics: A Comparative Study in Three Portuguese Cities," *enviornments.*

[4] "Improving access to greenspace A new review for 2020," *Public Health England.*

[5] S. L. Donghwan Ki, "Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning," *Landscape and Urban Planning.*

[6] H. C. C. A. D. P. L.-J. I. A. V.-M. a. M. S. S.-C. Marco A. Moreno-Armendáriz, "Deep Green Diagnostics: Urban Green Space Analysis Using Deep Learning and Drone Images," *sensors,* 2019.

[7] Z. R. a. N. Z. Jaloliddin Rustamov, "Green Space Quality Analysis Using Machine Learning Approaches," *sustainability.*

[8] ∗. D. R. Y. L. X. S. Y. Z. W. Z. T. Z. Yu Yea, "Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices," *Landscape and Urban Planning.*