From Above and Beyond: Decoding Urban Aesthetics with the Visual Pollution Index

Advait Gupta, Manan Padsala, Devesh Jani, Tanmay Bisen, Aastha Shayla, Susham Biswas

ABSTRACT Urban landscapes, emblematic of modernization and growth, are increasingly faced with the intricate challenge of visual pollution. This nuanced form of pollution, often overshadowed

Advait Gupta
Department of Computer Science, Rajiv Gandhi Institute of Petroleum Technology, Jais, UP, India
email: 20cs3072@rgipt.ac.in

Manan Padsala
Department of Computer Science, Rajiv Gandhi Institute of Petroleum Technology, Jais, UP, India
email: 20cs3038@rgipt.ac.in

Devesh Jani
Department of Electronics and Communication, Chandubhai S Patel Institute of Technology, Charotar University of Science and technology, Anand, Gujarat, India
email: 20ec026@charusat.edu.in

Tanmay Bisen
Department of Computer Science, Rajiv Gandhi Institute of Petroleum Technology, Jais, UP, India
email: 20cs3064@rgipt.ac.in

Aastha Shayla
Department of Computer Science, Rajiv Gandhi Institute of Petroleum Technology, Jais, UP, India
email: 20cs3001@rgipt.ac.in

Dr Susham Biswas
Department of Computer Science, Rajiv Gandhi Institute of Petroleum Technology, Jais, UP, India
email: susham@rgipt.ac.in

environmental discussions, profoundly influences the aesthetic harmony and mental well-being of urban inhabitants. In this research, we present an innovative methodology to detect visual pollution using drone-captured imagery. Our distinctive dataset captures a spectrum of visual pollutants, from graffiti, faded signage, and potholes to more complex issues like cluttered sidewalks and unkempt facades. Leveraging this dataset, we fine-tuned pre-trained object detection models, specifically YOLOv6, achieving remarkable accuracy in detecting these visual pollutants from images. Central to our study is the introduction of the Visual Pollution Index (VPI), a metric formulated through the multiplicative integration of the Counting Categories Ratio (CCR) and the Severity-Weighted Score (SWS). To provide a spatial representation of visual pollution levels, we further introduce heatmap visualizations. These heatmaps, overlaid on urban maps, offer a vivid depiction of pollution hotspots, enabling city planners and stakeholders to pinpoint areas of concern. Grounded in real-world perceptions, our approach offers a comprehensive lens to assess, visualize, and address visual pollution in urban environments.

Key words: Visual Pollution Index, VPI, urban aesthetics, drone imagery, object detection, YOLOv6, visual pollutants, heatmap visualization, urban planning, environmental impact, Counting Categories Ratio, CCR, Severity-Weighted Score, SWS, urban well-being, geospatial analysis, automated detection, urban environmental management

1 INTRODUCTION

Urban environments, as the epicenters of human activity and innovation, have witnessed unprecedented growth over the past few decades. While this growth has brought about numerous advancements and opportunities, it has also introduced a myriad of challenges, one of which is visual pollution. Visual pollution, a term that encompasses unsightly and out-of-place man-made objects within public and private spaces, has become a growing concern for urban planners, environmentalists, and city dwellers alike [1].

The concept of visual pollution is not new; however, its significance has grown in tandem with rapid urbanization. Visual disturbances, ranging from graffiti, faded signage, and potholes to cluttered sidewalks and unkempt facades, can degrade the aesthetic appeal of urban areas, impacting not only the visual harmony but also the psychological well-being of residents [2]. Such disturbances can lead to decreased property values, reduced tourist interest, and even adverse health effects due to stress and mental fatigue [3].

With the advent of technology, particularly in the realms of drone imagery and machine learning, there exists an opportunity to address this issue in a more systematic and data-driven manner. Drones, with their ability to capture high-resolution images from vantage points previously inaccessible, offer a unique perspective on urban landscapes [4]. When combined with advanced object detection algorithms, such as YOLOv6, these images can be analyzed to detect and quantify visual pollutants with remarkable accuracy [5].

This paper introduces a novel approach to quantify visual pollution using the Visual Pollution Index (VPI), a metric derived from drone-captured imagery and object detection techniques.

Furthermore, we present heatmap visualizations to spatially represent visual pollution levels, providing a tool for urban planners and stakeholders to make informed decisions.

2 LITERATURE SURVEY

Urban environments are increasingly being scrutinized for their aesthetic appeal, given the rapid urbanization and the subsequent challenges it brings. Visual pollution, an often-overlooked aspect, plays a crucial role in determining the aesthetic quality of a place. The term "visual pollution" refers to the entire set of unsightly and visually unpleasing elements in an environment. This can range from graffiti, billboards, overhead power lines, and even the architectural design of buildings.

For instance, a 2016 study by Chmielewski et al. delves into the commercialization of public space by outdoor advertising and its potential negative impact on the quality of life and enjoyment of public spaces [6]. The research illustrates that visual pollution can be quantified by correlating public opinion with the number of visible advertisements. Using a 2.5D outdoor advertisement dataset from a busy urban street in Lublin, Poland, the study translates visibility into visual pollution. The findings suggest that streetscape views with more than seven visible advertisements result in visual pollution in the studied context. Our study extends this methodology by incorporating a more comprehensive set of visual pollutants and applying advanced deep learning techniques for detection.

Building on this, a 2019 study by Ahmed et al. delves deeper into the realm of visual pollution [7]. The researchers propose a novel approach to detect visual pollutants using deep learning techniques. More importantly, they suggest the potential of creating a "Visual Pollution Index" (VPI) in the future. This index would serve as a tool for urban planners and professionals in urban environmental management, allowing them to evaluate and compare the visual aestheticism of different geographic regions. While they only proposed the idea, our research has taken the initiative to design and implement such an index.

A distinguishing feature of our research is the utilization of a unique and novel high-quality dataset, which stands in contrast to the commonly used publicly available datasets of lower quality. This innovative dataset enhances the accuracy, detection capabilities, and Intersection over Union (IoU) metrics in our study. The superior quality of our dataset ensures more reliable results, setting our research apart from previous studies in the domain.

Further, a 2019 study by Wakil et al. titled "A Hybrid Tool for Visual Pollution Assessment in Urban Environments" provides a systematic approach for the development of a robust Visual Pollution Assessment (VPA) tool [8]. The research introduces a methodology that integrates both expert and public perspectives to rank Visual Pollution Objects (VPOs). Using empirical decision-making techniques, the VPA tool produces a point-based visual pollution scorecard. After extensive testing in Pakistan, the tool offers regulators a consistent method for visual pollution assessment and equips policy makers with a foundation for evidence-based strategies. Our research builds on this, refining methods to evaluate visual pollutants across urban settings.

Another 2022 study by Alharbi and Rangel-Buitrago focuses on the visible deterioration and negative aesthetic quality of the landscape in coastal environments [9]. Factors such as erosion, marine wrack, litter, sewage, and beach driving are identified as contributors to visual pollution, particularly in the Rabigh coastal area of the Kingdom of Saudi Arabia. The research employs the Coastal Scenery Evaluation System (CSES) to assess the scenic quality of 31 coastal sites.

In addition, a 2016 study by Madleňák and Hudák discusses the concept of "visual smog," which has emerged as a social concern in recent decades [10]. This refers to the contamination of public spaces by aggressive and often illegally placed advertisements that are not proportionate in size. The study aims to measure the level of visual smog on selected road communications, taking into account the number of ads and billboards near roads, the distance between billboards and roads, and the density of these billboards. The research also incorporates an analysis of traffic accidents on the chosen road communications.

Lastly, a study by Yilmaz titled "In the Context of Visual Pollution: Effects to Trabzon City Center Silhouette" examines the influence of visual pollution on city silhouettes [11]. Silhouettes, indicative of a city's history and structure, are becoming markers of visual pollution. The research underscores that many cities now feature buildings that lack harmony and environmental consideration, leading to uniform concrete façades that eclipse their historical essence. Using Trabzon as a case study, a coastal city renowned for its historical richness, the study contrasts old and new city images to assess the aesthetic shifts, highlighting the interventions that have reshaped its distinctive silhouette.

In conclusion, the aforementioned studies collectively highlight the growing importance of addressing visual pollution in urban settings. Our research builds on these foundational studies, offering a more comprehensive and practical solution to the problem. By designing the Visual Pollution Index, we aim to provide urban planners and environmentalists with a robust tool to assess and mitigate visual pollution in our cities.

3 PROBLEM STATEMENT

As urban environments undergo rapid transformation, the issue of visual pollution has become increasingly prominent, impacting the aesthetic and overall quality of life in cities. While various studies have attempted to address this concern, there is a distinct lack of a standardized, universally applicable metric like the Visual Pollution Index (VPI) for quantifying and addressing visual pollutants comprehensively. The development and refinement of such an index, backed by advanced technological methodologies, is crucial for enabling urban planners and environmentalists to make informed decisions and implement effective strategies to combat visual pollution.

4 METHODOLOGY

The entire methodology adopted for this study, including data collection, processing, and analysis, is depicted in a comprehensive flowchart presented in Fig 1.

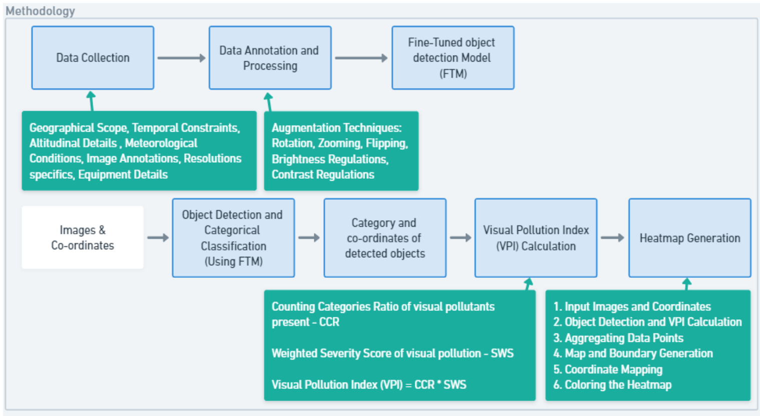

Fig. 1 Detailed Methodology of our proposed research

4.1 Data Collection

The backbone of our study is the caliber and pertinence of its data. We employed a dedicated drone application to systematically collect aerial photos of urban areas. This tool was thoughtfully created to snap photos during the drone's journey, sending them instantly to our exclusive servers. This approach safeguarded data accuracy and immediate accessibility for study.

4.1.1 AREA OF FOCUS.

Primarily, our data encompasses the urban zones of Newtown, Kolkata. The choice of this location was influenced by its urban expansion and the variety of visual contaminants, making it an apt subject for our study [12].

4.1.2 TIMING.

For the best lighting and image clarity, drone operations were planned between 10AM and 4PM. This slot was selected after initial assessments highlighted the difficulties of capturing under variable light conditions.

4.1.3 HIEGHT CONSIDERATIONS.

Image capture altitude varied with the desired view. For top-down visuals, drones hovered between 30 to 60 meters. For other angles, a uniform 10-meter height was maintained, ensuring crisp and detailed photos.

4.1.4 WEATHER CONCERNS.

Recognizing weather's role in image quality, we chose clear days for data collection. Previous studies have emphasized the importance of weather factors in aerial photography, backing our choice.

4.1.5 IMAGE ANOTATIONS.

After collection, each photo underwent a thorough manual annotation procedure using QGIS software [13]. This step was crucial to ensure data accuracy, especially for object spotting and further study.

4.1.6 RESOLUTION DETAILS.

Image clarity was determined by the drone's height. Photos from 30 to 60 meters had clarity ranging from 2.5 to 10 centimeters per pixel. In contrast, those from 10 meters displayed a sharper clarity of under 1 centimeter per pixel, giving detailed views of visual contaminants.

4.1.7 GEAR INFORMATION.

The drones selected for our research were picked for their superior image capture and dependability. We utilized the Phantom 4 PRO, Mavic Mini, and Phantom 4 Pro V2.0, all known for their exceptional image quality in various investigations. In summary, our meticulous data collection method guaranteed the gathering of top-notch, relevant data, establishing a solid base for our subsequent assessments.

4.1.8 DATASET DETAILS.

Post data gathering and tagging, our dataset was completed. It comprises high-definition photos of different urban visual pollutants, each being 700x700 pixels in size. The dataset is divided into specific categories: 600 images of Graffiti, 600 of Faded Signage, 1200 of Potholes, 1250 of Garbage, 600 of Under-Construction Roads, 600 of Damaged Signage, 600 of Malfunctioning Street Lights, 650 of Poorly-Maintained Billboards, 600 of Sand Patches on Roads, 600 of Overcrowded Sidewalks, and 600 of Neglected Building Facades.

Fig. 2 Sample images from the dataset

4.2 Data Preprocessing

To bolster the variety and resilience of our dataset, we incorporated data enhancement methods. Data augmentation serves as a tactic to synthetically increase the training dataset's volume by generating altered renditions of pre-existing images. Given the diverse visual characteristics of pollutants and the model's requirement to identify them in assorted conditions and angles, this approach holds significant pertinence for our research [14].

The augmentation methods we adopted encompass:

- **Rotation:** Images underwent rotation at multiple angles, mirroring the myriad orientations construction materials might be captured in.
- **Zooming**: Certain images experienced zoom-in or zoom-out effects, emulating the varying distances from which construction materials could be photographed.
- **Flipping:** We flipped images both horizontally and vertically to introduce additional variability.
- **Brightness and Contrast Adjustments**: These modifications were made to mimic a range of lighting conditions, ensuring the model's adaptability to diverse environmental settings

It is imperative to emphasize that with each image enhancement, the relevant annotation file was refreshed. This meticulous step ensures that the bounding box details are a true reflection of the object's position and scale in the enhanced images. Automating the task of refreshing these details guaranteed uniformity and precision across the enhanced dataset. By adopting a judicious data enhancement strategy, we not merely expanded our dataset's size but also endowed it with an array of conditions, paving the way for a model that is both steadfast and versatile.

4.3 Object Detection and Visual Pollutant Classification.

For the complex endeavor of pinpointing visual pollutants in overhead photos, we harnessed the prowess of the YOLO v6 framework [5]. Recognized for its instantaneous object detection ability, this architecture was diligently adapted to our specialized dataset brimming with hand-labeled images of various visual pollutants.

4.3.1 YOLO v6 Structure.

YOLO, standing for "You Only Look Once," represents a cutting-edge technique for real-time object identification. Contrasting conventional methods which first generate region suggestions followed by classification, YOLO embodies a distinct methodology. It operates through a unified neural network that scrutinizes the whole image in one go. This network segments the image into several areas, deducing the position and likelihood of objects within these zones. The significance of each inferred location is gauged by its tied probability. The YOLO v6, a subsequent version in the YOLO lineage, introduces further refinements. It's crafted to discern objects within visuals and categorize them on-the-fly, making it particularly suitable for tasks like ours that demand swift and sharp detection. The structure is lauded for its rapidity and precision, made possible through the integration of Darknet platforms and fine-tuned network tiers.

- **Training Details**: The model was trained rigorously for 50 epochs, using a batch size of 64. The hyperparameters used were:
- **Learning Rate**: 0.001 (complemented by a step decay)
- **Momentum**: 0.9
- **Weight Decay**: 0.0005
- **Loss Function**: A combination of Mean Squared Error for bounding box regression and Cross-Entropy for class prediction.

4.4 Visual Pollution Index (VPI) Calculation.

The Visual Pollution Index (VPI) is a pioneering metric introduced in this research to quantify the extent and intensity of visual pollution in urban landscapes using aerial images. The VPI is meticulously crafted to offer a comprehensive perspective on visual pollution, factoring in both the diversity of visual pollutants and their respective severities. For a group of images belonging to a particular area, we detect and store the coordinates and labels of each item detected from all the images. Based on this aggregated data, we calculate the VPI value for that specific area, ensuring a more holistic representation of the visual pollution present. The formulation of VPI is rooted in two primary components: the Counting Categories Ratio (CCR) and the Severity-Weighted Score (SWS).

4.4.1 Counting Categories Ratio (CCR).

The CCR is a measure of the diversity of visual pollutants present in an image. It calculates the ratio of the number of distinct visual pollution categories detected to the total number of possible categories. This ensures that the CCR value lies between 0 (indicating no visual pollutants detected) and 1 (indicating all possible visual pollutants detected). It is calculated as:

$$CCR = \frac{Number\ of\ Distinct\ Categories\ Detected}{Total\ number\ of\ categories}$$

4.4.2 Mapping of Visual Pollutants to Severity Categories

Initially, we categorized visual pollutants into five distinct severity levels: Very Low Severity, Low Severity, Medium Severity, High Severity, and Very High Severity. Each severity level was then associated with a numerical value: 0.2 for Very Low Severity, 0.4 for Low Severity, 0.6 for Medium Severity, 0.8 for High Severity, and 1 for Very High Severity. Using these values, each of the visual pollutants was mapped to one of the severity categories based on established literature and empirical observations. This mapping has been illustrated in Table 1.

Table 1 Severity Categorization of Visual Pollutants

Severity	**Pollutants**
Very Low	**Graffiti, Sand on Road**
Low	**Faded Signage, Cluttered Sidewalk**
Medium	**Bad streetlight, Broken Signage**
High	**Construction Road, Bad billboard, Unkempt Facade**
Very High	**Potholes, Garbage**

This mapping is based on established literature and empirical observations. For instance, graffiti, often perceived as a form of urban art, can sometimes be seen as a sign of neglect or decay, especially when it's unsanctioned or in inappropriate places [15][16]. Sand on roads, while causing minor visual disruption, is generally a temporary issue often resulting from nearby construction or natural causes [17]. Faded signage, while not as intrusive, can give an impression of negligence and can be a safety concern in certain contexts [18]. Cluttered sidewalks, indicative of disorganization, can impede pedestrian movement and give a sense of disorder [19]. Bad street lights not only affect the aesthetics of an area but also raise concerns about safety during nighttime [20]. Broken signage can be confusing for drivers and pedestrians, leading to potential safety risks. Construction roads, indicative of ongoing development, can sometimes be visually disruptive and indicate major infrastructural changes. Bad billboards, especially those that are oversized or have inappropriate content, can dominate the visual landscape and detract from the natural or architectural beauty of an area [21]. An unkempt facade can be a sign of neglect and can significantly downgrade the visual appeal of a building or structure [22]. Potholes are not just visually jarring but also pose significant safety risks, especially in high-traffic areas [23]. Accumulated garbage is a direct indicator of poor sanitation, environmental neglect, and can have health implications [24].

By mapping each visual pollutant to a severity category, we aim to provide a more nuanced understanding of the visual quality of urban landscapes. This ensures that the VPI captures the essence of real-world perceptions and concerns related to visual pollution.

4.4.3 Severity Weight Score

The SWS is designed to quantify the severity of visual pollution by taking into account the severity value assigned to each category and the number of instances of that category in the image. The score is then normalized to ensure its value ranges between 0 (indicating minimal severity) and 1 (indicating maximum severity). It is calculated as:

$$SWS = \frac{\Sigma(\boldsymbol{Severity\,Value} \times \boldsymbol{Number\,of\,Instances\,for\,Each\,Category})}{Total\,units\,belonging\,to\,all\,categories}$$

Here, the Severity Value corresponds to the value of severity against which a particular visual pollutant is mapped, e.g., 0.2 for Very Low Severity. It's worth noting that this severity value can be replaced with custom weights for each category based on specific user needs, allowing for a more tailored assessment of visual pollution in different contexts.

4.4.4 Visual Pollution Index (VPI) Calculation.

The VPI is derived by integrating both the CCR and the SWS, emphasizing the combined importance of pollutant diversity and their respective severities. The formula for VPI is;

$$VPI = CCR \times SWS$$

This results in a VPI value ranging from 0 (indicating minimal or no visual pollution) to 1 (indicating maximum visual pollution).

4.5 Rationale Behind the Multiplicative Approach.

The decision to calculate the Visual Pollution Score/Index (VPSI) using a multiplicative approach (CCR * SWS) stems from the following considerations:

- ➔ **Interplay of Diversity and Severity:** The multiplicative approach ensures that both the diversity of visual pollutants (CCR) and their severity (SWS) are equally emphasized. An area with a singular, yet highly severe pollutant might not be perceived as negatively as an area with multiple pollutants of moderate severity. This approach captures that nuance.

- ➔ **Real-World Perception of Pollution:** In real-world scenarios, the presence of multiple issues, even if individually less severe, can often be perceived as indicative of broader systemic inefficiencies or negligence. Conversely, a singular but severe issue might be seen as an isolated incident. Our multiplicative approach captures this perception effectively. When there's a combination of diverse pollutants, even if they have a slightly lower severity, the overall perceived pollution can be higher than an area with a singular, more severe pollutant.
- ➔ **Comprehensive Assessment:** By considering both the variety and intensity of visual pollutants, the VPSI offers a more holistic view of visual pollution. This ensures that areas with diverse pollutants, even if individually less severe, are given due attention, reflecting the broader public perception and concerns.

The VPI offers a comprehensive perspective on visual pollution, making it an invaluable tool for urban planners, environmentalists, and policymakers. By considering both the variety and severity of visual pollutants, the VPI provides a more nuanced understanding of the visual quality of urban landscapes. In our study, the VPI was computed for a set of aerial images, and these VPI values were then used to generate heatmaps, offering a visual representation of visual pollution intensity across different urban regions.

4.5.1 Heatmap Generation

The creation of a heatmap to visualize the Visual Pollution Index (VPI) across a designated region is an intricate procedure that seamlessly merges the outcomes of object detection with geospatial information. The following elucidates the step-by-step process involved in the heatmap generation:

1) **Grouped Image Input and Coordinates:** For a set of aerial images corresponding to a specific area, we input their collective geographical coordinates (latitude and longitude). This ensures that the VPI, calculated based on the combined data from these images, aligns accurately with its real-world geographical location.
2) **Object Detection and VPI Computation:** Each image within the group undergoes processing by the object detection model to identify and classify visual pollutants. Leveraging the detected pollutants and their respective severities, the VPI for the group of images is determined. This VPI value, in conjunction with the group's coordinates, is archived for the subsequent heatmap creation.
3) **Data Point Aggregation:** As multiple groups of images undergo processing, a repository of VPI values paired with their corresponding coordinates is established. This collective data forms the backbone for crafting a detailed heatmap spanning a vast geographical expanse.
4) **Map and Boundary Formulation:** With the aid of geospatial tools, such as GeoPandas [25], a map of the targeted region (be it a city, district, or a custom-defined area) is crafted based on delineated boundaries or coordinates. This map acts as the foundational layer upon which the heatmap is overlaid.

5) **Coordinate Integration:** For the constructed map, the corner coordinates are ascertained, facilitating the precise overlay of VPI data points onto the map.

Heatmap Coloring: Each VPI data point is depicted on the map using a color spectrum, transitioning from light green (indicating a VPI close to 0) to deep maroon (indicating a VPI nearing 1). The gradation in color intensity offers a visual cue about the severity of visual pollution in different regions.When the VPI is visualized for a limited set of images covering a smaller area, the resultant heatmap might appear more localized and detailed as depicted in Figure 3. Conversely, with a denser array of data points covering a broader region, the heatmap manifests as a more continuous and expansive visual representation as depicted in Figure 4.

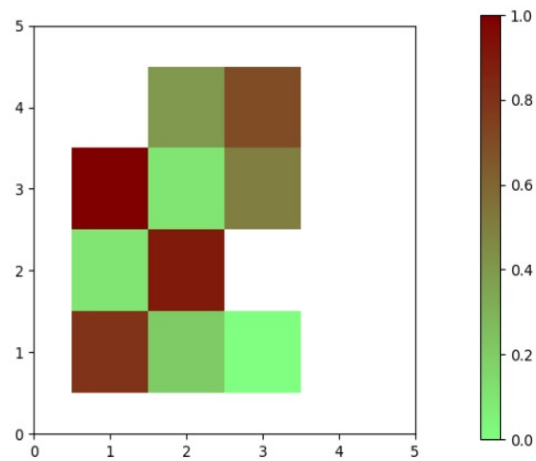

Fig. 3 Pixelated heatmap of VPI for a limited dataset. (Color bar represents VPI values)

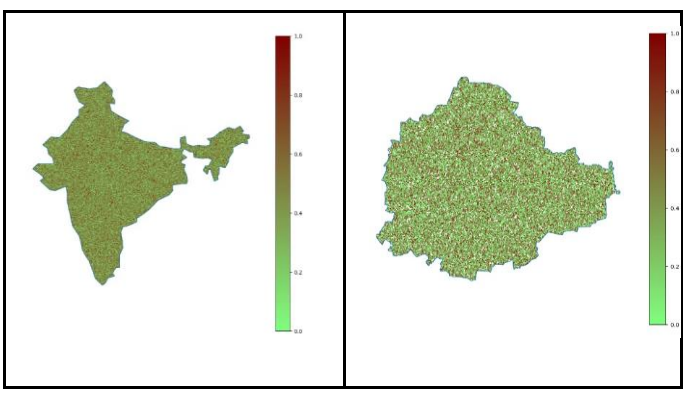

Fig. 4 Artist's illustrations of potential VPI heatmaps for India and Bengaluru.

5 EXPERIMENTAL ANALYSIS AND ON-SITE VALIDATION

5.1 Object Detection Techniques and Evaluation.

In this research, we implemented cutting-edge object recognition methods to detect and categorize building materials from images taken by drones. Specifically, we made use of the YOLOv6 [5] framework for detecting all types of materials. This model has shown outstanding results in object recognition across multiple fields.

Evaluation Criterion - Intersection over Union (IoU):

Intersection over Union (IoU) is a renowned standard for gauging the precision of object identification algorithms [26]. This metric evaluates the congruence between the model's predicted area (P) and the actual object's ground truth area (G). The equation to compute IoU is:

$$IOU = \frac{Area\ of\ Overlap\ (P \cap G)}{Area\ of\ Union\ (P \cup G)}$$

Where:

- $(P \cap G)$ represents the overlapping area between the predicted and ground truth bounding boxes.
- $(P \cup G)$ represents the combined area of both bounding boxes.

Elevated IoU scores suggest superior model accuracy. A score of 1 denotes an ideal match, while a score of 0 signifies no correspondence. We have achieved IoU values of 0.83 for **Graffiti**, 0.78 for **Faded Signage**, 0.88 for **Potholes**, 0.92 for **Garbage**, 0.82 for **Construction Road**, 0.71 for **Broken Signage**, 0.79 for **Bad streetlight**, 0.77 for **Bad Billboard**, 0.89 for **Sand on Road**, 0.78 for **Cluttered Sidewalk**, and 0.79 for **Unkempt Facade**. Overall, our models posted an average IoU rating of 0.81, which suggests a praiseworthy degree of precision in identifying building materials from the visuals.

5.2 PHYSICAL SURVEY OF VPI.

Elevated Pollution Zones: Upon analyzing several drone-acquired images of a specific area, Figure 5 presents a subset that represents the zone. The overall VPI for the entire area, based on all images including those not depicted in Figure 5, exceeded 0.7, signaling pronounced visual pollution. On-site evaluations and discussions with residents confirmed the visual disturbances, with many lamenting the area's "displeasing aesthetic."

Minimal Pollution Zones: In a contrasting scenario, we identified mere instances of visual pollution in a specific area, resulting in a VPI below 0.2, as depicted in Figure 6. To substantiate our model's observations, an on-ground survey was executed. Beyond the physical manifestations, interactions with local residents provided an insightful perspective. Most inhabitants expressed a sense of pride in their surroundings, often describing the area as "aesthetically pleasing" and "refreshingly clear." Such feedback underscores the low visual disruption observed in our VPI findings.

Fig. 5 Aerial images of an elevated pollution zone chosen for physical survey.

Fig. 6 Aerial images of some minimal pollution zones chosen for physical survey

6 COMPARATIVE ANALYSIS.

The quantification and visualization of visual pollution in urban landscapes using aerial imagery and computational methods have garnered attention in the urban planning and environmental aesthetics domain. Several studies have delved into understanding the impact of visual pollutants on urban aesthetics and the well-being of residents. Here, we present a comparative analysis of our work with notable contributions in the literature:

1. **Chmielewski, S. (2020):** Chmielewski's research titled "Chaos in Motion: Measuring Visual Pollution with Tangential View Landscape Metrics" delved into the concept of visual pollution (VP) in the form of outdoor advertisements (OA) as a threat to landscape physiognomy. The study proposed a methodological framework for measuring VP using tangential view landscape metrics, backed by statistically significant proofs. The research utilized raster products derived from aerial laser scanning data to characterize areas in Lublin, East Poland. The study highlighted the lack of consensus on the definition of VP and the need for a quantified approach to address this challenge [27].
2. **Zaeimdar, M., Khalilnezhad Sarab, F., & Rafati, M. (2019):** This study titled "Investigation of the relation between visual pollution and citizenry health in the city of Tehran" focused on the impact of visual contamination on the health of citizens in two urban areas of Tehran.

The research revealed a significant relationship between visual contamination and various health indicators of citizens, including physical signs, social function, anxiety, insomnia, and depression [28].

3. **Nahian Ahmed, M. Nazmul Islam, Ahmad Saraf Tuba, M.R.C. Mahdy, Mohammad Sujauddin (2019):** In their study titled "Solving visual pollution with deep learning: A new nexus in environmental management," published in the Journal of Environmental Management, the authors introduced methods for detecting visual pollution using deep learning [7]. They emphasized the potential of their study to be used in the future to design a visual pollution index or metric. Their research highlighted the need for automated visual pollutant classification and showcased the applicability of deep learning in achieving this. While they proposed the potential to develop such a metric, they did not actualize it. Our study has taken this forward by introducing the Visual Pollution Index (VPI), thereby realizing the potential they identified and providing a tangible metric for quantifying visual pollution in urban landscapes.

Our Contribution:

Our work distinguishes itself in several key aspects:

- We have introduced a comprehensive Visual Pollution Index (VPI) that not only detects visual pollution but also quantifies it, filling the gap identified in Chmielewski's study.
- Our approach is holistic, considering both the aesthetic and health impacts of visual pollution on urban populations, building on the findings of Zaeimdar and colleagues.
- Our unique dataset, combined with advanced computational methods, allows for a more granular and accurate assessment of visual pollution in various urban settings.

In conclusion, while several studies have delved into the concept and impacts of visual pollution, our approach offers a comprehensive, quantifiable, and actionable perspective. By integrating diverse methodologies and insights from previous research, our study promises to significantly advance the field of visual pollution assessment and mitigation.

7 USE CASES.

1. Urban Aesthetics and Revitalization:

- Heatmap Insights: Urban designers can harness the power of the generated heatmaps to discern visual pollution intensity across varied locales. This empowers them to pinpoint areas needing aesthetic enhancements or rejuvenation.

- Beautification Initiatives: Recognizing regions with pronounced visual pollution can guide city planners in orchestrating targeted beautification drives, green initiatives, and public art installations.

2. Governmental Supervision and Policy Formulation:

- Regulation and Enforcement: Authorities can leverage the VPI to ensure adherence to urban aesthetic standards, initiating corrective measures in areas with high visual pollution.

- Public Awareness and Campaigns: Governments can use VPI data to launch awareness campaigns, educating citizens about the importance of maintaining visual aesthetics and the role they can play.

3. Real-time Urban Monitoring for Community Stakeholders:
- Community Engagement: Local communities can utilize the VPI to monitor the visual health of their neighborhoods, rallying together for cleanup drives or community beautification projects.

- Feedback Mechanism: With the VPI as a reference, residents can provide feedback to municipal bodies about specific areas of concern, ensuring a collaborative approach to urban aesthetics.

4. Data-Informed Strategies for Investors and Businesses:
- Location Decisions: Entrepreneurs and investors can consult the VPI and associated heatmaps to determine suitable locations for new ventures, especially in sectors like hospitality or real estate, where aesthetics play a pivotal role.

- Market Analysis: Businesses can use VPI data to gauge the visual appeal of areas, helping them tailor marketing strategies or product placements in regions that align with their brand image.

8 CONCLUSION

In this research, we have unveiled a pioneering methodology to assess visual pollution in urban environments by harnessing aerial imagery and cutting-edge object detection mechanisms. Our Visual Pollution Index (VPI) emerges as a holistic metric, encapsulating both the variety and gravity of visual pollutants. With the support of our distinctive dataset and severity categorization, our results highlight the transformative potential of technology in reshaping our understanding of urban aesthetics. The congruence between our system's evaluations and real-world perceptions attests to the precision and dependability of our methodology. As cities worldwide grapple with the challenges of urbanization, instruments like ours will be instrumental in guiding efforts towards creating visually harmonious urban landscapes. The horizon looks promising, with opportunities for future research to refine, enhance, and broaden the scope of our approach, weaving in real-time data streams and more intricate analytical tools.

9 FUTURE SCOPE

- Integration with Smart City Infrastructure: As urban landscapes evolve into smart cities, there's potential to integrate the VPI with sensors and cameras placed throughout the city. This would facilitate real-time monitoring of visual pollution, enabling swift interventions and continuous urban beautification efforts.
- Predictive Analysis for Urban Aesthetics: By harnessing the power of artificial intelligence and machine learning, future versions of this system could anticipate areas prone to visual pollution. This predictive capability could be based on urban growth patterns, historical data, and socio-economic factors, allowing for proactive measures.
- Virtual Reality and Augmented Reality Enhancements: The next frontier could involve the use of Virtual Reality (VR) to simulate the visual experience of different urban areas based on their VPI scores. Additionally, Augmented Reality (AR) can be employed to superimpose potential solutions or improvements on existing urban landscapes, providing stakeholders with a futuristic vision of possible enhancements.

- Adaptive Urban Design Frameworks: With the insights derived from VPI, urban designers can develop adaptive design frameworks. These would be dynamic urban design strategies that evolve based on real-time VPI data, ensuring cities remain visually appealing amidst rapid urbanization.

ACKNOWLEDGMENT: We extend our heartfelt gratitude to Kesowa Infinite Ventures Pvt Ltd for their diligent efforts in collecting the invaluable dataset used in this study. Their commitment to quality and precision has been instrumental in the success of our research. We are deeply appreciative of their generosity in granting us exclusive access to this data and allowing its use for our study.

REFERENCES

[1] K. Lynch, "The Image of the City," MIT Press.

[2] J. L. Nasar, "Urban Design Aesthetics: The Evaluative Qualities of Building Exteriors," Environment and Behavior, vol. 26, no. 3, pp. 377–401, 1994

[3] R. S. Ulrich, "View through a window may influence recovery from surgery," Science, vol. 224, no. 4647, pp. 420-1, Apr. 1984.

[4] K. Anderson and K. J. Gaston, "Lightweight unmanned aerial vehicles will revolutionize spatial ecology," Frontiers in Ecology and the Environment, vol. 11, no. 3, pp. 138-146, Apr. 2013.

[5] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W. Nie, Y. Li, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei, and X. Wei, "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications," arXiv preprint arXiv:2209.02976, 7 Sep. 2022.

[6] S. Chmielewski, D. J. Lee, P. Tompalski, T. J. Chmielewski, and P. Wężyk, "Measuring visual pollution by outdoor advertisements in an urban street using intervisibilty analysis and public surveys," International Journal of Geographical Information Science, vol. 30, no. 4, pp. 801-818, 2016.

[7] N. Ahmed, M. N. Islam, A. S. Tuba, M.R.C. Mahdy, and M. Sujauddin, "Solving visual pollution with deep learning: A new nexus in environmental management," Journal of Environmental Management, vol. 248, 2019.

[8] K. Wakil, M.A. Naeem, G.A. Anjum, A. Waheed, M.J. Thaheem, M.Q. Hussnain, and R. Nawaz, "A Hybrid Tool for Visual Pollution Assessment in Urban Environments," Sustainability, vol. 11, 2019.

[9] O. A. Alharbi and N. Rangel-Buitrago, "Scenery evaluation as a tool for the determination of visual pollution in coastal environments: The Rabigh coastline, Kingdom of Saudi Arabia as a study case," Marine Pollution Bulletin, vol. 181, 2022.

[10] R. Madleňák and M. Hudák, "The Research of Visual Pollution of Road Infrastructure in Slovakia," in Challenge of Transport Telematics. TST 2016. Communications in Computer and Information Science, vol. 640, J. Mikulski, Ed. Cham: Springer, 2016.

[11] D. Yilmaz and A. Sagsöz, "In the context of visual pollution: Effects to Trabzon city Center Silhoutte," Asian Social Science, vol. 7, no. 5, pp. 98, 2011.

[12] Mitra, D. and Banerji, S., "Urbanisation and changing waterscapes: A case study of New Town, Kolkata, West Bengal, India," Applied Geography, vol. 97, pp. 109-118, 2018.

[13] https://www.qgis.org/en/site/about/index.html

[14] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J. Big Data, vol. 6, no. 1, p. 60, 2019.

[15] J. S. Armstrong, "The Graffiti Problem," General Economics and Teaching, University Library of Munich, Germany, 2004.

[16] N. Wardhana and E. Ellisa, "Youth tactics of urban space appropriation: case study of skateboarding and graffiti," Journal of Asian Architecture and Building Engineering, 2023.

[17] V. Stabnikov, J. Chu, A. N. Myo, et al., "Immobilization of Sand Dust and Associated Pollutants Using Bioaggregation," Water Air Soil Pollut, vol. 224, 2013.

[18] E. B. Oladumiye, "Urban Environmental Graphics: Impact, Problems and Visual Pollution of Signs and Billboards in Nigeria," International Journal of Education and Research, vol. 1, p. 89, 2013.

[19] S. Sarkar, "QUALITATIVE EVALUATION OF COMFORT NEEDS IN URBAN WALKWAYS IN MAJOR ACTIVITY CENTERS," Transportation Quarterly, vol. 57, no. 4, pp. 39-59, 2003.

[20] T. Ścieżor, “Effect of Street Lighting on the Urban and Rural Night-Time Radiance and the Brightness of the Night Sky,” Remote Sensing, vol. 13, no. 9, p. 1654, Apr. 2021, doi: 10.3390/rs13091654.

[21] S. Andjarsari, A. T. Subadyo, and N. Bonifacius, "Safe Construction And Visual Pollution Of Billboards Along Main Street," in IOP Conf. Ser.: Earth Environ. Sci., vol. 999, no. 1, p. 012015, 2022.

[22] Z. Q. Lafta and M. M. H. Al-Shamry, "The Spatial Variation of the Aspects of Environmental Pollution in Al-Hilla City and its Environmental Effects," Baltic Journal of Law & Politics, vol. 15, no. 2, 2022.

[23] Naveen, N., S. Mallesh Yadav, and A. Sontosh Kumar. "A study on potholes and its effects on vehicular traffic." Int. J. Creat. Res. Thoughts 6.1 (2018): 2320-2882.

[24] A. Latypova, A. Lenkevich, D. Kolesnikova, and K. Ocheretyany, “Study of Visual Garbage as Visual Ecology Perspective”, gmd, vol. 4, no. 2, pp. 153-172, Jun. 2022.

[25] https://geopandas.org/en/stable/

[26] M. Everingham et al., "The PASCAL Visual Object Classes (VOC) Challenge," International Journal of Computer Vision, vol. 88, no. 2, pp. 303–338, 2010.

[27] S. Chmielewski, “Chaos in Motion: Measuring Visual Pollution with Tangential View Landscape Metrics,” Land, vol. 9, no. 12, p. 515, Dec. 2020, doi: 10.3390/land9120515.

[28] M. Zaeimdar, F. Khalilnezhad Sarab, and M. Rafati, "Investigation of the relation between visual pollution and citizenry health in the city of Tehran (case study: municipality districts No.1 & 12 of Tehran)," Anthropogenic Pollution, vol. 3, no. 1, pp. 1-10, 2019