

Mapping Urban Visual Pollution at Global Scale Using Street-View Imagery and Deep Learning

BMVC 2026 Submission # 1444

Abstract

Visual pollution — encompassing man-made elements such as advertisements, infrastructure, and urban clutter — degrades perceived environmental quality and adversely affects human wellbeing. We propose a scalable framework for detecting and quantifying visual pollution from street-view imagery using deep-learning-driven object detection. To enable analysis at global scale, we curate a dataset of over two million images spanning 179 countries, drawn from multiple geospatial sources. We define eight visual pollutant categories and an auxiliary road-sign class, introducing finer-grained distinctions beyond prior work — notably separating shop signage and mobile advertisements from billboards. We train a YOLO-based detector on a manually annotated subset of this corpus to localise pollutant instances across diverse geographic and cultural contexts. We validate the framework through a series of experiments: measuring the regulatory impact of outdoor advertising bans in São Paulo, Brazil and selected Polish municipalities; and analysing land-use-driven variation across the city of Edinburgh, UK. In all settings, we observe interpretable spatial and temporal patterns in visual pollution, demonstrating the potential of this pipeline as a practical, scalable tool for urban environmental assessment.

1 Motivation

Visual pollution refers to the degradation of a visual environment by man-made elements such as billboards, wires, poles, and litter [19]. Although it does not have the same direct environmental or physical health impacts as other forms of pollution, it has been shown to negatively affect well-being, perceived livability, and mental health [14]. Some governments have recognised this issue and introduced regulatory measures. For example, São Paulo implemented the *Clean City Law*, which banned most forms of outdoor advertising [23], while Poland introduced the *Landscape Act*, enabling municipalities to regulate signage and advertisements [9]. Despite these efforts, visual pollution remains relatively under-addressed globally.

The term *visual pollution* originally referred to the visible presence of polluting substances e.g. exposed sewage pipes, garbage [12]. In more recent decades, however, the term has come to denote the diminishing of the visual or aesthetic quality of a scene by unappealing or intrusive man-made objects [14, 27, 29], moving from a primarily environmental concern to a more aesthetic one. Many researchers interpret the visual quality of a scene as

tied to its degree of order or coherence. Weber’s work on architectural aesthetics emphasises that spatial regularity, symmetry, alignment, and visual continuity enhance aesthetic judgment [43]. Scenes exhibiting high compositional harmony are generally perceived as of higher visual quality. In contrast, scenes lacking a governing structure or organising principle tend to appear chaotic or cluttered, yielding lower perceived quality [24]. From this perspective, visual pollution may be conceived as visual disorder, excessive contrast, incongruity, or unpredictability that undermines perceptual harmony.

Visual pollution has been linked to negative mental health outcomes, with cluttered and chaotic scenes contributing to stress and reduced comfort in urban spaces [12]. In addition, certain forms of visual clutter, such as billboards, have been associated with distracted driving and increased safety risks [11]. Early regulations, such as the *Bonus Act of 1958*, offered financial incentives to U.S. states to control outdoor advertising within 660 feet (200 metres) of interstate highways [28], and later legislation, most notably the *Highway Beautification Act of 1965*, required states to establish and enforce standards for billboard size, spacing, and lighting [13, 52]. Together, these acts reflect governmental awareness of the consequences of unchecked visual advertising. Similar concerns have recently been documented in the developing nations, where rapid urbanisation and unchecked outdoor advertising have contributed to visual pollution; for example, in Addis Ababa, Ethiopia, where 72% of surveyed local residents perceived outdoor advertisements as diminishing the city’s visual appeal [15]. Billboards and advertisements are not the only visual pollutants to consider. Overhanging telephone and utility wires may be less of an issue in developed countries, where many have been placed underground, but in developing countries such as India, Pakistan, and Bangladesh, they remain a significant source of visual clutter [11, 14]. In Indian case studies, utility wiring is often found alongside hoardings and unplanned signage, constituting a major visual pollutant [16]. Earlier work on overhead wires also regards them as a long-recognised source of aesthetic disturbance [53], and more recent environmental studies have quantified their negative visual impacts in both urban and rural landscapes [5].

Efforts to assess visual pollution face persistent challenges, mainly because visual perception is inherently subjective and shaped by individual experiences and cultural context [11, 29, 40]. Nevertheless, reliable assessment requires some degree of quantification to minimise subjectivity and enable consistent comparison across studies. To this end, researchers have employed a wide range of approaches: from human judgement and expert evaluation [11, 40, 41] to computer-based analysis and automated detection [11, 57], to evaluate and quantify visual pollution. For example, Ahmad et al. [11] built a custom CNN to classify four pollutant categories in Bangladesh, on a smaller 800 image dataset. Tasnim et al. (2023) [56] compared EfficientDet, Faster R-CNN, and YOLOv5 to classify textile-specific visual pollutants in Bangladesh, on a dataset of around 2,663 images comprising of images from web-crawling and manual collection from factories. Gupta et al. [17] used a YOLOv6-based Visual Pollution Index on drone imagery from Kolkata, but remains constrained by localised data collection and a limited category set. Almalki and Algethami (2025) [8] combined an ensemble of YOLOv5 and EfficientDet with BLIP-2 captioning, to produce textual descriptions of detected pollutants to reduce manual reporting errors. Rahman et al. (2026) [81] in their scoping review of deep learning architectures used for visual pollution detection and monitoring, highlights the issue of localization of datasets and geographic bias in existing work. They propose a multi-layered real-time monitoring framework that integrates detection, segmentation, and classification, drawing on diverse data sources such as street-view imagery, CCTV, and crowdsourced captures. See Table 1.

Although existing research has attempted to detect and quantify visual pollution, sev-

eral limitations persist including reliance on localised datasets, manual or semi-automated quantification methods, and a narrow focus on a limited set of pollutant categories making large-scale, globally representative analysis challenging. We aim to further advance the field with four contributions as follows. First, we construct a large-scale global street-view dataset spanning 179 countries, enabling the study of visual pollution across diverse geographic and cultural contexts. Second, we extend the scope of visual pollutants to eight classes: billboards, graffiti, utility poles, bins, potholes, barriers, mobile advertisements, and shop signs. An auxiliary road sign class is included to disambiguate billboard detections. Third, we develop a detection model trained on a geographically diverse annotated subset of sourced imagery, with the aim of improving robustness and generalisation performance across diverse environments. Fourth, we explore the relationship between estimated visual pollution level with policy changes, and land use for validation.

Table 1: Summary of computer vision literature on visual pollution estimation.

Paper	Country	CV model(s)	Labelled data	Fine-tuned?	Pollutant types covered									
					B	W	T	L	G	P	S	F	C	X
[1]	BGD	Custom CNN (5 layers)	~800	No	✓	✓	✓	✓	—	—	—	—	—	—
[2]	BGD	FRCNN, YOLOv5, EfficientDet	~2,663	Yes	✓	—	—	—	—	—	—	—	—	✓
[3]	BGD	YOLOv5/7, Faster SegFormer, EfficientDet	Custom	Yes	✓	✓	✓	✓	—	—	—	—	✓	✓
[4]	KSA	YOLOv5/7, FRCNN, MobileNetSSDv2, EfficientDet, Detectron2	34,460	Yes	—	—	—	—	—	✓	—	✓	✓	—
[5]	KSA	Dataset paper	34,460	—	—	—	—	—	—	✓	—	✓	✓	—
[6]	KSA	YOLOv5 + BiLSTM	Kaggle	Yes	—	—	—	✓	✓	✓	✓	✓	✓	—
[7]	KSA	YOLOv5 + EfficientDet + BLIP-2 captioning	~44,426	Yes	✓	—	—	✓	✓	✓	✓	✓	✓	—
[8]	IND	YOLOv6	Custom	Yes	✓	—	—	✓	✓	✓	✓	✓	—	—

✓ = pollutant type covered — = not addressed NR = not reported.

Categories: B: Billboards, W: Comm. wires/cables, T: Comm. towers/poles, L: Street litter/waste, G: Graffiti, P: Potholes/road defects, S: Broken/faded signage, F: Cluttered sidewalk/facade, C: Construction barriers, X: Textile/industrial waste.

Countries: BGD: Bangladesh, KSA: Saudi Arabia, IND: India.

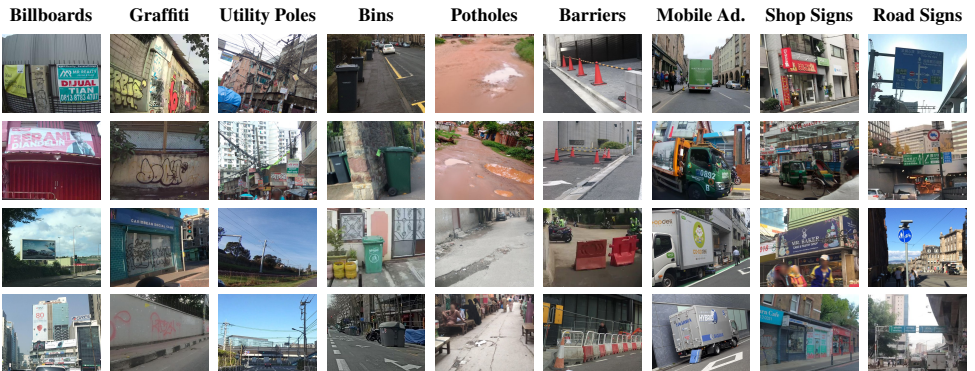


Figure 1: Example images for each of the nine urban street-scene classes used in this study.

2 Methodology

Data Sources We used three primary data sources: MAPILLARY, KARTAVIEW, and OPEN-STREETMAP. MAPILLARY and KARTAVIEW provide street-level imagery captured by con-

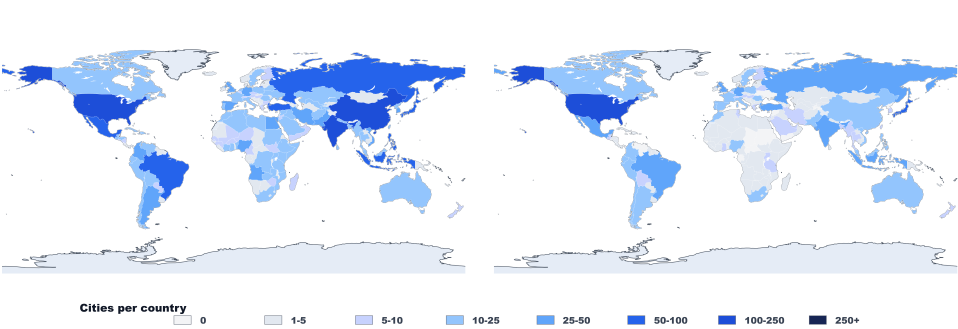


Figure 2: (Left) Number of selected cities per country with population above 100,000. (Right) Number of cities per country with at least 300 available images.

tributors worldwide, whereas OPENSTREETMAP provides geographic information on infrastructure and objects in the environment. **MAPILLARY** is a crowdsourced street-view imagery platform operated by META. Users upload geo-tagged images captured using smartphones, dash cameras, or other recording devices. These images are then made publicly available through an API. **KARTAVIEW**, formerly known as OPENSTREETCAM, is another crowdsourced street-view imagery platform operated by GRAB HOLDINGS, and provides street-level imagery uploaded by contributors worldwide. Both MAPILLARY and KARTAVIEW ensure privacy is protected by blurring faces and license plates in their images [24, 25]. **OPENSTREETMAP** is an open, collaboratively maintained geospatial database operated by the OPENSTREETMAP FOUNDATION. It provides structured geographic data describing roads, infrastructure, and points of interest contributed by volunteers worldwide. The selected features included bins, billboards, power infrastructure, barriers, and traffic signs [24, 25, 26, 27].

To achieve broad geographic coverage while keeping the dataset manageable, a global city dataset provided by SIMPLEMAPS [34] was used containing approximately 48,000 entries, including city–country pairs, geographic coordinates (latitude and longitude), and population estimates. For computational efficiency, we filtered to include only cities with populations exceeding 100,000 and available shapefiles for city boundaries, reducing the dataset to 5,414 cities. Figure 2 (left) illustrates the initial distribution of selected cities across countries. Images were annotated using LABEL STUDIO [38]. An initial YOLOv12 model was trained using a small subset of images to speed up the annotation process. Detections from this model were converted to Label Studio’s format and attached to the corresponding images, allowing them to be adjusted. To ensure geographic diversity, we attempted to sample 10 images per country for manual annotation where available and suitable, yielding 1,319 images. All annotations were manually reviewed and corrected before inclusion in the final training dataset.

Visual Pollutants Billboards and Mobile Advertisements Billboards are a well-known form of visual pollution and are among the most widely studied categories in the literature [4, 41]. We make a distinction between static billboards, such as Figure 1 (Billboards), and advertisements displayed on vehicles, such as Figure 1 (Mobile Ad.). This distinction is im-

portant because vehicle-based advertisements introduce a temporal component when using street-view imagery, and therefore do not reliably represent the persistent visual environment of a location. We place less emphasis on mobile advertisements as a visual pollutant (Table 2). **Graffiti** Graffiti has been widely identified as a form of visual pollution in existing studies [4, 19]. It is included due to its association with perceived disorder and reduced quality of life in urban environments. **Utility Poles** Utility poles are treated as an umbrella category, encompassing structures such as telephone poles, electricity poles, and pylons. These structures have been considered in previous studies of visual pollution [4, 53, 57], while other work has examined the visual impact of overhead wiring [40]. We focus on utility poles rather than wires, as poles are more consistently visible and easier to detect in street-view imagery. Street lamps and lighting infrastructure were not annotated in the training data as part of this category. **Bins** Although bins are necessary for waste management, they can contribute to visual pollution when poorly maintained, for example, when overfilled or placed carelessly. We have included all bins, while previous work has focused specifically on the impact of overflowing bins [10]. **Potholes** Potholes are a widely studied form of visual pollution [4, 8, 22] and are often associated with deteriorating road conditions and poor infrastructure. As with other visual pollutants, detecting potholes enables analysis of spatial patterns, co-occurrence with other pollutant types, such as graffiti, and associations with socioeconomic indicators such as urban deprivation. **Barriers** Barriers, such as concrete roadside barriers, have been considered in previous studies of visual pollution [8] and are prominently featured in existing datasets [2] In addition to permanent barriers, temporary structures, such as construction barriers and traffic cones, were included. **Shop signs** Shop sign detection has been explored in prior work outside the context of visual pollution [8, 22]. In this project, shop signs were included as a separate category primarily to maintain a clear distinction from billboards. **Road signs** Road signs are not considered visual pollutants in this project, as they serve essential functional roles in navigation and safety within urban environments. They are included to reduce conflation with billboards during detection.

Model A YOLO26 [20] model from the Ultralytics YOLO family was selected for object detection. YOLO-based architectures are widely used for visual pollution detection due to their strong performance in real-time object detection and a favourable balance between accuracy and computational efficiency [18]. Previous work has shown that YOLO frequently outperforms other common architectures, often with measurably greater accuracy [8, 57]. YOLO26, released in 2026, introduces architectural and optimisation improvements over earlier YOLO variants, with a particular emphasis on efficiency and small-object detection performance, directly relevant to visual pollutants, such as utility poles and distant signage, which often occupy only a small portion of the image. The medium-scale variant, yolo26m, was selected to balance detection accuracy (mAP50-95) against inference speed, as the dataset scale demands processing millions of images at reasonable computational cost.

Training was performed on a Google Compute Engine virtual machine of type g2-standard-4, which provides 4 vCPUs and 16 GB of system memory and an NVIDIA L4 GPU with 24 GB of GDDR6 memory. A train-validation-test split of 70%-20%-10% was used with stratified sampling. Training hyperparameters are shown in Supplementary Table 1. Training was conducted for 220 epochs to ensure sufficient model convergence. Performance metrics, including precision, recall, and mAP50, stabilised between epochs 180 and 190, indicating that the model had largely converged while avoiding premature termination. Input images were resized to a resolution of 1024×1024 to preserve fine-grained details

Pollutant	Severity	Detections	Cities	Countries	Avg. conf	Recall
Billboard	0.9	446,645	3,707	176	0.787	0.776
Graffiti	0.4	20,152	1,906	156	0.711	0.625
Utility Pole	0.3	1,136,380	3,770	176	0.776	0.780
Bin	0.2	13,587	1,791	145	0.701	0.739
Pothole	0.1	2,824	1,180	146	0.634	0.591
Barrier	0.1	42,748	2,486	163	0.753	0.780
Mobile Ad.	0.1	14,122	2,458	158	0.723	0.783
Shop Sign	0.1	78,880	3,074	171	0.707	0.688
Road Sign	0.0	592,092	3,581	176	0.796	0.841

Table 2: Severity weights and detection statistics by pollutant class. Bold values from [40].

present in street-view imagery. This is particularly important for small or thin objects, such as utility poles and distant road signs, which may be difficult to detect at lower resolutions. A batch size of 16 was chosen to balance computational efficiency and GPU memory constraints. The learning rate schedule was configured with an initial learning rate of 0.003 and a final scaling factor of 0.01, allowing for rapid initial learning followed by gradual convergence. A warmup period of 5 epochs was used to stabilise early training and prevent large gradient updates. Mosaic augmentation was applied with a probability of 0.5, and was disabled during the final 15 epochs to allow the model to fine-tune on more natural image distributions. Early stopping with a patience of 50 epochs was used to prevent overfitting while ensuring that training was not terminated prematurely. Finally, 8 data loader workers were used to maximise data throughput and minimise input bottlenecks during training.

Visual Pollution Index The VPI is computed at an image or city level as:

$$\begin{aligned} \text{CCR} &= \frac{\text{Number of Distinct Categories Detected}}{\text{Total Number of Categories}} \\ \text{SWS} &= \frac{\sum(\text{Severity Value} \times \text{Number of Instances for Each Category})}{\text{Total Number of Instances}} \\ \text{VPI} &= \text{CCR} \times \text{SWS} \end{aligned} \tag{1}$$

The scoring method is based on the Visual Pollution Index (VPI) proposed by Gupta et al. [40], with several severity weights derived from Wakil et al. [40]. It first calculates the Counting Categories Ratio (CCR), which measures the diversity of pollutant classes within an image. Images containing a more diverse range of pollutant types receive a higher CCR. The Severity Weighted Score (SWS) then applies class-specific severity weights to reflect the relative visual impact of each pollutant. For example, billboards are perceived as more visually polluting than potholes, so they contribute more heavily to the final score. Additional pollutant weights were assigned based on their perceived visual impact relative to other categories. Bins were assigned a slightly lower score than utility poles, as they are typically smaller and less visually dominant within a scene. Road signs were assigned a score of zero due to their functional necessity in urban environments. Shop signs and mobile advertisements were assigned low, non-zero values to account for both their occasional visual impact and minor model confusion with billboards. While the VPI system was intended for use on individual images, it can easily be extended to entire cities. Regions in the dataset contain varying numbers of images, leading to inconsistent sample sizes. As a result, regions with fewer images may yield unstable or unrepresentative scores, whereas regions with more im-



Figure 3: Examples of images in lowest (Q1) to highest (Q5) quantiles of VPIs.

ages provide more reliable estimates. To mitigate this, only cities containing more than 300 (2,300 cities) images were included in city-level comparisons (see Figure 2 (right)).

3 Results

Detection In total, 2,199,851 images were collected across 5,414 cities in 179 countries. Although most countries had at least some regions that returned images, the distribution was highly uneven. The majority of images originated from Western countries, with the United States alone contributing over 350,000 images. A small number of images were recorded with capture dates prior to 2010 (Supplementary Figure 1).

In total, 2,347,430 detections were produced. Of these, approximately 48.4% (1,136,380) correspond to the `utility_pole` class, followed by `road_sign`, with 592,092 detections. This distribution is expected, as both categories represent essential infrastructure that is prevalent in most urban environments. The next most frequent class is `billboard`, with 446,645 detections. The remaining classes appear significantly less frequently, indicating a clear class imbalance within the dataset. See Table 2. Figure 3 shows examples of images with different VPI scores.

OPENSTREETMAP To analyse how well the pipeline is able to capture visual pollutants, the detections are compared against OPENSTREETMAP features for alignment. In total, 6,546,984 OPENSTREETMAP features were collected. The majority of these were categorised under the `power` tag, which includes electrical infrastructure such as utility poles and related components. More than one million bin-related features were identified, supporting validation between image-based detections and map-based infrastructure data. Figure 4 compares the detections from street view images with reported objects in OPENSTREETMAP stratified by economic development. We observe similar trends in both cases, e.g., detection or presence of barrier and bins increases with socioeconomic status while utility poles decrease.

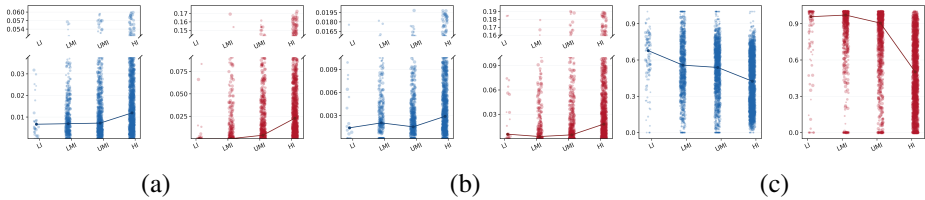


Figure 4: OSM and imagery rates: (a) Barrier, (b) Bin, (c) Utility pole.

Temporality To examine temporal patterns in detections, the detection rate and raw detection counts were analysed by time of day (Figure 5). As expected, images are captured predominantly during daytime. Detection rates for mobile advertisements, potholes, and road signs were comparatively stable throughout the day. However, detection rates for shop signs, billboards, utility poles, and barriers were higher at night relative to daytime, likely because artificial illumination makes these objects easier to detect in the image. Graffiti exhibited a similar pattern during the evening hours, although the mechanism is less clear.

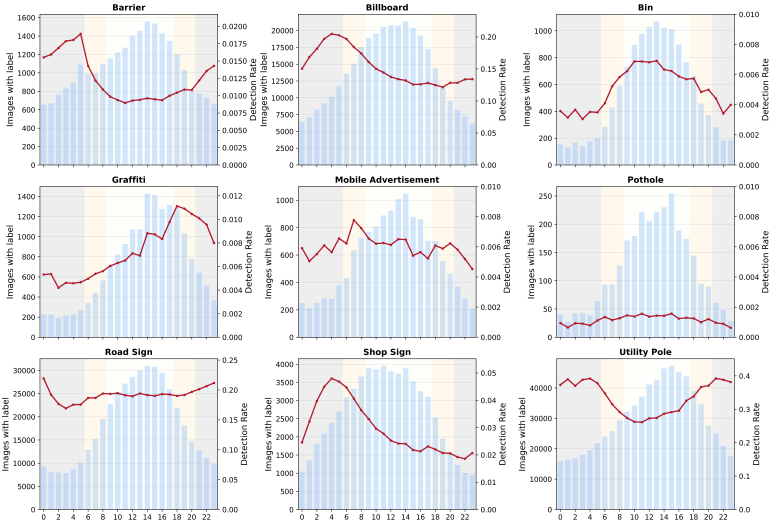


Figure 5: Detection rate and raw detection count by time of day

Visual Pollutants in Edinburgh To analyse how visual pollution varies across land-use types, the city of Edinburgh was selected as a case study. OPENSTREETMAP data was retrieved using the Python package `osmnx`, which provided polygon boundaries and associated land-use classifications for different areas of the city. Since most images were taken on the road rather than explicitly within a classified area, detections were assigned to the closest classified area. Five land-use categories were considered: commercial, residential, industrial, retail, and park/leisure/green. Figure 6 shows how pollutant detections vary across different land-use types in the city. It can be seen that billboards have the highest density in commercial (0.21) and retail (0.23) areas, and the lowest in residential areas (0.05). Furthermore, residential and park areas have a higher density of bins than other regions. This aligns with the expected distribution of urban infrastructure, providing qualitative validation

that the pipeline can capture meaningful variation in visual pollution across different urban areas.

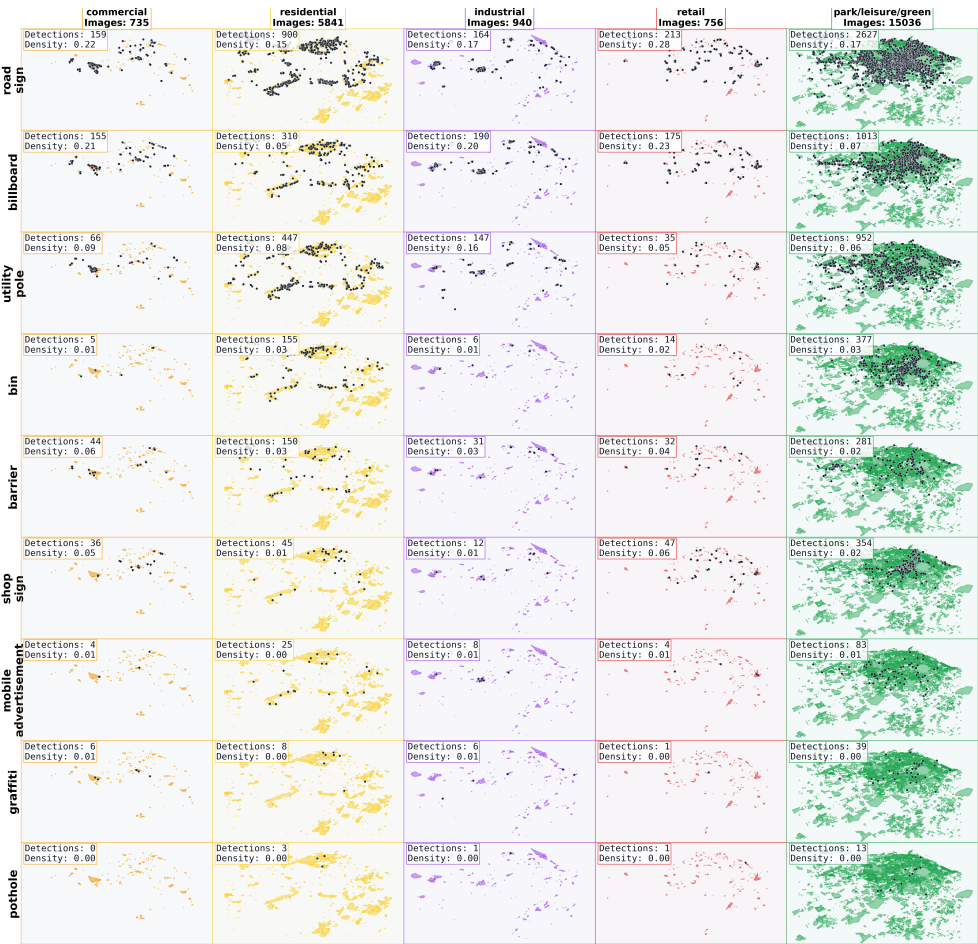


Figure 6: Detections in each land use type for each pollutant

Visual Pollution prevalence in United Kingdom The United Kingdom was selected for exploring visual pollution at a national level. A grid-fitting algorithm yielded 9,998 approximately uniform spatial cells, from which 10,000 points were sampled and queried against the MAPILLARY API to retrieve street-view imagery. This resolution ensured most cities contained multiple cells. Figure 7 presents the VPI calculated from these images, showing higher VPI values in the urban centres of the country, notably the areas in and around London, Birmingham, Manchester, and Leeds. This aligns well with the expectation that busier urban areas are more visually polluted, while less busy areas and protected spaces, such as the Scottish Highlands or Snowdonia tend to exhibit lower visual pollution.

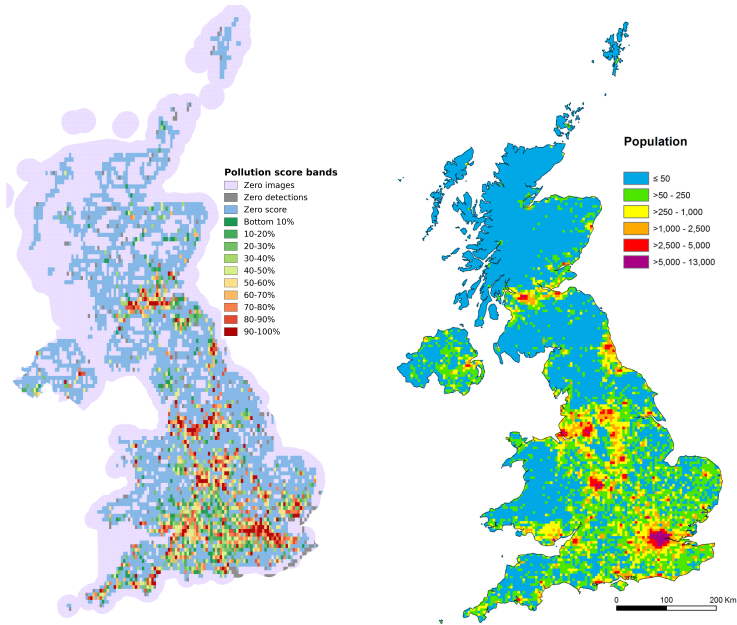


Figure 7: (Left) VPI in 9998 regions in UK. (Right) UK population density. [69]

Global prevalence of visual pollution index When looking at the visual pollution index globally, the highest-scoring cities are concentrated in South Asia, Southeast Asia, and parts of Africa (Figure 9 and Table 3). These regions are often associated with high urban density, rapid development, and less restrictive control of outdoor advertising and street infrastructure. As a result, scenes frequently contain a greater diversity and frequency of pollutant categories, which increases both the CCR and SWS components of the VPI. In contrast, the lowest-scoring cities are primarily located in Europe, North America, and East Asia. These regions typically benefit from stricter urban planning policies, more consistent infrastructure management, and regulation of signage and advertising. Consequently, both the diversity and frequency of detected pollutants are reduced, leading to lower VPI values. To explore this aspect further, we assessed the prevalence of visual pollution stratified by development indicators of Low Income, Lower Middle Income, Upper Middle Income, and High Income countries. Of the nine pollutant classes, the ones that showed the most change were billboards, road signs and utility poles. As shown in Figure 8, higher socioeconomic areas have lower rates of billboards and utility poles compared to lower socioeconomic areas. One reason for this could be stricter regulations on placement and style of billboards and underground power infrastructure, which requires fewer overground poles.

Evaluating Policy Changes Various countries and cities have implemented policies to combat visual pollution. This study attempts to assess their efficacy in Brazil and Poland.

São Paulo, Brazil The *Lei Cidade Limpa* (Clean City Law) implemented in 2006 in São Paulo provides a natural case study to evaluate the impact of regulation on visual pollution. As the law led to the removal of a large number of billboards, it is expected that São Paulo would exhibit a lower proportion of billboard detections than other major Brazilian cities.

Top 10% VPI			Bottom 10% VPI		
Country	Cities/Total	%	Country	Cities/Total	%
Indonesia	38/56	67.9	United States	64/323	19.8
India	32/54	59.3	Spain	17/50	34.0
Philippines	16/35	45.7	Japan	13/157	8.3
South Africa	12/33	36.4	Russia	13/82	15.9
Pakistan	8/29	27.6	Finland	8/8	100.0
Vietnam	6/23	26.1	Pakistan	6/29	20.7
Nigeria	6/15	40.0	Sweden	6/11	54.5
Thailand	5/10	50.0	Turkey	4/51	7.8
Peru	4/13	30.8	Italy	4/36	11.1
Ghana	4/7	57.1	Netherlands	4/26	15.4

Table 3: Countries of cities with top and bottom 10% of VPIs. For brevity, we show 10 countries in each group. “Cities/Total” shows grouped cities over total qualifying cities.

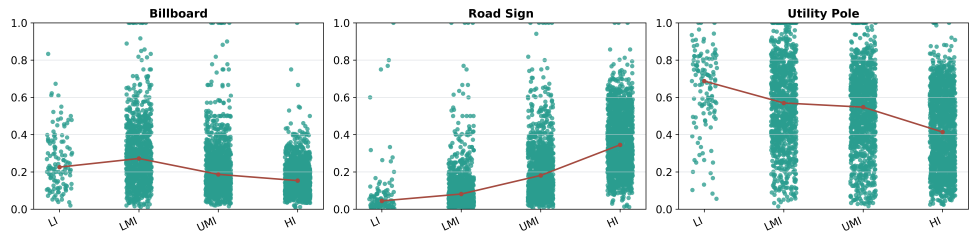


Figure 8: Detection rates for selected pollutants by economic classification.

To evaluate this, São Paulo was compared with Rio de Janeiro, Belo Horizonte, Fortaleza, Salvador, and Curitiba. The primary metric used was the Billboard Detection Rate, defined as the number of detected billboards divided by the total number of images within each city. This metric provides a simple proxy for the prevalence of large-scale advertising structures within each city. Differences in this ratio can indicate the effectiveness of regulatory interventions in reducing specific types of visual pollution. The results show a clear, consistent reduction in billboard density associated with regulation compared to Rio de Janeiro, Fortaleza, Salvador and Curitiba. Belo Horizonte exhibits an opposite pattern, displaying a lower billboard detection rate than São Paulo. This is likely attributable to Belo Horizonte’s own regulatory history: the city established outdoor advertising controls *Código de Posturas* (Law 8.616/2003), with further restrictions introduced in 2010 which prohibited billboards in parts of the city [50]. Combined with the city’s smaller scale and lower commercial density relative to São Paulo, these factors may explain why Belo Horizonte exhibits a comparatively lower billboard prevalence.

Poland In 2015, Poland introduced the *Landscape Act*, which enabled municipalities to regulate the aesthetics of public spaces, including advertising and signage [51]. This provides an opportunity to evaluate changes in visual pollution over time under varying regulatory conditions. As noted by Brecz-Kupiec (2024) [51], implementation of the act was decentralised, with municipalities introducing restrictions at different times, while some did not introduce regulations at all. This variation enables a natural comparison between pre-

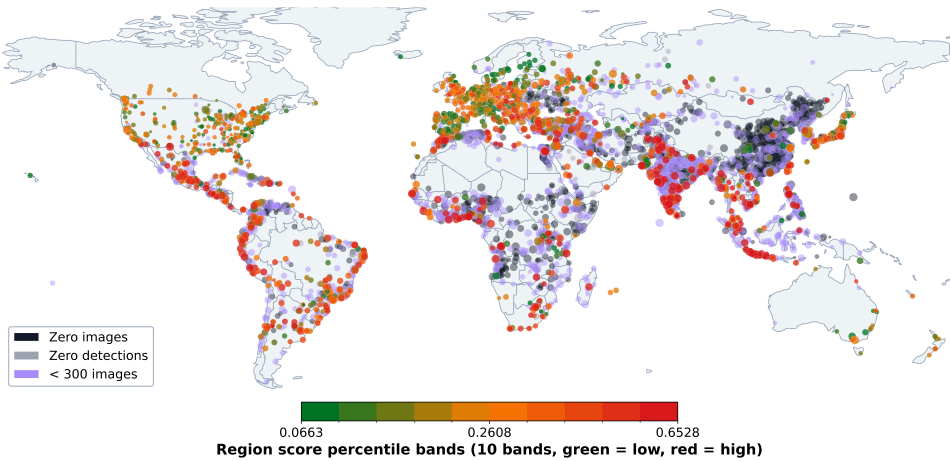


Figure 9: VPI scores across all regions. Dot size denotes population density.

City	Pop. Density (/km ²)	Images	Billboards	Detection Rate	Δ vs São Paulo
São Paulo	9,420	10,247	1,061	0.104	—
Rio de Janeiro	8,046	8,490	1,460	0.172	+66.09%
Fortaleza	8,556	9,888	1,636	0.165	+59.79%
Salvador	11,737	4,608	851	0.185	+78.36%
Belo Horizonte	7,216	12,261	1,117	0.091	-12.02%
Curitiba	4,472	19,255	2,527	0.131	+26.75%

Table 4: Comparison of billboard detection rates across selected Brazilian cities.

and post-regulation environments. Six cities were selected for analysis: Warsaw, Kraków, Gdańsk, Wrocław, Łódź, and Poznań. Łódź and Gdańsk introduced restrictions in 2016 and 2018, respectively, while Kraków and Wrocław implemented regulations in 2020; Poznań introduced them in 2023. Warsaw is included as a city with no restrictions to date. For each city, billboards were detected between 2014 and 2026. The analysis focuses on changes in billboard prevalence, measured using the Billboard Detection Rate defined previously. Figure 10 shows that detection rates generally decline over time, and remain lower after regulation. Table 5 shows the aggregated billboard detection rates before and after regulation. The results indicate a reduction in billboards following the introduction of advertising regulations across the five regulated cities. Warsaw, included as an unregulated comparison city, also shows a reduction, albeit smaller in magnitude, which may be consistent with the effectiveness of restrictions. Although this does not establish causality, the five regulated cities show substantial post-regulation reductions, ranging from 41.4% to 54.7%. The consistent reductions across all selected cities indicate that the approach can capture meaningful changes aligned with real-world regulatory interventions.

4 Discussion

Pollution is a major concern in modern society. Among its various forms—air, water, light, noise, and others—visual pollution remains comparatively underexplored, largely owing to

City	Regulation	Period	Images	Billboards	Detection Rate	Δ %
Warsaw	None	2014–2019	18,327	5,358	0.292	-21.36
		2020–2026	22,901	5,265	0.230	
Kraków	High	2014–2019	8,565	3,361	0.392	-54.60
		2020–2026	14,982	2,669	0.178	
Gdańsk	High	2014–2017	3,527	1,171	0.332	-54.69
		2018–2026	4,308	648	0.150	
Wrocław	Moderate	2014–2019	7,155	2,189	0.306	-41.38
		2020–2026	6,646	1,192	0.179	
Łódź	High	2014–2015	665	353	0.531	-53.55
		2016–2026	6,323	1,559	0.247	
Poznań	High	2014–2022	6,986	2,328	0.333	-42.08
		2023–2026	6,347	1,225	0.193	

Table 5: Billboard detection rates before and after regulation in selected Polish cities.

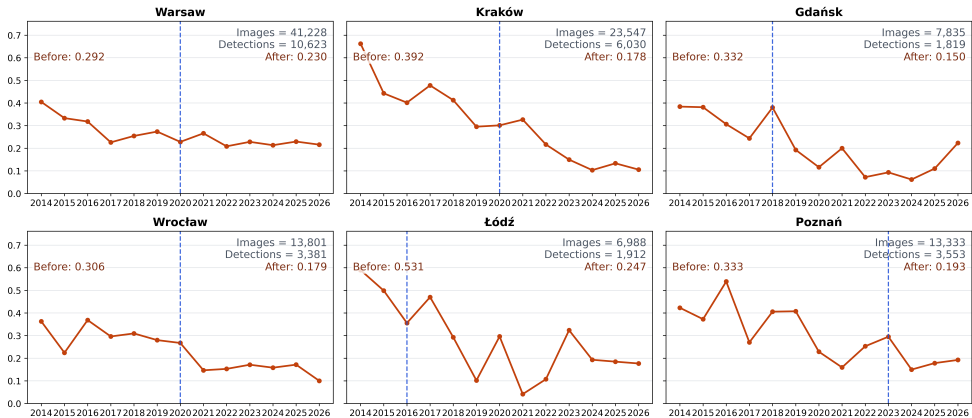


Figure 10: Billboard detection rates in cities before and after regulation.

its subjective nature. We contribute to this area by assessing the effectiveness of deep learning for detecting and quantifying visual pollution at a global scale, across both time and space. We evaluate the quality of our pipeline through both qualitative and quantitative analyses. We find that the detected visual pollutants align reasonably, though not perfectly, with reported points of interest in OPENSTREETMAP. This imperfect alignment can be attributed to inconsistencies in OPENSTREETMAP reporting, particularly in developing nations. Detection trends follow expected socioeconomic patterns — for example, a higher prevalence of utility poles in developing nations. The resulting visual pollution maps reveal meaningful global patterns, with visual pollution concentrated in busier areas. Finally, we observe encouraging temporal patterns before and after a policy change, demonstrating that the pipeline yields consistent results both spatially and temporally. Thus, the pipeline has several strengths, the most notable being its reliance on publicly available AI tools and a global dataset, which makes it both accessible and reproducible. This design also allows the pipeline to be extended to additional visual pollution classes, including those reflecting local

attributes and those in data-scarce regions.

A key limitation is the uneven distribution of data across regions: cities in India and China have low image availability, while Europe and North America exhibit substantially greater coverage, as seen in Figure 9. Additional limitations arise from the labelling process, e.g., through ambiguity in category definitions. The definition of a billboard was intentionally broad to capture a wider range of advertising formats, including posters and banners. While this improves coverage, it introduces inconsistency when comparing results against real-world regulations, where definitions are typically more precise. As a result, the detected “billboard” category does not always align directly with policy definitions. Labelling across a global dataset also introduces challenges related to language and cultural context. In non-English-speaking regions, it is difficult to accurately interpret signage. Consequently, some shop signs may have been misclassified as billboards, also considering that visual styles vary significantly between countries. This introduces label noise, which may bias the model in specific regions, affecting both detection reliability and cross-regional comparisons. The labelled dataset also exhibits class imbalance, with some categories appearing far more frequently than others. This can lead to stronger performance on dominant classes and weaker performance on under-represented ones, reducing the reliability of the overall scoring.

Another limitation concerns image quality. There was no scalable mechanism to automatically verify that retrieved images were suitable for analysis. In some cases, images were blurred, poorly framed, or primarily comprised close-up views of surfaces such as walls and therefore uninformative. These cases represented a small proportion of the dataset and were therefore retained. We also note that street-view imagery may be captured at different times and dates. As a result, certain features such as billboards or bins may no longer be present or may have changed location since the image was recorded. Although timestamps were stored where available, temporal changes in infrastructure were not explicitly analysed in this work.

The proposed pipeline can be extended in various ways. The current VPI captures the presence of various pollutants but not how their spatial distribution within the image. The ‘entropy’ of the pollutants can be meaningful in capturing the nature of visual pollution. Additionally, the current pipeline is limited to a few visual pollutants while many others, e.g., dilapidated buildings or litter, can be integrated given suitable training data. Finally, the current study does not explore causality, for example, in assessing the effect of a policy change on visual pollution or effect of visual pollution, e.g., on mental health. Exploring this relationship would provide more actionable insight into the impact of visual pollution on our wellbeing and addressing this pollution in our community.

The proposed work has several research and policymaking implications. To the best of our knowledge, this is among the first attempts to capture visual pollution at a global scale, although the notion of visual pollution can itself vary across regions. The pipeline is equally applicable at a local level where assessing visual pollution can inform urban planning and potentially benefit mental health. That said, integrating this into policymaking requires careful attention to context. In particular, visual pollution scores do not directly correspond to aesthetic quality: regions such as deserts may receive low scores simply because few detectable pollutants are present, rather than because the environment is visually desirable. Even so, the framework enables meaningful comparisons between regions with similar characteristics, particularly where data coverage is comparable. In such cases, it can reveal which objects contribute most significantly to visual pollution across different urban environments.

References

[1] Nahian Ahmed, M. Nazmul Islam, Ahmad Saraf Tuba, M.R.C. Mahdy, and Mohammad Sujaudinn. Solving visual pollution with deep learning: A new nexus in environmental management. *Journal of Environmental Management*, 248: 109253, 2019. ISSN 0301-4797. doi: <https://doi.org/10.1016/j.jenvman.2019.07.024>. URL <https://www.sciencedirect.com/science/article/pii/S0301479719309557>.

[2] Mohammad AlElaiwi, Hafiz Ahmad, Jamil Hussain, Mugahed Al-antari, Badar Almarri, and Areeba Azhar. Saudi arabia public roads visual pollution dataset, 2023. URL <https://doi.org/10.17632/bb7b8vtwry.5>.

[3] Mohammad AlElaiwi, Mugahed A. Al-antari, Hafiz Farooq Ahmad, Areeba Azhar, Badar Almarri, and Jamil Hussain. Vpp: Visual pollution prediction framework based on a deep active learning approach using public road images. *Mathematics*, 11(1), 2023. ISSN 2227-7390. doi: 10.3390/math11010186. URL <https://www.mdpi.com/2227-7390/11/1/186>.

[4] Mohammad AlElaiwi, Mugahed A. Al-antari, Hafiz Farooq Ahmad, Areeba Azhar, Badar Almarri, and Jamil Hussain. Visual pollution real images benchmark dataset on the public roads. *Data in Brief*, 50:109491, 2023. ISSN 2352-3409. doi: <https://doi.org/10.1016/j.dib.2023.109491>. URL <https://www.sciencedirect.com/science/article/pii/S2352340923005917>.

[5] Haya Almalki and Nahlah Algethami. Ensemble deep learning with image captioning for visual pollution detection, classification, and reporting. *Scientific Reports*, 15(1): 31516, August 2025. ISSN 2045-2322. doi: 10.1038/s41598-025-17200-0. URL <https://www.nature.com/articles/s41598-025-17200-0>.

[6] Mrouj Almuhajri and Ching Y. Suen. Shop signboards detection using the shos dataset. In *Pattern Recognition and Artificial Intelligence: Third International Conference, ICPRAI 2022, Paris, France, June 1–3, 2022, Proceedings, Part II*, page 235–245, Berlin, Heidelberg, 2022. Springer-Verlag. ISBN 978-3-031-09281-7. doi: 10.1007/978-3-031-09282-4_20. URL https://doi.org/10.1007/978-3-031-09282-4_20.

[7] Javier Rozo Alzate, Marta S. Tabares, and Paola Vallejo. Graffiti and government in smart cities: a deep learning approach applied to medellín city, colombia. In *International Conference on Data Science, E-learning and Information Systems 2021, DATA'21*, page 160–165, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 978-1-4503-8838-2. doi: 10.1145/3460620.3460749. URL <https://dl.acm.org/doi/10.1145/3460620.3460749>.

[8] Ji-Won Baek and Kyungyong Chung. Pothole classification model using edge detection in road image. *Applied Sciences*, 10(19), 2020. ISSN 2076-3417. doi: 10.3390/app10196662. URL <https://www.mdpi.com/2076-3417/10/19/6662>.

[9] Sara Brecz-Kupiec. Evaluating a decade of landscape resolutions in poland: Progress, challenges, and effectiveness in public space management. *European Research Studies*, XXVII(3):888–907, December 2024. ISSN 1108-2976.

- [10] Szymon Chmielewski, Danbi J. Lee, Piotr Tompalski, Tadeusz J. Chmielewski, and Piotr Wężyk. Measuring visual pollution by outdoor advertisements in an urban street using intervisibility analysis and public surveys. *International Journal of Geographical Information Science*, 30(4):801–818, 2016. doi: 10.1080/13658816.2015.1104316. URL <https://doi.org/10.1080/13658816.2015.1104316>.
- [11] Szymon Chmielewski, Marta Samulowska, Michał Lupa, Danbi Lee, and Bogdan Zagajewski. Citizen science and webgis for outdoor advertisement visual pollution assessment. *Computers, Environment and Urban Systems*, 67:97–109, 2018. ISSN 0198-9715. doi: <https://doi.org/10.1016/j.compenvurbsys.2017.09.001>. URL <https://www.sciencedirect.com/science/article/pii/S019897151730234X>.
- [12] Oxford English Dictionary. “visual pollution”, n.d. URL https://www.oed.com/dictionary/visual-pollution_n. Accessed October 2025.
- [13] Federal Highway Administration. How the highway beautification act became a law, June 2017. URL <https://www.fhwa.dot.gov/infrastructure/beauty.cfm>. Accessed October 2025.
- [14] Hangyu Gao, Shamsul Abu Bakar, Suhardi Maulan, Mohd Johari Mohd Yusof, Riyadh Mundher, Yu Guo, and Benxue Chen. A systematic literature review and analysis of visual pollution. *Land*, 13(7), 2024. ISSN 2073-445X. doi: 10.3390/land13070994. URL <https://www.mdpi.com/2073-445X/13/7/994>.
- [15] Eshetu Gelan. Assessing visual pollution: The impact of urban outdoor advertisements in addis ababa, ethiopia. *Architecture*, 5(1), 2025. ISSN 2673-8945. doi: 10.3390/architecture5010009. URL <https://www.mdpi.com/2673-8945/5/1/9>.
- [16] Vasudha Gokhale. Examining impact of visual pollution on city environment: Case study of pune, india. *Spandrel Journal of School of Planning and architecture Bhopal*, 1:10, 01 2011.
- [17] Advait Gupta, Manan Padsala, Devesh Jani, Tanmay Bisen, Aastha Shayla, and Gargi Srivastava. *From Above and Beyond: Decoding Urban Aesthetics with the Visual Pollution Index*, page 101–119. Springer Nature Switzerland, Cham, 2024. ISBN 978-3-031-56388-1. doi: 10.1007/978-3-031-56388-1_8. URL https://doi.org/10.1007/978-3-031-56388-1_8.
- [18] Md. Yearat Hossain, Ifran Rahman Nijhum, Md. Tazin Morshed Shad, Abu Adnan Sadi, Md. Mahmudul Kabir Peyal, and Rashedur M. Rahman. An end-to-end pollution analysis and detection system using artificial intelligence and object detection algorithms. *Decision Analytics Journal*, 8:100283, 2023. ISSN 2772-6622. doi: 10.1016/j.dajour.2023.100283.
- [19] Chuanbo Hu, Shan Jia, Fan Zhang, Changjiang Xiao, Mindi Ruan, Jacob Thrasher, and Xin Li. Updexplainer: An interpretable transformer-based framework for urban physical disorder detection using street view imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 204:209–222, October 2023. ISSN 0924-2716. doi: 10.1016/j.isprsjprs.2023.08.017.

[20] Glenn Jocher and Jing Qiu. Ultralytics yolo26, 2026. URL <https://github.com/ultralytics/ultralytics>.

[21] KartaView. Privacy policy. <https://kartaview.org/privacy-policy>. Accessed: 2026-05-27.

[22] Christian Koch and Ioannis Brilakis. Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, 25(3):507–515, 2011. ISSN 1474-0346. doi: <https://doi.org/10.1016/j.aei.2011.01.002>. URL <https://www.sciencedirect.com/science/article/pii/S1474034611000036>. Special Section: Engineering informatics in port operations and logistics.

[23] Kurt Kohlstedt. Clean city law: Secrets of são paulo uncovered by outdoor advertising ban, February 2016. URL <https://99percentinvisible.org/article/clean-city-law-secrets-sao-paulo-uncovered-outdoor-advertising>. Accessed October 2025.

[24] J.T. Lang. *Creating Architectural Theory: The Role of the Behavioral Sciences in Environmental Design*. Role of the Behavioural Sciences in Environmental Design. Van Nostrand Reinhold Company, 1987. ISBN 9780442259815. URL <https://books.google.co.uk/books?id=lHlwQgAACAAJ>.

[25] Mapillary. Privacy. <https://help.mapillary.com/hc/en-us/articles/115001770349-Privacy>. Accessed: 2026-05-27.

[26] Alanoud Al Mazroa, Mohammed Maray, Abdullah M. Alashjaee, Faiz Abdullah Alotaibi, Ahmad A. Alzahrani, Abdulwhab Alkharashi, Shoayee Dlaim Alotaibi, and Mrim M. Alnfai. Computer Vision With Explainable Artificial Intelligence for Visual Pollution Detection in the Kingdom of Saudi Arabia. *IEEE Access*, 12:193014–193027, 2024. ISSN 2169-3536. doi: 10.1109/ACCESS.2024.3513696. URL <https://ieeexplore.ieee.org/document/10786207/>.

[27] John Copeland Nagle. Cell phone towers as visual pollution. *Notre Dame JL Ethics & Pub. Pol’y*, 23:537, 2009.

[28] Outdoor Advertising Association of America. Laws & regulations, n.d. URL <https://oaaa.org/policy-advocacy/general-information/laws-regulations>. Accessed October 2025.

[29] Adriana Portella. *Visual Pollution: Advertising, Signage and Environmental Quality*. Routledge, 1st edition, 2014. doi: 10.4324/9781315547954. URL <https://doi.org/10.4324/9781315547954>.

[30] Prefeitura Municipal de Belo Horizonte. Decreto n.º 14.060, de 6 de agosto de 2010, 2010. URL <https://www.legisweb.com.br/legislacao/?id=172609>. Regulamenta a Lei n.º 8.616, de 14 de julho de 2003, que contém o Código de Posturas do Município de Belo Horizonte.

[31] Mohammad Masudur Rahman, Md Rashedur Rahman, Ashraful Islam, Saadia B. Alam, and M. Ashraful Amin. A scoping review of deep learning for urban visual pollution and proposal of a real-time monitoring framework with a visual pollution index. (arXiv:2602.09446), February 2026. doi: 10.48550/arXiv.2602.09446. URL <http://arxiv.org/abs/2602.09446>. arXiv:2602.09446 [cs].

- [32] Scenic America. About the highway beautification act, n.d. URL <https://www.scenic.org/why-scenic-conservation/billboards-and-sign-control/hba/>. Accessed October 2025. 782 783 784 785
- [33] Robert J. Schodorf. A study of visual pollution from overhead wires and associated structures. Master's thesis, Western Michigan University, August 1973. URL https://scholarworks.wmich.edu/cgi/viewcontent.cgi?article=3701&context=masters_theses. Accessed October 2025. 786 787 788 789
- [34] Simplemaps. World cities database, 2026. URL <https://simplemaps.com/data/world-cities>. Accessed: 9 March 2026. 790 791 792
- [35] Andreas Sumper, Oriol Boix-Aragones, Joan Rull-Duran, Joan Amat-Algaba, and Joachim Wagner. Assessment of the visual impact of existing high-voltage lines in urban areas. *Electricity*, 2(3):285–299, 2021. ISSN 2673-4826. doi: 10.3390/electricity2030017. URL <https://www.mdpi.com/2673-4826/2/3/17>. 793 794 795 796
- [36] Najia Hasan Tasnim, Sadia Afrin, Barna Biswas, Arifa Akter Anye, and Riasat Khan. Automatic classification of textile visual pollutants using deep learning networks. *Alexandria Engineering Journal*, 62:391–402, 2023. ISSN 1110-0168. doi: <https://doi.org/10.1016/j.aej.2022.07.039>. URL <https://www.sciencedirect.com/science/article/pii/S111001682200494X>. 797 798 799 800 801
- [37] Md Fahim Shahoriar Titu, Abdul Aziz Chowdhury, S. M. Rezwanul Haque, and Riasat Khan. Deep-learning-based real-time visual pollution detection in urban and textile environments. *Sci*, 6(1), 2024. ISSN 2413-4155. doi: 10.3390/sci6010005. URL <https://www.mdpi.com/2413-4155/6/1/5>. 802 803 804 805 806
- [38] Maxim Tkachenko, Mikhail Malyuk, Andrey Holmanyuk, and Nikolai Liubimov. Label Studio: Data labeling software, 2020-2025. URL <https://github.com/HumanSignal/label-studio>. Open source software available from <https://github.com/HumanSignal/label-studio>. 807 808 809 810
- [39] M. Vieno, Mathew Heal, Martin Williams, Edward Carnell, J. Stedman, and Stefan Reis. Sensitivities of uk $\text{pm}_{2.5}$ concentrations to emissions reductions. *Atmospheric Chemistry and Physics Discussions*, 15:20881–20910, 08 2015. doi: 10.5194/acpd-15-20881-2015. 811 812 813 814 815
- [40] Khydija Wakil, Malik Asghar Naeem, Ghulam Abbas Anjum, Abdul Waheed, Muhammad Jamaluddin Thaheem, Muhammad Qadeer ul Hussnain, and Raheel Nawaz. A hybrid tool for visual pollution assessment in urban environments. *Sustainability*, 11(8), 2019. ISSN 2071-1050. doi: 10.3390/su11082211. URL <https://www.mdpi.com/2071-1050/11/8/2211>. 816 817 818 819 820
- [41] Khydija Wakil, Ali Tahir, Muhammad Qadeer ul Hussnain, Abdul Waheed, and Raheel Nawaz. Mitigating urban visual pollution through a multistakeholder spatial decision support system to optimize locational potential of billboards. *ISPRS International Journal of Geo-Information*, 10(2), 2021. ISSN 2220-9964. doi: 10.3390/ijgi10020060. URL <https://www.mdpi.com/2220-9964/10/2/60>. 821 822 823 824 825
- [42] Bin Wang and Li Jing. *First-place Solution for Streetscape Shop Sign Recognition Competition*. January 2025. doi: 10.48550/arXiv.2501.02811. 826 827

[43] Rolf Weber. *On the Aesthetics of Architecture: A Psychological Approach to the Structure and the Order of Perceived Architectural Space*. Routledge, 1st edition, 1995. doi: 10.4324/9781003502883. URL <https://doi.org/10.4324/9781003502883>.

[44] OpenStreetMap Wiki. Tag:amenity=waste basket — openstreetmap wiki,, 2025. URL https://wiki.openstreetmap.org/w/index.php?title=Tag:amenity%3Dwaste_basket&oldid=2900169. [Online; accessed 29-May-2026].

[45] OpenStreetMap Wiki. Key:power — openstreetmap wiki,, 2025. URL <https://wiki.openstreetmap.org/w/index.php?title=Key:power&oldid=2902447>. [Online; accessed 29-May-2026].

[46] OpenStreetMap Wiki. Key:advertising — openstreetmap wiki,, 2026. URL <https://wiki.openstreetmap.org/w/index.php?title=Key:advertising&oldid=3040446>. [Online; accessed 29-May-2026].

[47] OpenStreetMap Wiki. Key:barrier — openstreetmap wiki,, 2026. URL <https://wiki.openstreetmap.org/w/index.php?title=Key:barrier&oldid=2968970>. [Online; accessed 29-May-2026].