Graffiti and government in smart cities: a Deep Learning approach applied to Medellín City, Colombia

Javier Rozo Alzate*
Universidad EAFIT
Medellín, Colombia
jarozoa@eafit.edu.co

Marta S. Tabares
Universidad EAFIT
Medellín, Colombia
mtabares@eafit.edu.co

Paola Vallejo
Universidad EAFIT
Medellín, Colombia
pvallej3@eafit.edu.co

ABSTRACT

Graffiti is an element of graphic expression that manifests different states of the human being. However, for many governments worldwide, it has been an element of discord between them and the communities that express themselves through graffitis. This article proposes identifying graffiti and concentration zones through Computer Vision and object detection and localization to support public policy management in smart cities. ASUM-DM methodology is used to achieve the aim. Initially, the current problems faced by municipal governments in the management of public graffiti policy are identified. Then available datasets of images from Google Street View (GSV) and other acquired datasets are identified for the case study carried out in the city of Medellín (Colombia) and border municipalities. A training dataset of 1,395 images and a production dataset of 71,100 panoramas is placed on strictly using the experimental method of the division of training data, validation, and a production sample, to make a correct estimation of the generalization error. As a result of the training process, we obtained an Average Precision of 69,14%, which presented a high precision Tag of 89.23%, and low precision of 59.13% in Mural. Finally, it is possible to build heat maps of graffiti concentration areas that could guide rulers to create or improve public policies related to graffiti expression.

CCS CONCEPTS

• **Applied computing → Law, social and behavioral sciences**;
• **Computing methodologies** → *Computer vision.*

KEYWORDS

Deep Learning, Object Detection and Localization, Smart cities, Public policies, Graffiti.

ACM Reference Format:
Javier Rozo Alzate, Marta S. Tabares, and Paola Vallejo. 2018. Graffiti and government in smart cities: a Deep Learning approach applied to Medellín City, Colombia. In *ACM International Conference on Data Science, E-learning and Information Systems 2021, April 05–07, 2021, Petra, Jordan.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Around the world, some of the elements of urban tourism that becomes more relevant are the *Graffiti Tour* and *Street-Art Tours*[1]. Specifically, Colombian cities like Medellín and Bogotá have been changing Graffiti's view into the city: Graffiti Tour[2], Bogotá Graffiti Tour[3]. Nevertheless, graffiti is and will be a social element under discussion for a long time. It is because "writing on the walls" could be an artistic expression or human vent by nature. Its analysis has been done from different social perspectives such as culture, art, politic, human rights, and economics (e.g., tourism sector). Some approaches support graffiti as a positive expression for the city transformation [19] [26] while others show how might still it be a vandalic expression [23] [22].

In the last decade, the rulers of the city of Medellín have been strategic in transforming the violence left by drug trafficking, the guerrillas, and the displacement of villagers into several citizen opportunities. One of them is graffiti and street art as an expression of culture that provides a new perspective of life for the most vulnerable citizens, and an economic axis of growth for them and consequently for the city [25]. However, many efforts or challenges on that type of activities can be frustrated by the spontaneous "artists"' intervention in different city areas, whose authorship is unknown. It is not easy to identify the graffiti's location and much less the intensions with those graphic expressions. Consequently, this makes these "unofficial" staging areas difficult for local governments to manage. It could increase the likelihood of graffiti proliferation in unregulated spaces, particularly related to vandalism' acts or that "dirty" the city.

Nowadays, information technologies have begun to support this type of social concern from the perspective related to the treatment of images [18] [17]. Besides, Artificial Intelligence tools and techniques guide the analysis of information on countless problems. For instance, Gordienko *et al.* [8] train Machine Learning and Deep Learning models for the recognition of image characters. Munsberg *et al.* [15] validated the ability to use a convolutional classification network, trained on ImageNet to classify graffiti images taken from Flickr. Nahar *et al.* [16] proposed a tool based on an uncrewed aerial vehicle (UAV) to remove graffiti in a city. The detection is made from a convolutional network. In the same way, Wang *et al.* [27]

*All authors contributed equally to this research.

[1]https://www.bombingscience.com/graffiti-street-art-walking-tours-around-world/
[2]https://https://www.comuna13grafititour.com//
[3]http://bogotagraffiti.com/

DATA 21, April 05–07, 2021, Petra, Jordan

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

presented a UAV for cleaning and graffiti removal. For image prediction, a TensorFlow API-based model pre-trained with the COCO dataset is used.

In this article, we propose a Computer Vision model supported by transfer learning and object detection techniques. These use GSV images (Google Street View) and acquired datasets (i.e., a training dataset of 1,395 images and a production dataset of 71,100 panoramas) to recognize graffiti in Medellín city and some south border municipalities. We emphasize scenarios where the image contains additional elements to the graffiti (e.g., mailboxes, signs, walls, etc.), where predictive capacity is significantly affected.

Our proposal uses ASUM-DM methodology adapted by Angée *et al.* [1]. We emphasize the core activities of the stages Analyze-Design-Configure&Build and Deploy. Thus, we accomplish to be robust in the stages of training and evaluating the model. Initially, we identify current problems that must be faced by municipal governments in the management of the public graffiti policy (i.e., conflicts of interest that may arise between author, authority, and owner concerning property rights and public space others). Then, we define requirements related to identifying graffiti, zones (i.e., geographical area such as neighborhood, street) where they are concentrated, and the way to achieve high performance by applying object detection and localization technique.

To achieve that, we search available datasets for the case study in the city of Medellín (Colombia) and its south border municipalities like Itagüí, Envigado, and Sabaneta (i.e., "our focus area"). Using an experimental method to the division of training data, its validation, and the test sample production to find a correct estimation of the generalization error. As a result of the training process, an Average Precision of 69,14% is obtained. Finally, it provided visualization results that will be the input for public political's guidelines to help municipalities manage this social manifestation type.

The article is organized as follows. Section 2 presents the Literature Review related to the most relevant concepts and related work. Section 3 presents the proposal development step by step and the results obtained. Finally, section 4 concludes and presents future work.

2 LITERATURE REVIEW

Ulmer [26] mentions a formal meaning of Graffiti, which is "*Creating "tags" and "pieces" through stylized words and text, often with spray paint in aerosol cans*". It is also a communicative expression that merges with the creative and artistic related to community and territory. Ulmer also presents several graffiti' categories, of which we will work with three of the most representative: "Tag", "Writing/Block", and "Street Art". **Tag** is defined as *"The personal signature of an artist or crew, often written, repetitively, to gain recognition or mark territory"*. It is a unicolor signature that is usually painted with spray paint. **Writing/Block** consists of letters or adaptation of writings to design and elaborate graffiti. It is a category that groups several styles in which the designs are mostly composed of letters; in turn, there are several sub-classifications based on shapes, time of execution of the painting, or size. **Street Art** is *"Visual art placed in public spaces, usually without permission; it includes traditional graffiti; also referred to as post-graffiti or urban art"*. Likewise, **Mural** which can be classified as part of Street Art and can be *"a large painting or visual artwork situated on a wall"*. Examples of those kinds of images located in our focus area can be seen in Figure 1.

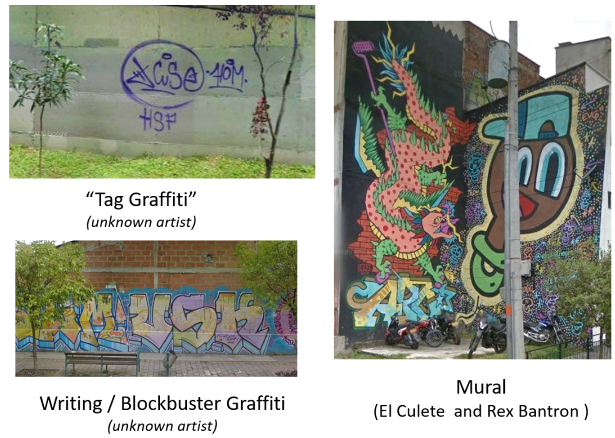

Figure 1: Examples of graffiti by category.

Public policy around graffiti is a complex issue to deal with. It depends on various political positions defined for the municipalities [19] [23], the state, or even the country, and how the artistic expression contributes to the value generation in a city and its society. The public policy regarding graffiti in Colombia is Law 1809 of 2016, which issues the Police Code, regulates the duties regarding public space, and specifically mentions graffiti's proceeds [13]. The Police Code prohibits writing or posting on public or private space "captions, drawings, or graffiti" without proper permission. Likewise, the law gives graffiti space to be carried out in designated areas, with the competent authority's authorization. There is no regulation or explicitly established policy around graffiti in our focus area that expands or contextualizes what is specified in the Police Code.

Computer Vision is the field of study used to achieve our approach. It seeks through techniques based on algorithms and mathematical modeling to provide machines with the ability to perceive, understand, and take action on the information contained in image or video files [11]. Specifically, transfer learning is the most used method for handling images; it allows the use of a previous training parameter so that the new training does not start from scratch but rather appropriates the learning of a different problem [9]. The basic approach to the Computer Vision problem is that of a classification problem. One of the first reference used to evaluate or compare the accuracy and efficiency with models is ImageNet [4] [10], which is a large-scale image ontology built on the backbone of the WordNet structure [14]. Then, since 2012, with the introduction of AlexNet [12], different convolution network architectures become the approach with the best results, growing in efficiency while increasing in size [24].

Several classification image techniques use Deep Learning; some are image preprocessing, semantic segmentation, feature extraction and training, object detection, and localization. Specifically, the last one is our interest because it seeks to identify images, considering that there can be several classes and several objects of the same class in the same image. Unlike the "classification and location" approach, the problem to be solved is more complex and requires more robust architectures [7]. The Object Detection Architecture uses the R-CNN method (Regions with CNN (Convolutional Neural Networks) features); however, Fast R-CNN and Faster R-CNN provide much better results in accuracy and efficiency [6][21]. Specifically, the Faster R-CNN model modifies the R-CNN architecture, proposing regions through a "region proposal network" (RPN). It is part of the 2-step object detection networks because it divides the task into two main parts; the first is an algorithm to propose some regions to be analyzed, followed by a classifier. This type of architecture's main challenge lies in the algorithm for proposing regions, as it is usually very computationally intensive and is considered a "bottleneck". In this way, a significant reduction in the predictions' computational intensity is achieved [21].

In the object detection task, it is important to measure that the prediction class corresponds to the "true" class and measure that the prediction bounding box coincides with the "true" bounding box. It is calculated with the "Intersection over Union" (IoU). When a prediction achieves an IoU above a defined criterion, and the prediction class corresponds, it is considered a true positive. With this information, it is possible to calculate accuracy and recall metrics to interpret the model's effectiveness. Once the indicators per image are calculated, the average precision (AP) is estimated to add a single indicator of the model's effectiveness [20].

3 PROPOSAL

We use ASUM-DM methodology adapted by Angée *et al.* [1]. Figure 2 shows how the experimental process is achieved for our proposal.

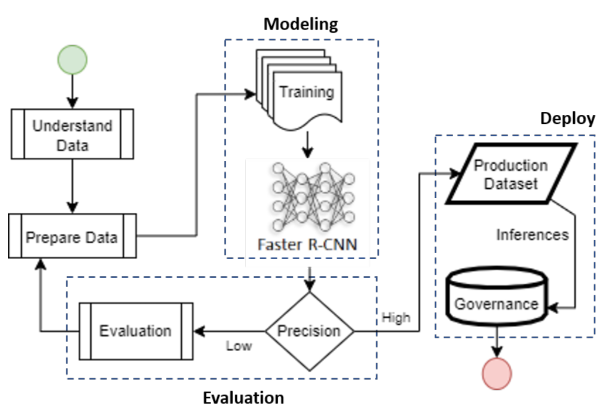

Figure 2: ASUM methodology adaptation to support the experimental process.

3.1 Understand Data

We collected Google Street View (GSV) images from our focus area. City images usually have much information in a single shot; for example, different objects located on the streets, and in the presence of graffiti, it can be located in different zones with different sizes, shapes, and angles. These characteristics make that object detection model a better approach. For our case, diverse datasets are used: STORM, Extended dataset, Production dataset.

STORM is a dataset of graffiti structured for object detection training; the West Attica University collected it in Greece [3]. It comprises 1,022 images with their annotations, bounding boxes, and Class. It is mainly composed of Tag type graffiti and will be part of the training process dataset.

Extended dataset is created as part of our research process during an experimental process. It contains 373 GSV images with Mural and Blockbuster graffiti. These were selected, preprocessed, and mixed with the STORM images to obtain an extended dataset that would add up to 1,395 images from our focus area. It will be part of the training process dataset too.

Production dataset is a collection of GSV images of our focus area. This dataset was obtained using the "Street View download 360" tool. The collected images are called "panorama," a 360^o view of a specific point. To select them, we create a grid of geographical coordinates, with a distance of 26 meters between each point; this is done to reduce the number of images downloaded. We took 71.000 panoramas, each with geolocation information (latitude and longitude). It will allow us to make the geolocalization analysis of the graffiti found.

The collection of this number of images required a significant effort since these images have an accumulated weight in the storage of approximately 100 GB; this means that the primary data engineering effort was invested in constructing this dataset and its associated metadata.

3.2 Prepare Data

We define three tasks to prepare data: Labeling, Size Standardization, and Panoramas transformation.

Labeling. Due to the Faster R-CNN method is part of the supervised models, image labeling is required. It includes the bounding boxes and class labels for the training images. This process is carried out with the LabelerImg tool [5]. Thus, we obtained a file with the Pascal VOC format labels for each image of both training and evaluation datasets.

Size Standardization. In this task, the images are rescaled to guarantee that all images have the same size in the training and evaluation datasets. A dimension of 960x720 pixels is defined.

Panoramas transformation. A panorama image consists of a 360^o view of a specific point; then, several images can be extracted from the same panorama, each corresponding to the perspective in one direction. For each panorama, we extracted six images that correspond to the following angles from the north point: -120^o, -60^o, 0^o, 60^o, 120^o, 180^o (see Figure 3). These images are the input to the prediction model (we decided not to store the images on disk to optimize resources). This preprocessing activity is carried out in conjunction with the inference process.

3.3 Modeling and Evaluation

The Detecto [2] library was used for the training; this is a specialized library in Computer Vision and object detection; it is based on Pytorch, which is one of the leading Deep Learning frameworks. A Faster R-CNN network is trained over a pre-trained ResNet-50 classifier on ImageNet. During the training process, the network

Figure 3: Panorama image extraction.

parameters are adjusted to minimize the loss function; it is calculated with the model's prediction and the test data. Two training are carried out, one with the STORM dataset and one with the extended dataset. Then, we built the Validation/Evaluation dataset, which was composed of data (images and metadata) with the same distribution of the Production dataset. This dataset comprises 164 GSV images located in our focus area, i.e., graffiti type Tag, Block, Mural. This dataset is used to measure the performance's model; in other words, the prediction results on the Validation/Evaluation dataset.

In the object detection task, many possible bounding boxes should not be detected within an image; this means that using a metric that contains a True Negative rate may lead to biased results. Therefore, we use the Average Precision (AP) metric since a regular Accuracy might not be the best metric to evaluate the performance due to the misdirection caused by the True Negative rate [20]. Thus, from the Validation/Evaluation dataset, we generated four sub-datasets to analyze where the model presented better and worse results (i.e., minimum learning guarantee).

Easy Tag. It comprises 33 images with medium-sized Tag type graffiti, with few tags per image, without other types of graffiti.

Hard Tag. It comprises 31 images; unlike the previous dataset, it includes tagged images of different sizes, including small tags, additionally images of facades with lots of graffiti (many predictions on the same image).

Block. This 36-image dataset contains block-type graffiti, some medium-sized, and another full wall size.

Mural. It containts 25 wall images, most of which are full size. Some are partially hidden by other objects such as trees or vehicles. The model results are the prediction over the class "Graffiti", the probability of belonging, and the vector with the bounding boxes information.

3.3.1 STORM dataset training evaluation. Once the first iteration with the STORM dataset's data is done, the Average Precision (AP) obtained is 58,3%. When we evaluate the prediction on the four subsets of images, the best results are obtained on images with Tag (easy Tag AP 73%; hard Tag AP 66%); the results on other graffiti types are significantly lower (Block AP 59%, Mural 37%).

3.3.2 Training evaluation with the Extended dataset. When analyzing the validation results of the trained model in STORM vs. the model trained with the Extended dataset, it is observed that the results improve significantly: STORM AP 58.3% vs. Extended dataset AP 69,14% (see Figure 4).

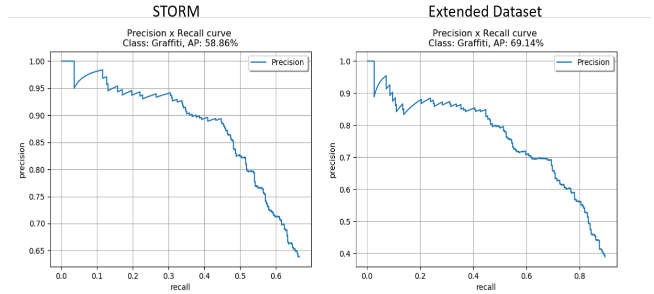

Figure 4: STORM AP vs. Extended dataset AP.

Figure 5) shows the extended dataset results model, which is better in all subsets. Particularly the Tag's type results (Easy Tag 89.23%; Hard Tag 72.4%). Also, Block type graffiti's results improved significantly (Block STORM 59%; Block extended 75%), even when the results on Murals improved (37.4% to 59.13%) still are the lowest of the different datasets. The Extended dataset results model is better in all subsets, particularly the Tag's type results (Easy Tag 89.23%; Hard Tag 72.4%). Also, the results on Block type graffiti improved significantly (Block STORM 59%; Block extended 75%), even when the results on Murals improved (37.4% to 59.13%) still are the lowest of the different datasets (see Figure 5). Specifically, this presents a significant challenge since the concept of Mural or artistic graffiti is ambiguous, so it is difficult to define its latent characteristics that allow the classification with a model like the one proposed. With these results, the model is considered good enough to be deployed.

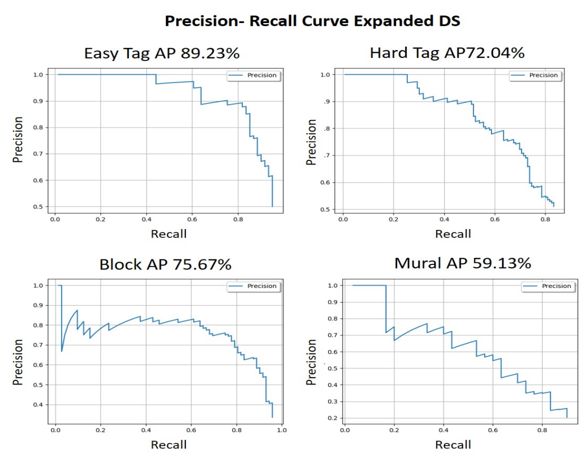

Figure 5: Extended dataset Recall Precision curves.

3.4 Deploy

At this stage, we propose a way to guide and transfers knowledge through image inference. Given the results obtained in the previous section, the model trained with the Extended dataset is chosen to deploy with the Production dataset. In this process, we create a dictionary with the geolocation information and the number of geographical position detections. Seventy-one thousand one hundred geographical positions were evaluated, each with six images, which means that the inference was made on 426,600 images. As a result, we identified 2,869 positions with a concentration of graffiti (i.e., 4.03% of the total positions analyzed) and 19,696 graffiti. With this information, we presented the results related to zone concentration using heat map visualization. To achieve that, We propose two categories of graffiti concentration zones High and Low.

High. It means the zone could have high graffiti concentration, which is shown using a large proportion of red, orange, or yellow colors. It could be viewed from two perspectives: 1) where images are concentrated in almost the whole zone; 2) where the images are scattered in different zone points.

Low. It means the zone could have low graffiti concentration, which is shown using a large proportion of blue color or absence of color.

Figure 6 shows Medellín's downtown (Z1), which is identified as a zone with high graffiti concentration. It gathers vast commercial areas that could provoke social and security concerns. On the other hand, the west side of the city (Z2) presents a low concentration graffiti zone because it comprises middle-income residential neighborhoods, except for these zones' main access roads.

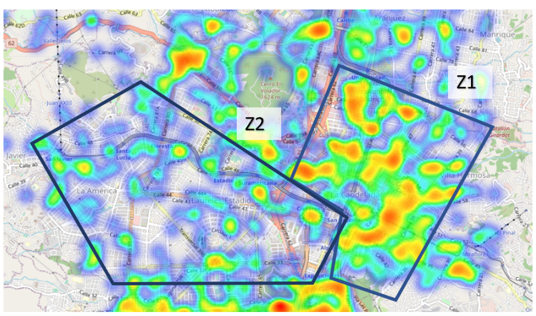

Figure 6: Downtown Medellín (Z1). Residential neighborhoods at the west of Medellín (Z2).

Figure 7 shows upper-middle-income residential zones located at the southeast of Medellín, such as El Poblado. Specifically, that zone has suffered substantial transformation to commercial and entertainment places in the last ten years, leading to cultural changes from vandalism graffiti to Block and Mural graffiti. Besides, the north of Envigado municipality is a residential zone; however, it is possible to see a yellow point where there are Tag or Block graffiti outside the private residential complex wall.

Figure 8 highlight two zones. Z3 has a dense concentration of industrial and commercial areas and one of the main Medellín avenues. Precisely, it is occupied by homeless people at not-working hours, provoking Tag and Block graffiti as part of either vandalic or cultural expressions. Big places characterize Z4 as the city Zoo and a recreational sports center, which have fomented the Mural graffiti.

Likewise, in the municipalities in the south of the focus area, we identified zones with high graffiti concentration surrounding the town squares.

Figure 7: Residential zones on the east side of the focus area.

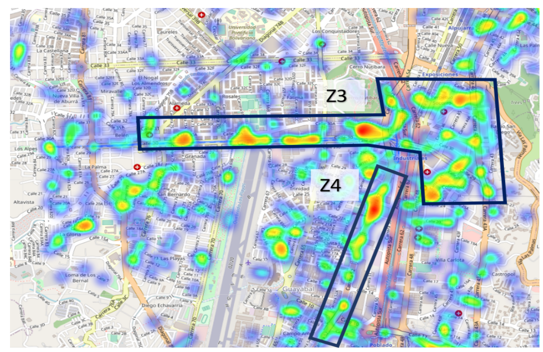

Figure 8: Industrial/Commercial zone (Z3). Cultural zone (Z4).

Guidelines to create or improve Public Policies from the proposal. Once a high graffiti concentration zone has been identified, we suggest the concerned authorities take into account the following guidelines: 1) Analyze the graffiti from their content and impacts on the community. Some suggested parameters are the purpose (vandalic, cultural, political, etc.), geolocalization, concentration level, urban zone characteristics, etc. 2) Analyze the current regulations related to security, culture, touristic, or other proposals applied to the zone. 3) Identify possible zone transformation, creating or improving public policies.

For instance, areas with a high Tag type concentration or vandalic graffiti located in industrial/commercial zones could generate negative impacts such as property deterioration and safety risks; then, restoration interventions should be promoted. Otherwise, in areas with a high concentration of graffiti and where content analysis identifies an artistic vocation, then, tourist activities must be promoted. To this end, communities and companies dedicated to guiding tourists and public administration should coordinate promotional activities to guarantee safety conditions. Commercial areas must coordinate with companies or government joint calls where artists can intervene in the walls with promotional activities.

4 CONCLUSION AND FUTURE WORK

The proposal presented in this article identified graffiti and concentration zones through Computer Vision and object detection and localization techniques to support public policy management in smart cities. Thus, the implementation guidelines were defined from the resulting models to formulate governance concerning the graffiti treatment. Specifically, this article presented the Medellín and some border municipalities case study. A visualization demo can be seen at https://github.com/BaoxiticDC/GraffitiData.

In the particular case of Medellín and border municipalities, it was identified that there is no formal policy beyond what is regulated in the Police Code regarding graffiti. Still, there is evidence of interest in promoting and integrating Street Art into the city's aesthetic, such as calls for painting in specific areas defined for this purpose.

Regarding the Deep Learning models application, we found evidence that our proposal of using object detection and localization technique for graffiti classification was suitable, as shown in the results in section 3. Furthermore, object detection architecture Faster R-CNN instead of simple convolutional networks significantly improved the results. On the other hand, the methodology used to estimate the proposed model's generalization error supported the model's predictive capacity.

As part of the exercise's implementation, a dataset of 426,600 images of Medellín and some border municipalities was formed. The prediction results allowed for analyses that could enrich public policies' definition with the images' metadata. Besides, the visualization tool is built, identifying graffiti concentration zones through heat maps. That allowed the definition of guidelines to create or improve sectorized policies, classify these zones to carry out preservation activities in artistic graffiti or restore in cases where the graffiti has a negative connotation or deteriorates public or private property.

From the identification of graffiti concentration zones proposal, we will work in extending the possibilities of images analysis through a more in-depth classification to achieve new graffiti classes based on its shape (e.g., Tag, Block, Mural, among others) or by its purpose (e.g., artistic, politic, educational, among others).

To complement the current approach, we will enhance the analysis of concentration zones with additional information, such as economic level, lighting, criminality level, etc. It could lead to finding correlations or causal effects that enable a better understanding of the phenomenon of graffiti.

REFERENCES

[1] Santiago Angée, Silvia I. Lozano-Argel, Edwin N. Montoya-Munera, Juan-David Ospina-Arango, and Marta S. Tabares-Betancur. 2018. Towards an Improved ASUM-DM Process Methodology for Cross-Disciplinary Multi-organization Big Data & Analytics Projects. In *Knowledge Management in Organizations*. Springer International Publishing, 613–624.

[2] ALAN BI. 2019. Detecto. Alan Bi. https://github.com/alankbi/detecto

[3] Patrikakis Charalampos, Kasnesis Panagiotis, Toumanidis Lazaros, and Tzitamidis Alexandros. 2019. STORM graffiti/tagging detection dataset. Zenodo. https://doi.org/10.5281/zenodo.3238357

[4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. 2009. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*.

[5] Bryan Russell Antonio Torralba William T. Freeman. 2013. labelImg. MIT, Computer Science and Artificial Intelligence Laboratory. https://github.com/tzutalin/labelImg/blob/master/LICENSE

[6] R. Girshick. 2015. Fast R-CNN. In *2015 IEEE International Conference on Computer Vision (ICCV)*. 1440–1448.

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik. 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 580–587.

[8] Nikita Gordienko, Peng Gang, Yuri Gordienko, Wei Zeng, Oleg Alienin, Oleksandr Rokovyi, and Sergii Stirenko. 2018. Open Source Dataset and Machine Learning Techniques for Automatic Recognition of Historical Graffiti. In *Neural Information Processing*, Long Cheng, Andrew Chi Sing Leung, and Seiichi Ozawa (Eds.). Springer International Publishing, Cham, 414–424.

[9] Nathaniel Huber-Fliflet, Fusheng Wei, Haozhen Zhao, Han Qin, Shi Ye, and Amy Tsang. 2019. Image Analytics for Legal Document Review: A Transfer Learning Approach.

[10] Mi-Young Huh, Pulkit Agrawal, and Alexei A. Efros. 2016. What makes ImageNet good for transfer learning? *CoRR* abs/1608.08614 (2016). http://arxiv.org/abs/1608.08614

[11] Asharul Khan. 2020. Machine Learning in Computer Vision. *Procedia Computer Science* 167 (04 2020). https://doi.org/10.1016/j.procs.2020.03.355

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. [n.d.]. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25*.

[13] ley 1801. 2016. Código Nacional de Policiía y Convivencia. Gaceta del Congreso de la República de Colombia. https://www.policia.gov.co/sites/default/files/ley-1801-codigo-nacional-policia-convivencia.pdf

[14] George A Miller. 1998. *WordNet: An electronic lexical database*. MIT press.

[15] Pedro Munsberg, Glauco R.and Ballester, Marco F. Birck, Ulisses B. Correa, Virginia O. Andersson, and Ricardo M. Araujo. [n.d.].

[16] P. Nahar, K. Wu, S. Mei, H. Ghoghari, P. Srinivasan, Y. Lee, J. Gao, and X. Guan. 2017. Autonomous UAV forced graffiti detection and removal system based on machine learning. In *2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation*. 1–8.

[17] Zhang Nan, Zhang Fan, and Enmao Liu. 2020. Design of A Shared Platform for Interactive Public Art from Perspective of Dynamic Vision. In *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 37–42.

[18] Tessio Novack, Leonard Vorbeck, Heinrich Lorei, and Alexander Zipf. 2020. Towards Detecting Building Facades with Graffiti Artwork Based on Street View Images. *ISPRS International Journal of Geo-Information* 9, 2 (2020), 98.

[19] Gabriel Ortiz van Meerbeke and Bjørn Sletto. 2019. 'Graffiti takes its own space' Negotiated consent and the positionings of street artists and graffiti writers in Bogotá, Colombia. *City* 23, 3 (2019), 366–387.

[20] R. Padilla, S. L. Netto, and E. A. B. da Silva. 2020. A Survey on Performance Metrics for Object-Detection Algorithms. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. 237–242.

[21] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *CoRR* abs/1506.01497 (2015).

[22] Jeffrey Ian Ross and Benjamin S Wright. 2014. "I've Got Better Things to Worry About": Police Perceptions of Graffiti and Street Art in a Large Mid-Atlantic City. *Police Quarterly* 17, 2 (2014), 176–200.

[23] Hunter Shobe and Tiffany Conklin. 2018. Geographies of graffiti abatement: Zero tolerance in Portland, San Francisco, and Seattle. *The Professional Geographer* 70, 4 (2018), 624–632.

[24] Mingxing Tan and Quoc V. Le. 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *CoRR* abs/1905.11946 (2019). arXiv:1905.11946 http://arxiv.org/abs/1905.11946

[25] Natalia Pérez Torres. 2018. The Graffitour of the 13: an aesthetic, political and historical trajectory through Medellín. *Vibrant: Virtual Brazilian Anthropology* 15, 1 (2018).

[26] Jasmine B Ulmer. 2017. Writing urban space: Street art, democracy, and photographic cartography. *Cultural studies critical methodologies* 17, 6 (2017), 491–502.

[27] Shuqin Wang, Jerry Gao, Weiyi Li, Yanning Li, Kaixuan Wang, and Shengqiang Lu. 2019. Building Smart City Drone for Graffiti Detection and Clean-up. https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00337