A Generative Approach to Light Placement for Street Lighting

I. Evangelou[1], N. Vitsas[1], G. Papaioannou[1], A. Gkaravelis[1]

[1]Department of Informatics, Athens University of Economics and Business, Greece

Abstract

The design of plausible and effective street lighting configurations for arbitrary urban sites should attain predetermined illuminance levels and adhere to specific layout intentions and functional requirements. This task can be time consuming, even for automated solutions, since there exists an one-to-many mapping between illumination goals and lighting options. In this work, we propose a generative approach for this task, based on an adversarial optimisation scheme. Our proposed method effectively overcomes these task-specific limitations by providing a range of viable solutions that adhere to the input constraints and can be generated within an interactive design life cycle.

CCS Concepts

• ***Computing methodologies*** → *Graphics systems and interfaces;* ***Artificial intelligence;***

1. Introduction

Adequate and appropriate illumination of any closed or open environment is a core concept in civil engineering. It defines and controls the way a person navigates, experiences and functions in a particular space. In an urban environment, proper nighttime illumination, via street lighting, is essential for enhancing its safety, security and aesthetic appearance. Street lighting planning is a particularly difficult task because of the scale that one has to operate in.

Several methods have been proposed over the years for automatic generation of urban lighting configurations. Schwarz and Wonka [SW14] proposed a procedural model based on constrained optimisation with application to architectural lighting. However, their method requires the placement of user-defined illumination goals on the actual geometry and does not scale well with respect to the number of light sources to be applicable in city-scale street lighting. Zou and Li [ZL10] proposed a genetic algorithm that optimises the luminous intensity distribution of a fixed arrangement of road lights, in a single road segment. Our approach is designed to work on a large scale and more importantly, without requiring complex constraint definition other than the target illumination goals themselves. We adopt a generative approach to street lighting design in order to estimate multiple valid alternative lighting configurations fast, while learning layout aesthetics and practices by example, instead of trying to impose hard rules.

2. Method

In this work we propose a goal-driven method for the generative placement of street lights based on conditional Generative Adversarial Networks (GAN) theory. We use a combination of supervised and unsupervised learning in order to extract meaningful information for both layout characteristics of street lighting configurations (supervised) and illumination level targets (unsupervised).

Generative Adversarial Networks [GPAM*14] have demonstrated their exceptional expressive power to model complex distributions even in the conditional setting [MO14]. Their applicability spans diverse fields, such as image synthesis, image-to-image translation, style transfer and many others. Here, we shift our focus on the class of adversarial games subject to conditional input and optimised using an alternate gradient descent scheme. Briefly, the objective function is defined as:

$$\mathcal{L}(G,D) = \mathbb{E}_{(x,y)}\left[\log(D(x,y))\right] + \mathbb{E}_{(z,y)}\left[\log(1-D(G(z,y),y))\right] \tag{1}$$

where (G,D) are distinct neural networks, (x,y) is a tuple of an element sampled from the real distribution with its associated condition. Likewise, tuple (z,y) encapsulates a sample seed drawn from a predefined distribution along with a condition variable. The generator network G is trained in order to learn a mapping $(z,y) \mapsto x'$ through minimisation of Equation 1, consequently generating samples that closely resemble the source distribution subject to the same conditional variable. On the contrary, the discriminator network D is trained to distinguish between real and generated samples (x,x'), subject to condition y, thereby maximising Equation 1.

Dataset. Our case study is Washington DC. The District Department of Transportation (DDOT) maintains a publicly available source containing information on street light positions and their corresponding wattage [Cit23]. Additionally, we leveraged the Open Street Map (OSM) [Ope17] public API to obtain corresponding building and street geometry data.

To integrate the combined source data, we partitioned the city

DOI: 10.2312/egs.20241031

landscape into overlapping tiles of size $128m^2$. These city tiles were rendered using ray-traced direct illumination and purely diffuse materials. We used a top-down orthographic view of the same resolution (1 pixel per m^2), since we were only interested in street-level illuminance. We observed that a full global illumination pass with varying materials for building and street geometry contributed very little to the resulting illuminance, especially when compared to real aerial photographs. This choice also greatly simplified the evaluation during training (see next). The luminaries in the scene were represented as hemispherical point light sources, pointing towards the ground with a vertical offset of 4 meters. To summarise, our dataset consists of about 71K luminaries ranging from 10000lm to 40000lm (LED elements with an efficacy of 100), distributed and rendered into 33K residential tiles.

For each rendered tile, a segmentation mask is also computed, which classifies individual pixels according to function (building, street or general background). Additionally, for the street regions, further sub-classification is carried out based on the tags reported by OSM [Ope17]. For each road section, we determine the illumination goal for a given tile as the average illuminance (lm/m^2) registered in segment's pixel during the lighting simulation. We also assume that non-street pixels have a zero target illuminance. Overall, a single sample from the processed dataset consists of the rendered image $\mathbf{X}_r$, the segmentation mask $\mathbf{S}$, and the target illumination objective $\mathbf{Y}$, all being images of 128^2 resolution.

Network architecture. Contrary to the majority of existing adversarial models targeting image synthesis, our approach does not predict individual pixel values. Instead, we are interested in sampling from the marginal distribution of valid light configurations under specific illumination constraints. Consequently, the generator model, in addition to the standard random noise input, is conditioned on the input illumination goal $\mathbf{Y}$. The output is a regular grid representing candidate light positions and corresponding light intensities. Typically, the grid size is smaller than the resolution of the image. In this work, we set the grid size to 24^2, effectively providing candidate light positions with a spacing of 5 meters.

Candidate light intensities are inferred from a trainable linear output. To constrain the output predictions, we apply two filtering operators on the inferred light grid. First, we linearly suppress intensities of lights that are placed far away from the road segments. For this, we define the distance of a luminary from a road segment as the frequency of pixels categorised as streets in a 11^2 region centred at the candidate light position.Then, define the decay factor of $p = 1$ (no suppression) if frequency is $\geq 1/3$, which linearly decays to zero and effectively disables lights in regions with no road segments. Second, we set a minimum threshold to the luminous intensity of each light to better match the real world dataset and avoid degenerate lighting configurations containing many faint lights. This is achieved by using a shifted-ReLU activation function. This parameter is a constant for all our experiments and set to 100. To account for the zero gradients due to clipping in the backward pass, we approximate the shifted-ReLU activation using the Softplus function with the elasticity hyperparameter set to $b = 20$.

For every generated light grid, the resulting illumination must be computed, in order to compare the illumination with respect to the given goal $\mathbf{Y}$. Since we only employ diffuse direct illumination with punctual light sources, the illuminance estimation can be reduced to a vectorised operation between precomputed illuminance transfer tensors, which are estimated on the GPU during initialisation, using image (tile) space ray marching. Since there is a fixed number of light positions on each tile, we can precompute a two-dimensional illuminance transfer matrix $\mathbf{T}^{n\times\ell}$, where $n = H \times W$, H, W are the tile dimensions and ℓ is the number of lights. Each matrix element is the product of the geometric, albedo and visibility terms combined. To compute the final illuminance, a matrix multiplication with a light grid in column vector $G(z, \mathbf{Y}) = \mathbf{L}$ of size $\ell \times 1$ must be performed. This final operation is the output of the generator network which yields a directly illuminated image $\mathbf{X}_f$.

The discriminator network is given a tuple of real and fake samples along with the conditional goal $\{\mathbf{X}_{r/f}, \mathbf{Y}\}$ and is trained to predict the per-pixel source distribution and conformance to the illumination goal. It operates directly on rendered results rather than some raw sequence of lights. This allows to learn the actual impact of each light relative to the goal $\mathbf{Y}$ on a per pixel basis. Predictions are restricted to tile pixels (i, j) categorised as streets in $\mathbf{S}$. The discriminator must evaluate two criteria simultaneously: the per-pixel deviation from the input luminance levels $\mathbf{Y}$ and the light placement style (distribution), always prioritising the first. We define $m_{ij} = 1$ iff $|x_{ij} - y_{ij}|/y_{ij} \leq \alpha$ and 0 otherwise. The error tolerance α is set to 0.2 for all our experiments. We use this factor to guide the training to respect the illumination goal for pixels in the receptive field that this is attained. This is a necessary modification in order to capture the light position "style" in the training set, without being affected by the achieved illumination in the input images, which might deviate from the objective goal. The first term in Eq. 1 becomes:

$$\mathbb{E}_{(\mathbf{X},\mathbf{Y})}\left[\sum_{(i,j)\in\mathbf{R}} m_{ij}\log(D_{ij}(\mathbf{X},\mathbf{Y})) + (1-m_{ij})\log(1-D_{ij}(\mathbf{X},\mathbf{Y}))\right] \tag{2}$$

where, $\mathbf{R} = \{(i, j)|s_{ij} = \text{street}, s_{ij} \in \mathbf{S}\}$. The second term in Eq. 1, related to the generator update term remains unchanged, treating all pixels belonging to the street category as fake.

An overview of our network architecture is summarised in Figure 2. The generator network is modelled as a style-based decoder through a series of spatially adaptive normalisation layers [PLWZ19], followed by our light placement regression, filtering and direct illumination calculation module. The discriminator part of the network is conditioned on the input sample $\mathbf{X}_{r/f}$ through a channel-wise concatenation following the principles established in [MO14]. Its architecture bears resemblance to the U-Net approach originally introduced in [SSK20]. Given that we operate in High Dynamic Range images, and as per established literature, it is customary to rescale the discriminator's input values to lie close to the unit interval. This rescaling is achieved by dividing the inputs including the goals $\mathbf{Y}$, with a constant factor set to $s = 20$. The scaling factor is approximately equivalent to the illuminance received at a shading point located at the center of a 3×3 light grid with maximum lumen output, a spacing of 1 meter, and a height of 4 meters. We also train our GAN model using the non-saturating loss variant. Finally, the generator outputs the light placement in candela units where during the training phase we use a unit efficacy, while during the inference stage we convert and scale values to proper lumen units.

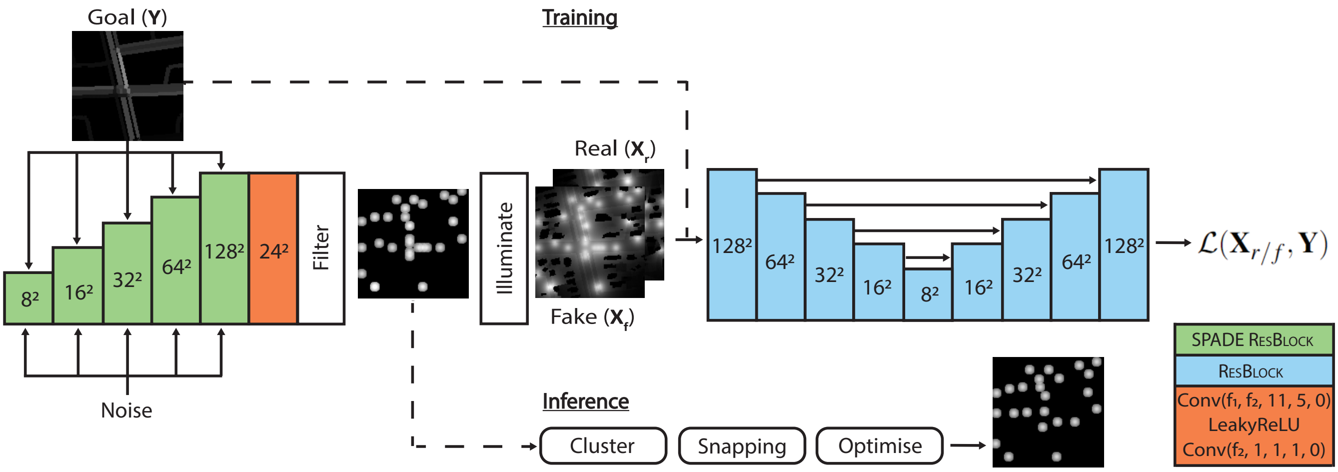

Figure 1: *Our proposed network architecture for the generative inference of street light configurations. We train a conditional generator network [PLWZ19] (green) to transform a sample seed followed by a regression head (orange) to output a fixed size (24^2) light grid of intensity values. These are subsequently filtered to suppress unimportant lights and rendered to produce the final illuminated block (see Section 2). During training (top-right), we employ a per-pixel discriminator [SSK20] (blue). In the inference phase (bottom-right), we post-process the predicted light intensities and positions, yielding the final configuration in lumen units.*

Post processing. We incorporate a series of essential post-processing steps on the generator's output light grid, to further refine the proposed light placement with respect to road segment boundaries. Here we assume that lights have a coverage radius of 5 meters in real units or equivalently, 5 pixel units. Based on empirical observations from the real dataset distribution, we anticipate that, in the general case, overlapping coverage among individual light sources is rarely needed. For our grid resolution, this implies that light source contributions should not overlap within their immediate 1-neighbourhood. To impose the observed pattern, we implement a clustering scheme, where only the most dominant source in a 3×3 neighbourhood centred at each grid sample remains active. The other light sources are deactivated. Furthermore, our approach takes into account the fact that lights tend to be aligned with road segment boundaries (e.g. sidewalks). Consequently, we snap the luminary positions to the closest road edge, based on the image spatial gradient extracted from **S**. The aforementioned operations may introduce an error relative to the goal. To account for this, we apply a limited number of gradient descent steps in order to optimise the luminous intensity of the updated light configuration in terms of the target goal. For this step, we utilise the direct illumination module discussed in the previous section and apply the mean square error loss function relative to the actual image goal. It is worth noting that these gradient descent steps, are employed to adjust the luminous intensity of existing light sources and therefore do not activate or deactivate any light source.

3. Evaluation

We implemented our method in PyTorch and trained the generator and discriminator networks with the Adam optimiser. We used a learning rate of $1.e-4$ and $2.e-4$ respectively, with $\beta_1 = 0$ and $\beta_2 = 0.999$. Additionally, we used multi-level noise injection [LHK19] sampled from a zero mean Gaussian distribution. For the training process, we used 31K rendered blocks, from which we pick 1K samples at random for each epoch and augment them using flipping and rotational transformations as proposed by [ZLL*20]. We evaluated the quality of inference on a separate test set which consists of 2K blocks. All our experiments were run on an NVIDIA RTX 3080Ti with 12GB of VRAM.

We quantify the quality of light placement configurations on 16 distinct seeds of each input residential block. First we use the lighting uniformity formula, defined as the average ratio of minimum over the average illuminance value for each road section present in the block. According to established design rules used for residential blocks, the uniformity value should usually be greater than 0.2 [dl10]. In our test subset this value is on average 0.33 for each block, and including the post-processing step is 0.34, which is about 5% better than the measurements on the reference subjects, indicating that every road section is well and evenly lit. Additionally and for the same setup, we also measure the average per pixel absolute error (*MAE*) between the predicted illuminance and the conditional goal on the corresponding road pixels. This evaluates to 17.76lx and 16.06lx for the generated and post-processed blocks respectively. The actual target values for road sections, range from 100lx to 1500lx in our dataset. In Figure 2 we demonstrate some indicative test cases.

Finally, our generator network consists of roughly 13M trainable parameters, while the discriminator network has about 4M. In terms of inference performance, our model can generate conditional light placement configurations in roughly 9ms per input random seed, while the whole post-processing phase requires an additional 30ms, on average. Since the generator network is trained to infer well-separated and aligned light placement settings, our clustering phase is employed for an average of 2 to 3 rounds and the local optimisation step is also terminated after about 15 iterations.

Proceedings published by Eurographics - The European Association for Computer Graphics.

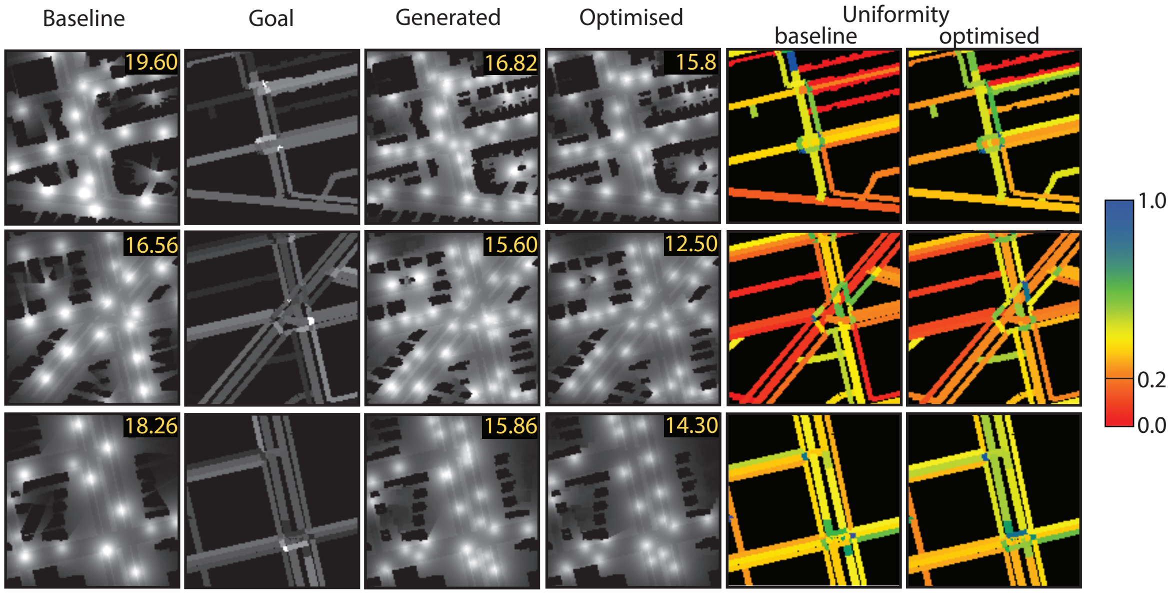

Figure 2: *Goal-driven street lighting results for different city tiles. From left to right, illuminance from the original lights (baseline), the illumination goal for each road section (in the interval of 100lx to 1000lx), illuminance from the light configurations generated by the GAN, illuminance after post-processing, measured uniformity for the baseline and our results (larger values are better). Yellow values in the top right corner indicate per pixel MAE (deviation from the goal).*

4. Conclusions and future work

In this work, we have introduced an interactive generative approach aiming to address the task of goal-driven optimisation in the context of street lighting within urban environments. Our methodology achieves this objective by sampling from a distribution of feasible light placement configurations, while ensuring that these configurations faithfully adhere to the specific conditional lighting goals. We have leveraged the conditional adversarial theory and adapted the adversarial loss function to better suit the unique requirements for this task. An interesting future task is the integration of our method into a general system, that is capable of simultaneously tackling larger-scale inputs that encompasses multiple strides of residential blocks, ultimately yielding a neighbourhood-scale lighting solution.

Data Availability. Our code and dataset is publicly available at https://github.com/cgaueb/streetGAN.

Acknowledgements

This research was funded by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "3rd Call for H.F.R.I. Research Projects to support Post-Doctoral Researchers" (Project No: 7310).

References

[Cit23] CITY OF WASHINGTON: Street lights, 2023. data retrieved from Open Data DC, opendata.dc.gov/datasets/DCGIS::street-lights/about. 1

[dl10] DE L'ÉCLAIRAGE C. I.: Lighting of roads for motor and pedestrian traffic: Cie 115: 2010. CIE. 3

[GPAM*14] GOODFELLOW I., POUGET-ABADIE J., MIRZA M., XU B., WARDE-FARLEY D., OZAIR S., COURVILLE A., BENGIO Y.: Generative adversarial nets. In *Advances in Neural Information Processing Systems* (2014), Ghahramani Z., Welling M., Cortes C., Lawrence N., Weinberger K., (Eds.), vol. 27, Curran Associates, Inc. 1

[LHK19] LEE S., HA J., KIM G.: Harmonizing maximum likelihood with GANs for multimodal conditional generation. In *International Conference on Learning Representations* (2019). URL: openreview.net/forum?id=HJxyAjRcFX. 3

[MO14] MIRZA M., OSINDERO S.: Conditional generative adversarial nets. *CoRR abs/1411.1784* (2014). arXiv:1411.1784. 1, 2

[Ope17] OPENSTREETMAP CONTRIBUTORS: Planet dump retrieved from https://planet.osm.org . www.openstreetmap.org, 2017. 1, 2

[PLWZ19] PARK T., LIU M., WANG T., ZHU J.: Semantic image synthesis with spatially-adaptive normalization. *CoRR abs/1903.07291* (2019). arXiv:1903.07291. 2, 3

[SSK20] SCHONFELD E., SCHIELE B., KHOREVA A.: A u-net based discriminator for generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 8207–8216. 2, 3

[SW14] SCHWARZ M., WONKA P.: Procedural design of exterior lighting for buildings with complex constraints. *ACM Trans. Graph. 33*, 5 (sep 2014). 1

[ZL10] ZOU J., LI L.: Optimization of luminous intensity distribution of roadway lighting luminaire based on genetic algorithm. In *2010 Second WRI Global Congress on Intelligent Systems* (2010), vol. 1, pp. 327–330. 1

[ZLL*20] ZHAO S., LIU Z., LIN J., ZHU J.-Y., HAN S.: Differentiable augmentation for data-efficient gan training. In *Conference on Neural Information Processing Systems (NeurIPS)* (2020). 3