Counterterrorism for Cyber-Physical Spaces: A Computer Vision Approach

Giuseppe Cascavilla
Universiteit Jheronimus Academy of Data Science
Netherlands
g.cascavilla@tue.nl

Johann Slabber
Universiteit Jheronimus Academy of Data Science
Netherlands
j.slabber@tilburguniversity.edu

Fabio Palomba
SeSa Lab
University of Salerno
Italy
fpalomba@unisa.it

Dario Di Nucci
Universiteit Jheronimus Academy of Data Science
Netherlands
d.dinucci@uvt.nl

Damian A. Tamburri
Universiteit Jheronimus Academy of Data Science
Netherlands
d.a.tamburri@tue.nl

Willem-Jan van den Heuvel
Universiteit Jheronimus Academy of Data Science
Netherlands
W.J.A.M.v.d.Heuvel@jads.nl

ABSTRACT

Simulating terrorist scenarios in cyber-physical spaces—that is, urban open or (semi-) closed spaces combined with a cyber-physical systems counterparts—is challenging given the context and variables therein. This paper addresses the aforementioned issue with AL_{Ter} a framework featuering computer vision and Generative Adversarial Neural Networks (GANs) over terrorist scenarios. We obtained the data for the terrorist scenarios by creating a synthetic dataset, exploiting the Grand Theft Auto V (GTAV) videogame, and the Unreal Game Engine behind it, in combination with OpenStreetMap data. The results of the proposed approach show its feasibility to predict criminal activities in cyber-physical spaces. Moreover, the usage of our synthetic scenarios elicited from GTAV is promising in building datasets for cybersecurity and Cyber-Threat Intelligence (CTI) featuring simulated videogaming platforms. We learned that local authorities can simulate terrorist scenarios for their own cities based on previous or related reference and this helps them in 3 ways: (1) better determine the necessary security measures; (2) better use the expertise of the authorities; (3) refine preparedness scenarios and drills for sensitive areas.

CCS CONCEPTS

• **Human-centered computing → Visualization systems and tools.**

KEYWORDS

Cyber-Physical Spaces, Counterterrorism, Computer Vision, Generative Adversarial Neural Networks

ACM Reference Format:
Giuseppe Cascavilla, Johann Slabber, Fabio Palomba, Dario Di Nucci, Damian A. Tamburri, and Willem-Jan van den Heuvel. 2018. Counterterrorism for Cyber-Physical Spaces: A Computer Vision Approach. In *AVI'20: International Conference on Advanced Visual Interfaces, June 08–12, 2020, Ischia, Italy.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/1122445.1122456

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AVI'20, June 08–12, 2020, Ischia, Italy
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

While the world has never been safer than today, the news seems to be dominated by headlines of terrorist attacks, whether it is a vehicle ramming in a large crowd or a mass shooting [35]. Consequently, authorities trying to secure urban spaces are faced with countless different possibilities of attacks. At the same time, such authorities are limited to the experience and imagination of local law-enforcement, which, especially in smaller cities, lacks the knowledge and infrastructure to be fully-prepared for these events as well as the speculative technology that would allow them for better-instrumented preparedness plans [21]. To support such endeavours, we investigate how computer vision approaches can help law-enforcement agencies and municipalities in the protection of urban spaces against terrorism [29].

On the one hand, our goal is to inform law-enforcement agents (LEAs) about the consequences of these attacks on the locations they protect.

On the other hand, we aim at simulating the behaviour that the agents should have.

To this aim, we designed and prototyped AL_{Ter}—**A**dversarial **L**earning for counter**Ter**rorism—a new framework to simulate complex terrorism scenarios based on computer vision and deep learning. The purpose of AL_{Ter} is two-fold. On one side, the AL_{Ter} solution helps LEAs to identify vulnerable locations and prepare appropriate responses and contingency plans [13]. On the other side, AL_{Ter} provides city policy-makers a simulation of the consequences that attacks could have in specific areas of public spaces, hence giving the LEAs and municipalities the possibility to predict, manage, and avoid terrorist attacks with a surgically-precise lens of analysis. Finally, AL_{Ter} would allow local law-enforcement agencies to provide landmark images from their location and simulate different scenarios that occurred in other locations, thus enabling the actionable use of counter-terrorism information sharing for the purpose of preparedness and urban contingency-planning beyond terror-understanding.

To this end, in the scope of designing AL$_{Ter}$, we also proposed a novel approach to elicit terrorist scenarios from images and transferring them to different landmarks [34]. In particular, we adopt a computer vision approach that trains Generative Adversarial Networks [18] by means of images of terrorist attacks extracted from video-gaming scenarios. In the scope of AL$_{Ter}$, GANs are used as a key design artifact since they were previously used to reproduce fake images which are almost indistinguishable to real ones (e.g., deep fakes) [17] and therefore we assumed they may be suitable to terrorist attack simulation as well.

To the best of our knowledge, AL$_{Ter}$ is the first attempt in this direction.

To demonstrate the feasibility of our approach we implement a proof-of-concept and simulate terrorist attacks via gamification [19]. In particular, we extract video recordings of terrorist attacks from Grand Theft Auto V[1] (GTAV), a well-known action-adventure videogame. We conducted several controlled-gaming sessions and we recorded the fragments in which criminal and terrorist activity are perpetrated. We use such gaming environment because of the lack of data available about these scenarios and the privacy sensitivity of the data [37]. To extract the location of the events where to simulate the attacks, we rely on OpenStreetMap[2] (OSM), a collaborative project to create a free editable map of the world. Both these source of information were used to create a synthetic dataset which could be reused in the future for other purposes as well. Finally, we use StyleGAN [24] to map the features of the extracted terrorist attack scenes to the latent space. Afterwards, we change such latent space to extract terrorist attacks (e.g., a vehicle ramming humans in an image) and transfer them to another locations.

We evaluate our proof-of-concept in the real-world scenario of the city of Malaga (ES) in which we simulated a small scale terrorist attack (i.e., a terrorist starting a fire in main square of the city). The primary goal of our evaluation was to assess the extent to which we could transfer the scenarios from a simulated environment to the real-world cyber-physical spaces counterparts existing in the same city. From these experiments, we learned several lessons concerning scope and limitations of GANs, for example, that training GANs on cyber-physical spaces is challenging because of the stochastic variance of the real world [31]. More specifically, the virtual world encoded by programmers clashes with the countless variables which need specific approximation systems and assumptions in real-life cyber-physical spaces [31].

The contributions of this paper are:

(1) a computer vision framework to simulate terrorist attacks in other locations;
(2) a proof-of-concept featuring StyleGAN to implement such framework;
(3) an initial evaluation of the StyleGAN architecture [25] in the context of terrorist attack simulation.

The remaining of the paper is organized as follows. Section 2 presents the technologies used in this paper. Section 3 outlines the problem definition, the framework, and the data required by our approach. Section 4 reports on the previously introduced proof-of-concept, while Section Section 5 concludes the paper.

2 BACKGROUND

In the last few years, computer vision—namely, the interdisciplinary study of algorithms that train models to gain high-level understanding from digital images or videos—has received increasing attention given the multitude of application fields where it can play a role [36]. Indeed, computer vision is a branch of Artificial Intelligence (AI), which focuses on partially replicating the complexity of human vision systems to enable computers to identify and process objects in images and videos in the same way that humans do. In this research paper, we present a computer vision approach to give to municipalities and local law enforcement agencies a new tool to prevent and minimize the effects of terrorist attacks. In this section, we describe (i) how artificial intelligence can be used to model criminal scenarios and (ii) generative adversarial networks for computer vision.

2.1 Artificial Intelligence to Model Criminal Scenarios

Conte et al. [8] use Artificial Intelligence to model complex scenarios. To generate data regarding criminal activities, usually multi-agent-based simulations are used [4, 9, 20, 28]. Bosse et al. [4] and Hao et al. [20] identify targets for burglary and other criminal activity. Devia and Weber [9] evaluated the performance of current practices for reduce the resources required throughout a city. More recently, Birks et al. [3] criticized the top-down approach for analyzing crime. They argued that computational models could allow researchers to evaluate the interactions of the different components in a dynamic social systems. Their purpose was to investigate the effect of manipulating one or more components to on the social system. The current state-of-the-art for analyzing terrorism scenarios focuses on risk assessment, which is based on regression models and probability theory [12, 26]. However, this approaches can lead to misleading results as terrorist activities are surrounded by high uncertainty [6]. Currently, there is no successful approach to simulate terrorist scenarios by considering all the different parameters that entail a terrorist attack.

2.2 Generative Adversarial Networks and Computer vision

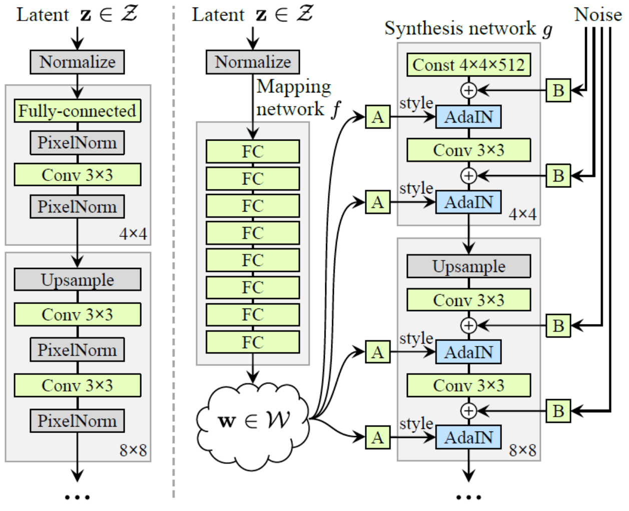
To simulate terrorist scenarios, we rely on Generative Adversarial Networks [18] (GANs). The ability of GANs in generating realistic images has rapidly increased in recent years. The concept of transferring the content of an image to another one has been initially introduced by Gatys et al. [16]. However, their framework has limited practical application [23]. To solve this issue Huang et al. [23] proposed AdaIN, a new approach to enable arbitrary style transfer in real-time. For the same purpose two networks able to generate high-quality images were proposed: BigGANs [5] and StyleGANs [24].

BigGANs have been introduced by Brock et al. [5] and improve GANs in terms of scalability, robustness, and stability. BigGANs are suitable to encode diverse pictures by isolating the different aspects of the images. This approach is suitable for modelling complex scenarios such as terrorist attacks. However, the network size prevented us to apply this network due to the huge computing power needed to train the model.

[1] https://www.rockstargames.com/V/
[2] https://www.openstreetmap.org/

Table 1: Parameters needed to simulate a counterterrorism event.

Environment	
Location	3D composition of the location of the main site, including all the buildings and access areas
# of Entries	number of places that people are able to enter the main site
# of Exits	number of places that people are able to exit the main site
# Access Control	number of secured entry points
# of Barriers	number of access points with barricades (including natural ones) to protect against the attack
Surveillance	if there are surveillance systems in place for early alerting against an attack
Event	
Activity	the type of event that is taking place
Attack	the type of attack
Crowd Density	how many people per square meter will be at the event
Agents	
Type	citizens = people attending the event or which are in close proximity, these are the targets of terrorists terrorist = a criminal attack on citizens, with the intent of inflicting death or serious bodily injury police = police officers (either on foot or in a vehicle) which are responsible for protecting citizens fire and rescue = firefighters/rescue personal which are part of emergency response health = ambulances and hospitals which are responsible for emergency response in case when people are injured
Groups	groups of agents (e.g., police agents or families) should be together
Speed	the speed of the agent
Vision Radius	the distance in which an agent is able to detect another agent
Route	route that the agent follows
Response Time	the response time of agent



(a) Traditional (b) Style-based generator

Figure 1: Difference a traditional and a style-based generator architecture from Karras et al. [24].

For this reason, we relied on StyleGANs, a new architecture proposed by Karras et al. [24] that combines GANs and AdaIN and was successfully applied to other domains than paintings [15].

Figure 1 depicts the differences between the style-based generator architecture and the traditional GANs. StyleGAN extends the progressive training with a mapping network with to goal of encoding the input into a feature vector whose elements control different visual features and styles that translate the previous vector into its visual representation. By using separate feature vectors for each level, the model can combine multiple features. For example, given two images, the model can use coarse granularity features from the first, fine granularity features from the second to generate a third image that combines the two. With respect to other equally effective GANs (e.g., BigGANs [5]), this model requires less training time to produce high-quality realistic-looking images [38].

3 TOWARDS AL$_{Ter}$: COMPUTER VISION FOR CONTERTERRORISM SIMULATION

As previously mentioned, this paper elaborates ALTer, a framework to describe the elements needed to simulate the protection of a public cyber-physical space against terrorism. Previous work already established that due to the high dimensionality of the involved variables this problem is complex [30]. In particular, we describe the *data* that should feed the *simulation* implemented by Generative Adversarial Networks.

3.1 Framework Parameters

Based on previous work [4, 28], we analyzed the factors that should be considered in the simulation of counterterrorism on public spaces. Table 1 lists the parameters we elicited from the state of the art grouped in categories namely: (i) the *environment* where the terrorist attack happen, (ii) the *event* happening in the environment, and (iii) the *agents* that move in the environment.

Environment. The environment reflects the place where the event takes place. It entails the entire cyber-physical terrain, the security measures, and the cyber-tech which are in place and interacting with the terrain itself (e.g., fixed-cams, drones, aerial recognition intelligence, etc.). Most of such data can be accurately gathered using online data. However, it is worth considering that terrorist attacks can come from everywhere (e.g., from the sky, buildings etc.) Therefore, a two-dimensional representation is useless while a more complex three-dimensional representation should be adopted.

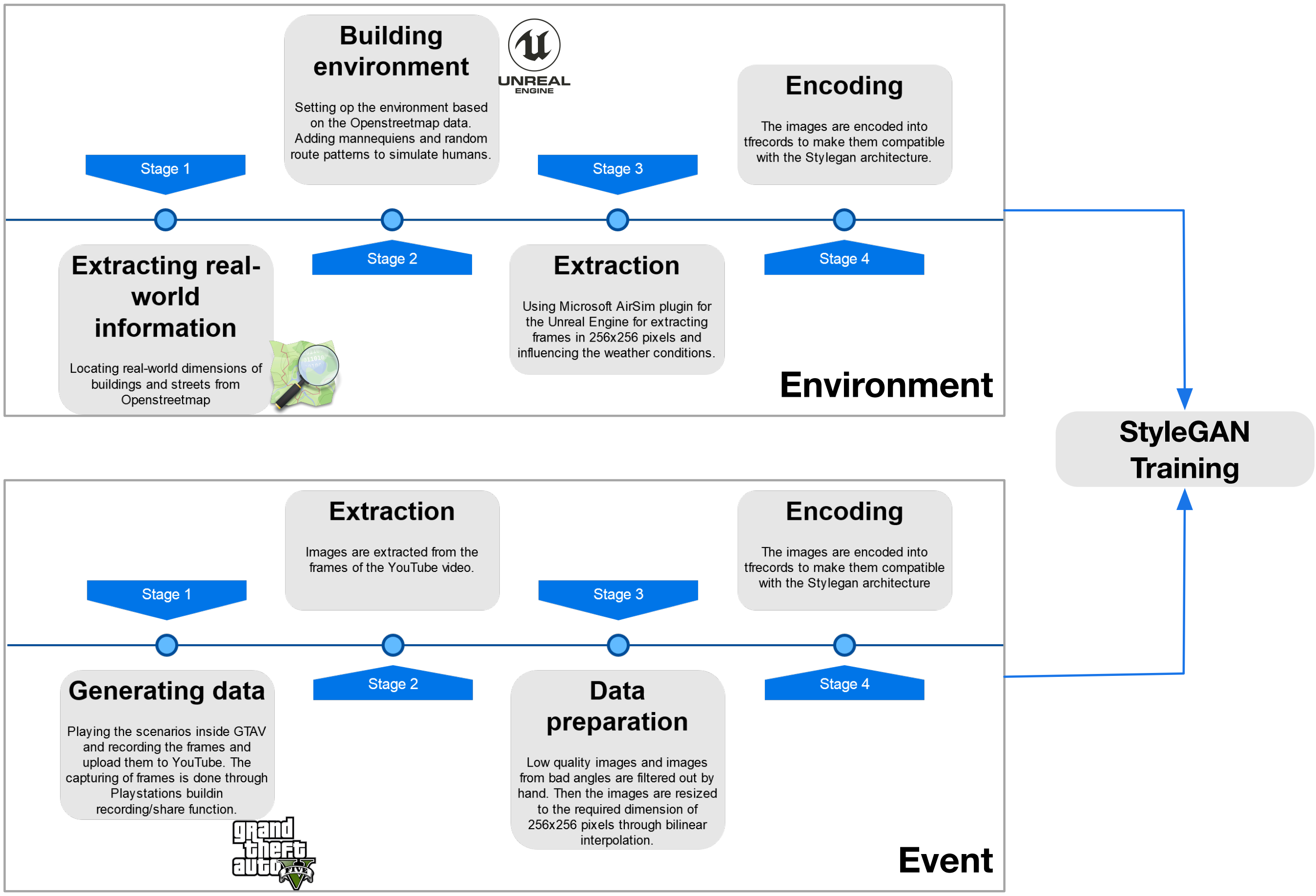

Building environment
Setting op the environment based on the Openstreetmap data. Adding mannequiens and random route patterns to simulate humans.
UNREAL ENGINE
Encoding
The images are encoded into tfrecords to make them compatible with the Stylegan architecture.
Stage 1
Stage 3
Stage 2
Stage 4
Extracting real-world information
Locating real-world dimensions of buildings and streets from Openstreetmap
Extraction
Using Microsoft AirSim plugin for the Unreal Engine for extracting frames in 256x256 pixels and influencing the weather conditions.
Environment
StyleGAN Training
Extraction
Images are extracted from the frames of the YouTube video.
Encoding
The images are encoded into tfrecords to make them compatible with the Stylegan architecture
Stage 1
Stage 3
Stage 2
Stage 4
Generating data
Playing the scenarios inside GTAV and recording the frames and upload them to YouTube. The capturing of frames is done through Playstations buildin recording/share function.
Data preparation
Low quality images and images from bad angles are filtered out by hand. Then the images are resized to the required dimension of 256x256 pixels through bilinear interpolation.
Event

Figure 2: AL_{Ter} Pipeline; extracting environments (top), events (bottom), and train the Generative Adversarial Network.

Event. The event describe the type of terrorist scenario, the activity that was taking place during the attack, and the crowd density. For example, when a concert is playing, people might have loud music which would prevent people hearing/seeing when something happens but digital technology can be present to track and direct people towards specific event location areas (e.g., emergency exits).

Agents. Agents are all the different individuals that we consider in the simulation. As described in Table 1, there are five types of agents. A single scenario should be composed at least of (i) the citizens who are in the environment, (ii) the terrorist who is are attacking the main site, and (iii) the emergency services (i.e., police, fire and rescue, and health). The main difference between the emergency services and the other agents is that the former have a response time before acting if they are not at the main site as they need to be notified first. For the terrorist, we created another subclass which defines the attack they are committing, the kind of weapon they are wielding. Each agent has a certain speed that for example could be determined by the means of transportation. Regarding the route, it generally predetermined as Malleson and Birkin [28] already indicated, as police generally have a certain route and citizens, as well as terrorists, have a goal which they work towards.

3.2 Using Generative Adversarial Networks to Protect Cyber-Physical Spaces

After defining the framework parameter, we need to successfully train the Generative Adversarial Network to encode the world inside the network. With this respect Bau et al. [2] showed that it is possible to map specific components of images back to causal units inside the layers of a GAN. They used segmentation masks to effectively distinguish between trees or humans and map them back to the different layers of the network. They also showed that by adjusting the network they could force it to generate a door an place it on a building. Once the GAN has been successfully trained, the next step is to reverse the GAN to map the image features back to their variables in their latent representation. Donahue et al. [10] and Dumoulin et al. [11] proposed an encoder to map the created images back to their representation in the latent space. Alternatively, Luo et al. [27] introduced inverse mapping and reformulated the problem as a minimization problem. For sake of simplicity, please consider that neural networks have three basic components:

(i) input, (ii) parameters, and (iii) output of the network. In our framework, we keep the input and output fixed, and we map the parameters through backpropagation from the input to output differently from Luo et al. [27] who keep the output and parameters fixed and map the output images back to the latent input vector. To implement scalable GAN, we could rely on BigGANs [5] and StyleGANs [24]. However previous research [1, 14, 33] showed that StyleGAN are more suitable for large datasets, hence motivating our choice towards such networks.

4 PROOF OF CONCEPT

An initial part of framework defined above has been implemented in a proof of concept, which we describe and preliminary evaluated in the following sections. The goal of this proof of concept is to demonstrate that it is possible to use Generative Adversarial Networks to transpose an *event* in a different *environment*.

4.1 Prototype Implementation

Figure 2 shows the AL_{Ter} pipeline to extract the data needed to feed and train the Generative Adversarial Network: the environment and the event.

4.1.1 Retrieving the Environment. To retrieve the data concerning the environment, we collect and assemble multiple images that represent real-world sites, i.e., the environment on which our approach should be applied to predict criminal activities. To this aim, we create a low-quality representation of the environment using the Unreal engine[3]. We start with the lowest level of detail by using the Unreal Engine through which we can simulate a real-world environment as a set of blocks. To this end, we first acquire the information about the building size and dimensions from Openstreetmap[4], an open-source project whose aim it to create and distribute editable maps of the whole world. Afterwards, to simulate the presence of humans, we add some basic mannequins that walk around although their interactions are extremely basic, they represent typical real-world interactions. Examples of these simulations can be found in the four images in the centre of Figure 3, which represent squares in the city of Malaga (Spain). One of the drawback of using OpenStreetMap is that it is based on volunteers, thus data could be incomplete. In the third step, frames of 256x256 pixels are extracted using the Microsoft AirSim plugin [32]. Finally, these frames are encoded within the StyleGan architecture, i.e., they represent the first input required for AL_{Ter}, our computer vision approach to simulate counterterrorism events.

4.1.2 Generating and Collecting Events. The second step toward this proof of concept concerns the events. As explained in the previous sections, to the best of our knowledge there are no publicly available sources reporting terrorist scenarios that can be used without violating privacy constraints. As such, we have to find an alternative through the creation of a synthetic dataset. In particular, we exploit GTAV and collect data coming from several gameplay sessions recorded from different perspectives. Overall, we collect around $10,000$ images which have elements involving terrorist attacks from multiple views which even includes a helicopter filming a truck ramming into people/buildings. For testing whether the network is able to distinguish the features of terrorist attacks, we use another dataset which contains sequences of urban spaces of GTAV and was previously employed by Huang et al. [22]. This dataset provide to us unseen images that we could use in the interpolation operations to test the disentanglement. Figure 2 shows the steps required to extract the data. After playing the scenarios and recording them by using the built-in capabilities of Playstation (stage 1), we uploaded them on YouTube: this allowed us to extract the video frames (stage 2) Then, we analyze them with the aim of discarding low-quality images as well as images from bad angles (stage 3). Therefore, we could clean the dataset from noise and resize the remaining images to the required dimension (256x256 pixels). These images represent the second input for the StyleGAN architecture i.e., they represent the second input required for AL_{Ter}, our computer vision approach to simulate counterterrorism events.

4.1.3 Training the Generative Adversarial Network. To model our *simulation*, we rely on the StyleGAN architecture. We consider this model as a black box as we do not have control over how it learns the elements of the scenario. As for the loss function, apriori we do know which one works best for a two-fold reason: there is no previous research which encoded real world images to synthetic dataset and previous experiments were not in the context of terrorist scenarios. Therefore, we considered a combination of several loss functions considered in the state-of-the-art that are depicted in Table 2: (i) Pixel-wise loss, (ii) MS-SSIM loss, (iii) LPSIS loss, and (iv) VGG loss.

To converge, a GAN requires that the output of the loss function for both the generator and the discriminator are around -0.5 and 0.5 in terms of Mean Squared Error (MSE). During the training phase, we observed large problems with the network convergence for both the discriminator and the generator. Generally, we observed that training low resolution layers is easier than training higher quality layers. An example is provided in Figure 4, where the Y-axis is the loss and the X-axis is the amount of steps the network trained. Here the loss gradually moves towards the targeted loss until the network switches to an higher resolution: from that point the loss moves in the opposite direction.

We analyzed the reasons behind the issue and we discovered that the network is not able to converge due to the diversity of the images in the event dataset. Hence we decided to train the network using only the images from the dataset by Huang et al. [22]. We observed that when increasing the quality of the images, the training session tends to break down as well. Finally, we retrained the network on a dataset created from the Unreal engine in which some mannequins representing people were running around the environment. To simulate an attack, we included a fire in a random location of the square. The images of this dataset were less diverse and generally of the same shape. Furthermore, the initial location was encoded as block. To diversify the dataset, we included several images of squares in the city of New York as well. The network did converge smoothly to the target loss of -0.5/0.5 as is shown by the loss graphs in Figure 5.

[3] https://www.unrealengine.com/

[4] https://www.openstreetmap.org/

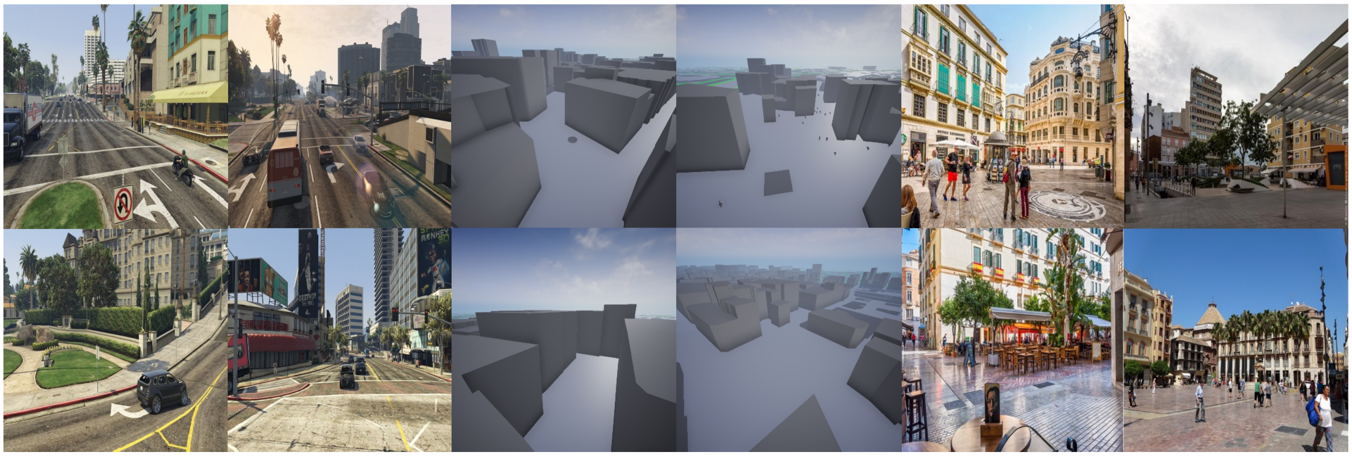


Figure 3: GTAV urban spaces images (left), skeleton representations Malaga (centre), real world locations Malaga (right). For the real-world locations, we choose 4 squares in the city centre of Malaga: Plaza de FéLix Sáenz, Plaza Enrique Garcia-Herrera, Plaza de Las Flores, and Plaza de la Constitución.

Table 2: The loss functions used to train the StyleGAN.

Loss function	Formula
Pixel-wise loss	$L_{Pixel-wise} = \frac{1}{N} * log(cosh(G(w) - I))$
	N is the number of scalars (height, width and channel in our cases; $N = hxwx3$) $G(.)$ is the generator you trained w^* is the input latent code you try to optimize I is the target image you trained Note: $(G(w)$ is the generated image by the generator with the latent code w
MS-SSIM loss	$L_{ms-ssim} = [l_m(x, y)]^{\alpha M} * \prod_{j=1}^{M} [c_j(G(w), I)^{\beta_j}][s_j(G(w), I)]^{\gamma_j}$
	- $l(x, y) = \frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1}$ - $c(x, y) = \frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2}$ - $s(x, y) = \frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3}$ $x = G(w^*)$, $y = I$ represent two images that you compare
LPSIS loss	$L_{lpsis} = \sum_{j=1}^{l} \frac{1}{N_j} \|F_j(G(w) - F_j(I)\|_2^2$
	l is the number of layers j is the j-th layer of the network N_j is the number of scalars in the feature map of the j-th layer
VGG loss	$L_{VGG} = \frac{1}{N_j} \lvert F_j(G(w)) - F_j(I) \rvert$
	j is the j-th layer of the network N_j the number of scale in the feature map of the j-th layer

4.2 Limitations and Lessons Learned

This section describes the limitations and the lesson learned from our initial implementation.

Feature Spatial Location. The preliminary observations on the results of the proof of concept that not always the network is able to effectively disentangle all features within the latent code. Current StylegGAN applications enroll datasets which contain images whose features have similar spatial location. For example, a face can have different forms (e.g., round or diamond shape) but yet the locations of the eyes, mouth and ears are generally on the same place. In our case the locations of buildings, streets, cars, or people moving around vary considerable. This constitues an additional challenges for the Generative Adversarial Networks=.

Alternatives to StyleGANs. The network is not able to cope with high diversity. StyleGANs seem not being able to isolate the different aspects of the images. Indeed, as outlined in Section 2, one of the drawbacks of the StyleGANs is the diversity they can encode; Therefore, BigGANs [5] might be a better choice for modelling something as complex and diverse as terrorist scenarios because of the size of the network, but such size could also prevent to correctly

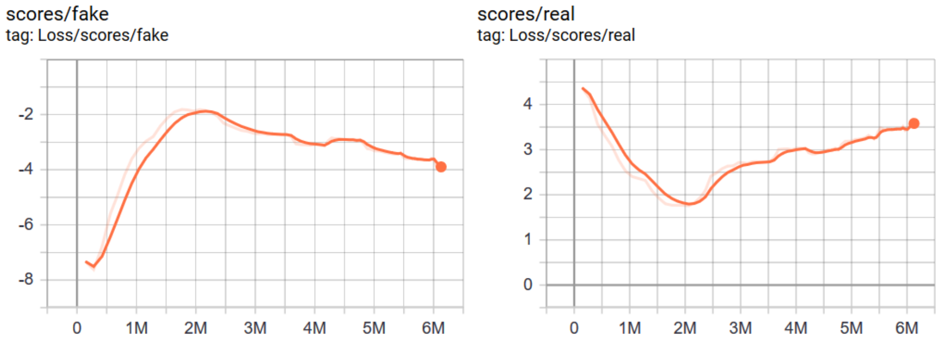


Figure 4: Loss progress while training on the events generated with GTA V with a resolution of 256px.

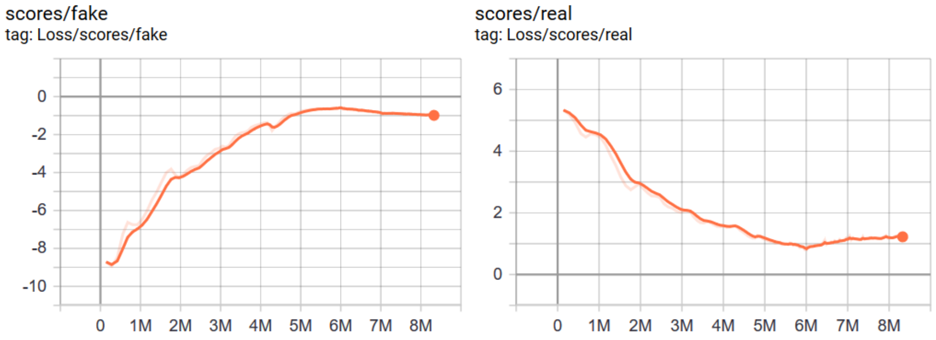


Figure 5: Loss progress while training on the events generated with the Unreal Engine.

reverse engineering the features to the latent space. However, BigGANs could be a suitable alternative that we will analyse as part of our future agenda.

Loss Functions. We could conclude that StyleGANs are not suitable to model complex simulations such as terrorist attack. Indeed, the StyleGAN was not able to successfully encode the images. However, we also showed that currently there is no perceptual loss function which can be used to effectively encode the urban spaces in the latent space. Therefore, future efforts should be devote in experimenting and analysing different loss functions to be used for real-world scenarios.

5 CONCLUSIONS

In this paper we introduced AL_{Ter}, a framework for Law-Enforcement Agencies to simulate real-life terrorist attack scenarios in their own cities to support a number of counter-terrorism and crime-fighting activities, e.g., a better-intrumented preparedness plan to counter terrorist attacks. Our approach and tool features a novel way for simulating terrorism scenarios by using computer vision and GANs, that is, generative adversarial neural networks that learn automatically the features of terrorist scenarios and are able to simulate such features in the scope of real-life cyber-physical systems and spaces.

Our proof-of-concept shows we were able to simulate real-life cyber-physicial spaces as subject to terrorist threats, but we were not able to do so effectively and with a low margin of error. However, our proof-of-concept shows that the future of this approach is promising and further experimentation is needed, possibly with different architectures, more data and better quality data as well as data-fusion approaches between real and simulated data. Furthermore, as part of this study, we show a new way of generating data for terrorism scenarios based on video-gaming, thus addressing the lack of data and the sensitivity criticalities around terrorism data.

To conclude, we introduced AL_{Ter} a framework to use GANs for simulation of terrorist threats in the context of cyber-physical spaces. In the context of AL_{Ter}, we trained GANs on multiple datasets, varying their diversity significantly. We evaluated in total four different datasets: (a) images directly reflecting real terrorist scenarios; (b) on a dataset that combines the images of urban space and terrorist scenarios where the top view was filtered out; (c) on a dataset containing first-person view of urban spaces, e.g., the view of a law-enforcement agent occurring on the terror scene; and (d) on a dataset with skeletal representations of the real world with a minimum amount of detail. Results show that GANs are able to converge on the dataset with the lowest amount of detail—i.e., dataset (d) from the enumeration above—but even for this the accuracy shows evident traces of overfitting. This makes us conclude that: (1) the Stylegan architecture may not be capable of modelling the complexity which comes with simulating terrorist or any other real-life scenario; (2) the data available might be poor quality or too little to offer conclusive learning evidence to GANs; (3) the experimental setup is still too preliminary and requires further work.

In the future we plan to address the aforementioned gaps with further more concrete and larger-scale and scope data over both simulated and real-life terrorist scenarios. On one hand, we aim at testing further architectures for GANs underlying the core of AL_{Ter}. On the other hand, we plan to apply data-fusion approaches which could, in principle, add multiple and diverse sources of data to compound the overfitting risks of GANs themselves. Finally, we plan to blend the AL_{Ter} framework with trusted-computing technologies [7] such that its operations may be made more explainable and its results more traceable.

REFERENCES

[1] Rameen Abdal, Yipeng Qin, and Peter Wonka. 2019. Image2StyleGAN: How to Embed Images Into the StyleGAN Latent Space? arXiv:cs.CV/1904.03189

[2] David Bau, Jun-Yan Zhu, Hendrik Strobelt, Bolei Zhou, Joshua B. Tenenbaum, William T. Freeman, and Antonio Torralba. 2018. GAN Dissection: Visualizing and Understanding Generative Adversarial Networks. arXiv:cs.CV/1811.10597

[3] Daniel Birks, Michael Townsley, and Anna Stewart. 2012. Generative explanations of crime: Using simulation to test criminological theory. *Criminology* 50, 1 (2012), 221–254.

[4] Tibor Bosse and Charlotte Gerritsen. 2009. Comparing crime prevention strategies by agent-based simulation. In *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, Vol. 2. IEEE, 491–496.

[5] Andrew Brock, Jeff Donahue, and Karen Simonyan. 2018. Large Scale GAN Training for High Fidelity Natural Image Synthesis. arXiv:cs.LG/1809.11096

[6] Gerald G Brown and Louis Anthony Cox, Jr. 2011. How probabilistic risk assessment can mislead terrorism risk analysts. *Risk Analysis: An International Journal* 31, 2 (2011), 196–204.

[7] Bert-Jan Butijn, Damian A. Tamburri, and Willem-Jan Van Den Heuvel. 2019. Blockchains: a Systematic Multivocal Literature Review. arXiv:cs.CR/1911.11770

[8] Rosaria Conte, Rainer Hegselmann, and Pietro Terna. 2013. *Simulating social phenomena*. Vol. 456. Springer Science & Business Media.

[9] Nelson Devia and Richard Weber. 2013. Generating crime data using agent-based simulation. *Computers, Environment and Urban Systems* 42 (2013), 26 – 41. https://doi.org/10.1016/j.compenvurbsys.2013.09.001

[10] Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. 2016. Adversarial feature learning. *arXiv preprint arXiv:1605.09782* (2016).

[11] Vincent Dumoulin, Ishmael Belghazi, Ben Poole, Olivier Mastropietro, Alex Lamb, Martin Arjovsky, and Aaron Courville. 2016. Adversarially learned inference. *arXiv preprint arXiv:1606.00704* (2016).

[12] Barry Charles Ezell, Steven P Bennett, Detlof Von Winterfeldt, John Sokolowski, and Andrew J Collins. 2010. Probabilistic risk analysis and terrorism risk. *Risk Analysis: An International Journal* 30, 4 (2010), 575–589.

[13] Stanley Friedman. 1982. Contingency and disaster planning. *Computers & Security* 1, 1 (1982), 34–40. http://dblp.uni-trier.de/db/journals/compsec/compsec1.html#Friedman82

[14] Aviv Gabbay and Yedid Hoshen. 2019. Style Generator Inversion for Image Enhancement and Animation. *arXiv preprint arXiv:1906.11880* (2019).

[15] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2015. A Neural Algorithm of Artistic Style. *CoRR* abs/1508.06576 (2015). http://arxiv.org/abs/1508.06576

[16] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2016. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2414–2423.

[17] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 2672–2680. http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[18] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Networks. http://arxiv.org/abs/1406.2661 cite arxiv:1406.2661.

[19] J. Hamari, J. Koivisto, and H. Sarsa. 2014. Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In *2014 47th Hawaii International Conference on System Sciences*. 3025–3034. https://doi.org/10.1109/HICSS.2014.377

[20] Mengmeng Hao, Dong Jiang, Fangyu Ding, Jingying Fu, and Shuai Chen. 2019. Simulating Spatio-Temporal Patterns of Terrorism Incidents on the Indochina Peninsula with GIS and the Random Forest Method. *ISPRS International Journal of Geo-Information* 8, 3 (2019), 133.

[21] Isaias Hoyos, Bruno Esposito, and Miguel Núñez del Prado. 2018. DETECTOR: Automatic Detection System for Terrorist Attack Trajectories.. In *SIMBig (Communications in Computer and Information Science)*, Juan Antonio Lossio-Ventura, Denisse Muñante, and Hugo Alatrista-Salas (Eds.), Vol. 898. Springer, 160–173. http://dblp.uni-trier.de/db/conf/simbig/simbig2018.html#HoyosEN18

[22] Po-Han Huang, Kevin Matzen, Johannes Kopf, Narendra Ahuja, and Jia-Bin Huang. 2018. DeepMVS: Learning Multi-View Stereopsis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[23] Xun Huang and Serge Belongie. 2017. Arbitrary Style Transfer in Real-Time With Adaptive Instance Normalization. In *The IEEE International Conference on Computer Vision (ICCV)*.

[24] Tero Karras, Samuli Laine, and Timo Aila. 2019. A Style-Based Generator Architecture for Generative Adversarial Networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. https://github.com/NVlabs/stylegan

[25] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2019. Analyzing and Improving the Image Quality of StyleGAN. *CoRR* abs/1912.04958 (2019). http://dblp.uni-trier.de/db/journals/corr/corr1912.html#abs-1912-04958

[26] Jingyu Liu, Walter W Piegorsch, A Grant Schissler, and Susan L Cutter. 2018. Autologistic models for benchmark risk or vulnerability assessment of urban terrorism outcomes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 181, 3 (2018), 803–823.

[27] Junyu Luo, Yong Xu, Chenwei Tang, and Jiancheng Lv. 2017. Learning inverse mapping by autoencoder based generative adversarial nets. In *International Conference on Neural Information Processing*. Springer, 207–216.

[28] Nick Malleson and Mark Birkin. 2012. Analysis of crime patterns through the integration of an agent-based model and a population microsimulation. *Computers, Environment and Urban Systems* 36, 6 (2012), 551 – 561. https://doi.org/10.1016/j.compenvurbsys.2012.04.003 Special Issue: Advances in Geocomputation.

[29] Vivek Menon, Bharat Jayaraman, and Venu Govindaraju. 2011. The Three Rs of Cyberphysical Spaces. *IEEE Computer* 44, 9 (2011), 73–79. http://dblp.uni-trier.de/db/journals/computer/computer44.html#MenonJG11

[30] Il-Chul Moon and Kathleen M. Carley. 2007. Modeling and Simulating Terrorist Networks in Social and Geospatial Dimensions. *IEEE Intelligent Systems* 22, 5 (2007), 40–49. http://dblp.uni-trier.de/db/journals/expert/expert22.html#MoonC07

[31] Hideyuki Nakanishi, Toru Ishida, and Satoshi Koizumi. 2009. Virtual Cities for Simulating Smart Urban Public Spaces. In *Handbook of Research on Urban Informatics*, Marcus Foth (Ed.). IGI Global, 257–269. http://dblp.uni-trier.de/db/books/collections/FO2009.html#Nakanishi0K09

[32] Shital Shah, Debadeepta Dey, Chris Lovett, and Ashish Kapoor. 2017. AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles. In *Field and Service Robotics*. arXiv:arXiv:1705.05065 https://arxiv.org/abs/1705.05065

[33] Yujun Shen, Jinjin Gu, Xiaoou Tang, and Bolei Zhou. 2019. Interpreting the latent space of gans for semantic face editing. *arXiv preprint arXiv:1907.10786* (2019).

[34] Zbigniew Skolicki, Tomasz Arciszewski, M. H. Houck, and Kenneth A. De Jong. 2008. Co-evolution of terrorist and security scenarios for water distribution systems. *Advances in Engineering Software* 39, 10 (2008), 801–811. http://dblp.uni-trier.de/db/journals/aes/aes39.html#SkolickiAHJ08

[35] Dimitris Spiliotopoulos, Costas Vassilakis, and Dionisis Margaris. 2019. Data-driven country safety monitoring terrorist attack prediction.. In *ASONAM*, Francesca Spezzano, Wei Chen, and Xiaokui Xiao (Eds.). ACM, 1128–1135. http://dblp.uni-trier.de/db/conf/asunam/asonam2019.html#SpiliotopoulosV19

[36] Richard Szeliski. 2011. Computer vision algorithms and applications. http://dx.doi.org/10.1007/978-1-84882-935-0

[37] Damian A Tamburri. 2019. Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation. *Information Systems* (2019), 101469.

[38] Kayo Yin. 2019. How to Train StyleGAN to Generate Realistic Faces. https://bit.ly/2FQ0CU7.