# BACHELOR’S THESIS

Warsaw University of Technology

Faculty of Electronics and Information Technology

Department of Computer Science

**Generating new spiders using Generative Adversarial Network (GAN)**

Valery Burau

# Under The Supervision Of

**Prof. dr hab. Andrzej Pacut**

**Table of contents**

[1. Goal & main assumptions of the thesis 3](#_Description_of_file)

[2. Main purposes of the dissertation 3](#_Description_of_file)

[3. Usage of expected results 3](#_Description_of_file)

[4. Abbreviations 3](#_Description_of_file)

[5. Selected model: introduction to GAN 3](#_Description_of_file)

[5.1 Basic idea behind GANs 3](#_Toc45882527)

[5.2 Characteristics 3](#_Toc45882527)

[5.3 Usage 3](#_Toc45882527)

[5.4 Steps to train a GAN 3](#_Toc45882527)

[5.5 Types of GANs 3](#_Toc45882527)

[6. Libraries, data & preprocessing part 3](#_Description_of_file)

[1. Optimization functions 3](#_Description_of_file)

[1. Optimization functions 3](#_Description_of_file)

[1. Optimization functions 3](#_Description_of_file)

[2. Activation functions 3](#_Description_of_file)

[3. Literature 3](#_Description_of_file)

[4. Bibliography 3](#_Description_of_file)

# **Goal & main assumptions of the thesis**

The main goal of the thesis is to create images of fake spiders using artificial intelligence.

Main assumptions:

* Create the model that can be used for generating new spiders.
* The newly created spiders should look as realistic as possible.
* The new generated spiders should look like a new kind of spiders or similar to a training one. Due to about 7 different spider species in the training dataset it can be created a new 'mixed' version (=kind) of spider.

# **Main purposes of the dissertation**

Some aspects should be described in the documentation:

1. Ideas behind the used model for spider generation.
2. Methods for improving the quality of an object in an image.
3. Mathematical background of used algorithms and structures.
4. Achieved results & conclusions.

# **Usage of expected results**

Necessary assumption:

* Spider generation refers to a broader issue: generating new images for different objects (e.g. vehicles, animals, buildings, etc.).

Generation of new images can be used in:

* Creating new videos (different locations, different objects in any position). Moreover, the rights of the copyright holder will not be violated.
* Creating a new kind of animal, a new type of transport, furniture (for example, a new car model), an unusual view of building.
* Creating a large database for a particular object.

# **Abbreviations**

AI - Artificial Intelligence

NN – Neural Network

GAN – Generative Adversarial Network

# **Selected model: Introduction to GAN**

Artificial intelligence gains great popularity and necessity every day. Lots of industries start to incorporate AI into their projects. Currently, neural networks are used for various tasks such as classification, prediction, precision, video and image generation, etc. Generating new images is one of aforementioned problems. GAN is regarded as the best and most popular model for this kind of task. It's the reason why this model has been used in my dissertation.

Generative Adversarial Network was invented by Ian Goodfellow and his colleagues in 2014.

## **5.1 Basic idea behind GANs**

There are two most important parts while creating this type of network:

* Generator: replicates real data to produce fake data.
* Discriminator: distinguishes real data from fake data.

In this way generator and discriminator are playing mini-max game.

## 

## **5.2 Characteristics** [[1]](https://en.wikipedia.org/wiki/Generative_adversarial_network#cite_note-GANnips-1)

* Class of machine learning systems.
* Given a training set, the technique learns to generate new data with the same statistics as the training set.
* The *generative network* generates candidates while the *discriminative network* evaluates them.
* The generator trains based on whether it succeeds in fooling the discriminator.
* The generator is typically a deconvolutional neural network, and the discriminator is a [convolutional neural network](https://en.wikipedia.org/wiki/Convolutional_neural_network).

## **5.3 Usage**[[1]](https://en.wikipedia.org/wiki/Generative_adversarial_network#cite_note-GANnips-1)

* Creating fake and realistic photos (e.g. [*portraits, landscapes and album covers*](https://www.artbreeder.com/)*)* and videos.
* Increasing amount of data.
* Precision of some images (e.g. [*improve*](https://en.wikipedia.org/wiki/Image_restoration)*ment of* [*astronomical images*](https://en.wikipedia.org/wiki/Astrophotography) *and simulation gravitational lensing for dark matter research*).
* [Reconstruction 3D models of objects from images](https://en.wikipedia.org/wiki/3D_reconstruction_from_multiple_images) and model patterns of motion in video.
* Aging face photographs to show how an individual's appearance might change with age.
* Visualization of the effect that climate change will have on specific houses.
* Model called Speech2Face can reconstruct an image of a person's face after listening to their voice.
* Generating new ways to solve various problems (e.g. in construction) that people are even not able to foresee.
* Generating fake voices.

## 

## **5.4 Steps to train a GAN**[[2]](https://www.youtube.com/watch?v=O8LAi6ksC80&amp%3Blist=PLTl9hO2Oobd-1jxZ01__NjibEY6h15Kha&ab_channel=CodeEmporium)

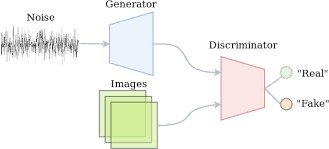
1. Define the problem (e.g. synthesize images from a caption, audio synthesis from sentences and etc.)
2. Define GAN architecture (depends on the complexity of the problem) (multi-layer perceptron, neural network or a simpler model).
3. Train the discriminator to distinguish “real” Vs “fake” data (data labelled as real data, generated data labelled as fake data).
4. Train the generator to synthesize data (parameters of the generative model should be modified to maximize a loss of the discriminator).
5. Repeat steps 3 and 4 N times (N epochs).
6. Final result: the discriminator will not be able to distinguish the real and the fake samples.
7. Once training is complete, synthesize data from the generator.

## **5.5 Types of GANs**

There are many various types of GANs. Each model is usually used to solve a specific problem. The basic idea [[4.1]](#_4.1_Basic_idea) of all GANs is the same, but the structure is different in all models.

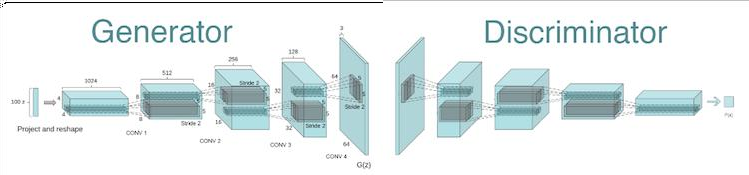
The following popular GAN models are represented below to see the difference in structure:

1. Original Vanilla GAN
   * A representation of original GAN.



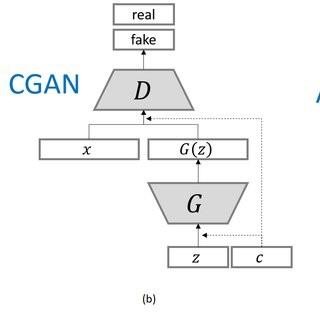
*Picture 1: Original Vanilla GAN* [*(Jos van de Wolfshaar (2018)) Semi-supervised learning with GANs)\_[3]*](https://medium.com/@jos.vandewolfshaar/semi-supervised-learning-with-gans-23255865d0a4)

1. Deep Convolutional GAN (DCGAN)
   * CNNs used in unsupervised learning.
   * Generators are Deconvolutional Neural Networks.
   * Discriminators are CNNs.



*Picture 2: DCGAN* [*(Roman Trusov GAN DEEP LEARNING ARCHITECTURES – REVIEW)\_[4]*](https://sigmoidal.io/beginners-review-of-gan-architectures/)

1. Conditional GAN(CGAN)
   * Dictate the type of data generated through a condition.



*Picture 3: Conditional GAN (CGAN)* [*(Gerasimos Spanakis* (2018)*. “Fig 1- uploaded by Gerasioms (Jerry) Spanakis”) [5]*](https://www.researchgate.net/figure/GAN-conditional-GAN-CGAN-and-auxiliary-classifier-GAN-ACGAN-architectures-where-x_fig1_328494719)

# **Libraries, data & preprocessing part**

# **Literature**

* Goodfellow, Ian; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). [*Generative Adversarial Networks*](https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf) (PDF). Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014). pp. 2672–2680
  + *Salimans, Tim; Goodfellow, Ian; Zaremba, Wojciech; Cheung, Vicki; Radford, Alec; Chen, Xi (2016). "Improved Techniques for Training GANs"*

[*.arXiv:*](https://en.wikipedia.org/wiki/ArXiv)[*1606.03498*](https://arxiv.org/abs/1606.03498) [*[cs.LG*](https://arxiv.org/archive/cs.LG)*].*

* + DCGAN. WRITTEN BY Li Yin[. Mar 8, 2017](https://medium.com/%40liyin2015/dcgan-79af14a1c247?source=post_page-----79af14a1c247----------------------).
  + GANS — PART2: DCGANs (deep convolution GANS) for generating images. Manish Chablani. [Jun 28, 2017](https://towardsdatascience.com/gans-part2-dcgans-deep-convolution-gans-for-generating-images-c5d3c7c3510e?source=post_page-----c5d3c7c3510e----------------------)
  + GAN — DCGAN (Deep convolutional generative adversarial networks). Jonathan Hui. [Jun 18, 2018](https://medium.com/%40jonathan_hui/gan-dcgan-deep-convolutional-generative-adversarial-networks-df855c438f?source=post_page-----df855c438f----------------------)

# **Bibliography**

1. Wikipedia. “[Generative adversarial network](https://en.wikipedia.org/wiki/Generative_adversarial_network#cite_note-GANnips-1)”.
2. YouTube - CodeEmporium (2018). “[Generative Adversarial Networks – FUTURISTIC & FUN AI !](https://www.youtube.com/watch?v=O8LAi6ksC80&amp;list=PLTl9hO2Oobd-1jxZ01__NjibEY6h15Kha)”.
3. Jos van de Wolfshaar (2018). “[Semi-supervised learning with GANs](https://medium.com/%40jos.vandewolfshaar/semi-supervised-learning-with-gans-23255865d0a4)”.
4. Roman Trusov. “[GAN Deep Learning Architectures - review](https://sigmoidal.io/beginners-review-of-gan-architectures/)”.
5. Gerasimos Spanakis (2018). “[Fig 1- uploaded by Gerasioms (Jerry) Spanakis](https://www.researchgate.net/figure/GAN-conditional-GAN-CGAN-and-auxiliary-classifier-GAN-ACGAN-architectures-where-x_fig1_328494719)”.