

Image-to-Image Translation: Machine Learning Magic that Converts Winter Photos Into Summer

A magician can make his trick with just a wave of a magic wand, our engineers can make their magic with just one click! Interested how the same winter landscape would look in summer, then keep on reading. The comprehensive research was completed by our R&D department with the purpose of investigating current possibilities of Machine Learning generators and what quality they can provide for a task of translating winter to summer image as an analogy to automatic language translation. We have conducted the investigation of different generative models and application methods, prepared a solid dataset, and the most suitable model we have adopted to satisfy our requirements. The idea of transitioning the photo taken in winter into a summery image could be applied in a landscape design or related spheres. That would give an opportunity to have a clear vision of the same image in different seasons. The implementation of the same technologies can also enable the various color changes of images.

The secret to our magic is CGAN neural network

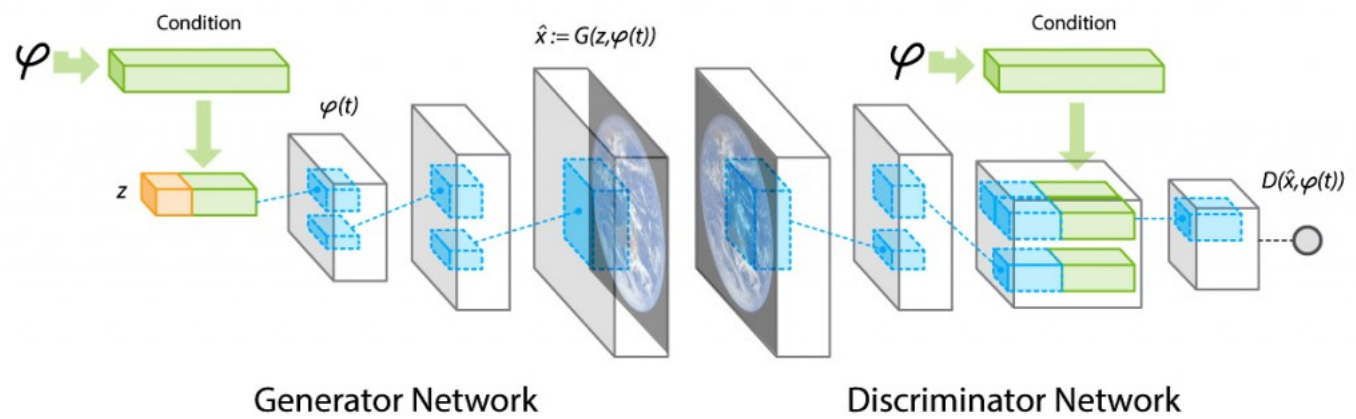
A lot of approaches and processing models were researched, analyzed and tested, i.e. neural networks: CGAN, CVAE; pixel-to-pixel translation. CGAN appeared to be the most suitable for our image-to-image translation task because of the following reasons:

- Sharp & realistic image synthesis
- Semi-supervised learning
- Excellent test of our ability to use high-dimensional complicated data
- Bleeding edge ML technology

Conditional GAN is a subset of GANs – Generative adversarial networks. GAN is a type of generative neural networks that consists of two networks: discriminator – D and generator – G. We can easily understand the key concept behind them by imagining a team of counterfeiters as a G and police as D. Counterfeiters (G) is constantly trying to produce fake currency (images), while police (D) is trying to estimate whether it's real or fake. As time passes both of them are becoming better in their jobs. Conditional GAN has the same principle, but we have some additional info – condition. So now, instead of asking counterfeiters to just produce some fake currency, we ask them to produce, let's say, fake 100\$. Similarly, instead of just asking the police to judge if it's fake or real, we ask them to judge if it's 100\$. It allows us to teach G to generate even more realistic currency (images in our case). Figure 1. cGAN Structure

This site uses cookies to offer you better services. By using our site, you acknowledge that you have read and understand our [Privacy Policy \(/legal/privacy-policy\)](#). Your use of Abtosoftware website is subject to these policies.

ACCEPT



(<https://www.abtosoftware.com/assets/Scheme-1.jpg>)

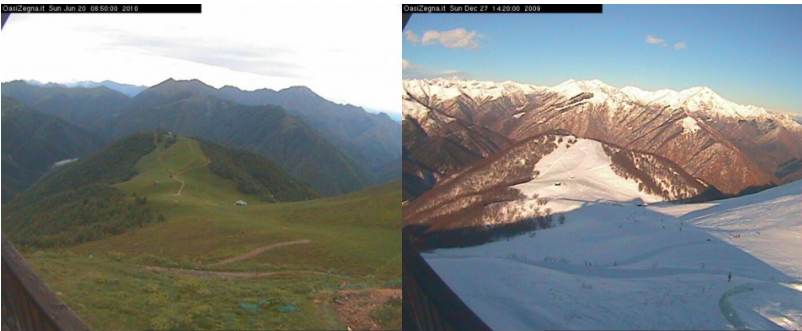
While working on the image-to-image translation of the winter photos, our image processing engineers have faced numerous challenges, like:

- dataset aggregation and preparation;
- long iterations
- the cost of experiments;
- the complexity of CGAN, the configuration of many knobs

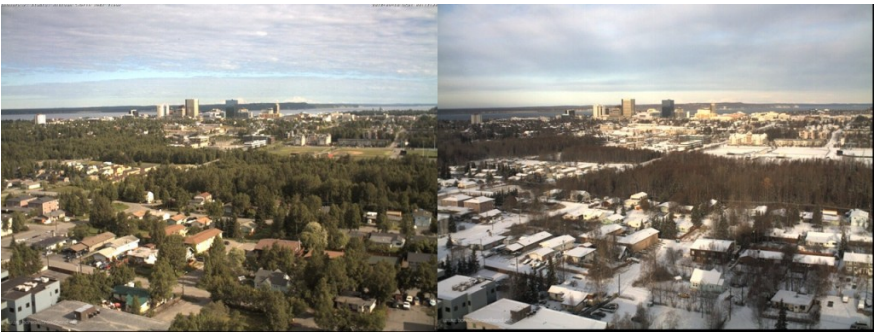
They continued their research on image-to-image translation with the huge and comprehensive dataset of images collection. At first, our R&D team worked with Transient Attribute Dataset. However, due to the lack of images and in order to reach better results and accuracy, the additional image dataset was needed. For this purpose, we selected “Nordland Line” train video footage. The details and samples of both datasets are provided below.

Dataset №1. Annotated photos from 101 webcams with outdoor scenes.

Annotated by people during crowdsourcing campaign (downloaded).



(<https://www.abtosoftware.com/assets/image3.jpg>)



(<https://www.abtosoftware.com/assets/image8.jpg>)

Filtered: picked ones that have “most summer” and “most winter” as pairs (Pandas)

This site uses cookies to offer you better services. By using our site, you acknowledge that you have read and understand our [Privacy Policy](#) ([/legal/privacy-policy](#)).
3000 image pairs of 2 seasons / 640×480 scaled to 256 x 256
Your use of AbtoSoftware website is subject to these policies.

Dataset №2. Four 10-hour Full HD videos of footage recorded from the train during “Nordland Line” trip in Norway in all four seasons.

Time-sync (downloaded)



Cut into frames (FFmpeg)

Auto align best nearby frames (Python + Hugin)

9000 image pairs of 2 seasons / 1000×1000, scaled to 256×256.



(<https://www.abtosoftware.com/assets/image12.jpg>)



(<https://www.abtosoftware.com/assets/image16.jpg>)









(<https://www.abtosoftware.com/assets/image4.jpg>)



(<https://www.abtosoftware.com/assets/image13.jpg>)

Image-to-image translation results

Our engineers have adopted the model to satisfy their requirements and share their best results with you.

Input	Output	Ground truth
		
(https://www.abtosoftware.com/assets/image15.jpg)	(https://www.abtosoftware.com/assets/image9.jpg)	(https://www.abtosoftware.com/assets/image10.jpg)
		
(https://www.abtosoftware.com/assets/image22.jpg)	(https://www.abtosoftware.com/assets/image5.jpg)	(https://www.abtosoftware.com/assets/image11.jpg)

ACCEPT



(<https://www.abtosoftware.com/assets/image2.jpg>)

(<https://www.abtosoftware.com/assets/image14.jpg>)

(<https://www.abtosoftware.com/assets/image14.jpg>)



(<https://www.abtosoftware.com/assets/image6.jpg>)

(<https://www.abtosoftware.com/assets/image20.jpg>)

(<https://www.abtosoftware.com/assets/image20.jpg>)



(<https://www.abtosoftware.com/assets/image17.jpg>)

(<https://www.abtosoftware.com/assets/image1.jpg>)

(<https://www.abtosoftware.com/assets/image1.jpg>)

Training set: 11.500 image pairs, testing set: ~600 image pairs.Here are the best setup details:

- 286×286 -> 256×256 jitter, horizontal mirroring
- Conditional D model
- PatchGAN.

Variations


We would also like to share some samples of variations that happened during the image-to-image translation process. The most typical were the following:

- Dataset manipulations and mixing
- Jittering, random mirroring, amount of epochs
- Unconditional/conditional D model
- PatchGAN/PixelGAN/ImageGAN


Input	No jitter: Same detailed patterns everywhere	Jitter 900×900 -> 256×256: Blurred
-------	---	---------------------------------------

This site uses cookies to offer you better services. By using our site, you acknowledge that you have read and understand our [Privacy Policy \(/legal/privacy-policy\)](#).
Your use of Abtosoftware website is subject to these policies.


ACCEPT




(<https://www.abtosoftware.com/assets/image25.jpg>)




(<https://www.abtosoftware.com/assets/image131.jpg>)




(<https://www.abtosoftware.com/assets/image201.jpg>)




(<https://www.abtosoftware.com/assets/image202.jpg>)




(<https://www.abtosoftware.com/assets/image203.jpg>)




(<https://www.abtosoftware.com/assets/image91.jpg>)




(<https://www.abtosoftware.com/assets/image41.jpg>)



(<https://www.abtosoftware.com/assets/image21.jpg>)




(<https://www.abtosoftware.com/assets/image21.jpg>)




(<https://www.abtosoftware.com/assets/image21.jpg>)

Input

Too much dataset №1:
Mixed illumination, artifacts




(<https://www.abtosoftware.com/assets/image181.jpg>)




(<https://www.abtosoftware.com/assets/image31.jpg>)

Too much dataset №2:
railway everywhere!



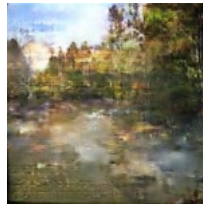
(<https://www.abtosoftware.com/assets/image182.jpg>)



(<https://www.abtosoftware.com/assets/image32.jpg>)



(<https://www.abtosoftware.com/assets/image91.jpg>)



(<https://www.abtosoftware.com/assets/image141.jpg>)(<https://www.abtosoftware.com/assets/image141.jpg>)



(<https://www.abtosoftware.com/assets/image111.jpg>)(<https://www.abtosoftware.com/assets/image111.jpg>)



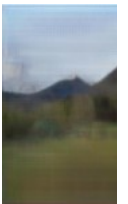
Input

ImageGAN: a bit distorted

PixelGan: Very blurry



(<https://www.abtosoftware.com/assets/image25.jpg>)



(<https://www.abtosoftware.com/assets/image121.jpg>)(<https://www.abtosoftware.com/assets/image121.jpg>)



(<https://www.abtosoftware.com/assets/image91.jpg>)(<https://www.abtosoftware.com/assets/image121.jpg>)(<https://www.abtosoftware.com/assets/image121.jpg>)

Technical Information

Hardware:

AWS p2.xlarge (NVIDIA GK210 GPU)
12Gb GPU memory. CUDA.

Software:

pix2pix – Lua/Torch implementation of cGAN.
Hugin – photo-aligning tool
Python, pandas, numpy.

To get an objective assessment of the output pictures quality, we need a few people as independent experts to compare 2 images and indicate which of them is a real photo, one of the images will be a real photo and another one just a generated picture. Would you like to take part in the experiment? Get in touch with us!

Tools & Technologies: Deep learning, Convolutional neural networks, adversarial networks, Computer vision, lua, torch, OpenCV, Linux, git.

Related Posts

This site uses cookies to offer you better services. By using our site, you acknowledge that you have read and understand our [Privacy Policy \(/legal/privacy-policy\)](#). Your use of Abtosoftware website is subject to these policies.

ACCEPT

How to Automatically Remove Background from Your Videos in Real-time

(<https://www.abtosoftware.com/blog/how-to-automatically-remove-background-from-your-videos>)

Science is the new Black, or Physics and Art in Computer Vision

(<https://www.abtosoftware.com/blog/science-is-the-new-black-or-physics-and-art-in-computer-vision>)

Signal processing. Extracting useful features from electric signal

(<https://www.abtosoftware.com/blog/signal-processing>)

Contact Us

To find out more about Abto Software expertise, request a quote or get a demo of your custom solution.

Full Name

Email Address

Phone Number (optional)

Message

I'm not a robot

reCAPTCHA

Privacy - Terms

ACCEPT

This site uses cookies to offer you better services. By using our site, you acknowledge that you have read and understand our [Privacy Policy \(/legal/privacy-policy\)](#). Your use of Abtosoftware website is subject to these policies.

Clicking this button, I agree to the processing of my personal data.

SEND MESSAGE

New York, US

+1 (650) 250-4555 (tel:1-650-250-4555)

347 Fifth Avenue, Ste 1009, New York, NY
(<https://goo.gl/maps/6UL8YXRZPfN2>)

EMAIL

office@abtosoftware.com (<mailto:office@abtosoftware.com>)

Lviv, Ukraine

+38 097 952 1622 (tel:38-097-952-1622)
+38 032 242 3307 (tel:38-032-242-3307)

77/T77 Heroiv UPA St.,
Lviv, Ukraine, 79015 (<https://goo.gl/maps/W7mz9HBHbSS2>)

FOLLOW US

<https://www.linkedin.com/company/abto-software/>
(<https://www.facebook.com/abtosoftware/>)

Copyright © 2007-2019 Abto Software. All rights reserved.

[Terms of Use \(/legal/terms-of-use\)](/legal/terms-of-use) [Privacy Policy \(/legal/privacy-policy\)](/legal/privacy-policy)

This site uses cookies to offer you better services. By using our site, you acknowledge that you have read and understand our [Privacy Policy \(/legal/privacy-policy\)](/legal/privacy-policy).
Your use of Abtosoftware website is subject to these policies.

ACCEPT