GENERATION OF ARTIFICIAL INFRARED CAMERA IMAGES FOR VISUAL NAVIGATION SIMULATION

Krystian Zyśk[1], Michał Hałoń[2], Kacper Kaczmarek[3], Marcin Kasprzyk[4], Piotr Rodo[5], Olgierd Skromak[6], Mateusz Sochacki[7]

[1]*Students' Space Association, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Nowowiejska 21/25, 00-665 Warsaw, Poland, krystian.zysk.stud@pw.edu.pl*
[2]*Institute of Control and Computation Engineering, Faculty of Electronics and Information Technology, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland, michal.halon@pw.edu.pl*
[3]*Students' Space Association, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Nowowiejska 21/25, 00-665 Warsaw, Poland, kacper.kaczmarek3.stud@pw.edu.pl*
[4]*Division of Automation and Aeronautical Systems, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Nowowiejska 24, 00-665, Warsaw, Poland, marcin.kasprzyk@pw.edu.pl*
[5]*Students' Space Association, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Nowowiejska 21/25, 00-665 Warsaw, Poland, piotr.rodo.stud@pw.edu.pl*
[6]*Students' Space Association, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Nowowiejska 21/25, 00-665 Warsaw, Poland, olgierd.skromak.stud@pw.edu.pl*
[7]*Division of Automation and Aeronautical Systems, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Nowowiejska 24, 00-665, Warsaw, Poland, mateusz.sochacki@pw.edu.pl*

ABSTRACT

This paper describes two methods of generating artificial infrared images based on visual-light spectrum images. The methods were developed as part of the FOK project – rocket guidance, navigation and control research and development platform of the Students' Space Association at the Warsaw University of Technology. The current mission goal of the FOK rocket is to guide the rocket towards ground-based visible or infrared markers. The infrared image generation methods were developed to facilitate software-in-the-loop tests of infrared vision-based navigation during which the rocket's flight is simulated using an in-house developed rocket flight simulation software SKA RFS. During the simulated flight, artificial infrared images are generated and fed into image recognition algorithms, and then the information about the detected target is fed into the rocket's control algorithm. Two methods of artificial infrared image generation, based on available visible-light spectrum images, are described within the paper: one based on a linear correlation between red, green, blue and infrared channels, and the second one based on neural networks.

1 INTRODUCTION

Visual navigation is a cornerstone of many modern navigation systems, from computer mice sensors through cars, guided missiles, spacecraft proximity operations and landers. In particular, infrared-based guidance is an essential feature of guided missiles [1]. Since 2017 the Students' Space Association at the Warsaw University of Technology has been developing the FOK rocket - a guidance,

navigation and control research and development platform [2]. Currently, the team is developing a vision-based navigation system composed of visible-light and infrared seekers with the goal of guiding the rocket towards ground-based visible or infrared markers [3].
As part of the project, the team works on software-in-the-loop tests of developed GNC algorithms. The tests are conducted within the SKA RFS software - an in-house developed tool for simulating a variety of sounding rockets [4]. It is a 6 degrees of freedom flight simulation tool able to simulate multi-staged rockets, equipped with multiple engines, multiple parachutes, and aerodynamic control systems. It also allows for conducting simulations of a video stream from the rocket-mounted camera. As the next iteration of the FOK project aims to use an infrared camera, the team faced the challenge of simulating infrared camera images within the RFS software to test the developed vision navigation algorithms.
This paper presents techniques for generating artificial infrared images for visual navigation simulation. Artificial images are generated based on source images taken in the visible-light spectrum. Two methods for artificial image generation are investigated: the first using a simple correlation between the visible and infrared images, and the second utilising neural networks.
The correlation method compares a set of reference infrared and visible-light spectrum image pairs of the same scenes to determine a correlation between red, green and blue channels and an infrared channel. Several correlation functions are used: simple linear correlation between single visible and infrared pixel intensities and linear correlations of pixel grids of various sizes. The correlation coefficients are determined using the least squares method. Although an assumption of linear correlation between red, green, blue and infrared channels might seem naive, this approach was investigated due to its simplicity.
To implement neural network models capable of generating artificial infrared images, research on existing solutions in similar cases was carried out. The research focused on methods that use paired images for training. Papers found can be divided into two types:
– methods of translating images from outside the visible into the visible spectrum,
– translation of images from the visible to images outside the visible spectrum.
The first category example solution is the SAR-to-Optical Image (SAR - Synthetic Aperture Radar) method [5], which was implemented using modified conditional generative adversarial network (cGAN), whose generator is based on the U-net architecture [6]. 10,000 pairs of SAR/optical images were used for training while 3,000 pairs were used for testing cGAN model.
Solutions belonging to the second category (which also includes the subject of the presented work) include methods for the generation of NIR images from RGB (visual spectrum), which were trained on one of the most popular datasets in this field – Sentinel-2 [7]. These solutions also use mostly cGAN networks with generators based on the U-net structure [8], [9].
Despite the great similarity of these methods to the issue described in this work, it is worth mentioning two major differences:
– different spectrum of infrared radiation (near infrared vs thermal infrared),
– different sets of training data (Sentinel-2 vs smaller - own dataset).
The most commonly used solutions for the problem of translating images between different spectra (which simultaneously use paired datasets for training) are those based on modified GAN network versions [10] (cGAN, cycleGAN, pix2pix).
Solutions providing source code for implemented models allow to easily repeat the presented works or adapt them to one's needs. Unfortunately, only a small part of analyzed papers provide such a feature. Therefore, further work focused on ready-to-use neural cGAN-based network models that could be easily used for training and fine-tuning on own dataset. These are described in detail in Neural network section. A detailed survey of the most important methods and solutions for Image-to-Image Translation can be found in [11].

The source material, pairs of visible and infrared images of the same scene, were acquired using an integrated infrared-visible spectrum camera mounted on an unmanned aerial vehicle (UAV). Images were acquired over relevant scenes – fields, forests, heath – landscapes usually encountered during the prospective flight of the FOK rocket.
The paper concludes with a detailed comparison of the developed methods both in terms of quantitative metrics describing the similarity of artificially generated images and original infrared images, as well as in terms of subjective perception of the generated images.

2 FOK PROJECT

The FOK project was started in 2017 by the Students' Space Association at the Warsaw University of Technology. Its goal was to build a small rocket which would serve as a testbed for control systems and algorithms development [2]. The aerodynamically-controlled rocket has 4 canards and is propelled by a solid rocket motor which enables it to reach velocities up to Mach 0.6. It is 1.2 metres long and weighs just shy of 5 kilograms. Its fuselage is made of aluminium while the nose cone is made of glass fibre composite due to radio signals permittivity. The construction was iteratively improved thanks to 4 flights (see Fig. 1) whose main goal was to gather data on all subsystems' performance, as well as to perform aerodynamic model identification [12].

Figure 1: FOK during one of its launches.

The original design of FOK assumed utilisation of inertial and satellite sensors to execute its missions. The latest version of the rocket, FOK 2, is additionally equipped with a seeker housing a camera placed at the top of the nose cone (see Fig. 2). Its primary objective is to enable the rocket to guide itself towards a ground-based target using proportional navigation algorithms [3], [13], [14]. Two types of imaging nose cones were designed – one with a regular, visible-light-based camera; and another one which will carry a long-wave infrared (LWIR) camera to enable imaging in the infrared spectrum. A series of tests was performed to establish an optimal optical setup of the infrared system [15].

Figure 2: FOK 2 with the new seeker.

3 SOURCE MATERIAL AND PREPROCESSING

The source material used in this research was a set of pairs of infrared and visible-light spectrum images of the same scene. The images were acquired using a DJI Zenmuse XT 2 combined infrared-visible light spectrum camera [16] mounted on Matrice 210 V2 RTK UAV [17]. Images were acquired during several flights, from altitudes up to 120 metres, over grass, buildings, heath, sand and forest areas. Table 1 presents the specifications of Zenmuse XT 2 cameras, while Fig. 3 presents raw images from thermal and infrared cameras.

Table 1: Specification of Zenmuse XT 2 thermal and visible cameras. Source: [16].

	Thermal camera	Visible camera
sensor	Uncooled VOx Microbolometer	CMOS, 1/1.7”
resolution	640×512	4000×3000
pixel pitch	17 µm	1.9 µm (estimated)
spectral band	7.5–13.5 µm	–
focal length	13 mm	8 mm
field of view	45°×37°	57.12°×42.44°

As the field of view of the visible camera was larger than that of the thermal camera, a correlation procedure was performed. The visible light spectrum images were transformed to match the thermal images’ resolution and field of view. The process accounted for radial distortions (as the original visible images seemed not to be properly calibrated), scaling, rotation, translation and cropping. Fig. 4 presents the preprocessed visible light spectrum image, which matches the thermal infrared image (compare it with Fig. 3).

Furthermore, in order to allow for comparison of data acquired at different conditions – particularly thermal infrared images acquired at different ambient temperatures – it was decided to scale the data stored in each image’s pixels to the 0-1 range. For visible light spectrum images, the scaled pixel intensity (for red, green and blue channels) was obtained by dividing the raw pixel intensity (in this case stored as an unsigned 8-bit integer) by 255 (see Eq. 1). For thermal infrared images, first, an aggregated histogram of infrared intensity values was determined (see Fig. 5). Then, based on the

(a) Visible light spectrum image.

(b) Thermal infrared image.

Figure 3: Raw images from Zenmuse XT 2 cameras.

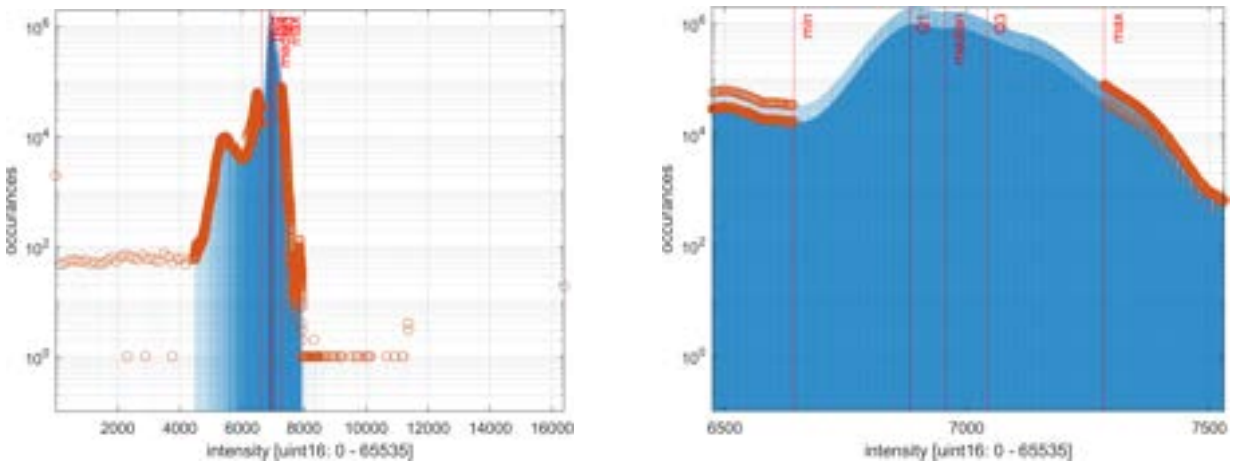
Figure 4: Preprocessed visible light spectrum image from Zenmuse XT 2 camera.

histogram, the 1.5 interquartile range method was used to determine the minimum and maximum bounding values (I_{minBound} and I_{maxBound}; those are commonly used to identify outliers). Finally, the scaled pixel intensity was obtained by subtracting the minimum bounding value from the raw pixel intensity (stored as an unsigned 16-bit integer) and dividing the obtained difference by the difference between the maximum and minimum bounding values (see Eq. 2).

$$R = \frac{R_{\text{uint8}}}{255}, \quad G = \frac{G_{\text{uint8}}}{255}, \quad B = \frac{B_{\text{uint8}}}{255} \tag{1}$$

$$I = \frac{I_{\text{uint16}} - I_{\text{minBound}}}{I_{\text{maxBound}} - I_{\text{minBound}}} \tag{2}$$

Out of every parent image, four 256 by 256 square images of random, uniformly distributed coordinates have been extracted. This has been done due to the limitations of neural networks of choice. The resulting dataset (5016 image pairs) has been split into training (3511), validation (1003) and test (502) sets. The correlation method used both training and validation sets to calculate its coefficients.



(a) Entire range of the histogram.

(b) Histogram zoom-in.

Figure 5: Aggregated infrared images histogram. Minimum and maximum bounding values and outliers labelled.

4 CORRELATION

The first investigated method of generating artificial infrared images is based on linear correlation between visual and infrared-light spectrum images of the same scene. As each pixel in the visual-light spectrum image is a combination of three visual light channels: red, green and blue (RGB); and in the infrared image, each pixel has single value of registered wavelength intensity; it is possible to formulate an equation which linearly correlates the RGB channels intensities to infrared intensity:

$$I_{i,j} = \alpha \cdot R_{i,j} + \beta \cdot G_{i,j} + \gamma \cdot B_{i,j}, \tag{3}$$

where:
$I_{i,j}$ — normalized infrared intensity of the (i, j) pixel,
$R_{i,j}$ — normalized red wavelength intensity of the (i, j) pixel,
$G_{i,j}$ — normalized green wavelength intensity of the (i, j) pixel,
$B_{i,j}$ — normalized blue wavelength intensity of the (i, j) pixel,
α, β, γ — correlation parameters.
Considering the entire image or a set of images, the same equation can be formulated for each pixel, composing a system of linear equations which can be written in a matrix form as follows:

$$\mathbf{AX} = \mathbf{I} \tag{4}$$

where:

$$\mathbf{A} = \begin{bmatrix} R_{11} & G_{11} & B_{11} \\ R_{12} & G_{12} & B_{12} \\ R_{13} & G_{13} & B_{13} \\ \vdots & \vdots & \vdots \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} I_{11} \\ I_{12} \\ I_{13} \\ \vdots \end{bmatrix}. \tag{5}$$

Three correlation parameters α, β, and γ, which describe the relations between the intensities of RGB and infrared channels are being calculated by the least square method. After calculating those parameters, it is possible to generate an artificial infrared image from an arbitrary RGB image by multiplying each pixel's RGB channels' intensities of the source image by the respective correlation parameters and summing up the products. The result is the infrared intensity of the artificial image pixel.

The above method (from now on called the 1-to-1 correlation) can be expanded by incorporating more correlation coefficients. In particular, one can deduce the intensity of the infrared channel not solely based on the RGB channels intensities of the corresponding pixel but based on both the corresponding pixel and its neighbours. In this case, the intensity of the artificial infrared pixel is determined based on a 3 by 3 grid of pixels centred on the corresponding RGB pixel – this method is called the 9-to-1 correlation.
As a result, there are 27 parameters describing the relation between visual and infrared images, 9 for the green channel, 9 for the red channel and 9 for the blue channel. It is worth noting that because of the 3 by 3 grid, correlation cannot be determined for the pixels on the edges of the image. As a result, images reconstructed by this method have a smaller resolution than the source images from which they were created. For instance, from a 256x256 visual-light image, a 254x254 infrared image is created.
The equation describing the correlation in the 9-to-1 method for a single (i, j) pixel of the infrared image is:

$$I_{i,j} = \sum_{k=0}^{2}\sum_{l=0}^{2} (\alpha_{i+k,j+l} \cdot R_{i+k,j+l} + \beta_{i+k,j+l} \cdot G_{i+k,j+l} + \gamma_{i+k,j+l} \cdot B_{i+k,j+l}) \tag{6}$$

Similar to 1-to-1 method, it is also very convenient to use the matrix form. Equation 4 still represents the corresponding system of linear equations, only the $\mathbf{A}$ and $\mathbf{X}$ matrices change:

$$\mathbf{A} = \begin{bmatrix} \mathbf{R} & \mathbf{G} & \mathbf{B} \end{bmatrix}, \tag{7}$$

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & R_{21} & R_{22} & R_{23} & R_{31} & R_{32} & R_{33} \\ R_{12} & R_{13} & R_{14} & R_{22} & R_{23} & R_{24} & R_{32} & R_{33} & R_{34} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_9 \\ \beta_1 \\ \vdots \\ \beta_9 \\ \gamma_1 \\ \vdots \\ \gamma_9 \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} I_{22} \\ I_{23} \\ I_{24} \\ \vdots \end{bmatrix}. \tag{8}$$

Matrices $\mathbf{G}$ and $\mathbf{B}$ are analogous to $\mathbf{R}$.
For both methods, the parameters were calculated based on 4514 images. Obtained correlation parameters for the 1-to-1 and 9-to-1 methods were presented in Equations 9.

$$\begin{aligned} \mathbf{X}_{\mathrm{1to1}} &= \begin{bmatrix} 1.750 \\ 0.236 \\ -1.033 \end{bmatrix}, & \boldsymbol{\alpha}_{\mathbf{9to1}} &= \begin{bmatrix} 0.680 & -0.089 & 0.853 \\ -0.299 & -0.973 & -0.317 \\ 0.991 & -0.071 & 1.112 \end{bmatrix}, \\ \boldsymbol{\beta}_{\mathbf{9to1}} &= \begin{bmatrix} 0.789 & -0.157 & 0.511 \\ -0.302 & -0.872 & -0.451 \\ 0.544 & -0.235 & 0.329 \end{bmatrix}, & \boldsymbol{\gamma}_{\mathbf{9to1}} &= \begin{bmatrix} -1.314 & 0.322 & -1.201 \\ 0.671 & 1.873 & 0.842 \\ -1.380 & 0.382 & -1.279 \end{bmatrix}. \end{aligned} \tag{9}$$

5 NEURAL NETWORK

Two variants of the pix2pix architecture were selected for training on custom dataset and further analysis.

Pix2pix [18] is a conditional generative adversarial network (cGAN). Such networks consist of a generator that is trained to generate images and a discriminator that is trained to distinguish between real and fake images. Pix2pix is a very flexible network that can be used to conduct versatile image manipulation tasks such as: translation between image styles, translation of label maps into images, colourization of images, generation of images from edges and filling missing parts of images, to name a few. Pix2pix with its default training settings was used, that is with a patch (70×70 discriminator) and L1+cGAN loss function. Pix2pix network has been trained with 2 different batch sizes.
Pix2pixHD [19] is an evolution of pix2pix network that allows for generating images with finer detail and of higher resolutions. Its main use is to translate label maps to images but it also is capable of translating one style of image into another. Its downside compared to pix2pix is a need for greater resources required to train it.
Table 2 presents details of training parameters performed on selected architectures.

Table 2: Networks' training parameters

Network	pix2pix	pix2pix-b25	pix2pixHD
No. training images	3511	3511	3511
No. validation images	1003	1003	–
No. epochs	200	300	200
Batch size	1	25	1
Epoch time	175 s	115 s	790 s
GPU	RTX 3060 Max-Q (6GB)	Google Colab Tesla T4	RTX 3060 Max-Q (6GB)

6 RESULTS

To compare the results obtained by various methods, source infrared scaled images and artificially generated infrared images were converted into temperature ranges assuming that the minimum and maximum bounding values obtained in Section 3 corresponded to minimum and maximum surface temperature on the day of data acquisition (14 °C and 24 °C, respectively), and with the assumption of simple black body radiation model defined by Planck's law. This was done to allow comparison in terms of somewhat meaningful units (in this case degrees Celsius) rather than arbitrary normalized pixel intensities. It should be noted that the resulting temperature is still considered arbitrary but what was of interest to the authors is the difference between the temperature on the true infrared image and the artificially generated images.
Tables 3-6 presents a selection of artificially generated images from each method, along with true infrared images, source RGB images and error images (difference between artificially generated image and source infrared image). Table 7 presents the comparison of temperature error statistics - the data were determined based on a sample of 502 test images.
Looking at the statistical comparison in Table 7 it can be noted that overall the neural networks offer smaller errors than the correlation methods. The pix2pix and pix2pixHD yielded the smallest mean error, followed by the pix2pix-b25, while the correlation methods yielded the largest error, about four to five times larger than pix2pix and pix2pixHD. The pix2pix-b25 and the correlation methods had the mean absolute error and the standard deviation about twice that of the pix2pix and pix2pixHD. It needs to be mentioned that although pix2pix-b25 has managed to achieve smaller errors than correlation methods, it tends to generate visually unsatisfying results.

The superiority of the neural networks' methods in regards to chosen metrics is better visible in Tables 3-6. Correlation methods struggle to properly assign infrared values to a broad range of objects. One of its features is edge preservation. This is a direct result of the method's construction. Edge preservation can be both an upside, like when processing most man-made objects (image 1 in Table 3, image 8 in Table 4, images 10 and 11 in Table 5), and a downside, in case of processing of solar panels or greenhouses (image 2 in Table 3 and images 5 and 6 in Table 4). Chosen neural networks do not have this property. This leads to generally worse shape reproduction (images 1 and 2 in Table 3, images 5, 6, 8 in Table 4). Neural networks, however, achieved good results when translating images of roads.
The major downside of used correlation methods is their general inability to map multiple RGB colours to one infrared. This can be clearly seen in images 11 and 12 in Table 5 and in the transformation of images of trees in Table 6.
It is worth mentioning that neural networks can generate artefacts, especially when presented with types of objects that were underrepresented in the training dataset. In image 1 in Table 3 transformed rectangular features seem to be translated. In images presenting tracks in the sand (image 4 in Table 3, image 9 in Table 5, and images 13 and 14 in Table 6) it is visible that neural networks struggled to translate them faithfully.
Regarding the comparison between 1-to-1 and 9-to-1 (3x3 grid) pixel correlation methods, the latter provided only tiny improvements. This could be expected as the application of pixel grids is useful when the signal gradients play a role. In general, the temperature of an object will depend on the object type (its colour) rather than the gradient of its colour. Therefore, no larger pixel grids (5x5, 7x7 etc.) were investigated.
Concerning computational complexity, it must be noted that the correlation methods were computed in less than a minute on an average-class PC, while neural networks training took several days and required access to a cloud computing platform or a high-end GPU.

7 SUMMARY

The paper presented several methods for generating artificial infrared images based on visible light spectrum images. Although it can be determined that the neural networks-based methods offered overall better results than correlation methods, they are subject to several important constraints. As the utilized neural networks must be trained using 256 by 256 images and can only output images of that size, their output images cannot be directly used while simulating image frames from an infrared camera of arbitrary resolution. One way to overcome this problem would be to divide visible light spectrum image corresponding to the simulated infrared camera field of view into tiles appropriate for the neural network, and then stitch the resulting artificial infrared images back together to form an image of appropriate size. Another solution would be to develop a new neural network capable of working on images of the required size; however, this would require retraining of the network each time the required image's resolution changes.
This restriction does not apply to the correlation methods as they operate on a pixel-by-pixel basis. Therefore, each image used either during training or during later application, might be of a different size. To increase the quality of results offered by the correlation method, one might increase the number of coefficients used. This might be done, for example, by switching from determining the correlation between radiation bands (red, green, blue and thermal infrared) to determining the correlation between specific colours and thermal radiation – dividing visible colourspace into a finite number of distinct colours and assigning a specific value of thermal radiation to each colour.
Neural network architectures selected for training on created dataset transform images of natural terrains with relatively high accuracy. However, they perform worse with images of man-made objects,

Table 3: Comparison of results from various artificial infrared image generation methods. Part 1/X

No.		true	corr. 1-to-1	corr. 9-to-1	pix2pix	pix2pix-b25	pix2pixHD
1	IR image						
	RGB or error						
2	IR image						
	RGB or error						
3	IR image						
	RGB or error						
4	IR image						
	RGB or error						

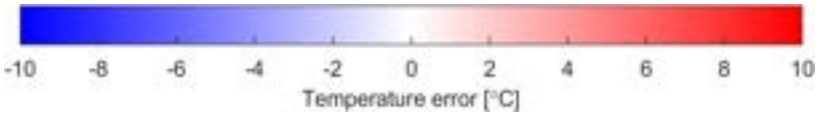
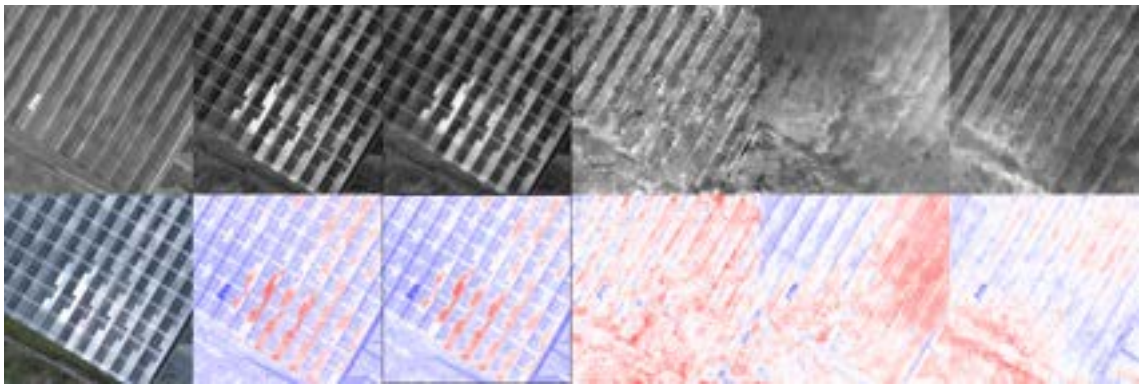
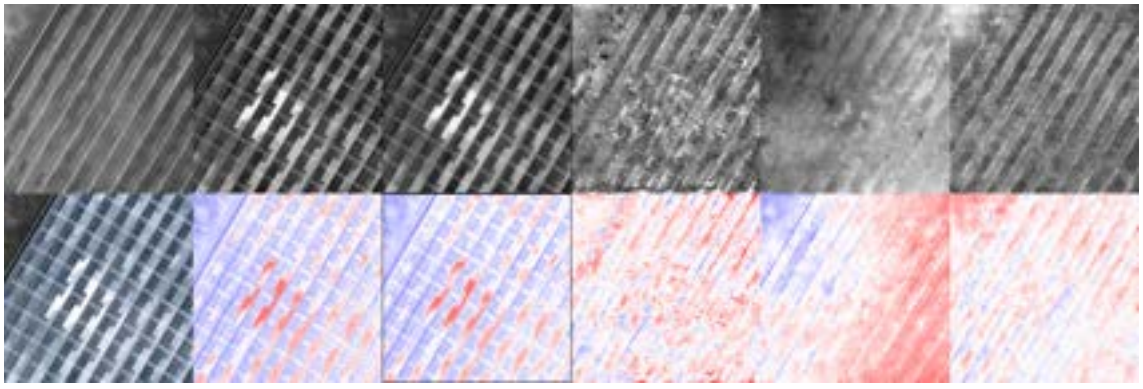
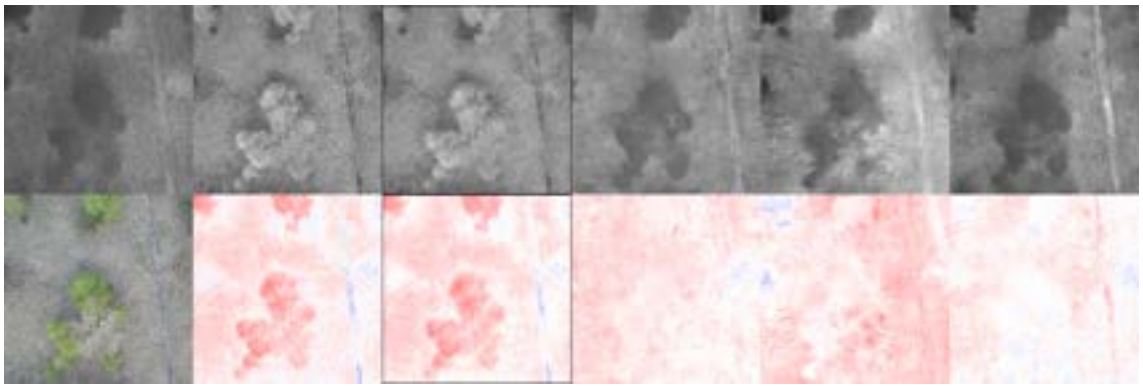
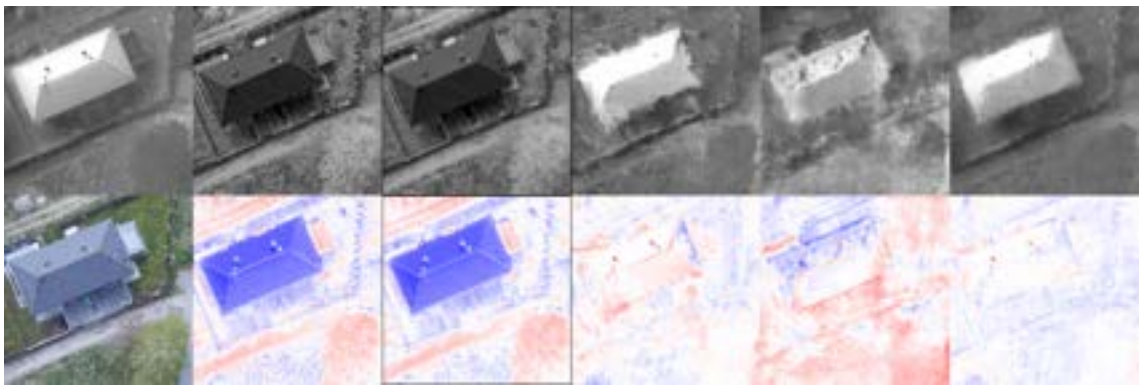


Table 4: Comparison of results from various artificial infrared image generation methods. Part 1/X

No.		true	corr. 1-to-1	corr. 9-to-1	pix2pix	pix2pix-b25	pix2pixHD
5	IR image						
	RGB or error						
6	IR image						
	RGB or error						
7	IR image						
	RGB or error						
8	IR image						
	RGB or error						

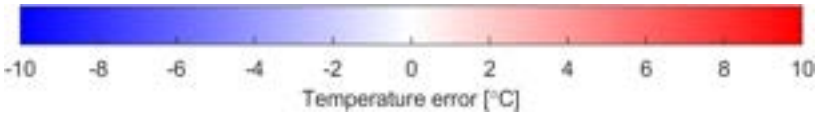


Table 5: Comparison of results from various artificial infrared image generation methods. Part 1/X

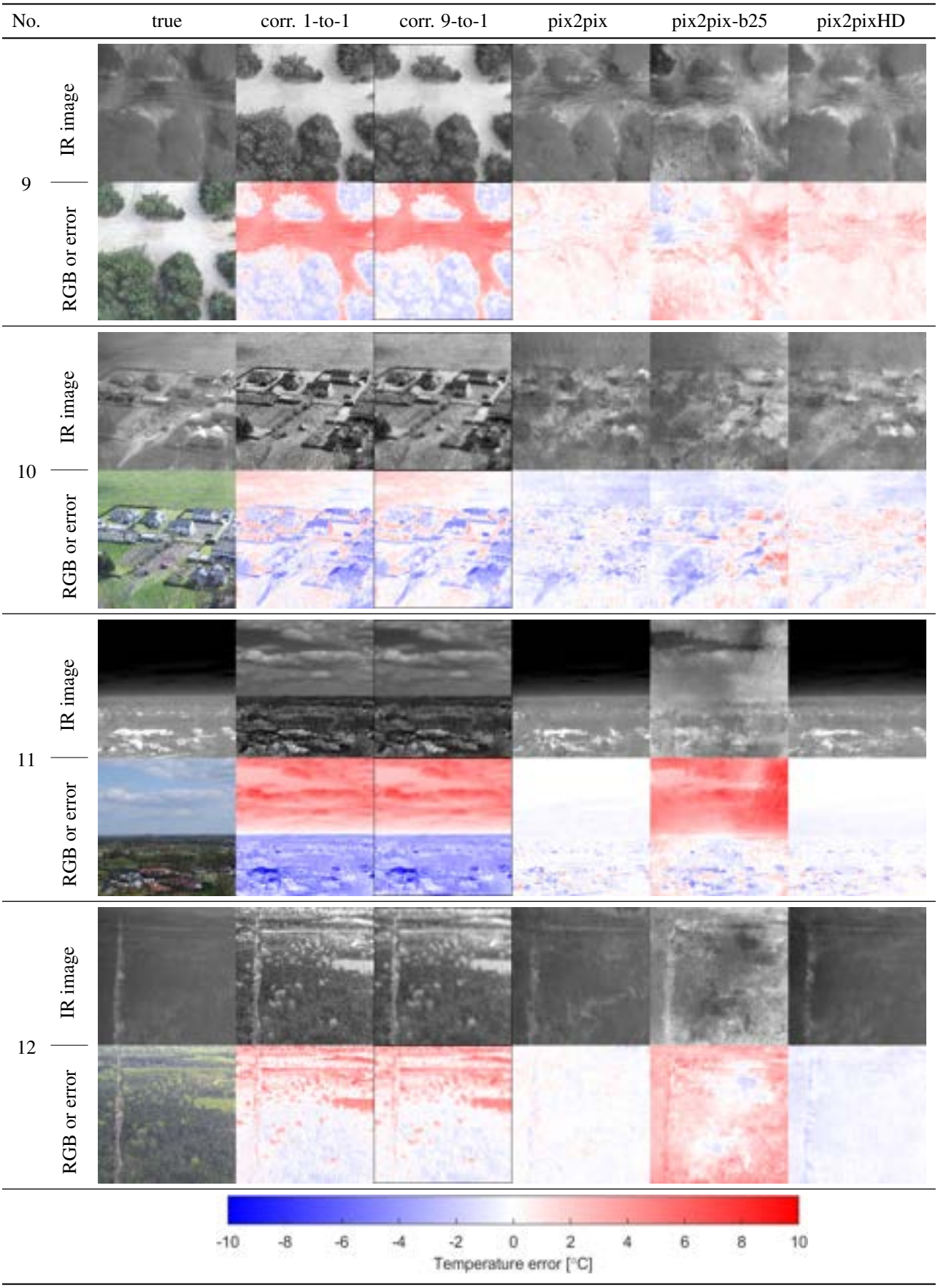


Table 6: Comparison of results from various artificial infrared image generation methods. Part 1/X

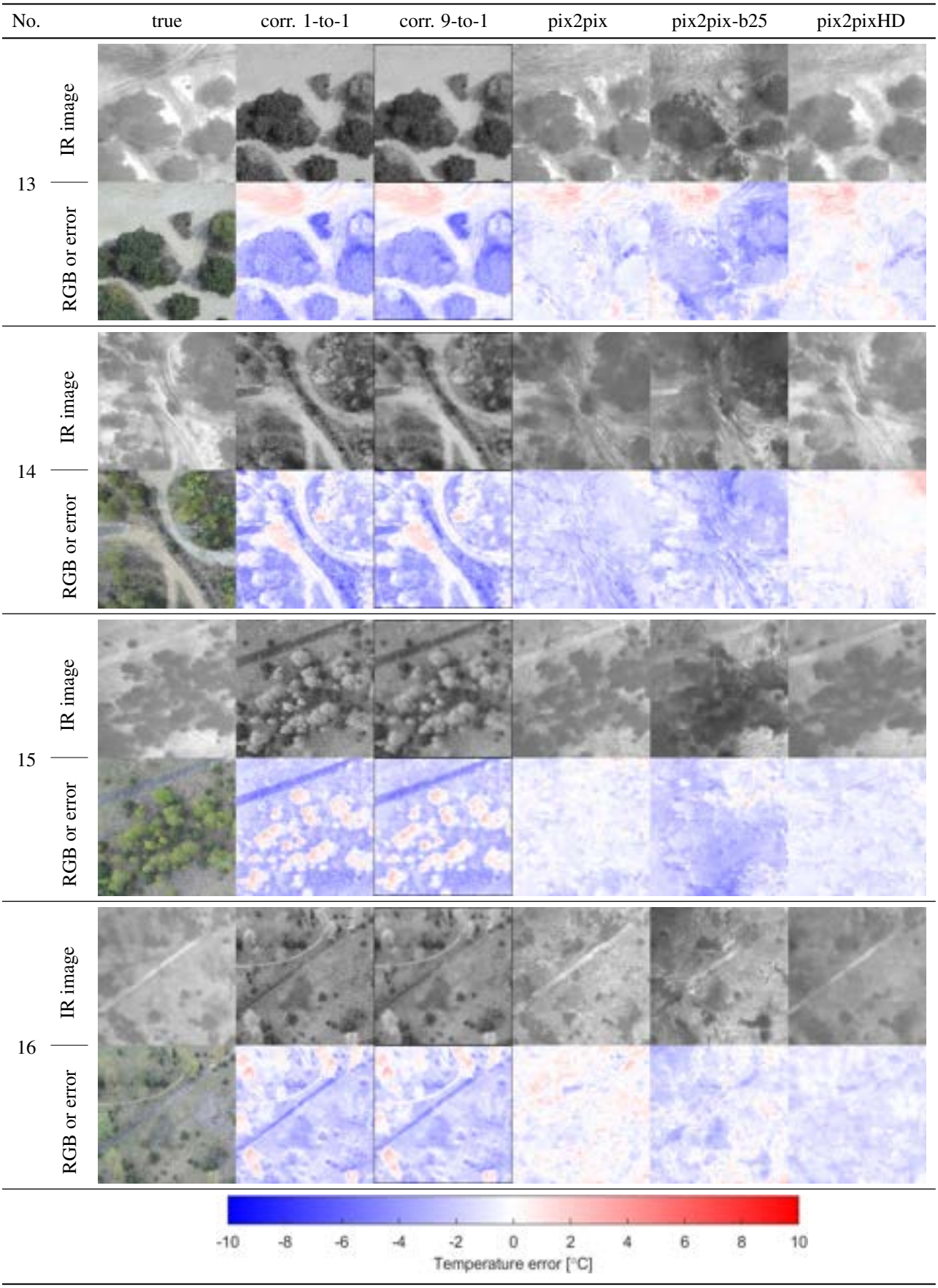


Table 7: Statistical comparison of results from various artificial infrared image generation methods.

	correlation 1-to-1	correlation 9-to-1	pix2pix	pix2pix-b25	pix2pixHD
Mean error (°C)	-0.46	-0.4	-0.086	0.20	0.11
Mean absolute error (°C)	1.84	1.78	0.73	1.6	0.69
Standard deviation (°C)	2.23	2.16	0.99	2.1	0.96

such as buildings and solar panels. This may be due to the relatively small representation of the latter objects in the dataset. It is probable though, that more refined neural networks would achieve better results. Another possible solution to address this issue would be to use networks trained on unpaired datasets. Nevertheless, the main scope of this research remain images of terrain that the rocket flies above, which is why man-made objects are of lesser concern.
The aim of this paper was not to find a correlation between the visible light spectrum and thermal radiation – if it existed, it would almost certainly have been found before. The task was to develop an algorithm that would produce artificial infrared images "good enough" for prospective application - simulation of infrared-based rocket guidance. As there are still challenges to overcome and possible improvements to be made, the research is far from over.

8 ACKNOWLEDGEMENTS

This research was supported by the Boeing Company as part of the Higher Education Relations Grant awarded to the Warsaw University of Technology.

REFERENCES

[1] E. Fleeman, *Tactical Missile Design* (AIAA education series). American Institute of Aeronautics and Astronautics, 2006, ISBN: 9781563473852. [Online]. Available: https://books.google.pl/books?id=iSsfAQAAIAAJ.

[2] D. Miedziński, K. Bresler, S. Małecki, A. Mochol, M. Sochacki, and K. Wojciechowski, "Low cost rocket guidance and control development platform," in *Proceedings, International Astronautical Congress*, 2020.

[3] D. Miedziński, K. Bresler, S. Małecki, *et al.*, "Preliminary design of a homing rocket using image recognition," in *Proceedings, International Astronautical Congress*, 2021.

[4] M. Michałów, K. Kaczmarek, and P. Umiński, "SKA Rocket Flight Simulation," in *Proceedings, Studencka Konferencja Kosmiczna*, 2020.

[5] Y. Li, R. Fu, X. Meng, W. Jin, and F. Shao, "A sar-to-optical image translation method based on conditional generation adversarial network (cgan)," *IEEE Access*, vol. 8, pp. 60 338–60 343, 2020. DOI: 10.1109/ACCESS.2020.2977103.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," vol. 9351, Oct. 2015, pp. 234–241, ISBN: 978-3-319-24573-7. DOI: 10.1007/978-3-319-24574-4_28.

[7] M. Schmitt, L. Hughes, C. Qiu, and X. Zhu, "Sen12ms - a curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion," vol. IV-2/W7, Sep. 2019, pp. 153–160. DOI: 10.5194/isprs-annals-IV-2-W7-153-2019.

[8] X. Yuan, J. Tian, and P. Reinartz, "Generating artificial near infrared spectral band from rgb image using conditional generative adversarial network," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. V-3-2020, pp. 279–285, 2020. DOI: 10.5194/isprs-annals-V-3-2020-279-2020. [Online]. Available: https://isprs-annals.copernicus.org/articles/V-3-2020/279/2020/.

[9] X. Yuan, J. Tian, and P. Reinartz, "Learning-based near-infrared band simulation with applications on large-scale landcover classification," *Sensors*, vol. 23, no. 9, 2023, ISSN: 1424-8220. DOI: 10.3390/s23094179. [Online]. Available: https://www.mdpi.com/1424-8220/23/9/4179.

[10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, *et al.*, "Generative adversarial networks," *Advances in Neural Information Processing Systems*, vol. 3, Jun. 2014. DOI: 10.1145/3422622.

[11] Y. Pang, J. Lin, T. Qin, and Z. Chen, "Image-to-image translation: Methods and applications," *CoRR*, vol. abs/2101.08629, 2021. arXiv: 2101.08629. [Online]. Available: https://arxiv.org/abs/2101.08629.

[12] D. Miedziński, K. Kaczmarek, P. Rodo, N. Sahbon, M. Sochacki, and M. Łukasiewicz, "Missile Aerodynamics Model Identification Using Flight Data," in *2023 IEEE Aerospace Conference*, 2023, pp. 1–12. DOI: 10.1109/AERO55745.2023.10115928.

[13] K. Kaczmarek, M. Hałoń, S. Małecki, D. Miedziński, M. Sochacki, and P. Rodo, "Mission design of FOK guided rocket," in *Proceedings, 2nd International Conference on Flight Vehicles, Aerothermodynamics and Re-entry Missions & Engineering (FAR)*, 2022.

[14] D. Miedziński, M. Hałoń, K. Kaczmarek, P. Rodo, and M. Sochacki, "Proportional navigation missile guidance using image recognition," in *Proceedings, 9th European Conference for Aeronautics and Space Sciences (EUCASS)*, 2022. DOI: 10.13009/EUCASS2022-7335.

[15] Y. Leshchenko, K. Kaczmarek, F. Łabaj, S. Małecki, D. Miedziński, and M. Sochacki, "Imaging infrared seeker design and tests for FOK guided missile," in *Proceedings, SPIE Defense + Commercial Sensing 2023*, 2023.

[16] DJI, "Zenmuse XT 2 User Manual," 2018. [Online]. Available: https://dl.djicdn.com/downloads/Zenmuse%5C%20XT%5C%202/Zenmuse_XT_2_User_Manual_v1.0_en_.pdf.

[17] DJI, "Matrice 200 Series V2 User Manual," 2020. [Online]. Available: https://dl.djicdn.com/downloads/m200_v2/20200630/M200_Series_V2_User_Manual_en4.pdf.

[18] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *CoRR*, vol. abs/1611.07004, 2016. arXiv: 1611.07004. [Online]. Available: http://arxiv.org/abs/1611.07004.

[19] T. Wang, M. Liu, J. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," *CoRR*, vol. abs/1711.11585, 2017. arXiv: 1711.11585. [Online]. Available: http://arxiv.org/abs/1711.11585.