ISP pipeline analysis

Sahan Nelundeniyalage

1. Introduction

1.1. Introduction and context of the study

Cameras typically capture images by converting analog voltages generated when light interacts with photodiodes on the image sensor into a digital format. The resulting digital image, termed the RAW image, undergoes further refinement based on its intended application. Adjustments, such as alterations in gain, white balance, contrast, and histogram equalization, are applied through a sequence known as the Image Signal Processing (ISP) Pipeline.

In the realm of cameras, thermal cameras have garnered significance, especially in scenarios characterized by limited visibility, such as complete darkness. Thermal cameras utilize infrared radiation imaging to observe and detect potential threats, finding applications in crucial areas like perimeter protection, military systems, and pollution detection. However, the processing and interpretation of thermal images present unique challenges distinct from visible light images. This work highlights the complexities inherent in thermal image processing, encompassing crucial aspects like non-uniformity correction, bad pixel correction, and image enhancement through established algorithms and methodologies.

1.2. Research questions and objectives of the project

Given that a camera's ISP comprises various functional blocks tasked with enhancing or diminishing specific features in the input RAW image, there exists a potential for obtaining significantly different ISP outputs by manipulating the input RAW image features. This variability may lead to inaccurate object detection in applications reliant on object detection, thereby posing a severe security concern. In this project, we explore three types of ISP pipelines employed in Automotive Cameras—traditional color camera ISP, deep-learning-based ISP, and thermal ISP—and assess the performance of object detection under various ISP attack scenarios through simulation.

1.3. Objectives and milestones of the task

In this study, two primary tasks are involved. The primary task involves the implementation and analysis of existing thermal ISPs. This will encompass a detailed examination of thermal ISP pipeline stages, such as bad-pixel correction, non-uniformity correction, and image enhancement, through implementing a chosen thermal camera ISP.

The secondary task involves dataset handling and preparation. I am responsible for acquiring and organizing both RAW and processed images suitable for the pipeline. The manual implementation of bounding boxes on processed images, utilizing YOLOv8 or similar object detection algorithms, will facilitate a comparative analysis with images produced at the pipeline's end, while RAW images serve as the original input data.

The set of milestones includes:

1. Literature Survey: Conduct an exhaustive literature survey on existing thermal ISP pipeline implementations to identify common pipeline stages. This milestone will establish a robust foundation for subsequent tasks.
2. Implementation: Implement selected existing thermal ISP pipelines and reproduce their results to gain comprehensive insights into each stage of the pipeline.
3. Dataset Preparation: Prepare the dataset with the required RAW images and processed images, incorporating bounding boxes for a comparative analysis of the images obtained at the end of the pipeline.

2. Background and related work

A thermal camera is a device that captures and visualizes infrared (IR) radiation emitted by objects, enabling the detection of temperature variations. Unlike traditional cameras that rely on visible light, thermal cameras operate in the infrared spectrum, making them capable of capturing heat signatures. Infrared radiation (IR) is electromagnetic radiation with longer wavelengths than visible light, and thermal cameras use these infrared signals to create images based on temperature differences in the observed scene.

Thermal cameras are increasingly used in diverse fields, including integration into military environments for object detection, identification, and tracking, security and surveillance systems, inspection systems for defect identification, pollution detection, and more [1]. The key to their effectiveness across these applications lies in the effective processing of infrared data to guarantee the faithful representation of the observed situation. Thermal cameras achieve this generally using three essential modules [1]–[3]:

- Focal Plane Array (FPA) Module: This module contains the detector array that converts infrared radiation into electrical signals and the analog-to-digital converter (ADC) that transforms the signals into digital format for subsequent processing.
- Control and Digital Processing Module: This module is responsible for sending synchronization and control signals to the detector array module. This module also executes image processing and analysis methods to enhance the quality and usability of the thermal image.
- Imaging Module: This module displays the thermal image on a video screen or other output device,

allowing users to adjust the camera parameters and settings.

While the domain of digital camera ISP has seen extensive research, it is noteworthy that thermal camera ISP implementation has received comparatively less attention. However, in works [2], [3], the authors propose a thermal ISP pipeline designed for a field-programmable gate array (FPGA) implementation. These works detail the design and implementation of a digital module capable of real-time operations on infrared images, including non-uniformity correction (NUC), bad pixel correction (BPC), display interface driving, and additional information imposition. The functional diagram of ISP introduced in the papers mentioned above serves as a foundation in [4] for the implementation of a real-time thermal ISP in a System-On-Chip (SoC)-FPGA.

Building upon [2]–[4], [1] extends the research, providing a comprehensive and detailed explanation of thermal camera ISP. This work systematically explains each essential functional block within the thermal ISP, making it a valuable resource for those seeking to implement thermal ISP independently. Notably, [1] includes example images for every functional block, accompanied by clear explanations, making it an instrumental blueprint for our work. In addition to detailing the thermal ISP, the authors of [1] delve into detection, recognition, and object tracing algorithms. Hence, [1] serves as a comprehensive reference for those interested not only in thermal ISP but also in associated algorithms for detection, recognition, and object tracing. However, our work only focuses on the ISP since the object detection part is done separately.

[5] also provides a flow diagram for a thermal ISP, which can be used in bolometers IR detectors. In [6], a comprehensive digital circuit design for the acquisition and pre-processing of infrared images using FPGA is proposed by advantages offered by uncooled focal plane arrays. This paper also consists of a functional explanation of the thermal ISP. [7], [8] provide application-specific thermal ISP, respectively explaining the applicability of thermal cameras in system defect inspection and agriculture fields. These works also provide insights into essential thermal ISP functional blocks.

As elaborated in the Methodology section, NUC, BPC, and advanced image processing are crucial functional blocks in any thermal ISP. [9] proposes one of the widely adopted NUC algorithms, utilizing two sets of gain and offset for each pixel. Another commonly used NUC algorithm, employing the simple two-point correction, is proposed in [10]. Our implementation follows the approach outlined in [11], which leverages midway infrared equalization for NUC.

Concerning BPC, numerous algorithms for detecting bad pixels have been documented in the literature, such as [12]–[14]. The BPC in OpenISP [15], an open-source ISP implementation, is integrated with our ISP. [1], [15]–[17] offer in-depth insights into advanced image processing. Following a thorough literature analysis, we opted to include noise filtering, edge enhancement, and brightness and contrast control as functional blocks in the advanced image processing module.

3. Methodology

As explained in the Background and Related Work section, the functionality of a thermal camera can be divided into three modules: FPA, Control and Digital Processing Module, and Imaging Module. This project specifically targets the Control and Digital Processing Module, focusing on the digital processing part commonly referred to as the ISP. According to [1], image processing in a thermal camera can be categorized into three groups. The first group includes algorithms necessary for the thermal imager's operation, such as NUC of detector responses and BPC. The second group comprises algorithms that enhance image quality to facilitate interpretation by operators or vision systems. The third group involves image processing algorithms for data analysis, enabling automatic object detection, tracking, and scene interpretation, often associated with machine vision. The remainder of this section elaborates on implementing the algorithms from the first two types, as depicted in Figure 1.

3.1. Non-Uniformity Correction (NUC)

Imaging sensors, predominantly charge-coupled devices (CCDs), capture photons and convert them into charges. Each pixel exhibits a unique response in CCDs, resulting in non-uniformity (NU) in sensor response. This non-uniformity often appears as structured noise, known as fixed pattern noise or NU noise, manifesting as a row or line pattern in images, as shown in the input RAW image in Figure 1. This NU noise significantly compromises image quality. Uncooled infrared cameras face specific challenges in addressing this issue, as their detector response undergoes rapid changes over time, preventing a one-time correction by the manufacturer.

For this implementation, we employed the NUC algorithm proposed in [11], known as MIRE, which stands for Midway Infrared Equalization. The algorithm involves the following main steps:

- For each column of the image, compute the cumulative histogram of the pixel values.
- For each column, compute a local midway histogram by averaging the inverse histograms of the neighboring columns using Gaussian weights.
- For each column, map the pixel values to the midway histogram to obtain the corrected image.

The MIRE algorithm is fully automatic, operates on a single image, and is parameter-free. Notably, it does not require any registration, motion compensation, or closed aperture calibration. This approach ensures a straightforward and efficient non-uniformity correction process for thermal images.

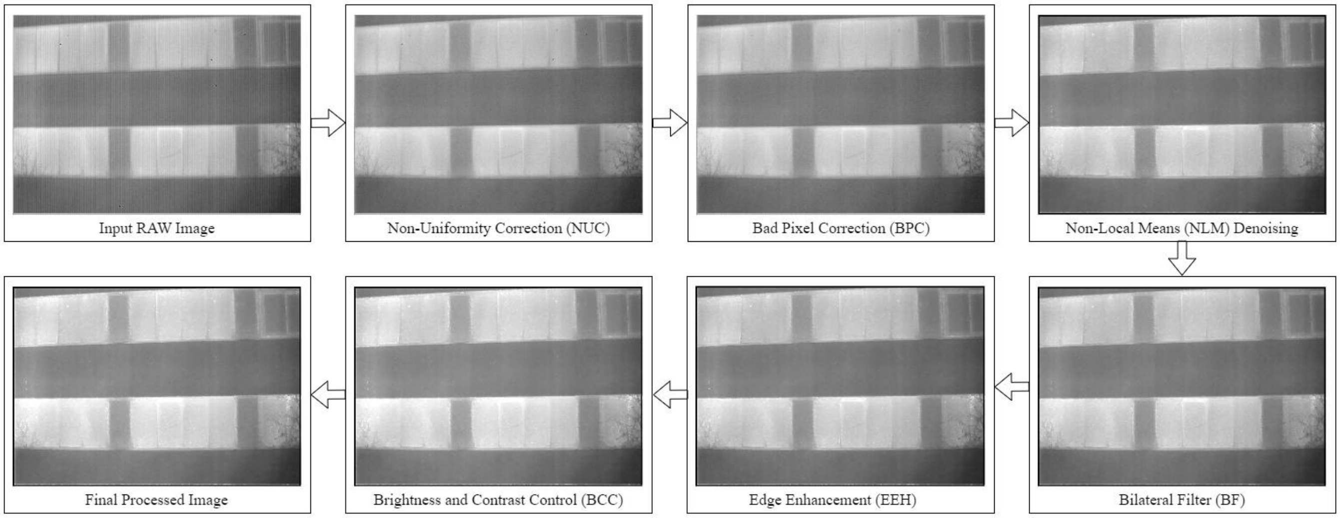

Figure 1: Thermal ISP flow diagram with example images

3.2. Bad Pixel Correction (BPC)

Infrared images may exhibit "bad pixels," characterized as bright or dark spots with fixed coordinates that remain unchanged with the target. In the top right side of the input RAW image in Figure 1, several dark spots are evident, which are caused by bad pixels. These pixels result from the diverse response rates of individual detection elements in the infrared detector to infrared radiation and encompass overheated and dead pixels. Also known as invalid pixels, these anomalies can adversely impact image quality. Imaging sensors commonly contain a small percentage of invalid pixels with abnormal sensitivity, providing no useful information about the observed scene. These defects, typically distributed evenly across the sensor, appear as impulse noise on the image, potentially disrupting processing operations. Therefore, the use of noise mitigation techniques becomes necessary.

These defects can be static (always constant) or dynamic (changed by conditions like exposure or temperature). The methods for detecting these defects vary. For static defects, black or white images are captured, and the locations of known defects are stored. For dynamic defects, the pixel value is estimated using the differences between the center pixel and its neighbors and then compared with a threshold. Once detected, the correction methods include a mean filter, which replaces the defect with the average of the neighboring pixels, or a gradient-based filter, which replaces the defect with the average of the adjacent pixels in the direction with the minimum gradient. This process ensures the integrity of the image data despite the presence of defective pixels. In this project, we adopt dynamic defects detection and both mean filter-based and gradient-based filter correction implementation in [15].

3.3. Image Quality Enhancement

The image quality enhancement process doesn't increase the information in the image but modifies certain features like contrast to make the visual image easier to interpret. The techniques used are mostly heuristic and include several image transformations such as noise reduction, brightness enhancement, contrast enhancement, and edge detection and sharpening [17]. The theoretical background for thermal image enhancement is digital image processing theory [1]. Nevertheless, methods effective for processing visual images may not translate directly to thermal image processing, given the inherent distinctions between visual and thermal imagery. The developed implementation includes three stages: noise removal, edge detection and sharpening, and brightness and contrast enhancement. The sequence of these operations is crucial for achieving the desired level of image enhancement.

3.3.1 Noise Filtering

The initial stage of the image quality enhancement process is noise reduction, aiming to decrease the number of pixels with significantly different brightness levels from their close neighborhood. This prevents such pixels from being emphasized by subsequent procedures in the pipeline, thereby avoiding the creation of unwanted artifacts [16], [17]. We have implemented two noise filtering techniques: Non-Local Means (NLM) denoising and Bilateral Filter (BF) denoising. NLM denoising relies on the principle that similar patches within an image share statistical characteristics [18]. By averaging pixel values based on the similarity of these patches, NLM effectively reduces noise while preserving image details. On the other hand, Bilateral Filter (BF) denoising considers both spatial and intensity information, smoothing an image while preserving its edges [19]. These methods play a crucial role in enhancing the quality of IR images, providing effective means to reduce noise artifacts and enhance the overall clarity of thermal visuals.

3.3.2 Edge Enhancement

Edge enhancement is a technique designed to improve the apparent sharpness of an image by intensifying the contrast around its edges. This process involves identifying edge boundaries and applying a gain to the pixels near these edges. Adjusting the gain is possible through parameters like radius, threshold, and the direction of the edge. The most commonly used algorithm for edge enhancement is unsharp masking, employing a linear filter to subtract a blurred version of the image from the original. The unsharp masking in [15] is used in the implemented thermal ISP. A more sophisticated algorithm, gradient-corrected bilinear interpolation, uses a nonlinear filter to rectify luminance changes along the edge direction. While edge enhancement can enhance the visual quality of an image, it may introduce artifacts such as overshoot and undershoot, necessitating clipping or suppression.

3.3.3 Contrast and Brightness Control

Brightness/contrast control is a process that fine-tunes the intensity and range of an infrared (IR) image. Brightness reflects the relative expression of the light output intensity from a source, while contrast represents the ratio between the luminance of the brightest and darkest colors. Manipulating brightness and contrast can either enhance or diminish the visibility of details in an image. This adjustment involves adding or subtracting a constant value to alter brightness and multiplying or dividing a ratio to adjust contrast [1], [15], [17]. These constant values and ratios serve as configurable parameters, allowing users to fine-tune them based on their specific preferences and the desired visual effect. In the context of the developed thermal ISP, we specifically implement the brightness and contrast control algorithm detailed in [15].

3.3.3 Dataset Preparation

The secondary task of the project involves dataset handling and preparation. A literature survey was conducted to find a dataset with both RAW and processed images. Since our overall goal is to evaluate the object detection effectiveness of attacks on the ISP, we require processed images with bounding boxes. In object detection, a bounding box is a rectangular frame that precisely encloses and delineates the spatial extent of a detected object within an image. It provides a structured representation of the object's location, aiding in the localization and identification of objects in computer vision applications. We opted to use YOLO as the object detection algorithm because it is particularly valued for its speed and efficiency, enabling rapid and accurate object detection in diverse contexts, such as surveillance, autonomous vehicles, and image analysis.

PASCALRAW [20] is a RAW image database for object detection applications; however, it does not contain processed images. Dataset [21] contains both RAW and processed images without bounding boxes on processed images. This dataset can be utilized by manually implementing the bounding boxes. [22] includes a car images dataset with both RAW and processed images along with bounding boxes. The given bounding boxes were not in the format used in YOLO. Therefore, we converted all the bounding boxes into the YOLO format. After this process, we selected [22] as our dataset for traditional and deep learning-based ISP testing.

4. Experimental Results

In this section, we present the experimental results of our thermal ISP. First, we describe our experimental setup. Next, we outline the results of our experiments. For the experiments, we used the same IR image dataset used in [16], [23], [24]. The implementation of the thermal ISP was done in the C++ programming language. Our experiments were conducted on a Linux environment, utilizing a 2.9 GHz 8-core AMD Ryzen 7 4800HS processor with 16 GB of RAM. As illustrated in Figure 1, we generate the output of each functional block in the ISP, as explained in the Methodology section. Each image in Figure 1 represents the output of the labeled functional block. Subsequently, we compare the outcome images with those from previously published works and evaluate the desired functionality of each block in the ISP. After the evaluation, we found that we achieved almost the expected results for the NUC and BPC blocks. To test the denoising blocks (NLM and BF), we provided a RAW image with artificially added Gaussian noise. After several calibrations of parameters used in the blocks, these denoising blocks were able to remove noise to some extent. Similar to the NLM and BF blocks, the BCC block also required several attempts to find suitable calibration parameters. After identifying the suitable calibration parameters, we obtained acceptable results for the BCC block.

5. Analysis and discussion

This project comprises two main tasks with three milestones. The first milestone involved conducting an exhaustive literature survey on existing thermal ISP pipeline implementations to identify common pipeline stages, establishing a robust foundation for subsequent tasks. During this milestone, it was observed that compared to other ISPs, such as traditional camera ISPs and deep learning-based ISPs, the number of accessible literature on thermal cameras is relatively limited. This limitation could be attributed to the predominant application of thermal cameras in the fields of security and defense. Additionally, within the accessible literature, there are only a few open-source implementations. However, despite these challenges, the existing literature provided insights to identify all the essential and common functionalities in thermal ISPs.

The second milestone was to implement selected existing thermal ISP pipelines and reproduce their results. Leveraging the knowledge gathered during the first milestone and utilizing open-source implementations, a functional thermal ISP with essential functional blocks was successfully created. Experimental results demonstrate the efficiency of these implementations, as illustrated in Figure 1. When it comes to testing, it is worth mentioning that there wasn't a standard benchmark available to evaluate the outputs of the thermal

ISP functional blocks. The final milestone involved preparing a dataset with the required RAW images and processed images, incorporating bounding boxes for a comparative analysis of the images obtained at the end of the pipeline. This dataset serves as the basis for comparing object detection accuracy under attacks on traditional and deep learning-based ISPs. We were able to find several datasets from previous works, as explained in the Methodology section, and we made some modifications and adopted the selected dataset into our project.

5.1. Takeaways

The tasks of implementing the thermal ISP and preparing the dataset are integral components of a project that assesses object detection accuracy under attacks on ISP. Throughout these subtasks, we delved into the domain of image processing, particularly in the context of thermal imagery. The literature survey and the implementation task provided comprehensive insights into the functionality, purpose, and design of ISPs.

5.2. Limitations

One of the limitations of the implemented thermal ISP is that it currently supports only 8-bit PNG as the input RAW image. Another limitation, as explained in the Methodology section, is that certain functional blocks, such as NLM, BF, and BCC, possess calibration parameters that significantly impact the final output image quality. To achieve optimal results from the thermal ISP, calibration of these parameters is essential. Regarding the dataset, the currently employed dataset features only one type of object class (cars). If additional classes are required for other applications, the dataset in [20] can be utilized, and processed images can be generated using the implemented traditional ISP or any other ISP.

As a final note, although the implemented thermal ISP yields acceptable results, we acknowledge that further fine-tuning and optimization could enhance its performance for better results in future works.

REFERENCES

[1] T. Sosnowski, G. Bieszczad, and H. Madura, "Image Processing in Thermal Cameras," in *Studies in Systems, Decision and Control*, vol. 106, Springer International Publishing, 2018, pp. 35–57.

[2] T. Sosnowski, G. Bieszczad, M. Kastek, and H. Madura, "Digital image processing in high resolution infrared camera with use of programmable logic device," in *Optics and Photonics for Counterterrorism and Crime Fighting VI and Optical Materials in Defence Systems Technology VII*, SPIE, Oct. 2010, p. 78380U. doi: 10.1117/12.865026.

[3] G. Bieszczad, T. Sosnowski, H. Madura, M. Kastek, and J. Barela, "Adaptable infrared image processing module implemented in FPGA," in *Infrared Technology and Applications XXXVI*, SPIE, Apr. 2010, p. 76603Z. doi: 10.1117/12.851146.

[4] G. Bieszczad, "SoC-FPGA embedded system for real-time thermal image processing," in *Proceedings of the 23rd International Conference Mixed Design of Integrated Circuits and Systems, MIXDES 2016*, Institute of Electrical and Electronics Engineers Inc., Aug. 2016, pp. 469–473. doi: 10.1109/MIXDES.2016.7529788.

[5] T. Ducol, T. Rahaga, and B. Louvat, "Image signal processor for bolometers IR detectors," SPIE-Intl Soc Optical Eng, Apr. 2021, p. 31. doi: 10.1117/12.2588758.

[6] E. Lielāmurs, A. Cvetkovs, R. Novickis, and K. Ozols, "Infrared Image Pre-Processing and IR/RGB Registration with FPGA Implementation," *Electronics (Switzerland)*, vol. 12, no. 4, Feb. 2023, doi: 10.3390/electronics12040882.

[7] M. Siami, T. Barszcz, J. Wodecki, and R. Zimroz, "Design of an Infrared Image Processing Pipeline for Robotic Inspection of Conveyor Systems in Opencast Mining Sites," *Energies (Basel)*, vol. 15, no. 18, Sep. 2022, doi: 10.3390/en15186771.

[8] K. A. Steen, A. Villa-Henriksen, O. R. Therkildsen, and O. Green, "Automatic detection of animals in mowing operations using thermal cameras," *Sensors (Switzerland)*, vol. 12, no. 6, pp. 7587–7597, Jun. 2012, doi: 10.3390/s120607587.

[9] A. F. Milton, F. R. Barone, and M. R. Kruer, "Influence of non-uniformity on infrared focal plane array performance," *Optical Engineering*, vol. 24, no. 5, pp. 855–862, 1985.

[10] H. Budzier and G. Gerlach, "Calibration of uncooled thermal infrared cameras," *Journal of Sensors and Sensor Systems*, vol. 4, no. 1, pp. 187–197, 2015, doi: 10.5194/jsss-4-187-2015.

[11] Y. Tendero, S. Landeau, and J. Gilles, "Non-uniformity Correction of Infrared Images by Midway Equalization," *Image Processing On Line*, vol. 2, pp. 134–146, Jul. 2012, doi: 10.5201/ipol.2012.glmt-mire.

[12] C.-W. Chen, C.-Y. Cho, Y.-F. Sun, T.-M. Chen, and C.-L. Su, "Low complexity photo sensor dead pixel detection algorithm," in *2012 IEEE Asia Pacific Conference on Circuits and Systems*, 2012, pp. 360–363. doi: 10.1109/APCCAS.2012.6419046.

[13] C.-H. Chan, "Dead pixel real-time detection method for image." Google Patents, Sep. 2009.

[14] G. T. Tchendjou and E. Simeu, "Detection, Location and Concealment of Defective Pixels in Image Sensors," *IEEE Trans Emerg Top Comput*, vol. 9, no. 2, pp. 664–679, 2021

[15] "Open Image Signal Processor (openISP)," https://github.com/cruxopen/openISP?tab=readme-ov-file#open-image-signal-processor-openisp.

[16] G. Yu and G. Sapiro, "DCT Image Denoising: a Simple and Effective Image Denoising Algorithm," *Image Processing On Line*, vol. 1, pp. 292–296, Oct. 2011, doi: 10.5201/ipol.2011.ys-dct.

[17] R. Dulski, P. Powalisz, M. Kastek, and P. Trzaskawka, "Enhancing image quality produced by IR cameras," in *Electro-Optical and Infrared Systems: Technology and Applications VII*, SPIE, Oct. 2010, p. 783415. doi: 10.1117/12.864979.

[18] F. Russo, "Performance Evaluation of Non-Local Means (NLM) Algorithms for Grayscale Image Denoising," in *2017 European Conference on Electrical Engineering and Computer Science (EECS)*, 2017, pp. 204–210. doi: 10.1109/EECS.2017.46.

[19] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, 1998, pp. 839–846.

[20] A. Omid-Zohoor, D. Ta, and B. Murmann, "PASCALRAW: Raw Image Database for Object Detection. Stanford Digital Repository," http://purl.stanford.edu/hq050zr7488.

[21] A. Ignatov, L. Van Gool, and R. Timofte, "Replacing Mobile Camera ISP with a Single Deep Learning Model," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 2275–2285. doi: 10.1109/CVPRW50498.2020.00276.

[22] Ekdeep Singh Lubana, "RAW cars dataset," https://github.com/EkdeepSLubana/raw_dataset/tree/master.

[23] Y. Tendero and J. Gilles, "ADMIRE: a locally adaptive single-image, non-uniformity correction and denoising algorithm: application to uncooled IR camera," in *Infrared Technology and Applications XXXVIII*, SPIE, May 2012, pp. 83531O-83531O–16. doi: 10.1117/12.912966.

[24] Z. He, Y. Cao, Y. Dong, J. Yang, Y. Cao, and C.-L. Tisse, "Single-image-based non-uniformity correction of uncooled long-wave infrared detectors: a deep-learning approach," *Appl. Opt.*, vol. 57, no. 18, pp. D155–D164, Jun. 2018, doi: 10.1364/AO.57.00D155.