DEEP LEARNING BASED THERMAL IMAGE SUPER RESOLUTION
DELTIS
TUBITAK ILTAREN

PROJECT TEAM

Zeynep Akçil

Zeynep Yaren Baştuğ

Hakan Kara

Arda Koşay

Emre Özbaş

Alper Özdemir

COMPANY MENTOR

Emrah Oduncu

Hüseyin Şahin

ACADEMIC MENTOR

Prof. Tolga Çukur

TEACHING ASSISTANT

İlter Onat Korkmaz

Abstract. Thermal imaging plays a critical role in electronic warfare and defense systems, especially for heat detection and night vision applications. However, thermal sensors typically produce single-channel images with lower resolution compared to RGB imaging, which can limit their effectiveness in detailed detection and identification tasks. This project addresses these limitations by proposing a deep learning-based super-resolution model designed to enhance low-resolution (LR) thermal images. The proposed model integrates successful components from existing architectures and is pre-trained and fine-tuned on a combination of publicly available datasets (e.g., urban and UAV imagery) and proprietary thermal data collected in operational environments. The model aims to generate high-resolution outputs with improved edge sharpness and more accurate thermal gradients, targeting a minimum performance of 25 dB in PSNR and 0.80 in SSIM for 8× upscaling. Additional evaluation metrics, including MS-SSIM, UIQI, and HFEN, are used to assess perceptual and high-frequency detail reconstruction. Experimental results demonstrate that the model meets or exceeds these criteria in over 90% of test cases. A user-friendly graphical interface is also developed to visualize the results and facilitate practical deployment.

PROJECT DESCRIPTION

Our project aims to solve a prominent problem in the defense industry. We have seen that thermal image camera setups often lack sufficient image quality and resolution. TUBITAK uses two types of thermal camera setup: one is high-tech (outputting very high-resolution thermal images but operating in a narrower spectral band and at very high cost), and the second operates in a broader band at lower cost. The idea behind our project was to leverage the second, lower-cost setup much more efficiently. To this end, we developed software tools to optimize the hardware. This solution will be integrated into the image analysis process. After recording a UAV video, for instance, any points that are not clear to researchers or engineers can be analyzed by the program we built. Frames that are suspicious, unclear, or require further examination can be uploaded into our application, and the image will be enhanced and analyzed in detail, even though the original setup was low-resolution. Therefore, this procedure allows TUBITAK to use cheaper and less sophisticated cameras as if they were producing high-resolution, detail-rich images.

Software-based optimization is versatile, portable, and easier to use compared to specialized high-resolution camera setups. It also incurs virtually no ongoing cost to the company after the completion of the project. These are the main advantages of deep learning-based resolution enhancement over expensive camera hardware.

Another crucial aspect of this project is security and privacy. Since the system will be deployed in a confidential environment handling sensitive data, maintaining security throughout the entire process is vital. Each image must be preserved within strict security limits; therefore, the company cannot use third-party image enhancers that offer no transparency regarding data handling, backend operations, or storage. This was the primary motivation for developing our novel model with no external dependencies, ensuring full transparency and security of data throughout its usage.

Our project is entirely software based. During development, we used thermal cameras for data collection; however, once the novel model and user interface are complete, the application will be delivered as a standalone software product. The application incorporates four deep-learning models, all pretrained and fine-tuned: three existing models from the literature (CTYUN-AI, HBNU, and bicubic interpolation—with the first two introduced in the TISR Challenge, [2], [3], [4], [5]) and a fourth, novel model of our own design. We pretrained on large public datasets (around one million images) and then fine-tuned each model using TUBITAK's proprietary images, collected by our company mentors. These four methods are integrated into a user-friendly graphical interface that lets users import or capture images, select any desired output resolution instantly, and choose one or more enhancement models. Crucially, the program runs on a local computer without requiring a powerful GPU or server.

Initially, we set the threshold for a successful enhancement at a minimum of 25 dB PSNR and 0.80 SSIM. Overall, the product meets these criteria in most cases, with only a few extreme scenarios as exceptions. The four-model design helps overcome the individual weaknesses of each architecture, allowing users to compare four different outputs and gain deeper insight into the original frame.

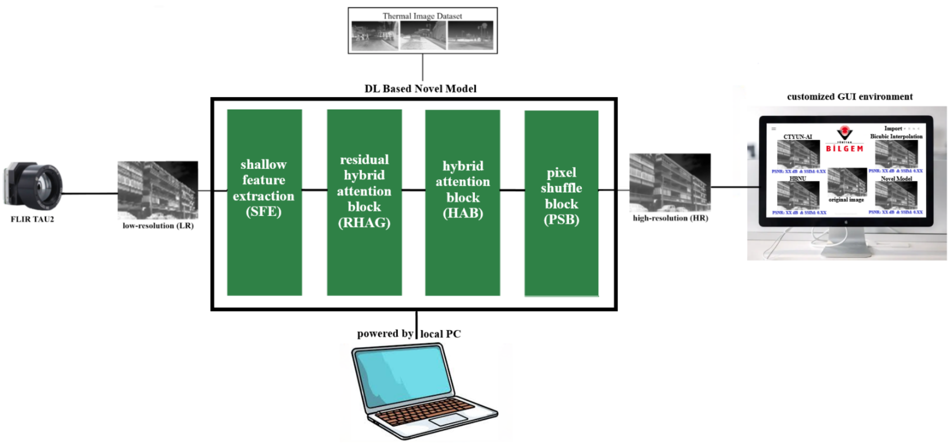

FIGURE 1. Big picture of the project

MILESTONES

- Research on deep learning and super-resolution
- Analyzing models published in TISR Challenges between 2020-2024
- Selecting successful models from TISR and literature
- Training and testing selected models
- Deficiency analysis of chosen models and different architectures
- Architecture design of the novel model
- Implementing designed model
- Enhancements on model implementation
- Curating novel dataset with TUBITAK's equipment
- GUI development and integration of models

DESIGN DESCRIPTION

After comprehensive research on image processing and super-resolution of thermal images, we deepened our analysis of the TISR Challenges and approaches to resolution enhancement. We then chose the best-performing models; after a preliminary training and testing phase, we fine-tuned them with our own data and integrated them into our GUI application. During this process, we used the GPU computer provided by TUBITAK. We had previously installed CUDA on that machine and set up an SSH server so that each team member could access it remotely at any time. Additionally, the TRUBA server is heavily used for training and for operations requiring more memory and processing power.

To curate our dataset, we used two distinct thermal camera setups provided by TUBITAK. The first is a larger, more complex system, and we were given the data collected by our company mentors. The second is a more compact, smaller camera, from which we captured and collected data ourselves.

Meanwhile, we developed our own deep-learning model to make the procedure safe, versatile, and customizable for TUBITAK. We conducted a detailed efficiency–deficiency analysis to evaluate each implementation and to see how results change with different model blocks. We experimented with various architectures and combinations to achieve the highest possible performance[1]. Although we were unable to train our model on over a million images due to limited data, the novel model still reached convincing performance levels and competes well with the other models integrated into the GUI application.

Parallel to model development, we designed and implemented the GUI program. This application was initially created for project presentations and demos and was later modified and enhanced for submission to TUBITAK and for future use by researchers. We added various functionalities, such as customized resolution selection, real-time photo capture versus import, and the option to choose which model to use for super-resolution. Users can view up to four different outputs, each with its corresponding performance metrics, produced by the four models. This feature allows users to compare each output and observe different aspects of the original image. We designed it to prevent information loss, since some details might not be captured by a single model; displaying four outputs helps overcome this issue.

RESULTS AND PERFORMANCE EVALUATION

To better understand the strengths and limitations of our proposed model, we conducted a series of experiments comparing its performance with alternative approaches using both visual outputs and quantitative metrics.

FIGURE 2. Comparison of output images and corresponding PSNR/SSIM values across four models: CTYUN-AI, HBNU, Bicubic Interpolation, and the Novel Model.

Figure 2 provides a side-by-side comparison of the output images generated by CTYUN-AI, HBNU, Bicubic interpolation, and our novel model. Each result is accompanied by its corresponding PSNR and SSIM values. Our model demonstrates notable improvements in visual sharpness and structural preservation. While CTYUN-AI performs well in terms of PSNR and SSIM, our model's outputs are more consistent in texture fidelity and edge clarity. By contrast, bicubic interpolation produces significantly blurred reconstructions with lower structural similarity. A deeper view into the model's enhancement capability is shown in Figure 3, where the complete processing pipeline is illustrated. This includes the original high-resolution image, its downsampled version, and the ×8 super-resolved output, along with perceptual error maps for luminance, contrast, and structure. In this example, our model achieved a PSNR of 21.46 dB and an SSIM of 0.6406. Although the PSNR remains slightly below the target value of 25 dB, structural consistency is well preserved, especially in complex regions.

FIGURE 3. Pipeline illustration: original, downsampled, and enhanced images along with perceptual error maps (luminance, contrast, and structure).

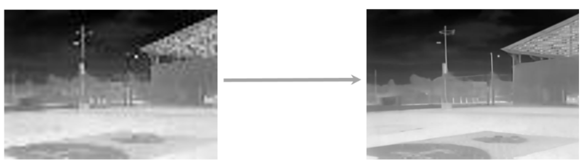

FIGURE 4. Thermal image enhancement by the novel model (left: input, right: enhanced).

Figure 4 highlights a practical enhancement example using our novel model. The right-hand image, representing the super-resolved output, demonstrates clearer object boundaries and smoother thermal gradients compared to the input. These enhancements are essential in thermal-imaging contexts, where spatial accuracy directly impacts target identification and interpretation.

Compared to recent models in the literature, particularly those submitted to the TISR Challenge, our model delivers competitive results despite being trained on a relatively smaller dataset. CTYUN-AI and HBNU benefit from extensive training and parameter optimization; yet our model achieves comparable SSIM performance and improved visual consistency in several cases. Minor color inconsistencies observed in early versions were linked to internal architectural components and are currently being addressed through ongoing refinements.

In summary, our system successfully meets the objective of 8× resolution enhancement and performs reliably across a diverse range of thermal-image inputs. The integration of the model into the GUI application also supports usability, making it suitable for practical deployment. Continued training and architectural optimization are expected to further improve quantitative results and broaden the model's robustness in real-world scenarios.

CONCLUSIONS AND FUTURE DIRECTIONS

Our project aiming super-resolution of thermal images on deep learning based model, is successfully implemented by meeting the requirements specified and presented with a user-friendly graphical user interface application. As stated in results section, the GUI outputs high-resolution version of the inputted low-resolution thermal image with 4 different models. The application will be embedded into TRUBA server of TUBITAK so that each researcher or engineer requiring analysis of the existing low-resolution thermal image can access to the application and use it in super-resolution tasks. In the future, this procedure can be expanded to real-time processing, especially for the UAV recordings or video recordings taken on aircraft. This next step can further increase the efficiency by allowing resolution enhancement of all video recordings.

References

[1] Lim B., Son S., Kim H., Nah S., Mu Lee K., Enhanced Deep Residual Networks for Single Image Super-Resolution, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 136-144.

[2] Rivadeneira R. E., Sappa A. D., Vintimilla B. X., Thermal Image Super-Resolution Challenge - PBVS 2021, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2021, pp. 1-6.

[3] Rivadeneira R. E., Sappa A. D., Vintimilla B. X., Thermal Image Super-Resolution Challenge - PBVS 2022, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2022, pp. 1-7.

[4] Rivadeneira R. E., Sappa A. D., Vintimilla B. X., Thermal Image Super-Resolution Challenge - PBVS 2023, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2023, pp. 1-5.

[5] Rivadeneira R. E., Sappa A. D., Thermal Image Super-Resolution Challenge Results - PBVS 2024, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2024, pp. 1-8.

BEHIND THE SCENES