

ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻՆ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏ

Հայկ Առնակի Գասպարյան

ԱՐԵՎԱՅԻՆ ՊԱՆԵԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅԱՆ ԲԱՐԵԼԱՎՈՒՄԸ
ՕԳՏԱԳՈՐԾԵԼՈՎ RGB-ԻՑ ՄՈՒԼՏԻՍՊԵԿՏՐԱԼ ՏՐՈՂՈՒՄ ԵՎ ՀԱՐՄՈՆԻԿ ՑԱՆՑԵՐ

Ե.13.05 - «Մաթեմատիկական մոդելավորում, թվային մեթոդներ և ծրագրերի համալիրներ»
մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի զիտական աստիճանի համար

ՍԵՂՄԱԳԻՐ

Երևան 2025

INSTITUTE FOR INFORMATICS AND AUTOMATION PROBLEMS OF THE NAS RA

Hayk Gasparyan

ENHANCING SOLAR PANEL ANALYTICS THROUGH RGB-MULTISPECTRAL
DECOMPOSITION AND HARMONIC NETWORKS

SYNOPSIS

of the dissertation for obtaining a Ph.D. degree in Technical Sciences on specialty 05.13.05
“Mathematical modelling, numerical methods and program complexes”

Yerevan 2025

Ատենախոսության թեման հաստատվել է ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում:

Գիտական ղեկավար՝

Ֆիզ. մաթ. գիտ. դոկտոր **Ս. Ս. Աղայան**

Պաշտոնական ընդդիմախոսներ՝

Ֆիզ. մաթ. գիտ. դոկտոր **Լ. Ն. Ասլանյան**

Ֆիզ. մաթ. գիտ. թեկնածու **Կ. Ի. Կարապետյան**

Առաջատար կազմակերպություն՝

Հայաստանի ազգային պոլիտեխնիկական համալսարան

Ատենախոսության պաշտպանությունը տեղի կունենա 2025թ. հուլիսի 10-ին ժամը 14:00-ին՝ ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտի մասնագիտացված խորհուրդ 037 «Ինֆորմատիկա» մասնագիտական խորհրդի նիստում հետևյալ հասցեով՝ Երևան, 0014, Պ. Սևակ փող. 1:

Ատենախոսությանը կարելի է ծանոթանալ ՀՀ ԳԱԱ ԻԱՊԻ գրադարանում:

Մեղմագիրն առաքված է 2025թ. հունիսի 10-ին:

Մասնագիտական խորհրդի գիտական քարտուղար ֆիզ. մաթ. գիտ. դոկտոր՝



Մ. Գ. Հարությունյան

The topic of the dissertation was approved at the Institute of Informatics and Automation Problems of NAS RA.

Scientific supervisor:

Doctor of phys-math sciences **S. S. Agaian**

Official opponents:

Doctor of phys-math sciences **L. H. Aslanyan**

Candidate of phys-math sciences **K. I. Karapetyan**

Leading organization:

National Polytechnic University of Armenia

The dissertation defense will take place on July 10, 2025, at 14:00, at the Specialized Council 037 «Informatics» at the Institute of Informatics and Automation Problems of NAS RA. Address: 1 P. Sevak Street, Yerevan 0014.

The Dissertation is available at the library of IIAP NAS RA.

The abstract is delivered on 10 June, 2025.

Scientific Secretary of the Specialized Council Doctor of phys-math sciences



M. E. Harutyunyan

Relevance of the Research

Solar Photovoltaic (PV) energy is now among the leading renewable energy sources worldwide. The global installed capacity for solar PV increased from roughly 40 Gigawatt (gW) in 2010 to more than 1.6 Terawatt (tW) by 2023, driven primarily by the decreasing cost of solar modules, favorable clean energy policies, and growing investments from both the government and private sectors. Despite this growth, the profitability of large-scale solar farms remains sensitive to different factors. The performance of PV modules largely depends on environmental factors (such as dust, snow, shades), their inner conditions (such as bad connections) and outer physical conditions (such as cracks, hot-spots), as demonstrated in Figure 1. Field data and financial analyses show that these hidden and environmental faults translate into three impacts:

- **Tiny hidden faults** including dust shading, hair-line cracks, bad connectors, and hot-spots trim about **5-9% of annual energy yield**. Field inspections show that roughly **12%** of modules already have such defects within the first two years, and unchecked hot-spots may even ignite fires.
- Only **1%** annual decrease in solar plant performance can translate into **€3 billion** lifetime revenue losses, significantly affecting the financial viability of large-scale installations. Current studies put the global bill for “invisible” performance loss at roughly **€3–15 billion** each year.
- Well-kept modules typically fade by only **0.5%** per year, staying close to **90%** of their power after **20 years**. Dirt, moisture and thermally stressed hot-spots speed that annual degradation beyond **1%**, cutting useful life and forcing earlier, costly replacements.

Therefore, monitoring and maintaining optimal solar panel performance is crucial to minimizing energy production losses. Currently, three primary approaches exist for monitoring PV installations, each with unique costs, accuracy, and operational constraints (as described in Table 1).



Figure 1: Solar PV fault examples.

Table 1: Comparison of PV-array monitoring approaches

Metric	Manual Inspection	Embedded Sensors	UAV Monitoring
Typical cost	$\leq \$1,500 \text{ MW}^{-1}$	$\approx \$10,000 \text{ MW}^{-1}$	$\approx \$150\text{--}500 \text{ MW}^{-1}$
Inspection time	$\leq 25 \text{ h MW}^{-1}$	Real-time, continuous	$\approx 3.75 \text{ h MW}^{-1}$
Maintenance cycle	Crew on demand	Sensor swap 2–5 yr	Battery recharge
Accuracy / error	20–30% miss-rate	$\pm 2\text{--}3\%$ sensor drift	1–5% false detections
Energy impact	$\approx 4\text{--}6\%$ loss yr^{-1}	$\approx 4\text{--}7\%$ loss yr^{-1}	$\leq 1\%$ loss yr^{-1}

Human manual inspection (visual, thermal and electroluminescence checks) can diagnose individual modules in depth but remains costly, about $\$1,500 \text{ MW}^{-1}$ and labour-intensive, requiring $\sim 25 \text{ h}$ per 1,000 panels. As crews typically sample only 10–25 % of modules and miss-rates run 20–30 %, residual energy losses stay at **4–6 % yr^{-1}** (residual losses are the portion energy yield still lost due to missed faults during inspection).

Embedded sensor networks offer automated continuous monitoring by embedding sensors directly into PV arrays. Despite their automation advantages, these systems have substantial initial costs, reaching approximately $\$10,000$ per MW. Moreover, sensors require regular replacement every 2–5 years due to sensor degradation and operational wear, resulting in additional expenses. Measurement inaccuracies from these sensors can reach more than 2%, potentially leading to significant annual residual losses of around 4–7%.

UAV-based monitoring is a promising alternative that addresses many limitations of the previous methods. It covers 1 MW of panels in minutes, cutting operations cost by 50–90 % ($\approx \$150 - 500 \text{ MW}^{-1}$) and field time by $\sim 85\%$ ($\approx 3.75 \text{ h MW}^{-1}$). Through advanced deep learning analytics, UAV systems achieve fault detection accuracies between 95 % and 99 %, which translates into estimated annual losses of just about 0.5.

These advancements, however, critically depend on the effectiveness of computer vision algorithms deployed on UAVs. The greater the accuracy and speed of these algorithms, the more reliable and efficient the overall PV monitoring process becomes, directly impacting the reduction of undetected faults and residual energy losses.

Camera Types. The first vital factor to consider when designing a UAV-based monitoring pipeline is the choice of camera. Different cameras significantly impact efficiency, performance, accuracy, and cost. In this field, four primary camera categories are used, each possessing distinct strengths and limitations. Table 2 summarizes the advantages and disadvantages of these camera types.

1. **Visible Red-Green-Blue (RGB) camera** captures three discrete spectral bands corresponding to red, green, and blue light. These images typically offer high spatial resolution but limited spectral detail. RGB cameras are also sensitive to noise and varying lighting conditions. However, their widespread availability and relatively low cost make them a common choice.
2. **Thermal camera** records emitted infrared radiation in a single spectral band, making it effective for detecting temperature variations on solar panels. Despite their effectiveness, thermal images typically possess lower spatial resolution and might miss fine structural details. Additionally,

high-quality thermal cameras are significantly more expensive than standard RGB cameras.

3. **Electroluminescence (EL) camera** generates single-band images capturing electroluminescent emissions from electrically biased panels. EL imaging effectively identifies micro-cracks, inactive cells, and contact defects. However, EL cameras require dark conditions and an external power source to energize panels. Thus, EL cameras cannot operate independently on UAV platforms and typically serve as supplementary devices. Additionally, their relatively large size and higher cost further constrain their UAV applicability.
4. **Multispectral camera** captures numerous narrow spectral bands (up to hundreds in the case of hyperspectral cameras), providing detailed spectral information, which is crucial for precise material discrimination. While multispectral imaging is highly effective for monitoring tasks due to its comprehensive spectral detail, these cameras are heavier and considerably more expensive, limiting their suitability for UAV applications.

Table 2: Comparison of camera types for solar panel monitoring

Camera Type	Bands	Advantages	Disadvantages
Visible RGB	3 (Red, Green, Blue)	High spatial resolution; widely available; low cost.	Limited spectral detail; sensitive to noise and lighting variations.
Thermal	1 (Infrared)	Effective at detecting temperature differences and hot spots.	Lower spatial resolution; high-end units are relatively expensive.
Electroluminescence (EL)	1 (Emitted under electrical bias)	Reveals micro-cracks, inactive cells, and contact defects.	Requires dark conditions and external power; bulky and expensive equipment.
Multispectral	4-400+ narrow bands	Rich spectral information; Non-sensitive to environmental effects.	Heavy and costly; less practical for UAV deployment.

Multispectral Cameras. Multispectral cameras record a scene with N narrow wavelength bands, rather than a standard sensor's three broad RGB channels. The raw data form a spectral cube $I(x, y, \lambda_k)$, $k = 1, \dots, N$, where each band index k captures a small wavelength window

$$\left[\lambda_k - \frac{\Delta\lambda}{2}, \lambda_k + \frac{\Delta\lambda}{2} \right],$$

centered at λ_k (for example, $\lambda_1 = 450$ nm, $\lambda_2 = 550$ nm, and so on). Each pixel is therefore described by an N -dimensional spectral vector

$$\mathbf{r}(x, y) = [I(x, y, \lambda_1), I(x, y, \lambda_2), \dots, I(x, y, \lambda_N)]^\top,$$

whose k -th component stores the reflected light in that band. Every element of $\mathbf{r}(x, y)$ corresponds to one narrow wavelength centered at λ_k (please see Figure 2). Since different materials peak at different wavelengths, they stand out in different bands, giving computer vision algorithms richer information than ordinary color images.

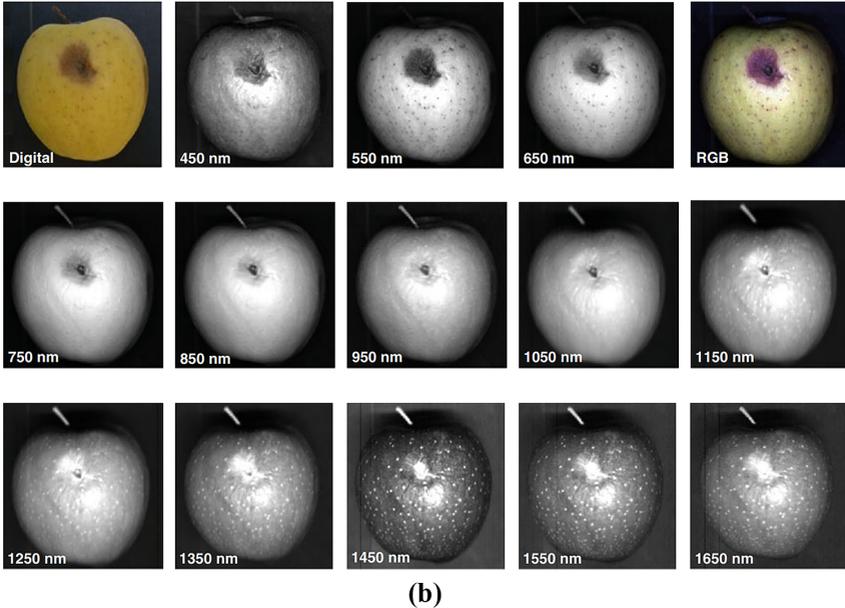
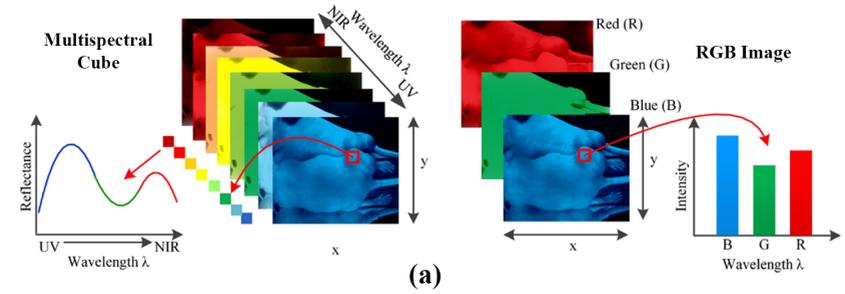


Figure 2: (a) Comparison of Multispectral and RGB images with their corresponding wavelengths, (b) example of multispectral image. Each band corresponds to a specific wavelength, representing different types of features

Figure 2 (a) shows the difference between this concept and a standard RGB image. On the left, the stack of colored slices represents the N individual bands that make up a multispectral cube; the curve represents the intensities of wavelengths λ_k in a small patch or pixel. On the right, an RGB camera compresses the same scene into only three wide channels, shown by the blue, green, and red bars; spectral information is averaged across bands into three channels.

Figure 2 (b) demonstrates the practical benefit: the healthy skin of an apple is bright in the near-infra-red (NIR) band at $\lambda = 750$ nm, whereas a hidden bruise absorbs NIR and appears dark in that slice. In the short-wave range (1350 nm to 1650 nm) additional small surface defects become visible. Once the bands are separated, these faults are easy to detect.

The same principle works for solar panel inspection. Silicon cells (solar panels are made by silicon) show characteristic absorption in the short-wave range (1100 nm to 1500 nm), so panels can be visible better in just two or three specific λ_k bands while remaining almost invisible in raw RGB. Algorithms that work on $\mathbf{r}(x, y)$ instead of (R, G, B) reach 10-20 % higher F_1 scores for segmentation and fault classification. Despite all these advantages, multispectral cameras can bring additional complexity and challenges.

The first **challenge** is their non-real-time operations, as the majority of multispectral systems depend on *spatial scanning*. The reason is that the sensor must traverse the scene or wavelength range sequentially.

The second **challenge** is more prolonged exposure or scan times, which also make it difficult to capture scenes that contain *motion*. Some solutions exist, such as mosaic filter arrays and light-field-based cameras. They bring a speed close to real-time multispectral capture. Still, they decrease the spectral and spatial resolutions, which are the main advantages of multispectral cameras. There is no reason to use these cameras without spectral and spatial high resolution.

The third **challenge** is the *cost*. Even low-cost multi/hyper-spectral cameras remain expensive (typically \$10-\$100k), heavy and power consuming, limiting their usage on UAV platforms. These factors limit the widespread deployment of multi-/hyperspectral cameras in drone-based applications.

The fourth **challenge** is the band size of multispectral images, which can have up to 400 channels (called hyperspectral camera), and their processing requires a lot of computational power.

Other **challenges** include accurate solar panel localization from the image, fault classification, and their efficiency and computational complexity. Balancing the high accuracy of algorithms under resource-constrained environments (e.g., on UAV platforms) remains a challenge in existing methods.

To overcome these challenges and limitations, this thesis introduces a Multispectral Decomposition (MD) method that generates multispectral bands from a standard RGB image. MD “decomposes” RGB data into separate spectral channels, offering multispectral-like features with affordable, lightweight RGB cameras. We demonstrate that MD significantly improves performance on critical monitoring tasks, addressing real-world challenges in solar panel inspection. Furthermore, new harmonic networks are proposed to reduce computational complexity, memory usage, lower energy consumption, and optimize latency, thereby enabling near real-time analysis in practical UAV inspection scenarios.

The Goal of the Thesis

The goal of this thesis is to address the challenges presented above and develop reliable, fast, and accurate deep learning based Computer Vision (CV) algorithms/models essential for an advanced automated solar PV monitoring system. To achieve this goal, the following technical tasks are set:

1. Develop a general reflectance extraction and multispectral decomposition network
2. Design an efficient spectral band selection strategy
3. Develop an efficient solar panel segmentation framework
4. Implement an efficient and accurate harmonic fault classification networks
5. Evaluate the system on key metrics, including accuracy, computational complexity, and show the generalization and practical applications of the proposed methods.

Structure of the Thesis

Figure 3 illustrates the structure of the thesis, its Chapters, their connections, and corresponding contributions in the overall system.

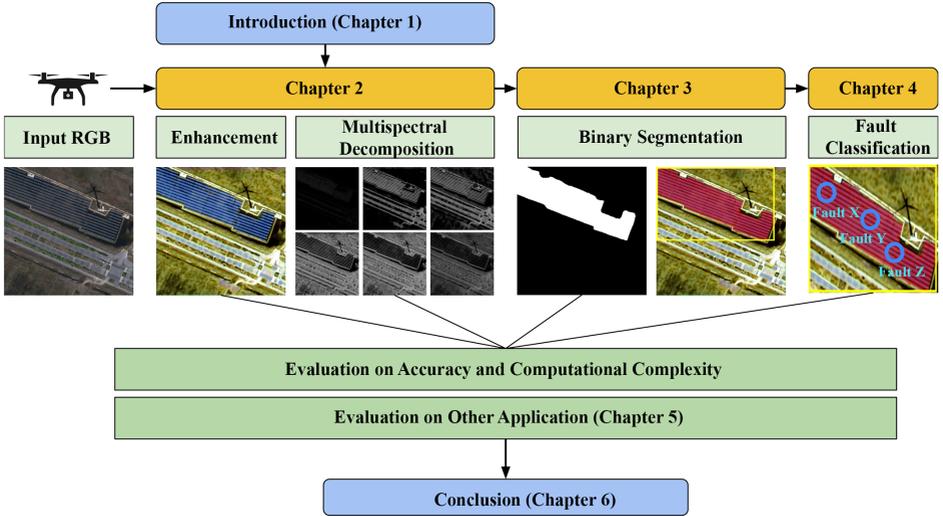


Figure 3: Structure of the thesis.

Chapter 1 serves as an introduction. It gives the definition and the motivation of the problem, highlights main challenges, sets the goal of the thesis and technical objectives, provides main contributions, lists the publications in the scope of this thesis, and emphasizes the impact of the thesis.

Chapter 2 aims to develop a generalized multispectral decomposition framework that addresses critical limitations of existing RGB-to-spectral reconstruction methods. Despite achieving very low pixel-wise reconstruction errors, current deep learning-based approaches still suffer from poor generalization across diverse lighting conditions and varying camera sensors. Besides, they cannot decompose a variable number of spectral bands, resulting in redundant, blurred, and less informative outputs. Moreover, they lose crucial details and lack evaluation on downstream remote sensing tasks. To overcome these challenges, this Chapter proposes **Retinex-based** spectral reconstruction pipeline. First, an illumination-invariant enhancement step is introduced, ensuring consistent performance across diverse real-world scenarios. The step is designed to extract intrinsic *reflectance* R and scene *illumination* L from UAV imagery. Here R contains the materials' colors and fine textures of the scene and is invariant to lighting, whereas L is a smooth, single-channel map that captures overall brightness. According to Retinex theory, an observed RGB image I can be expressed pixel-wise as

$$I(x, y) = L(x, y) R(x, y), \quad (x, y) \in \Omega. \quad (1)$$

The problem is that both L and R are unknown, and it makes this extraction an *ill-posed* problem: infinitely many pairs (L, R) satisfy the same product. To find an optimal solution, this Chapter

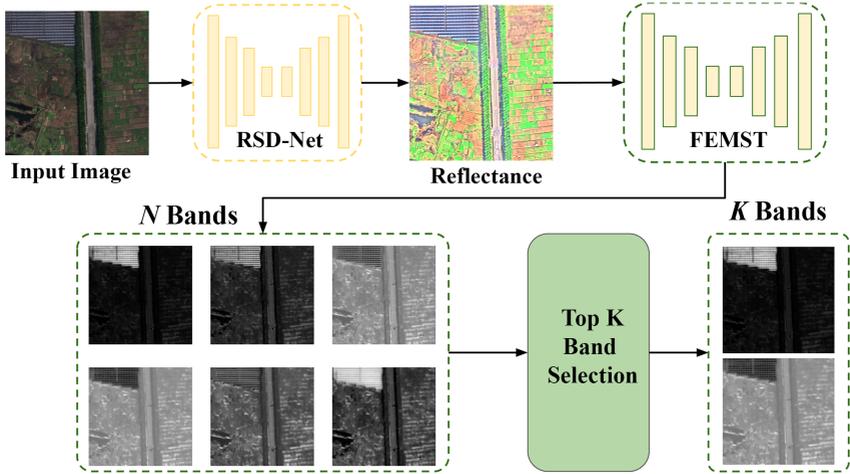


Figure 4: Overall pipeline of Multispectral Decomposition (MD) framework

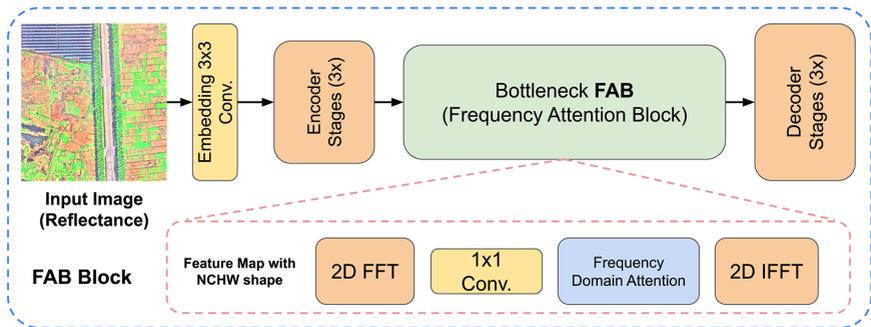


Figure 5: FEMST architecture.

introduces an RSD-Net [1], a two-branch network trained on matched *low-* and *normal-illumination* views of the same scene. Next, a developed Frequency Enhanced Multi Stage Transformer (FEMST) network decomposes 256 spectral bands. Figure 4 gives a high-level view of the proposed pipeline. Starting from a raw UAV RGB image, the framework proceeds through four sequential modules, each designed to solve one of the challenges stated above:

1. **RSD-Net (Retinex Decomposition).** The input image is first fed to *RSD-Net*, a lightweight 2-branch network that decomposes the scene into an illumination map L and an illumination-invariant reflectance R . Then we use reflectance for further processing, ensuring that subsequent stages operate on a physically grounded signal robust to changes in sun angle, cloud cover, or exposure/camera settings.

2. FEMST Core Network. The reflectance map enters the FEMST network (illustrated in Figure 5). The architecture of FEMST is inspired by, which uses multi-stage visual transformer blocks to expand the feature dimension progressively. Instead of their Spectral Attention Block (SAB) we use our proposed Frequency Attention Block (FAB). Learnable frequency attention masks select informative spectra in the frequency domain, and an inverse FFT returns the features to the spatial domain. This operation reduces redundant channels and keeps the model simple, enhancing generalization in data-scarce settings.

3. Adaptive Band Selection. An attention-based scoring block ranks the B output bands and selects the top K channels (user-specified) for downstream tasks.

The main results in this chapter are the improved quality of reflectance extraction and the better general decomposition of multispectral bands from RGB images. Strong computer simulations prove the superiority of RSD-Net and FEMST against other state-of-the-art methods. Key image similarity and reconstruction metrics (such as PSNR, SSIM, RMSE) are utilized for evaluation. Figure 6 demonstrates some bands predicted by the FEMST network. For each image, both reflectance and its original RGB components are used for the decomposition. It is observed that band variability significantly increased when reflectance is used as an input. Additionally, this chapter proposes an entropy-based (**ES**) measure for quantifying spectral information in bands, which is a gap in existing methods. Two

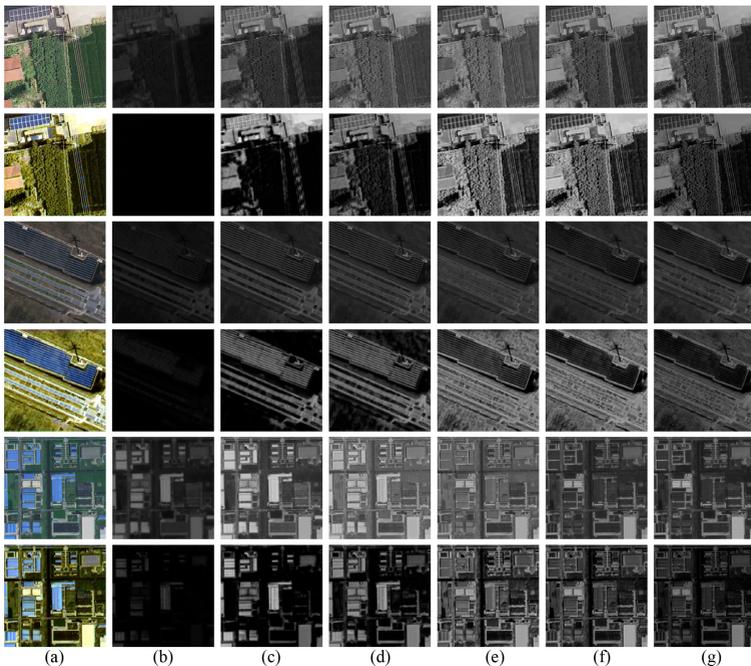


Figure 6: Band decomposition results for some band indexes: **a)** input and reflectance images **b)** index 1 **c)** index 5, **d)** index 10 **e)** index 15, **f)** index 25, **g)** index 30

Table 3: NTIRE 2022 HSI validation results.

Method	Params (M)	FLOPs (G)	MRAE ↓	RMSE ↓	PSNR ↑	ES ↑
HSCNN+	4.65	304.45	0.3814	0.0588	26.36	0.802
HRNet	31.70	163.81	0.3476	0.0550	26.89	0.812
EDSR	2.42	158.32	0.3277	0.0437	28.29	0.845
AWAN	4.04	270.61	0.2500	0.0367	31.22	0.861
HDNet	2.66	173.81	0.2048	0.0317	32.13	0.870
HINet	5.21	31.04	0.2032	0.0303	32.51	0.882
MIRNet	3.75	42.95	0.1890	0.0274	33.29	0.886
Restormer	15.11	93.77	0.1833	0.0274	33.40	0.887
MPRNet	3.62	101.59	0.1817	0.0270	33.50	0.890
MST-L	2.45	32.07	0.1772	0.0256	33.90	0.894
MST++	1.62	23.05	0.1645	0.0248	34.32	0.903
FEMST (Ours)	1.72	19.9	0.1405	0.0197	35.12	0.912

Table 4: Retinex decomposition results on LOL dataset.

Model	PSNR	SSIM	RMSE	MRAE
R2RNet	20.21	0.816	0.115	0.105
Retinex-2021	16.77	0.562	0.248	0.129
Deep Retinex	16.77	0.425	0.275	0.272
KinD++	21.80	0.829	0.102	0.098
RSD-Net (Ours)	22.49	0.845	0.085	0.083

benchmark datasets (**LOL** and **ARAD-1k**) are used for the training and comparison of the proposed 2 networks with existing state-of-the-art methods. Tables 3 and 4 summarize the results and show that the proposed networks outperform existing methods across different metrics.

Chapter 3 creates a Solar Panel Segmentation (SPS) framework, called MSS-Net (Multispectral Segmentation Network) [2], which is **first** to utilize Multispectral Decomposition (MD) for the segmentation task. The MD (from Chapter 2) solves challenges that most segmentation methods face during remote sensing image analytics. They often fail to adequately consider the intrinsic physical characteristics of solar panels, such as color and texture, which often translates into false positive errors. Besides that, remote sensing aerial images commonly have low resolution and various degradations, which pose a challenge in differentiating small panels from their surroundings. Moreover, their high computational demands and large trainable parameter size limit real-time applications and their generalization to different complex scenes. To address these challenges, this Chapter integrates a multispectral decomposition framework, introduced in Chapter 2. An efficient band selection mechanism is designed to select the optimal bands, containing rich information about solar panels. This minimizes the possible false positive errors of other SOTA methods. Moreover, a Chebyshev Transformation (CHT) layers are introduced and integrated in the network to keep it efficient and reduce trainable parameters, thus reducing overfitting and generalization errors. Figure 7 illustrates the architecture of MSS-Net, and Figure 8 shows some bands (j-n), their corresponding weight maps (c-g), the final guide image (i), and the final prediction of the solar panel mask (h).

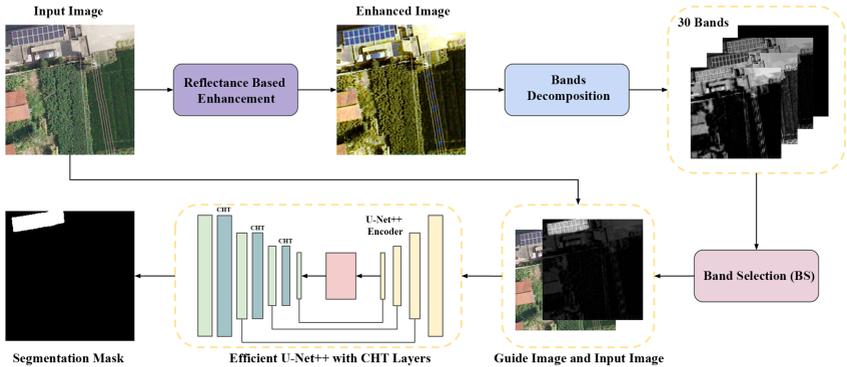


Figure 7: Overall architecture of MSS-Net

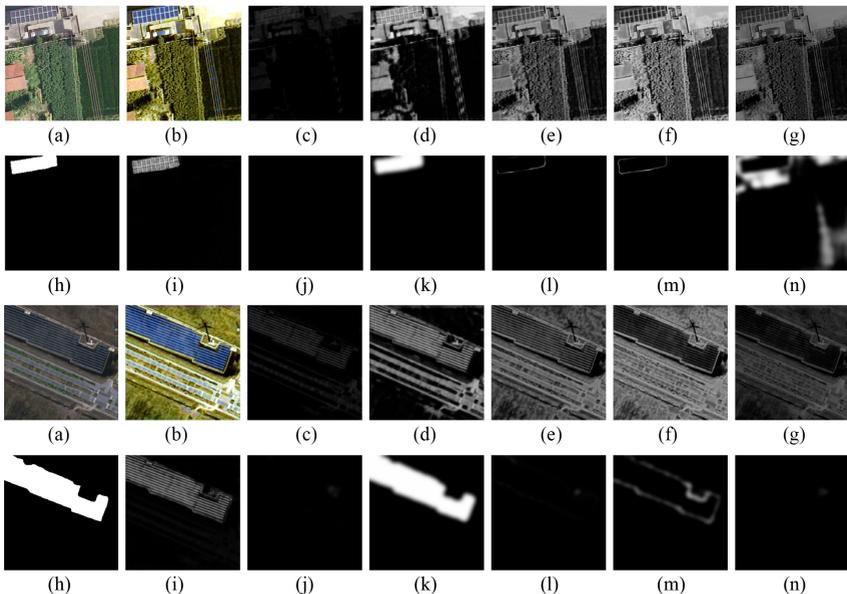


Figure 8: Band decomposition (c-g) with their corresponding weight maps (j-n), predicted by band selection module. a) and b) are original image and its reflectance respectively, i) is the combined guided image and h) is the final mask.

The presented method is validated on three publicly available SPS benchmark datasets (BDAPPV, PV, and DeepSolar). The comparison of the performance of MSS-Net is made against other methods, including CNN-based and transformer-based networks, showing that the proposed framework outperformed all SOTA methods across several key evaluation metrics, while reducing the trainable parameter size multiple times.

Chapter 4 solves the main fault classification task in the monitoring pipeline. Existing Visual Transformer (ViT)s are considered as SOTA models in classification tasks, but have limitations such as quadratic computational complexity and a large training dataset requirement. As demonstrated in this Chapter, some redundancy of learned features also arises from self-attention blocks. To address these limitations, this Chapter aims to develop lightweight and efficient Fast Fourier Transform Power Coefficient (FFT-PC) and Slant Fast Orthogonal Transformation (SFOT) modules ([3]) to replace existing self-attention layers of visual transformers, achieving comparable accuracy to vision transformers while significantly enhancing computational efficiency. In parallel, the Spatial Power Coefficient (S-PC) module uses architectural concepts from to enhance edges in the spatial domain, fusing its output with FFT-PC in the frequency domain.

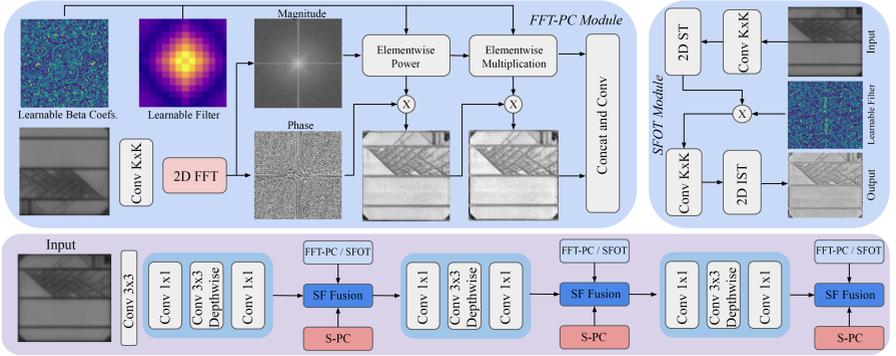


Figure 9: Overall architecture of MobileFFT or MobileSFOT.

Two new networks are developed based on FFT and Slant transformations, called *MobileFFT* and *MobileSFOT*, which are illustrated in Figure 9. The MobileFFT is based on the FFT-PC module, which begins by applying a convolution to spatial dimensions to adjust and reduce the channel size of the input feature map for the frequency transformations. Final output I_{out} is calculated by the following formulas:

$$F_{\beta}(I)(u, v) = |\mathcal{F}(I)(u, v)|^{\beta(u, v)} e^{j \arg(\mathcal{F}(I)(u, v))} \quad (2)$$

$$I_{out}(x, y) = \text{IFFT}(\alpha(u, v) \cdot F_{\beta}(I)(u, v)) \quad (3)$$

Where $\arg(\mathcal{F}(I)(u, v))$ is the phase of the Fourier transform, and $|\mathcal{F}(I)(u, v)|$ is the magnitude of the Fourier transform. Figure 9 illustrates the overall architecture of the FFT-PC block. In the final step, an additional convolution is used to restore the channel size to its original dimension before reduction. Similarly, MobileSFOT uses Slant fast orthogonal transformation.

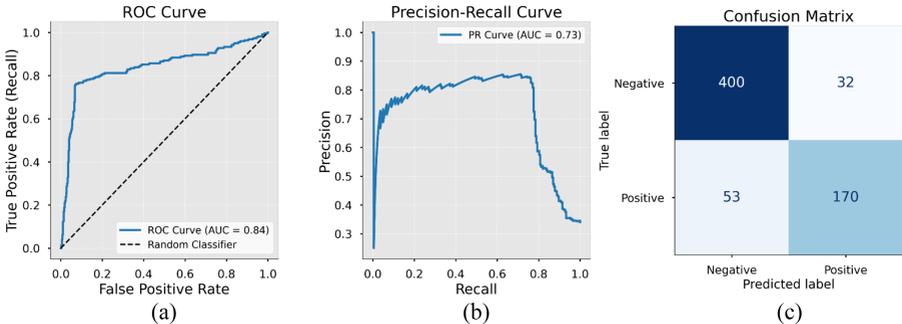
Compared to standard transformer-based attention blocks, the proposed approach achieves approximately $4\times$ fewer Giga Floating Point Operation (GFLOP)s (1.26 GFLOPs), $\sim 2.5\times$ fewer parameters (1.45 M), reduced inference latency, reduced Graphics Processing Unit (GPU) memory usage, and lower energy consumption (as shown in Table 5). This allows near-real-time Central Processing Unit (CPU) performance with low classification error (Figure 10, Table 6).

Table 5: Comparison of computational complexities

Model	Params.	GFLOPS	Thrp. GPU	Thrp. CPU	Energy	Power
<i>MobileNetV3</i>	1.50	0.15	98	875	5.17	99.2
<i>EfficientNet</i>	7.70	1.75	48	393	24.30	109.5
<i>DenseNet121</i>	7.00	7.46	37.6	292	31.90	124.5
<i>DaViT-T</i>	27.50	12.94	52.1	419	49.10	133.5
<i>GCViT-xxT</i>	11.48	3.90	49	300	30.50	112.0
<i>MobileViT-xs</i>	2.00	1.86	59	420	14.80	121.7
<i>DFFformer-s18</i>	28.00	9.93	24	164	66.40	112.8
<i>GFNet</i>	7.10	10.00	54	345	34.70	121.4
<i>MobileFFT</i>	1.45	1.59	97	558	9.80	105.4
<i>MobileSFOT</i>	1.45	1.26	71	506	8.40	102.7
<i>MobileFFT-light</i>	0.70	0.14	194	915	3.40	91.5
<i>MobileSFOT-light</i>	0.70	0.11	168	850	3.20	89.9

Table 6: Quantitative comparison of proposed method against others

Models	ELPV				Solar Faulty			
	Acc.	Prec.	Recall	F1	Acc.	Prec.	Recall	F1
MobileViT3	0.826	0.805	0.776	0.788	0.737	0.742	0.719	0.718
EfficientNet	0.829	0.820	0.766	0.784	0.709	0.721	0.691	0.688
DenseNet121	0.862	0.844	0.831	0.837	0.804	0.829	0.782	0.793
DaViT_T	0.833	0.810	0.795	0.802	0.759	0.747	0.747	0.746
GCViT_xxT	0.824	0.795	0.796	0.796	0.743	0.717	0.722	0.719
MobileViT_xs	0.841	0.818	0.806	0.811	0.774	0.786	0.752	0.767
DFFformer_s18	0.836	0.834	0.772	0.792	0.743	0.784	0.717	0.734
GFNet	0.823	0.794	0.791	0.793	0.737	0.725	0.703	0.703
MobileFFT-light	0.855	0.842	0.813	0.827	0.758	0.762	0.735	0.748
MobileSFOT-light	0.843	0.829	0.796	0.812	0.769	0.783	0.755	0.812
MobileFFT	0.874	0.866	0.851	0.858	<u>0.785</u>	<u>0.809</u>	<u>0.783</u>	<u>0.795</u>
MobileSFOT	<u>0.861</u>	<u>0.847</u>	<u>0.824</u>	<u>0.835</u>	0.828	0.846	0.836	0.840

**Figure 10:** Results of MobileFFT network on ELPV binary classification dataset. (a) ROC curve, (b) Precision-Recall curve, and (c) confusion matrix.

Chapter 5 evaluates the generalization and performance of the multispectral decomposition-based pipeline on other tasks. It proves the practical applicability of the methods in real-world applications. This Chapter proposes a novel solution (*MSSOD-Net*) [4] to solve Salient Object Detection (SOD) problem. SOD aims to identify the most visually prominent objects in images, crucial for tasks like image segmentation, visual tracking, autonomous navigation, and photo cropping (input and expected output examples are demonstrated in Figure 11).

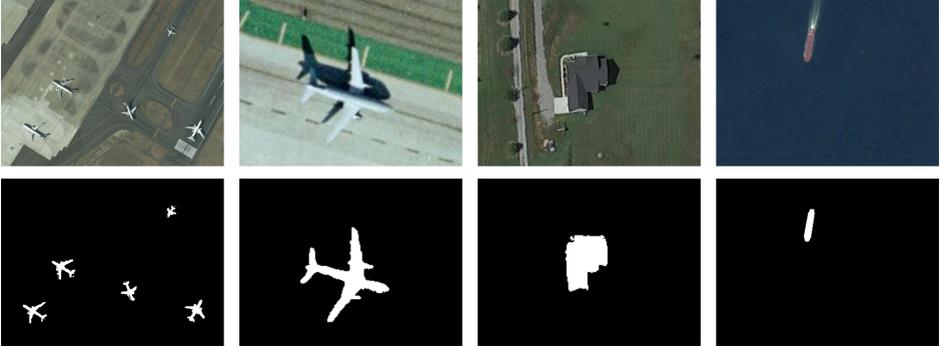


Figure 11: Examples of SOD. Input and expected outputs are in the first and second rows, respectively.

Overall architecture of the proposed pipeline is illustrated in Figure 12. Initially, the RGB image is enhanced and decomposed into multiple spectral bands, enhancing the representation of salient features by capturing richer spectral information. Next, the bands containing the most salient information are

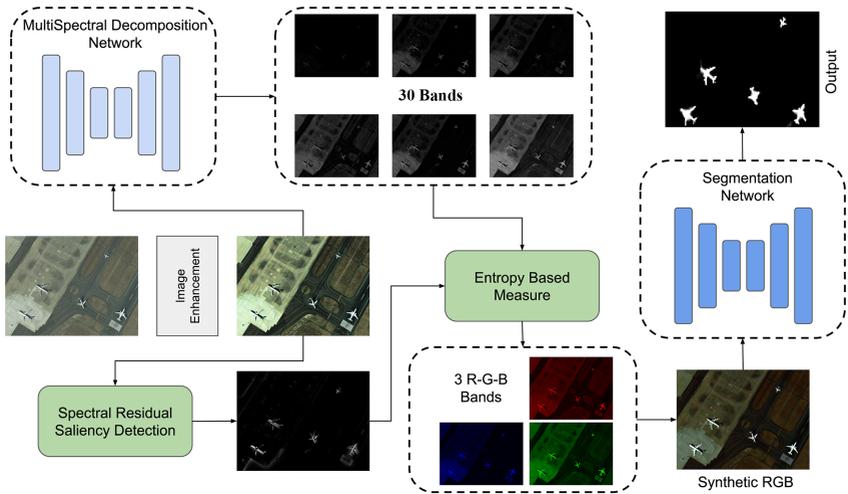


Figure 12: Overall architecture and workflow of MSSOD-Net.

identified and selected using a newly developed **entropy-based measure** operating in the frequency domain. A new synthetic RGB image is generated through the chosen bands, emphasizing salient objects more distinctly than the original input. Finally, a segmentation model processes the fused input (original and synthetic RGB), significantly improving the accuracy and reliability of salient object segmentation, especially in complex remote sensing scenarios.

To efficiently identify the most informative spectral bands, for each block B_{ij} of the image (with $i = 0, 1, \dots, H - 1$ and $j = 0, 1, \dots, W - 1$, where H and W are the image's height and width), a Fourier transform is performed to obtain the DC and AC components:

$$F_{ij} = \text{FFT}(B_{ij}), \quad F_{ij}^{\text{shift}} = \text{FFTShift}(F_{ij}).$$

$$\text{DC}_{ij} = |F_{ij}^{\text{shift}}(0, 0)|^\beta, \quad \text{AC}_{ij} = \sum_{x=1}^k \sum_{y=1}^k |F_{ij}^{\text{shift}}(x, y)|^\alpha.$$

α and β coefficients are selected experimentally at 0.6 and 2, respectively. A probability value p_{ij}^k is computed from the ratio of the AC and DC components from the k -th band:

$$p_{ij}^k = \frac{\text{AC}_{ij}^k}{\text{AC}_{ij}^k + \text{DC}_{ij}^k}.$$

Finally, we compute the entropy-based band selection measure H for each band k using the following formula:

$$H_{ij}^k = 20(-p_{ij}^k \log(p_{ij}^k)).$$

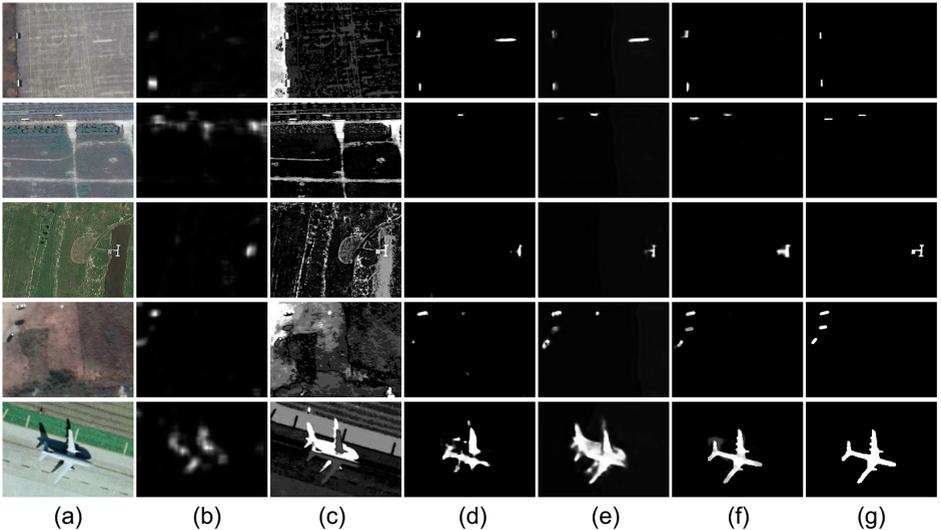


Figure 13: Comparison of MSSOD-Net with others. (a) input image, (b) SRS, (c) GCR, (d) DeepLabV3, (e) GCANet, (f) MSSOD-Net, (g) ground truth.

Comprehensive experiments on publicly available benchmark datasets validate the superior performance of MSSOD-Net compared to state-of-the-art approaches. Visual comparison of the proposed framework with other methods is presented in Figure 13. On the contrary, MSSOD-Net has successfully detected the salient objects and has better boundaries than others, which have unclear object boundaries, false positive detected pixels, and miss some parts of the objects. The experiments demonstrate that the proposed multispectral decomposition method effectively generalizes to a broader range of remote sensing applications beyond solar panel monitoring tasks.

Chapter 6 concludes the thesis and summarizes the results and key contributions.

Contributions of the Thesis

The key contributions of the proposed methods and frameworks are:

1. Introducing a novel Multispectral Decomposition (MD) framework featuring **RSD-Net** and **FEMST** for reflectance extraction and multispectral decomposition.
2. Developing **MSS-Net**, the first multispectral decomposition and **Chebyshev** transformation-based segmentation pipeline.
3. Proposing lightweight transformer modules based on **harmonic** transforms instead of self-attention layers in visual transformers, significantly reducing computational complexity, energy consumption, and redundant information in transformers.
4. Extensive benchmarking demonstrating the superiority of developed frameworks across performance and precision key metrics compared to existing state-of-the-art models

Practical Contributions of the Thesis

The methods proposed in this thesis have been practically applied and validated through extensive experimental evaluations, confirming their real-world utility in two key domains:

1. The proposed RGB-multispectral decomposition and harmonic networks were applied to UAV-based remote sensing tasks, specifically **salient object detection (SOD)** [4]. Rigorous experimentation on publicly available UAV imagery benchmark datasets (such as EORSSD) confirmed the methods' effectiveness, efficiency, and ability to surpass existing state-of-the-art algorithms.
2. The newly developed Retinex-based decomposition network was successfully employed for practical **low-light** and **nighttime visibility** restoration [1]. Experiments conducted under realistic conditions clearly demonstrated significant visibility improvements, establishing the network's superiority over traditional and contemporary state-of-the-art visibility enhancement methods.

This thesis addresses real-world problems approved and funded by the **FAST** Foundation under the **ADVANCE** Research Program, in cooperation with the Higher Education and Science Committee of the Republic of Armenia (Project No. 25FAST-1B001).

Potential Application Domains

This thesis makes a strong contribution to industry and academia by presenting a general framework that extracts multispectral-level insight from an RGB camera. Besides the main problems this thesis addresses, it can impact other domains as well:

1. **Bridges, pipelines, and power-line corridors** and other structures can be surveyed in a single pass with a low-cost camera.
2. **In agriculture**, UAVs can map water deficits, dry areas, yield variations, plant diseases, and animals. The system can easily detect and segment every category of interest using multispectral bands.
3. **Military** applications have tasks that rely on night-vision, infrared, or near-infra-red (NIR) imagery. The proposed methodology can enhance low-light scenes and extract infrared-like bands from the multispectral decomposition.
4. **In computer vision research**, a paired RGB-multispectral dataset can be generated for training and benchmarking.
5. **Biotech and healthcare** can benefit from the proposed framework by getting spectral channels from standard microscopes. This can be combined with automatic segmentation in tissue and cell studies.

List of Author's Publications

- [1] H. A. Gasparyan, S. A. Hovhannisyanyan, S. V. Babayan, and S. S. Agaian, "Iterative retinex-based decomposition framework for low light visibility restoration," *IEEE Access*, vol. 11, pp. 40 298–40 313, 2023.
- [2] H. A. Gasparyan, T. A. Davtyan, and S. S. Agaian, "A novel framework for solar panel segmentation from remote sensing images: Utilizing chebyshev transformer and hyperspectral decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–11, 2024.
- [3] H. Gasparyan, S. Agaian, and S. Wu, "Efficient lightweight networks for solar panel fault classification using el and rgb imagery," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–11, 2025.
- [4] H. A. Gasparyan, "A multispectral decomposition and frequency-based framework for salient object detection in remote sensing images," *Mathematical Problems of Computer Science*, vol. 62, pp. 93–111, 2024.

Ամփոփում

Հայկ Առնակի Գասպարյան

Արևային Պանելների Վերլուծության Բարելավումը Օգտագործելով RGB-ից Մուլտիսպեկտրալ Տրոհում և Հարմոնիկ Ցանցեր

Աշխատանքը նվիրված է արևային պանելների ավտոմատացված մշտադիտարկմանը՝ օգտագործելով RGB պատկերների մուլտիսպեկտրալ տրոհում և հարմոնիկ ցանցեր: Այն ուղղված է անօդաչու թռչող սարքերով (**ԱԹՍ**) իրականացվող տեսահսկման գործընթացի արդյունավետության, ճշգրտության և կիրառելիության զգալի բարելավմանը:

Աշխատանքի արդիականությունը պայմանավորված է այն փաստով, որ արևային պանելների շահագործման ընթացքում առաջացող բազմաթիվ խնդիրների (փոշի, ճեղքեր, միացման խափանումներ, ստվերներ) հայտնաբերումը և վերլուծությունը բարդ գործընթաց է, ինչը հանգեցնում է զգալի էներգետիկ կորուստների: Մինչ այժմ կիրառվող մեթոդները (մարդկային վերահսկում, ներկառուցված սենսորային համակարգեր, **ԱԹՍ**-ներ) ունեն բազմաթիվ թերություններ՝ բարձր գին, ցածր արագություն և մեծ սխալանք:

Աշխատանքի հիմնական նպատակն է մշակել արդյունավետ, արագ և հուսալի խորքային ուսուցման տեսողական ալգորիթմներ, որոնք ապահովում են արևային պանելների ավտոմատ մշտադիտարկման բարձր մակարդակ՝ հաղթահարելով գոյություն ունեցող մեթոդների սահմանափակումները:

Գիտական նորությունը և հիմնական լուծված խնդիրները հետևյալն են՝

1. Առաջարկվել է **RSD-Net** նոր երկնյուղ ցանց, որը կատարում է **RGB** պատկերից արտացոլման կոմպոնենտի առանձնացում: Այնուհետև կատարվում է արտացոլման պատկերից մուլտիսպեկտրալ գոտիների տրոհում: Մուլտիսպեկտրալ տրոհման համար օգտագործվել է **FEMST** ցանց, որը պարունակում է ձևափոխված Ֆուրյեի ձևափոխության հիման վրա ուսուցանվող նոր բլոկներ:
2. Մշակվել է **MSS-Net** մուլտիսպեկտրալ հատվածավորման (segmentation) ալգորիթմ՝ Չեքիշևի ձևափոխության և գոտիների ընտրության նոր մեխանիզմով, որը զգալիորեն նվազեցնում է կեղծ-դրական սխալանքը:
3. Առաջարկվել են **MobileFFT** և **MobileSFOT** հարմոնիկ ցանցեր՝ հիմնված հաճախականության տիրույթում աշխատող նոր տեսողական «**ուշադրության**» մեխանիզմների (attention mechanisms) վրա, որոնք զգալիորեն կրճատում են գոյություն ունեցող տրանսֆորմերներ մոդելների հաշվարկային և ալգորիթմիկ բարդությունները, ինչպես նաև հզորության և էներգիայի ծախսերը՝ պահպանելով բարձր ճշգրտություն:

4. Ներդրվել է նոր **էնտրոպիայի** վրա հիմնված ընտրության ցուցիչ, որը թույլ է տալիս ավտոմատ կերպով ընտրել առավել տեղեկատվական գոտիներ՝ մոլտիսպեկտրալ տրոհման արդյունքներից:
5. Ցույց է տրվել առաջարկվող ալգորիթմների ամբողջական շղթայի **ընդհանրացման** կարողությունը այլ խնդիրներում, որոնք կիրառվում են ԱԹՍ-երով հեռազննման մեջ:

Աշխատանքի կիրառական արդյունքներ

Այս թեզում առաջարկվող մեթոդները գործնականում կիրառվել և վավերացվել են լայնածավալ փորձարարական գնահատումների միջոցով՝ հիմնավորելով դրանց կիրառելիությունն ու արդյունավետությունը իրական պայմաններում.

- Առաջարկվող RGB-մոլտիսպեկտրալ տրոհումն ու հարմոնիկ նեյրոնային ցանցերը գործնականում կիրառվել են ԱԹՍ-երի վրա հիմնված հեռազննման խնդիրներում՝ մասնավորապես աչքի ընկնող օբյեկտների հայտնաբերման (SOD) առաջադրանքում: Հանրային հասանելի ԱԹՍ պատկերների տվյալների հավաքածուների (օրինակ՝ EORSSD) վրա իրականացված փորձարկումները հիմնավորել են առաջարկվող մեթոդների արդյունավետությունը, արտադրողականությունը և գոյություն ունեցող ժամանակակից ալգորիթմներին գերազանցելու կարողությունը:
- Retinex տեսության վրա հիմնված նոր մշակված տրոհման ցանցը հաջողությամբ կիրառվել է **թույլ լուսավորության և գիշերային տեսանելիության** վերականգման խնդիրներում: Իրական պայմաններում կատարված փորձարկումները հստակ արձանագրել են տեսանելիության շոշափելի բարելավում՝ հաստատելով ցանցի առավելությունը ինչպես դասական, այնպես էլ ժամանակակից տեսանելիության բարելավման մեթոդների նկատմամբ:

Տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի ստացման համար ներկայացվող ատենախոսության հիմնական արդյունքներն ու ներդրումները ասպահովում են արևային պանելների մշտադիտարկման ոլորտում նոր որակի ավտոմատացված վերլուծություն՝ գերազանցելով մինչ այժմ գոյություն ունեցող մեթոդները ինչպես ճշգրտությամբ, այնպես էլ արագությամբ և գործնական կիրառելիությամբ:

Заключение

Гаспарян Айк Арнакович

Улучшение анализа солнечных панелей с помощью многоспектральной декомпозиции RGB и гармонических сетей

Работа посвящена автоматизированному мониторингу солнечных панелей с использованием мультиспектральной декомпозиции RGB-изображений и гармонических сетей. Целью исследования является существенное повышение эффективности, точности и практической применимости видеомониторинга, осуществляемого с помощью беспилотных летательных аппаратов (БПЛА).

Актуальность работы обусловлена тем, что обнаружение и анализ многочисленных дефектов, возникающих в процессе использования солнечных панелей (таких как пыль, микротрещины, нарушения соединений, затенение), представляет собой сложную задачу, напрямую влияющую на эффективность их работы и приводящую к значительным потерям энергии. Существующие на сегодняшний день методы мониторинга (ручной осмотр, встроенные сенсорные системы и использование БПЛА) обладают рядом ограничений и недостатков: высокая стоимость, ограниченная скорость работы, низкая точность и другие недостатки.

Основная цель работы — разработка эффективных, высокоскоростных и надёжных визуальных алгоритмов глубокого обучения, обеспечивающих высокий уровень автоматизированного мониторинга солнечных панелей и способных преодолевать ограничения существующих подходов.

Научная новизна и основные решенные задачи заключаются в следующем:

1. Предложена новая двухветвевая сеть **RSD-Net**, выполняющая выделение компоненты отражения из RGB-изображения. Затем выполняется мультиспектральное разложение изображения отражения на каналы. Для мультиспектрального разложения использована сеть **FEMST**, содержащая новые обучаемые блоки на основе модифицированного преобразования Фурье.
2. Разработан алгоритм мультиспектральной сегментации (segmentation) **MSS-Net** с новым механизмом выбора каналов на основе преобразования Чебышева, который значительно снижает количество ложноположительных ошибок.
3. Предложены гармонические сети **MobileFFT** и **MobileSFOT** с новыми визуальными механизмами «внимания» (attention mechanisms), работающими в частотной области, которые значительно сокращают вычислительную и алгоритмическую сложность существующих трансформеров, а также затраты мощности и энергии, сохраняя при этом высокую точность.
4. Внедрен новый критерий выбора на основе **энтропии**, позволяющий автоматически выбирать наиболее информативные каналы из результатов мультиспектрального разложения.
5. Продемонстрирована **обобщаемость** полной цепочки предложенных алгоритмов на другие задачи, применяемые в дистанционном зондировании с помощью БПЛА.

Практические результаты работы

Методы, предложенные в этой диссертации, были практически применены и проверены посредством обширных экспериментальных оценок, подтверждающих их реальную полезность в двух ключевых областях:

- Предложенные RGB-мультиспектральная декомпозиция и гармонические сети были применены к задачам дистанционного зондирования на основе БПЛА, в частности, к обнаружению заметных объектов (SOD). Строгие эксперименты на общедоступных наборах данных эталонных изображений БПЛА (таких как EORSSD) подтвердили эффективность, действенность и способность методов превосходить существующие современные алгоритмы.
- Недавно разработанная сеть декомпозиции на основе теории Retinex была успешно использована для практического восстановления видимости при слабом освещении и в ночное время. Эксперименты, проведенные в реалистичных условиях, наглядно продемонстрировали значительное улучшение видимости, установив превосходство сети над традиционными и современными современными методами улучшения видимости.

Основные результаты и научный вклад диссертации, представленной на соискание учёной степени кандидата технических наук, обеспечивают качественно новый уровень автоматизированного анализа в области мониторинга солнечных панелей, превосходящий существующие методы по точности, скорости обработки и практической применимости.

