ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏ

Ստեփան Վահրամի Բաբայան

ԱՆԲԱՎԱՐԱՐ ՏԵՍԱՆԵԼԻՈՒԹՅԱՆ ՊԱՅՄԱՆՆԵՐՈՒՄ ՍՏԱՑՎԱԾ ՊԱՏԿԵՐՆԵՐԻ ՄՇԱԿՄԱՆ ՀԱՄԱԿԱՐԳ

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

ՍԵՂՄԱԳԻՐ

Երևան 2026

Institute for Informatics and Automation Problems

Stepan Vahram Babayan

A SYSTEM FOR PROCESSING IMAGES OBTAINED IN CONDITIONS OF INSUFFICIENT VISIBILITY

E.13.05 – “Mathematical Modeling, Numerical Methods, and Software Complexes”
Dissertation submitted for the degree of
Candidate of Sciences in Technical Sciences

SYNOPSIS

Yerevan 2026

Ատենախոսության թեման հաստատվել է Հայաստանի ազգային պոլիտեխնիկական համալսարանում:

Գիտական ղեկավար՝
տ.գ.թ., դոցենտ Ռ. Գ. Հակոբյան

Պաշտոնական ընդդիմախոսներ՝
տ.գ.դ., պրոֆեսոր Դ. Գ. Ասատյան
տ.գ.թ., ասիստենտ Գ. Վ. Սեֆիլյան

Առաջատար կազմակերպություն՝
Երևանի Կապի Միջոցների Գիտահետազոտական Ինստիտուտ

Ատենախոսության պաշտպանությունը տեղի կունենա 2026թ. հուլիսի 17-ին ժամը 14:00-ին՝ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտի 037 «Ինֆորմատիկա» մասնագիտացված խորհրդի նիստում, հետևյալ հասցեով՝ Երևան, 0014, Պ. Սևակ փող. 1:

Ատենախոսությանը կարելի է ծանոթանալ ԻԱՊԻ գրադարանում:

Սեղմագիրն առաքված է 2026թ. հունիսի 17-ին:

Մասնագիտական խորհրդի գիտական քարտուղար,
ֆիզ. մաթ. գիտ. դոկտոր՝ Մ. Ե. Հարությունյան

The topic of the dissertation was approved at the National Polytechnic University of Armenia.

Scientific Supervisor:
Candidate of Technical Sciences, Associate Professor R. G. Hakobyan

Official Opponents:
Doctor of Technical Sciences, Professor D. G. Asatyan
Candidate of Technical Sciences, Assistant G. V. Sefilyan

Leading Organization: Yerevan Telecommunication Research Institute

The dissertation defense will take place on July 17, 2026, at the Specialized Council 037 «Informatics», at the Institute of Informatics and Automation Problems. Address: Sevak Street, Yerevan 0014.

The Dissertation is available at the library of IIAP.

The synopsis is delivered on July 17, 2026.

Scientific Secretary of the Specialized Council,
Doctor of Phys-Math Sciences: M. E. Haroutunian

Relevance of the Research

Autonomous vehicles navigating urban and highway environments depend on object detectors that fail to recognize pedestrians and other vehicles under dense haze or in dimly lit conditions: proper dehazing can restore a 12% improvement in downstream detection performance, while methods that fail to generalize to real-world haze distributions introduce artifacts that actively degrade detection by over 20% relative to leaving the original hazy image unprocessed (Table 2). Surveillance cameras deployed at payment terminals, transit stations, and public infrastructure produce unusable nighttime footage in which face detection without enhancement misses more than 80% of individuals that become recoverable after appropriate low-light processing. Drone and unmanned aerial vehicle (UAV) platforms conducting photogrammetric surveys and aerial mapping encounter haze-induced color corruption that invalidates 3D reconstruction, spectral classification, and change-detection pipelines, forcing costly mission rescheduling. These are not aesthetic deficiencies: they are safety-critical, operationally costly failures that directly motivate the development of principled visibility restoration systems.

Automated analysis of visual data captured under adverse conditions is one of the central challenges in modern computer vision. Two degradation families account for the vast majority of real-world visibility failures: atmospheric haze and low-light illumination. Haze, fog, and smoke arise from the scattering of light by suspended particles, attenuating scene radiance and drastically reducing image contrast. Low-light conditions introduce severe underexposure, noise, color distortion, and, in nighttime scenes, additional complications from artificial light sources whose glow saturates nearby regions while leaving the rest of the scene in near-total darkness. Figure 1 illustrates the spectrum of atmospheric degradation; Figure 2 illustrates the components of the nighttime formation model.

Figure 1: Atmospheric scattering degradation. From left to right: (a) clear scene; (b) haze; (c) mist; (d) fog. Contrast and color are progressively attenuated as the proportion of scattered airlight increases.

Classical prior-based methods, such as the Dark Channel Prior, the color attenuation prior, and haze-line clustering, encode hand-crafted statistical regularities that break down under dense or spatially non-uniform real-world haze. Supervised deep learning replaced them with learned mappings, achieving high benchmark scores, but these models depend entirely on the training distribution. Since real paired training data is practically impossible to collect at scale, networks are trained on synthetically generated degradations that are substantially simpler and more uniform than what they encounter in deployment. This mismatch is apparent in downstream performance: methods that rank highest on the synthetic Haze4K benchmark, namely WDMamba and CasDyF-Net, achieve only $+5.6\%$ and

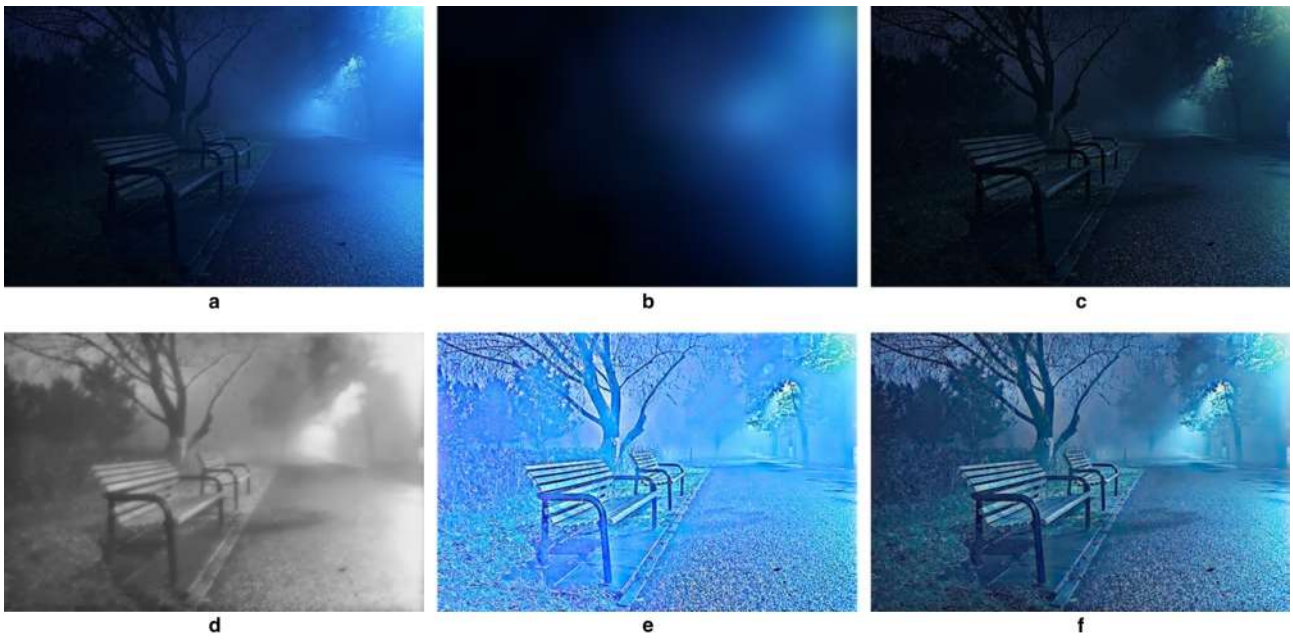

Figure 2: Visual decomposition of a nighttime low-light image: (a) input with artificial light source glow; (b) isolated glow component $G(x)$; (c) background after glow removal; (d) estimated illumination $L(x)$; (e) denoised reflectance $R(x)$; (f) final enhanced output.

comparable object detection gains on KITTI (Table 2), demonstrating that synthetic benchmark rank does not predict real-world utility. Neither method family incorporates explicit modeling of artificial light source glow, and their computational demands of 4,000–33,000 GFLOPs correspond to throughputs of 0.1–5 FPS, making real-time 4K deployment unattainable.

Two structural problems motivate this dissertation. The **synthetic-to-real generalization gap** arises because training data is almost exclusively synthetic: deep learning models overfit to the specific statistics of synthetic degradation and fail on real-world images. Leading low-light enhancement Transformers collapse from 25 dB on synthetic test sets to below 11 dB on real camera captures. The root cause is not architectural: no purely data-driven model can generalize beyond the statistical distribution it was trained on, and synthetic data cannot faithfully represent the full diversity of real-world atmospheric and illumination conditions.

The **quality-speed gap** is the second structural problem. The methods that come closest to bridging the generalization gap require 4,000–33,000 GFLOPs per image at throughputs of only 0.1–5 FPS on current hardware, placing real-time operation at 4K resolution entirely out of reach.

The present dissertation addresses both gaps through a unified strategy built on a common principle: physically-grounded probabilistic modeling. The physical formation laws governing haze and low-light are well established (the Atmospheric Scattering Model for haze and Retinex theory for low-light), and embedding them directly into the learning framework provides inductive biases that remain valid across all real-world conditions, not only those represented in the training distribution. Both teacher networks, described in Chapter 4, achieve state-of-the-art real-world generalization trained exclusively on synthetic data (Haze4K for dehazing; LOLv2-Synthetic for low-light enhancement), and knowledge distillation transfers this to compact EfficientViT students enabling real-time inference

at 4K resolution. The result is the first system to simultaneously achieve high real-world quality and real-time throughput across both degradation families, a combination the studied prior work has been unable to meet.

Purpose and Objectives

The purpose of this research is to develop, evaluate, and deploy a coherent system of algorithms for restoring visibility in images degraded by atmospheric haze or low-light conditions, addressing both the *quality* dimension (accuracy and real-world generalization) and the *efficiency* dimension (real-time inference at high resolution).

The following four specific objectives guide this work:

1. **Variational Bayesian dehazing for synthetic-to-real generalization.** Formulate single-image dehazing as a Bayesian inverse problem over the ASM, derive a multi-scale prior on the haze-free image *that promotes synthetic-to-real generalization*, and validate on real-world benchmarks without real-world paired training data.

2. **Iterative Retinex-based low-light image enhancement.** Design a physically motivated multi-stage architecture based on the formation model $I(x) = G(x) + R(x)\,L(x)$ that explicitly decomposes artificial glow, performs Retinex illumination and reflectance separation, and applies guided denoising.

3. **Knowledge-distilled real-time dehazing.** Extend the variational Bayesian framework with a Gaussian output layer and space-variant atmospheric light estimation, and distil its capabilities into a lightweight EfficientViT student that processes 4K images in real time *while preserving the teacher's synthetic-to-real generalization*.

4. **Knowledge-distilled real-time low-light enhancement.** Derive a spatially-varying per-pixel uncertainty from the Retinex illumination ratio and use it to construct an uncertainty-weighted distillation loss that trains a compact student matching teacher quality at real-time throughput.

Scientific Contributions

1. A multi-scale variational Bayesian framework for single-image dehazing is proposed that, within the studied literature, is the first to model all ASM components, the haze-free image, transmission, and atmospheric light, in a joint Bayesian model. The ASM serves as the physical likelihood, and a novel three-component prior on the haze-free image, combining a spatial ℓ_1 term preserving structural edges, a frequency-domain FFT term correcting spectral distortion, and a multi-scale quadratic term suppressing diffuse noise residuals, achieves superior synthetic-to-real generalization **without any real-world paired training data**, closing the generalization gap at the modeling level rather than through data augmentation or domain adaptation, where data-driven methods trained on synthetic distributions routinely fail. [1, 3]

2. RSD-Net introduces the **first supervised glow decomposition module** in the studied low-light enhancement literature that explicitly separates artificial light source glow from the scene background prior to Retinex decomposition. All prior approaches absorb glow into the illumination estimate, causing artifacts that propagate through the enhancement pipeline; the proposed iterative Retinex architecture, physically motivated by $I(x) = G(x) + R(x)\,L(x)$, with dedicated stages for glow removal, illumination and reflectance separation, and guided denoising, yields demonstrably better results on nighttime scenes than all existing unsupervised decomposition approaches. [2]

3. A **novel physics-grounded knowledge distillation framework for dehazing** is proposed, the first in the studied literature to combine knowledge distillation with a variational Bayesian dehazing framework, integrating KL divergence between student and teacher variational posteriors directly into the ELBO objective. It extends the ASM to space-variant per-pixel atmospheric light estimation to handle non-homogeneous real-world haze, and uses Gaussian teacher sampling as implicit data augmentation during student training, departing fundamentally from prior KD methods that transfer only deterministic point estimates. [4]

4. This work introduces the **first physically-grounded, spatially-varying distillation loss** for low-light image enhancement reported in the studied literature, derived analytically from the Retinex illumination ratio $\hat{t}(x) = \hat{L}(x)/\hat{L}_{\text{enh}}(x)$. Every prior knowledge distillation approach for low-light enhancement weights all spatial locations equally, despite dark pixels near zero admitting many plausible enhanced values while bright pixels have tightly constrained outputs; the per-pixel predictive variance $\hat{\sigma}^2(x)$ produced by this physically-grounded formulation drives a $1/\hat{\sigma}^2(x)$-weighted distillation loss that concentrates the student's gradient on reliably recoverable bright pixels and attenuates it in ambiguous dark regions. This mechanism has no precedent in the studied LLIE knowledge distillation literature.

Practical Significance

The methods of this dissertation are designed for settings where visibility degradation directly compromises operational safety or mission success, and where conditions cannot be controlled or rescheduled.

In autonomous driving, haze and low-light are among the primary causes of perception failure: object detectors miss pedestrians and vehicles when image quality degrades, and the consequences are safety-critical. A dehazing system that generalizes to real-world atmospheric conditions and runs at 4K resolution in real time can be integrated directly into the vehicle perception stack without additional hardware or offline post-processing.

In surveillance, nighttime CCTV footage at payment terminals, transit stations, and public infrastructure is often unusable for face detection without enhancement. The low-light framework restores the operational capability of the downstream detection system under the exact conditions it is deployed for, turning previously unusable nighttime footage into actionable data.

In aerial survey and photogrammetry, drone platforms cannot always reschedule missions to avoid haze: the dehazing framework enables 3D reconstruction and spectral analysis to proceed on imagery that would otherwise be discarded, directly reducing mission-rescheduling costs for operators.

Across all three domains, real-time throughput is not an optimization target; it is a deployment requirement. The knowledge-distilled student networks developed in Chapter 4 are what make the framework operationally viable, translating the quality of the teacher models into systems that run on live video streams.

Real-World Applications

OCUTECH LLC has integrated the Chapter 4 dehazing student into AutoWaypoints, a drone mapping and photogrammetry automation platform, enabling their enterprise clients to conduct aerial surveys in hazy conditions.

Ayatech has deployed the Chapter 4 knowledge-distilled low-light enhancement student within the security surveillance pipeline of their payment terminal network. The student processes live nighttime CCTV footage to enable face detection where unenhanced imagery yields zero or near-zero detections.

Parsl.ai, a startup co-founded by Stepan Babayan, has reached a verbal agreement with UP42, a geospatial data platform for Earth Observation, to integrate the dehazing pipeline; this agreement has not been formalized in a signed contract at the time of writing.

Structure of the Dissertation

The dissertation comprises four chapters preceded by an introduction and followed by a conclusion. **Chapter 1** reviews the state of the art in image dehazing and low-light enhancement, introduces the Atmospheric Scattering Model and Retinex theory, and formally states the four research gaps. **Chapter 2** presents the multi-scale variational Bayesian dehazing framework. **Chapter 3** presents RSD-Net, the iterative Retinex decomposition framework for nighttime low-light enhancement. **Chapter 4** develops the unified uncertainty-weighted knowledge distillation framework that bridges both teacher networks to compact real-time EfficientViT students.

Variational Bayesian Framework for Image Dehazing

Framework Overview

Single-image dehazing is formulated as a Bayesian inverse problem over the Atmospheric Scattering Model (ASM):

$$I = J \odot t + A \odot (1 - t), \tag{1}$$

where I is the observed hazy image, J is the latent haze-free image, t is the transmission map, and A is the atmospheric light. To handle the ill-posedness principally, a joint probability distribution is placed over all variables: $p(J, t, A, I) = p(J)\, p(t)\, p(A)\, p(I \mid J, t, A)$. The likelihood is Gaussian with noise precision λ_η; flat priors are assigned to t and A.

The informative prior on the haze-free image is the central design element:

$$p(J) \propto \exp\{-\lambda_q \mathcal{L}_q(J) - \lambda_l \mathcal{L}_l(J) - \lambda_F \mathcal{L}_F(J)\}\,, \tag{2}$$

where $\mathcal{L}_l$ (spatial ℓ_1, from a Laplacian prior) preserves structural edges; $\mathcal{L}_F$ (FFT-domain ℓ_1) corrects spectral distortion introduced by haze; and $\mathcal{L}_q$ (multi-scale quadratic) suppresses diffuse low-amplitude noise residuals that the ℓ_1 term under-penalizes. All three terms are applied at three spatial scales, enforcing multi-scale consistency that prevents overfitting to the uniform statistics of synthetic haze. Variational inference with Dirac delta posteriors yields a tractable Evidence Lower Bound (ELBO) training objective.

Three dedicated subnetworks implement the inference: **JNet** (CasDyF-Net backbone with a Dynamic Filter Network refinement stage) estimates the clean image; **TNet** (GCANet) estimates the transmission map; **ANet** (global average pooling) estimates the atmospheric light. All three are trained jointly so that ASM consistency couples their parameters. Figure 3 shows the complete architecture.

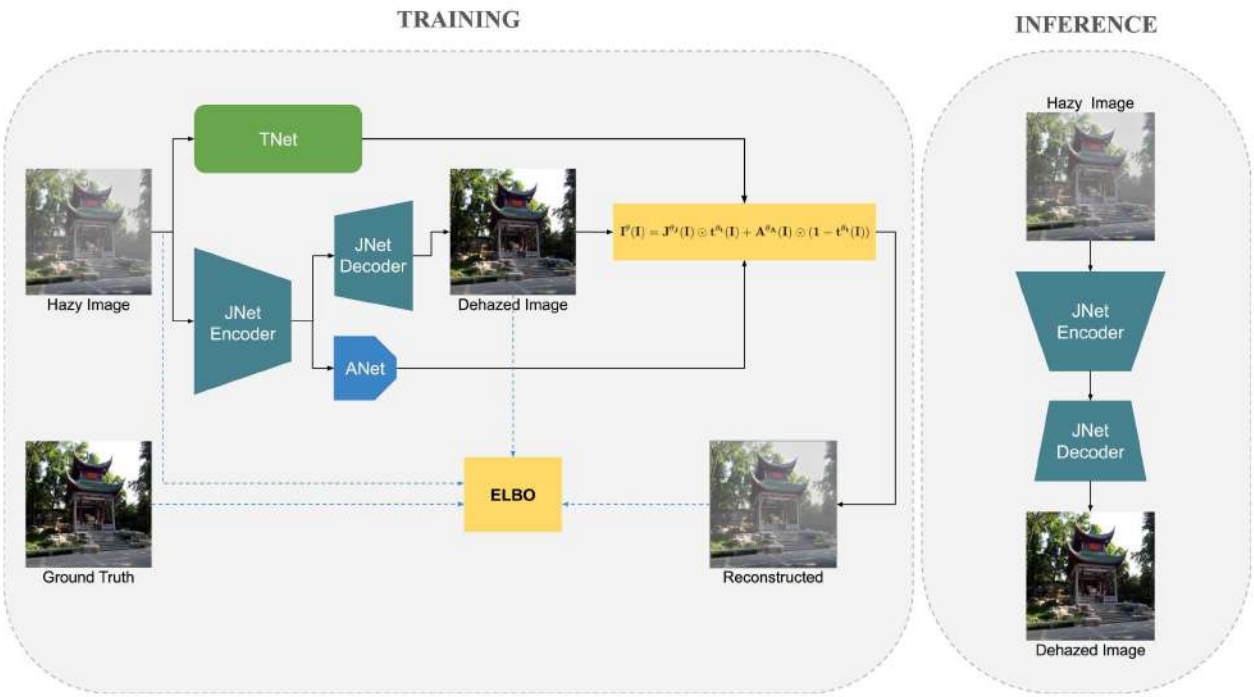

Figure 3: Architecture of the variational Bayesian dehazing framework. JNet estimates the clean image J; TNet estimates the transmission map t; ANet estimates the atmospheric light A. Joint optimization enforces ASM consistency. At inference, only JNet is used.

Experimental Results

All models were implemented in PyTorch and trained on NVIDIA A100 GPUs, using the 3,000 training pairs of the synthetic Haze4K dataset, and evaluated on its 1,000-image test set together with real-world benchmarks, without any fine-tuning. Table 1 reports PSNR/SSIM on four datasets. The proposed method ranks third on the synthetic Haze4K test set, behind WDMamba and CasDyF-Net; this reflects an intentional trade-off: the multi-scale prior regularization sacrifices a small amount

of in-distribution performance in exchange for physics-grounded constraints that prevent overfitting to synthetic rendering statistics. Those same top-Haze4K methods rank among the lowest on real-world benchmarks, and WDMamba achieves only +5.6% object detection improvement on KITTI versus the proposed method's +12.0%, confirming the overfitting hypothesis. The proposed method achieves first rank on O-Haze and NH-Haze and second rank on I-Haze, demonstrating the strongest synthetic-to-real generalization among all compared methods. Table 2 reports downstream object detection on KITTI: the proposed method achieves the highest improvement (+12.0%) among all dehazing methods, while methods optimized for synthetic benchmarks (FSNet, CasDyF-Net) actively degrade detection performance. Figure 4 shows representative qualitative comparisons.

Table 1: PSNR / SSIM on synthetic (Haze4K) and real-world benchmarks. All methods trained on Haze4K. **Bold**: best; underlined: second.

Method	**Haze4K** (synthetic)	**O-Haze** (real)	**I-Haze** (real)	**NH-Haze** (real)
WDMamba	**35.88** / 0.990	17.76 / 0.738	16.78 / 0.733	11.41 / 0.524
CasDyF-Net	35.71 / **0.992**	17.79 / 0.744	15.44 / 0.706	11.43 / 0.531
MB-TaylorFormer	34.93 / 0.989	18.26 / 0.666	16.24 / 0.709	11.82 / 0.509
FSNet	34.12 / 0.990	18.45 / 0.806	15.88 / 0.755	12.02 / 0.556
GridFormer	33.27 / 0.986	18.51 / 0.684	15.36 / 0.720	11.79 / 0.518
Ours	35.28 / 0.991	**18.91** / **0.815**	16.80 / 0.779	**12.26** / **0.576**

Table 2: Object detection on KITTI (200 synthetically hazed images, YOLOv8n).

Method	**Avg Det**	**Improv. vs Hazy (%)**	**Total Det**
Hazy Input	4.90	0.0	980
FSNet	3.86	−21.3	771
CasDyF-Net	4.77	−2.7	954
WDMamba	5.16	+5.6	1035
FFA-Net	5.40	+10.1	1079
GridFormer	5.42	+10.6	1084
DEA-Net	5.44	+10.9	1087
MB-TaylorFormer	5.44	+11.1	1089
Ours	**5.49**	**+12.0**	**1098**
Ground Truth	5.79	+18.2	1158

Iterative Retinex-Based Low-Light Visibility Restoration

Framework Overview

Nighttime low-light images require handling three simultaneous degradations: severe underexposure, spatially non-uniform sensor noise, and artificial light source glow. The image formation model is extended with an explicit glow term:

$$I(x) = G(x) + R(x)\,L(x), \quad R(x) = R'(x) + N(x), \tag{3}$$

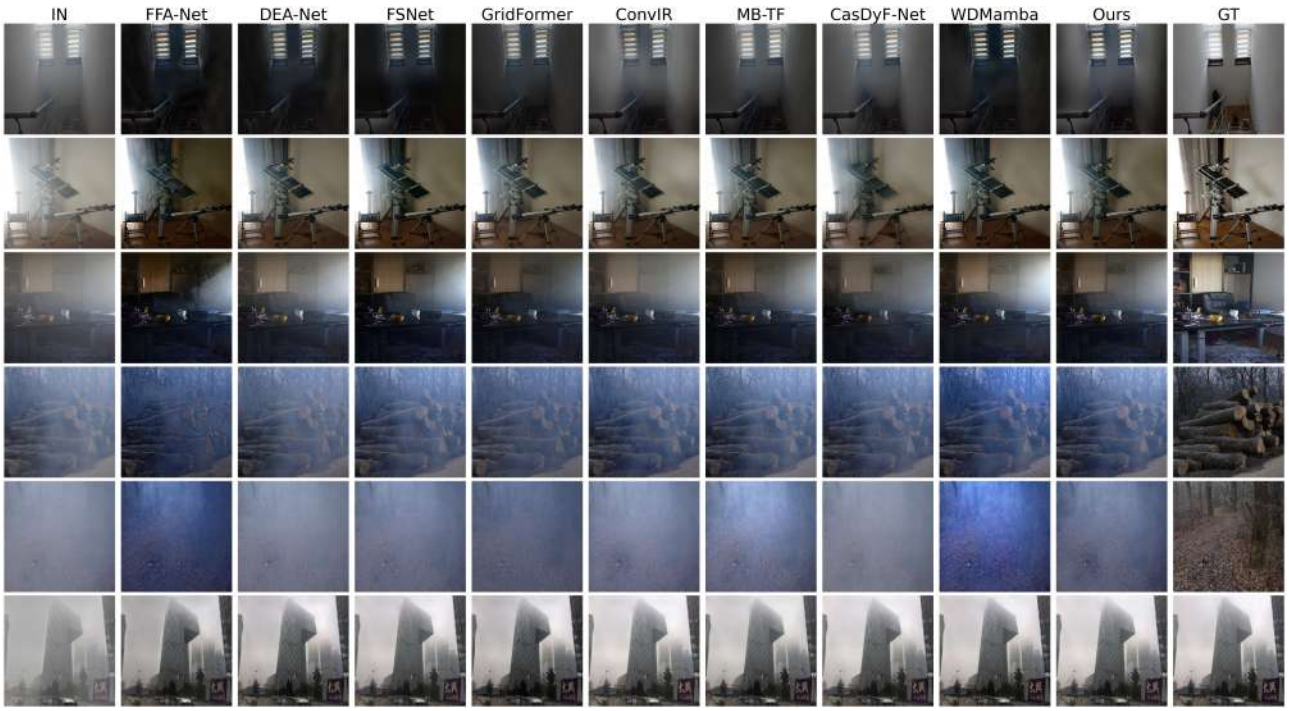

Figure 4: Qualitative dehazing comparison on I-Haze (rows 1–2), O-Haze (rows 3–4), and Haze4K (row 5). IN: hazy input; GT: ground truth.

where $G(x)$ is the glow, $L(x)$ is the illumination, $R'(x)$ is the clean reflectance, and $N(x)$ is noise. The target enhanced image is $J(x) = R'(x)\, L_{\text{enh}}(x)$.

RSD-Net (Retinex Supervised Decomposition Network) processes the input through four sequential stages (Figure 5): (1) a supervised **DeGlow** module (RCAN without long skip connections, trained on synthetically generated glow labels) separates the glow layer from the background; (2) a first **Retinex decomposition** (modified U-Net for illumination + RCAN for reflectance) decomposes the glow-removed image into illumination and reflectance; (3) a **joint enhancement network** applies Detail Reconstruction Module (DRM) processing and an attention-guided illumination enhancer, with a parallel attention-generation module that maps the illumination difference to a per-pixel denoising weight; (4) a **second Retinex decomposition** iteratively corrects residual illumination inconsistencies from the first stage, contributing the largest single-stage quality improvement.

Experimental Results

RSD-Net is trained on the LOL dataset (LOLv1, 485 real paired captures) and evaluated on unpaired real-world benchmarks without any fine-tuning. Table 3 reports no-reference perceptual quality (NIQE, BRISQUE) on three unpaired real-world datasets. RSD-Net achieves the best NIQE on all three and the best BRISQUE on two, confirming that the iterative decomposition produces outputs conforming to natural image statistics across diverse real-world conditions. Performance on the Real Nighttime Haze dataset, which combines glow, haze, heavy noise, and underexposure simultaneously, is particularly strong (NIQE 3.44), a consequence of the dedicated glow decomposition and the iterative two-stage design. Figure 6 shows qualitative comparisons on this most challenging benchmark.

Face detection on the DARK FACE benchmark confirms practical utility: RSD-Net reveals the

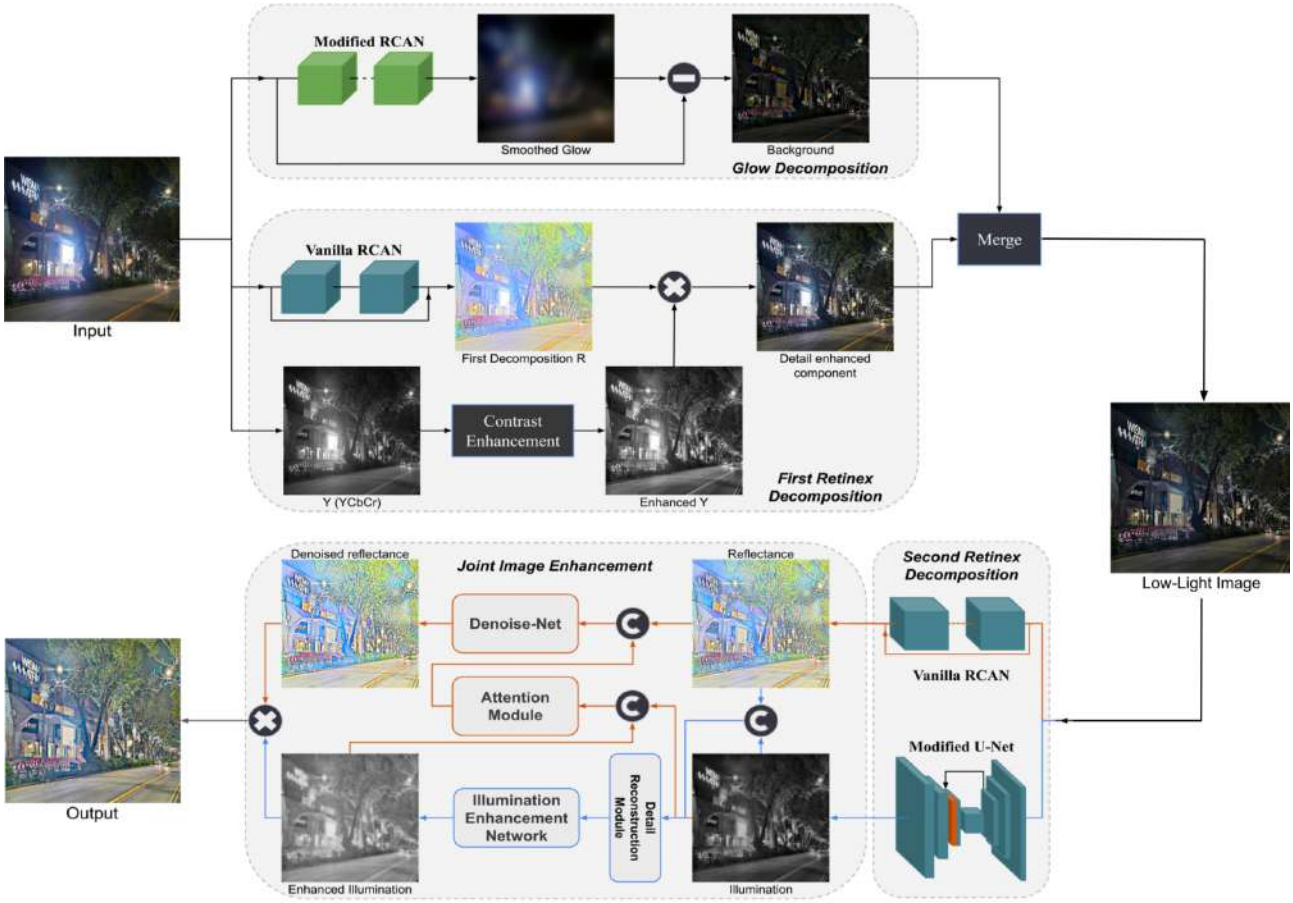

Figure 5: RSD-Net overall architecture: (a) glow decomposition block; (b) first Retinex decomposition; (c) second Retinex decomposition; (d) joint image enhancement and denoising.

most faces across all three evaluated images (8, 9, and 4 detections vs. 0, 3, and 0 for unenhanced input), outperforming all compared methods. Figure 7 illustrates the detection results visually.

Table 3: No-reference comparison (Avg / Std) on three real-world benchmarks. NIQE and BRISQUE: lower is better. **Bold**: best.

	TM-DIED		DICM		Nighttime Haze	
Method	NIQE↓	BRISQUE↓	NIQE↓	BRISQUE↓	NIQE↓	BRISQUE↓
KinD++	4.38	18.42	3.81	26.07	3.51	22.07
MIRNet	4.26	20.88	3.80	21.01	4.52	18.15
night-enh.	4.32	17.80	5.23	28.24	5.05	31.13
IAT	4.61	17.53	4.42	29.03	4.33	19.97
Bread	4.35	19.20	4.23	27.58	4.58	24.62
RSD-Net	**4.21**	**15.46**	**3.48**	**15.62**	**3.44**	**16.99**

Efficient Large-Scale Processing via Knowledge Distillation

Paradigm: Uncertainty-Weighted Knowledge Distillation

Chapters 2 and 3 establish high-quality teacher frameworks, but both are computationally prohibitive for real-time deployment: the dehazing teacher requires 4,890 GFLOPs at 4K resolution (0.5 FPS) and the LLIE teacher runs at approximately 1 FPS. Figure 8 illustrates the quality-speed trade-off

Figure 6: Visual comparison on the Real Nighttime Haze dataset. RSD-Net simultaneously handles glow, haze, and noise, where competing methods amplify or fail to remove these artifacts.

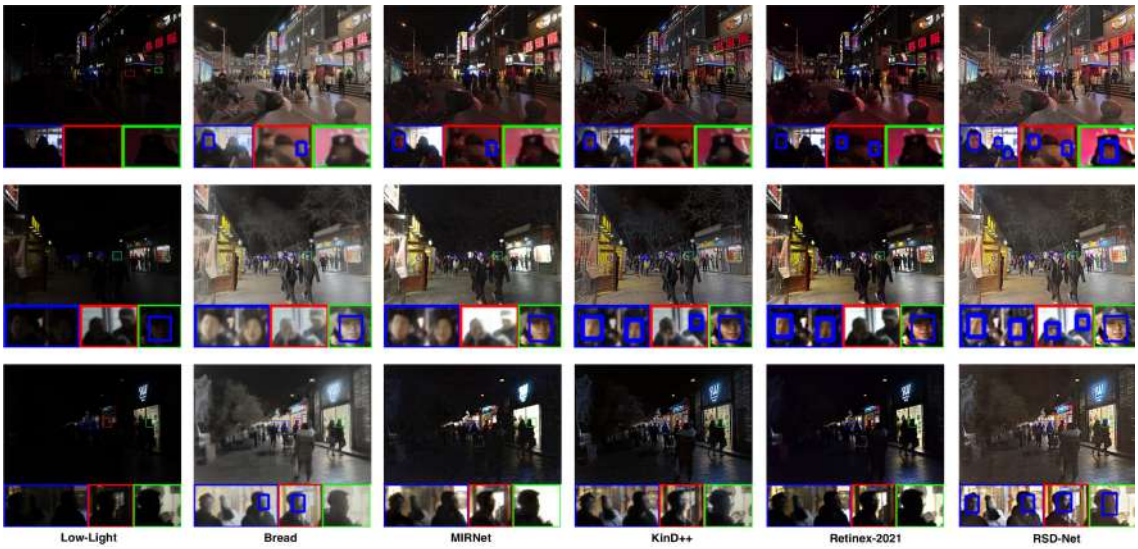

Figure 7: Face detection results on DARK FACE. RSD-Net is the only method that recovers faces in all three evaluated images simultaneously, achieving 8, 9, and 4 detections vs. 0, 3, and 0 for the unenhanced input.

that motivates this chapter: no existing method simultaneously achieves high real-world quality and real-time throughput.

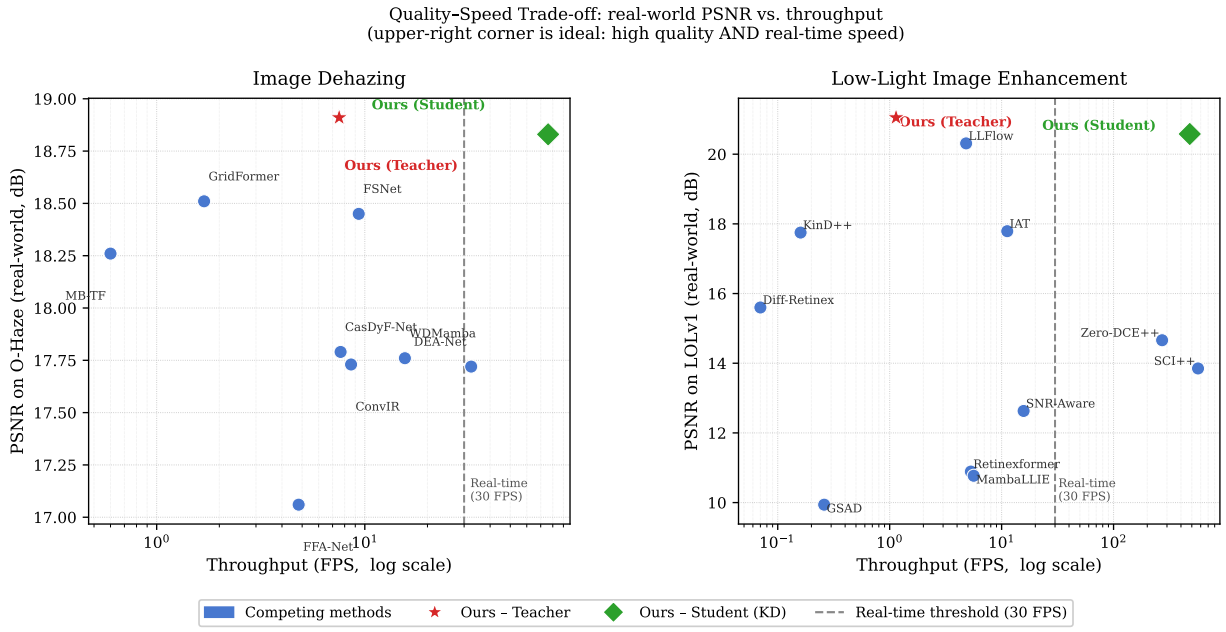

Figure 8: Quality–speed trade-off across dehazing (left) and low-light enhancement (right). Real-world PSNR on O-Haze and LOLv1 respectively. Heavy SOTA models (MambaLLIE, Retinexformer) score below lightweight baselines on real-world benchmarks while also being slow. The proposed Teacher–Student approach (star/diamond) is the only configuration achieving both high real-world quality and real-time throughput.

The key insight is that *probabilistic* teachers provide richer supervision than deterministic ones. By attaching a Gaussian output layer to each teacher and sampling from the resulting distribution, the student receives diverse training targets (implicit data augmentation) and a physically grounded uncertainty map. The $1/\hat{\sigma}^2(x)$-weighted distillation loss concentrates student gradients on reliably recoverable pixels and attenuates them where the teacher is uncertain, constituting a novel physically-grounded distillation mechanism. In both cases the student adopts an **EfficientViT** backbone, designed for memory-efficient processing at arbitrary resolution without architectural modification.

Dehazing: Space-Variant Teacher and Compact Student

The dehazing teacher is extended with per-pixel atmospheric light estimation (replacing global average pooling with a 4-layer dilated convolutional ANet) and a constant-variance Gaussian output layer ($\sigma^2 = 0.05$). The resulting 54-GFLOP EfficientViT student achieves **76 FPS at 4K resolution**, over $150\times$ faster than the teacher.

LLIE: Illumination-Ratio Uncertainty and Compact Student

For low-light enhancement the per-pixel predictive variance is derived analytically from the Retinex illumination ratio:

$$\hat{t}(x) = \frac{\hat{L}(x)}{\max(\hat{L}_{\text{enh}}(x),\, \varepsilon)}, \qquad \hat{\sigma}^2(x) = \sigma^2_{\text{min}} + (\sigma^2_{\text{max}} - \sigma^2_{\text{min}})(1 - \hat{t}(x)), \tag{4}$$

with calibrated values $\sigma^2_{\text{min}} = 0.002$, $\sigma^2_{\text{max}} = 0.06$. Dark pixels ($\hat{t}(x) \approx 0$) receive high variance; bright pixels receive low variance. The resulting 28.59-GFLOP EfficientViT-M0 student achieves **479 FPS at** 512×512, over $90\times$ the throughput of Retinexformer.

Experimental Results

Table 4 compares PSNR, SSIM, and NIQE on paired benchmarks; all methods are trained on the 900-pair LOLv2-Synthetic training split and evaluated on the 100-image LOLv2-Synthetic and 15-image LOLv1 test sets. The most notable result is the cross-domain behaviour: MambaLLIE and Retinexformer collapse from ≈25 dB on LOLv2-Synthetic to below 11 dB on the real-world LOLv1 benchmark (a 15 dB drop), while the Teacher suffers only a 2 dB drop (23.12 → 21.05 dB). The Student retains this cross-domain stability (20.58 dB on LOLv1) despite a $420\times$ throughput advantage over the Teacher.

Table 4: PSNR / SSIM / NIQE on LOLv1 and LOLv2-Synthetic (all methods trained on LOLv2-Synthetic). **Bold**: best; underlined: second.

Method	LOLv1	LOLv2-Synthetic
Zero-DCE++	14.66 / 0.47 / 7.80	17.58 / 0.81 / 4.42
IAT	17.79 / 0.69 / 6.37	23.50 / 0.82 / 3.97
LLFlow	20.31 / 0.84 / 5.43	23.42 / 0.93 / 4.33
MambaLLIE	10.77 / 0.47 / 5.86	**25.87** / **0.94** / 4.09
Retinexformer	10.89 / 0.46 / 6.31	25.67 / 0.93 / 4.01
Teacher	**21.05** / **0.84** / 4.58	23.12 / 0.89 / **3.85**
Student	20.58 / 0.82 / **4.57**	22.47 / 0.86 / 3.87

Table 5 quantifies the efficiency advantage. The Student (479 FPS) is $\approx 90\times$ faster than Retinexformer and $> 1800\times$ faster than diffusion-based methods, with only 28.59 GFLOPs and 2.30M parameters.

Table 5: Efficiency comparison at 512×512 on NVIDIA A100.

Method	FPS↑	Params (M)↓	Mem (GB)↓	GFLOPs↓
SCI++	**568**	**0.0003**	**0.0029**	0.13
Zero-DCE++	272	0.011	**0.0029**	**0.05**
MambaLLIE	6	4.39	0.018	62.38
Retinexformer	5	1.61	0.009	136.15
Diff-Retinex	0.07	58.91	0.317	794.91
Teacher	1	15.87	0.064	2776
Student	479	2.30	0.006	28.59

Downstream Task Evaluations

Face detection on DARK FACE. Table 6 reports face detection results on 300 randomly sampled DARK FACE images using YOLOv12n-face, with all methods trained on LOLv2-Synthetic. The Teacher achieves the highest face count (+170.7% over unenhanced input). The Student ranks third

(+164.4%), trailing the Teacher by only 2.3%, directly validating that the uncertainty-weighted distillation preserves downstream practical utility at 479 FPS. Figure 9 shows qualitative LLIE comparisons.

Table 6: Face detection on DARK FACE (300 images, YOLOv12n-face). **Bold**: best; underlined: second.

Method	**Avg Det**	**Improv. vs Dark (%)**	**Total Det**
Dark Input	0.80	0.0	239
KinD++	1.39	+74.5	417
IAT	1.50	+87.9	449
SNR-Aware	1.73	+116.7	518
GSAD	1.79	+124.3	536
SCI++	1.87	+135.1	562
LLFlow	1.93	+142.7	580
Retinexformer	1.99	+150.2	598
MambaLLIE	2.08	+161.5	625
Diff-Retinex	2.09	+162.0	626
Zero-DCE++	2.12	+166.1	636
Teacher	**2.16**	**+170.7**	**647**
Student	2.11	+164.4	632

Figure 9: Qualitative LLIE comparison on DICM (rows 1–2) and TM-DIED (rows 3–4). Teacher and Student produce consistently brightened outputs with natural colour across the full tonal range.

Object detection on KITTI. Table 7 reports object detection results after dehazing on 200 KITTI images. The Teacher achieves the highest improvement (+13.9%); the Student maintains 89% of the Teacher's detection gain (+12.4%) at 76 FPS, making it the only method simultaneously achieving top-tier detection improvement and real-time throughput. Figure 10 shows qualitative dehazing comparisons for the KD framework.

Table 7: Object detection on KITTI after dehazing (200 images, YOLOv8n). **Bold**: best; underlined: second.

Method	**Avg Det**	**Improv. vs Hazy (%)**	**Total Det**
Hazy Input	4.90	0.0	980
FFA-Net	5.40	+10.1	1079
GridFormer	5.42	+10.6	1084
DEA-Net	5.44	+10.9	1087
Base (Ch. 2)	5.49	+12.0	1098
Student	5.51	+12.4	1101
Teacher	**5.58**	**+13.9**	**1116**
Ground Truth	5.79	+18.2	1158

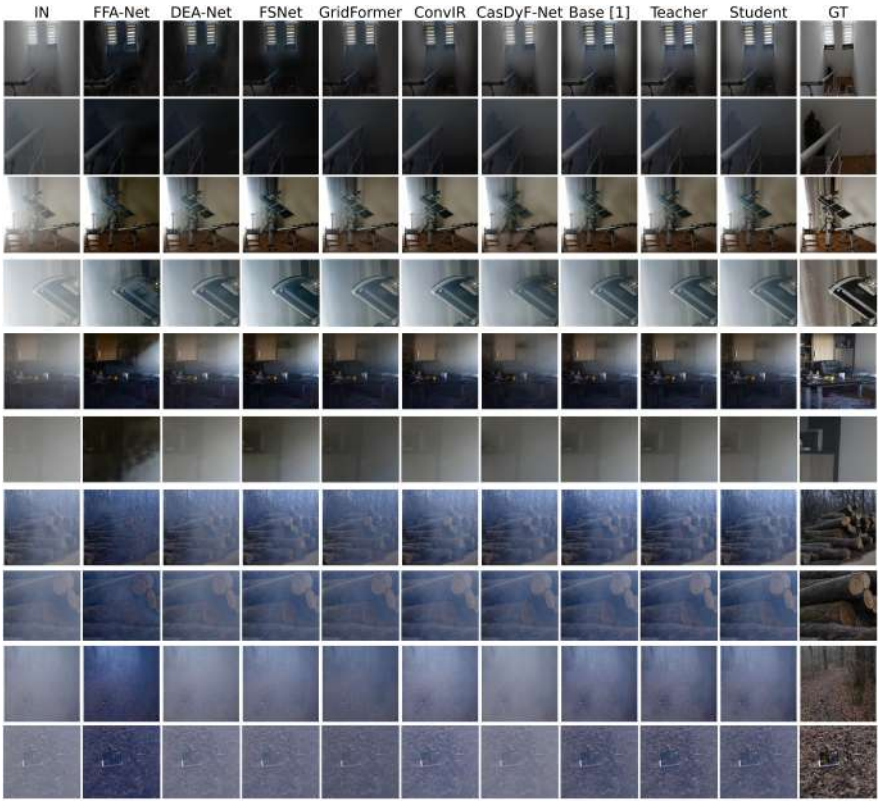

Figure 10: Qualitative dehazing comparison on O-Haze, I-Haze, and Haze4K for the KD framework. Teacher and Student closely match in visual quality while achieving over 150× speedup.

Main Contributions of the Work

1. A multi-scale variational Bayesian framework for single-image dehazing, the first in the studied literature to model all components of the Atmospheric Scattering Model in a single Bayesian framework, embedding the ASM as a physical likelihood and regularizing the solution through spatial, frequency-domain, and multi-scale priors that close the synthetic-to-real generalization gap, achieving state-of-the-art real-world performance without any real-world paired training data. [1, 3]

2. RSD-Net, an iterative Retinex decomposition architecture for nighttime low-light enhancement, incorporating the first supervised glow decomposition module in the studied literature that explicitly separates artificial light source glow prior to Retinex processing. [2]

3. A physics-grounded knowledge distillation framework for dehazing that integrates KL divergence between student and teacher posteriors into the ELBO, extending to space-variant atmospheric light estimation and producing a compact EfficientViT student running at 76 FPS at 4K resolution. [4]

4. An uncertainty-weighted knowledge distillation loss for low-light enhancement derived analytically from the Retinex illumination ratio, producing a compact student running at 479 FPS while preserving the teacher's real-world generalization.

List of Author's Publications

1. S. Babayan, F. J. Sáez-Maldonado, J. Mateos, and R. Molina, "Bridging the Synthetic-to-Real Gap in Single Image Dehazing," *Digital Signal Processing*, vol. 177, 2026.

2. H. Gasparyan, S. Hovhannisyan, S. Babayan, and S. Agaian, "Iterative Retinex-Based Decomposition Framework for Low Light Visibility Restoration," *IEEE Access*, vol. 11, pp. 40298–40313, 2023.

3. S. Babayan, F. J. Sáez-Maldonado, J. Mateos, and R. Molina, "Variational Bayesian Dehazing with Atmospheric Scattering Model for Synthetic-to-Real Generalization," in *Proc. 2025 IEEE ICIPW*, pp. 380–385, 2025.

4. S. Babayan, "Knowledge-Distilled Variational Bayesian Framework for Efficient Large-Scale Image Dehazing," *Proc. YSU, Phys-Math Sciences*, vol. 60, pp. 63–82, 2026.

Ամփոփում

Ստեփան Վահրամի Բաբայան

Անբավարար տեսանելիության պայմաններում ստացված պատկերների մշակման համակարգ

Աշխատանքը նվիրված է մթնոլորտային մշուշով կամ անբավարար լուսավորությամբ վատթարացած պատկերներում տեսանելիության վերականգնման ալգորիթմների համակարգի մշակմանը: Առաջարկվող մեթոդներն ուղղված են այն խնդիրներին, որտեղ տեսանելիության վատթարացումն ուղղակիորեն ազդում է անվտանգության կամ համակարգերի շահագործման արդյունավետության վրա, ինչպիսիք են ինքնավար տրանսպորտային միջոցների նավիգացիան և անվտանգության համակարգերում գիշերային տեսահսկողությունը: Հետազոտությունն ուղղված է երկու կենտրոնական խնդիրների միաժամանակյա լուծմանը՝ սինթետիկ տվյալների վրա սովորած մոդելների՝ իրական պայմանների համար ընդհանրացմանը և մոդելների՝ իրական ժամանակում աշխատելու ապահովմանը:

Արդիականություն. Մթնոլորտային մշուշը և անբավարար գիշերային լուսավորությունը տեխնիկական համակարգերի «տեսողության» խափանման հիմնական պատճառներն են իրական պայմաններում: Խիտ մշուշի կամ թույլ լուսավորության պայմաններում ինքնավար տրանսպորտային միջոցների օբյեկտների հայտնաբերման համակարգերը չեն ճանաչում հետիոտներին և այլ մեքենաներին: Գիշերային տեսահսկման տեսախցիկներից ստացված պատկերները օգտագործելի չեն, քանի որ դրանցում դեմքի ճանաչումը գրեթե անհնար է:

Սինթետիկ տվյալների վրա ուսուցված ժամանակակից խորը ուսուցման մեթոդները զգալի որակի անկում են ապրում իրական պայմաններում տեղակայվելիս (PSNR-ի ավելի քան 15 dB անկում իրական տվյալների հավաքածուներում)՝ պայմանավորված սինթետիկ ուսուցման պայմանների և իրական մթնոլորտային ու լուսավորության պայմանների միջև հիմնարար անհամապատասխանությամբ: Այս բացը հաղթահարելուն ամենամոտ մեթոդները պահանջում են 4000--33000 GFLOPs մեկ պատկերի համար՝ 0.1--5 FPS արագությամբ, ինչը պատկերների իրական ժամանակում մշակումն անիրագործելի է դարձնում: Գոյություն ունեցող մոդելներից ոչ մեկը **միաժամանակ** չի բավարարում սինթետիկից՝ իրական տվյալների ընդհանրացման, ինչպես նաև իրական ժամանակում մշակման պահանջները:

Գիտական նորույթ.

1. Առաջարկվել է բազմամակարդակ վարիացիոն Բայեսյան շրջանակ պատկերների մշուշազերծման համար, որը մեր ուսումնասիրած գրականության մեջ առաջինն է, որ Բայեսյան միասնական համակարգում մոդելավորում է մթնոլորտային ցրման (ASM) մոդելի բոլոր բաղադրիչները՝ մառախուղից զերծ պատկերը, թուլացման գործակիցը և մթնոլորտային լույսը: ASM-ը ներառվում է որպես ֆիզիկական հավանականության ֆունկցիա (physical likelihood), և առաջարկվում է եռաբաղադրիչ նախնական հավանականություն (prior)՝ ապահովելով **ընդհանրացում սինթետիկից՝ իրական տվյալների վրա առանց իրական պատկերներով ուսուցման օրինակների**:

2. Մշակվել է **իտերատիվ** Retinex դեկոմպոզիցիայի ճարտարապետությամբ համակարգ՝ մեր ուսումնասիրած գրականության մեջ առաջին վերահսկվող՝ **արհեստական լույսի մշակման մոդուլով** (DeGlow), որը հստակ առանձնացնում է արհեստական լույսի աղբյուրների փայլը Retinex դեկոմպոզիցիայից առաջ:

3. Մշակվել է գիտելիքի թորման շրջանակ **ֆիզիկապես հիմնավորված հավանականային ուսուցչով**՝ մշուշազերծման համար, որը մեր ուսումնասիրած գրականության մեջ առաջինն է համատեղում գիտելիքի թորումը վարիացիոն Բայեսյան մշուշազերծման շրջանակի հետ՝ ինտեգրելով Կուլբեկ--Լայբլերի շեղման կորստի ֆունկցիան ուսանողի և ուսուցչի ելքային բաշխումների միջև կառուցված ELBO նպատակային ֆունկցիայի մեջ, ինչպես նաև մթնոլորտային լուսավորության (Atmospheric Light) **տարածականորեն անհամասեռ գնահատումը**՝ իրական տարասեռ լուսավորության պայմաններում համակարգերի աշխատանքի համար:

4. Առաջարկվել է գիտելիքի թորման նոր $1/\hat{\sigma}^2(x)$ **գործակցով կշռված կորստի ֆունկցիա**, թույլ լուսավորության պայմաններում պատկերի բարելավման համար, որն անալիտիկորեն ստացվում է Retinex մոտեցման լուսավորության կոմպոնենտների հարաբերակցությունից՝ $\hat{t}(x) = \hat{L}(x)/\hat{L}_{\text{enh}}(x)$: Այս մեխանիզմը անալոգ չունի թույլ լուսավորությամբ պատկերների բարելավման համար օգտագործվող գիտելիքի թորման համակարգերի՝ մեր ուսումնասիրած գրականության մեջ:

Գործնական նշանակություն. Ատենախոսության շրջանակում առաջարկվող մեթոդները ներդրվել են առևտրային համակարգերում: Մշուշազերծման մոդուլն ինտեգրվել է AutoWaypoints օդային ֆոտոգրամետրիայի հարթակում (OCUTECH LLC): Թույլ լուսավորությամբ պատկերների բարելավման մոդուլը (Գլուխ 4) ինտեգրվել է Ayatech ընկերության՝ դրսում տեղադրվող վճարային սարքավորումների անվտանգության վերահսկման համակարգում: Ստեփան Բաբայանը, իր համահիմնադրմամբ ստեղծված Parsl.ai ստարտափի միջոցով նաև ձեռք է բերել բանավոր համագործակցության պայմանավորվածություն UP42 աշխարհատարածական տվյալների հարթակի հետ:

- Մշուշազերծման կոմպակտ ուսանողը մոդելն աշխատում է 76 FPS արագությամբ 4K UHD պատկերների դեպքում (ուսուցչից ավելի քան $150\times$ արագ)՝ բարելավելով օբյեկտների հայտնաբերումը KITTI հավաքածուի վրա $+12.4\%$-ով:

- Լուսավորության բարելավման կոմպակտ ուսանողը հասնում է 479 FPS-ի 512×512 չափի պատկերների վրա (Retinexformer-ից ավելի քան $90\times$ արագ) և հայտնաբերում է 164.4%-ով ավելի շատ դեմքեր DARK FACE տվյալների բազայի պատկերներում՝ ուսուցչին զիջելով ընդամենը 2.3%-ով:

- Երկու ուսանողներն էլ զբաղեցնում են առաջին կամ երկրորդ տեղը NIQE ցուցանիշով յուրաքանչյուր փորձարկված հավաքածուի վրա, ներառյալ իրական DICM և TM-DIED հավաքածուները՝ առանց լրացուցիչ ուսուցման (fine-tuning):

Заключение

Степан Ваграмович Бабаян

Система обработки изображений, полученных в условиях недостаточной видимости

Работа посвящена разработке системы алгоритмов восстановления видимости на изображениях, деградированных атмосферной дымкой или недостаточным освещением. Предложенные методы ориентированы на задачи, где ухудшение видимости непосредственно влияет на безопасность или операционную эффективность: навигация автономных транспортных средств и ночное видеонаблюдение в системах безопасности. Исследование направлено на одновременное решение двух ключевых проблем: обобщение с синтетических данных на реальные условия и достижение производительности реального времени.

Актуальность. Атмосферная дымка и недостаточное ночное освещение являются главными причинами отказа систем технического зрения в условиях реальной эксплуатации. Системы обнаружения объектов в автономных транспортных средствах не распознают пешеходов и автомобили в условиях плотной дымки или при слабом освещении; ночные камеры видеонаблюдения формируют непригодные изображения, при которых распознавание лиц невозможно. Современные методы глубокого обучения, обученные на синтетических данных, испытывают значительное падение качества при развёртывании в реальных условиях (снижение PSNR более чем на 15 dB на реальных наборах данных): это обусловлено фундаментальным несоответствием между синтетическими условиями обучения и реальным многообразием атмосферных и световых явлений. Методы, ближе всего подходящие к преодолению данного разрыва, требуют 4000--33000 GFLOPs на одно изображение при скорости 0.1--5 FPS, что делает обработку в реальном времени при разрешении 4K неосуществимой. Ни одна из существующих моделей не удовлетворяет одновременно требованиям к обобщению на реальные данные и к скорости обработки.

Научная новизна.

1. Предложена многомасштабная вариационная байесовская основа для устранения дымки на одном изображении, которая в изученной литературе первой моделирует все компоненты модели атмосферного рассеяния (ASM), чистое изображение, коэффициент пропускания и атмосферный свет, в рамках единой байесовской модели. Модель атмосферного рассеяния встраивается как физическая функция правдоподобия, и предлагается трёхкомпонентный априорный закон (ℓ_1 пространственный, FFT частотный, квадратичный многомасштабный), обеспечивая **обобщение на реальные данные без реальных парных обучающих примеров**.

2. Разработана **итеративная** архитектура декомпозиции Retinex (RSD-Net) с первым в изученной литературе контролируемым **модулем устранения свечения (DeGlow)**, явно отделяющим артефакты искусственных источников света до разложения по модели Retinex.

3. Разработана основа дистилляции знаний с **физически обоснованным вероятностным учителем** для устранения дымки, которая первой в изученной литературе объединяет дистилляцию знаний с вариационной байесовской основой устранения дымки, интегрируя расхождение Куллбека--Лейблера между апостериорными распределениями ученика и учителя в целевую функцию ELBO, а также используя **пространственно-неоднородную оценку** атмосферного освещения для работы в неоднородных условиях.

4. Предложена новая $1/\hat{\sigma}^2(x)$ **взвешенная функция потерь** дистилляции знаний для улучшения изображений при слабом освещении, аналитически выведенная из отношения освещённостей Retinex $\hat{t}(x) = \hat{L}(x)/\hat{L}_{\text{enh}}(x)$. Данный механизм не имеет аналогов в изученной литературе по дистилляции знаний для задач улучшения изображений при слабом освещении.

Практическая значимость. Методы диссертации внедрены в коммерческие системы: модуль устранения дымки интегрирован в платформу аэрофотограмметрии AutoWaypoints (OCUTECH LLC); модули улучшения изображений при слабом освещении (Главы 3--4) развёрнуты в системе безопасности платёжных терминалов компании Ayatech; достигнута договорённость о сотрудничестве с платформой геопространственных данных UP42 через стартап Parsl.ai.

- Компактный ученик устранения дымки работает со скоростью 76 FPS при разрешении, превышающем 4K UHD (более чем в $150\times$ быстрее учителя), улучшая обнаружение объектов на наборе KITTI на $+12.4\%$.

- Компактный ученик улучшения освещённости достигает 479 FPS при разрешении 512×512 (более чем в $90\times$ быстрее Retinexformer) и обнаруживает на 164.4% больше лиц на 300 изображениях DARK FACE по сравнению с необработанным входом, уступая учителю лишь на 2.3%.

- Оба ученика занимают первое или второе место по показателю NIQE на каждом протестированном наборе данных, включая непарные реальные наборы DICM и TM-DIED, без дополнительной настройки.