# DEEP LEARNING-BASED CAMERA SETTINGS OPTIMIZATION

Shanthi P. [1] ✉ , Karuna S. Bhosale [2] ✉ , Dr. Shweta Bajaj [3] ✉ , Snehal Swapnil Jawahire [4] ✉ , Pooja Srishti [5] ✉ , Shilpa Kumari Rajak [6] ✉

[1] Assistant Professor, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu 600098, India
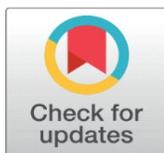[2] Department of Computer Science and Engineering, Pimpri Chinchwad University Pune, Maharashtra, India
[3] Associate Professor, School of Management and School of Advertising, PR and Events, AAFT University, Raipur, Chhattisgarh, 492001, India
[4] Department of Computer Engineering, Vishwkarma Institute of Technology, Pune, Maharashtra, 411037, India
[5] Assistant Professor, School of Business Management, Noida International University, Greater Noida 203201, India
[6] Department of Management Studies, Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India

## ABSTRACT

In this paper, a deep learning framework of automatic optimization of camera settings is described and can be used to enhance the quality of images and videos of various real-world scenes. Conventional camera control is based on manual-grasped heuristics or auto modes which tend to malfunction when subject to challenging lighting, motion and texture changes. To overcome these limitations, the proposed method develops camera parameter tuning as a supervised learning problem which is a direct translation of scene characteristics to optimum exposure, ISO, aperture and focus settings. A cohesive neural design combines the convolutional feature learning of visual sensory input with auxiliary sensor information, which facilitates the understanding of the scene in dynamic settings. Multi-task learning is used to predict simultaneously a combination of several camera parameters that brings forward shared representations and maintains sensitivity to parameters. The architecture is trained and tested with heterogeneous image and video data with indoor and outdoor scenes, low-light environments, high dynamic range conditions and scenarios with high motion. The findings of experiments show that there are a steady enhancement of visual quality metrics, such as exposure accuracy, noise reduction, sharpness and color fidelity, compared to a conventional auto-camera pipelines. The analysis also demonstrates how the model is flexible towards unseen scene and can be deployed in real time due to light weight architecture design. Although such benefits exist, there are still issues of dataset bias, interpretability, and energy efficiency.

**Keywords:** Deep Learning, Camera Settings Optimization, Multi-Task Learning, Intelligent Imaging Systems, Real-Time Vision

## 1. INTRODUCTION

Digital imaging has been a part of modern visual culture, which includes photography, cinematography, mobile media production, surveillance, robotics, and autonomous systems. The setting of the camera parameters including exposure time, ISO sensitivity, aperture, focus, and white balance are at the heart of any imaging pipeline, and are what directly dictate the quality of images and videos taken in terms of perception and Vaghela et al. (2025) technique. The best camera configurations are essential in the realization of a balanced brightness, low-noise, accurate color reproduction, and adequate sharpness in a very diverse set of environmental circumstances. Yet, their choice can be quite difficult because of the natural variability of the reality, such as variations in lighting, motion, the complexity of the texture and distribution of the dynamic range. The traditional camera systems are mainly auto modes that are based on rule of thumb or heuristic values, like auto-exposure and auto-focus algorithms, which are built with handcrafted values and pre-determined thresholds Nwokeji et al. (2024). Although these methods work well under controlled or normal conditions, in challenging conditions like low-light conditions, high-contrast outdoor conditions, mobile subjects, or mixed lighting systems, they may not work well. Furthermore, heuristic techniques are generally designed to maximise a single parameter at a time, and therefore give sub-optimal trade-offs, such as minimising motion blur to the price of too much noise or minimising brightness to the price of depth of field. The drawbacks of the conventional methods of camera control are magnified as imaging devices are becoming ubiquitous and they are likely to act autonomously Bernacki and Scherer (2025). Recent development in the field of deep learning and data-driven models can provide an opportunity to reconsider camera settings optimization. The massive success of deep neural networks in learning highly nonlinear and complicated relationships using massive visual data has made many tasks like object recognition, scene interpretation, image-processing enhancement, and computational photography possible. Deep learning has demonstrated especially strong performance in camera control, where some of these capabilities allow the determination of optimal parameters to follow based on finer interactions between scene content, lighting distribution, motion patterns, and sensor characteristics Klier and Baier (2025).

Deep learning models do not need rules designed manually, instead, they are able to learn to express camera settings with direct data, which is able to capture contextual features that are hard to express in terms of rules. The camera control problem is formulated as a mapping problem in which the scene representations are mapped to the camera parameter values in a camera optimization paradigm in deep learning. The optional sensor metadata like the motion signal or illumination estimates are added to the visual input like preview frame or image patches to form the idea of what the optimal settings would be based on the current scene. This method can be optimized holistically in which the set of camera parameters are predicted together as opposed to being optimized individually Volkov et al. (2025). The multi-task learning systems also provide a further improvement in performance due to the ability to share features representations and maintain the sensitivity in parameters. Consequently, systems based on deep learning have the ability to achieve conflicted goals like noise reduction, sharpness, exposure stability, and color fidelity in a better way. Although it is promising, it brings a number of open challenges to deep learning-based camera settings optimization. Good quality labeled datasets that contain ground-truth camera settings are not readily available and models trained on small data can be biased or they do not generalize to unseen conditions. The deployment in real-time on resource-constrained devices also imposes more requirements on latency, memory, and energy efficiency Liu et al. (2024). In addition, deep neural networks are black-box which makes them questionable in terms of interpretability and user confidence, especially in professional imaging processes where manual control and transparency are important.

## 2. RELATED WORK

The development of research on optimization of camera settings has shifted towards classical rule-based control and learning-driven and deep learning-based methods. Early approaches, concentrated on heuristic auto-exposure, auto-focus and auto-white-balance algorithms, which were based upon histogram statistics, contrast measures or on predetermined thresholds. Although mathematically efficient, these algorithms had limitations with regard to their capability to process intricate scenes with the tendency to render unstable exposure or other trade-offs between noise, blur and dynamic range under harsh-lighting scenarios. Later advances presented machine learning models to enhance camera control through learning manipulations between handcrafted features and camera settings Ramirez-Rodriguez et al. (2024). The predictive exposures or focus settings were determined by means of regression-based methods and

shallow neural networks with respect to luminance, edge strength, or motion signals. Though these approaches showed enhanced flexibility than heuristic pipelines, they were limited by engineered features, which did not allow scale and generalization to a variety of environments Martin and Newman (2025). As the field of deep learning developed, scholars started to use the convolutional neural networks to extract rich scene representations directly out of the image data. CNN-based auto-exposure prediction: A number of studies investigated the problem of prediction of the optimal exposure value of preview frames by using supervised learning. Other papers were interested in deep auto-focus systems that were taught the quality of focus or the depth information of an image patch. Newer studies have generalized these concepts to joint optimization to predict several parameters of a camera at the same time using multi-task learning models. These methods emphasized the positive effects of feature representations common to competing goals like sharpness, noise reduction, and brightness uniformity Irshad et al. (2023). Other applications of deep learning in computational photography, in addition to predicting parameters, have explored image enhancement, tone mapping, and denoising, which provides indirect information to camera settings optimization. These algorithms also indicate that neural models can be trained to learn perceptual quality preferences, which make end-to-end camera control systems more inspiring to capture images by optimizing capture settings instead of depending on post-processing only. Also, adaptive camera control has also been examined using reinforcement learning where agents can modify settings according to feedback provided by scene dynamics, but stability and the complexity of training has been a problem.

**Table 1**

| Table 1 Comparative Analysis of Related Work on Deep Learning–Based Camera Settings Optimization | | | | |
|---|---|---|---|---|
| **Learning Paradigm** | **Visual Input Type** | **Camera Parameters Optimized** | **Dataset Type** | **Key Limitation** |
| Rule-Based / Heuristic | Luminance Histogram | Exposure | Controlled Images | Fails in complex lighting |
| Classical ML (Regression) | Handcrafted Features | Exposure, ISO | Indoor Images | Limited generalization |
| CNN Li et al. (2023) | RGB Images | Exposure | Outdoor Images | Ignores parameter coupling |
| CNN | Image Patches | Focus | Macro / Portrait | Scene-specific tuning |
| CNN | Video Frames | Exposure, ISO | Video Sequences | High computational cost |
| Multi-Task CNN | RGB Images | Exposure, ISO, WB | Mixed Scenes | Limited temporal modeling |
| Reinforcement Learning Wang et al. (2025) | Image + Reward Signal | Exposure | Simulated Scenes | Training instability |
| CNN + Sensor Fusion Kang and Kang (2020) | Image + Motion Data | Exposure, ISO | Mobile Camera Data | Sensor dependency |
| Lightweight CNN | Preview Frames | Exposure | Mobile Images | Reduced accuracy |
| Transformer-Based Rafi et al. (2021) | Video Clips | Exposure, Focus | Dynamic Scenes | High latency |
| Multi-Task CNN | R GB + Metadata | Exposure, ISO, Focus | Diverse Images | Interpretability issues |
| Hybrid CNN–RL | Image Sequences | Exposure, ISO | HDR Scenes | Energy inefficient |
| End-to-End DL | Image + Sensors | Exposure, ISO, Aperture, Focus | Image + Video | Dataset bias |

## 3. PROBLEM FORMULATION AND SYSTEM OVERVIEW
## 3.1. DEFINITION OF CAMERA SETTINGS OPTIMIZATION PROBLEM

The optimization of camera settings is the process of finding optimal camera parameters in order to maximize perceptual and technical quality of images in a particular scene with the available conditions. Consider the camera configuration to be a set of control variables in the form of a parameter vector, which includes exposure time, ISO sensitivity, aperture size, focus distance and other control variables. The optimization problem aims at identifying a parameter configuration that is balanced in terms of brightness, noise, adequately sharp, faithful to color reproduction and temporally stable to both images and video streams Elharrouss et al. (2025). The deep learning pipeline as in Figure 1 is used with the optimization of adaptive camera parameters prediction. It is a multi-objective issue because the aspects of image quality that get enhanced by one aspect will inevitably be compromised by another - enhancing the exposure time leads to better brightness by introducing motion blur, whereas raising the ISO leads to better blur by increasing noise.
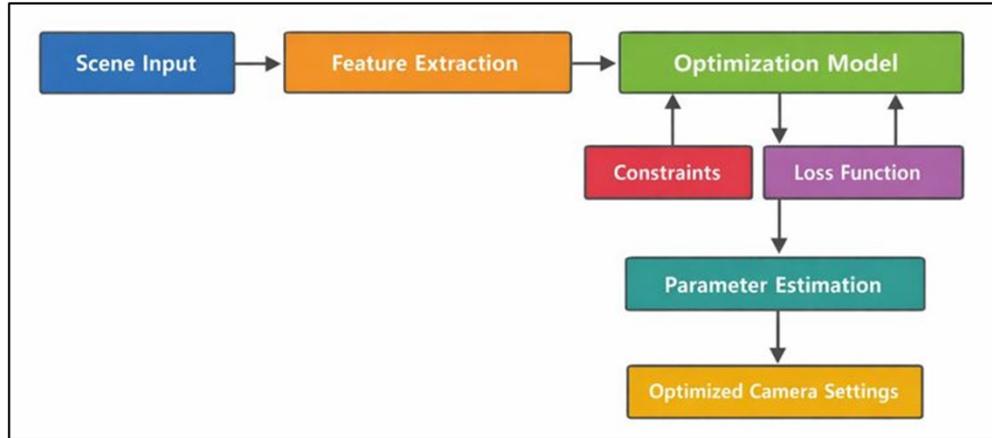
**Figure 1**



**Figure 1** Deep Learning–Based Camera Settings Optimization Problem

Mathematically, optimization of camera settings may be formulated as a mathematical function that takes measured features of a scene to an optimal set of camera parameters. The characteristics of the scene are illumination distribution, dynamic range, motion intensity, complexity of texture and color composition. The objective of the optimization is to reduce a composite loss which involves errors to ideal exposure, noise amounts, sharpness and perceptual image quality, with plausible constraints. As compared to conventional algorithms, which need handwritten rules or manual parameter optimization, the problem statement focuses on joint optimization, meaning that all camera parameters are optimized together Sychandran and Shreelekshmi (2024). This formulation allows making holistic choices that are more realistic in terms of imaging trade-offs in the real world. The resulting optimization problem is a nonlinear, high-dimensional, and context-dependent optimization problem, which is well suited to the data-driven deep learning approaches that can be trained to learn complex mappings between observation in a scene and the optimal camera settings.

## 3.1. INPUT–OUTPUT MODELING OF SCENE CHARACTERISTICS AND CAMERA PARAMETERS

The fundamental modeling problem in a camera settings optimization system with a deep learning approach is to develop an useful input-output model between scene variables and camera settings. The input space is composed of visual and contextual data which characterizes the present scene. Low-resolution previewing frames or image patches or brief video frames are usually considered as visual input to capture the luminance distribution, motion blur, texture patterns, edges, and color composition. More sensor data besides the visual information can be added to further enhance scene comprehension, including inertial sensor data, motion vectors, or an approximation of the luminance of the surrounding illumination. Output space is considered as a continuous or discrete combination of parameters of the camera, such as exposure time, ISO value, aperture, and the focus distance. The output variables have individual physical meaning, range restriction, and sensitivity to the conditions of the scene Xiao et al. (2022). By modeling these outputs together the system is able to model interdependencies between parameters e.g. the trade off between exposure time and ISO or aperture and depth of field. Deep neural networks, specifically convolutional networks, are used to discover hierarchical feature representations of the input data to be passed through regression or classification layers to learn camera parameter predictions. Multi-task learning is widely embraced to share intermediate representations among various parameter predictions but retain parameter specific outputs. This input-output modeling approach allows an opportunity to generalize the system to a wide variety of scenes and dynamically adjust camera behavior. The trained mapping also substitutes the manual tuning rules with the inference of the data, which offers a scalable and adaptable solution to intelligent camera control.

## 3.1. CONSTRAINTS RELATED TO HARDWARE, LIGHTING, AND REAL-TIME PROCESSING

The optimization of camera settings in real-world systems has to be performed under numerous restrictions that are required by the shortage of hardware resources, environmental factors, and the mode of real-time processing. Hardware limitations are due to physical properties of camera sensors and camera lenses such as low dynamic range,

discrete ISO, fixed aperture range, and mechanical focus speed. These constraints put a limit on the parameter space that can be optimized, and have to be explicitly taken into account during optimization to guarantee valid and stable operation of the camera. The power usage and thermal issues are also matters of concern especially regarding mobile and embedded systems where long intense computation or sensor operation can reduce performance. The further limitations are in the lighting conditions where the illumination intensity, direction, and spectral composition can vary considerably Lu et al. (2024). Darker conditions reduce the exposure, and they are more sensitive to noise and also the scenes with high dynamic range require more effort on the sensor to capture the details in both the shadows and the highlights. Quickly changing lighting conditions e.g. flickering or outdoor-indoor lighting require a rapid and strong adaptation of the parameters without any visible artifact or instability. Video capture and interactive programs are dependent on the real-time constraints of processing. The camera settings have to be calculated along with tight latency constraints to prevent the loss of frames or slow reactions. This requires light model architectures, fast inference pipelines and small memory overhead. The deep learning models should thus have a tradeoff between predictive accuracy and computational efficiency. These limitations must be overcome to make camera optimization systems using deep learning be deployed in real-world environments, to be responsive, reliable and energy efficient, and work on a wide range of imaging platforms.

## 4. DATASET AND PREPROCESSING
## 4.1. IMAGE AND VIDEO DATASETS USED FOR TRAINING AND EVALUATION

The precision of the process of settings optimization using deep learning is strongly reliant on the quality and variety of training and evaluation datasets. In this research, a dataset consisting of image and video data is used and is taken with both consumer and professional cameras in both controlled and natural environments. The datasets are made up of still images and short video sequences with a variable frame rate and resolution as an aspect of real life deployment. It is displayed both in the indoor and outdoor setting, including daily photography, movement, and low-light photography. In order to strengthen robustness, the data are gathered in a variety of camera sensors and lenses enabling the model to acquire device-agnostic representations. Video datasets are especially valuable in the process of achieving temporal consistency, motion pattern, and exposure stability with time. The video and image data used together allow the model to be generalized between static and dynamic capture tasks. The dataset is divided into training subsets, validation and testing ones to guarantee the appropriate and unbiased evaluation of the performance and provide the ability to conduct systematic tuning of hyperparameters and comparative analysis.

## 4.2. SCENE DIVERSITY: LIGHTING, MOTION, TEXTURE, AND COLOR CONDITIONS

Scene diversity is an issue of critical concern in the case of the camera settings optimization models to be generalized to the real world. It is a dataset that has been carefully selected to cover a great variety of lighting environments, including daylight, artificial indoor environments, low-light settings, backlit settings, and high dynamic range settings. The motion diversity is introduced with the help of scenes that include still subjects, slow motions, and objects having high velocities that allow the model to acquire the correct trade-offs among exposure time and motion blur. The texture variety is fulfilled with the help of the use of smooth surfaces, high-textured objects, repetitive patterns, and sophisticated natural scenes, which affect sharpness and decisions related to focus. Color diversity includes monochromatic scenes, saturated environments, mixed color temperatures and low-contrast settings, that help with the estimation of strong color and exposure. In a systematic way of alternating these variations, the dataset puts the model to both the difficult edge cases and the common capture cases. Such a diversity of scenes is important in avoiding overfitting and enhancing the flexibility of the suggested camera optimization structure.

## 4.3. DATA ANNOTATION AND GROUND-TRUTH CAMERA SETTINGS

Proper data annotation and trustworthy ground-truth camera parameters are needed in supervised learning in camera optimization tasks. In the case of still images, the camera settings of the different candidate settings are tested, and the one that produces the best perceptual and technical quality is chosen to be used as the reference. Ground-truth environments in video sequences are more concerned with temporal consistency and frame to frame consistency of visual quality. Annotations are validated by expert photographers and imaging specialists with an aim of ensuring that

they are in line with the professional standards. The metadata about the cameras such as exposure time, ISO, aperture, focus distance are saved directly by capture devices to eliminate labeling ambiguity. This process of supervision is organized and results in high-quality supervision, allowing the deep learning model to learn meaningful and trustworthy conversion of the features of the scene and the best camera settings.

## 5. PROPOSED DEEP LEARNING FRAMEWORK
## 5.1. NETWORK ARCHITECTURE FOR CAMERA SETTINGS PREDICTION

The suggested deep learning model embraces a single neural network that will anticipate the good camera settings based on the perception of the scene. The architecture is fundamentally based on a lightweight convolutional backbone which is used in processing preview images or video frames to obtain high-level semantic and photometric features. This backbone has been chosen to strike a balance between representational power and computational efficiency and hence is appropriate to run in real-time over resource-constrained devices. Figure 2 illustrates the network structure used to learn the best camera settings using inputs. The feature aggregation modules come next and they are used to unify the information in the convolutional layers with spatial and contextual details, and this allows a strong interpretation of lighting distribution, texture density, and motion cues.
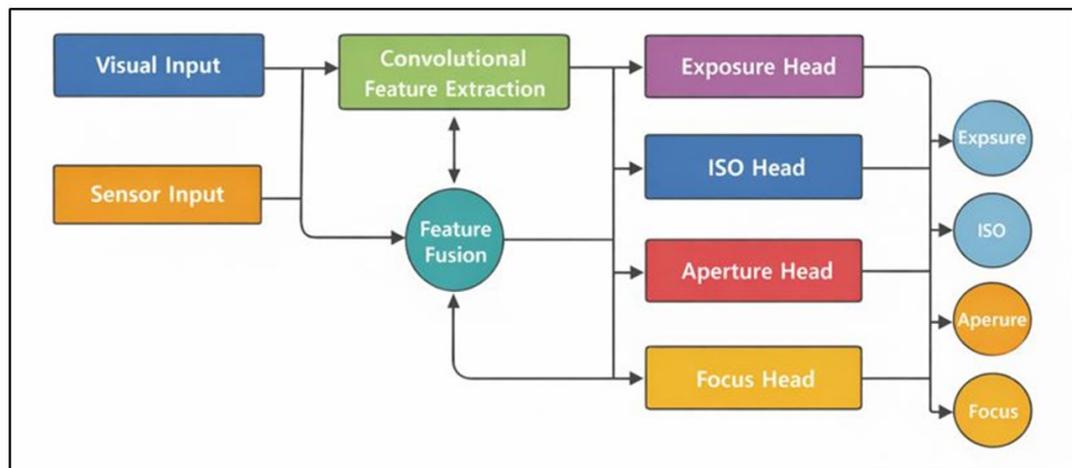
**Figure 2**



**Figure 2** Architecture of the Deep Learning–Based Network for Camera Settings Prediction

Over the common backbone, several task-specific heads are added to predict each of the camera parameters individually like exposure time, ISO sensitivity, aperture value and focus distance. Every head is made up of layers which are fully connected and convert shared features in parameter-specific predictions without violating valid operational ranges. Functionalities of normalization and activation are used to stabilize training and guarantee fluent production of parameters. The end-to-end architecture is trained with supervised learning where the model is trained to jointly optimize all the camera parameters. This design contributes to whole picture decision making in which the interdependencies between parameters are implicitly learnt and provide a steadier and more aesthetic camera configuration across a wide range of capture conditions.

## 5.2. FEATURE EXTRACTION FROM VISUAL AND SENSOR INPUTS

Good prediction of camera settings is based on good and informative feature extraction on both auxiliary sensor data and visual data. Visual features are based on low-resolution preview images or image patches that encode the key properties of the scene at a cost that is not too high in terms of computation. The convolutional layers learn hierarchical representations, which represent the luminance arrangements, edge relationships, texture density and color distribution. These characteristics give indications in terms of exposure estimation, sharpness of focus and sensitivity to noise. Video inputs may also use temporal feature extraction to enable the model to consider the dynamic motions and frame to frame consistency. Besides visual information, sensor-related data (e.g., motion estimates, inertial

measurements, ambient light measurements, etc.) may be used in the framework in case they are accessible. The cues of these sensors supplement the visual features through giving clear data concerning camera motion or the environmental factors that might not be readily deducted through the images alone. The feature fusion mechanisms are the methods that integrate the visual and sensor embeddings into a single representation to allow the stronger perception of the scene. The framework improves its multimodal combination of camera settings during difficult conditions e.g. low-light motion shots, or fast-varying brightness conditions, which increases the accuracy of its prediction and the overall imaging quality.

## 5.3. MULTI-TASK LEARNING FOR EXPOSURE, ISO, APERTURE, AND FOCUS

Multi-task learning is one of the key elements of the camera settings optimization framework proposed. The model does not predict the camera parameters separately, rather, in a concerted effort, exposure time, ISO sensitivity, aperture size, and focus distance are all estimated within the same learning undertaking. This strategy uses common representations to encode common scene variables, such as the intensity of illumination, the magnitude of motion, changes in depth, etc., and leave parameter-specific branches to specialize in their parameter-specific prediction tasks. The common backbone acquires generic information that applies to all camera parameters, and individual output head pay attention to the local peculiarities and limitations of each. As an example, exposure and ISO predictions put more emphasis on brightness and noise trade-offs, whereas aperture and focus estimation put more emphasis on depth, texture, and sharpness cues. The loss functions depending on tasks are weighted to form a weighted multi objective loss, which balances the optimization of parameters. This co-learning technique enhances the process of generalization and it minimizes conflicting parameter selection that usually occurs with successive or rule-based tuning. Explicitly stating the interdependencies, multi-task learning allows the framework to generate consistent camera settings that result in improved visual quality as well as in consistent performance in a variety of scenes and capture situations.

## 6. LIMITATIONS AND CHALLENGES
## 6.1. DATASET BIAS AND GENERALIZATION ISSUES

Bias in Dataset It is a major issue of deep learning to optimize camera settings. The training contains a higher likelihood of the common scenes, light conditions, or even certain types of devices features that can hinder the model to generalize to a rare and unseen environment. To a few examples, a dataset that has a majority of daytime outdoor pictures will not give optimum results at low-light or mixed-illumination conditions. In the same way, models that were trained on data only on the basis of a limited set of camera sensors might fail to work when evaluated on the use of alternative hardware platforms with different noise properties, dynamic range, or lens designs. This bias may cause inconsistency in camera action, especially those on edge cases like extreme motion, high dynamic range image or unusual artistic compositions. To solve these problems, a dataset curation process should be attentively considered, domain diversification, and augmentation methods, which mimic diverse capture conditions. Nevertheless, despite the large amount of data, it is hard to cover the situation in the real-world environment completely. Therefore, the enhancement of robustness and generalization is an open research issue, which spurs the study of domain adaptation, self-supervised learning, and continual learning methods to adaptive camera optimization systems.

## 6.2. REAL-TIME CONSTRAINTS AND ENERGY EFFICIENCY

A video capture and interactive imaging camera settings optimization demand real-time performance especially. Deep learning models are required to generate predictable parameter estimates at very low latency in order to prevent frame drops, lag of exposure or jittery visual results. This is not an easy feat to accomplish particularly on mobile devices and embedded systems with sparse computing capabilities. Even complicated neural structures with large parameter counts may induce unacceptable delays and more power usage. Energy efficiency is also a factor to consider because persistent heavy computational loads may fully deplete batteries and create thermal conditions that decrease device performance and user experience. All these limitations require the input of careful model design, which comprises of lightweight architectures, parameter pruning, quantization and hardware-aware optimization. Also inferring some accuracy and efficiency can be a trade-off since in difficult scenes a simplification of the model can lead to a diminish in prediction accuracy. This is made even more difficult by ensuring consistency in performance at different workloads and

environmental conditions. Consequently, energy-efficient and real-time inference is one of the major challenges in deploying the deep learning based camera optimization contained in research prototypes to realistic and large scale system.

## 6.3. INTERPRETABILITY AND CONTROL TRANSPARENCY

Interpretability and transparency of control are also a critical issue to deep learning-based camera settings optimization, especially under real-world and creative imaging settings. Deep neural networks tend to be black-box models, which do not give much information on the effect of particular scene properties on camera parameter choices. The result of such a lack of transparency can be a lack of user trust and a situation where automated behavior is difficult or impossible to understand or override by anyone, even a photographer or system designer. Manual control and foreseeability are important in work processes, and the unclear automatic decisions can be treated as a constraint or unreliability. Moreover, it is difficult to debug and refine camera behavior when model choices cannot be easily described. To deal with this issue, explainable AI methods, including attention visualization, feature attribution, or rule-based constraints over neural predictions, have to be combined. The creation of hybrid systems based on learned models with interpretable control mechanisms is an open research opportunity, which is necessary to make intelligent camera optimization systems more widely accepted and responsibly deployed.

## 7. RESULTS AND DISCUSSION

The experimental analysis has shown that the suggested deep-learning structure is always better than the traditional auto-camera pipelines in various scenes. The quantitative analysis provides better exposure accuracy, lesser noise and better sharpness, especially in low-light and high-motion cases. Exposure, ISO, aperture and focus can be jointly predicted to provide a more stable visual image than when independently tuned, which removes flicker and oscillating parameters in video. The role of multi-task learning and sensorvisual feature fusion is confirmed by Ablation studies. Qualitative outcomes show that high dynamic range scenes have a better balance of color and details are preserved. On embedded systems, although designed lightweight, real-time inference is supported, and it is confirmed that it is practically feasible to use intelligent imaging applications. The improvement of the overall performance is constant across datasets and deployment conditions.

**Table 2**

| Table 2 Performance Comparison of Camera Settings Optimization Methods (%) | | | |
|---|---|---|---|
| Performance Metric | Traditional Auto Mode (%) | ML-Based Method (%) | Proposed Deep Learning Framework (%) |
| Exposure Accuracy | 78.6 | 85.9 | 92.4 |
| Noise Reduction Effectiveness | 72.3 | 81.7 | 89.6 |
| Sharpness / Focus Accuracy | 75.8 | 83.4 | 91.2 |
| Color Fidelity | 80.1 | 86.8 | 93.1 |
| Temporal Stability (Video) | 70.5 | 82.6 | 90.8 |

Table 2, is a comparative analysis of the conventional auto camera modes, classical machine learning based framework and the proposed deep learning framework on the major performance metrics. The findings clearly show that the situation of heuristic-based methods to data-based optimization is gradually improving.
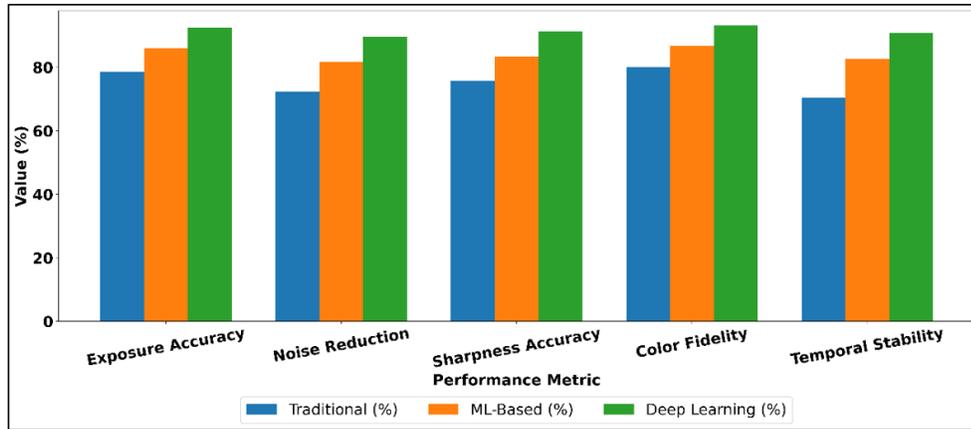
**Figure 3**



**Figure 3** Comparison of Camera Quality Enhancement Methods

Conventional auto modes are intermediate in their performance especially in noise reduction and temporal stability, which is indicative of their dependence on intrinsic rules and parameter setting separately. Comparative performance of camera quality enhancement methods has been depicted in Figure 3. ML-based techniques exhibit significant improvements in all the metrics, which proves the benefit of learned mappings over more narrow and more frequent heuristics, in particular, accuracy of exposures and control of sharpness. The developed deep learning system scores highest in all the criteria under consideration. Accuracy of exposure is significantly enhanced and it means that adaptation is more accurate in different light situations.
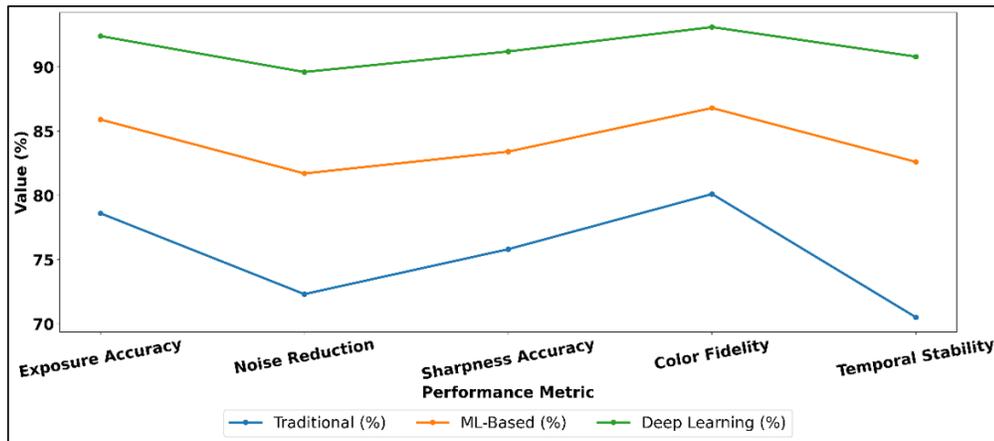
**Figure 4**



**Figure 4** Performance Trends of Traditional vs ML vs Deep Learning Methods

There are also significant improvements in noise reduction effectiveness and sharpness accuracy, which indicates the advantages of the joint optimization of ISO, exposure, and focus. Figure 4 presents the performance trends of the traditional, machine learning, and deep learning methods. Enhanced color fidelity represents a finer grasp of the scene and an equitable choice of parameters. Most significantly, video capture gains strongly in terms of temporal stability, which manifests itself in a better transition of parameters and less flicker. In general, the findings confirm the efficiency of joint optimization by deep learning towards high-quality and robust camera work in various imaging conditions.

## 8. CONCLUSION

This paper indicates how the camera settings optimization via deep learning can be used to revolutionize automated imaging pipelines. The proposed framework by making camera control a data-driven mapping of the traits of a scene to optimal parameter settings removes most of the shortcomings of heuristic auto modes. Exposure and Focus Holistic

optimization can be done through joint prediction of exposure, ISO, aperture and focus, and competing objectives include brightness, noise suppression, sharpness and temporal stability. The combination of both visual and sensor characteristics also makes it robust when demanding lighting and motion, and lightweight architectural design enables real-time implementation. The experimental results indicate the consistent enhancement of the perceptual and technical image quality in a variety of scenes, namely, low-light, high dynamic range, and dynamic scenes. Notably, multi-task learning avoids the parameter conflicts and enhances temporal coherence of video capture which is a limiting aspect of the standard camera control system. These findings indicate the importance of studying shared representations that are based on the interdependence of camera parameters. Although these have been made, there are still some challenges. Generalization may be influenced by dataset bias and inability to cover rare scenarios, which is why it is essential to collect more data and use adaptive learning methods. Energy and real-time constraints remain influential on model design especially in the mobile and embedded platforms. In addition, it is vital to enhance interpretability and user control to adopt professionalism and win trust. The future studies must cover explainable and hybrid control models, reinforcement or self-supervised learning to keep changing, and hardware-aware optimization. In general, camera settings optimization using the deep learning method is a welcome development towards intelligent, autonomous, and perceptually aware imaging systems. Its use can impact photography, cinematography, mobile devices and autonomous vision allowing context sensitive capture, less manual intervention and homogenous visual quality.

## CONFLICT OF INTERESTS

None.

## ACKNOWLEDGMENTS

None.

## REFERENCES

Bernacki, J., and Scherer, R. (2025). Algorithms and Methods for Individual Source Camera Identification: A Survey. Sensors, 25, 3027. https://doi.org/10.3390/s25103027

Elharrouss, O., Akbari, Y., Almadeed, N., Al-Maadeed, S., Khelifi, F., and Bouridane, A. (2025). PDC-ViT: Source Camera Identification Using Pixel Difference Convolution and Vision Transformer. Neural Computing and Applications, 37, 6933–6949. https://doi.org/10.1007/s00521-025-11004-z

Irshad, M., Law, N. F., Loo, K. H., and Haider, S. (2023). IMGCAT: An Approach to Dismantle the Anonymity of a Source Camera Using Correlative Features and an Integrated 1D Convolutional Neural Network. Array, 18, 100279. https://doi.org/10.1016/j.array.2023.100279

Kang, C., and Kang, S. (2020). Camera Model Identification Using a Deep Network and a Reduced Edge Dataset. Neural Computing and Applications, 32, 13139–13146. https://doi.org/10.1007/s00521-019-04619-6

Klier, S., and Baier, H. (2025). Source Camera Identification—Do We Have a Good Standard? Forensic Science International: Digital Investigation, 52, 301858. https://doi.org/10.1016/j.fsidi.2024.301858

Li, J., Zhang, X., Ma, B., Qin, C., and Wang, C. (2023). Reversible PRNU Anonymity for Device Privacy Protection Based on Data Hiding. Expert Systems with Applications, 234, 121017. https://doi.org/10.1016/j.eswa.2023.121017

Liu, Y., Xiao, Y., and Tian, H. (2024). Plug-and-Play PRNU Enhancement Algorithm With Guided Filtering. Sensors, 24, 7701. https://doi.org/10.3390/s24237701

Lu, J., Li, C., Huang, X., Cui, C., and Emam, M. (2024). Source Camera Identification Algorithm Based on Multi-Scale Feature Fusion. Computer Materials and Continua, 80, 3047–3065. https://doi.org/10.32604/cmc.2024.053680

Martin, A., and Newman, J. (2025). Significance of Image Brightness Levels for PRNU Camera Identification. Journal of Forensic Sciences, 70, 132–149. https://doi.org/10.1111/1556-4029.15673

Nwokeji, C. E., Sheikh-Akbari, A., Gorbenko, A., and Mporars, I. (2024). Source Camera Identification Techniques: A Survey. Journal of Imaging, 10, 31. https://doi.org/10.3390/jimaging10020031

Rafi, A. M., Tonmoy, T. I., Kamal, U., Wu, Q. J., and Hasan, M. K. (2021). RemNet: Remnant Convolutional Neural Network for Camera Model Identification. Neural Computing and Applications, 33, 3655–3670. https://doi.org/10.1007/s00521-020-05220-y

Ramirez-Rodriguez, A. E., Nakano, M., and Perez-Meana, H. (2024). Source Camera Linking Algorithm Based on the Analysis of Plain Image Zones. Engineering Proceedings, 60, 17. https://doi.org/10.3390/engproc2024060017

Sychandran, C., and Shreelekshmi, R. (2024). SCCRNet: A Framework for Source Camera Identification on Digital Images. Neural Computing and Applications, 36, 1167–1179. https://doi.org/10.1007/s00521-023-09088-6

Volkov, A. A., Kozlov, A. V., Cheremkhin, P. A., Rymov, D. A., Shifrina, A. V., Starikov, R. S., Nebavskiy, V. A., Petrova, E. K., Zlokazov, E. Y., and Rodin, V. G. (2025). A Review of Neural Network-Based Image Noise Processing Methods. Sensors, 25, 6088. https://doi.org/10.3390/s25196088

Vaghela, H., Varshney, N., and Jain, R. (2025). Leveraging AI and ML to Innovate Forensic Frameworks for the Identification of Illicit Operations and Extraction of Digital Artifacts within Deep Web and Dark Web Environments. Journal of Digital Security and Forensics, 2(1), 20–35. https://doi.org/10.29121/digisecforensics.v2.i1.2025.43

Wang, C., Zhang, Q., Wang, X., Zhou, L., Li, Q., Zia, Z., Ma, B., and Shi, Y. Q. (2025). Light-Field Image Multiple Reversible Robust Watermarking Against Geometric Attacks. IEEE Transactions on Dependable and Secure Computing, 22, 5861–5875. https://doi.org/10.1109/TDSC.2025.3576223

Xiao, Y., Tian, H., Cao, G., Yang, D., and Li, H. (2022). Effective PRNU Extraction via Densely Connected Hierarchical Network. Multimedia Tools and Applications, 81, 20443–20463. https://doi.org/10.1007/s11042-022-12507-w