Received: 4 January 2024 | Revised: 27 May 2024 | Accepted: 2 July 2024

IET Image Processing

DOI: 10.1049/ipr2.13182

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

ExposureNet: Mobile camera exposure parameters autonomous control for blur effect prevention

Abdelwahed Nahli[1] | **Dan Li**[1] | **Rahim Uddin**[1] | **Muhammad Irfan**[1,2] | **Mohamed Oubibi**[3] | **Qiyong Lu**[1] | **Jian Qiu Zhang**[1]

[1]Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai, China

[2]Department of Computing, University of Turku, Turku, Finland

[3]Smart Learning Institute, Beijing Normal University, Beijing, China

Correspondence
Abdelwahed Nahli and Dan Li, Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai 200433, China.
Email: nabdelwahed21@m.fudan.edu.cn and lidan@fudan.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 12374431

Abstract

The quality of images we perceive visually is heavily impacted by the settings used for camera exposure. When these settings are imbalanced, it can result in an undesired prominent phenomenon known as blur effects. To address this problem, an ExposureNet project has been undertaken, which aims to develop an autonomous camera exposure settings control system for blur effects prevention. The proposed ExposureNet model is a CNN/Transformer hybrid neural structure, created and trained in a comprehensive manner to effectively predict the ideal exposure settings based on the semantic features of the scene being captured. This system is designed to learn the necessary steps for processing, such as identifying relevant scene features, using only two camera exposure parameters (shutter speed (SHS) and ISO) as training signals. As a result, this system can associate the semantic features of a scene with the appropriate exposure parameter adjustments, customized to the scene's dynamics and lighting conditions. By simultaneously optimizing all processing steps and bypassing traditional post-processing stages, the proposed system is designed to achieve faster performance, reduced computational cost, and lower power consumption. Experimental results demonstrate that the proposed system significantly outperforms existing methods and achieves cutting-edge performance.

1 | INTRODUCTION

Auto-exposure is an essential function in cameras for capturing high-quality photos. However, the conventional exposure parameters control algorithms typically rely on simple metering that is independent of scene content and dynamics, where the exposure parameters are only adjusted based on the amount of light reaching the camera sensor, without taking into consideration the scene content and motion.

Knowing the fact that the occurrence of the blur effect phenomenon is simultaneously related to the scene's amount of motion and its lighting conditions. Thus, the lake of awareness of the scene-level features and dynamics makes of the conventional auto-exposure algorithms ineffective and unreliable approaches for blur effects prevention.

Besides auto-exposure mode, exposure parameters manual control mode is also available in most of the nowadays smartphone cameras, allowing users to manually adjust exposure parameters and find the optimal settings for the current scene before pressing the shutter. However, this slow manual parameter tweaking process can only work for static scenes. In dynamic scenes, where fast motion is very dominant, this slow process can lead to missing the right moment to capture the desired image.

In light of the aforementioned limitations of the conventional camera auto-exposure algorithms and the manual-exposure mode, an urgent need has arisen for a blur-aware autonomous camera exposure parameters control system.

This study aims to introduce a data-driven solution for preventing the occurrence of the blur effect phenomenon and tackle the limitations of conventional handcrafted techniques. Therefore, we developed ExposureNet, a well-trained deep learning model, as an autonomous camera exposure parameters control system for blur effect prevention. The newly proposed approach robustly performs the blur-aware auto-exposure task and bypasses the need for any manual interventions or

post-processing steps. To our knowledge, there is no existing research that utilizes generic scene information for blur effect prevention.

The major three contributions of this study can be listed as below:

1. We have successfully developed a data-driven autonomous camera exposure parameters control system for blur effects prevention.
2. We introduce Blur-Prevent dataset as a brand-new dataset to address the most dissatisfactory aspects of the pre-existent datasets. It consists of sharp and blurry images, alongside with their corresponding exposure metadata (Exposure time, ISO).
3. As far as our knowledge extends, this study is the first to introduce a blur effect prevention system based on an autonomous control of the exposure parameters. We argue that this system can serve as inspiration for tackling similar computer vision problems in new and innovative ways.

2 | RELATED WORKS

The auto-exposure function in cameras has been a subject of extensive research and development, aiming to enhance the quality of photographs by ensuring optimal exposure settings under varying lighting conditions. This section reviews a selection of seminal works and recent advancements in the field of camera auto-exposure.

Authors in [1] provide a concise explanation of how auto-exposure operates in Nikon cameras when set to Auto(P) mode. Their work emphasizes the importance of setting a lower limit for the shutter speed to prevent blur caused by object motion and shaking. They suggest that this threshold should not exceed $\frac{1}{40}$ s. The Auto(P) mode maintains a fully open aperture until the shutter speed exceeds $\frac{1}{40}$ s. If the shutter speed is very low, the aperture is slightly closed. When the fully open aperture is insufficient to achieve the desired shutter speed, the ISO parameter is increased. To achieve auto exposure, one method is to divide the frame into several blocks and calculate the luminance (BV) in each block [2]. This technique is commonly employed by various camera brands such as Nikon [1] and Canon [3]. The suggested approach involves determining the maximum (BV_{max}) and minimum (BV_{min}) luminance in each block and calculating the average luminance in the lower (BV_{lower}), upper (BV_{upper}), and all blocks (BV_{mean}). Hence, the scene luminance can be expressed using the following equation:

$$BV = aBV_{max} + bBV_{min} + cBV_{mean} + dBV_{upper} + eBV_{lower} \quad (1)$$

Determining the shutter speed and aperture requires setting coefficients a, b, c, d and e, which should be defined by the camera's firmware. In references [4] and [5], an auto-exposure algorithm utilizing the False-position approach is presented and implemented. This algorithm is specifically designed for industrial applications, aiming to accurately expose leather samples to detect defects. Prioritizing shutter speed, the algorithm controls the camera to capture properly exposed leather samples based on the lighting conditions.

In [6], researchers propose an auto-exposure algorithm that detects lighting conditions. Their method identifies high-contrast lighting situations and enhances the dynamic range of images. The lighting conditions are estimated by calculating the difference between the mean value and median value of the brightness levels in the captured images. High contrast is determined by comparing the average brightness level of the entire image with the median brightness level of all pixels. If the difference between the mean and median exceeds a certain threshold, it indicates high contrast lighting. A similar approach for detecting high-contrast lighting conditions using differences between lighting conditions is proposed in reference [7]. The authors suggest dividing the frame into smaller areas and calculating the brightness level based on the average brightness of the area and its surrounding areas, by considering the corresponding weights. Experimental results show that their approach successfully detects high-contrast scenes with an acceptable level of exposure error.

However, despite providing real-time auto-exposure functionality, these approaches may result in motion blur because they are not primarily designed for blur prevention. Most of them rely on simple metering derived from mathematical models with underlying assumptions, analysing the lighting conditions regardless of the content and dynamics of the scene. This limitation hinders these approaches from fully addressing the blur issue.

To tackle the blur effect phenomenon, conventional approaches in [8, 9], mainly rely on recursive strategies as post-processing solutions that operate in the RGB domain, whereas, as opposed to that, in this study, we promote a novel blur preventive strategy by using the camera's exposure parameters as training signals alongside with their corresponding image data.

Kuno et al. introduced early concepts of automatic exposure control in digital cameras, discussing techniques that formed the foundation for subsequent research in the area [10]. Thompson and Williams conducted a comprehensive survey of auto-exposure algorithms in consumer cameras, highlighting the evolution of techniques and their impact on image quality [11]. Onzon and Mannan presented an adaptive auto-exposure technique for high dynamic range (HDR) imaging, addressing challenges related to capturing scenes with extreme luminance variations [12]. Jiang and Kuhnert proposed a real-time exposure adjustment method based on histogram analysis, contributing to improved image quality by considering pixel distribution [13]. Additionally, Su and Lin introduced a dynamic auto-exposure method that utilized scene classification for enhanced performance in different scenarios [14].

Hybrid approaches also gained attention, as evidenced by the work of Torres and Menéndez, who developed a multi-zone analysis-based auto-exposure algorithm for smart cameras [15]. Odinaev and Chin conducted a comparative study of various auto-exposure techniques in low-light conditions, shedding light on their respective strengths and weaknesses [16]. Görmer explored auto-exposure strategies specific to HDR

photography, outlining approaches that optimize image exposure across varying lighting conditions [17]. The integration of learning techniques in auto-exposure was addressed by Kuang and Jin, who presented a learning-based control system for unmanned aerial vehicles [18]. Torres and Menéndez focused on real-time surveillance cameras, proposing an efficient auto-exposure algorithm tailored for video surveillance applications [19].

In the realm of smartphone photography, Yang and Wang leveraged deep learning to achieve improved auto-exposure control in smartphones, showcasing the benefits of machine learning in consumer devices [20]. Reinforcement learning was explored by Liu and Wang, who enhanced auto-exposure techniques using reinforcement learning to adapt to dynamic lighting conditions [21]. The fusion of multiple sensors was investigated by Shen and Mulgaonkar in the context of surveillance networks, demonstrating the advantages of sensor cooperation in auto-exposure [22]. Nia and Hu introduced a fuzzy logic-based algorithm for industrial cameras, addressing exposure control in specific application domains [23].

Lu and Zhang employed genetic algorithms to optimize auto-exposure settings for astrophotography, catering to the unique challenges of capturing astronomical scenes [24]. Energy efficiency in camera networks was considered by Wang and Kim, who proposed an energy-efficient auto-exposure approach for wireless camera networks [25]. Han and Jia tackled auto-exposure challenges in HDR video recording, introducing adaptive strategies for preserving image quality [26]. Neural network ensembles were explored by Wu and Tsotsos [27], who demonstrated their potential in enhancing auto-exposure control. Microscopic imaging in life sciences was the focus of Skandarajah and Reber's work, where auto-exposure techniques were tailored for specific imaging requirements [28]. Lastly, Li and Zou provided insights into auto-exposure algorithms in UAV photogrammetry, highlighting their significance in remote sensing applications [29].

The field of camera auto-exposure has seen remarkable research efforts, with numerous studies contributing to the advancement of auto-exposure techniques in various multimedia contexts.

Early work by Gu and Al Noman introduced adaptive auto-exposure control through histogram analysis, emphasizing the importance of utilizing image histograms for exposure adjustment [30]. Piao and Xu expanded auto-exposure capabilities to high dynamic range (HDR) imaging, proposing real-time adjustment methods that catered to scenes with extreme luminance variations [31].

In the domain of video surveillance, Zhang and Lee addressed challenges in surveillance systems, developing content-aware auto-exposure methods that accounted for dynamic surveillance environments [32]. Yang and Wang harnessed deep learning for auto-exposure control in smartphone cameras, substantially enhancing image quality [33].

Responding to demands for high-quality multimedia streaming, Li and Handa proposed dynamic auto-exposure techniques for HDR video streaming, ensuring optimal exposure settings during real-time streaming scenarios [34]. Dasari explored sensor fusion's potential in adaptive auto-exposure for visual sensor networks, leveraging multiple sensors to enhance exposure control [35].

Recent advancements in image deblurring have seen notable innovations. An improved DeblurGAN for blind motion deblurring has significantly enhanced image restoration [36]. Additionally, joint learning techniques for both motion and defocus deblurring, using real-world datasets, have been developed to tackle various types of blur [37]. Furthermore, new methods for reduced-reference image deblurring quality assessment, utilizing multi-scale feature enhancement, have been proposed for more precise quality evaluation [38]. These developments represent the diverse efforts to effectively address image blur.

These works collectively underscore the evolution of auto-exposure techniques, ranging from foundational concepts to sophisticated machine learning-based approaches, across diverse multimedia domains. These contributions have significantly enriched the field, reshaping how cameras adapt to changing lighting conditions, and ultimately enhancing the quality of multimedia content.

Attention mechanisms [39] are recently utilized very often in computer vision to capture long-range dependencies within features. In the context of image deblurring, these mechanisms are effective in learning cross-pixel relationships, which are crucial for handling non-uniform blur. Therefore, a CNN/Transformer hybrid model equipped with multi-head self-attention, which can assess both local and global correlations, is well-suited for this deblurring application.

In this study, beyond conventional auto-exposure methods, we describe a CNN/Transformer hybrid neural network model that can learn the entire processing pipeline required to autonomously control camera parameters for blur effects prevention, regardless of the type of motion or lighting scenarios. To the best of our knowledge, there have been no prior deep learning-based investigations on autonomously controlling camera exposure parameters specifically for blur prevention in dynamic scenes with very heavy motion.

3 | BLUR-PREVENT DATASET

In contrast to amateur photographers, expert photographers possess comprehensive knowledge of various photography techniques. Their extensive experience enables them to capture artistic and sharp photographs in any type of scene, lighting conditions, or motion scenarios. Experts often utilize manual mode, which grants them full control over the elements of the exposure triangle (ISO, shutter speed, and aperture) shown in Figure 1. By carefully balancing and adjusting these three key values, they can skilfully manipulate their cameras and prevent issues like under-exposure, over-exposure, noise, and motion blur.

Recognizing that the majority of the world population consists of amateur photographers, we initiated this project to develop a deep neural network capable of extracting underlying patterns from the expert photographers shooting data,

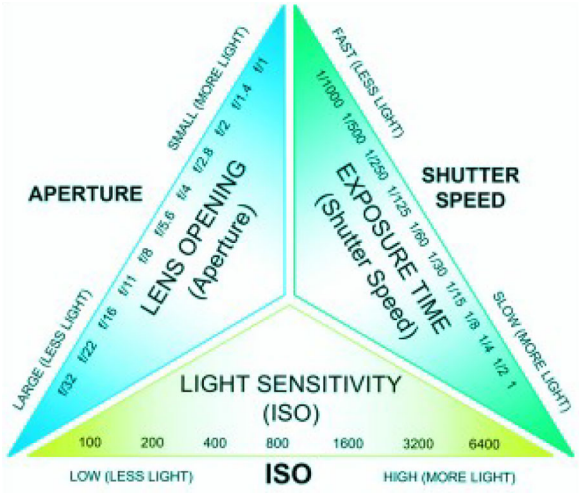

FIGURE 1 Camera triangle of exposure.

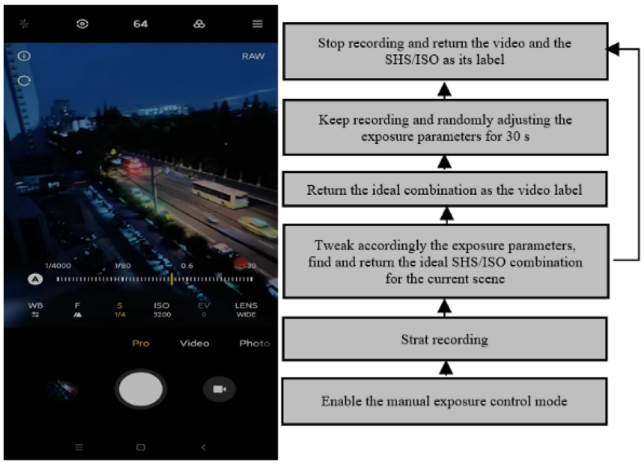

FIGURE 2 A simplified block diagram of the Blur-Prevent dataset collection procedure.

and learning from their skills and accumulated experience. By deploying the trained ExposureNet model for a mobile device, we expect that this model will demonstrate the ability to operate an autonomous blur effects prevention and artistic shooting just like expert photographers do.

Blur-Prevent dataset was collected by recording videos in a wide range of scenes, encompassing various lighting and weather conditions. The majority of the scene videos were captured in Shanghai, Casablanca, and Paris. Dynamic scenes were filmed in public parks, zoos, forests, jungles, public roads, and highways. We collected data in different weather conditions, including clear, cloudy, foggy, snowy, and rainy conditions, spanning both day and night.

Figure 2 depicts a simplified block diagram of the procedure used to collect the Blur-Prevent dataset. Our data collection tool was a mobile application installed on a Huawei Nova Youth smartphone, which allowed full control of the camera's manual mode. This particular smartphone model features a single 11.8 Megapixels rear camera with an aperture of f/2.2. Six volunteer users, all utilizing the same smartphone brand and model, received training from a professional photographer on how to manually adjust the camera's exposure parameters to avoid motion blur in various lighting and dynamic conditions. The volunteers were tasked with recording on-scene videos, with each video labelled with the ideal exposure parameters for preventing motion blur in a specific scene type.

The training data consisted of individual images sampled from the collected video data, each labelled with the corresponding shutter speed and ISO parameters.

The constructed Blur-Prevent dataset has been publicly released and is currently accessible through the following link: https://github.com/nahliabdelwahed/BLUR-PREVENT-DATASET.

4 | EXPOSURENET MODEL ARCHITECTURE

Convolutional and transformer neural networks have had a revolutionary impact on pattern recognition. Previously, pattern recognition tasks typically involved a two-step process: manually extracting features followed by classification. In contrast, modern neural network models [40–43] can automatically learn features from training data. Their power lies in their ability to extract meaningful insights as image features through the convolution operation, and producing attention-weighed features by the mean of an incorporated attention mechanism.

Additionally, by scanning an entire image with convolution kernels, only a small number of parameters need to be learned compared to the total operations.

ExposureNet is a CNN/Transformer hybrid model trained by adjusting the weights to minimize the mean squared error between the predicted exposure parameters and the target parameters.

As depicted in Figure 3a, ExposureNet consists of a total of 12 layers. These layers include a normalization layer, 7 convolutional layers, 2 Transfo-attention blocks, and 6 fully connected layers. Table 1 presents more details about ExposureNet architecture. Although its structure appears to be slightly similar to that of Bojarski and Del [44], but they differ on so many levels, especially, in terms of specific design choices, optimizations, and adaptations made to tailor for camera exposure parameters control.

The input to the network is an RGB image with dimensions of $512 \times 512 \times 3$. This image is processed through the network's layers, resulting in two output values representing the predicted camera exposure parameters.

The first layer of our network performs batch normalization, which remains fixed and unaltered during the learning process.

Applying batch normalization within the network gains more stability and become more robust to changes in input distribution, allowing for faster convergence and improved generalization.

The convolutional layers in our network are specifically designed for feature extraction and were selected based on empirical experimentation with various layer configurations. Each of the first three convolutional layers uses a stride of 2×2

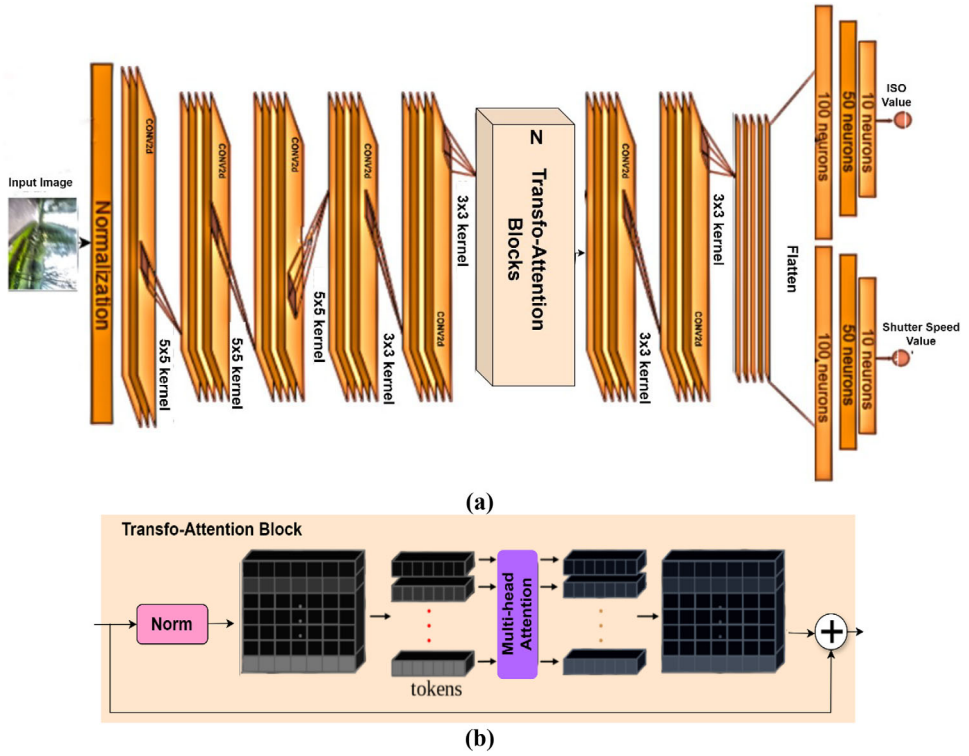

FIGURE 3 (a) ExposureNet architecture: A CNN/Transformer hybrid model structure. (b) Transfo-Attention (TA) block.

TABLE 1 ExposureNet architecture: FC-1 and FC-2 denote the first and the second branches of fully connected layers, respectively.

Layer	Channel	Stride	Kernel size	Padding	Output size
(W, H, Channels)	–	–	–	–	224 × 224 × 3
Normalization	–	–	–	–	224 × 224 × 3
Conv2D, ReLu	512	2 × 2	5 × 5	0	112 × 112 × 512
Conv2D, ReLu	256	2 × 2	5 × 5	0	56 × 56 × 256
Conv2D, ReLu	128	2 × 2	5 × 5	0	28 × 28 × 128
Conv2D, ReLu	64	2 × 2	3 × 3	0	14 × 14 × 64
Conv2D, ReLu	32	1 × 1	3 × 3	0	14 × 14 × 32
TA Block	–	–	–	–	14 × 14 × 32
Conv2D, ReLu	32	1 × 1	3 × 3	0	14 × 14 × 32
Conv2D, ReLu	32	1 × 1	3 × 3	0	14 × 14 × 32
Flatten	–	–	–	–	6272
FC-1	–	–	–	–	10
FC-2	–	–	–	–	10

and a kernel size of 5 × 5, whereas the very next two convolutional layers, on the other hand, employ no striding with a kernel size of 3 × 3.

Following the very first five feature extraction convolutional layers, we incorporate two Transfo-Attention (TA) Blocks, producing attention-weighed features. These resultant features are once again passed through other last two convolutional layers, then flattened and fed into two parallel branches of three subsequent fully connected layers. These fully connected layers aim to map the high-level semantic features learned by the network to the desired camera exposure parameters, resulting in two output values.

Transfo-Attention (TA) block: TA block is a core component of the introduced ExposureNet model. We detail its design as follows.

As illustrated in Figure 3b, a TA block primarily includes a multi-head self-attention layer, where parallel tokens of the extracted features are received as input.

Let the input features of a TA block be $X \in R^{H\times W\times C}$, where H, W, and C stand for the height, the width, and the number of channels, respectively. Initially, these features are passed through a normalization layer (Norm) followed by a 1 × 1 convolution layer (Conv) with C filters to obtain the input features, described as

$$X^b = \text{Conv}\left(\text{Norm}\left(X\right)\right) \tag{2}$$

where $X^b \in R^{H\times W\times C}$ stand for the input features for TA block, where $D = \frac{C}{2}$.

We split the input features X^b into H horizontal non-overlapping strips $X_i^b \in R^{W\times D}$, $i = \{1, 2, \ldots, H\}$. Each strip X_i^b contains W tokens with D dimensions. Next, we generate queries, keys, and values associated with X_i^b as Q_{ij}^b, K_{ij}^b, and $V_{ij}^b \in R^{W\times\frac{D}{m}}$ for the multi-head attention mechanism as follows:

$$\left(Q_{ij}^b, K_{ij}^b, V_{ij}^b\right) = \left(X_i^b P_j^Q, X_i^b P_j^K, X_i^b P_j^V\right) \tag{3}$$

where P_j^Q, P_j^K and $P_j^V \in R^{W\times\frac{D}{m}}$, $j \in \{1, \ldots, m\}$, are linear projection matrices for the query, key, and value with the multi-head attention. We set the number of heads to five, $m = 5$. The multi-head attended feature $O_{ij}^b \in R^{W\times\frac{D}{m}}$ for one strip is calculated as

$$O_{ij}^b = SoftMax\left(\frac{Q_{ij}^b\left(K_{ij}^b\right)^T}{\sqrt{D/m}}\right)V_{ij}^b \tag{4}$$

whose space complexity is $O(W^2)$. We concatenate the multi-head horizontal features $O_{ij}^b \in R^{W\times\frac{D}{m}}$ along the channel dimension to generate $O_i^b \in R^{W\times D}$ and fold all of them into three-dimensional tensors $O^b \in R^{H\times W\times D}$.

We then feed O^b into a 1×1 convolution layer with a residual connection to the original input features X to obtain the attended features $O_{attn} \in R^{H\times W\times C}$ as follows:

$$O_{attn} = \text{Conv}\left(O^b\right) + X \tag{5}$$

5 | TRAINING DETAILS

The initial step in training a neural network involves the selection of frames to utilize. In our case, we have labelled data that corresponds to the exposure parameters for various scene types, weather conditions, and lighting conditions. To train a real-time blur effect prevention convolutional neural network (CNN), we opted to sample our video data at a rate of 15 frames per second (FPS). Choosing a higher sampling rate would result in including images that are very similar, offering limited additional valuable information. For training purposes, we allocated 80% of our collected data for CNN training, reserving the remaining 20% for testing the performance of our model. The training process for our model was carried out on a server equipped with an NVIDIA Tesla T4 GPU. We implemented our framework using the TensorFlow platform [45].

To ensure consistency, all input images were resized to dimensions of 512×512, and a batch size of 16 was utilized. The scalar elements of each label vector were normalized within the range of -1.0 to 1.0. The network underwent training for 150 epochs, employing a stepped learning rate policy where the rate was reduced by half every 15 epochs. We used the Adam optimizer with an initial learning rate of 0.0001, the target objective is to minimize a joint loss function (6), where Δ_{ISO_i} and Δ_{SbS_i} are the CNN computed ISO and Shutter speed values while Y_{ISO_i} and Y_{SbS_i} are the target labels. This function averages the mean square error of each network head, creating the below joint loss between the network outputs.

$$L_j = \sum_{i=1}^{N}\left(\Delta_{ISO_i} - Y_{ISO_i}\right)^2 + \sum_{i=1}^{N}\left(\Delta_{SbS_i} - Y_{SbS_i}\right)^2 \tag{6}$$

ALGORITHM 1 ExposureNet training pseudocode

```
    Input: path to the training dataset
    Output: a trained ExposureNet model
1:  Trainet(x, e, b, lr)
2:      prepare the input data x, as a list of tensors
3:      e is the number of iteration epochs
4:      b is the number of batches
5:      lr is the learning rate
6:      fit the model to the input data x
7:      compile the model
8:      for 0 to e do
9:          for 0 to b do
10:             use the L1 loss function in (3)
11:             train ExposureNet model
12;         end for
13:         return the trained ExposureNet
14      end for
15: end Trainet()
```

Weight adjustment is accomplished through backpropagation, which has been implemented using the TensorFlow machine learning package. The entire process of training the model can be summarized in Algorithm 1.

There is a notable distinction between training a convolutional neural network (CNN) from scratch and employing transfer learning, which leverages pre-trained CNN models to transfer their acquired knowledge from a general domain to a specific sub-domain.

Drawing inspiration from [46], which explored and compared various pre-trained CNN models using the ImageNet dataset to assess their transfer learning performance for a new task, we replace the last SoftMax layers of four cutting-edge pre-trained deep learning image classifiers. Instead, we employ two parallel fully connected networks that adapt their architectures for regression tasks rather than image classification tasks. The VGGNet [47], ResNet [48], Inception [49], and Xception [50] models serve as the backbones for these networks, resulting in four network architectures that appear identical to the one depicted in Figure 3 but differ in the number and configuration of their hidden convolutional layers. Subsequently, we retrain all the networks using our training dataset. This retraining process fine-tunes the hidden convolutional layers and fully connected networks to effectively perform the autonomous task of preventing blur effects. Figures 4 and 5 visualize an overall description of the proposed model training and inference stages, respectively.

6 | PERFORMANCE EVALUATION

Our system evaluation is conducted in two stages. Initially, we quantitatively assess the performance of our network using our dedicated testing dataset. Subsequently, we proceed to a series of on-scene tests in real-world production settings to evaluate its practical performance.

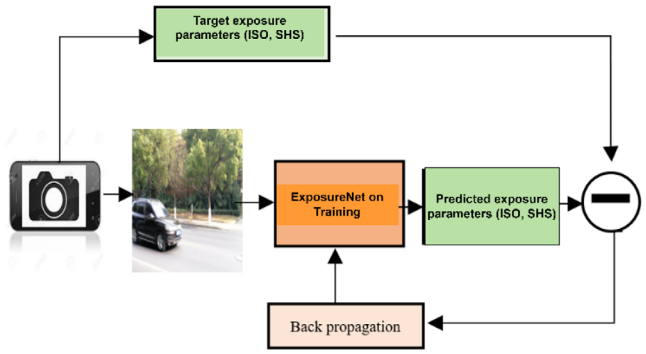

FIGURE 4 ExposureNet training stage.

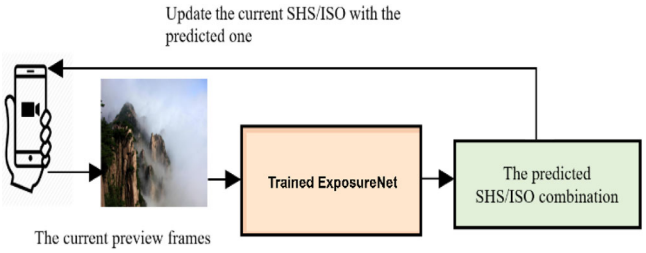

FIGURE 5 ExposureNet inference stage.

6.1 | Quantitative evaluation

To assess the trained model's performance in a quantitative manner, we utilize the testing data subset and execute the trained ExposureNet model for inference. This inference process generates a list of predicted exposure parameters, which allows us to evaluate the model's performance. Since we are dealing with a regression task, we choose the mean square error (MSE) in (7) as well as the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) as evaluation metrics.

$$MSE = \sum_{i=1}^{N} \frac{\left(\Delta_{EV_i} - Y_i\right)^2}{N} \tag{7}$$

where N is the number of testing images, and Δ_{EV_i} and Y_i are the predicted and the target shutter speed/ISO values of image i. The target values are in the range $1 \leq Y_i \leq 10{,}000$ for the shutter speed, whereas in the range $50 \leq Y_i \leq 3200$ for the ISO.

To compare ExposureNet's visual performance with baseline methods, we deployed the trained models on five mobile devices with the same camera aperture and hardware configuration. The devices were mounted side-by-side on a rig and taken on a test to challenge dynamic scenes with intense motion and actions. Each device collected 150 images, which are consequently used as testing set to assess the proposed model performance. Table 2 presents the performance evaluation of ExposureNet using different network structures and training strategies, measured by the MSE assessment metric. An ExposureNet model trained from scratch achieves the best score in terms of MSE while employing the transfer learning strategy which significantly reduces training time and yields comparable performance even with less training data. These findings suggest that superior performance in the ImageNet image classification task does not necessarily translate to regression tasks like blur effect prevention. Table 3 presents the proposed model's quantitative results in terms of the PSNR and SSIM assessment metrics on two benchmarked datasets, GoPro dataset [54] and RSBlur dataset [55], respectively. On GoPro dataset, ExposureNet performance favourably surpasses those of the baselines in terms of SSIM, while it falls behind IRNeXt model [52] with a drop of 2 dB. Similarly, on RSBlur dataset, it excels all the baselines on terms of PSNR, but slightly drops behind MPRNet model [53]. The testing results validate our system design and training strategy.

TABLE 2 The trained system performance on our testing dataset in terms of MSE and number parameters.

Models	Predictions	MSE	Parameters
VGGNet [47]	Exposure time	23.015	80.31 M
	ISO	28.998	
ResNet-50 [48]	Exposure time	21.459	192 M
	ISO	24.349	
MobileNet-V3 [49]	Exposure time	19.012	5.4 M
	ISO	22.008	
EficientNet [50]	Exposure time	17.670	480 M
	ISO	18.921	
Uform [51]	Exposure time	14.301	114 M
	ISO	18.032	
IRNeXt [52]	Exposure time	16.103	N/A
	ISO	**10.874**	
MPRNet [53]	Exposure time	14.972	N/A
	ISO	12.78	
ExposureNet (ours)	Exposure time	**13.070**	25 M
	ISO	15.321	

6.2 | Qualitative evaluation

As a proof-of-concept for a blur effect prevention system, we developed and deployed a test app on the Huawei Nova Youth smartphone. The predicted exposure parameters from the model are used to adjust the current exposure parameters through Google Camera API2. ExposureNet achieves an optimal balance between accuracy and speed, predicting the appropriate exposure parameters at a speed of 26.37 FPS. Figure 6 illustrates the quantitative findings in terms of speed (FPS) and MSE. To visually compare performance, we deployed the trained ExposureNet on three mobile devices (Xiaomi, Google Nexus, and Huawei Nova) and mounted them side-by-side on a rig, similar to that shown in Figure 7. We simultaneously recorded videos in challenging dynamic scenes, aligned and cut them to the same duration, and sampled frames for comparison.

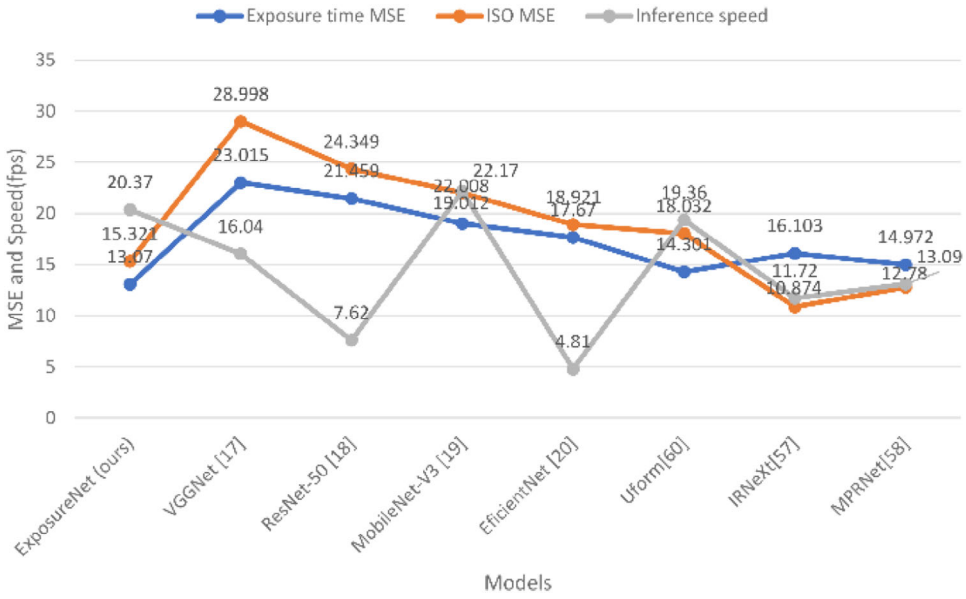

FIGURE 6 ExposureNet performance against the baselines in terms of MSE and speed (FPS) evaluation metrics.

TABLE 3 Quantitative results: Models evaluation in terms of PSNR and SSIM metrics on two benchmarked datasets.

Datasets	Models	PNSR (dB)	SSIM
GoPro dataset [54]	VGGNet [47]	28.60	0.886
	ResNet-50 [48]	27.85	0.819
	MobileNet-V3 [49]	29.24	0.931
	EficientNet [50]	26.58	0,812
	Uform [51]	33.06	0.967
	IRNeXt [52]	**36.09**	0.973
	MPRNet [53]	34.93	0.952
	ExposureNet (ours)	35.76	**0.986**
RSBlur dataset [55]	VGGNet [47]	30.08	0.894
	ResNet-50 [48]	27.85	0.841
	MobileNet-V3 [49]	29.68	0.906
	EficientNet [50]	25.81	0,838
	Uform [51]	30.90	0.953
	IRNeXt [52]	33.97	0.971
	MPRNet [53]	35.02	**0.983**
	ExposureNet (ours)	**37.03**	0.967

To address the issue of viewpoint differences, we followed the approach proposed by [56] and [57], which involves aligning images based on local features and cropping for the best comparison. Figures 8 and 9 display the visual comparison results against the baseline methods.

ExposureNet, as opposed to the baseline methods, was able to prevent the blur effect phenomenon by autonomously adjust-

FIGURE 7 Rig setup for the on-scene test.

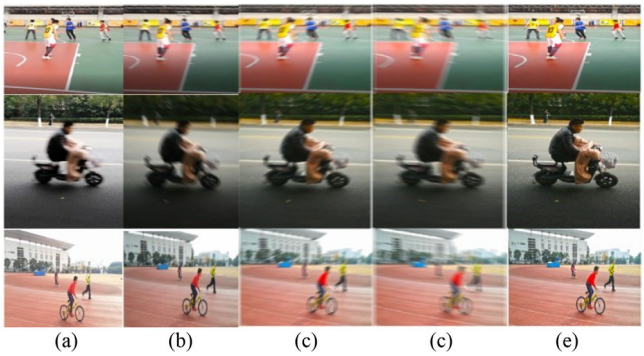

FIGURE 8 Visual comparison of ExposureNet's results against the baseline methods. (a–e) The captured images correspondent to the exposure parameters predicted by VGGNet [47], ResNet-50 [48], MobileNet-V3 [49], EficientNet [50], and ExposureNet (ours), respectively.

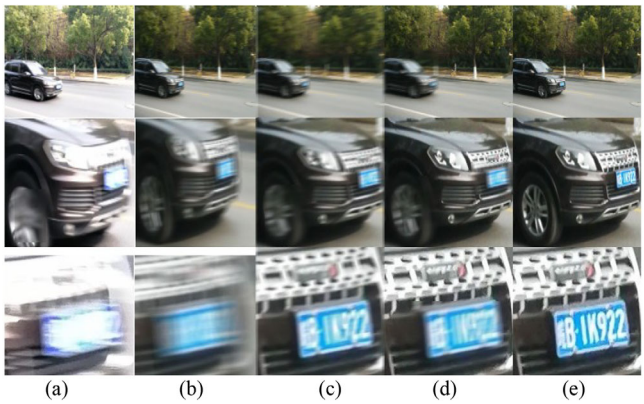

FIGURE 9 Visual comparison of ExposureNet 's results against the baseline methods. (a–e) The captured images correspondent to the exposure parameters predicted by VGGNet [47], ResNet-50 [48], MobileNet-V3 [49], EficientNet [50], and ExposureNet (ours), respectively.

TABLE 4 Ablation results for ExposureNet main components, (a) batch normalization, (b) attention mechanism, (c) stridings, and (d) loss functions.

Normalize	MSE	PSNR
With-norm	**13.070**	**37.03**
Without-norm	18.725	34.28

(a)

Attention	MSE	PSNR
Without TA	15.061	33.07
With TA	**13.070**	**37.03**

(b)

Stride	MSE	PSNR
No stride	**13.070**	**37.03**
2x stride	15.930	35.82
4x stride	21.716	34.14

(c)

Loss	MSE	PSNR
L1-loss	14.831	33.73
Joint-Loss (6)	**13.070**	**37.03**
Huber-Loss	17.115	32.02

(d)

ing the camera exposure parameters to better fit the scene dynamics and lighting conditions. On the other hand, the baseline methods fail to set the appropriate exposure parameters for blur effect prevention in such high-motion scenarios. Thus, producing blur effect corrupted images. These examples demonstrate the ability of our model to learn semantic representations of scenes for blur effect prevention.

6.3 | Ablation study

A set of thorough experiments were performed to evaluate the effectiveness of the key elements of ExposureNet and its training approach. In order to maintain fairness, all models underwent training using identical configuration and settings. The results of these experiments are presented in Table 4. In this section, besides the concise summary mentioned in the main text, a comprehensive explanation of certain crucial network components is provided. These ablation studies were carried out on the recently constructed Blur-Prevent dataset.

6.3.1 | Spatial striding

Pooling or striding [58, 59] techniques are commonly used in neural networks, but they can unintentionally remove important fine details from images, which are crucial for regression tasks such as preventing blur effects. To experimentally validate this, we incorporated 2×2 and 4×4 striding in the network architecture. The findings in Table 4b confirmed our hypothesis, as they demonstrated a decrease in the mean square error (MSE) from 22.215 to 16.430 when no striding was employed. Consequently, we decided to utilize a 2×2 stride in the initial convolution layers and refrain from using striding in the last two convolution layers.

6.3.2 | Batch normalization

When the training is conducted without batch normalization, the performance of ExposureNet is impacted by several challenges. Firstly, the network may suffer from the internal covariate shift problem, where the distribution of inputs to each layer changes during training, making it difficult for the network to converge. Additionally, without batch normalization, the network may require careful initialization and tuning of hyperparameters to ensure stable training and prevent vanishing or exploding gradients.

In contrast, when batch normalization is applied, it significantly improves performance. Batch normalization helps mitigate the internal covariate shift problem by normalizing the activations of each layer with respect to the mini-batch statistics. This normalization process makes the network more robust to changes in input distribution, allowing for faster convergence and improved generalization. Additionally, batch normalization acts as a regularizer, reducing the dependence on dropout and other regularization techniques. By stabilizing the training process, batch normalization enables deeper and more complex neural architectures to be trained effectively, leading to improved accuracy and minimal prediction errors, findings in Table 4a support this hypothesis.

6.3.3 | Loss function

While using solely L1 or L2 loss may enhance the mean square error (MSE) metric, it can occasionally lead to inaccurate predictions. Nevertheless, according to the outcomes outlined in Table 4c, introducing an additional loss function such as Huber loss did not yield any improvement in the MSE metric. In our case, the introduced joint loss function in Equation (6) yields minimal prediction errors.

6.3.4 | Attention mechanism

As depicted in Table 4, the multi-head self-attention mechanism plays a significant role in the network. Notably, the model trained with Transfo-Attention (TA) block demonstrates a focused attention on regions with significant motion, resulting in more confident predictions for blur prevention. This is supported by the observation that the PSNR value in Table 4a decreases from 37.03 to 33.07 when training without the TA block, further validating the effectiveness of

TABLE 5 Blind test result, user votes per system.

Devices	Models	Blur prevention	Night performance	Video stabilization
Xiaomi	VGGNet [47]	12%	13%	6%
	ResNet-50 [48]	4%	9%	17%
	MobileNet-V3 [49]	16%	11%	3%
	EficientNet [50]	7%	15%	5%
	ExposureNet (ours)	61%	52%	69%
Huawei	VGGNet [47]	14%	8%	14%
	ResNet-50 [48]	9%	19%	6%
	MobileNet-V3 [49]	12%	3%	18%
	EficientNet [50]	17%	4%	8%
	ExposureNet (ours)	48%	66%	54%

the multi-head self-attention mechanism for this deblurring task.

Ablation results revealed the importance and the influence of each designed component. By taking in consideration the findings of this ablation study, we settle on the final model structure in Figure 3, which enables a robust motion blur effect prevention.

6.4 | User study

To compare the performance of our system with related approaches, we conducted a user study. To facilitate this study, we constructed a rig that enabled us to simultaneously film with three smartphones, capturing nearly identical angles. The complete setup can be seen in Figure 7. We mounted three mobile devices, namely Google Nexus, Xiaomi, and Huawei Nova, running the ExposureNet app, side-by-side on the rig.

We recorded multiple videos across various scenes. In the blind comparison test, we invited over two hundred volunteer users to vote for the best system performance in terms of real-time blur prevention, night performance, and video stabilization. Each user was presented with five videos: one from our system and the other four from the baseline methods. The videos were randomly arranged, and the users were asked to select the best among the five choices labelled "A," "B", "C", "D", and "E". The results of this A–E test are summarized in Table 5 and Figure 10, confirming the effectiveness of ExposureNet's design and training strategy for autonomous blur effect prevention.

7 | DISCUSSION

In this research work, we explored the potential and the effectiveness of a newly developed auto-exposure metadata-driven approach for blur effect prevention.

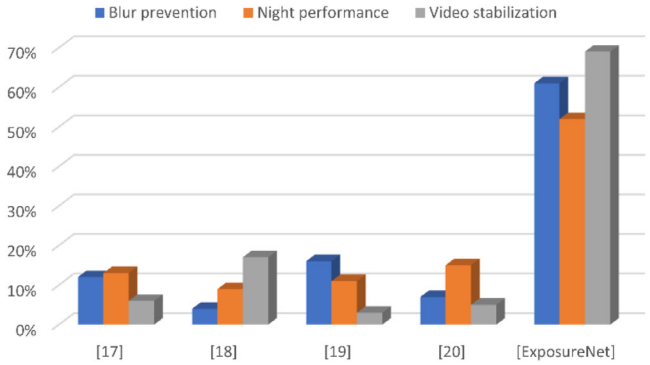

FIGURE 10 Blind user study result: Users vote for the best performance.

The introduced model in this paper is solely designed for mobile devices with adjustable software-based aperture and fixed hardware-based aperture, which is the trend for the wide majority of the novel generation of smartphones. It is also worth mentioning that the trending adjustable software-based aperture is no more than a post-processing stage, which aims to mimic the effects of a real camera aperture, which does not really interfere with the proposed method in the image acquisition stage. However, in future work, we plan to expand this work to cover other image acquisition accessories with adjustable hardware-based aperture, for instance, DSLR cameras.

The constructed Blur-Prevent dataset is still the first dataset of its kind, which holds promise for blur effect prevention. It facilitates end-to-end model training by taking advantage of camera exposure metadata compressed in images, enhancing the potential for advancements in this field.

Beyond blur effect prevention, the proposed ExposureNet model can easily be deployed into a core element of a blur-aware camera ISP system, as it can perform one of the essential subtasks of an ISP system in an end-to-end fashion. In future work, we plan to investigate the possibility of extending ExposureNet capabilities, to cover dual and multi-camera-equipped devices, and involving a wider range of functions, like, video interpolation and high dynamic range reconstruction.

8 | CONCLUSION

In this research article, we introduced an autonomous method for controlling camera exposure parameters to prevent blur effects. We demonstrated through empirical experiments that ExposureNet, a CNN/Transformer hybrid model, can learn the entire task of blur effect prevention without the need for manual decomposition or handcrafted processing steps. Surprisingly, even with a small amount of training data, our system effectively enables smartphone cameras to operate artistically in various conditions while being fully aware of image corruption phenomena that need to be prevented. By using just two camera exposure parameters (shutter speed and ISO) as training signals, ExposureNet successfully learns meaningful scene features.

ExposureNet is designed to be compact, low-latency, and low-power, meeting the resource constraints of mobile devices

and striking a balance between latency, size, and accuracy. During the training process, the system learns to identify scene types (dynamic or static) as well as lighting and weather conditions without relying on explicit labels. We conducted extensive experiments on our testing dataset, and the results were highly promising, confirming the effectiveness of the system's design and training strategy.

Furthermore, our system is not limited to mobile devices, it can also be deployed on a wide range of camera-equipped devices, including DSLR and mirrorless cameras. Therefore, future work will explore the deployment of this system in other real-world applications. This study has opened up new possibilities for promising research opportunities in the field.

AUTHOR CONTRIBUTIONS

Abdelwahed Nahli: Conceptualization; methodology; software; validation; formal analysis; investigation; resources; data curation; writing—original draft preparation; writing—review and editing; visualization; project administration. **Dan Li**: Resources; writing—review and editing; supervision; funding acquisition. **Rahim Uddin**: Validation. **Muhammad Irfan**: Validation. **Mohamed Oubibi**: Writing—review and editing. **Qiyong Lu**: Writing—review and editing. **Jian Qiu Zhang**: Resources; writing—review and editing; supervision; funding acquisition. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGEMENTS

This study was funded by Fudan University and the National Natural Science Foundation of China. The authors express appreciation to their entire laboratory team for engaging discussions and access to the ultra-fast computing resources.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author.

ORCID

Abdelwahed Nahli https://orcid.org/0000-0001-5281-9809

REFERENCES

1. Thomas, J.D.: Nikon D800 & D800E Digital Field Guide. John Wiley & Sons, Hoboken, NJ (2012)
2. Sampat, N., Venkataraman, S., Yeh, T., Kremens, R.L.: System implications of implementing auto-exposure on consumer digital cameras. Proc. SPIE 3650, 100–107 (1999)
3. Brierley, J.: Digital Bird Photography-A Comprehensive Tutorial. Lulu.com, Morrisville, NC (2005)
4. Kiran, B.R., Vardhan, G.H.: A fast auto exposure algorithm for industrial applications based on false-position method. In: Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA). pp. 509–515 (2013)
5. Kiran, B.R., Krishna, K.M., Basha, S.S., Vardhan, H., Raju, N., Kumari, V.V.: False position based auto exposure algorithm for properly exposing the leather samples in the leather industries. Int. J. Comput. Appl. 83(14), 7–9 (2013)
6. Vuong, Q.K., Yun, S.H., Kim, S.: A new auto exposure and auto white-balance algorithm to detect high dynamic range conditions using CMOS technology. In: Proceedings of the World Congress on Engineering and Computer Science. San Francisco, pp. 22–24 (2008)
7. Liang, J., Qin, Y., Hong, Z.: An auto-exposure algorithm for detecting high contrast lighting conditions. In: 2007 7th International Conference on ASIC. pp. 725–728. (2007)
8. Nahli, A., Cao, S., Jia, Z., Ma, R., Xu, S.: Dataset and network structure: Towards frames selection for fast video deblurring. IEEE Access 9, 61369–61382 (2021)
9. Nahli, A., Cao, Y., Xu, S.: Face image deblurring: A data-driven strategy. In: Proceedings of the GCAI Global Conference of Artificial Intelligence. 59–69 (2020)
10. Kuno, T., Sugiura, H., Matoba, N.: A new automatic exposure system for digital still cameras. IEEE Trans. Consum. Electron. 44(1), 192–199 (1998)
11. Tedla, S., Yang, B., Brown, M.S.: Examining Autoexposure for Challenging Scenes. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 13076–13085 (2023)
12. Onzon, E., Mannan, F., Heide, F.: Neural auto-exposure for high-dynamic range object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7710–7720 (2021)
13. Jiang, T., Kuhnert, K.D., Nguyen, D., Kuhnert, L.: Multiple templates auto exposure control based on luminance histogram for onboard camera. In: 2011 IEEE International Conference on Computer Science and Automation Engineering. pp. 237–241 (2011)
14. Su, Y., Lin, J.Y., Kuo, C.C.: A model-based approach to camera's auto exposure control. J. Visual Commun. Image Represent. 36, 122–129 (2016)
15. Mosbrucker, A.R., Major, J.J., Spicer, K.R., Pitlick, J.: Camera system considerations for geomorphic applications of SfM photogrammetry. Earth Surf. Processes Landforms 42(6), 969–986 (2017)
16. Odinaev, I., Chin, J.W., Luo, K.H., Ke, Z., So, R.H., Wong, K.L.: Optimizing camera exposure control settings for remote vital sign measurements in low-light environments. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6085–6092 (2023)
17. Görmer, S., Hold, S., Kummert, A., Iurgel, U., Meuter, M.: Multi-exposure image acquisition for automotive high dynamic range imaging. In: 13th International IEEE Conference on Intelligent Transportation Systems. pp. 1881–1886 (2010)
18. Kuang, Q., Jin, X., Zhao, Q., Zhou, B.: Deep multimodality learning for UAV video aesthetic quality assessment. IEEE Trans. Multimedia 22(10), 2623–2634 (2019)
19. Torres, J., Menéndez, J.M.: Optimal camera exposure for video surveillance systems by predictive control of shutter speed, aperture, and gain. Proc. SPIE 9400, 238–251 (2015)
20. Yang, H., Wang, B., Vesdapunt, N., Guo, M., Kang, S.B.: Personalized exposure control using adaptive metering and reinforcement learning. IEEE Trans. Visual Comput. Graphics 25(10), 2953–2968 (2018)
21. Liu, X., Wang, Y., Xie, S., Zhang, X., Ma, Z., McDuff, D., Patel, S.: Mobilephys: Personalized mobile camera-based contactless physiological sensing. Proc. ACM Interact. Mobile Wearable Ubiquitous Technol. 6(1), 1–23 (2022)
22. Shen, S., Mulgaonkar, Y., Michael, N., Kumar, V.: Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV. In: 2014 IEEE International Conference on Robotics and Automation (ICRA). pp. 4974–4981 (2014)
23. Nia, H.F., Hu, H., Gan, J.Q.: A novel fuzzy logic approach to online exposure time calculation of line scan cameras in industrial inspection. Int. J. Modell. Identif. Control 21(1), 8–16 (2014)
24. Lu, H., Zhang, H., Yang, S., Zheng, Z.: Camera parameters auto-adjusting technique for robust robot vision. In: 2010 IEEE International Conference on Robotics and Automation. pp. 1518–1523 (2010)
25. Wang, H., Kim, B., Xie, J.L., Han, Z.: LEAF+ AIO: Edge-assisted energy-aware object detection for mobile augmented reality. IEEE Trans. Mob. Comput. 22(10), 5933–5948 (2022)
26. Han, B., Jia, X., Song, R., Ran, F., Rao, P.: Auto complementary exposure control for high dynamic range video capturing. IEEE Access 9, 144285–144299 (2021)

27. Wu, Y., Tsotsos, J.: Active control of camera parameters for object detection algorithms. arXiv preprint arXiv:1705.05685 (2017)
28. Skandarajah, A., Reber, C.D., Switz, N.A., Fletcher, D.A.: Quantitative imaging with a mobile phone microscope. PLoS One 9(5), e96906 (2014)
29. Li, H., Zou, Y., Chen, N., Lin, J., Liu, X., Xu, W., Zheng, C., Li, R., He, D., Kong, F., Cai, Y.: MARS-LVIG dataset: A multi-sensor aerial robots SLAM dataset for LiDAR-visual-inertial-GNSS fusion. Int. J. Rob. Res. 02783649241227968 (2024)
30. Gu, Q., Al Noman, A., Aoyama, T., Takaki, T., Ishii, I.: A fast color tracking system with automatic exposure control. In: 2013 IEEE International Conference on Information and Automation (ICIA). pp. 1302–1307 (2013)
31. Piao, Y., Xu, W.: Method of auto multi-exposure for high dynamic range imaging. In: 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering. pp. 93–97 (2010)
32. Zhang, X., Lee, J.Y., Sunkavalli, K., Wang, Z.: Photometric stabilization for fast-forward videos. Comput. Graphics Forum 36(7), 105–113 (2017)
33. Yang, H., Wang, B., Vesdapunt, N., Guo, M., Kang, S.B.: Personalized attention-aware exposure control using reinforcement learning. arXiv preprint arXiv:1803.02269, 53 (2018)
34. Li, S., Handa, A., Zhang, Y., Calway, A.: HDRFusion: HDR SLAM using a low-cost auto-exposure RGB-D sensor. In: 2016 Fourth International Conference on 3D Vision (3DV). pp. 314–322 (2016)
35. Dasari, R.: Multi-sensor fusion for fast and robust computer vision applications. Doctoral dissertation, State University of New York, Buffalo. (2020)
36. Ji, W., Chen, X., Li, Y.: Blind motion deblurring using improved DeblurGAN. IET Image Proc. 18(2), 327–347 (2024)
37. Li, Y., Shu, X., Ren, D., Li, Q., Zuo, W.: Joint learning of motion deblurring and defocus deblurring networks with a real-world dataset. Neurocomputing 565, 126996 (2024)
38. Hu, B., Wang, S., Gao, X., Li, L., Gan, J., Nie, X.: Reduced-reference image deblurring quality assessment based on multi-scale feature enhancement and aggregation. Neurocomputing 547, 126378 (2023)
39. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. In: Proceedings of Neural Information Processing Systems. 30 (2017)
40. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. Neural Comput. 1(4), 541–551 (1989)
41. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25(2), (2012). http://doi.org/10.1145/3065386
42. Jackel, L.D., Sharman, D., Stenard, C.E., Strom, B.I., Zuckert, D.: Optical character recognition for self-service banking. AT&T Tech. J. 74(4), 16–24 (1995)
43. Russakovsky, O., Deng, J., Su, H.: Imagenet large scale visual recognition challenge. Int. J. Comput. Vision. 115, 211–252 (2015)
44. Bojarski, M., Testa, D.D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L.D., Monfort, M., Muller, U., Zhang, J., Zhang, X.: End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316 (2016)
45. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S.: Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016)
46. Kornblith, S., Shlens, J., Le, Q.V.: Do better imagenet models transfer better? In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2019. pp. 2661–2671 (2019)
47. Ding, X., Zhang, X., Ma, N., Han, J., Ding, G., Repvgg, S.J.: Making vgg-style convnets great again. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 13733–13742 (2021)
48. Bello, I., Fedus, W., Du, X., Cubuk, E.D., Srinivas, A., Lin, T.Y., Shlens, J., Zoph, B.: Revisiting resnets: Improved training and scaling strategies. Adv. Neural Inf. Process. Syst. 34, 22614–22627 (2021)
49. Howard, A., Sandler, M., Chu, G., Chen, L.C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q.V.: Searching for mobilenetv3. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1314–1324 (2019)
50. Touvron, H., Vedaldi, A., Douze, M., Jégou, H.: Fixing the train-test resolution discrepancy. Adv. Neural Inf. Process. Syst. 32, 8250–8260 (2019)
51. Wang, Z., Cun, X., Bao, J., Zhou, W., Liu, J., Li, H.: Uformer: A general u-shaped transformer for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17683–17693 (2022)
52. Cui, Y., Ren, W., Yang, S., Cao, X., Knoll, A.: IRNeXt: Rethinking convolutional network design for image restoration. In ICML (2023)
53. Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H., Shao, L.: Multi-stage progressive image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14821–14831 (2021)
54. Nah, S., Kim, T.H., Lee, K.M.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3883–3891. (2017)
55. Rim, J., Kim, G., Kim, J., Lee, J., Lee, S., Cho, S.: Realistic blur synthesis for learning image deblurring. In: European Conference on Computer Vision. pp. 487–503. (2023)
56. Wang, B., Yu, Y., Xu, Y.Q.: Example-based image color and tone style enhancement. ACM Trans. Graphics 30(4), 1–2 (2011)
57. Hwang, S.J., Kapoor, A., Kang, S.B.: Context-based automatic local image enhancement. In: Computer Vision–ECCV 2012: 12th European Conference on Computer Vision. Florence, Italy, pp. 569–582 (2012)
58. He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans. Pattern Anal. Mach. Intell. 37(9), 1904–1916 (2015)
59. Hodosh, M., Young, P., Hockenmaier, J.: Framing image description as a ranking task: Data, models and evaluation metrics. J. Artif. Intell. Res. 47, 853–899 (2013)

How to cite this article: Nahli, A., Li, D., Uddin, R., Irfan, M., Oubibi, M., Lu, Q., Zhang, J.Q.: ExposureNet: Mobile camera exposure parameters autonomous control for blur effect prevention. IET Image Process. 1–12 (2024). https://doi.org/10.1049/ipr2.13182