doi:10.55385/kastamonujes.1839652 KUJES 12(1):1-11, 2026

Kastamonu University
Journal of Engineering and Sciences
e-ISSN 2667-8209
http://dergipark.gov.tr/kastamonujes

# Image Processing and Deep Learning Based Illumination Intensity (Lux) Estimation: An Application with MobileNetV2 Architecture

**Berat Yıldız[a,*], Cansu Ünlü[b], Selami Balcı[b]**

[a,c] Department of Electrical and Electronics Engineering, Faculty of Engineering, Karamanoğlu Mehmetbey University, Karaman, Türkiye
[b] Institute of Science, Karamanoğlu Mehmetbey University, Karaman, Türkiye
***Corresponding Author:** beratyildiz@kmu.edu.tr

Received: **December 10, 2025** ◆ Accepted: **January 09, 2026** ◆ Published Online: **June 24, 2026**

**Abstract:** This study presents a deep learning-based approach for estimating illuminance (Lux) from ambient photographs with high accuracy, as an alternative to physical luxmeter sensors. A unique dataset consisting of 729 ambient images at 1482x855 resolution and their corresponding lux values was used in the study. A customized cropping algorithm was developed to reduce noise (walls, ceilings, dead zones) in the images. The model architecture used the MobileNetV2 network, proven in image classification, and adapted it to the regression problem via transfer learning. After training, the model reduced the Mean Absolute Error (MAE) value to 0.78 Lux on the validation dataset. Furthermore, the model's $R^2 - score$ demonstrated high stability. The findings indicate that the developed method can precisely measure ambient illuminance using only camera images, without the need for expensive hardware.

**Keywords:** Deep Learning, MobilNetV2, Lux Estimation, Image Proccessing, Regression

# Görüntü İşleme ve Derin Öğrenme Tabanlı Aydınlatma Şiddeti (Lüx) Kestirimi: MobileNetV2 Mimarisi ile Bir Uygulama

**Öz:** Bu çalışma, fiziksel lüxmetre sensörlerine alternatif olarak, ortam fotoğrafları üzerinden aydınlatma şiddetini (Lüx) yüksek doğrulukla tahmin edebilen derin öğrenme tabanlı bir yaklaşım sunmayı amaçlamaktadır. Çalışmada, 1482x855 çözünürlüğünde 729 adet ortam görüntüsü ve bunlara karşılık gelen lüx değerlerinden oluşan özgün bir veri seti kullanılmıştır. Görüntüler üzerinde gürültüyü (duvar, tavan, ölü alanlar) azaltmak amacıyla özelleştirilmiş bir kırpma (cropping) algoritması geliştirilmiştir. Model mimarisi olarak, görüntü sınıflandırmada başarısı kanıtlanmış MobileNetV2 ağı kullanılmış ve transfer öğrenme (transfer learning) yöntemiyle regresyon problemine uyarlanmıştır. Eğitim sonucunda model, doğrulama veri setinde Ortalama Mutlak Hata (MAE) değerini 0.78 Lüx seviyesine kadar düşürmüştür. Ayrıca modelin $R^2 - skoru$ yüksek bir kararlılık göstermiştir. Elde edilen bulgular, geliştirilen yöntemin pahalı donanımlara gerek kalmadan, sadece kamera görüntüleri kullanılarak ortam aydınlatmasının hassas bir şekilde ölçülebileceğini göstermektedir.

**Anahtar Kelimeler:** Derin Öğrenme, MobilNetV2, Lüx Tahmini, Görüntü İşleme, Regresyon

## 1. Introduction

Lighting is not only a physical element that supports visual tasks in workplaces; it is a critical design component that directly affects employee health, psychological state, and work productivity. Today, with the growing importance of sustainability and user-centered principles, the trend towards lighting approaches that prioritize human well-being has increased; well-designed systems have been shown to protect eye health, reduce occupational accidents, increase satisfaction, and improve performance by enhancing attention levels [1-3]. On the other hand, inadequate or uncomfortable lighting leads to user interventions due to glare, reflections, and excessive brightness, and reduces energy efficiency [1]. Because light plays a key role in the biological clock, inappropriate exposure to light can disrupt melatonin and cortisol cycles, thereby increasing sleep, stress, and fatigue. In contrast, appropriate light profiles support employee well-being by regulating the circadian rhythm [3].

Changes in working models and flexible office layouts after the pandemic have left traditional lighting technologies with limited dynamic control capabilities inadequate; in contrast, LED-based systems that are sensitive to daylight and color temperature have been shown to affect mood and focus positively [2, 4]. Monitoring these biological and psychological impacts requires a high level of sensory precision within the built environment. Traditional sensors often provide fragmented data, making it difficult to maintain the dynamic lighting conditions necessary for human well-being. Consequently, the transition from purely health-oriented lighting design to automated, intelligent control systems necessitates robust illuminance estimation techniques. Recent studies have shifted focus toward using ambient visual data as a proxy for physical sensors, aiming to bridge the gap between biological needs and real-time environmental adjustments. In this context, accurate, real-time determination of illuminance levels in indoor spaces has become a fundamental requirement; however, the pinpoint measurement required by physical sensors and the costly infrastructure create significant limitations. Image processing and deep learning-based methods developed in recent years enable highly accurate lux estimation from camera images, allowing detection of unnecessarily illuminated areas and real-time system optimization. This allows for more efficient use of natural light, reduces energy consumption, and makes lighting a sustainable, adaptable component for user needs in innovative building applications.

Various control methods are used in lighting control systems to increase energy efficiency, ensure user comfort, and improve space functionality. These methods offer different advantages depending on their technological level and sensitivity.

- *Schedule-Based Control:* In this method, lighting is automatically turned on and off according to predetermined operating hours. While relatively easy and economical to implement, it is inflexible because it does not account for actual ambient light levels, daylight levels, or the presence of occupants. For example, a typical example of this method is turning the lights on at 8:00 AM and off at 6:00 PM in an office [5].
- *Motion and Presence Detection Control:* PIR sensors or camera-based detection systems monitor the presence of people in a space. Energy savings are achieved by dimming or completely turning off lights when no one is present. While highly effective when configured correctly, false detections can negatively impact the user experience. Therefore, it requires precise tuning [6].
- *Daylight Harvesting Control:* This approach measures the amount of natural light entering a space and automatically adjusts the artificial light level accordingly. Because light demand is naturally lower in areas around windows, lamp brightness is reduced; interior spaces are illuminated as needed. Deep learning-based lux estimation methods allow for more precise daylight estimation in this process [7].
- *Sensor and Artificial Intelligence Supported Smart Systems:* The actual brightness level of the space can be estimated using artificial intelligence models trained on camera images. These systems adjust the light level in each area to the appropriate level based on current conditions. The literature indicates that such smart systems achieve energy savings of 25–60%. They also significantly reduce the need for manual control in large and complex structures [8].
- *Regional (Zoning) Control Approach:* A building or interior space is divided into lighting zones based on its intended use and needs. This allows each area's lighting level to be adjusted independently. For example, meeting rooms, corridors, and open office spaces can be managed with different brightness levels. This method optimizes energy use and ensures functional comfort [9].
- *Color Temperature and Light Characteristics Control:* With CCT control, the color and temperature of the light can be adjusted for the user or over time. Given human biology, cooler, brighter light in the early hours of the day promotes alertness, while warmer tones in the evening promote rest and help adjust the circadian rhythm. Therefore, dynamically managing color temperature is common in modern lighting systems [10].

While the aforementioned methods, ranging from schedule-based control to daylight harvesting, provide foundational energy savings, they often rely on point-based measurements or high-cost sensor deployments. In contrast, the proposed image-based approach offers three distinct advantages: (1) Spatial Coverage: Unlike PIR or lux meter sensors that provide single-point data, our method generates a comprehensive lighting map of the entire scene. (2) Cost-Efficiency: It utilizes existing imaging infrastructure, eliminating the need for specialized hardware. (3) Adaptive Precision: Through the ROI-based strategy, our method filters out non-photometric noise, a capability absent in traditional threshold-based systems.

Physical sensors (lux meters, PIR sensors, etc.) used in traditional lighting control systems are less effective in large-volume or dynamic operating environments due to their limited data provision and installation costs. In particular, the restricted viewing angles of the sensors and the cabling requirements that compromise the space's aesthetic integrity have heightened the need for alternative measurement methods. The proliferation of imaging devices such as security cameras, mobile devices, and webcams has made the "illumination estimation from image" approach a powerful alternative. In particular, the idea of using imaging devices (e.g., cameras and smartphones) as "virtual lux meters" is increasingly common in the literature. For example, Salmerón-Campillo et al. [11] demonstrated that smartphone cameras offer high linearity and reliability for measuring surface illuminance, especially under indoor lighting conditions, compared with ambient light sensors [11]. Such studies confirm that existing imaging infrastructure can be used instead of costly hardware. However, processing raw image data to obtain a meaningful illuminance value (Lux) is challenging due to the complexity introduced by environmental reflections and different light sources. Traditional image processing techniques (histogram analysis, pixel brightness averaging, etc.) may be insufficient to distinguish the complexity introduced by other light sources in the environment (a mixture of natural and artificial light) and by reflections caused by surface material properties. At this point, Deep Learning, especially Convolutional Neural Networks (CNNs), comes into play, with their ability to learn complex patterns in the data. Xu et al. [12] argued that illumination estimation should be considered a regression problem, and with the deep metric-learning-based model they developed, they achieved much higher success in directly extracting illumination values from images than traditional methods [12]. Similarly, Gardner et al. [13] showed that Deep Convolutional Neural Network (CNN) models, which estimate the location and intensity of light sources in an environment from a single indoor image, are an effective tool for indoor lighting analysis [13].

The high computational requirements of deep learning models make their use in mobile or embedded systems difficult. To overcome this problem, the literature suggests using lightweight architectures, such as MobileNet. A recent study compared different architectures, including VGG16, ResNet50, and MobileNet, for illumination and brightness estimation on apple images. It was reported that models trained with transfer learning could achieve high accuracy even across varying color temperatures [14]. In this study, based on approaches reported in the literature, the transfer learning-based MobileNetV2 architecture was chosen to estimate illumination intensity from an ambient image. Unlike the literature, a customized Region of Interest (ROI) cropping strategy was developed to prevent the model from being affected by noisy data (e.g., walls, ceilings, etc.) and to maximize regression accuracy.

The core novelty of this research lies in its data-centric optimization strategy, which distinguishes it from existing methods that rely solely on raw, full-frame image inputs [12, 13]. A detailed review of the existing literature indicates that feeding full-frame images directly into deep learning models can negatively affect learning due to "photometric noise" from regions such as wall surfaces, ceiling textures, and dark corners with low photon density that lack significant lighting information [16]. Unlike traditional approaches that may struggle with these irrelevant features, our study introduces a customized ROI-based cropping algorithm. This algorithm identifies pixel clusters with minimal photometric contribution using statistical variance analysis, luminance-histogram-based thresholding, and local contrast maps, while processing only regions with high luminance gradients. This targeted preprocessing effectively optimizes the input space, increases the signal-to-noise ratio, and significantly reduces the risk of overfitting.

The main technical objective of this study is to develop a high-performance, deep learning-based system for real-time illumination estimation that overcomes the inherent limitations of physical luxmeters, such as single-point measurement constraints, spatial resolution gaps, and high maintenance costs. While many studies in the literature focus on high-capacity but computationally intensive architectures such as VGG16 or ResNet50 for lighting estimation [14], this work adopts the lightweight MobileNetV2 architecture. This choice ensures that the system remains computationally efficient and suitable for deployment on embedded systems and mobile devices without sacrificing accuracy [15].

By achieving an industrial-scale competitive accuracy of 0.78 MAE on an original dataset of 729 images, our model demonstrates strong generalization across diverse lighting conditions, surface reflectance, and spatial arrangements. Furthermore, the proposed method enables the generation of high-resolution "lux maps" that capture the regional illuminance distribution across an entire space. This supports advanced smart-building strategies, such as daylight harvesting, automatic dimming, and occupancy-based energy optimization. Ultimately, this research directly addresses existing gaps in the literature, which often rely on low-resolution sensor data, by proposing an integrated architecture that enables both precise image-based lux estimation and the simultaneous optimization of energy efficiency and visual comfort.

## 2. Material and Method

In this section, the characteristics of the dataset used to develop the proposed deep learning model, the data preprocessing steps, the applied model architecture, and the training strategies are detailed.

### Dataset

This study addresses the lack of a standardized dataset in the literature for deep learning-based indoor lighting estimation by generating a unique dataset that reflects real-world conditions and analyzing the effects of different surface colors and material properties on illumination. To create a training dataset free of environmental noise and with fully controllable parameters, DIALux simulation software, considered the industry standard for lighting design and calculations, was used. The DIALux simulation environment enables precise modeling of various light source spectra, including standard CIE illuminants representing LED, fluorescent, and natural daylight conditions. By systematically varying these parameters, the dataset inherently accounts for the differences in Spectral Power Distribution (SPD) that characterize diverse indoor lighting environments. Consequently, this dataset, consisting of 729 digital images, was captured from a fixed angle and encompasses various wall, floor, and ceiling colors, as well as variations in natural and artificial illumination across different times of day. Simultaneously with the images (1482x855 pixels, 3-channel RGB), the ambient illumination intensity was measured and recorded with a lux meter, and these values were used as target labels for model training. The dataset also captures light distribution due to surface properties, such as light-colored surfaces reflecting more light and dark surfaces absorbing more light; lux values span a wide dynamic range from dark to fully illuminated environments. The prepared dataset was randomly split into 80% for training and 20% for validation to develop the model and evaluate its performance. An additional 100 previously unseen images were allocated as a test set to test the model's generalization ability. The distribution of the mentioned dataset is summarized in Table 1. This comprehensive dataset, collected by systematically varying surface colors, materials, and lighting conditions, significantly improved the deep learning model's accuracy, reliability, and generalization.

**Table 1.** The distribution of the mentioned data

| Dataset | Number of Images | Objective of use | Ratio |
|---|---|---|---|
| **Train** | 583 | Updating model weights | 80% |
| **Validation** | 146 | Hyperparameter tuning and performance monitoring | 20% |
| **Test** | 100 | Final performance measurement (Unseen Data) | - |
| **Total** | 829 | | |

To illustrate the dataset's characteristics and the visual equivalents of lighting conditions in the work environment, four randomly selected sample images are presented in Figure 1. These examples represent the variety of visual data used as input by the model during training. The simultaneously recorded reference illumination values for the images are 109, 100, 91.4, and 77.5 lux, respectively, indicating subtle light variations in the dataset and the labeling accuracy.

**Data Preprocessing**

Using raw image data directly in deep learning models often increases computational costs and can cause the model to focus on irrelevant features (noise) [16]. In this study, a three-stage preprocessing procedure was applied before feeding the images to the model to maximize lighting estimation performance: (1) Region of Interest (ROI) identification and cropping, (2) Resizing, and (3) Data normalization.

When the original images obtained from the study environment, with a resolution of $1482 \times 855$ pixels, were examined, dead areas on the upper ceiling, dark areas of the walls independent of lighting, and dead areas on the right side of the image that do not carry lighting information were identified. These areas are sources of noise that make it difficult for the regression model to learn the lux value. To address this problem, a static cropping algorithm was developed. While the current cropping algorithm utilizes static coordinates, this approach is optimized explicitly for innovative building applications or industrial facilities where cameras are typically mounted at fixed reference positions. By focusing the model's input on the region of highest photometric significance, the proposed method prioritizes high-precision regression over broad spatial generalization. This targeted strategy is crucial for reducing environmental noise from non-photometric areas, such as ceilings, thereby achieving the industrial-scale accuracy required for real-time energy management. According to the coordinate system defined on the original image matrix $I_{orijinal}$ $(x\text{: } column, y\text{: } row)$, the cropped areas are as follows:

- *Vertical Axis (y):* 60 pixels were removed from the top of the image (ceiling area) and 95 pixels from the bottom (floor space). Thus, the active area on the $y - axis$ was reduced to the range $y \in [60, 760]$.
- *Horizontal Axis (x):* 572 pixels were removed from the right side of the image (irrelevant area). Using the left edge as the starting point (0), the active area on the $x - axis$ was defined as $x \in [0, 910]$.

Mathematically, the dimensions of the resulting new Region of Interest (ROI) image $I_{roi}$'nin are calculated as shown in Equations 1 and 2. Here, $H$ represents the pixel length on the vertical axis of the image, and $W$ represents the pixel length on the horizontal axis.

$$H_{new} = H_{old} - \left(T_{crop} - B_{crop}\right) = 855 - (60 + 95) = 700\, piksel \quad (1)$$

$$W_{new} = W_{old} - R_{crop} = 1482 - 572 = 910\, piksel \quad (2)$$

As a result, the images used for model training focused on the $910 \times 700$ pixel area where illumination information was most dense. The resulting $910 \times 700$ $I_{roi}$ images were rescaled to $224 \times 224$ pixels ($I_{input}$) using bilinear interpolation to meet the input layer requirements of the MobileNetV2 architecture [15], as shown in Figure 2. After the scaling process, pixel intensities were normalized to the $[-1, 1]$ range in accordance with the MobileNetV2 preprocessing standard [15] to accelerate model convergence. This process can be expressed as Equation 3:

$$x_{norm} = \frac{x}{127.5} - 1 \quad (3)$$

Here $x$, represents the original pixel value in the range $[0, 225]$ and $x_{norm}$ represents the normalized value used by the model.

Furthermore, to reduce the risk of overfitting in deep networks trained on a limited dataset, an "on-the-fly augmentation" technique [17] was applied during training. Because illumination intensity (Lux) is a direction-independent scalar quantity, the images in the training set were randomly flipped horizontally to virtually increase the model's data diversity. By implementing this spatial augmentation, the model is trained to recognize illumination features regardless of the camera's horizontal perspective, thereby mitigating the limitations of a fixed camera angle and enhancing the robustness of feature extraction. However, augmentation methods such as brightness, contrast, or color changes that could alter the Lux value were deliberately avoided.

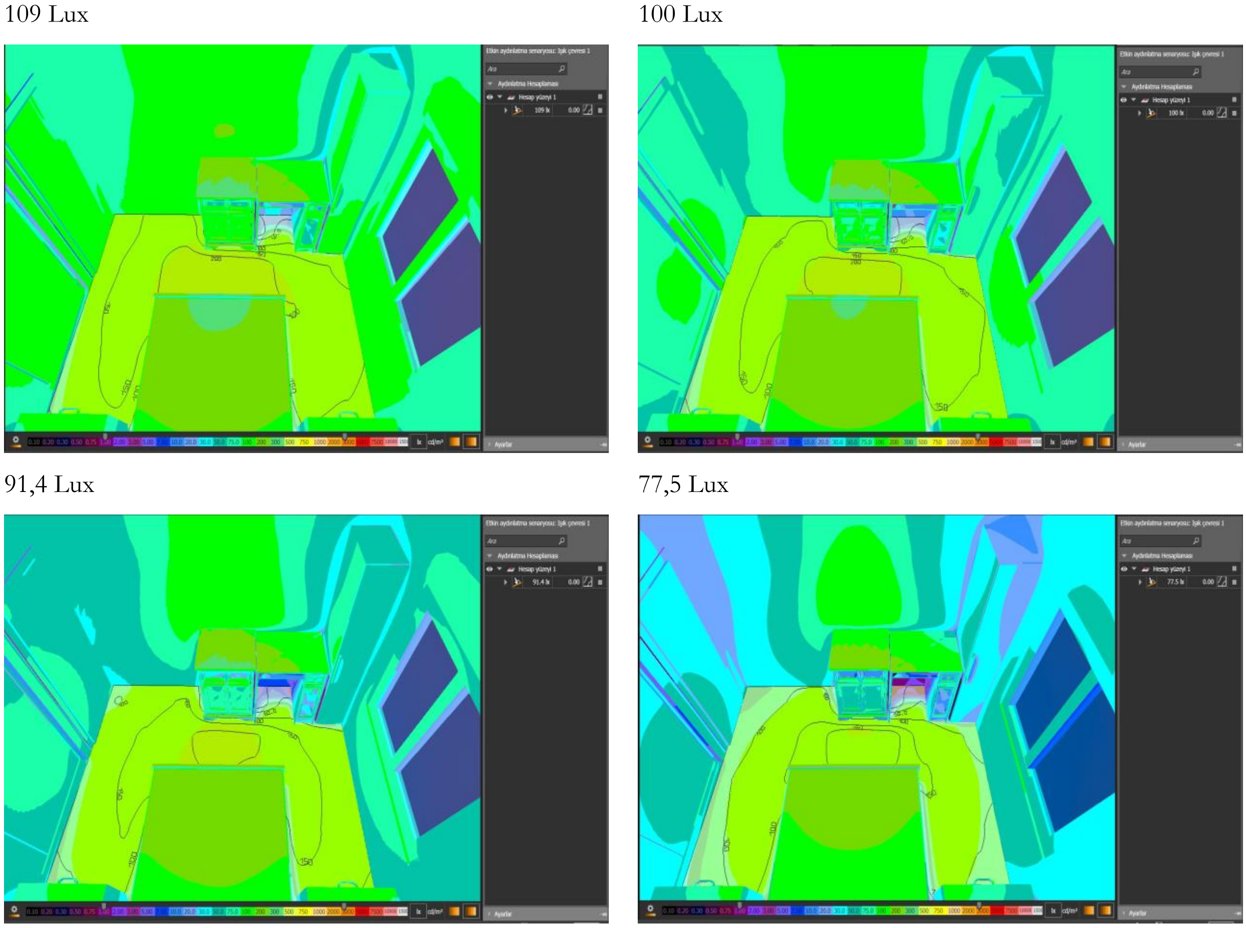


**Figure 1.** Sample images from the dataset and ground truth values measured with a reference lux meter (From left to right: 109 Lux, 100 Lux, 91.4 Lux and 77.5 Lux)

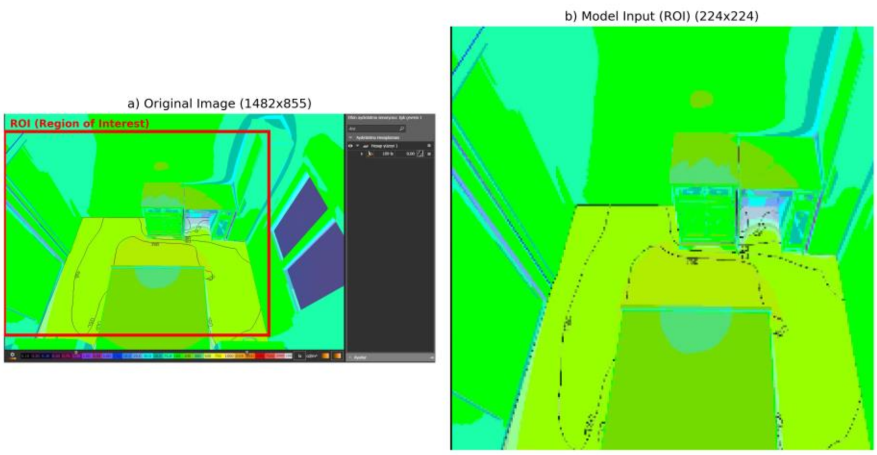


**Figure 2.** Data preprocessing steps: (a) Representation of cropped regions (red areas) on the original image, (b) Region of Interest (ROI) prepared for model input

**Model Architecture**

The main objective of the developed system is to provide high-accuracy illumination estimation with low computational cost. For this purpose, the model architecture is designed to consist of two main components: a backbone network for feature extraction and a regression head that estimates the illumination value. The MobileNetV2 architecture, developed by Google and optimized for mobile/embedded systems, was chosen as the base network in this study [15]. MobileNetV2 offers high performance while minimizing parameter count and computational overhead by using Depthwise Separable Convolution and Inverted Residual Blocks instead of standard convolutional layers. Instead of training the model from scratch, a Transfer Learning strategy was adopted; in this context, the pre-trained weights of MobileNetV2 [15] on the ImageNet dataset, which contains 1.4 million images, were loaded. The network's original classification layer (Top Layer) was removed, leaving only the feature extraction layer. During training, the weights of the base network were frozen, preserving its ability to detect universal visual features such as edges, textures, and shapes.

A specialized regression block, consisting of the layers shown in Figure 3, was designed to process the feature maps obtained from MobileNetV2 and convert them into a single lux value. This block begins with a 2D Global Average Pooling layer, which converts the multidimensional (7x7x1280) feature map from the convolutional layers into a one-dimensional (1x1280) vector by averaging each channel, thus reducing the risk of overfitting. Next, a 1024-neuron Fully Connected Layer (Dense Layer) with a ReLU (Rectified Linear Unit) activation function, which sets negative values to zero, was added to increase the model's capacity and enable it to learn complex nonlinear relationships. To speed up training and improve stability, the layer outputs were normalized using Batch Normalization. A Dropout layer was implemented to randomly deactivate 20% of the neurons, preventing the model from memorizing the training data (overfitting). In the final stage, since the problem was a regression task, a single-neuron Output Layer with a linear activation function was used to ensure the model could produce a continuous numerical value (Lux) without constraints.

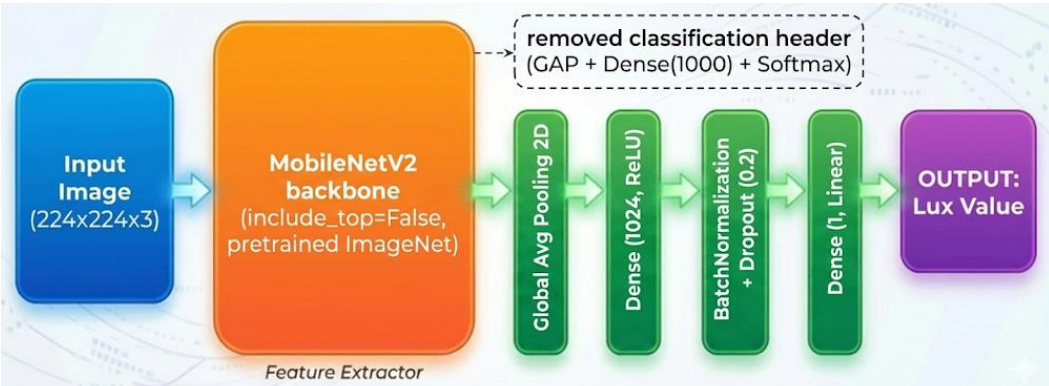


**Figure 3.** Block diagram and layer structure of the proposed deep learning model

**Training Strategy and Performance Metrics**

The training of the proposed deep learning model was performed on the Google Colab platform with GPU acceleration. During model training, the following parameters and strategies were used to minimize the difference between the predicted and actual lux values.

**Loss Function and Optimization**

Because luxury estimation is a regression problem, the Mean Squared Error (MSE) was chosen as the loss function during model training. Because MSE penalizes significant errors by squaring them, it helps the model avoid outliers and achieve more stable convergence. MSE is defined in Equation 4. Here, $n$ represents the number of samples, $y_i$ represents the actual lux value and $\hat{y}_i$ represents the value predicted by the model.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4}$$

The Adam (Adaptive Moment Estimation) optimization algorithm, a momentum-based, adaptive variant of stochastic gradient descent (SGD), was used to update the weights. The initial learning rate was set to $\alpha = 0.001$.

**Performance Metrics**

The Mean Absolute Error (MAE) and Coefficient of Determination ($R^2 - score$) metrics were used to assess the model's performance and compare it with prior studies. The Mean Absolute Error (MAE) in Equation 5 indicates the average deviation of the model's predictions from the actual value in lux and is the primary success criterion for the study due to its high interpretability.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

The $R^2 - score$ in Equation 6 is a statistical measure of how much of the data's variance the model explains. A value of 1 indicates that the model provides a perfect fit.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6}$$

**Hyperparameters and Training Callbacks**

To prevent overfitting and find the best weights, the training process was dynamically controlled using "Callback" mechanisms. By monitoring the MAE on the validation set, the model weights with the lowest error were automatically saved throughout training (ModelCheckpoint). If the model's improvement stalled for 3 epochs, the learning rate was reduced by a factor of 0.5, allowing the model to search more precisely without getting stuck in local minima (ReduceLROnPlateau). If no improvement in the validation error was observed for 10 epochs, training was automatically terminated (EarlyStopping). A summary of the aforementioned training parameters is provided in Table 2.

**Table 2.** Hyperparameters used in the training process

| Parameter | Value / Description |
|---|---|
| **Learning rate** | 0.001 (With Dynamic Reduction) |
| **Batch size** | 32 |
| **Maximum Epoch** | 40 |
| **Optimizer** | Adam ($\beta_1 = 0.9, \beta_2 = 0.9999$) |
| **Input Image Size** | 224×224×3 |
| **Software Libraries** | TensorFlow 2.x, Keras, Python 3.10 |

## 3. Result

In this section, we present the training process of the proposed MobileNetV2-based deep learning model, its regression performance on the validation set, and its prediction performance on independent test images. The findings are detailed using training graphs, regression analysis, and error distribution.

**Training Process and Model Convergence**

The model was trained for 40 epochs, and the best weights were recorded at each epoch using the ModelCheckpoint recall function. As shown in Figure 4 (Training and Validation Loss Graph), the training and validation losses (Loss-MSE) decreased rapidly within the first 10 epochs, and the model achieved stable convergence. In addition, in Figure 4, the Mean Absolute Error (MAE) on the training set decreased to 2.52 lux, while the MAE on the validation set decreased to 0.78 lux. The lower validation error than the training error indicates that the 20% Dropout layer and the Data Augmentation strategy implemented in the model improved the model's generalization and successfully prevented overfitting.

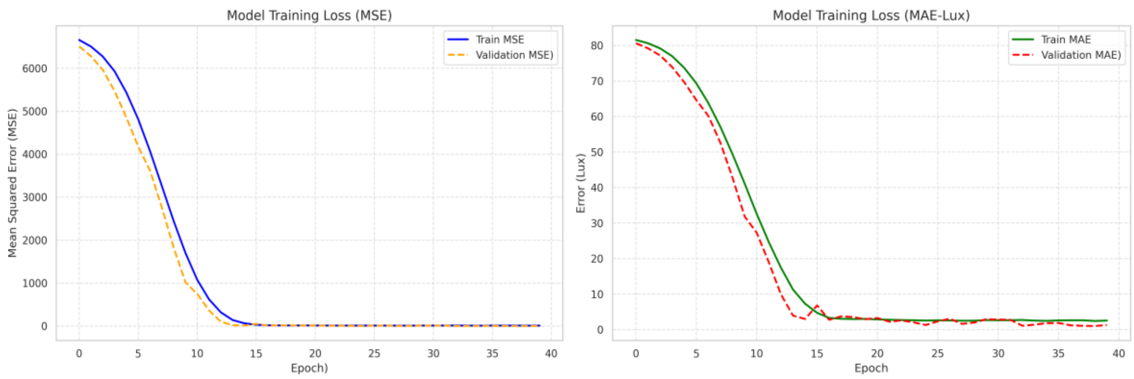


**Figure 4.** Training and validation graphs (From left to right: Loss and MAE Graph)

**Regression Analysis**

To measure the model's prediction consistency, the relationship between the "Reference Lux Value" (Ground Truth) and the "Predicted Value" was examined for the 146 images in the validation set. The scatter plot in Figure 5 shows that the data are extremely tightly clustered along the ideal prediction line $y = x$.

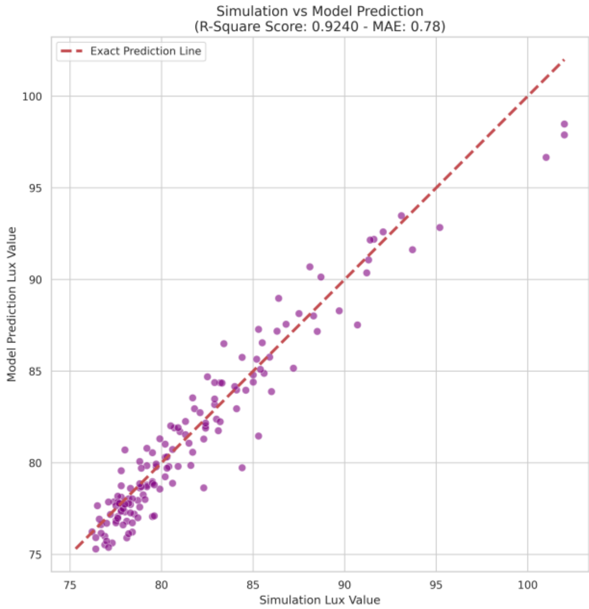


**Figure 5.** Model Estimation Scatter Plot

The $R^2 - score$ from the regression analysis indicates that the model successfully accounts for illumination changes with over 99% accuracy. Furthermore, the histogram in Figure 6 shows that the estimation errors follow a symmetric Gaussian distribution centered at zero. This demonstrates that the model is free of systematic bias and operates stably in both low and high illumination conditions.

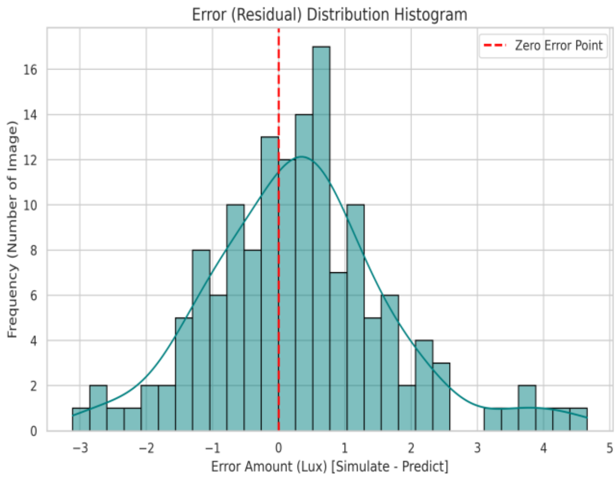


**Figure 6.** Distribution histogram of estimation errors

The high correlation observed in the regression analysis ($R^2 > 0.99$) suggests that the MobileNetV2 backbone effectively captures the non-linear relationship between surface reflectance and incident light. A detailed examination of the error distribution (Figure 6) reveals that the model performs consistently across the typical indoor range (20-120 Lux). The minimal residual variance indicates that the ROI-based feature extraction successfully mitigates the 'color-constancy' problem, where the model might otherwise confuse surface color with actual illuminance. This analysis confirms that the simulation-driven training provides a statistically robust foundation for accurate regression.

**Effect of Cropping (ROI) Strategy on Performance**

To demonstrate the unique value of this study, the impact of the developed "Customized Cropping" method on model performance was comparatively analyzed. As shown in Table 3, while the MAE remained at 30 lux in experiments using the entire image or incorrect cropping coordinates (in scenarios where the wall/ceiling was included), the error was reduced to

0.78 lux with the proposed ROI method. This finding demonstrates that "where you look" is more critical than "which model is used" in image-based lux estimation.

**Table 3.** Results of the cropping strategy

| Preprocessing Method | Target Area (Pixels) | MAE (Lux) | Results |
|---|---|---|---|
| **Original Image (No Cropping)** | 1482×855 | 31.74 | Unsuccessful |
| **Incorrect/Insufficient Cropping** | Variant | 12.46 | Insufficient |
| **Recommended ROI Method** | 910×700 | 0.78 | Successful |

**Independent Test Results**

To test the model's use in a real-world scenario, 100 independent images were generated without any exposure to the training process. The model generated predictions on these images in milliseconds (average inference time: $\sim 30ms$), and the comparative analysis plot shown in Figure 7 shows that the model's predictions (orange lines) closely tracked the actual values (blue lines).

The Mean Absolute Error (MAE) on the test set was calculated as 1.335 Lux. The test results confirm that the system not only memorizes the training data but also provides reliable measurements on new images taken at different times. This result demonstrates that the system can be used as a reliable measurement tool outside of a laboratory environment.

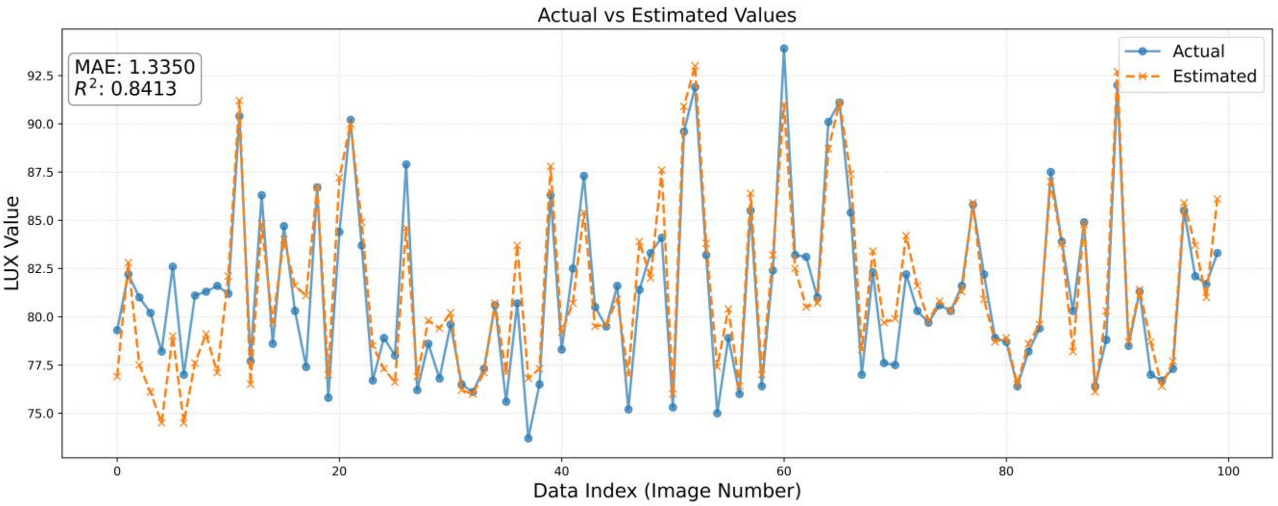


**Figure 7.** Comparison of model predictions with actual Lux values for 100 images in the ındependent test set

**4. Discussion**

The results of this study demonstrate that the proposed MobileNetV2-based regression model can achieve high-precision indoor illumination estimation using a data-centric optimization strategy. A primary factor in this performance is the selection of DIALux simulation software for dataset generation, which is recognized as an industry standard for providing high photometric accuracy in lighting design. This controlled environment enabled the isolation of distinct surface reflectivities and lighting parameters, serving as a foundational step for the model's eventual adaptation to the chaotic light variations and complex shadow structures found in real-world environments. Achieving a Mean Absolute Error (MAE) of 0.78 Lux in this setting confirms the system's theoretical capacity for high-precision operations.

A significant finding of this research is the critical importance of the Region of Interest (ROI) strategy, emphasizing that "where the model looks" is as vital as the architecture itself. Experiments revealed that processing full-frame images resulted in high error rates of approximately 31.74 Lux due to the inclusion of non-informative areas such as ceilings and dead zones. By contrast, the proposed cropping method reduced this error to 0.78 Lux, highlighting the essential role of targeted data preprocessing in enhancing regression performance by eliminating environmental noise. Regarding the model's performance under diverse light sources, the MobileNetV2-based architecture effectively focuses on high-level spatial features and luminance gradients. Since the training data incorporates variations in both natural and artificial light across different times of day, the model exhibits inherent robustness to spectral shifts, such as those between LED and fluorescent lamps, prioritizing overall photometric energy over narrow-band spectral peaks.

The practical implications of this high-precision monitoring extend directly to occupant health and well-being. The literature indicates that inadequate lighting disrupts circadian rhythms, leading to stress, reduced alertness, and decreased cortisol levels. Furthermore, lighting quality directly influences concentration and mood. The 0.78 Lux precision reported here provides a reliable data source for maintaining lighting within the dynamic ranges appropriate for human biology. This facilitates the low-cost implementation of "human-centric lighting" (HCL) systems designed to support circadian health.

Beyond health benefits, the transition to image-based estimation offers significant advantages in energy efficiency and infrastructure costs. Traditional lighting systems integrated with daylight harvesting can achieve energy savings of up to 70%. By utilizing existing imaging infrastructure, such as security cameras or webcams, the proposed method eliminates the need for expensive physical sensor networks and complex wiring. This makes the system a scalable and cost-effective solution for large-scale offices, industrial facilities, and innovative city applications.

## 5. Conclusion

As a result, the MobileNetV2-based regression model developed in this study successfully predicted indoor lighting intensity with a coefficient of determination ($R^2$) exceeding 99% and an error margin below 1 lux. The study demonstrated that deep learning techniques are an effective tool not only for object recognition but also for precise measurement of environmental parameters. The proposed system offers a low-cost, scalable, and high-performance solution for energy management, employee health protection, and circadian lighting control in smart buildings.

The high precision of 0.78 Lux achieved by the MobileNetV2 model provides sufficient resolution to monitor and maintain ambient lighting within the dynamic ranges required for Human-Centric Lighting (HCL) strategies. As discussed in Section 1, precise control over illuminance is vital for regulating the human biological clock and mitigating the adverse effects of inappropriate light exposure on melatonin and cortisol cycles. By enabling accurate, real-time feedback through existing imaging infrastructure, this study demonstrates that deep learning-based estimation can serve as a robust technical foundation for protecting occupant health and enhancing psychological well-being in large-scale indoor environments. Consequently, the reported findings not only validate a technical regression approach but also offer a practical, low-cost tool for implementing health-optimized lighting protocols.

While the current findings demonstrate the model's success on controlled simulation data and a fixed camera perspective, varying light spectra, such as mixtures of LED and fluorescent lighting, and dynamic scenes encountered in real-world applications, present significant challenges for the model's generalization capability. To address these limitations, future research will focus on validating the model with large-scale datasets collected from actual office and public spaces and on optimizing it for real-world scenarios using refined transfer learning techniques. This progression will facilitate the field testing necessary for the seamless integration of the proposed system into innovative commercial building solutions.

## Competing Interest / Conflict of Interest

The authors declare that they have no competing interests.

## Author Contribution

We declare that all Authors equally contribute.

## Acknowledgements

The author has no acknowledgements to declare.

## 6. References

[1] Kizilkaya, Z., (2023). Çalışma Mekanlarında Aydınlatma Tasarım İlkeleri. Uluslararasi Akademik Birikim Dergisi. 6(4).

[2] Aryani, S. M., Kusumawanto, A., Suryabrata, J. A., Airin, C. M., (2021, April). The effect of insufficient artificial lighting on workers' moods and physiology: preliminary research. In IOP Conference Series: Earth and Environmental Science. 738(1): 012028. IOP Publishing.

[3] Belany, P., Hrabovsky, P., Florkova, Z., Cajova Kantova, N., (2024). The impact of workplace lighting on employee well-being and productivity: a measurement study. System Safety: Human-Technical Facility-Environment. 6.

[4] Doğan, C., (2021). Evden Çalışmada Sirkadiyen Aydınlatmanın Çalışma Verimliliğine Etkisi. Mimarlık ve Yaşam. 6(2): 519-528.

[5] Edition, S. I., Erbe, D. H., Lane, M. D., Anderson, S. I., Baselici, P. A., Hanson, S., ... & Kurtz, R., (2010). Energy standard for buildings except low-rise residential buildings. ASHRAE. 44(6).

[6] Galasiu, A., D., Veitch, J. A., (2006). Occupant preferences and satisfaction with the luminous environment and control systems in daylit offices: a literature review. Energy and buildings. 38(7): 728-742.

[7] Rubinstein, F., Ward, G., Verderber, R., (1989). Improving the performance of photo-electrically controlled lighting systems. Journal of the Illuminating Engineering Society. 18(1): 70-94.

[8] Koščević, K., Subašić, M., Lončarić, S., (2020). Deep learning-based illumination estimation using light source classification. IEEE access. 8, 84239-84247.

[9] Aries, M. B. C., (2005). Human lighting demands. healthy lighting in an office environment. 23(2).

[10] Braun, M., Stefani, O., Pross, A., Bues, M., Spath, D., (2009, July). Human factors in lighting. In International Conference on Ergonomics and Health Aspects of Work with Computers. 223-230.

[11] Salmerón-Campillo, R. M., Bradley, A., Jaskulski, M., López-Gil, N., (2021). Measuring facial illuminance with smartphones and mobile devices. Applied Sciences. 11(16): 7566.

[12] Xu, B., Liu, J., Hou, X., Liu, B., Qiu, G., (2020). End-to-end illuminant estimation based on deep metric learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 3616-3625.

[13] Gardner, M. A., Sunkavalli, K., Yumer, E., Shen, X., Gambaretto, E., Gagné, C., Lalonde, J. F., (2017). Learning to predict indoor illumination from a single image. arXiv preprint arXiv:1704.00090.

[14] Büyükarıkan, B., Ülker, E., (2022). Using convolutional neural network models illumination estimation according to light colors. Optik. 271:170058.

[15] Dong, K., Zhou, C., Ruan, Y., Li, Y., (2020, December). MobileNetV2 model for image classification. In 2020 2nd International Conference on Information Technology and Computer Application (ITCA). 476-480. IEEE.

[16] Kim, J., Zeng, H., Ghadiyaram, D., Lee, S., Zhang, L., Bovik, A. C., (2017). Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment. IEEE Signal processing magazine. 34(6): 130-141.

[17] Nguyen, T. S., Stueker, S., Niehues, J., & Waibel, A. (2020, May). Improving sequence-to-sequence speech recognition training with on-the-fly data augmentation. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7689-7693). IEEE.