Journal of Autonomous Intelligence (2024) Volume 7 Issue 3
doi: 10.32629/jai.v7i3.562

REVIEW ARTICLE

Street view images blur detection

Masoud Moeini*, Ehsan Yaghoubi, Simone Frintrop

Department of Informatics, Universität Hamburg, 20148 Hamburg, Germany

*** Corresponding author:** Masoud Moeini, masod.atc@gmail.com

ABSTRACT

Blurred regions in images can hinder visual analysis and have a notable impact on applications such as navigation systems and virtual tours. Many existing approaches in the literature assume the presence of blurred regions in an image and process the entire image, even when no blurred regions are actually present. This approach leads to unnecessary computational overhead, resulting in inefficiency and resource consumption. In this paper, we introduce a Street-view images Blur Detection Network (SBDNet), consisting of two interconnected subnetworks: the Classifier network and the Identifier network. The Classifier network is responsible for categorizing street-view images as either blurred or not blurred. Once the Classifier network determines that an image is blurred, the Identifier network is then activated to estimate the specific areas that are blurred within the image. High-level semantic features from the Classifier network are used to construct the blur map estimation in the Identifier network, when necessary. The algorithm was trained and evaluated using the Street-View Blur Images dataset (SVBI) and three publicly available blur detection datasets: CUHK, DUT, SZU-BD. Our quantitative and qualitative results demonstrate that SBDNet competes with state of the arts in blur map estimation.

Keywords: street-view; deep learning; blur detection; blur classification

ARTICLE INFO

Received: 30 March 2023
Accepted: 18 June 2023
Available online: 19 January 2024

1. Introduction

Blur detection in images plays a critical role in various domains of image processing and computer vision, such as photography, virtual reality, and medical imaging. The presence of a blur detection system could be essential as a preprocessing step for many applications where assessing and enhancing image quality are important. Therefore, the computational efficiency of the blur detection system could be equally vital as its performance (see **Figure 1**). However, current approaches often assume the existence of blurred regions in an image, resulting in unnecessary computational burden when processing images without any actual blurring. This inefficiency restricts the real-time deployment of blur detection algorithms and poses challenges in integrating them into resource-constrained systems. Additionally, the performance of these methods may fall short when faced with complex image content or varying degrees of blur. Therefore, it is imperative to explore novel approaches that can effectively address these challenges and significantly improve the accuracy and reliability of blur detection.

Many studies have been conducted on image blur detection and segmentation of blurred regions based on gradient and frequency methods (handcrafted feature extraction methods)[1–7]. In some particular cases, hand-crafted methods are unable to properly segment blurred regions in images, while using deep Convolutional Neural

Networks (CNNs) could assist in solving the issue of using hand-crafted approaches[8]. The main challenges of the deep learning based blur detection methods are: coarse object boundaries for non-blurred regions, background clutter, and misclassification of low contrast regions[9]. The study of Zhang et al.[10] is particularly similar to our research. The authors of this study present a model that predicts a blur map and assesses the image quality as good, medium, or bad. However, the model incorporates the predicted blur map to assign the image to a desirability class, which results in computational intensity for images without blur.

This paper introduces a deep learning approach to address the challenges of blur detection and classification for Google street-view images. Our main idea is to develop a single deep learning algorithm that efficiently performs both tasks: blur classification, and blur map estimation. The proposed Street-view images Blur Detection Network (SBDNet) consists of two interconnected sub-networks: the Classifier network and the Identifier network. The Classifier network, utilizing ResNet50 or GoogLeNet as its backbone, performs blur classification. The Identifier network, comprising a hierarchical architecture with three subnetworks, aims to detect and identify the blurred regions by generating a blur map estimation. To train and evaluate SBDNet, we created a new dataset called Street-View Blur Images (SVBI), consisting of high-resolution street view images with blurred regions and corresponding ground truths. In addtion, we evaluated the proposed method on CUHK, DUT, SZU-BD datasets. The SBDNet architecture demonstrates competetive performance with a potential reduced complexity. Our main contributions are as follows:

1) We introduce a deep learning approach, SBDNet, for blur detection and classification in street-view images. SBDNet consists of two interconnected sub-networks: the Classifier network and the Identifier network.

2) We create a dataset called Street-View Blur Images (SVBI), consisting of high-resolution street-view images with blurred regions and corresponding ground truths. The SVBI dataset is used for training and evaluating SBDNet.

3) We evaluate the performance of SBDNet using three publicly available blur detection datasets: CUHK[1], DUT[8], and SZU-BD. The quantitative and qualitative results demonstrate that SBDNet achieves competitive performance compared to state-of-the-art methods for blur map estimation.

4) We provide the code implementation of SBDNet, making it accessible for further research and development[11].

Figure 1. Google street-view images blur detection and classification problem.

2. Related work

Distinguishing between in-focus and blurred regions is known as blur detection. Images may contain blurry regions for a variety of reasons; the most common types of blurry images are defocused blur and motion blur. Defocused blur images are statically taken to highlight targeted objects or infocus regions of the images. While motion blurred images are captured during motion and the captured scene is moving.

Recent advancements in deep learning methods[8–22] have significantly contributed to the domain of blur detection. The challenge in blur detection arises from the fact that blurred regions are not necessarily distinguished by sharp edges. In fact, the image pixels in blurred regions do not provide enough information for effective blur detection. In short, accurately distinguishing between in-focus and blurred regions can be challenging due to the complex nature of blur and the presence of various types of blurring, such as motion blur, defocus blur, and lens aberrations. This challenge has led to the development of various approaches.

For instance, Zhao et al.[9] represented Hierarchical Edge-aware Network (HEANet) method using ResNet50[23] as a backbone (pre-trained on ImageNet) to extract accurate boundary information for infocus regions and refining extracted features in a hierarchical progressive process. Zhao et al.[15] designed a hybrid CNN transformer architecture to address the problem of detecting low contrast regions and providing distinct infocus object boundaries.

Jonna et al.[13] represented a computationally efficient method for DBD using a synthetic generative adversarial network consisting of two generator networks (teacher network and student network) to generate a blur map from an input image and 3 discriminator networks to train adversarially generator networks. The proposed teacher generator network comprises three sequential components, at the first step it employs SE-ResNeXt-101[24] as a encoder network to extract features from an input image, in the next step, the extracted features from SE-ResNeXt-101[24] is refined by utilizing the DenseASPP[25] and self-attention modules in parallel and concatenating high-level features. The third component of the teacher generator network is a decoder network and produces the final blur map from exploited high-level features of the previous component. The generated blur map from the teacher generator is adversarially examined with ground-truth map by applying the first discriminator network. The student generator network[13] has encoder-decoder network architecture, using EicientNetB3[26] as an encoder network and 5 convolutional transposed layers as a decoder network to produce blur map.

Lin et al.[16] presented a residual attention learning approach to address DBD using feature information from a combination of infocus and defocus regions. Zhao et al.[14] designed a DBD network using a generative adversarial network that can be trained by eliminating manually pixel-level annotation. To address DBD problem, Guo et al.[17] represented an encoder-decoder in a heterogeneous architecture using nested pyramid pooling modules into encoders (5 encoder blocks) and nested u-shaped network in decoders (6 decoder blocks) combined with nested attention blocks to extract and fuse features with different scales productively. Guo et al.[17] method demonstrated noticeable evaluation results on CUHK[1], DUT[8] datasets.

Tang et al.[18] represented DefusionNet using VGG[27] network as a backbone and extracting high-level features from deeper-layers and low-level features form shallow layers by up-sampling and fusion technique. Tang et al.[18] method recurrently (in 3 steps) refines up-sampled and fused features to produce final out-of-focus map from an input image. Following[18] method, Tang et al.[19] proposed a deep neural network approach which recurrently refines up-sampled multi-scale residual features extracted from different layers of the basic pre-trained ResNeXt[28] network. Tang et al.[19] method achieved higher evaluation results on CUHK[1] and DUT[8] datasets compared to DefusionNet[18]. Sun et al.[12] presented deep pyramid network using a distinction enhanced block to combine high-level semantic features information with low-level information to identify concurrently blurred areas and clear regions. Sun et al.[12] created new blur detection dataset (SZU-BD) which

is accessible for public alongside with other two famous blur detection datasets: CUHK[1] and DUT[8] as benchmarks for blur detection. Zhao et al.[20] developed a cross-ensemble network. In this network instead of a single defocus blur detector network Zhao et al.[20] broke down the problem into a series of smaller parallel DBD networks that refine each other's errors and improve the estimated in-focus region. Zhao et al.[8] designed an end-to-end multi-stream bottom-top-bottom fully convolutional network (BTBNet) for DBD. Zhao et al.[8] employed modified VGG16[27] to extract and to combine multi-scales features from different levels.

Zhang et al.[10] represented ABC-FuseNet network for image blur understanding. The proposed ABC-FuseNet[10] network can automatically classify an input image whether it includes blurred areas or not, and if blurred regions exists, ABC-FuseNet can accurately estimate the blur map and specify the types of input blurred image. For input blurred images, the proposed ABC-FuseNet[10] produces jointly blur map estimation and categorization labels. Zhang et al.[10] created the SmartBlur dataset consisting of 1000 blurred images with corresponding ground-truth maps and annotated blur categories, for efficiently training and evaluating the proposed ABC-FuesNet. Zhang et al.[10] exclusively evaluated ABC-FuesNet on CUHK[1] dataset for blur map estimation.

Ma et al.[21] trained deep neural networks using VGG16[27] as a backbone network for DBD. Ma et al.[21] algorithm extracts high-level features from the last convolutional layer of adjusted VGG16[27] network and then the network produces the final blur map by up-sampling and fusing the extracted high-level semantic information. Kim et al.[22] presented a deep neural network with an encoder-decoder architecture to detect blurred images. Kim et al.[22] network uses residual skip-connections and multi-scale reconstruction loss functions to obtain high-level information and low-level structural features.

In conclusion, recent works have made significant progress in deep learning-based blur detection methods by employing various network architectures and techniques. However, challenges such as accurately distinguishing between in-focus and blurred regions, handling noisy images and artifacts, addressing complex image content and diverse scene conditions, ensuring computational efficiency, and creating comprehensive datasets remain areas of active research and exploration.

Among the above-mentioned challenges in blur detection, this work focuses on addressing the computational efficiency aspect. To achieve this, we propose a framework that incorporates a two-step process. Firstly, we classify the input image into either a blur or no blur class. This initial classification step allows us to quickly determine if further processing is required for estimating the blur map. By efficiently identifying non-blurred images, we can avoid unnecessary computations, thereby improving the overall efficiency of the algorithm. Once an image is classified as blur, we proceed to the second step, where we perform more detailed processing to estimate the blur map. In the next sections, we provide a detailed explanation of the proposed method and our collected dataset.

3. The street view blur images dataset

In this paper, we focus on the task of blur detection in street-view images, with applications in navigation systems and virtual tours. While publicly available datasets offer diverse images with various blur scenarios, they do not adequately cover the specific context of Google street-view images. Therefore, our first step is to create the Street-View Blur Images (SVBI) dataset. By curating high-resolution street-view images with deliberately blurred regions and providing corresponding ground truths, SVBI allows to supervisly train and evaluate our proposed Street-view images Blur Detection Network (SBDNet) comprehensively.

The blur level of images in the proposed SVBI dataset is artificially generated using two different methods, A and B. The dataset is composed of two classes, blur and no blur, each containing 5000 high-quality images from the UCF Google street-view dataset[29]. The blur level of images in the Blur class is created by applying

the GaussianBlur filter to selected regions in the street-view images.

Method A involves using the Haar-cascade object detection algorithm[30,31] to detect objects within the street-view images. The developed a code-agent to identify cars (as the main available object on the street) and place bounding boxes around them. The GaussianBlur filter is then applied to the selected regions within the bounding boxes to produce blurred images. The radius of the GaussianBlur filter determines the noise level (blurriness) added to the image. The code-agent generates corresponding ground-truth annotations associated with the blurred images. All the resulting images, including the original street-view image, the produced blurred image, and the ground-truth, are resized to 224 × 224 × 3.

In cases where the code-agent fails to find deterministic objects inside a street-view image, Method B is automatically utilized. Method B randomly chooses rectangular-shaped regions within the street-view images and applies the GaussianBlur filter to these regions to produce blurred images. Similar to Method A, the code-agent generates corresponding ground-truth annotations for the blurred images. All images generated using Method B are also resized to 224 × 224 × 3.

We selected an blurriens intensity of 15 for Method A and 20 for Method B. These intensities were chosen intuitively and based on experimentation, but can vary depending on the specific dataset and application. We resized the generated images as the resizing of all blurred and non-blurred images to the same dimensions ensures consistency and compatibility for further processing and training of machine learning models.

4. The proposed approach

4.1. SBDNet architecture

The proposed SBDNet aims to classify a given Google street-view image as either clear or blurred. Additionally, if the input image contains blurred regions, SBDNet efficiently detects the blur map for the image at the same resolution.

Figure 2 illustrates the overall architecture of SBDNet, which comprises two main networks: the Classifier network and the Identifier network. The Classifier network utilizes high-level features to classify input images, while the Identifier network estimates the blur map for input blurry images.

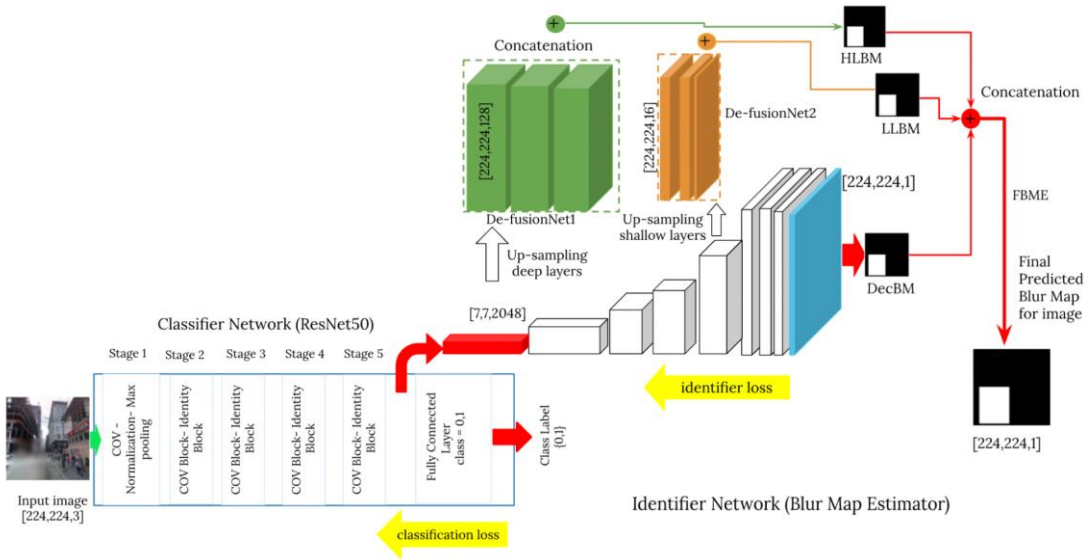

Figure 2. The proposed Street-view images Blur Detection Network (SBDNet) architecture. The SBDNet concurrently predicts a binary classification value (0 or 1) and 224 × 224-pixel gray-scale Final Blur Map Estimation (FBME) for input RGB images with size of 224 × 224 × 3. The FBME is generated by combining the blur map estimations from the Decoder network (DecBM), with up-sampled high-level semantic features (HLBM) and up-sampled low-level features (LLBM). The resulting FBME is a grayscale image with a resolution of 224 × 224 pixels.

In this architecture, the proposed Classifier network employs ResNet50[21]as its backbone. On the other hand, the Identifier network consists of a basic decoder network architecture and two streams of up-sampled blocks extracted from high-level feature maps (deeper layers) and low-level feature maps (shallow layers). These up-sampled blocks maintain the same size as the input image but differ in the number of channels.

The Identifier network initially produces a blur map estimation for input images (DecBM) using the decoder network. Subsequently, it concatenates and fuses the up-sampled blocks from the deeper layers (HLBM) and shallow layers (LLBM). Finally, the network combines all the estimated blur maps from different streams using a convolutional operator to produce the Final Blur Map Estimation (FBME).

4.2 Classifier and blur map estimator

The Classifier network in SBDNet is responsible for classifying input images as either in-focus or blurred. In our study, we demonstrate the overview of SBDNet using ResNet50 (**Figure 2**) and GoogleNet (**Figure 3**) Classifiers. However, it is worth noting that other deep learning models can also be employed. The input images are passed through the convolutional layers, and the feature maps from the final convolutional layer are extracted, capturing the high-level semantic features of the input images. Subsequently, these extracted features are fed into fully connected layers to accomplish the classification task. These layers are trained to learn the mapping of the high-level features to a binary classification value, indicating whether the input image is clear or blurred. The experimental results we present in the following sections demonstrate that utilizing ResNet50 or GoogleNet as the backbone for the Classifier network effectively enables the proposed framework to classify street view images as in-focus or blurred.

Inspired by Jonna et al.[13], our proposed Identifier network uses a decoder network to produce initial blur map estimation. The proposed decoder network fetches high-level semantic features from the last convolutional layer of the classifer network and employs sequences of eight transposed convolutional layers to generate initial blur map estimation. To simplify the explanation, decoder network layers are labeled as Dec_i, i $\in$ {1, ..., 8}, where Dec_1 represents the last convolutional layer of ResNet50 encoder[23] with size of 7 × 7 × 2048. The layers Dec_2 to Dec_8 in the decoder network are transposed convolutional layers with 3 × 3 kernel size, resulting in Decoder network Blur Map (DecBM) estimation with size of 224 × 224 × 1. Layers Dec_1 to Dec_3 contain high-level semantic features and Dec_4 to Dec_6 include low-level features information.

As shown in **Figure 2**, the decoder network generates a blur detection map named DecBM. Inspired by DefusionNet[18], we construct concurrently two subnetworks, shown as green (De-fusionNet1) and orange (De-fusionNet2) in **Figure 2**. The green subnetwork is built by up-sampling of layers Dec_1 to Dec_3, to the dimensions of 224 × 224 × 128. The proposed up-sampling technique uses transposed convolutional operator with 3 × 3 kernel and relu activation function. After up-sampling deeper layers, the up-sampled blocks will be concatenated and fused. The fusion technique is implemented by utilizing convolutional operator with 1 × 1 kernel, stride of 1 pixel and relu activation function. The green subnetwork (De-fusionNet1) produces High-Level Blur Map (HLBM) estimation with size of 224 × 224 × 1, based on high-level semantic features data. Similarly, The orange subnetwork (De-fusionNet2) is made by concatenating the up-sampled Dec_4 to Dec_5 layers with the Dec_6 layer. By applying fusion operator to the concatenated layers, the orange subnetwork produces Low-Level Blur Map (LLBM) estimation with size of 224 × 224 × 1 based on shallow layers information. The generation of HLBM and LLBM can be mathematically formulated as follows:

$$\text{HLBM} = \text{ReLU}(W_l \star Cat(UpDec_1, UpDec_2, UpDec_3) + b_1), \quad (1)$$

$$\text{LLBM} = \text{ReLU}(W_m \star Cat(UpDec_4, UpDec_5, Dec_6) + b_m), \quad (2)$$

where $UpDec_i$, $i \in \{1, 2, 3, 4, 5\}$ denotes up-sampled features from *i*-th layer of decoder network. W_l , W_m and b_1, b_m respectively, represent weights and bias for convolutional operation. *Cat* denotes concatenation

operation. Symbol ⋆ represents the convolutional operator.

The Final Blur Map Estimation (FBME) is then generated by fusing DecBM, HLBM, and LLBM at the end of the process:

$$\text{FBME} = \text{Sigmoid}(W_n \star Cat(DecBM, HLBM, LLBM) + b_n). \quad (3)$$

The aforementioned convolutional operator constructs FBME with dimensions of 224 × 224 × 1 using a 1 × 1 kernel, stride of 1 pixel, and sigmoid activation function. FBME will be supervised in the training phase against the corresponding ground-truth.

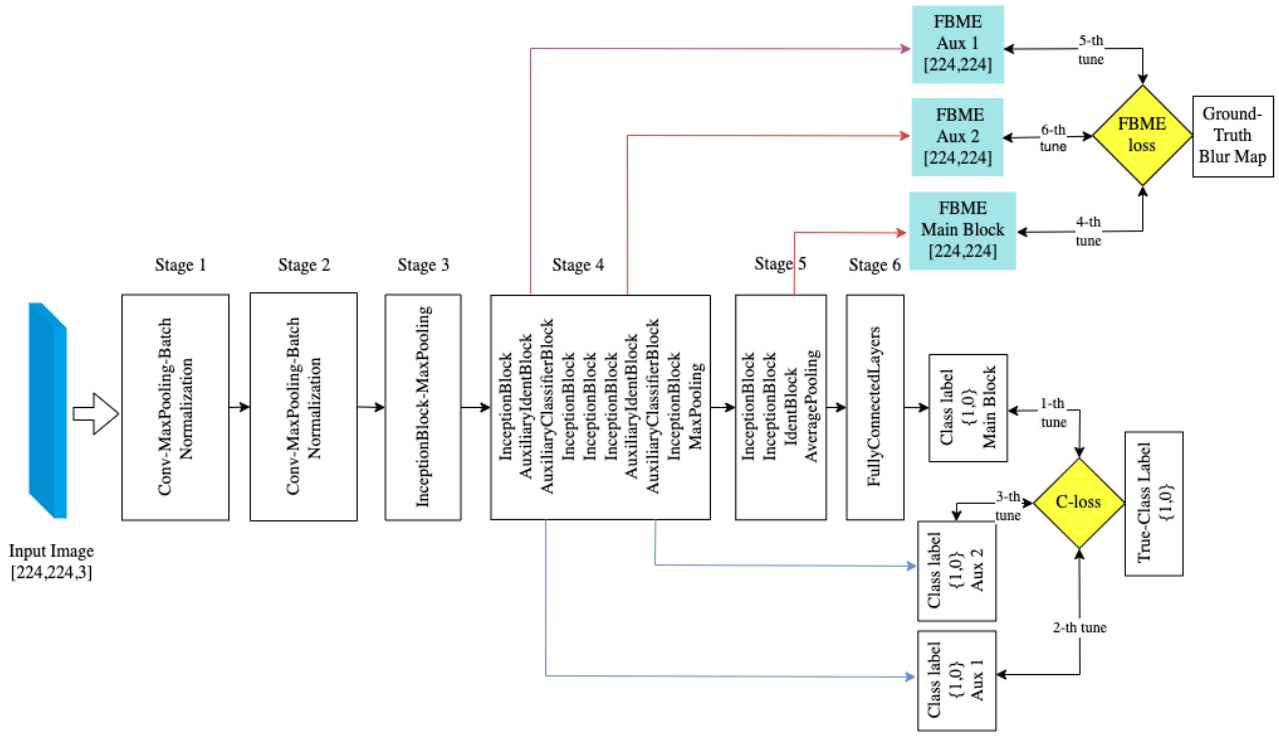

Figure 3. The proposed SBDNet using GoogLeNet[32] as a backbone network for classification and features extraction. Twofully-connected layers called Auxiliary Classifier blocks (AuxiliaryClassifierBlock) in stage 4 collaborate with the main fully-connected layers (stage 6) to predict classification labels for an input image. Two adjusted Identifier networks called Auxiliary Identifier blocks (AuxiliaryIdentBlock) in stage 4, alongside the main Identifier network (IdentBlock) in stage 5, produce FBMEs respectively for an input blurred image. The Classification loss (C-loss) and the Identifier loss (FBME loss) compute classification and blur map estimation errors. The proposed network entirely is tuned by backpropagation of the gradients from the corresponding errors.

4.3. Model training and testing

The proposed SBDNet is trained end-to-end. Throughout the training process, the network learns to predict both the blur map estimation and the classification label for an input image. By the end of training, the network is capable of jointly predicting these two outputs. The SBDNet network employs two loss functions during training: Binary Cross Entropy loss for the classification task (Equation (4)) and pixel-wise Mean Squared Error (MSE) loss for blur map estimation (Equation (5)):

$$Classifier_{Loss} = -(y * log(p) + (1-y) * log(1-p)), \quad (4)$$

$$Identifier_{Loss} = (1/(h*w)) * \sum(y - \hat{y})^2 \quad (5)$$

where in Equation (4), y is the true label (blur or no blur), and p is the predicted probability of the positive class (between 0 and 1). In Equation (5), h and w are the height and width of the maps, y is the true binary map, and $\hat{y}$ is the predicted binary map.

We employed the Adam optimizer with a learning rate of 0.001 for training the model. The training was

conducted over 100 epochs, with a batch size of 20 samples.

We also implemented the proposed SBDNetusing GoogLeNet[32] Classifier network (see **Figure 3**). The proposed architecture is similar to what we did for the ResNet50 encoder. As shown in **Figure 3**, the SBDNet includes 6 stages using 2 Auxiliary Classifier blocks (AuxiliaryClassifierBlock) and 2 Auxiliary Identifier blocks (AuxiliaryIdentBlock), connected to intermediate layers. The proposed auxiliary Classifiers increase the network's performance by efficiently backpropagate gradients of Classification loss to lower stages as well as the proposed auxiliary Identifiers expedite the network's training for blur map estimation. The aforementioned auxiliary Classifiers are two additional small, fully-connected layers that are connected to the network's middle layers and propagate the corresponding gradients related to classification errors back to the middle layers, allowing the network to effectively start training from earlier stages. The proposed 2 auxiliary Identifiers blocks, carry out Identifier network architecture without Dec_1 and Dec_2 layers and produce blur map estimations based on extracted features from the middle layers. At Stage 5, comparable with Classifier network architecture using ResNet50[23], the proposed architecture in **Figure 3** uses the main Identifier network for FBME. In the end, at Stage 6, network uses main fully-connected layers for classification.

As shown in **Figure 3**, all the predicted results from Auxiliary Classifier blocks, Auxiliary Identifier blocks, main Identifier block (Stage 5) and main fully-connected layer (Stage 6), in supervised manner will be optimized against the ground-truths and true class labels during the training process. The proposed SBDNet model displayed in **Figure 3**, contains approximately 19 million trainable parameters. Similar to the training of the SBDNet model using ResNet50[23], to train the SBDNet model using GoogLeNet[32], the related hyperparameters, loss functions, and optimizer algorithm were not changed, and only the learning rate coefficient was set to 0.001. During the training process, for both SBDNet models, 80% of the input images (800) were used for training, and 20% of the input images randomly were selected for validation.

In order to compare the proposed SBDNet with other state-of-the-art blur detection methods, the proposed SBDNet models were trained exclusively for blur map estimation task using images from Shi et al.[1] dataset (CUHK) without any classification components. To re-construct SBDNet models for blur map estimation, in the proposed SBDNet model using ResNet50[23] as a Classifier network, fully-connected layers were removed and the proposed Identifier network was kept in connection to the last convolutional layer of ResNet50[2]. Accordingly, in the proposed SBDNet model using GoogLeNet[32] as backbone, the Auxiliary Classifier blocks and fully connected layers (Stage 6) were disabled.

5. Experiments

5.1. Datasets

We carry out evaluational experiments on four datasets, including SVBI, CUHK[1], DUT[8] and SZU-DB[12] datasets, to assess the accuracy and performance of our proposed SBDNet models using ResNet50[23] and GoogLeNet[32] as a Classifier network.

The SVBI dataset is a self-generated dataset, that was artificially created in this paper (See Sec. 3) for the purpose of street-view images blur detection. The dataset consists of 10,000 images with corresponding ground-truths. Out of these 10,000 images, 5000 are fully clear images, and the rest 5000 are blurred images. The generated blurred images do not belong to the motion blur category, since they are partially blurred for the purpose of hiding some specific regions. They can be considered as defocused blurred images. The proposed SBDNet models were trained with 2000 images and evaluated with 1000 images from the SVBI dataset.

The CUHK[1] dataset is a famous dataset and publicly available for benchmarking and evaluation of blur detection approaches. This dataset includes 704 defocus blur images and 296 motion blur images with

corresponding manually annotated ground-truths. In the first experiment, a number of 800 images, encompassing both out-of-focus blur and motion blur categories, were used to train the proposed SBDNet models. The proposed SBDNet models were evaluated using the remaining 200 images. In the second experiment, SBDNet was trained with 604 defucos images out of 704 total defucos images and evaluated with the remaining 100 images.

The DUT dataset is relatively new and was created by Zhao et al.[8]. It contains 500 out-of-focus blurred images with related manually crafted ground-truths. The proposed SBDNet models were evaluated with 500 images.

The SZU-BD[4] dataset contains 784 blurred images with corresponding ground-truths. Defocus blur makes up 709 of the total 784 blurry images, while motion blur makes up the remaining 75. The proposed SBDNet models were tested on a total of 784 images, including defocused and motion-blurred images.

5.2. Evaluation metrics

To quantitatively assess our proposed SBDNet's performance and accuracy in relation to two key tasks: 1-Street-view images blur detection and classification, 2-Comparison with the most recent state-of-the-art blur detection approaches; two sets of evaluation metrics were implemented. Almost all well-known algorithms for blur detection, semantic segmentation, and classification employ these evaluation metrics. To assess the proposed SBDNet with relation to the task 1: accuracy and F1-score metrics were utilized to evaluate the reliability of SBDNet's prediction results for the classification task. Average Intersection over Union (IoU) score, Average F-measure score, and Mean Absolute Error (MAE), metrics were applied to measure the accuracy of the proposed SBDNet predictions outcomes for blur map estimation. Regarding Task 2, Mean Absolute Error (MAE) and Average F-measure score were used to evaluate SBDNet performance and robustness.

Average IoU: intersection over Union (IoU) is a quantitative metric to assess how much the predicted blur map and the actual ground truth overlap. IoU score measures the common pixels percentage among ground-truth and FBME. The average IoU shown in Equation (6), represents the mean value of calculated IoUs for all image from the evaluation-set.

$$Average\ IoU\ = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{I_i^{Groun\ Truth} \cap I_i^{FBME}}{I_i^{Groun\ Truth} \cup I_i^{FBME}}\right), \tag{6}$$

where, taking into account the corresponding index of the images from the evaluation-set as $i \in \{1, .., n\}$, $I_i^{Groun\ Truth}$ and I_i^{FMBE} represent ground-truth and FBME for *i*-th image of evaluation set.

F-measure: another popular metric to determine how accurate a predicted blur map is in relation to the associated ground-truth is the F-measure score. When compared to a ground-truth, the accuracy of each pixel from the predicted blur map is represented at the pixel-level by the F-measure score (Equation (7)). To produce the best F-measure score in the proposed experiments, the coefficient β^2 was given a value from {0.2, 0.3}. Equation (7) demonstrates how to compute F-measure scores at the pixel-level for each individual pixel of the estimated blur maps that correspond to the input images. For each input image, the mean F-measure score is computed for all of the pixels in the estimated blur map. Considering the entire images of the evaluation set, Average F-measure score equals to the average value for all FBME Fmeasure scores.

$$F_{measure} = \frac{(1+\beta^2)\times precision\times recall}{(\beta^2+precision)\times recall}. \tag{7}$$

5.3. Evaluations on SVBI dataset

Experiment setting: in this experiment, our proposed SBDNet models using ResNet50[23] and

GoogLeNet[32] as a Classifier network, were trained end-to-end with 2000 images from the SVBI dataset and tested with 1000 images from the SVBI dataset. To assess the accuracy and robustness of SBDNet, evaluation images were chosen randomly with corresponding annotated pixel-level ground-truths and classification labels. **Table 1** displays the details of the SVBI dataset images used in training and evaluation of the proposed SBDNet model. The proposed model was initially trained and validated using 2000 input images including blur and no blur types. The model was evaluated using 1000 images.

Table 1. The distribution of the SVBI dataset for training and testing of the SBDNet model.

Class	Blur	No blur	Number of samples (images)
Training & validation	1014	986	2000
Evaluation	487	513	1000
Total	1501	1499	3000

Experiment results on Blur Classification: the evaluation results of the proposed SBDNet model using images from the SVBI dataset with respect to the classification task displayed in **Table 2**. The obtained results illustrated in terms of Precision ↑, Recall ↑, F1-score ↑ and Accuracy ↑ metrics. The arrow ↑ stands for higher value is better (The ideal value for Precision, Recall, F1-score and Accuracy is 1.0).

Table 2. The quantitative evaluation results of the SBDNet for classification using 1000 images from the SVBI dataset in terms of Precision ↑, Recall ↑, F1-score ↑ and Accuracy ↑ metrics.

Class	Precision	Recall	F1-score	Number of samples (images)
Blur	0.97	0.95	0.96	487
No blur	0.96	0.97	0.97	513
Accuracy	0.96			1000

Experiment results on blur map estimation: three metrics of average IoU, average F-measure, and average MAE were employed to assess the performance and accuracy of the proposed SBDNet model, for blur map estimation.

Table 3 displays the quantitative evaluation of SBDNet for blur map estimation, using 1000 test images from SVBI dataset in terms of average IoU ↑, average F-measure ↑ and average MAE ↓ metrics. The arrow ↑ depicts the higher value is better and the arrow ↓ denotes the lower value is better. The ideal values for average IoU, average F-measure and average MAE metrics are 1.0, 1.0 and 0.0 respectively.

Table 3. The quantitative evaluation results of SBDNet for blur map estimation, using 1000 images from SVBI dataset in terms of IoU ↑, F-measure ↑ and MAE ↓ metrics.

Metric	SBDNet
IoU	0.933
F-measure	0.938
MAE	0.106

Figure 4 shows the SBDNet prediction results for 4 evaluation images from the SVBI dataset, including predicted classification label, FBME and Keras Class Activation Map (CAM)[33] against the corresponding true classification labels and ground-truth blur maps. The last row in **Figure 4,** represents the CAM visualization for input images, and it shows which regions of the input images were targeted by SBDNet for classification. The first, third and last columns in **Figure 4** show examples with blurred regions, which model correctly predicted classification labels (label 0 denoting class blur) and FBMEs for this input images. The corresponding CAMs for this examples display the focal attention point of the network around the blurred

regions with yellow mass. The second column in **Figure 4** shows an example image with no blurred regions and model correctly assigned label 1 (no blur) to this example. The predicted FBME for this image also includes no white color areas which means there are no blurred regions. The related CAM for this input image shows no specific attention region with intensive yellow mass that represents network looks at the whole regions to assign label 1 (no blur) to this example.

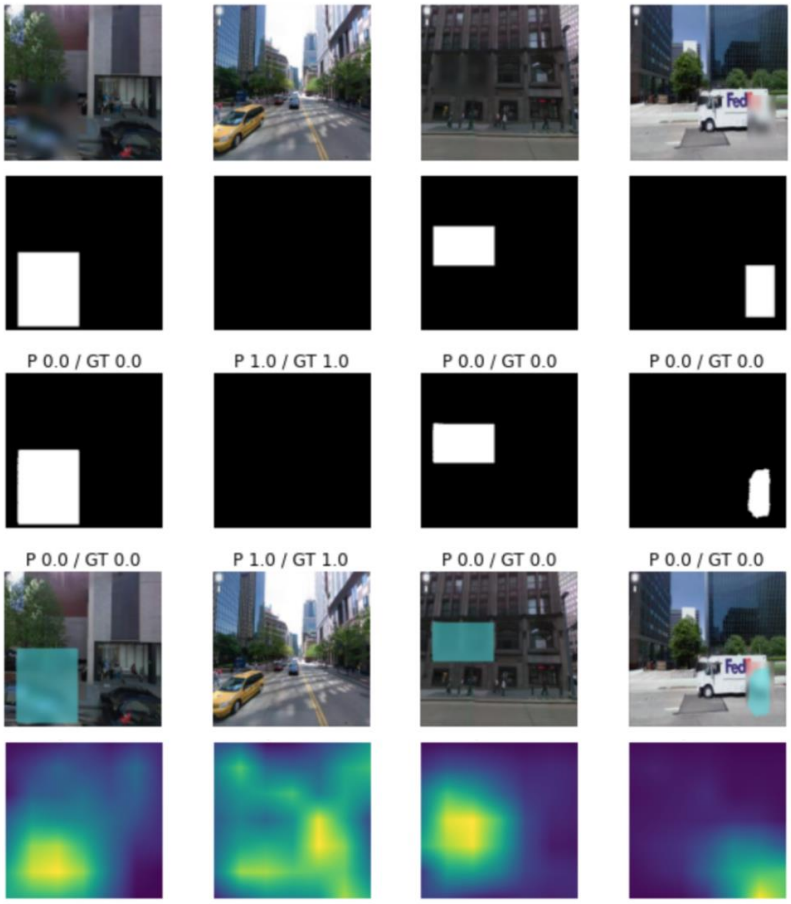

Figure 4. The SBDNet evaluation results for images from SVBI dataset. First Row: Selected input images from the SVBI dataset. Second Row: The corresponding ground-truth of input images. Third Row: SBDNet prediction outputs, including classification label and FBME of input images. Forth Row: Visualization of original images with a cyan colored mask over blurred regions based on the corresponding predicted FBME. Last Row: visualization of Keras Class Activation Map (CAM)[33] for input images.

5.4. Evaluation on CUHK, DUT, SZU-BD datasets

Unlike other state-of-the-art deep learning methods for image blur detection[9,12–19,34–37], our proposed SBDNet method was developed primarily for street-view images blur detection and classification. However, in this section, we examines the SBDNet method for blur detection and blur map estimation in a wide variety of blurred images from various categories, and not specifically street-view images. To test the validity of the SBDNet algorithm for general blur detection problem, the classification challenge was disregarded because the experiments in this section simply call for blur map estimation.

In order to compare the performance and efficiency of the SBDNet method with other cutting-edge blur detection methods, two experiments were carried out (see **Table 4**). In Experiment 1 and Experiment 2, the proposed SBDNet respectively is trained using images of the CUHK[1] dataset and evaluated by images from the CUHK[1], DUT[8] and SZU-BD[12] datasets.

Table 4. Train-test splits for training and evaluation of SBDNet on the CUHK[1], DUT[8], SZU-BD[12] datasets.

Experiments	Details
Experiment 1	Training: 800 images from CUHK[1] dataset
	Evaluation: 200 images from CUHK[1], 500 images from DUT[8], 784 images from SZU-BD[12] datasets
Experiment 2	Training: 604 defocus blur images from CUHK[1] dataset
	Evaluation: 100 defocus blur images from CUHK[1], 500 images from DUT[8], 784 images from SZU-BD[12] datasets

Experiment 1 setting: in Experiment 1, from the total of 1000 images in the CUHK[1] dataset, 800 images were selected for training and the remaining 200 images were utilized for evaluation. The evaluation phase was extended using 500 images of DUT[8] dataset and 784 images of SZU-BD[12] dataset.

Experiment 2 setting: following the completion of Experiment 1, for the purpose of an equivalent comparison between SBDNet and the state-of-the-art Defocus Blur Detection (DBD) methods, Experiment 2, was organized as follows: 604 images, out of 704 defocus blur images from CUHK[1] dataset were chosen for training of SBDNet, and the remaining 100 images were used for evaluation. The evaluation step was extended by using 500 images from the DUT[8] dataset and 784 images from the SZU-BD[12] dataset.

Table 5. The evaluation results of our proposed SBDNet model for blur map estimation, using CUHK[1], DUT[8], SZU-BD[12] datasets in terms of F-measure ↑ and MAE ↓ metrics. See **Table 4** for details of Experiment 1 and 2 (train and test splits).

Experiment	Dataset & metric					
	CUHK		DUT		SZU-BD	
	F-measure	MAE	F-measure	MAE	F-measure	MAE
Experiment 1	0.878	0.157	0.899	0.148	0.949	0.101
Experiment 2	0.916	0.146	0.910	0.139	0.948	0.112

Experiment results on blur map estimation: **Table 5** shows the best accomplished results from our SBDNet model. The results in **Table 5** were obtained from Experiment 1 and Experiment 2 using images from the CUHK[1], DUT[8] and SZU-BD[12] datasets in terms of the average F-measure ↑ and average MAE ↓ metrics. The arrow ↑ depicts the higher value is better and the arrow ↓ denotes the lower value is better. The ideal values for average F-measure and average MAE metrics are 1.0 and 0.0 respectively. It is important to highlight that the proposed SBDNet model trained and evaluated in Experiment 1 using both out-of-focus and motion-blurred images. Experiment 1 was conducted primarily to assess the performance and reliability of the SBDNet models in the face of motion blur and defocus blur images. The collected measurements from Experiment 1 involve all categories of blurred images. The results from Experiment 1, somehow can be quantitatively compared with the other DBD methods, noticing that Experiment 1 covers all blurred image categories. Experiment 2 mainly was performed to compare quantitatively the performance of the SBDNet models with other DBD methods. Since other deep learning algorithms employed the same Experimental setting for training, evaluation, and benchmarking on the CUHK[1], DUT[8], SZU-BD[12] datasets.

5.5. Comparisons with the state-of-the-art methods

We utilized the evaluation results of our proposed SBDNet model according to **Table 5** to compare our SBDNet method with other state-of-the-art methods. **Figure 5** shows the visual comparison of the SBDNet evaluated candidate images with other cutting-edge methods including Blur Detection-Enhanced Feature Pyramid Network (BD-EFPN) or Deep Pyramid Network (DPN)[12], deep blur mapping via Exploiting High-Level Semantics (EHS)[21], High-frequency multiscale Fusion and Sort Transform of gradient magnitudes (HiFST)[38], MVV[12],and Shi[1] for motion blur detection on CUHK[1] dataset. The illustrated comparison images in **Figure 5** has been collected from Sun et al.[12].

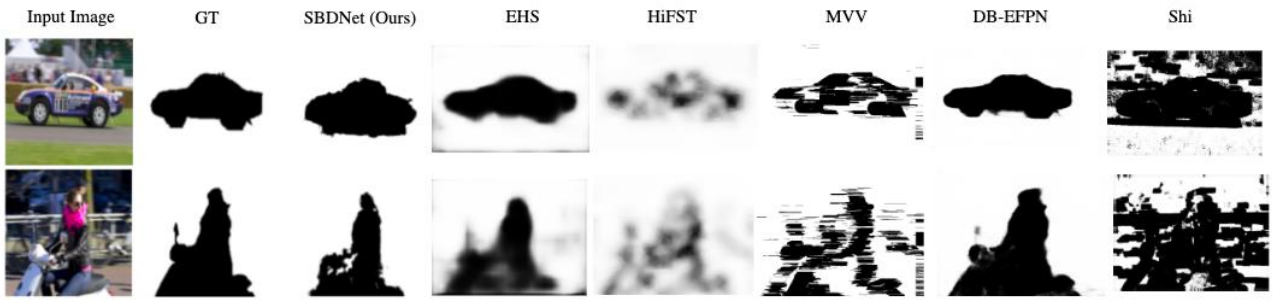

Figure 5. Visual comparison of SBDNet for motion blur detection on CUHK[1]. The illustrated results (EHS[21], HiFST[38], MVV[12], BD-EFPN[12], Shi[1]) were collected from Sun et al.[12].

Table 6 demonstrates the quantitative comparison of the obtained SBDNet model best results using ResNet50[2] as a Classifier network on the CUHK[1], DUT[8] datasets and using GoogLeNet[32] as a Classifier network on the SZU-BD[4] dataset in terms of average F-measure and average MAE metrics. In **Tabel 6,** we compared the performance of our proposed method with 6 state-of-the-art blur detection methods consisting of Distill-DBDGAN[13], Encoder Feature Ensemble Network (EFENet)[39], self-generated defocus blur detection via Dual Adversarial Discriminators network (DAD)[14], Defocus blur detection via depth distillation (DD)[40], HiFST[38], and LBP[41]. The results of the corresponding state-of-the-art methods have been collected from the work of Jonna et al.[13] which involves evaluations on three famous benchmarking datasets.

As shown in **Table 6**, our model achieves high F-measure scores across the CUHK, DUT, and SZU-BD datasets, ranging from 0.916 to 0.948. These results indicate that our model is competitive with state-of-the-art methods in identifying blurred and non-blurred regions in images under various scenarios.

Table 6. The quantitative comparisons of the proposed SBDNet model with cutting-edge blur detection methods on the CUHK[1], DUT[8], SZU-BD[12] datasets in terms of F-measure ↑ and MAE ↓ metrics. The other methods results have been collected from Jonna et al.[13].

Method	**Dataset & metric**					
	CUHK		**DUT**		**SZU-BD**	
	F-measure	**MAE**	**F-measure**	**MAE**	**F-measure**	**MAE**
LBP[41]	0.681	0.183	0.687	0.191	0.912	0.160
HiFST[38]	0.553	0.221	0.503	0.249	0.899	0.117
DD[40]	0.879	0.057	0.828	0.107	0.972	0.055
DAD[14]	0.916	0.172	0.884	0.079	0.794	0.153
EFENet[39]	0.914	0.053	0.854	0.094	0.968	0.073
D-DBDGAN[13]	0.926	0.041	0.901	0.068	0.969	0.064
SBDNet (Ours)	0.916	0.146	0.910	0.139	0.948	0.112

5.6. Ablation study

As it was stated, our proposed SBDNet method uses the Identifier network to produce FBME for an input image. We performed the ablation study to confirm the effectiveness of the De-fusionNet1 (green) and De-fusionNet1 (orange) subnetworks in the SBDNet model shown in **Figure 2**. In ablation analysis, SBDNet full-model including decoder network, De-fusionNet1 and DefusionNet2, is compared with three light-weight models descended from SBDNet full model which are nominated by SBDNet01, SBDNet02 and SBDNet03 (**Table 7**). The corresponding SBDNet01 is a descended version of SBDNet by disabling De-fusionNet1 and De-fusionNet2 for blur map estimation. The proposed SBDNet01 model generates FBME by employing decoder network. In the proposed SBDNet02 model, the De-fusionNet1 network is disabled, and the SBDNet02 model generates FBME using decoder network and De-fusionNet2 network. As shown in **Table 7**,

the third light-weight model is SBDNet03. The proposed SBDNet03 model produces FBME using decoder network and De-fusionNet1 network. To perform the proposed ablation study all models stated in **Table 7** were trained based on the experiment 1 setting (according to **Table 4**) by using 800 images of CUHK[1] dataset. All described models in **Table 7** were evaluated using 200 images from CUHK[1] and 500 images from DUT[8] datasets. All models used the same hyperparameters (e.g., learning rate coefficient, batch size). The training process was set to early stop upon reaching the specific threshold of accuracy value. **Table 8** shows the details of the ablation study in terms of average F-measure ↑, average MAE ↓, and Model Parameters ↓ metrics on CUHK[1] and DUT[8] datasets. In **Table 8**, Model Parameters ↓ indicates the total number of trainable parameters for corresponding models. The arrow ↓ denotes lower value is better. The proposed model with a lower Model Parameters value is more efficient.

According to the illustrated results in **Table 8**, SBDNet full-model has better performance compared to the smaller models in terms of the average F-measure metric on CUHK[1] and DUT[8] datasets. In terms of the average MAE metric, SBDNet full-model demonstrated better performance on the DUT[8] dataset. Considering, SBDNet01 as a base model with the lowest Model Parameters number (the proposed Identifier network uses only decoder network for blur map estimation), to produce FBME from an input image. The results demonstrate that the De-fusionNet1 network boosts the performance of SBDNet01 by increasing average F-measure score and reducing the average MAE score. By adding De-fusionNet2 to SBDNet01 without using Defusion-Net1, the performance of the SBDNet01 in terms of average F-measure and average MAE scores is reduced. The results shown in **Table 8** indicate that by applying De-fusionNet1 and De-fusionNet2 to SBDNet01 model, average F-measure score is improved by 0.006 on the CUHK[1] dataset and 0.007 on the DUT[8] dataset. Adding De-fusionNet1 and De-fusionNet2 to the SBDNet01 model causes better average MAE performance, by reducing 0.001 on the CUHK[1] dataset and 0.009 on the DUT[8] dataset.

Table 7. Settings of the ablation studies of the proposed SBDNet model. The ✓ symbol indicates that a component of the model has been incorporated, while the × symbol indicates that a component has been ablated.

Ablated model	**Components of the SBDNet model (See Figure 2)**
Full-model	✓ ResNet50 Classifier ✓ Decoder network ✓ De-fusionNet1 (green) ✓ De-fusionNet2 (orange)
SBDNet01	✓ ResNet50 Classifier × Decoder network ✓ De-fusionNet1 (green) ✓ De-fusionNet2 (orange)
SBDNet02	✓ ResNet50 Classifier ✓ Decoder network × De-fusionNet1 (green) ✓ De-fusionNet2 (orange)
SBDNet03	✓ ResNet50 Classifier ✓ Decoder network ✓ De-fusionNet1 (green) × De-fusionNet2 (orange)

Table 8. Ablation analysis of the proposed SBDNet model on the CUHK[1] and DUT[8] datasets using F-measure ↑, MAE ↓ and Model Parameters ↓ metrics. Details of ablated models are presented at **Table 7**.

Dataset	Metric	Abtaled model			
		SBDNet01	**SBDNet02**	**SBDNet03**	**Full-model**
	Model parameters	71,350,901	71,373,984	76,660,061	76,683,170
CUHK	F-measure	0.854	0.853	0.856	0.860
	MAE	0.198	0.191	0.186	0.197
DUT	F-measure	0.892	0.889	0.895	0.899
	MAE	0.157	0.164	0.152	0.148

6. Conclusion

This paper presents a novel deep learning algorithm for Google street-view images blur detection and classification. Classifier and Identifier networks are the two main architectural networks that make up the proposed SBDNet. Our proposed SBDNet collaboratively generates blur map estimation and the appropriate classification category label for each input street-view image. The Classifier network extracts high-level features and classifies input images. The Identifier network combines the obtained high-level semantic features from the Classifier network with low-level features to produce accurate blur map estimation. This paper also provided a new Street-View Blur Images (SVBI) dataset including 10,000 images for both types: blur and no blur, with the corresponding pixel-level, annotated ground-truths. The proposed SBDNet achieved outstanding experimental results over the SVBI dataset. Finally, SBDNet's accuracy and robustness for blur map estimation was evaluated and verified against the current existing baselines and datasets. The evaluation's results demonstrate that the SBDNet competes with the other cutting-edge methods.

Conflict of interest

The authors declare no conflict of interest.

References

1. Shi J, Xu L, Jia J. Discriminative blur detection features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 23–28 June 2014; Columbus, USA. pp. 2965–2972.
2. Tai YM, Brown MS. Single image defocus map estimation using local contrast prior. In: Proceedings of 2009 16th IEEE International Conference on Image Processing (ICIP); 7–10 November 2009; Cairo, Egypt. pp. 1797–1800.
3. Zhuo S, Sim T, Defocus map estimation from a single image. Pattern Recognition 2011; 44(9): 1852–1858. doi: 10.1016/j.patcog.2011.03.009
4. Su B, Lu S, Tan CL. Blurred image region detection and classification. In: Proceedings of the 19th ACM international conference on Multimedia; 28 November–1 December 2011; Scottsdale, Arizona, USA. pp. 1397–1400.
5. Couzinie-Devy F, Sun J, Alahari K, Ponce J. Learning to estimate and remove non-uniform image blur. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 23–28 June 2013; Portland, USA. pp. 1075–1082.
6. Tang C, Hou C, Song Z. Defocus map estimation from a single image via spectrum contrast. Optics Letters 2013; 38(10): 1706–1708. doi: 10.1364/OL.38.001706
7. Tang C, Hou C, Hou Y, et al. An effective edge-preserving smoothing method for image manipulation. Digital Signal Processing 2017; 63: 10–24. doi: 10.1016/j.dsp.2016.10.009
8. Zhao W, Zhao F, Wang D, Lu H. Defocus blur detection via multistream bottom-top-bottom fully convolutional network. IEEE Transactions on Pattern Analysis and Machine Intelligence 2020; 42(8): 1884–1897. doi: 10.1109/TPAMI.2019.2906588
9. Zhao Z, Yang H, Luo H. Hierarchical Edge-aware Network for defocus blur detection. Complex & Intelligent Systems 2022; 8: 4265–4276. doi: 10.1007/s40747-022-00711-y
10. Zhang S, Shen X, Lin Z, et al. Learning to understand image blur. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 18–23 June 2018; Salt Lake, USA. pp. 6586–6595.
11. Available online: https://github.com/MasoudMoeini/Google-Street-View-Images-Blur-Detection (accessed on 31

July 2023).
12. Sun X, Zhang X, Xiao M, Xu C. Blur detection via deep pyramid network with recurrent distinction enhanced modules. Neurocomputing 2020; 414: 278–290. doi: 10.1016/j.neucom.2020.06.068
13. Jonna S, Medhi M, Sahay RR. Distill-dbdgan: Knowledge distillation and adversarial learning framework for defocus blur detection. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 2022; 19(2): 87. doi: 10.1145/3557897
14. Zhao W, Shang C, Lu H. Self-generated defocus blur detection via dual adversarial discriminators. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 20–25 June 2021; Nashville, USA. pp. 6933–6942.
15. Zhao Z, Yang H, Luo H. Defocus blur detection via transformer encoder and edge guidance. Applied Intelligence 2022; 52: 14426–14439. doi: 10.1007/s10489-022-03303-y
16. Lin X, Li H, Cai Q. Hierarchical complementary residual attention learning for defocus blur detection. Neurocomputing 2022; 501: 88–101. doi: 10.1016/j.neucom.2022.06.023
17. Guo W, Xiao X, Hui Y, et al. Heterogeneous attention nested u-shaped network for blur detection. IEEE Signal Processing Letters 2021; 29: 140–144. doi: 10.1109/LSP.2021.3128375
18. Tang C, Liu X, Zheng X, et al. DefusionNet: Defocus blur detection via recurrently fusing and refining discriminative multi-scale deep features. IEEE Transactions on Pattern Analysis and Machine Intelligence 2020; 44(2): 955–968. doi: 10.1109/TPAMI.2020.3014629
19. Tang C, Liu X, Zhu X, et al. R^2MRF: Defocus blur detection via recurrently refining multi-scale residual features. Proceedings of the AAAI Conference on Artificial Intelligence 2020; 34(7): 12063–12070. doi: 10.1609/aaai.v34i07.6884
20. Zhao W, Zheng B, Lin Q, Lu H. Enhancing diversity of defocus blur detectors via cross-ensemble network. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 15–20 June 2019; Long Beach, USA. pp. 8905–8913.
21. Ma K, Fu H, Liu T, et al. Deep blur mapping: Exploiting high-level semantics by deep neural networks. IEEE Transactions on Image Processing 2018; 27(10): 5155–5166. doi: 10.1109/TIP.2018.2847421
22. Kim B, Son H, Park SJ, et al. Defocus and motion blur detection with deep contextual features. Computer Graphics Forum 2018; 37(7): 277–288. doi: 10.1111/cgf.13567
23. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 27–30 June 2016; Las Vegas, USA. pp. 770–778.
24. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 18–23 June 2018; Salt Lake, USA. pp. 7132–7141.
25. Yang M, Yu K, Zhang C, et al. Denseaspp for semantic segmentation in street scenes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 18–23 June 2018; Salt Lake, USA. pp. 3684–3692.
26. Tan M, Le Q. Efficientnet: Rethinking model scaling for Convolutional Neural Networks. Proceedings of Machine Learning Research 2019; 97: 6105– 6114.
27. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv 2014; arXiv:1409.1556. doi: 10.48550/arXiv.1409.1556
28. Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks. arXiv 2017; arXiv:1611.05431. doi: 10.48550/arXiv.1611.05431
29. Zamir AR, Shah M. Image geo-localization based on multiplenearest neighbor feature matching usinggeneralized graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence 2014; 36(8): 1546–1558. doi: 10.1109/TPAMI.2014.2299799
30. Jalled F, Voronkov I. Object detection using image processing. arXiv 2016; arXiv:1611.07791. doi: 10.48550/arXiv.1611.07791
31. Rastogi A, Ryuh BS. Teat detection algorithm: Yolo vs. Haar-cascade. Journal of Mechanical Science and Technology 2019; 33(4): 1869–1874. doi:10.1007/s12206-019-0339-5
32. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 7–12 June 2015; Boston, USA. pp. 1–9.
33. Gulli A, Pal S. Deep Learning with Keras. Packt; 2017.
34. Tang C, Zhu X, Liu X, et al. DefusionNet: Defocus blur detection via recurrently fusing and refining multi-scale deep features. IEEE Transactions on Pattern Analysis and Machine Intelligence 2019; 44(2): 955–968. doi: 10.1109/TPAMI.2020.3014629
35. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. Advances in Neural Information Processing Systems 2022; 63(11): 139–144. doi: 10.1145/3422622
36. Khmag A, Ramlee R, Blur removal in natural digital images using self- reference generative networks. Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 2021; 13(3): 61–65.
37. Khmag A, Ramli AR, Kamarudin N. Clustering-based natural image denoising using dictionary learning approach in wavelet domain. Soft Computing 2019; 23(17): 8013–8027. doi: 10.1007/s00500-018-3438-9
38. Golestaneh SA, Karam LJ. Spatially-varying blur detection based on multiscale fused and sorted transform coefficients of gradient magnitudes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern

Recognition; 21–26 July 2017; Honolulu, USA. pp. 5800–5809.
39. Zhao W, Hou X, He Y, Lu H. Defocus blur detection via boosting diversity of deep ensemble networks. IEEE Transactions on Image Processing 2021; 30: 5426–5438. doi: 10.1109/TIP.2021.3084101
40. Cun X, Pun CM. Defocus blur detection via depth distillation. In: European Conference on Computer Vision, Proceedings of 16th European Conference; 23–28 August 2020; Glasgow, UK. Springer Cham; pp. 747–763.
41. Yi X, Eramian M. Lbp-based segmentation of defocus blur. IEEE Transactions on Image Processing 2016; 25(4): 1626–1638. doi: 10.1109/TIP.2016.2528042