IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 December 2023, accepted 20 January 2024, date of publication 26 January 2024, date of current version 1 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358825

RESEARCH ARTICLE

Data and Model Uncertainty Aware Salient Object Detection

HEEJIN LEE[1], (Associate Member, IEEE), SEUNGHYUN LEE[1,2], (Associate Member, IEEE), AND BYUNG CHEOL SONG[1], (Senior Member, IEEE)
[1]Department of Electrical and Computer Engineering, Inha University, Incheon 22212, South Korea
[2]Upstage AI, Yongin 16942, South Korea

Corresponding author: Byung Cheol Song (bcsong@inha.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korea Government (MSIT) under Grant 2022R1A2C2010095 and Grant 2022R1A4A1033549, and in part by the Institute of Information & Communications Technology Planning & Evaluation (IITP) Grant funded by the Korea Government (MSIT) (Artificial Intelligence Convergence Innovation Human Resources Development (Inha University) and AI Innovation Hub) under Grant RS-2022-00155915 and Grant 2021-0-02068.

ABSTRACT In general, salient object detection (SOD) datasets have ambiguity due to annotation accuracy and human subjectivity in determining saliency. Since this data uncertainty causes inaccurate prediction, many techniques tackling data uncertainty have been proposed so far. Previous works estimated data uncertainty in terms of predictive inaccuracy and adjusted the learning contribution so that a given model can focus more on specific data. However, inaccurate predictions can occur due to not only data uncertainty but also model uncertainty in which the model does not fully explain the data. As a result, a region that is inaccurately predicted due to model uncertainty is considered a region with high data uncertainty, resulting in insufficient learning. To solve this problem, we propose a novel uncertainty-aware learning scheme where model uncertainty is decomposed from prediction uncertainty and it is minimized. Also, we propose a refinement method to further improve performance by correcting the prediction result using data uncertainty in the inference step. The proposed uncertainty-aware method excludes data uncertainty from learning step and inference step more effectively, making the model more accurately detect salient object(s). The experimental results prove that the proposed method achieves state-of-the-art performance on several SOD datasets and qualitatively detects salient objects more accurately than the prior arts. The code will be uploaded on Github.

INDEX TERMS Data uncertainty, model uncertainty, salient object detection.

I. INTRODUCTION

Salient object detection (SOD) is a technique that imitates the human cognitive system to detect and segment main objects or regions in real-world images. SOD has been an important tool for various computer vision tasks such as object recognition/detection [1], [2], [3], image/video segmentation [4], [5], [6], and visual tracking [7]. With the advent of convolutional neural networks (CNNs), SOD has been developed rapidly [8], [9].

SOD datasets are generally composed of pairs of input images and ground truth (GT) images in which salient objects are labeled in the form of binary masks. Since saliency is defined in various ways according to human subjectivity, the labeled salient object has data ambiguity [10]. In particular, since the contours of salient objects are difficult to be accurately labeled, data ambiguity due to such poor annotation quality is inevitable [11]. Thus, SOD datasets contain data uncertainty due to human subjectivity and annotation issues. Here, data uncertainty refers to ambiguity caused by human subjectivity and noise in the data itself, and is defined as aleatoric uncertainty [12]. Since data uncertainty is noise inherent in the data itself and cannot be solved by learning, some techniques have been proposed to minimize the effect of data uncertainty on learning [13], [14]. As the most representative approach, data uncertainty map was

The associate editor coordinating the review of this manuscript and approving it for publication was Zhongyi Guo.

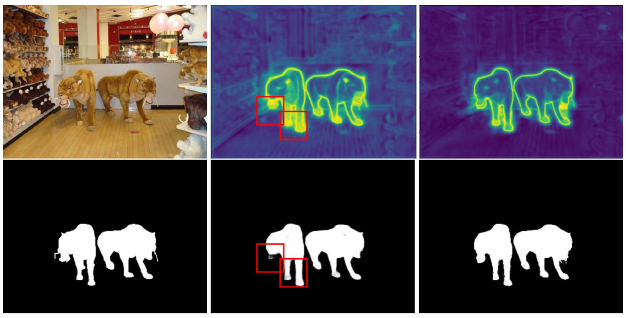

FIGURE 1. Data uncertainty and prediction according to whether model uncertainty is minimized in estimating data uncertainty. (a) Input data and ground truth (b) data uncertainty and salient objects estimated by previous methods (c) data uncertainty and salient objects obtained by the proposed method to which model uncertainty minimization process is applied.

estimated based on prediction inaccuracy, i.e., the distance of the prediction map from GT, and it was used as a loss attenuation [12].

Note that prediction inaccuracy results from not only data uncertainty but also model uncertainty. Model uncertainty is uncertainty in which the model does not sufficiently explain the data, and unlike data uncertainty, it can be minimized. Model uncertainty occurs when the model does not sufficiently learn data. So, if model uncertainty cannot be resolved due to insufficient learning, it can be reflected in the estimated data uncertainty. As a result, even regions with high model uncertainty are considered to have high data uncertainty. This results in a problem that the data is not sufficiently learned. To make this problem clear, let's qualitatively analyze the data uncertainty map of Fig. 1 (b) which does not consider model uncertainty. We can observe that the uncertainty is estimated even on a relatively easy (or accurately predictable) background. Also, the detection accuracy in the region with high data uncertainty is low (see the red boxes). Nevertheless, any technique that considers both uncertainties does not yet exist.

In this paper, we propose a novel method of estimating data uncertainty more accurately by minimizing the involvement of model uncertainty and using it for learning and inference of the SOD model. The proposed method consists of an uncertainty-aware learning step and a refinement step. First, we present a learning strategy that additionally considers model uncertainty to improve the existing data uncertainty-aware learning algorithm(s). The proposed model uncertainty-aware learning strategy is to decompose model uncertainty from the prediction uncertainty caused by the two uncertainties, and then minimize it. The higher learning weights are given in regions with high model uncertainty, so model uncertainty is resolved quickly, which results in more accurate data uncertainty estimation. In addition, we present a method to refine the prediction map based on data uncertainty in the inference step, enabling more accurate salient object inference even in unseen data.

Experiments show that the proposed uncertainty-aware method remarkably improves SOD algorithms in both quantitative and qualitative aspects. The proposed method added on to a recent SOD method [15] achieves state-of-the-art (SOTA) performance on various SOD datasets. Looking at the data uncertainty map of the proposed method that minimizes model uncertainty, the uncertainty in the background is greatly reduced (see Fig. 1(c)). We can also observe that the detection accuracy of the red box in Fig. 1(b) is improved because even data with low data uncertainty is sufficiently learned.

The contribution points of this paper are as follows.

- Uncertainty-aware learning that simultaneously considers data and model uncertainties is proposed for the first time in SOD task. Through the proposed uncertainty-aware learning, a pixel-wise data uncertainty map in which model uncertainty is constrained can be generated.
- In the inference step, a method for correcting the prediction map according to model uncertainty is proposed, which improves the detection accuracy of salient objects predicted by the model with certainty.
- The proposed method achieves SOTA performance on four SOD datasets, and is qualitatively superior to existing methods.
- The proposed method can be easily added-on to other SOD methods for further performance improvement.

II. RELATED WORKS

With the rapid development of deep learning, the performance of CNN-based SOD models has overtaken traditional techniques. CNN-based SOD methods extract semantic information or features of various levels and integrate them efficiently. Reference [16] proposed a network consisting of two branches: Supervised learning and unsupervised learning. The two branches model the interaction between human intuition and memory to detect salient object(s). Recently, several multi-scale techniques have been proposed to further improve the learning ability of semantic information. For example, MINet [17] is a network that fuses features of adjacent layers, which detects salient objects through a strategy to minimize noise caused by the difference in resolution of feature maps by using small up-/down-sampling rates. Chen et al. [18] proposes a novel method with designing a parallel multi-scale structure to integrate the salient features at each levels. Ji et al. [19] proposed a method for learning context between feature information of different scales by applying spatial and channel unit attention modules to multi-scale encoder-decoder networks. In addition, a few techniques for additionally learning contour (or edge) information of an input image as well as a prediction map have improved segmentation performance of SOD. Reference [20] proposed EGNet, which learns

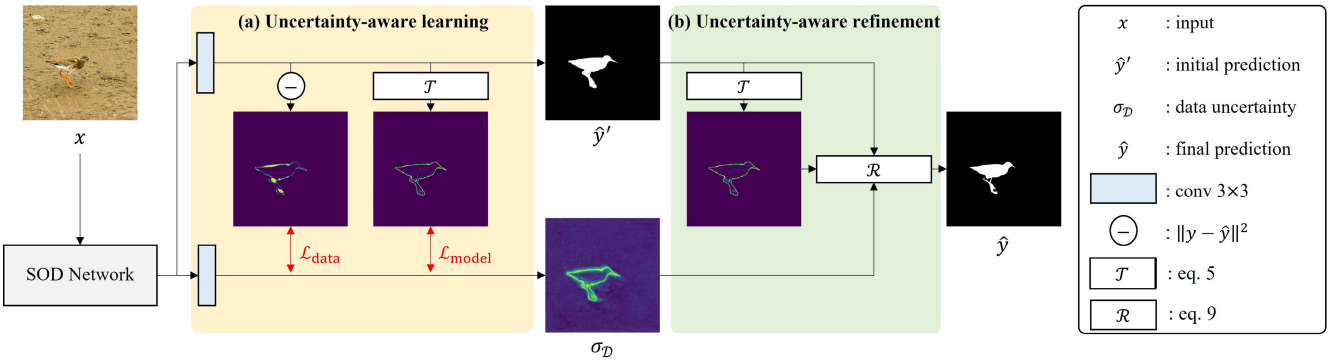

FIGURE 2. The proposed data and model uncertainty-aware method. Step (a) extracts pixel-wise data and model uncertainty map. Through Step (b), the uncertainty map of unknown data is used as a zero-masking map in the inference step.

edge information through supervised learning based on multiple branches and integrates the learned local edge information and global location information. Han et al. [21] added an edge convolution contraint to U-Net to predict more accurate saliency map. SOD-ID [22] detects salient object and estimate importance degree with object contours information. Reference [23] discovered a problem that the contour-saliency fusion of existing SOD techniques to learn contour information causes excessive false positives, and proposed the RCSBN that employs recursive CNNs to extract and learn contours and saliency. [24] proposed a binary cross entropy loss, i.e., contour loss, in which contour information was given a higher weight to effectively predict the contour of main object. DIMONet [25] extract both object and boundary features with dual-branch, and refines each features with mutual optimization and fusion modules. Meanwhile, many techniques for estimating and resolving uncertainty in SOD tasks have been devised. As mentioned above, uncertainty is classified into data and model uncertainty. Reference [10] proposed a generative model UCNet that generates multiple saliency predictions to estimate data (labeling) uncertainty due to subjective criteria for saliency. Reference [14] proposed a so-called adversarial learning network that newly defines the uncertainty that occurs in the process of labeling main objects or backgrounds and solves the uncertainty through joint-learning of the camouflaged object detection task and the salient object detection task. Reference [26] estimate out-of-distribution of SOD for the first time to investigate distributional uncertainty with deep ensemble and single-model uncertainty methods. Reference [13] argued that there is a limitation to precise labeling of training-purpose datasets, and proposed a data uncertainty loss function that overcomes the limitation as an alternative. Reference [27] used an adversarial decoder with a saliency network to extract a confidence map representing model uncertainty and use it for learning. Reference [12] estimated data uncertainty through prediction inaccuracy based on the difference between the prediction map and GT, which is known as the most general approach for estimating data uncertainty.

However, according to [12], inaccuracies in prediction can be caused not only by data uncertainty but also by model uncertainty. So, if model uncertainty is not considered, a region that is inaccurately predicted due to model uncertainty is considered a region with high data uncertainty, which may cause insufficient learning of the data. Thus, this paper proposes a novel learning strategy considering data uncertainty and model uncertainty at the same time. In addition, we propose a refinement method that further improves the detection accuracy of salient objects predicted by the model with certainty by applying the estimated uncertainty to the inference step.

III. PROPOSED METHOD

This section details the uncertainty-aware learning and refinement steps of the proposed method, which are conceptually visualized in Fig. 2. Note that the proposed method can be attached to the generic deep learning-based SOD network to further improve its performance.

The first step is (a) uncertainty-aware learning. As we point out that existing data uncertainty-aware learning does not consider model uncertainty, we propose the estimation and learning strategy of data and model uncertainties. We set auxiliary blocks to estimate data uncertainty and then apply $\mathcal{L}_{\text{data}}$ and $\mathcal{L}_{\text{model}}$ to consider both data and model uncertainty on learning. The second step is (b) uncertainty-aware refinement. To improve the reliability of the estimated salient object in the inference step, we apply a zero-masking refinement framework on high model uncertainty region that is difficult for the model to estimate as unknown data.

A. DATA UNCERTAINTY-AWARE LEARNING

This section mathematically defines the existing pixel-wise data uncertainty, and analyzes the previous learning approach considering data uncertainty [12] and its limitation. While model uncertainty can be resolved if only data for learning is sufficient, data uncertainty cannot be directly resolved due to various noises inherent in the data itself, such as sensor noise and motion noise. Especially, since SOD is a task of detecting salient objects from images captured in the real-world,

as mentioned in Section I, SOD datasets inevitably have data uncertainty. If we can effectively cope with this inevitable data uncertainty, stabilization of prediction performance will be achieved.

To solve this fundamental problem of SOD, [12] proposed a technique for estimating data uncertainty and learning it. Specifically, pixel-wise data uncertainty was estimated based on the Gaussian likelihood and it was applied to learning. Eq. 1 defines the Gaussian likelihood with respect to the prediction $\hat{y}'$ for an input image x, GT y, and data uncertainty $\sigma^{\mathcal{D}}$.

$$p(y|\hat{y}') = \mathcal{N}(\hat{y}', (\sigma^{\mathcal{D}})^2) \tag{1}$$

where $\mathcal{N}$ indicates Gaussian distribution. $\sigma_i^{\mathcal{D}}$ indicates pixel-wise data uncertainty, and the larger the value, the more difficult it is to accurately predict the data. So, [12] assumed that a region is difficult to predict as the data uncertainty increases, and introduced an auxiliary layer to estimate it. Then, based on the estimated uncertainty map, they employed a loss function that adjusts the learning contribution of the data, which is as follows:

$$\mathcal{L}_{\text{data}} = \frac{1}{P}\sum_{i\in P}\frac{1}{2(\sigma_i^{\mathcal{D}})^2}||y_i - \hat{y}_i'||^2 + \frac{1}{2}\log(\sigma_i^{\mathcal{D}})^2 \tag{2}$$

where P indicates a set of pixels, and $\sigma_i^{\mathcal{D}}$, y_i, $\hat{y}_i'$ indicates a pixel value of data uncertainty, GT, and initial prediction, respectively. Here, the first term adopts the inverse of the data uncertainty map as loss attenuation, which indicates that the higher the data uncertainty, the lower the weight. In other words, this term serves to decrease the learning weights in regions with high data uncertainty. The second term is a regularization term to prevent the data uncertainty map from diverging infinitely. If Eq. (2) is used as a loss function, the function diverges as $\sigma_i^{\mathcal{D}}$ approaches 0. Therefore, for numerical stability, we set $\log\sigma^{\mathcal{D}}$ as a trainable variable that allows stable learning considering the characteristics of logarithmic and exponential functions. Accordingly, $\mathcal{L}_{\text{data}}$ is defined by the following equations.

$$\exp\hat{\sigma}^{\mathcal{D}} = \exp\log\sigma^{\mathcal{D}} = \sigma^{\mathcal{D}} \tag{3}$$

$$\mathcal{L}_{\text{data}} = \frac{1}{P}\sum_{i\in P}\exp(-(\hat{\sigma}_i^{\mathcal{D}})^2)||y_i - \hat{y}_i'||^2 + (\hat{\sigma}_i^{\mathcal{D}})^2 \tag{4}$$

Although the estimated $\sigma^{\mathcal{D}}$ seems to effectively represent data uncertainty, the strong assumption that prediction errors occur due to data uncertainty may cause some errors. To analyze this phenomenon qualitatively, Fig. 3 shows the prediction map and data uncertainty map for an input image. The red box in the first row is regarded as a region where data uncertainty exists due to human subjectivity in judging main object(s). In fact, looking at $\sigma^{\mathcal{D}}$, we can find that the region is strongly activated, which means that data uncertainty is well estimated. On the other hand, looking at the red box of the second row, although the data uncertainty is not high, the prediction accuracy is greatly reduced, and the corresponding region is strongly activated in $\sigma^{\mathcal{D}}$.

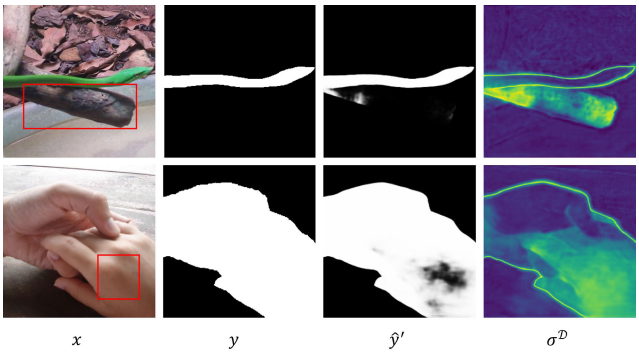

FIGURE 3. Prediction and data uncertainty map estimated through data uncertainty-aware learning.

In other words, an inaccurate prediction occurred due to another factor other than data uncertainty in the region, and accordingly, the data uncertainty map was also incorrectly estimated. Therefore, we argue that prediction inaccuracy is not necessarily caused by data uncertainty alone, and that we need to analyze and resolve another factor of prediction inaccuracy.

B. MODEL UNCERTAINTY-AWARE LEARNING

In this section, we introduce model uncertainty, another factor involved in estimating data uncertainty. And to improve the performance of uncertainty-aware learning by minimizing model uncertainty, we propose a strategy to decompose only data uncertainty from the prediction error map. $\sigma^{\mathcal{D}}$ is learned to be large in the region where the difference between GT y and prediction $\hat{y}'$, that is, the prediction inaccuracy is large. However, as in Fig. 3, prediction inaccuracy is not caused by data uncertainty alone. As in [12], we also define another factor of prediction inaccuracy as model uncertainty, which occurs because the model is not sufficiently trained. If regions with large prediction error are regarded as having large data uncertainty and their contribution to learning is reduced, regions with only high model uncertainty may not be properly learned. To solve this problem, we present a method to decompose model uncertainty from the estimated data uncertainty map and quickly minimize it. In other words, both data uncertainty and model uncertainty are learnable, but have the characteristics of being unreducible and reducible, respectively. Therefore, since the two uncertainties are independent of each other in terms of solvability, we assume that prediction uncertainty can be decomposed into data and model uncertainty. Note that this assumption was derived by referring to uncertainty decomposition techniques [28], [29].

To estimate model uncertainty, we first define the prediction uncertainty caused by data uncertainty $\sigma^{\mathcal{D}}$ and model uncertainty $\sigma^{\mathcal{M}}$. Since SOD is a binary classification, values far from 0 or 1, i.e., 0.5 is judged to have the highest prediction uncertainty $\mathcal{T}$, which is represented by

$$\mathcal{T} = 1 - (2\hat{y}' - 1)^2. \tag{5}$$

Here, to simplify this problem, we assume that the two uncertainties are independent. Then $\mathcal{T}$ can be decomposed into the product of $\sigma^{\mathcal{D}}$ and $\sigma^{\mathcal{M}}$. Thus, model uncertainty $\sigma^{\mathcal{M}}$ is defined by

$$\sigma^{\mathcal{M}} = \mathcal{T}/\sigma^{\mathcal{D}} \tag{6}$$

Also, based on model uncertainty, we define the model uncertainty constraint $\mathcal{L}_{\text{model}}$ as follows.

$$\mathcal{L}_{\text{model}} = \frac{1}{P}\sum_{i\in P}\exp(-\hat{\sigma}_i^{\mathcal{D}})\mathcal{T} \tag{7}$$

$\mathcal{L}_{\text{model}}$ is a loss function that is independent of data uncertainty, that is, allows the model to additionally learn sufficiently predictable data through the Eq. 7. The model uncertainty constraint strongly polarizes prediction uncertainty with low data uncertainty, which boosts the gradients in regions where learning is inhibited due to $\mathcal{L}_{\text{data}}$ so that they can be learned sufficiently. As a result, the model's decision-making ambiguity is minimized and the contribution to learning in regions where learning is inhibited due to data uncertainty loss is improved.

The total loss function $\mathcal{L}_{\text{total}}$ of the proposed method consists of the binary cross entropy $\mathcal{L}_{\text{bce}}$, which is a fundamental loss function for the SOD task, and two uncertainty-based constraints, i.e., $\mathcal{L}_{\text{data}}$ and $\mathcal{L}_{\text{model}}$. $\mathcal{L}_{\text{total}}$ is represented as follows:

$$\mathcal{L}_{\text{total}} = \lambda_1\mathcal{L}_{\text{bce}} + \lambda_2\mathcal{L}_{\text{data}} + \lambda_3\mathcal{L}_{\text{model}} \tag{8}$$

Both λ_1 and λ_2 are weights of binary cross entropy and data uncertainty loss that are learned based on GT and output. In this paper, they were set to the same parameters. Here, model uncertainty loss was estimated from prediction and data uncertainty, so λ_3 was set equal to λ_2. Therefore, the proposed uncertainty-aware learning strategy minimizes the involvement of model uncertainty to induce a more accurate estimation of the data uncertainty map, and adjusts the learning weight so that the model can intensively learn more meaningful data. As a result, the SOD network trained by the proposed method can provide qualitatively and quantitatively improved performance, as shown in Section IV.

C. UNCERTAINTY-AWARE REFINEMENT

However, even models that have learned both uncertainties can still make inaccurate predictions when encountering unseen data in the inference step. In other words, prediction uncertainty may increase for unseen data. This section presents a method to further improve performance by applying a specific refinement based on prediction uncertainty to the estimated prediction map.

Note that the main goal of SOD task is to predict the salient object with high precision by minimizing false positives [30]. So, we propose a refinement step which zero-masks a region with high data uncertainty and sets it as a background. In detail, we first apply a min-max normalization to make the range of $\sigma^{\mathcal{D}}$ be within [0, 1], and then build a mask with the region higher than a specific threshold. Here, although $\sigma^{\mathcal{D}}$ is learned to minimize the effect of model uncertainty, the accuracy of the unseen data may be somewhat lower because it is still estimated by the auxiliary network. Therefore, in order to increase the reliability of this refinement step, we adopt another condition based on prediction uncertainty $\mathcal{T}$, which is represented by

$$\mathcal{R} = \begin{cases} 0 & \text{if } \mathcal{T}_i > \theta^{\mathcal{P}} \text{ and } \sigma_i^{\mathcal{D}} > \theta^{\mathcal{D}} \\ \hat{y}_i' & \text{otherwise} \end{cases} \tag{9}$$

where $\theta^{\mathcal{P}}$ and $\theta^{\mathcal{D}}$ are set to 0.8. Even though the proposed refinement process was designed through somewhat heuristic decision-making, it is sufficiently reasonable in terms of algorithm configuration. In particular, note that it provides superior performance compared to the previous methods (see Section IV-D).

In summary, the proposed framework for uncertainty-aware learning and refinement improves salient object accuracy even in unseen data by minimizing the effect of data uncertainty that adversely affects SOD performance. In addition, the proposed method can be attached to all SOD learning frameworks as an add-on module without architecture dependency to improve its performance, which is verified experimentally in the next section.

IV. EXPERIMENTS

A. EXPERIMENTS CONFIGURATION

In this paper, the performance of the proposed method was evaluated on five benchmark datasets, i.e., DUTS-TE [31], DUT-OMRON [32], ECSSD [33], HKU-IS [34], and PASCAL-S [35]. DUTS-TE consists of 5,019 complex scenes, and DUTS-OMRON consists of 5,168 images composed of complex backgrounds and various contents. ECSSD consists of 1,000 complex nature images, and HKU-IS consists of 4,447 images composed of several separate objects and structurally similar foregrounds and backgrounds. PASCAL-S consists of 850 complex images [17].

We evaluated the performance of each model based on five metrics: S-measure [36] evaluates the structural similarity between prediction map and GT map while considering object-aware and region-aware structural similarities. Weighted F-measure F_β^ω [37] is an intuitive generalization of F-measure by alternately applying different weights ω to precision and recall calculations. MAE [38] calculates the average absolute value of the pixel-wise difference between the predicted map and the GT map. E-measure [39] simultaneously evaluates the global mean and local pixel agreement between the predicted and GT maps. F-measure [40] computes the weighted harmonic mean, taking both precision and recall into account comprehensively. Here, we set both E and F-measure as average values.

The proposed method was implemented in Pytorch. Except for the ablation study, ZoomNet [15] was adopted as the backbone SOD Network. As in [17], [41], and [42], ResNet-50 pre-trained with ImageNet was used as an encoder. The

TABLE 1. Quantitative results of the proposed and comparative methods on each dataset. Here, ↑ and ↓ indicate that the lower or higher, the better the performance, respectively. Also, red, blue, and bold indicate the first, second, and third place, respectively.

Method	DUTS-TE					DUT-OMRON					ECSSD				
	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$
BASNet	.843	.763	.059	.873	.783	.817	.721	.068	.846	.737	.907	889	.042	.937	.902
RCSBN	.873	.832	.036	.913	.854	.829	.745	.051	.854	.763	.903	.891	.041	.928	.909
PoolNet+	.890	.822	.037	.914	.844	.834	.730	.053	.851	.753	.926	.906	.036	.946	.920
HQSOD	**.891**	.859	.032	.933	.876	**.840**	.770	**.050**	.877	.787	.927	.927	**.029**	.957	.939
EDN	.890	.841	.035	.923	.860	.848	.767	.050	.875	.785	**.927**	.918	.033	.952	**.930**
ZoomNet	.899	**.853**	**.032**	**.928**	**.864**	.836	.745	.051	.854	.758	.933	**.924**	.027	**.957**	.930
Ours	.895	.860	.031	.929	.876	.840	**.762**	.048	**.865**	**.778**	.933	.931	.027	.959	.941

Method	HKU-IS					PASCAL-S				
	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$
BASNet	.901	.875	.036	.936	.888	.818	.761	.088	.853	.787
RCSBN	.904	.891	.031	.939	.910	.841	.797	.067	.877	.823
PoolNet+	.920	.897	.029	.949	.910	.863	.809	.065	.895	.834
HQSOD	.921	.919	.026	.960	.928	.859	**.832**	**.060**	**.905**	.853
EDN	**.924**	.908	**.026**	.955	.919	**.864**	.825	.061	.903	.846
ZoomNet	.929	**.915**	.026	**.958**	**.920**	.869	.834	.054	.913	**.851**
Ours	.926	.919	.023	.959	.930	.872	.843	.052	.916	.862

weight decay was set to 0.0005, stochastic gradient descent (SGD) was employed as the optimizer, and momentum was set to 0.9. The learning rate was set to 0.05. The model was trained for a total of 50 epochs at NVIDIA RTX A600 environment with the model batch size set to 22. As a training dataset, DUTS-TR [32] was used in the same way as SOTA methods [8], [15], [20], [43]. The size of the input image was set to 352 × 352.

B. QUANTITATIVE RESULTS

We quantitatively compared the proposed method with seven SOTA SOD methods: BASNet [41], MINet [17], RCSBN [23], PoolNet+ [44], HQSOD [13], EDN [43], and ZoomNet [15]. For a fair comparison, numerical figures for each method were obtained through prediction maps generated by source codes released by the authors. Here, both the learning dataset and the learning environment are the same.

The quantitative results based on the aforementioned dataset and evaluation criteria are shown in Table 1. The proposed method not only effectively improved the performance of the baseline ZoomNet [15], but also showed superior performance compared to the existing methods. Note that we achieved SOTA performance on four datasets: DUTS-TE, ECSSD, HKU-IS, and PASCAL-S. In detail, the MAE of the proposed method was the best in all the datasets. This proves that the proposed method more accurately detects both salient and non-salient regions than other ones [30]. In terms of F_m, the proposed method greatly improved ZoomNet by 0.01 to 0.02. This is because the proposed uncertainty-aware learning enabled to learn certain data. In addition, the proposed refinement scheme effectively removes false positives while minimizing the increase in false negatives. The quantitative results in the table seem to show a rather low performance improvement. Instead, it is observed that the proposed method is added-on to various learning-based baselines and contributes to greatly improving the performance of the baselines. Additional observations can be found in Section IV-D. Therefore, we claim that the proposed method effectively could improve SOD performance.

C. QUALITATIVE RESULTS

Figure 4 compares the previous SOD methods and the proposed method qualitatively. Previous methods and learning environment are the same as in Section IV-B. Each row of Fig. 4 shows salient object(s) detected from images sampled from DUTS-TE [31], DUT-OMRON [32], ECSSD [33], HKU-IS [34], and PASCAL-S [35], respectively. Here, as mentioned in Section III-B, the more it is not binarized to 0 or 1, the higher the prediction uncertainty, which is visualized as gray.

The first row is the result of complex image where it is difficult to distinguish between the salient object and the background. Other methods had a high prediction uncertainty in the region with fine details, but in the proposed method, the prediction uncertainty of the corresponding region was low, so the salient object and background were detected relatively accurately. The second row is a case in which the data is not so difficult for humans to predict, but other methods have limitations in accurately predicting the data. Even in this case, the proposed method accurately detected the red box through sufficient learning of the corresponding data with low data uncertainty. The third row is a case that is difficult for the model to predict accurately due to a fine structure. Here, existing methods suffered from high prediction uncertainty and high prediction inaccuracy. On the other hand, the proposed method detected the corresponding region similarly to GT and estimated the prediction uncertainty low. This

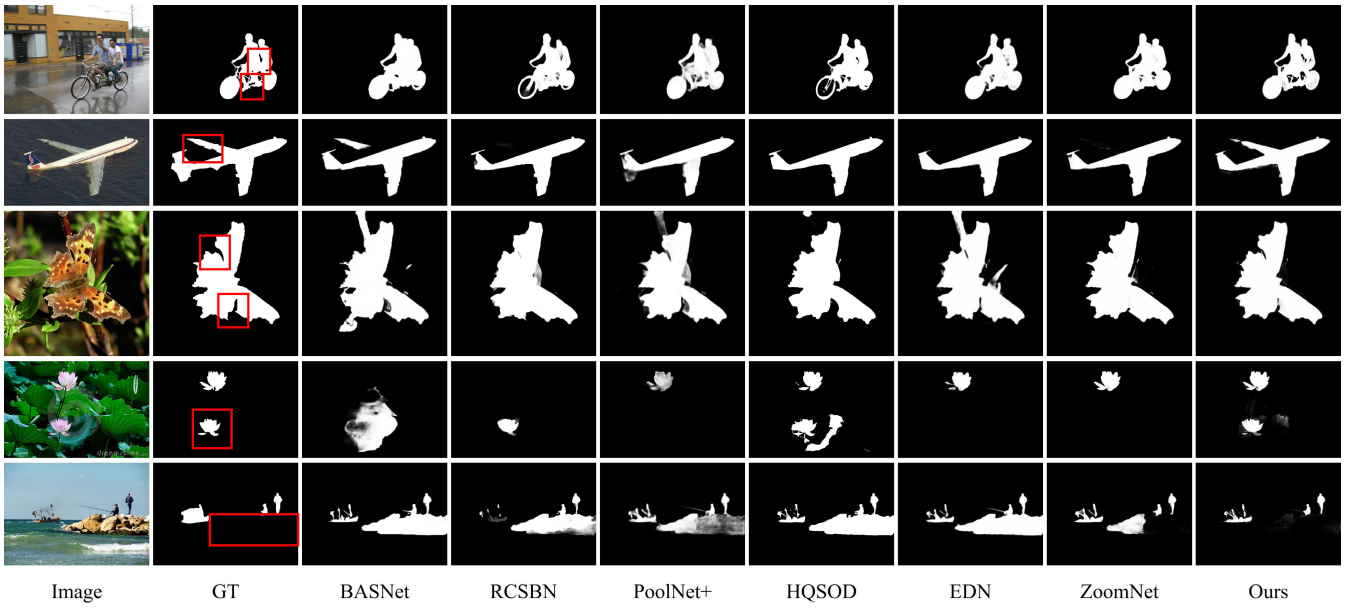

FIGURE 4. Qualitative test results.

supports that the learning and refinement of the proposed method are designed well. The fourth row is the case of detecting multiple salient objects. Most existing methods had low prediction accuracy because they detected only one object or estimated the noisy background around the salient object as a salient object. However, the proposed method had low data uncertainty and mainly learned data that can be predicted with certainty. Therefore, all salient objects were uniquely detected by the proposed method. The red box in the fifth row corresponds to the background. All of the existing methods detected it as a salient object or estimated the prediction uncertainty of the corresponding region highly. However, the proposed method effectively excluded region(s) with high data uncertainty and focused well on the actual salient object. Especially, it is noteworthy that the proposed method captures even the fine details of objects that are not well represented in GT. We compared performance on sample data from five datasets. Even in cases where it is difficult to distinguish between object and background, the proposed method distinguished salient objects with high accuracy. Additionally, because the proposed method considers areas with high data uncertainty, a tendency to detect salient objects with certainty rather than ambiguous areas is observed. As a result, the proposed method is expected to demonstrate the above advantages in tasks that require minimizing the rate of false negatives or false positives.

Through the above quantitative and qualitative comparison, we proved that the proposed uncertainty-aware learning and refinement makes it more effective to infer salient objects and capture fine details. In other words, the proposed method has strengths that are not well represented by quantitative indicators.

FIGURE 5. Comparison with and without model uncertainty-aware learning (MUAL) method. Each subcaption is the same as Table 2.

D. ABLATION STUDIES

1) MODEL UNCERTAINTY-AWARE LEARNING

This section verifies the performance according to the presence or absence of model uncertainty-aware learning. Figure 5 and Table 2 are the qualitative and quantitative results for the five datasets and DUT-OMRON dataset, respectively. Figure 5 (b) is a region where the model can sufficiently learn owing to low data uncertainty, but the predictive detection rate was remarkably low. On the other hand, in (c), the low data uncertainty region was well learned through the minimization of model uncertainty. In addition, data uncertainty was estimated more accurately than (b) through noise removal. In the quantitative aspect, a significant performance improvement was not achieved when only the existing data uncertainty-aware learning was used. However, note that applying the proposed model uncertainty-aware learning here improved the performance in all metrics.

TABLE 2. Quantitative results with and without model uncertainty-aware learning (MUAL) application. (a) Baseline (b) When MUAL is not applied in the proposed method (c) The proposed method applied with MUAL.

Method	DUTS-TE					DUT-OMRON					ECSSD				
	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$
(a) Baseline	**.899**	.853	.032	.928	.864	.836	.745	.051	.854	.758	.933	.924	.027	.957	.930
(b) Ours wo/ MUAL	.894	.857	.031	.929	.871	.833	.752	.049	.860	.767	.929	.926	.029	.955	.936
(c) Ours w/ MUAL	.895	**.860**	**.031**	**.929**	**.876**	**.840**	**.762**	**.048**	**.865**	**.778**	**.933**	**.931**	**.027**	**.959**	**.941**

Method	HKU-IS					PASCAL-S				
	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$
(a) Baseline	**.929**	.915	.026	.958	.920	.869	.834	.054	.913	.851
(b) Ours wo/ MUAL	.925	.918	.023	.959	.927	.869	.838	.054	.912	.854
(c) Ours w/ MUAL	.926	**.919**	**.023**	**.959**	**.930**	**.872**	**.843**	**.052**	**.916**	**.862**

TABLE 3. Quantitative results with and without uncertainty-aware learning refinement application. (a) Baseline as EDN (b) EDN with add on the proposed method.

Method	DUTS-TE					DUT-OMRON					ECSSD				
	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$
(a) EDN	**.890**	.841	.035	.923	.860	**.848**	.767	.050	.875	.785	**.927**	.918	.033	.952	.930
(b) EDN + Ours	.888	**.844**	**.034**	**.924**	**.867**	.845	**.770**	**.048**	**.879**	**.795**	.925	**.920**	**.031**	**.954**	**.938**

Method	HKU-IS					PASCAL-S				
	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$	$S_m \uparrow$	$F_\beta^\omega \uparrow$	MAE↓	$E_m \uparrow$	$F_m \uparrow$
(a) EDN	**.924**	.908	.026	.955	.919	.864	.825	.061	.903	.846
(b) EDN + Ours	.923	**.912**	**.024**	**.959**	**.924**	**.865**	**.827**	**.059**	**.907**	**.853**

TABLE 4. Performance comparison with and without uncertainty-aware refinement as well as θ in the DUTS-TE dataset. (a) A case considering only data uncertainty. Here, the equation of $\mathcal{R}$ is defined as Eq. 10. (b) Proposed refinement scheme.

Method	θ	S_m	F_β^ω	MAE	E_m	F_m
wo/ Refinement	-	.896	.854	.032	.926	.867
Refinement (a)	0.9	.891	.858	.031	.927	.875
	0.8	.873	.849	.033	.914	.874
	0.7	.846	.818	.038	.881	.854
Refinement (b)	0.9	**.897**	.859	.031	.929	.872
	0.8	.895	.860	**.031**	**.929**	**.876**
	0.7	.894	**.861**	.031	.929	.875

Such a qualitative evaluation result is demonstrated in Table 2. Comparing (a) and (b), the proposed method without MUAL showed a decrease in S-measure [36] compared to the baseline. In the ECSSD and PASCAL-S datasets, even E-measure [39] decreased slightly. Considering that S-measure and E-measure quantitatively evaluate structural similarity, we can find that the absence of MUAL causes to reduce structural similarity with GT. On the other hand, (c) with MUAL improved both S-measure and E-measure than (b). For DUT-OMRON, ECSSD, and PASCAL-S datasets, (c) scored higher than (a). For the remaining metrics, (c) showed better performance than (a) and (b).

Through performance analysis according to whether or not model uncertainty-aware learning was applied, we proved that consideration of model uncertainty can solve the insufficient learning issue of predictable data.

2) UNCERTAINTY-AWARE REFINEMENT

The proposed uncertainty-aware refinement masks both prediction and data uncertainties. So, we quantitatively demonstrate that considering both uncertainties is effective. First, refinement using only data uncertainty $\theta^{\mathcal{D}}$ is expressed as follows.

$$\mathcal{R} = \begin{cases} 0 & \text{if } \sigma_i^{\mathcal{D}} > \theta^{\mathcal{D}} \\ \hat{y}_i' & \text{otherwise} \end{cases} \tag{10}$$

Take a look at Table 4. When refinement was applied using only data uncertainty, performance drops were observed in most indicators. In particular, when a relatively low threshold of 0.7 was chosen, the overall performance was significantly degraded. That is, refinement does not work well. Instead, if prediction uncertainty is additionally considered here, performance is improved in most indicators, contrary to the previous result. Similar performance was observed when modulation is applied to the threshold, which means that the hyper-parameter sensitivity is low. In other words, the result

(b) of the proposed refinement technique considering both uncertainties shows high performance for the same threshold compared to the result (a) considering only data uncertainty. Additionally, it is observed that the proposed refinement is not sensitive to changes in threshold. As a result, the proposed method is proven to provide excellent performance even in various datasets and environments. Therefore, we can find that our strategy of additionally considering prediction uncertainty to solve the data uncertainty estimation error is effective.

3) MODEL GENERALITY

We show that the proposed method can be added on to conventional deep learning-based SOD networks to improve their performance. To verify this, we experimented with 5 datasets same as above by replacing the existing baseline ZoomNet [15] with EDN [43]. According to Table 3, 'EDN + Ours' to which the proposed method is applied to EDN shows quantitatively better performance than EDN, similar to IV-B. In particular, the performance of F_m was greatly improved. Each value in the table is the average of five experimental values. Note that the proposed method shows a significant performance improvement over the variability in the learning process. On the other hand, other techniques provide marginal performance improvement.

In terms of S_m, which is a metric that measures structural similarity, the performance of the proposed method was marginally degraded on both baselines. This might be due to a phenomenon in which some structural meanings differ from those of GT as pixels with low prediction uncertainty are forcibly removed when refinement is applied. However, note that the proposed method achieves very significant performance improvements in other quantitative/qualitative indicators except S_m.

V. CONCLUSION

This paper proposes a data and model uncertainty-aware method that considers both data and model uncertainty in the SOD task. We find that existing data uncertainty estimation methods have overlooked the fact that model uncertainty intervenes in prediction error, and propose a novel model uncertainty-aware learning to solve this problem. In addition, we present a refinement algorithm based on data uncertainty to reduce the prediction uncertainty due to unseen data input in the inference step. The proposed method achieved SOTA performance in most of the datasets and showed superior results qualitatively. Performance in low-contrast environments or with data where the distinction between object and background is ambiguous is still challenging. However, if techniques to decompose model uncertainty from data are further studied, it is expected that high performance can be achieved even with such challenging data. It is true that performance in low-contrast environments or on data where object and background classification is ambiguous is still challenging. However, if further research is conducted to decompose model uncertainty from data, it is expected that the proposed method will be able to achieve high performance even with such challenging data. Additionally, the generality of the proposed method was proven through quantitative and qualitative performance analysis on various datasets. On the other hand, the threshold method in real-world data should always be handled carefully. Therefore, when applying the proposed refinement method to real-world data, it needs to be accompanied by a parameter search optimized for each environment. Also, through intensive ablation studies, we show that the proposed method can be easily added to conventional SOD frameworks to further improve their performance.

REFERENCES

[1] H. Lee and D. Kim, "Salient region-based online object tracking," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jun. 2018, pp. 1170–1177.

[2] X. Qin, S. He, C. P. Quintero, A. Singh, M. Dehghan, and M. Jagersand, "Real-time salient closed boundary tracking via line segments perceptual grouping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 4284–4289.

[3] X. Qin, S. He, Z. Zhang, M. Dehghan, and M. Jägersand, "Real-time salient closed boundary tracking using perceptual grouping and shape priors," in *Proc. Brit. Mach. Vis. Conf.*, 2017.

[4] X. Qin, S. He, X. Yang, M. Dehghan, Q. Qin, and J. Martin, "Accurate outline extraction of individual building from very high-resolution optical images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 11, pp. 1775–1779, Nov. 2018.

[5] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 1915–1926, Oct. 2012.

[6] Y. Wei, J. Feng, X. Liang, M.-M. Cheng, Y. Zhao, and S. Yan, "Object region mining with adversarial erasing: A simple classification to semantic segmentation approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1568–1576.

[7] S. Hong, T. You, S. Kwak, and B. Han, "Online tracking by learning discriminative saliency map with convolutional neural network," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 597–606.

[8] A. Siris, J. Jiao, G. K. L. Tam, X. Xie, and R. W. H. Lau, "Scene context-aware salient object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 4136–4146.

[9] M. Zhuge, D.-P. Fan, N. Liu, D. Zhang, D. Xu, and L. Shao, "Salient object detection via integrity learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 3, pp. 3738–3752, Mar. 2023.

[10] J. Zhang, D. P. Fan, Y. Dai, S. Anwar, F. S. Saleh, T. Zhang, and N. Barnes, "UC-Net: Uncertainty inspired RGB-D saliency detection via conditional variational autoencoders," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 8582–8591.

[11] Y. Zeng, P. Zhang, Z. Lin, J. Zhang, and H. Lu, "Towards high-resolution salient object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7233–7242.

[12] A. Kendall and Y. Gal, "What uncertainties do we need in Bayesian deep learning for computer vision?" in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, p. 30.

[13] L. T. Bo Li, Y. Zhong, S. Ding, and M. Song, "Disentangled high quality salient object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 3560–3570.

[14] A. Li, J. Zhang, Y. Lv, B. Liu, T. Zhang, and Y. Dai, "Uncertainty-aware joint salient object and camouflaged object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 10071–10081.

[15] Y. Pang, X. Zhao, T.-Z. Xiang, L. Zhang, and H. Lu, "Zoom in and out: A mixed-scale triplet network for camouflaged object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 2160–2170.

[16] C. Pan, J. Liu, W. Q. Yan, F. Cao, W. He, and Y. Zhou, "Salient object detection based on visual perceptual saturation and two-stream hybrid networks," *IEEE Trans. Image Process.*, vol. 30, pp. 4773–4787, 2021.

[17] Y. Pang, X. Zhao, L. Zhang, and H. Lu, "Multi-scale interactive network for salient object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 9413–9422.
[18] Q. Chen, T. Liu, Y. Shang, Z. Shao, and H. Ding, "Salient object detection: Integrate salient features in the deep learning framework," *IEEE Access*, vol. 7, pp. 152483–152492, 2019.
[19] Y. Ji, H. Zhang, and Q. M. J. Wu, "Salient object detection via multi-scale attention CNN," *Neurocomputing*, vol. 322, pp. 130–140, Dec. 2018.
[20] J. Zhao, J.-J. Liu, D.-P. Fan, Y. Cao, J. Yang, and M.-M. Cheng, "EGNet: Edge guidance network for salient object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 8779–8788.
[21] L. Han, X. Li, and Y. Dong, "Convolutional edge constraint-based U-net for salient object detection," *IEEE Access*, vol. 7, pp. 48890–48900, 2019.
[22] Y. Umeki, I. Funahashi, T. Yoshida, and M. Iwahashi, "Salient object detection with importance degree," *IEEE Access*, vol. 8, pp. 147059–147069, 2020.
[23] Y. Y. Ke and T. Tsubono, "Recursive contour-saliency blending network for accurate salient object detection," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2022, pp. 2940–2950.
[24] Z. Chen, H. Zhou, J. Lai, L. Yang, and X. Xie, "Contour-aware loss: Boundary-aware learning for salient object segmentation," *IEEE Trans. Image Process.*, vol. 30, pp. 431–443, 2021.
[25] Z. Chen, Y. Zhan, and S. Gao, "Dual-branch network of information mutual optimization for salient object detection," *IEEE Access*, vol. 11, pp. 46120–46131, 2023.
[26] X. Tian, J. Zhang, M. Xiang, and Y. Dai, "Modeling the distributional uncertainty for salient object detection models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 19660–19670.
[27] Y. Lv, B. Liu, J. Zhang, Y. Dai, A. Li, and T. Zhang, "Semi-supervised active salient object detection," *Pattern Recognit.*, vol. 123, Mar. 2022, Art. no. 108364.
[28] S. Depeweg, J. M. Hernández-Lobato, F. Doshi-Velez, and S. Udluft, "Decomposition of uncertainty in Bayesian deep learning for efficient and risk-sensitive learning," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2018, pp. 1184–1193.
[29] B. Hou, Y. Liu, K. Qian, J. Andreas, S. Chang, and Y. Zhang, "Decomposing uncertainty for large language models through input clarification ensembling," 2023, *arXiv:2311.08718*.
[30] W. Wang, Q. Lai, H. Fu, J. Shen, H. Ling, and R. Yang, "Salient object detection in the deep learning era: An in-depth survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 6, pp. 3239–3259, Jun. 2022.
[31] L. Wang, H. Lu, Y. Wang, M. Feng, D. Wang, B. Yin, and X. Ruan, "Learning to detect salient objects with image-level supervision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 136–145.
[32] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3166–3173.
[33] J. Shi, Q. Yan, L. Xu, and J. Jia, "Hierarchical image saliency detection on extended CSSD," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 4, pp. 717–729, Apr. 2016.
[34] G. Li and Y. Yu, "Visual saliency based on multiscale deep features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5455–5463.
[35] Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille, "The secrets of salient object segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 280–287.
[36] D.-P. Fan, M.-M. Cheng, Y. Liu, T. Li, and A. Borji, "Structure-measure: A new way to evaluate foreground maps," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4548–4557.
[37] R. Margolin, L. Zelnik-Manor, and A. Tal, "How to evaluate foreground maps," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 248–255.
[38] F. Perazzi, P. Krähenbühl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 733–740.
[39] D.-P. Fan, C. Gong, Y. Cao, B. Ren, M.-M. Cheng, and A. Borji, "Enhanced-alignment measure for binary foreground map evaluation," 2018, *arXiv:1805.10421*.
[40] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1597–1604.
[41] X. Qin, Z. Zhang, C. Huang, C. Gao, M. Dehghan, and M. Jagersand, "BASNet: Boundary-aware salient object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7479–7489.
[42] Q. Zhai, X. Li, F. Yang, C. Chen, H. Cheng, and D.-P. Fan, "Mutual graph learning for camouflaged object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12997–13007.
[43] Y.-H. Wu, Y. Liu, L. Zhang, M.-M. Cheng, and B. Ren, "EDN: Salient object detection via extremely-downsampled network," *IEEE Trans. Image Process.*, vol. 31, pp. 3125–3136, 2022.
[44] J.-J. Liu, Q. Hou, Z.-A. Liu, and M.-M. Cheng, "PoolNet+: Exploring the potential of pooling for salient object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 887–904, Jan. 2023.

HEEJIN LEE (Associate Member, IEEE) received the B.S. and M.S. degrees in information and communication engineering, electrical and computer engineering from Inha University, Incheon, South Korea, in 2021 and 2023, respectively. Her research interests include computer vision and image processing.

SEUNGHYUN LEE (Associate Member, IEEE) received the B.S. and Ph.D. degrees in electronic engineering from Inha University, Incheon, South Korea, in 2017 and 2023, respectively. In 2023, he joined Upstage AI, as an AI Research Scientist. His research interests include model compression and lightweighting.

BYUNG CHEOL SONG (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 1994, 1996, and 2001, respectively. From 2001 to 2008, he was a Senior Engineer with the Digital Media Research and Development Center, Samsung Electronics Company Ltd., Suwon, South Korea. In 2008, he joined the Department of Electronic Engineering, Inha University, Incheon, South Korea, where he is currently a Professor. His research interests include image processing and computer vision.