IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 August 2023, accepted 13 September 2023, date of publication 25 September 2023, date of current version 29 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3319087

RESEARCH ARTICLE

Categorical Weighting Domination for Imbalanced Classification With Skin Cancer in Intelligent Healthcare Systems

VAN-DUNG HOANG[1], **XUAN-THUY VO**[2], **(Student Member, IEEE), AND KANG-HYUN JO**[2], **(Senior Member, IEEE)**

[1]Faculty of Information Technology, HCMC University of Technology and Education, Ho Chi Minh City 700000, Vietnam
[2]Department of Electrical, Electronic and Computer Engineering, University of Ulsan, Ulsan 44610, South Korea

Corresponding author: Van-Dung Hoang (dunghv@hcmute.edu.vn)

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.05-2021.04.

ABSTRACT In the field of dermatological diseases, especially for skin cancer, machine learning (ML) methods are used to classify melanoma and nevus using skin images. ML techniques result in high accuracy of diagnostic tasks since they are trained on balanced datasets. However, MLs working with imbalanced datasets produce erroneous results on precision, sensitivity, and specificity measured criteria. To deal with this problem, an augmentation approach combined with a category seesaw is used for the compensation factor. It increases the penalty for misclassified instances, thereby reducing the occurrence of false positives within the less common categories. This paper presents an approach to improve the efficiency of DCNN for classifying multi-class medical images on imbalanced datasets. The solution consists of three major contributions: (1) feature extraction based on some backbone models with customizing fully connected layers for classifier layers, (2) optimizing loss function (LF) and training parameters, (3) solving the problem of imbalanced samples using optimizing domination of weights between asymmetric classes with majority and minority categories. The method was evaluated and analyzed using the ISIC2018 benchmark and Chest X-ray dataset. Some well-known backbones were used for this study, e.g., EfficientNets, MobileNets, and DenseNets. The use of these backbones is to demonstrate that our methods are more efficient and stable in both light and heavy DCNN architectures. We also provide comparisons with existing methods that deal with the imbalance problem, e.g., data augmentation (AU), downsamples, customizing LF, and focal loss method (FL) for focusing on hard samples. Experimental results showed that these methods achieve good performance. However, there are several problems caused by generating new samples, and weighting samples, such as data overloading to train classifier models, a corrupt problem when applied to imbalanced data. Moreover, the FL method produced insufficient results on various DCNN backbones. Differently, our approach solves the imbalanced dataset based on boosting the sample weights of the minority and reducing the impact ratio of samples in majority categories. This strategy results in high precision and stable performance with various DCNN models without augmenting the dataset. Experiment results on ISIC2018 dataset demonstrated that our approach achieves more efficiency than other methods in some specific evaluation criteria as follows: higher than the FL method with 2.73% recall, 2.63% precision, 2.81% specificity, and 3.09% F1 using EfficientNet backbones; higher than AU method with 5.16% recall, 5.97% precision, 8.93% specificity, 6.16% F1 using DenseNet backbones.

INDEX TERMS Deep learning, feature extraction, imbalance data, DCNN, machine learning, disease diagnosis.

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

I. INTRODUCTION

Skin cancer is one of the most perilous diseases. According to recent reports, which illustrated that there are about more

0.1 million melanoma skin cancers, and 3 million non-melanoma skin cancers occur globally each year [1]. The correct diagnosis of the disease helps in effective treatment. Skin cancer disease can be predicted early which plays important most likely to be cured. Technology is continuously evolving, introducing new ideas in artificial intelligence and yielding innovative products that offer tangible benefits in the fields of medical and healthcare, and so on. Up to now, deep learning (DL) has shown remarkable success in various domains like medical diagnosis, healthcare, robotics, automation, and intelligent assistance systems. In particular, Deep Convolutional Neural Network (DCNN) based methods outperform traditional shallow learning methods, such as classical artificial neural networks (ANN), with their ability to handle complex tasks and large parameter sets.

One of the most essential problems affecting the capability of classification DL models is the problem of data imbalance between categories. That is, data samples belonging to the categories differ too much, leading to the classified model being biased in the majority of categories. Meanwhile, the misrecognition of the samples belongs to the minority categories. For instance, ISIC2018 dataset [3] is highly skewed in which sample numbers of the NV class are over-represented, i.e., this class appears very often (head class), while most of the other classes are under-represented, i.e., these classes appear more rarely. During training, the classified model is dominated by the NV class so that model features are not rich for other classes with smaller numbers. Therefore, the NV class performs better than other classes and it leads to degraded performance.

This paper contributes to customizing the loss estimated formulation, feature extraction backbone using DCNN architectures, and optimizing the fully-connected classifier, and extrinsic parameters to improve the efficiency. In our study, we propose a new approach for balanced domination classification on imbalanced datasets to harmonize the evaluation criteria of data categories, e.g. Accuracy (ACC), Precision (PRE), Recall (REC), Specificity (SPE), and F1. We have experimentally studied and conducted the effect of some approaches to improve the performance of disease recognition. Experimental results and analysis on imbalanced datasets demonstrate that this method reaches high robustness and stability, and balanced measurements in both kinds of light and heavy DCNN architectures.

II. RELATED WORKS

AlexNet [2] successfully trained the CNN model on large-scale datasets using GPU devices and achieved great improvements compared to classical models. According to this accomplishment, many researchers have introduced modern techniques to improve the performance of feature learning, efficiency, and optimization, such as residual connection [5], densely residual connection [6], multi-scale features [4], reducing model complexity by using depth-wise convolution [7], and searching optimal values of network depths and widths [8].

Recently, DCNN-based machine learning approaches have been widely acknowledged as state-of-the-art methods for medical image classification, for example, GoogleNet [4], Microsoft ResNet [5], DenseNet [6], MobileNet [7], and EfficientNet [8]. One notable area of development revolves around the selection of models. These approaches prioritize the unmanned selection of model types for recognition, eliminating the need for a specific default model. Classified models trained on specific datasets, these models improved accuracy. Furthermore, it enables the evaluation of data types and models, facilitating the automatic identification of the most suitable model for the task [9], [10]. In order to enhance the predictive capabilities of the system, some studies are exploring the potential of the CNN approach to learn and customize models that mimic the behavioral patterns of the human mind. This avenue of research aims to augment the system's prediction capacity and make more human cognitive processes. Several studies have been conducted on online tracking techniques that manipulate data to track objects and feature extraction [9], [11]. In this context, the contribution [12] suggests an adoption of adaptive learning in the object tracking process, which involves gathering suitable data and automatically retraining the recognition model. This proposed solution aims to improve the quality of automatic object recognition. CNN-based methods have gained widespread adoption in various industrial applications, including video surveillance systems [14]. These methods leverage the architecture of CNNs to enhance features through modified hourglass modules. Based on incorporating finer-resolution properties and utilizing lateral connections, these models produce more precise and accurate results. In intelligent agriculture application, authors in [13] use DCNN models for the classification of healthy and non-healthy classes of the oil palm tree disease dataset. The study demonstrated that the method based on fine-tuning the DenseNet121 model is the best-performing model.

In the domain of cancer disease diagnosis using medical images, dermoscopy emerges as a microscopic technique for capturing skin surface images. In an experimental setting, the ISIC2018 dataset was utilized and explored [3], [19]. This dataset comprises lesion images obtained from various dermatoscopy types, covering different anatomical sites and representing historical patients who underwent skin cancer screening across many organizations. Each image of lesions focuses on a single primary disease. The paper [20] provides an analysis of methods and validation outcomes from the ISIC Challenge 2018. The authors proposed a two-stage approach for segmenting lesion regions in medical images. They leverage an optimized training method and apply specific post-processing techniques. DCNNs have demonstrated outstanding performance in image recognition, surpassing human accuracy in certain cases involving large datasets. Another approach, a hybrid method was introduced to address the issue of class imbalance in skin disease classification [21]. The method combines both data-level and algorithm-level techniques to tackle the problem effectively. At the data level, a balanced mini-batch logic along with real-time image augmentation was employed to ensure a balanced representation of different classes during training. A new loss function is designed to further enhance the model's performance. Experimental results demonstrate that the EfficientNetB4 model achieves the highest accuracy among the compared models, including InceptionV3, ResNet50, and DenseNet169. This highlights the effectiveness of the hybrid approach in handling class imbalance and improving the overall performance of skin disease classification. Other studies have employed DCNNs for skin lesion classification [24], [26], [27], [30], [31], researches demonstrated that this topic still challenges persists due to limited data availability and imbalance issues. Furthermore, researchers have explored alternative approaches to medical image processing utilizing CNNs. For instance, lung cancer detection and classification have been addressed through enhanced Region-based RFCN with multilayer fusion RPN for automated decision-making [22]. Additionally, encrypted medical images have been subjected to a fast nearest neighbor search scheme based on CNNs [23]. In the study [25], Pham et al. presented a method for classifying skin lesions from dermoscopy images. The researchers conducted a comparison of classification results across six classifiers, such as support vector machine (SVM), random forest(RF), Logistic Regression, AdaBoost, Balanced FR, and others. They evaluated the performance of these classifiers in conjunction with seven different hand-crafted feature methods and four data preprocessing steps.

Several solutions have been proposed to address the challenge of imbalanced data, particularly in the field of object detection where there exists an imbalance between the foreground (minority) and background (majority) classes. Noteworthy solutions include Focal loss [32], Seesaw [33], BAGS [34], and EQL [35], which have demonstrated high accuracy in handling this issue. Focal loss [32] tackles the data imbalance problem by dynamically adjusting the contribution of samples to the classification loss based on their difficulty. It focuses on hard samples by assigning them higher weights and down-weights easy samples, thereby improving overall performance. In contrast, Seesaw loss [33] introduces a novel approach that balances the gradient magnitudes of both head and tail samples. Rather than directly manipulating sample weights, Seesaw loss aims to equalize the influence of different samples by adjusting the loss calculation, thus effectively addressing the data imbalance problem.

In contrast to existing approaches, we present a solution that utilizes boosted sample weights based on the relative sample rates across different categories. Our approach strengthens the weights of minority categories by an appropriate factor, ensuring that the classification model achieves balance while meeting evaluation criteria such as recall, accuracy (ACC), precision (PRE), specificity (SPE), and F1 score. Difference to the AU approach, which focuses on increasing samples of minority categories to achieve balance but may result in exceeding hardware system capacity or incurring high computational costs, our proposed approach achieves balanced influence by boosting minority categories with appropriate proportions. This leads to highly accurate classification results without the need to increase the number of training samples. The boosting approach effectively balances the capacity of the DCNN model. Additionally, we observe that the Focal loss (FL) method [32] is a well-known approach for improving the efficiency of classification models. Our experiments further demonstrate that both FL and the AU method contribute to improved results when used with certain DCNN models.

III. PROPOSED CLASSIFICATION MODELS

In this study, we aim to optimize the DCNN network architecture and customize the training parameters, and the loss formulation of models. Particularly, we investigate and customize the boosting mechanism that influences efficiency between data categories so that the classification model is trained to achieve as much balance as possible. DenseNet [6], MobileNet [7], EfficientNet [8]

A. OVERVIEW CLASSIFICATION SYSTEM

We present an overview of the classification architecture. Our research, instead of focusing on designing new DL architecture, is constructed using the popular CNN model combined with the customization of the loss function and fully connected layers for multiple-category classification. In this study, we focus on customizing some processing stages to improve the efficiency of classification systems. In the general flowchart of the training model illustrated in Fig. 1. In the architecture, there are three key stages to explore, such as feature extraction customization, classifier via the fully connected network, and error estimation with a loss function. Options for feature extraction include using pre-trained backbones or constructing DCNN architectures to find the best model. The resulting feature maps feed into the classification stage. To prove the efficiency of the proposed method, we evaluated some predefined DCNN backbones, which have been widely applied, such as DenseNets [6], MobileNets [7] and EfficientNets [8]. The backbones are representative of lightweight architectures such as MobileNets, more complex and dense connected architectures such as DenseNets, and another approach in very deep and wide architectures such as EfficientNets. In the classification phase, various methods are employed, including fully connected networks, SVM, and other ML techniques. Fully connected networks can be used for multi-classification tasks. In order to prevent the problem of overfitting, the dropout processing is incorporated into the ANN architecture. According to the previous mention, there is a hard problem with the classification task when applied to the imbalanced dataset. There are some solutions such as the AU method, and FL approach focusing on hard samples, and so on. This study has investigated and evaluated ISIC datasets using all the above approaches. The FL approach has been proven that is very useful for object detection problems however it is not stable with some backbones.

B. DESIGNING BACKBONES FOR FEATURE EXTRACTION

In this task, the traditional digital image processing (DIP) techniques can be applied to extract feature input to ML such as HOG, LBP, HOF,…In recent years, there are many approaches concentrated on DCNN techniques for feature extraction. The DCNN backbones are constructed following serial network architecture or directed graphs using common layers. In this paper, we have investigated using some types of feature extraction backbones with different numbers of parameters and architecture aspects to evaluate the influence of imbalanced handling solutions. We investigate the outstanding CNN architectures of DenseNets [6], MobileNets [7], and EfficientNets [8] with loss function, batch logic to ascertain a network in the training process. Obviously, these backbones can be defined using basic network architectures, or reused pre-trained models. To ensure objectivity, we used backbones that have been built and proven high capacity by the research communities. We evaluated multiple versions of its architecture to demonstrate comprehensiveness and suitability when applying it to solve the imbalanced dataset problem. The architecture of the MobileNet approach is known as a lightweight model that works efficiently on resource-limited systems. The transferred learning is applied to the pre-trained model. The feature map is extracted at the last layer of pre-trained model, named "multiply_56". In contrast, the DenseNet family of architecture represents densely connected convolutional networks that are more accurate and efficient. The transferred learning from the pre-trained model. The feature map is extracted at the last layer named "ReLU". They are deep and wide architectures with trainable parameters. The EfficentNet family represents large architectures with a balance between the depth and breadth of the network to achieve high efficiency. Especially, the EfficientNetB7 model has a very large number of parameters, requiring high hardware configuration. The EfficientNet network is transferred learning from the pre-trained model, the features map is taken at the last layer named "top_activation". Some backbones, the layer number, and the parameter details are demonstrated in Table 1.

C. CUSTOMIZING CLASSIFIER

Constructing an appropriate classifier aims to produce stable and more effective models in some application fields. In this subsection, we present the results of the customized fully connected network. This classified network is assembled in the tail of the feature extraction backbone. There are many ML techniques, which can be applied for the classified stage, e.g., SVM, RF, ANN, AdaBoost,…Recently, there are some solutions based on DSVM, Hybrid of SVM and ANN are proposed. In our experiment, some layers of a fully connected network are used for this purpose.

TABLE 1. The backbones of classified models are used for the experiment.

Backbone name	Number of layers	Number of parameters
MobileNet	90	3.757.255
MobileNetV2	158	2.917.447
MobileNetV3Large	280	4.885.895
MobileNetV3Small	247	2.058.359
DenseNet121	431	7.565.895
DenseNet169	599	13.498.951
DenseNet201	711	19.309.127
EfficientNetB0	241	4.709.034
EfficientNetB1	343	7.234.702
EfficientNetB4	478	18.595.430
EfficientNetB7	817	65.412.510

The classification block consists of some kinds of layers such as dense connection, global average polling, dropout, activation, and softmax layers. The dropout layer is used to prevent the overfitting occasion of the model. Therefore, this study concentrates on constructing the optimal network without underfitting due to small samples of some categories. Our method only focuses on boosting minority categories to balance the influence factor to the classifier without AU processing. The appropriate architecture is evaluated through trial and error approach. The architecture comprises two fully connected layers with 1,024 nodes and a next layer with 512 nodes, followed by an activation layer. The dropout layer with a ratio of 50% probabilities. The softmax activation function is used for the final output layer. This classification block combined with the feature extraction backbones was evaluated with some different loss functions.

IV. SOLUTIONS FOR IMBALANCED DATA PROBLEM

A. STATE OF THE ART SOLUTIONS

In classification literature, there are many approaches to solving this problem. Among them is the data augmented method, which augments data samples of minority categories by applying DIP techniques, for example, geometric transformation, artificial color transformation, and some other methods for concentration into samples of misclassification. The proposed learning architecture is appropriate to improve the performance of classified models in multiple skin diseases and solve the underfitting and overfitting problems.

■ The AU method: This approach is one of the simplest approaches to solving the imbalanced dataset problem. Besides that, some methods focus on making a balancing model for the classifier by customizing and optimizing the

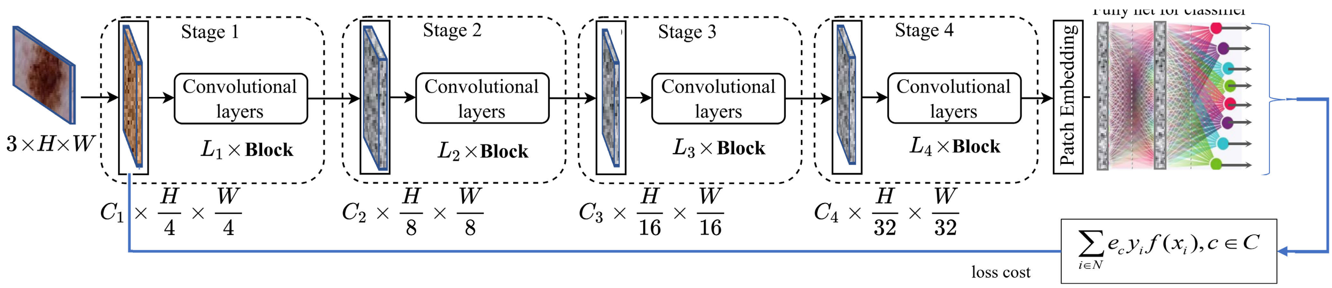

FIGURE 1. Training architecture of a DCNN-based classification model.

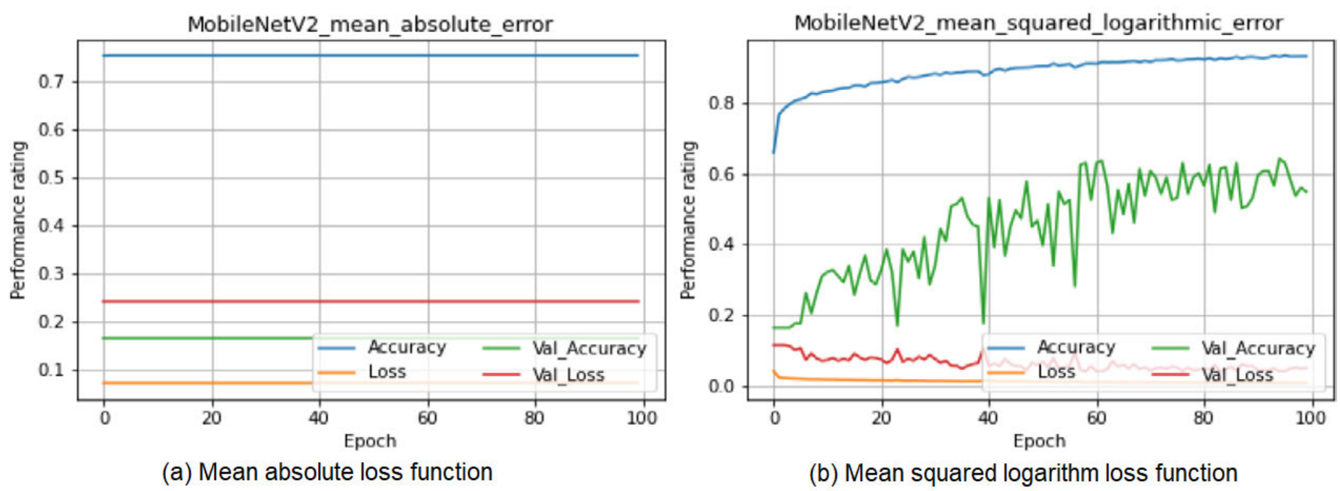

FIGURE 2. Some evaluation results of training progress with different LFs. Some LFs are not appropriately applied to imbalanced datasets.

loss functions. Meanwhile, other approaches try to customize the influent factor between each category in the dataset so that during processing the estimated errors are not biased towards some special categories. This approach aims to produce an efficient model without augmenting the dataset, meaning that the classification model works effectively and keeps a balance between the majority and minority categories. However, the practical applications also illustrated that each solution only achieves effective results in some particular application domains.

■ Customizing loss function: Generally, LFs are used to compute the quantity that the model should search and optimize the model parameters to minimize the error score during training progress. The LFs also have a great influence on classification results, especially in multi-class classification. Therefore, we also experimented with several different loss functions to select the best LF using trial and error methods. As a result, experiments show that DCNN using the cross-category entropy (CC) loss function reaches higher accuracy than other LFs such as mean absolute error, mean squared error, Poisson, and so on. In this study, we focus on solving the imbalance problem without augmenting the dataset. Previous research demonstrated that these loss functions work effectively on some practical benchmark datasets for classification and recognition tasks, such as Imagenet, CIFAR,…However, the practical evaluation shows that common loss functions do not provide good enough results with imbalanced datasets in some application domains. Some progress results of the training stage with different LFs are illustrated in Fig. 2. The amount of that categorical cross entropy LF is efficient for imbalanced dataset problems compared to other LFs. It is usually utilized for balancing and stable models in classification and recognition tasks.

In this study, some certain LFs used for multi-class classification have not resulted in lower effectiveness. One of the reasons is that the training task focuses on optimizing the parameters of the model for achieving the minimized loss cost on all datasets. This task aims to increase classification performance. This led to the seesaw problem. It means that the majority categories are more effectual meanwhile minority categories are less weighting to performance scores. In this subsection, we concentrate to present and analyze more details of the loss function using the categorical cross entropy metric (CC). The CC loss function is adopted in some existing frameworks briefly reviewed in [36] and [37]. The CC loss function has been widely used for multi-classification models which are two or more output categories. In this kind of LF, the output label is assigned to the one-hot category.

Generally speaking, CC LF is contributed for multiple classifications with categorical values of the confusion matrix. Since training labels are integer numbers, they are converted into categorical encoding. In the implementation, the Keras library supports LF named categorical-cross entropy. The classification block consists of a softmax activation layer and is followed by a CC measurement for multiple-category classification. The mathematical formulation of CC loss is written as follows:

$$H(p, q) = -\sum_{i \in C} y_i log\, p_i(x) \tag{1}$$

where x, y_i, and p_i are the input sample, target, and probability values in the output, respectively. C is a category domain with the equivalent of n classes of the model output. In this formula, the loss entropy value is smaller when the distributions get closer to each other.

Obviously, if the prediction value of the expected output with the high confident category reaches 1, then the probability of other categories is asymptotic to zero. During training processing, the samples are overfitting to a particular category meanwhile other ones reach zero, this value does not contribute to the loss score. Different from other LFs, the CC method is based on the probability of logarithm instead of the use of the distance measurement between the ground truth values and the predicted scores.

■ Focal loss solution: The original LF method was proposed by [32] to deal with the imbalanced dataset for object detection problems. The experimental results point out that the method significantly improves the precision of the detection models. In the case of object detection problems, there is an imbalance of data between the interesting objects and background. The loss function is estimated based on the idea of focusing on hard samples for balancing the ratio of sample weights. This approach is related to the performance of the one-stage detector in object detection problems. The focal loss can be computed following formulation with two impact parameters of α, γ.

$$F(q) = -\alpha_i (1 - q_i)^{\gamma} log\, q_i(x). \tag{2}$$

The label assignment in object detection is a binary classification problem, and the imbalance between positive and negative samples has appeared. The fact that the negative samples significantly outnumber the positive samples. Therefore, the cross entropy loss function often reduces accuracy for predicting bounding boxes of interesting objects. This task alters the influence of the sample on the loss score of the overall training dataset. The loss score distributes to the down-weight of the highly confident samples (called easy samples) and the up-weight of the miss classified samples (hard samples). In the formula of the FL method, the loss score with the multiplier of $(1 - q_i)^{\gamma}$, the balanced cross-entropy. However, the multiplier is significantly influenced by adjusting the impact ratio of each sample to the LF and gradient descent. During training progress, the trained model is simply to predict certainty the easy samples of majority categories The q_i probability of the easy sample is higher than the hard sample. Therefore, the factor $(1 - q_i)^{\gamma}$ is descended bias to a small value and negligible effecting on the loss scores of all training samples. In the case of hard samples, the q_i probability is descended. Therefore, the factor $(1 - q_i)^{\gamma}$ is increased asymptotically to one and greatly affects the loss score. As a result, the hard samples are more impact ratio than that of the easy samples. That means this modulating term approach applies to the cross-entropy loss in order to focus on hard negative examples during the training of the classified model.

B. PROPOSED BOOSTING INFLUENCE FACTOR OF CATEGORY

There are some solutions for boosting the multi-classification of imbalanced data problems. Recently, some boosting approaches are related to ML algorithms to deal with the problem of multi-class imbalanced data classification, as presented in [38]. The group of boosting algorithms focuses to improve the performance of the overall classifier by the combination of the set of individual weak classifiers, such as Adaboost, CatBoost, and LogitBoost. In our approach, we focus on boosting minority categories in imbalanced datasets to improve the performance of multiple classifications. Observably, the classification system usually biases to majority categories during the training processing due to the numerous samples, which are more effective to lose scores.

In this study, a new approach to solving the imbalanced data problem is investigated and presented. This approach is based on the ideal for balancing the impact ratio between samples in the majority and minority categories. The balanced solution is performed according to the affected coefficients of samples in each category to the total loss scores of a classifier. According to the fact that the influence coefficient between categories is linearly changed leading to over bias for some categories. This may cause failed models. The samples in each category have different prediction scores from the classified models. In medical image processing, there is a highly different level of mosaic distribution. Therefore, disease samples of the majority categories are more mosaic homology than the minority category. The DL techniques with a huge number of parameters are utilized for improving learning capability and avoiding overfitting to learn all training datasets. To deal with the imbalanced dataset, the boosting coefficient of each category is computed as the following formulation.

$$L_i = e^{-\upsilon\, p_i} \tag{3}$$

$$W_i = L_i + (1 - \max_{c \in C} L_c) \tag{4}$$

where υ and p_i are the influence ratio and probability distribution of samples to the category i^{th}. It is computed by the number of samples that belong to each category.

TABLE 2. Dataset details for experiment processing.

	MEL	NV	BCC	AKIEC	BKL	DF	VASC	Total
Training	890	5364	411	262	879	92	114	8012
Validation1	21	123	15	8	22	1	3	193
Validation2	223	1341	103	65	220	23	28	2003
Augmentation	5340	5364	5343	5240	5274	5336	5358	37255

Customizing an appropriate coefficient balances the weighting of minority and majority categories. In the training stage, these parameters are adjusted based on how to optimize the influence coefficients of the estimated loss scores and ground-truth scores. Thus, the influence ratios result in a balance between categories, which also produces more balancing of the precision, recall, specificity, f-score, and so on. Meanwhile, the AU stage aims to generate a new dataset by applying DIP methods, e.g., color normalization, geometry transformations, and artificial generation to make a balanced ratio of the sample number between minority and majority categories. In contrast, another balance approach is based on the down-sampling of majority categories. However, that approach does not utilize discriminated features for producing an efficient model due to removing the numerous training samples. That approach supports balancing the effective ratio of each category and results in efficient models in some particular applications, but it is not stable to be applied to a variety of feature extraction backbones. Therefore, it is not presented in this paper. Our boosting approach focuses on balancing the influenced ratio of all categories without increasing training samples. The method is more efficient and stable than the state of the arts.

V. EXPERIMENT AND EVALUATION

A. EXPERIMENTAL DATASET

The ISIC 2018 dataset is used for experiment and evaluation solutions for imbalanced datasets. This dataset consists of full train set, but small evaluation set, and additional testing set with ground truth labels. Therefore, there are only the training and validation datasets for experiment and comparison. In the original dataset, there are $10,015$ training samples and 193 evaluation samples. The dataset includes 7 disease categories: Actinic keratosis/ Bowen's disease (AKIEC), Basal cell carcinoma (BCC), Benign keratosis (BKL), Dermatofibroma (DF), Melanoma (MEL), Melanocytic nevus (NV), and Vascular lesion (VASC). The resolution image most is $450 \times 600 \times 3$ pixels. The ISIC training set is divided to 80% for training and 20% for validation as test dataset2, and the ISIC evaluation is used as validation dataset1. The dataset for this experiment is presented in Table 2.

The original validation set is used to evaluate and analyze a homologous characteristic and stable performance of the classifier on the training and evaluation dataset. The random method is used to separate the original training dataset into the training set and validation dataset2. These sets are fixed for all experiments and comparison processing with different methods and feature extraction backbones. The training and evaluation dataset is generated and fixed to ensure the objective property between methods. In this study, we only evaluate and analyze some methods for imbalanced data solutions on the same dataset and extrinsic parameters with a variety of DCNN backbones. We do not focus to compare with the results of other papers due to the data preprocessing, and parameter settings are different from our study, e.g., input image size, the number of epochs, and batch size.

1) DATA AUGMENTATION

In general, input image samples are preprocessed by using DIP methods. This research explores the influence factor of the augmented dataset on the performance ratio of classified models. The training dataset was augmented by color normalization, and affine transformation by rotation, flip, centering crop, skews, zoom. That means the new samples are generated by applying the DIP techniques to the original images for the new training dataset. These samples are generated anew from the original dataset for each epoch of training processing. The approach for data augmentation can be processed by the image data generators of different libraries such as TensorFlow and PyTorch. As a result, the training data is different when applied to each method, each training stage. The image data generator is also applied to avoid the overfitting problem. However, it may not objectively evaluate and compare because the training dataset is different for each time training a model. In our investigation, the AU is used as a method for dealing with the imbalanced data problem. It is compared to other state-of-the-art methods, such as the FL method and our approach. The AU method is utilized to make a balance of the dataset to improve the correct prediction. The main problem of the AU method generates a large training dataset, leading to excessive hardware demands and a substantial increase in computational time. The parameter setting to augment the training data is processed with randomized value of the rotation from -10 to 10 degrees, left-right and up-down flipping, change contrast value [0.7, 1], the width and heigh shifting from -10 to 10, shearing from -5 to 5 degrees, zoom in x, y direction from 0.8 to 1.

To provide furthermore results for assessing the proposed solutions, we conducted additional experiments on the Chest X-ray dataset [39]. This dataset exhibits an imbalance and contains fewer categories compared to ISIC2018. The most populous category is nearly three times the size of the least populous category. The dataset comprises 9,209 samples distributed across four categories as follows: Covid-19 (1,281 samples), Normal (3,271 samples), Pneumonia-Bacterial (3,001 samples), and Pneumonia-Viral (1,656 samples). The dataset has been normalized and resized to 224×224 dimensions. It was then randomly divided into distinct training and evaluation subsets with an 80% to 20% ratio respectively.

B. EVALUATION METRICS

There are some popular effectiveness measures the performance, such as Recall (REC), Accuracy (ACC), Precision (PRE), Specificity (SPE), F1, and so on. For the main task of comparison, we experimented with the multiple classifier problem on the ISIC dataset. Therefore, the measurement metric of the accuracy ratio is different from the binary classification problem. The accuracy is computed using the criterion of one versus all retain classes. That means the accuracy of each category, its samples are known as positive samples, and other retain samples as negative samples in binary classification formulation. As a result, the score of the accuracy criterion has a difference between binary and multiple categories. Other measurement metrics are the same as binary classifying problems. The measures of solution effectiveness are used as follows:

Accuracy (ACC) is the most intuitive performance measure, and it is a simplified ratio of correctly predicted samples to the total observations. The accuracy is a great measure but only when we have symmetric datasets where values of false positives and false negatives are almost the same. In the multiple classification tasks, the accuracy metric is computed by weighting the average of the accuracy ratios of all classes. The accuracy of the class i^{th} is computed as following form:

$$ACC_i = \frac{TP_i + TN_i}{Ns} \tag{5}$$

where Ns is the total number of samples in the dataset, where TP_i and FP_i are the number of true positive and false positive samples belonging to the class i^{th}, respectively; FN_i and TN_i are the number of false negative and true negative samples belonging to the class i^{th}, respectively. The number of negative samples N_i of class i^{th} is counted by the total sample Ns subtracted from the number of positive samples P_i.

$$ACC = \frac{1}{Ns} \sum_{i=1}^{c} n_i * ACC_i \tag{6}$$

where Ns is the number of samples in dataset, n_i is the number of samples of the class i^{th}.

There is another approach for computing the accuracy ratio of the multiple classification problem. In that approach, the accuracy of each class c^{th} is calculated by TP_c/total instances of the class c^{th}. This performance measurement is equivalent to the Recall ratio. Hence, the mentioned formula is used to calculate the accuracy rate.

Recall (REC) refers to the true positive rate, which is computed:

$$REC = TP/(TP + FN) \tag{7}$$

Precision (PRE) refers to the positive predictive value in both correction and mistake recognition as the following equation:

$$PRE = TP/(TP + FP) \tag{8}$$

Specificity (SPE) refers to the true negative rate, which is computed as follows:

$$SPE = TN/(TN + FP) \tag{9}$$

F1-Score (F1): It indicates the harmonic average of the precision and recall, which is computed as follows:

$$F1 = 2 * TP/(2 * TP + FP + FN) \tag{10}$$

C. EXPERIMENTAL RESULTS

The research conducted some experiment results and analyzed various methods for addressing imbalanced data problems. The methods included using loss functions based on category cross-entropy, FL, the AU method, and our proposed approach. Among these, the AU method incurs a higher computational cost for model training as it generates more substantial new samples to balance the training dataset. Additionally, we customized three types of feature extraction backbones: MobileNets, DenseNets, and EfficientNets. These backbones represent distinct approaches, with MobileNet being a compact architecture suitable for resource-constrained computing systems. The DenseNet backbone represents the densely connected network with a large number of trainable parameters. The EfficientNet backbone represents an approach to balancing CNN of the depth and breadth architecture. In each approach, we evaluated some versions of them with the smallest and largest architectures with different numbers of hidden layers and the number of trainable parameters. The details of CNN architectures are presented in Table 1.

1) EVALUATION RESULTS BASED ON MOBILENET BACKBONES

They are typically lightweight designs, containing just a few million trainable parameters. Despite their simplicity, they excel in accuracy across various applications, thanks to their efficient use of depth-wise separable convolutions. We evaluated four pretrained models as shown in Table 1. The experimental results on the evaluated dataset1 show that our proposed method reaches higher performance and more stable results compared to the AU method and other methods. Meanwhile, the results on the evaluated dataset2 show that the AU method is more efficient than our method. Finally, average results of both datasets, our method is the most efficient.

Fig. 3 shows the average of estimated results on dataset1 and dataset2 using the various feature extractors based on MobileNet models. The results illustrated that our method outperforms MobileNet and MobileNetV2 models in the REC, ACC, PRE, and F1 criteria. Meanwhile, the AU method is not stable to the SPE criterion. For the MobileNetV3Large and MobileNetV3Small, the FL, AU method, and our approach have similar quality of some measurements. The

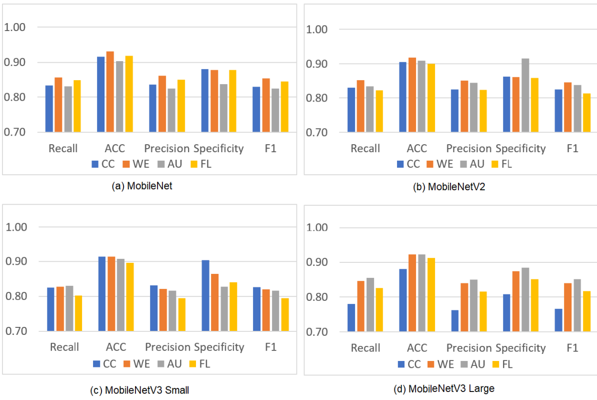

FIGURE 3. **Average on both evaluation datasets using MobileNet backbones.**

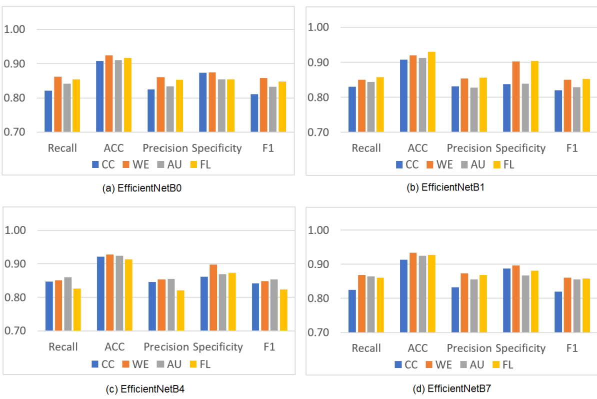

FIGURE 4. **Average results on dataset1 and dataset2 using EfficientNet backbones.**

CC method significantly increases the specificity score using MobileNetV3Small but it also largely reduces the effective score using MobileNetV3Large.

2) EVALUATION WITH EFFICIENTNET BACKBONES FOR FEATURE EXTRACTION

There are 4 versions of EfficientNets used for experiment and comparison. The EfficientNet architectures consist of more trainable parameters with very depth levels. The experimental results show that the AU, FL, and our method are better than the CC method in the most of measurement scores. Our method is the most efficient using EfficientNet7 and EfficientNet0 architectures. Meanwhile, the AU method has the best performance and the FL method has the worst results on EfficientNet4. In contrast, the FL method reaches the highest performance using EfficientNet1. In general, our method is more stable in comparison to other methods. Fig. 4 shows the average results on dataset1 and dataset2. The results illustrated that our method outperforms EfficientNetB7, EfficientNetB4, and EfficientNetB0 in almost all measurement metrics. Meanwhile, the FL method achieves the highest when using the EfficientNetB1 backbone but it is not stable for all feature extraction backbones. The CC method achieves the worst results in all measurement metrics and backbones.

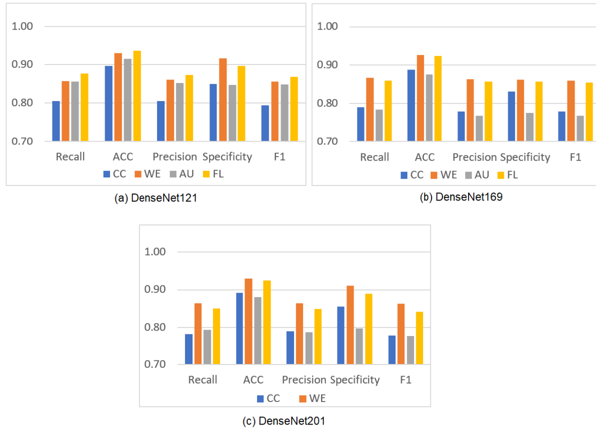

FIGURE 5. **Average of estimated results on dataset1 and dataset2 using DenseNets.**

3) EVALUATION WITH DENSENET BACKBONES FOR FEATURE EXTRACTION

We evaluated the proposed method and compared it to other approaches on 3 versions of DenseNets. The benefit is their ability to mitigate the vanishing-gradient problem. This problem occurs when gradients become too small during the backpropagation process, making it challenging for the network to learn effectively over many layers. DenseNets address this by creating dense connections between layers, ensuring that gradient information can flow more easily through the network. By using the DenseNet backbones, the experimental results demonstrate that the FL and the proposed method achieve higher scores than the AU and CC methods. In detail, our method is the best model using the DenseNet201 and DenseNet169 backbones. With DenseNet121, the FL reaches higher results for dataset1, but it is lower on dataset2. The interesting is that the AU method has the lowest results for all backbones. Particularly, the AU method using DenseNet201 and DenseNet169 is significantly lower accuracy (8% and more). The details of competitive results are shown in Fig. 5.

In summary, we have experimented to evaluate the performance using the backbone DCNNs with different approaches with different loss functions to deal with the problem of imbalanced classification data. MobileNets are small CNN architectures, DenseNets are dense connection CNN architectures and EfficientNets represents CNN architectures of balancing between depth and breadth level, which consists of a large number of layers and parameters. Experimental results show that the WE, AU, and LF approaches support improving classified performances for data imbalance problems. The AU method does not significantly improve performance scores with different backbones compared to other approaches. Meanwhile, it required high computational cost and hardware resources due to that generates more

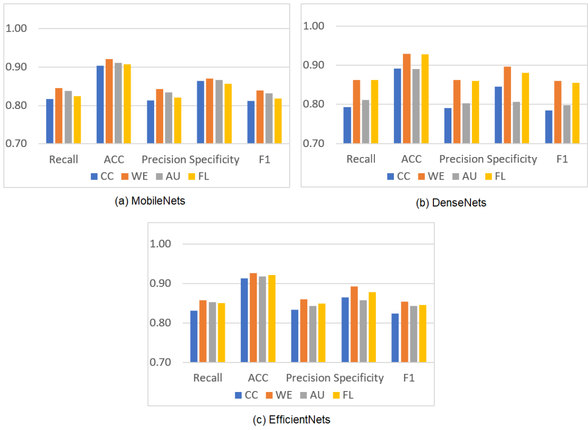

FIGURE 6. Summary classified results of the balancing solutions with the different backbones on the imbalanced dataset.

new samples. The FL method is a good approach to improve the classifier quality for the problem of imbalanced datasets. However, it does not stable with various feature extraction backbones and failed into a trap, e.g., MobileNets. In contrast, our approach focuses to customize the weight coefficients between categories with the appropriate factor, which supports a more balancing classified model with the CC loss function. Experimental results also point out that the classifier based on the balancing using our approach significantly improves the performance and fairly stable results with different backbone architectures. Fig. 6 shows the results of the CC, WE, AU, and LF methods based on DenseNets, MobileNets, and EfficientNets. The result values are estimated by the mean of evaluation products on all versions of each kind of CNN backbone and two validation datasets. Experiment results prove that our approach reaches the best results on most measurement metrics. The LF method is better than AU and CC methods with DenseNet and EfficientNet backbones. In contrast, the AU method is better than LF with MobileNet backbones. The studied results also show the AU and FL methods support improving efficient ratio however they are not stable with all backbones. The limitation of the proposed loss function is estimating appropriate the coefficient υ in formula (3) to calculate the category weighting is crucial and depends on each dataset. Therefore, a trial-and-error method is employed to estimate the value of υ for each particular dataset.

The experimental results on the Chest X-ray dataset reveal outcomes quite similar to those observed in the imbalanced dataset- ISIC2018. The experiments demonstrate that both the proposed solution and the augmentation method yield better results, which are compared to CC and FL on three different backbone models, as illustrated in Fig. 7. The data augmentation solution increases the number of training samples, which means higher computational costs for model training. On the other hand, the proposed approach supports enhancing classification precision without increasing the

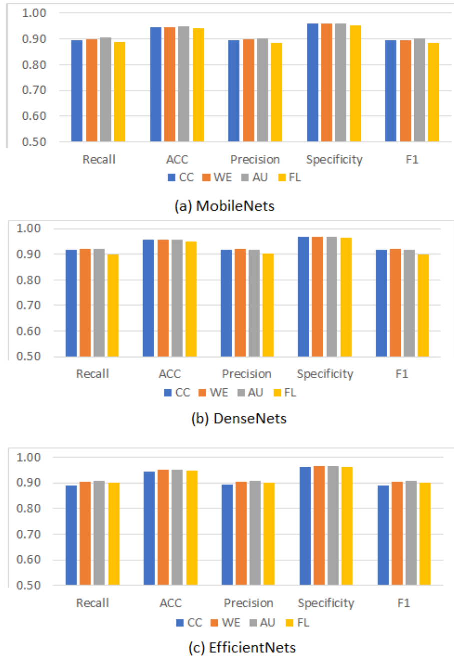

FIGURE 7. Average of estimated results on X-Ray dataset with the different backbones.

number of training samples, thus preventing a substantial rise in training computational expenses. Evaluation results on Chest X-ray dataset also consensus that the proposed method is utilized to improve accuracy without data augmentation and outperforms CC and FL loss functions. However, as above mention, the proposed solution requires the appropriate estimation of the parameter υ for improving the performance of the model.

VI. CONCLUSION

In this study, we introduced a novel approach for addressing imbalanced datasets without the need to create additional training samples. Our method focuses on equalizing the influence coefficient ratio between each category and the others, as opposed to the traditional methods of dealing with hard samples or generating new samples. Additionally, the CNN architecture is reformed and correctly customized by modifying fully connected layers and loss functions for multiple classification problems. The average experimental results on both evaluation datasets demonstrate that of follows: (1) The proposed method achieved the state of the art to other methods on some criteria of the (REC, ACC, PRE, SPE, F1). Evaluation on imbalanced dataset ISIC2018, using DenseNets backbone, our performance surpassed CC method by (6.99%, 3.67%, 7.11%, 5.05%, 7.50%) and AU method with (5.16%, 3.84%, 5.97%, 8.93%, 6.16%). On EfficientNets, our performance is higher CC and AU method with (2.73%, 1.36%, 2.63%, 2.81%, 3.09%), (0.54%, 0.82%, 1.71%, 3.59%, 1.14%), respectively; using MobileNets, our method is higher than the CC and FL method with (2.79%, 1.77%, 3.00%, 0.61%, 2.81%), and

(2.09%, 1.47%, 2.26%,1.30%, 2.21%), respectively. (2) Using DenseNet backbones combined with reconstructing fully connected layers to avoid overfitting problems, the proposed method reaches the highest results in REC, ACC, PRE, SPE, and F1 compared using EfficientNet and MobileNet backbones. The LF and AU methods are utilized to deal with the imbalanced problem of the dataset such as ISIC2018, Chest X-ray. However, these methods are not stable in some CNN models. In our approach, boosting the influence factor of minority categories with an appropriate weighting ratio supports improving performance without augmenting datasets, which requires expensive computational costs and hardware resources. Furthermore, the investigated results also show that the proposed approach is robust and stable with different backbones for feature extractions.

REFERENCES

[1] A. F. El-Yazbi, H. A. Khalil, T. S. Belal, and E. I. El-Kimary, "Inexpensive bioluminescent genosensor for sensitive determination of DNA damage induced by some commonly used sunscreens," *Anal. Biochem.*, vol. 651, Aug. 2022, Art. no. 114700.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 1106–1114.

[3] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC)," 2019, *arXiv:1902.03368*.

[4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[6] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

[7] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[8] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.

[9] L. Li and A. Talwalkar, "Random search and reproducibility for neural architecture search," in *Proc. Uncertainty Artif. Intell.*, 2020, pp. 367–377.

[10] H. Bertrand, R. Ardon, M. Perrot, and I. Bloch, "Hyperparameter optimization of deep neural networks: Combining hyperband with Bayesian model selection," in *Proc. Confrence sur lApprentissage Automatique*, 2017, pp. 1–5.

[11] T. Domhan, J. T. Springenberg, and F. Hutter, "Speeding up automatic hyperparameter optimization of deep neural networks by extrapolation of learning curves," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2015, pp. 1–9.

[12] D.-P. Tran and V.-D. Hoang, "Adaptive learning based on tracking and ReIdentifying objects using convolutional neural network," *Neural Process. Lett.*, vol. 50, no. 1, pp. 263–282, Aug. 2019.

[13] Y. H. Haw, Y. C. Hum, J. H. Chuah, W. Voon, S. Khairunniza-Bejo, N. A. Husin, P. L. Yee, and K. W. Lai, "Detection of basal stem rot disease using deep learning," *IEEE Access*, vol. 11, pp. 49846–49862, 2023, doi: 10.1109/ACCESS.2023.3276763.

[14] V.-T. Hoang, D.-S. Huang, and K.-H. Jo, "3-D facial landmarks detection for intelligent video systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 578–586, Jan. 2021.

[15] C. Huang, S. Lucey, and D. Ramanan, "Learning policies for adaptive tracking with deep feature cascades," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 105–114.

[16] M. Long, Y. Cao, Z. Cao, J. Wang, and M. I. Jordan, "Transferable representation learning with deep adaptation networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 12, pp. 3071–3085, Dec. 2019.

[17] N. Q. K. Le, T.-T. Huynh, E. K. Y. Yapp, and H.-Y. Yeh, "Identification of clathrin proteins by incorporating hyperparameter optimization in deep learning and PSSM profiles," *Comput. Methods Programs Biomed.*, vol. 177, pp. 81–88, Aug. 2019.

[18] J. Snoek, O. Rippel, K. Swersky, R. Kiros, N. Satish, N. Sundaram, M. Patwary, M. Prabhat, and R. Adams, "Scalable Bayesian optimization using deep neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2171–2180.

[19] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, no. 1, Aug. 2018, Art. no. 180161, doi: 10.1038/sdata.2018.161.

[20] C. Qian, T. Liu, H. Jiang, Z. Wang, P. Wang, M. Guan, and B. Sun, "A detection and segmentation architecture for skin lesion segmentation on dermoscopy images," 2018, *arXiv:1809.03917*.

[21] T.-C. Pham, A. Doucet, C.-M. Luong, C.-T. Tran, and V.-D. Hoang, "Improving skin-disease classification based on customized loss function combined with balanced mini-batch logic and real-time image augmentation," *IEEE Access*, vol. 8, pp. 150725–150737, 2020.

[22] A. Masood, B. Sheng, P. Yang, P. Li, H. Li, J. Kim, and D. D. Feng, "Automated decision support system for lung cancer detection and classification via enhanced RFCN with multilayer fusion RPN," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7791–7801, Dec. 2020.

[23] C. Guo, S. Su, K. R. Choo, and X. Tang, "A fast nearest neighbor search scheme over outsourced encrypted medical images," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 514–523, Jan. 2021.

[24] H. L. Gururaj, N. Manju, A. Nagarjun, V. N. M. Aradhya, and F. Flammini, "DeepSkin: A deep learning approach for skin cancer classification," *IEEE Access*, vol. 11, pp. 50205–50214, 2023, doi: 10.1109/ACCESS.2023.3274848.

[25] T. C. Pham, G. S. Tran, T. P. Nghiem, A. Doucet, C. M. Luong, and V.-D. Hoang, "A comparative study for classification of skin cancer," in *Proc. Int. Conf. Syst. Sci. Eng. (ICSSE)*, Jul. 2019, pp. 267–272.

[26] P. Tschandl, C. Rosendahl, B. N. Akay, G. Argenziano, A. Blum, R. P. Braun, H. Cabo, J. Y. Gourhant, J. Kreusch, A. Lallas, and J. Lapins, "Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks," *JAMA Dermatol.*, vol. 155, no. 1, pp. 58–65, 2019.

[27] T.-C. Pham, C.-M. Luong, V.-D. Hoang, and D. Antoine, "AI outperformed every dermatologist: Improved dermoscopic melanoma diagnosis through customizing batch logic and loss function in an optimized deep CNN architecture," *Scientific Reports vol.* , vol. vol. 11, p. 17485, pp. 1–9, 2021.

[28] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.

[29] H. Yi, Q. Jiang, X. Yan, and B. Wang, "Imbalanced classification based on minority clustering synthetic minority oversampling technique with wind turbine fault detection application," *IEEE Trans. Ind. Informat.*, vol. 17, no. 9, pp. 5867–5875, Sep. 2021.

[30] T. J. Brinker, A. Hekler, A. Hauschild, C. Berking, B. Schilling, A. H. Enk, S. Haferkamp, A. Karoglan, C. von Kalle, M. Weichenthal, E. Sattler, D. Schadendorf, M. R. Gaiser, J. Klode, and J. S. Utikal, "Comparing artificial intelligence algorithms to 157 German dermatologists: The melanoma classification benchmark," *Eur. J. Cancer*, vol. 111, pp. 30–37, Apr. 2019.

[31] T.-C. Pham, C.-M. Luong, M. Visani, and V.-D. Hoang, "Deep CNN and data augmentation for skin lesion classification," in *Proc. Asian Conf. Intell. Inf. Database Syst.*, 2018, pp. 573–582.

[32] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[33] J. Wang, W. Zhang, Y. Zang, Y. Cao, J. Pang, T. Gong, K. Chen, Z. Liu, C. C. Loy, and D. Lin, "Seesaw loss for long-tailed instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 9690–9699.

[34] Y. Li, T. Wang, B. Kang, S. Tang, C. Wang, J. Li, and J. Feng, "Overcoming classifier imbalance for long-tail object detection with balanced group softmax," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 10988–10997.

[35] J. Tan, C. Wang, B. Li, Q. Li, W. Ouyang, C. Yin, and J. Yan, "Equalization loss for long-tailed object recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11659–11668.

[36] Y. Luo, M. Jiang, and Q. Zhao, "Visual attention in multi-label image classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 820–827.

[37] A. Rusiecki, "Trimmed categorical cross-entropy for deep learning with label noise," *Electron. Lett.*, vol. 55, no. 6, pp. 319–320, Mar. 2019.

[38] J. Tanha, Y. Abdi, N. Samadi, N. Razzaghi, and M. Asadpour, "Boosting methods for multi-class imbalanced data classification: An experimental review," *J. Big Data*, vol. 7, no. 1, p. 70, Dec. 2020.

[39] U. Sait, K. G. Lal, S. Prajapati, R. Bhaumik, T. Kumar, S. Sanjana, and K. Bhalla, "Curated dataset for COVID-19 posterior-anterior chest radiography images (X-Rays)," Mendeley Data, 2020, vol. 1. [Online]. Available: https://data.mendeley.com/datasets/9xkhgts2s6/1

VAN-DUNG HOANG received the Ph.D. degree in electrical and computer engineering from the University of Ulsan, Ulsan, South Korea, in 2015. After a year of experience with Telecom Sudparis as a Postdoctoral Research Fellow, he joined the Faculty of Information Technology, HCMC University of Technology and Education, Ho Chi Minh City, Vietnam, where he is currently a Professor and the Head of the Artificial Intelligence Department. He has a total academic teaching experience of 18 years with more than 85 publications in reputed international conferences, journals, and book chapter contributions (Indexed By: ISI, WoS, Scopus, ACM DL, and DBLP). His research interests include computer vision, robotics, medical image processing, autonomous vehicle, and ambient intelligence.

XUAN-THUY VO (Student Member, IEEE) received the B.S. degree in electrical and electronic engineering from the University of Science and Technology—The University of DaNang, Da Nang, Vietnam, in 2018. He is currently pursuing the combined M.S. and Ph.D. degrees with the Department of Electrical, Electronic and Computer Engineering, University of Ulsan, South Korea. His current research interests include computer vision and deep learning, with a focus on object detection, object segmentation, and multiple people tracking.

KANG-HYUN JO (Senior Member, IEEE) received the Ph.D. degree in computer-controlled machinery from Osaka University, Japan, in 1997. After a year of experience with ETRI as a Postdoctoral Research Fellow, he joined the School of Electrical Engineering, University of Ulsan, Ulsan, South Korea, where he is currently the Faculty Dean. His current research interests include computer vision, robotics, autonomous vehicles, and ambient intelligence. He was the Director or an AdCom Member of the Institute of Control, Robotics, and Systems (ICROS), the Society of Instrument and Control Engineers (SICE), and IEEE Industrial Electronics Society (IEEE IES) Technical Committee on Human Factors Chair. He was an AdCom member and a secretary, until 2019. He is an Editorial Board Member of international journals, such as the *International Journal of Control, Automation, and Systems* and *Transactions on Computational Collective Intelligence*. He has also been involved in organizing many international conferences, such as the International Workshop on Frontiers of Computer Vision, International Conference on Intelligent Computation, International Conference on Industrial Technology, International Conference on Human System Interactions, and Annual Conference of the IEEE Industrial Electronics Society.