Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

Handling imbalanced class in melanoma: Kemeny–Young rule based optimal rank aggregation and Self-Adaptive Differential Evolution Optimization

Gaurav Srivastava, Nitesh Pradhan *

Department of Computer Science and Engineering, Manipal University Jaipur, 303007, Rajasthan, India

ARTICLE INFO

ABSTRACT

Melanoma, despite its relatively low incidence compared to other types of skin cancer, accounts for a significant proportion of skin cancer-related deaths. Early detection of melanoma is crucial for improving patient survival rates. Deep learning algorithms, which heavily rely on data, are widely used in melanoma detection. However, the performance of these algorithms is greatly influenced by the distribution of the dataset, particularly class imbalance. In this manuscript, the authors present a novel method based on the Kemeny–Young rule for optimal rank aggregation to address the class imbalance problem in melanoma detection. The proposed approach aims to reduce class bias and enhance overall classification accuracy. Furthermore, a cost-sensitive learning approach is introduced to improve the classifier's ability to handle class imbalance effectively. This novel cost-sensitive learning method utilizes Self-Adaptive Differential Evolution Optimization to determine optimal weights for each class. Our approach differs from traditional methods that assign weights based on predefined criteria. To evaluate the effectiveness of the proposed methods, extensive experiments and ablation studies are conducted on the highly imbalanced ISIC 2020 dataset, which is widely used in melanoma detection research. The Kemeny–Young rule-based majority voting achieves an overall error rate of 2.44%, while the cost-sensitive learning based on the Self-Adaptive Differential Evolution approach achieves an even lower error rate of 1.99%. Moreover, the proposed method achieves a sensitivity of 87.93% and a specificity of 98.19%. These experimental results demonstrate the competitiveness and effectiveness of the proposed methods in addressing the challenges posed by class imbalance and improving the accuracy of melanoma detection. By effectively mitigating class imbalance, these methods improve the accuracy and reliability of melanoma detection, thus offering valuable insights for developing advanced computer-aided diagnosis systems in dermatology. The relevant codes for our proposed approach are publicly available at: https://github.com/ctrl-gaurav/Handling-Imbalanced-Class-in-Melanoma.

1. Introduction

Melanoma is a type of skin cancer that has become increasingly widespread in the Western world. Despite being the least prevalent skin cancer, Melanoma is responsible for 75% of skin cancer fatalities (Jain et al., 2015). It is a cancer of melanocytes, which are neuroectodermal cells that create the pigment and can be found in the eye, skin, throat, nose, and rectum, among other places (Mayo Clinic, 2022). It can appear anywhere on the body and is most common in areas with a lot of sun exposure. However, persons with dark skin are likelier to have concealed melanomas in areas with little sun exposure, such as their palms, fingernails, and the soles of their feet (American Cancer Society, 2022). Melanoma is usually discovered when a patient is concerned about a growth on their skin (Skin Cancer Foundation, 2022). This might be anything from a mole to many forms of lesion problems. Only a close inspection can identify whether or not the anomaly is carcinogenic; if it is, additional testing can indicate whether it is Melanoma or not (NHS, 2022).

Even though Melanoma is a very lethal illness, it offers a severe threat to those affected. A substantial majority of its patients survive, given that they were discovered early in the disease's progress (Knackstedt et al., 2018). This is conceivable because our immune systems can fight this type of cancer, and modern therapies such as immunotherapy have dramatically improved survival rates (Nikolouzakis et al., 2020; Rajaram et al., 2010). This makes the melanoma diagnostic procedure even more critical, with the potential to save many lives in the future since the impact of solar radiation will only intensify as a result of heinous human acts against the earth's climate parameters (Kato et al., 2019; Fargnoli et al., 2012). With this in mind, several researchers have sought to use information from numerous disciplines to identify potential and rapid diagnostic treatments for Melanoma (Sánchez-Monedero et al., 2018; Yu et al., 2020).
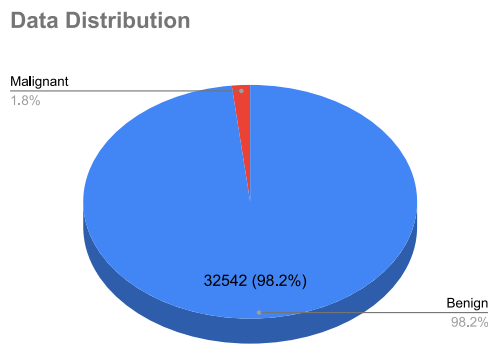
Data Distribution

Fig. 1. Pie chart depicting severe class imbalance in ISIC-2020 database.

Deep Learning (DL) is one such area that is a strong contender for automating the diagnosis of any medical ailment (LeCun et al., 2015; Srivastava et al., 2023a). It is a Machine Learning approach that uses neural networks, similar to human brains, to assist computers in learning and functioning intelligently. With body scans and images, deep learning researchers have made significant progress recognizing various cancer forms (Li et al., 2010). This aids in early illness detection, increasing the odds of survival exponentially (Smaoui and Bessassi, 2013). Unfortunately, in the case of Melanoma, the diagnostic process is currently unable to reach patients locally worldwide (Lallas et al., 2013). Instead, they are forced to undergo extensive examinations by highly qualified doctors and numerous body tests, which cost significant time and money in the diagnosis process alone (Aspinwall et al., 2008; Adepu et al., 2023).

Since deep learning is based on data, the data itself is sometimes the source of its limits. The model demands that the classes be perfectly balanced and the data be plentiful enough for the model to acquire those features quickly (Johnson and Khoshgoftaar, 2019; Pradhan et al., 2020, 2021). The leading cause of concern in the Melanoma dataset was class disparity. Melanoma-infected patients had a significantly lower number of images than non-melanoma-infected patients, as shown in Fig. 1. So, while training the model on this dataset, the model quickly picks up on the majority-class features and correctly classifies them. However, it struggles to distinguish between minority class samples. Because the minority class traits were not present at sufficient levels to accurately differentiate between them during the learning phase, the model typically classifies all those minority class samples to the majority class (Dong et al., 2018).

The dataset is extremely hard to discover every time a new disease emerges. For example, recently, when the COVID-19 pandemic hit worldwide, we had significantly fewer no. of COVID-19 infected images in the dataset (Oh et al., 2020; Srivastava et al., 2022b). The issue of class imbalance has been the subject of several research projects in the past (Wang et al., 2021; Khan et al., 2022; Tan et al., 2019). Researchers are continually addressing this subject of class inequality (Srivastava et al., 2022a; Chamseddine et al., 2022; Chakraborty et al., 2021; Sun et al., 2022). Various established methods such as oversampling, undersampling, and Synthetic Minority Oversampling Technique (SMOTE) have been developed to tackle class imbalance issues. Despite their respective merits and drawbacks, these strategies have been applied across various fields, such as medical imaging, which often involves a scarcity of afflicted individuals. In the realm of deep learning-based models, the issue of imbalanced data has been approached through techniques such as data sampling, class weighting, data augmentation, etc. (Saini and Susan, 2020). The authors have examined those strategies in this manuscript and briefly analyzed their benefits and drawbacks.

The present study introduces a novel approach to tackling class imbalance concerns. The authors' methodology entails utilizing the Kemeny–Young rule to aggregate the rankings of DCNN classifiers

efficiently. Specifically, distinct classifiers are employed to identify and differentiate between various dataset segments. In deep learning, ensemble learning algorithms have become popular for consolidating multiple classifiers' strengths to train an ensemble model that surpasses their individual performances (Chen et al., 2019). To address the class imbalance issue, it is common practice to assign different weights to various classes; however, determining the most effective weights presents a formidable challenge (Ren et al., 2018). Typically, weights are computed by dividing the number of class samples by the product of the number of classes and the class samples, yielding satisfactory outcomes. Nevertheless, alternative weight combinations can be employed to achieve superior results. The authors comprehensively analyzed various techniques to determine optimal weight values, detailing their respective merits and limitations.

Furthermore, the authors proposed a meta-heuristic approach called "weighted class training with Self-Adaptive Differential Evolution Optimization" as their second technique. Meta-heuristic algorithms, including evolutionary, genetic and swarm-based, are a few alternatives accessible for optimization (Salcedo-Sanz et al., 2014; Chopard and Tomassini, 2018; Maier et al., 2019). Evolutionary algorithms, in particular, are a global optimization approach that can handle higher-dimensional problems. As a result, an evolutionary algorithm-based optimization technique was selected in this study (Bozorg-Haddad et al., 2017). These algorithms can deal with evaluation functions that may not produce optimal results within a specific time frame, and they are resistant to noise in the evaluation process (Maier et al., 2014). Furthermore, they can be customized and tailored to specific issues. However, determining which evolutionary approach is ideal for a particular task has been the subject of numerous studies.

The proposed techniques exhibited superior accuracy and class imbalance mitigation capabilities compared to existing state-of-the-art methods. This accomplishment can be deemed a valuable contribution to addressing the challenge of class imbalance and boosting classifier performance in biomedical imaging. In addition, by removing the need for a first checkup and using a computerized test for a skin issue to determine whether the patient is likely to have Melanoma, this proposed model may be put into use practically.

Our research's primary contributions are:

1. A novel approach based on the Kemeny–Young rule was introduced for majority voting to address class biases. The authors leveraged this rule to optimally aggregate the ranks of DCNN classifiers, thereby enhancing the cumulative classifier performance.
2. A novel cost-sensitive learning method utilizing Self-Adaptive Differential Evolution optimization was proposed to determine optimal class weights for cost-sensitive learning, effectively mitigating the imbalance class problem. The authors extensively compared this approach with the conventional class weights calculation method to demonstrate its superiority.
3. The proposed approach has experimented with six different DCNN classifiers backbone. Multiple ablation studies were conducted, including the utilization of both standard loss function and modified loss function incorporating alpha-balanced focal loss, to ensure the robustness of the proposed methodology.
4. The authors conducted a comprehensive investigation of various strategies for mitigating class imbalance issues in melanoma detection. A comparative study was performed to analyze the effectiveness of existing methods compared to the proposed methodology.

The remaining study materials may be summed up as follows. Section 2 discusses the past researchers' work for melanoma detection and their strategies to address class inequality. Section 3 covers the theoretical background of the study. The proposed methods and the presented methodology in this study are described in Section 4. Finally, the experimentation and the outcomes of the aforementioned experiments are disclosed and covered in Section 5.

2. Related work

Melanoma, a prevalent disease in contemporary times, has been extensively studied with regard to early detection and simple therapies. To gain a comprehensive understanding of the issue, the authors conducted a meticulous examination of previous significant contributions made by numerous medical specialists, scientists, and researchers. Based on the literature review, various methods and approaches have been proposed in the past, which have been thoroughly investigated for their merits and limitations. These findings provided valuable insights and informed the authors' novel approach to the problem of melanoma diagnosis. In this study, the works analyzed are briefly summarized.

2.1. Imbalanced classification approaches

In the domain of machine learning and data mining, addressing imbalanced classification poses a significant challenge, necessitating the utilization of cost-sensitive learning strategies such as weighted extreme learning machines (W-ELM) to mitigate biases towards majority classes (Ling and Sheng, 2008). However, the effectiveness of W-ELM may be limited due to suboptimal weight assignments based solely on empirical costs. To overcome this limitation, Li et al. proposed a novel approach known as MOAC-ELM (Multi-Objective Optimization-based Adaptive Class-Specific Cost Extreme Learning Machine) for solving classification problems (Li et al., 2022). The primary objective of the MOAC-ELM method is to tackle the inherent issues associated with imbalanced datasets by incorporating class-specific information into the initial weight assignment process. By considering the unique characteristics of each class, the algorithm enhances the representation of minority classes through the introduction of penalty factors. Additionally, the authors employed an ensemble strategy to make optimal decisions after the optimization process. To evaluate the effectiveness of MOAC-ELM, extensive experiments were conducted on various benchmark datasets, as well as a real-world application dataset. The experimental results provided compelling evidence for the robustness and reliability of the proposed MOAC-ELM method.

While an extreme learning machine (ELM) is renowned for its rapid execution and superior generalization capabilities, it often lacks effective strategies for handling imbalanced datasets (Huang et al., 2011). Recognizing this limitation, Xiao et al. introduced a novel ELM model called Class-Specific Cost Regulation Extreme Learning Machine (CCR-ELM), along with its kernel-based extension (Xiao et al., 2017). CCR-ELM was specifically developed to address binary and multiclass classification problems characterized by imbalanced data distributions. The key contribution of CCR-ELM lies in the introduction of class-specific regulation costs for misclassification, effectively balancing the trade-off between structural risk and empirical risk. By assigning distinct regulation costs to different classes, CCR-ELM aims to alleviate the challenges associated with imbalanced data distributions. To evaluate the performance of CCR-ELM, extensive experiments were conducted using various benchmark datasets, as well as a real-world application dataset related to blast furnace status diagnosis. The experimental findings unequivocally demonstrated the superiority of CCR-ELM over the original ELM method and other existing ELM-based approaches specifically designed for imbalanced learning scenarios.

2.2. Deep learning based approaches for melanoma classification

Li et al. developed two deep learning models, named the Lesion Indexing Network (LIN) and Lesion Feature Network (LFN), to tackle the challenges of skin lesion image processing, including lesion classification, segmentation, and feature extraction from dermoscopy images (Li and Shen, 2018). They trained and evaluated these models on the ISIC 2017 dataset. In addition, two Fully Convolutional Residual Networks (FCRN-88) have been trained on subsets of the same dataset for coarse lesion classification and segmentation purposes (Codella et al., 2018).

To enhance the accuracy of lesion classification, the authors proposed using a Lesion Indexing Calculation Unit (LICU) that computes the significance of individual pixels. The FCRN-88 models were refined using a distance map generated by LICU. Furthermore, the authors proposed the LFN for extracting dermoscopic features, which achieved an average precision of 0.422 and a sensitivity of 0.693. For lesion segmentation, their proposed approach achieved an accuracy of 75.3%, while for lesion classification, it achieved 91.2% accuracy. Finally, for dermoscopic feature extraction of lesions, an accuracy of 84.8% was obtained.

Adegun et al. have introduced a novel approach to melanoma diagnosis using an encoder–decoder network combined with sub-networks of encoder and decoder, where skip pathways connect the semantic level of encoder and decoder feature maps for feature extraction (Adegun and Viriri, 2019). The proposed approach is based on a multi-scale, multi-stage approach that employs softmax as its classifier for pixel-wise classification. The approach has been tested and trained on two public datasets of skin lesion images, namely the ISIC 2017 and PH2 databases, without sample re-weighting in the loss function, which reduces system resources while enabling real-time melanoma diagnosis, with an average processing time of only 5 s per dermoscopy image. On both datasets, the proposed approach achieved an overall accuracy of 95% and a dice coefficient of 92% for the ISIC 2017 dataset, while the PH2 dataset produced an overall accuracy of 95% and a dice coefficient of 93%. In another study, Kassem et al. (2020) performed an 8-class classification on the ISIC-2019 dataset, using the pre-trained network GoogleNet to extract deep features from lesion images. With a transfer learning approach, they achieved an accuracy of 94.92%, including an "unknown class" for images that did not belong to any of the classes. Additionally, Kaur et al. (2022) designed a novel Deep Convolutional Neural Network (DCNN) architecture that is lightweight and efficient than other state-of-the-art approaches, utilizing a random over-sampling method with standard data augmentation techniques to handle the class imbalance problem. With this approach, the authors obtained an accuracy of 81.41%, 88.23%, and 90.42% on the ISIC-2016, ISIC-2017, and ISIC-2020 databases, respectively.

In another study, Esteva et al. proposed a single CNN model to classify skin lesion images using just disease and pixel labels as input (Esteva et al., 2017). This proposed model is trained on a large dataset of 129,450 medical images consisting of 2032 various diseases. This dataset is roughly double any database of clinical images discovered before them. The obtained 9-fold cross-validation accuracy was 72.1%. Further, Codella et al. proposed a deep learning model to segment, analyze and detect melanoma from skin lesion images (Codella et al., 2017). The proposed model is accompanied by an ensemble of well-established machine learning methods. It is trained and tested on a large public dataset consisting of 900 images for training and 379 images for testing the model. Finally, Xie et al. proposed a method to classify melanocytic tumors in their two categories, i.e., benign and malignant, by analyzing clinical dermoscopic images (Xie et al., 2016). The authors have identified 57 total descriptive features of the lesion objects, including 50 color and texture and another seven lesion border features. With this, they surpass the systems that only rely on standard shape features of these lesion objects.

2.3. Metaheuristic optimization and feature selection based approaches

Sayed et al. proposed a melanoma prediction model that combines the convolutional neural network (CNN) architecture with the Bald Eagle Search (BES) optimization technique (Sayed et al., 2021). They tackled the class imbalance problem using the random oversampling and standard data augmentation techniques, achieving an accuracy of 98.37% on the SIIM-ISIC 2020 dataset. İlkin et al. devised a hybrid classification algorithm by integrating the Support Vector Machine (SVM) algorithm with a heuristic optimization algorithm (İlkin et al., 2021). This approach, known as hybSVM, enhanced the SVM model utilizing a

Gaussian Radial Basis Function (RBF). The algorithm was evaluated on two distinct datasets, namely ISIC and PH2, employing a 10-fold cross-validation. The results demonstrated notable achievements, including AUC values of 98% and 97% for the respective datasets, along with efficient operation times of 26.5 s and 11.9 s.

Furthermore, Oliveira et al. proposed a computational method based on correlation-based feature selection for classifying skin lesions (Oliveira et al., 2019). They explored various color spaces for extracting color and texture-related characteristics, alongside several feature selection methods and classifiers. With the proposed feature selection method, an accuracy of 91.6% was attained, which further improved to 92.3% when utilizing the complete set of features. The study utilized a set of 50 features obtained through the proposed method. Bansal et al. introduced handcrafted feature extraction techniques specifically designed for dermoscopic images (Bansal et al., 2022). They presented two binary variants of the Harris Hawk Optimization (HHO) algorithm, namely BHHO-S and BHHO-V, for feature selection. The selected features were subsequently employed by a classifier to classify dermoscopic images as melanoma or non-melanoma. Experimental results revealed the superior performance of BHHO-S features compared to existing metaheuristic algorithms. Moreover, the study highlighted the efficacy of texture features extracted using local binary patterns and color features in achieving higher classification accuracy compared to global and local texture feature extraction techniques.

Deep learning is a data-driven method where the classes should be balanced appropriately enough for a model to learn those features and discriminate between classes. Many breakthroughs and promising results have been achieved in biomedical imaging, and researchers have worked to solve the imbalanced class problem. However, the majority of them have only explored the random over-sampling method. So, in this manuscript, the authors have proposed two novel methods to solve the class imbalance problem. The authors have also contrasted various currently used approaches with their suggestions and shown their benefits and drawbacks.

3. Presented methodology

3.1. Dataset preprocessing

As part of the data preprocessing step for this research, the images used were found to have varying sizes, necessitating a resizing operation to a standard size of 256×256 pixels, followed by RGB Reordering. This produced a final input of a $256 \times 256 \times 3$ matrix.

When downscaling images, careful consideration must be given to avoid the potential loss of vital information. Such a decision requires a thorough examination of the dataset. For instance, in the case of MRI scans for brain tumor classification, downsizing the images too much can cause the tumor to be nearly invisible, negatively impacting the training accuracy. However, increasing the image size to very large dimensions, such as 512×512, may lead to memory overload on the GPU. Thus, it is critical to choose the most appropriate image size based on experimental observations, to optimize memory efficiency while retaining all essential information in the image. The sample images from ISIC 2020 dataset is shown in Fig. 2.

3.2. Dataset division

In order to prevent model selection bias and mitigate overfitting, it is deemed advisable to partition the data into distinct training, testing, and cross-validation sets. Although a model may exhibit a training accuracy rate of 99%, its performance may not be as exceptional when evaluated real-time. The variability in parameter estimations is higher when the training dataset is smaller, and the same holds for performance measures when the testing dataset is less substantial. Therefore, it is imperative to split the data in a manner that does not result in excessive variances. In our study, we partitioned the complete dataset into three sections, with 80% allocated to training, and 10% each for validation and testing, as presented in Table 1.

Table 1
Dataset division.

	Training set	Validation set	Testing set	Total
Benign	26 034	3254	3254	**32 542**
Malignant	468	58	58	**584**

Table 2
Values and data augmentation techniques used to oversample the minority class.

Data augmentation technique	Value
Rotation range	15
Width shift range	0.15
Height shift range	0.15
Zoom range	0.15
Shear range	0.2
Horizontal flip	True

Table 3
8-fold dataset division for the Kemeny–Young optimal rank aggregation method.

Folds	Benign	Malignant
Fold 1	29 288/8 = **3661**	526 × 6 = **3156**
Fold 2	29 288/8 = **3661**	526 × 6 = **3156**
Fold 3	29 288/8 = **3661**	526 × 6 = **3156**
Fold 4	29 288/8 = **3661**	526 × 6 = **3156**
Fold 5	29 288/8 = **3661**	526 × 6 = **3156**
Fold 6	29 288/8 = **3661**	526 × 6 = **3156**
Fold 7	29 288/8 = **3661**	526 × 6 = **3156**
Fold 8	29 288/8 = **3661**	526 × 6 = **3156**

3.3. 8-fold packets dataset division

In our Proposed Method-I, we employ eight data packets to train eight sub-learners. To ensure a comprehensive evaluation, we first set aside 10% of the data as a separate testing set. This resulted in a training dataset consisting of 29,288 images in the benign class and 526 images in the malignant class. To address the class imbalance, we applied augmentation techniques specifically designed for the malignant class. The augmentation process, with detailed parameters and techniques provided in Table 2, resulted in the augmentation of the malignant class by a factor of six, yielding a total of 3156 augmented images. Next, we performed an 8-fold division of the benign class, ensuring an equal distribution of images among the folds. Each fold contained 3661 packets, resulting in a balanced representation of the benign class across the eight folds. Consequently, a single data packet contained 3156 images in the malignant class and 3661 images in the benign class. For clarity and reference, Table 3 provides a comprehensive overview of the data division for both classes after implementing the 8-fold packet division strategy.

This approach of dividing the dataset into separate folds serves two key purposes. First, it helps address the class imbalance issue by ensuring that the augmented malignant class is represented in each fold. Second, it enables each sub-learner to be trained on a diverse set of samples, including a balanced representation of both benign and augmented malignant images. By leveraging this balanced training data, our proposed method enhances the ability of the sub-learners to learn discriminative features and contribute effectively to the final classification decision. The 8-fold packet dataset division strategy is instrumental in promoting a more robust and accurate learning process, ultimately leading to improved melanoma detection performance.

3.4. Model training

In this study, we conducted experiments utilizing various pre-trained deep convolutional neural network (DCNN) models to extract features from histopathological images. As shown in Fig. 3, we froze all layers of the pre-trained DCNN models except for the final layer, and then flattened the extracted features. To mitigate overfitting,
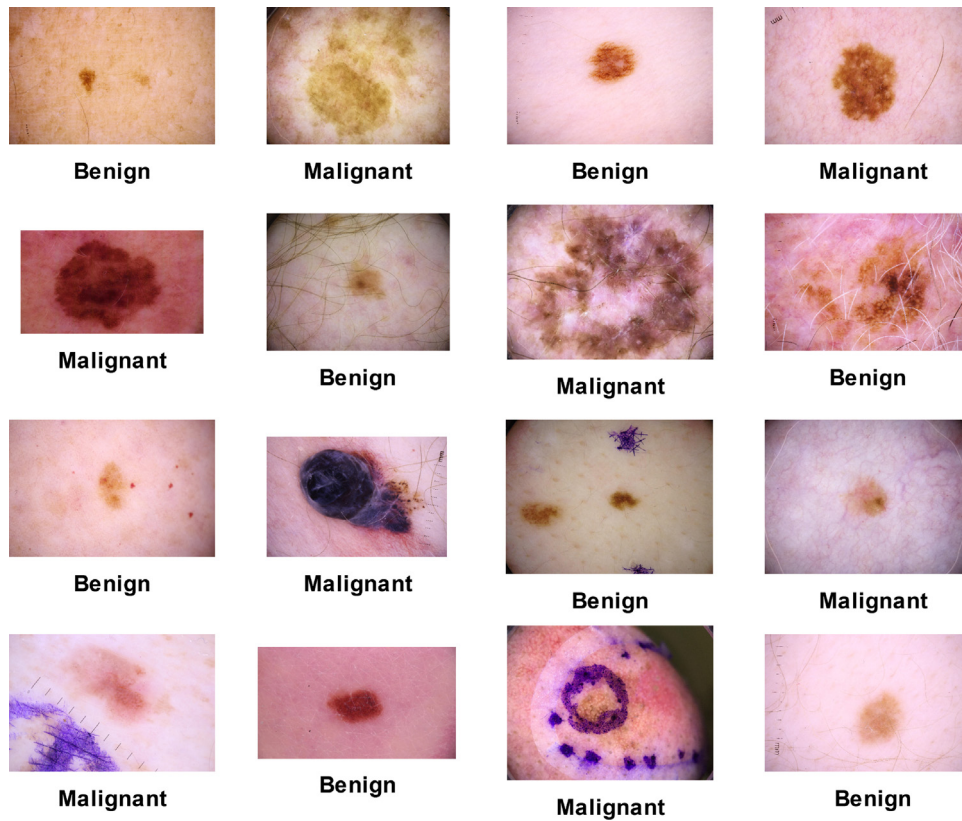
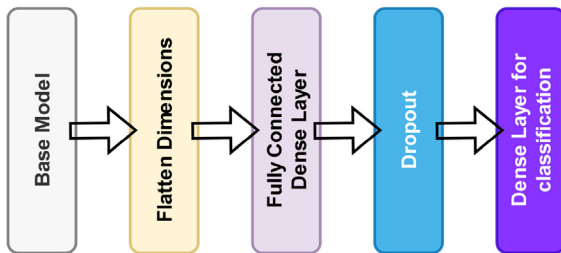**Fig. 2.** Sample benign and malignant images from the dataset ISIC 2020.

Fig. 3. Modified DCNN model architecture.

a dropout layer was introduced after a fully connected layer with 1024 neurons was added. Next, we included a dense layer with two neurons for classification purposes. The extracted deep features were subsequently used to train a multilayer perceptron network, with the sigmoid activation function employed for classification.

To optimize and enhance the deep learning models' feature extraction and classification performance, various methods and parameters were employed in this study. The loss function, optimizer, classifier, and learning rate scheduler were all systematically discussed to provide a comprehensive understanding of the models' training methods and parameters.

3.4.1. Loss function: alpha-balanced Focal Loss

Focal Loss is an improved version of Cross-Entropy Loss (CE) that addresses class imbalance by down-weighting simple cases and assigning more weight to hard or misclassified cases. The Focal Loss function is defined as shown in Eq. (1).

$$\text{FL}\left(p_t\right) = -\alpha_t \left(1 - p_t\right)^{\gamma} \log\left(p_t\right) \tag{1}$$

where p_t is defined as shown in Eq. (2).

$$p_t = \begin{cases} p; & \text{if } y = 1 \\ 1 - p; & \text{otherwise} \end{cases} \tag{2}$$

In the above equations, α_t is a weighting factor that emphasizes the importance of hard or misclassified cases, and γ is a focusing parameter that controls the rate at which easy cases are down-weighted. When $\gamma = 0$ and $\alpha_t = 1$, Focal Loss reduces to Cross-Entropy Loss.

Focal Loss has two primary properties that make it ideal for addressing class imbalance issues. First, the modulating factor is close to 1 when an example is misclassified, pt is modest, and the loss remains unaffected. However, the factor approaches 0 as pt approaches 1, leading to the down-weighting of the loss for well-classified samples. Second, the focusing parameter γ controls the rate at which easy cases are down-weighted, with higher γ values increasing the impact of the modulating factor.

The modulating factor reduces the loss contribution from simpler instances and increases the range where an example experiences minimal loss. For example, with $\gamma = 2$, an example with $p_t = 0.9$ would have a loss 100 percent less than CE, and with $p_t \approx 0.968$, it would have a loss of 1000 percent less. Thus, it is crucial to correct misclassified samples, whose loss is scaled down by at most 4x for $p_t \leqslant 0.5$ and $\gamma = 2$.

In this study, we utilized alpha-balanced Focal Loss which assigns weights to each class that are inversely proportional to their frequencies, ensuring that each class contributes equally to the loss function. It modifies the standard Focal Loss function by introducing an α-balance parameter that allows for greater control over the weights assigned to different classes. Specifically, this variant assigns a weight of $(1 - \alpha)$ to negative examples and a weight of α to positive examples, where α is a user-defined hyperparameter that determines the balance between the two classes. The modified Focal Loss function with the alpha-balance parameter is defined as shown in Eq. (3).

$$\text{FL}\alpha(pt) = -\alpha_t(1 - p_t)^{\gamma} \log(p_t) \tag{3}$$

where p_t and α_t is defined as shown in Eqs. (4) and (5) respectively.

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise} \end{cases} \tag{4}$$

and

$$\alpha_t = \begin{cases} \alpha & \text{if } y = 1 \\ 1 - \alpha & \text{otherwise} \end{cases} \tag{5}$$

Here, α_t represents the weight assigned to the true label y, and p_t is the predicted probability for the true label. The term $(1 - p_t)^\gamma$ is the modulating factor that amplifies the loss for hard-to-classify examples and reduces the loss for well-classified examples, as described in the standard Focal Loss function.

3.4.2. Optimizer: Adam

In the present research, we employ the Adam optimization algorithm to optimize our deep convolutional neural network model during training. This algorithm utilizes an exponential decay scheme to calculate the past gradients (m_t) and past squared gradients (v_t), as demonstrated by Eqs. (6) and (7) respectively. The hyperparameters β_1 and β_2 dictate the rate at which the mean and non-centered variance of the gradient are forgotten, respectively. Through keeping track of these two measures, the Adam optimizer estimates the first and second moments of the gradients and utilizes them to adaptively adjust the learning rate.

$$m_t = \beta_1 m_{t-1} + \left(1 - \beta_1\right)\left[\frac{\delta L}{\delta w_t}\right] \tag{6}$$

$$v_t = \beta_2 v_{t-1} + \left(1 - \beta_2\right)\left[\frac{\delta L}{\delta w_t}\right]^2 \tag{7}$$

where,

1. ϵ = a small positive constant to avoid denominator becoming zero when $\left(v_t -> 0\right) \cdot \left(10^{-8}\right)$
2. $\beta_1 \& \beta_2$ = decay rates of the average of gradients in the above two methods. $\left(\beta_1 = 0.9 \& \beta_2 = 0.999\right)$
3. α denotes the step size or learning rate, which is typically set to 0.001. It determines the magnitude of adjustment made to the model's parameters during each iteration of the optimization process.

3.4.3. Classifier: Softmax

In order to classify the extracted features, a Softmax classifier has been employed. The Softmax function is defined as shown in Eq. (8).

$$\sigma(\mathbf{z})i = \frac{e^{z_i}}{\sum j = 1^K e^{z_j}} \qquad \text{for } i = 1, \dots, K \tag{8}$$

where $\mathbf{z} = (z_1, \dots, z_K)$ represents the vector of class scores and $\sigma(\mathbf{z})_i$ represents the predicted probability of class i. This function converts a vector of real-valued scores into a probability distribution over K classes. The denominator in the equation ensures that the output is a valid probability distribution, where all the values range between 0 and 1 and the sum of all values is equal to 1.

3.4.4. Learning rate schedule : ReduceLROnPlateau

In this study, we employed the dynamic learning rate technique, specifically ReduceLROnPlateau. In the context of decay, the learning rate may be computed according to the formula given in Eqs. (9) and (10) respectively, which utilizes the decay parameter d, iteration step n, and learning rate η.

$$\eta_{n+1} = \frac{\eta_n}{1 + dn} \tag{9}$$

new learning rate = learning rate × factor $\tag{10}$

3.5. Kemeny–Young method

Given a set of candidates $C = c_1, c_2, \dots, c_m$ and a set of ranked lists $L_1, L_2, \dots, L_k$, where each ranked list L_i is a permutation of C, the goal of the Kemeny–Young method (Saari and Merlin, 2000; Levin and Nalebuff, 1995) is to find a consensus ranking π of the candidates that minimizes the total Kendall distance from the ranked lists as shown in Eq. (11).

$$\text{Kemeny}(\pi) = \sum_{i=1}^{k} d_K(L_i, \pi) \tag{11}$$

where $d_K(L_i, \pi)$ is the Kendall distance between the ranked list L_i and the consensus ranking π as shown in Eq. (12).

$$d_K(L_i, \pi) = \sum_{(c_a, c_b) \in \text{inv}(L_i)} \text{sgn}(\pi(c_a) - \pi(c_b)) \tag{12}$$

where $\text{inv}(L_i)$ is the set of all pairs (c_a, c_b) in the ranked list L_i such that c_a appears before c_b in L_i, and $\text{sgn}(x)$ is the sign function as shown in Eq. (13).

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{13}$$

To find the consensus ranking π, we can use the Kemeny–Young method based on majority voting, which involves finding the candidate that is preferred by the majority of voters in each pairwise comparison (Ahmed et al., 2017). Specifically, we can define the pairwise comparison matrix P as shown in Eq. (14).

$$P_{ab} = \sum_{i=1}^{k} [c_a \text{ is ranked above } c_b \text{ in } L_i] \tag{14}$$

where $[c_a \text{ is ranked above } c_b \text{ in } L_i]$ is the indicator function that is equal to 1 if c_a is ranked above c_b in ranked list L_i, and 0 otherwise.

Then, the consensus ranking π can be obtained by solving the integer linear programming problem shown in Eq. (15).

$$\min_{\pi} \quad \sum_{a < b} P_{ab} \, \text{sgn}(\pi(a) - \pi(b)) \tag{15}$$

$$\text{s.t.} \quad \pi(a) \in 1, 2, \dots, m \qquad \forall a \in C \tag{16}$$

$$\pi(a) \neq \pi(b) \qquad \forall a, b \in C, a \neq b \tag{17}$$

where $\pi(a)$ is the rank assigned to candidate c_a in the consensus ranking π. The objective function minimizes the weighted sum of the Kendall distances between the pairwise comparison matrix P and the consensus ranking π, and the constraints ensure that each candidate is assigned a unique rank in the consensus ranking. The Kemeny–Young Method is described in Algorithm 1.

3.6. Proposed Method-I: Kemeny–Young rule based majority voting

In the proposed Method-I, the entire dataset was partitioned into eight folds, and eight separate submodels were trained on these partitions. Each classifier was trained to differentiate between the 1/8th of the majority class images and those from the minority class. Thus, there were eight base models or voters in total. When an image is fed into the model, all eight voters predict and cast their votes. Finally, we aggregate the majority vote using the Kemeny–Young method-based majority voting and output the final predicted class as shown in Fig. 4.

Suppose we have a set of alternatives $\mathcal{A} = a_1, a_2, \dots, a_n$ and a set of m ranked preference orders $P_1, P_2, \dots, P_m$ over $\mathcal{A}$.

Let $d(a_i, a_j)$ denote the number of voters who rank a_i ahead of a_j, and $d(a_i, a_i) = 0$. Then, the pairwise disagreement distance D_{ij} between alternatives a_i and a_j is defined as shown in Eq. (18).

$$D_{ij} = \sum_{k=1}^{m} \mathbb{I}(a_i \text{ is ranked ahead of } a_j \text{ in } P_k) \tag{18}$$

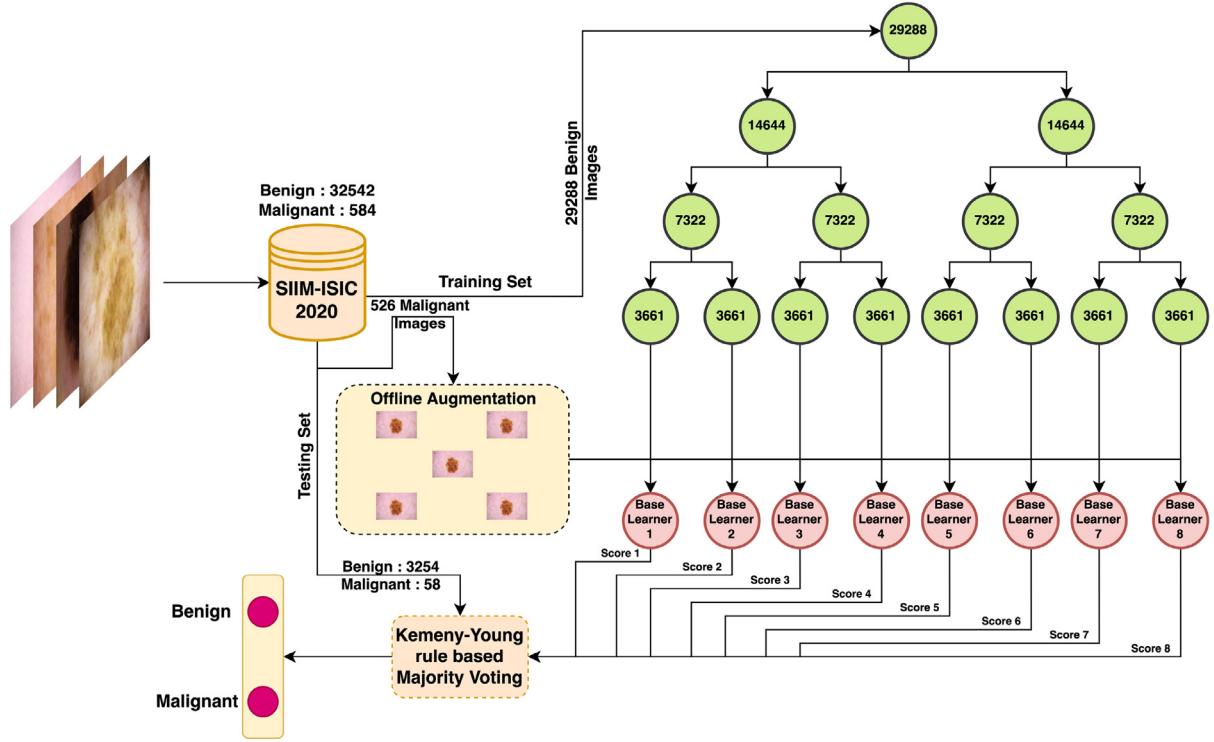where $\mathbb{I}$ is the indicator function.

Fig. 4. Flowchart for the majority voting based on Kemeny–Young optimal rank aggregation method.

Algorithm 1: Kemeny–Young Method

Input: D, a set of rankings

Output: A consensus ranking

Initialization: Set K to be the number of candidates;

1: Compute the pairwise disagreements d_{ij} between candidates i and j

$$d_{ij} = \sum_{k=1}^{n}(D_{ik} < D_{jk}) - \sum_{k=1}^{n}(D_{ik} > D_{jk})$$

2: Compute the Kendall tau distance between the pairwise disagreements d_{ij} and the pairwise disagreements of a candidate consensus ranking c_{ij}, where

$$\tau = \sum_{i<j} \text{sgn}(d_{ij} - c_{ij})$$

and $\text{sgn}(x)$ is the sign function;

3: Optimize the Kendall tau distance by finding the consensus ranking c^* that minimizes

$$\sum_{i<j} |\text{sgn}(d_{ij} - c_{ij}) - \text{sgn}(d_{ij} - c_{ij})|$$

4: Output the consensus ranking c^*;

The Kemeny distance $K(P, Q)$ between two preference orders P and Q is defined as the number of pairwise disagreements between P and Q shown in Eq. (19).

$$K(P, Q) = \sum_{i=1}^{n}\sum_{j=1}^{n} D_{ij} \cdot \mathbb{I}(P(a_i) \text{ is ranked ahead of } P(a_j)) \quad (19)$$

but $Q(a_i)$ is ranked ahead of $Q(a_j)$ where $P(a_i)$ denotes the rank of alternative a_i in P.

Given m ranked preference orders, the goal of the Kemeny–Young method is to find a consensus order C that minimizes the sum of

Kemeny distances to the input preference orders as shown in Eq. (20).

$$C = \text{argmin } P \in \mathcal{P} \sum k = 1^{m} K(P, P_k) \quad (20)$$

where $\mathcal{P}$ is the set of all possible preference orders over $\mathcal{A}$.

The algorithmic steps of the majority voting ensemble model based on the Kemeny–Young method are presented in Algorithm 2.

Algorithm 2: Kemeny–Young Method with Majority Voting

Input: Set of rankings $R_1, R_2, ..., R_k$

Output: Aggregated ranking R

1: Compute the pairwise disagreement matrix $D_{i,j}$ for all $i, j \in 1, 2, ..., k$, where $D_{i,j}$ represents the number of items where R_i and R_j disagree.

2: Compute the weighted pairwise disagreement matrix $W_{i,j} = \frac{1}{|S_{i,j}|} \sum_{x \in S_{i,j}} w(x)$, where $S_{i,j}$ is the set of items where R_i and R_j disagree, and $w(x)$ is the weight of item x.

3: Compute the sum of weights for each item, $W_i = \sum_{j \neq i} w_{i,j}$.

4: Compute the score for each item, $s_i = \sum_{j \neq i} w_{i,j} - \sum_{j \neq i} w_{j,i}$.

5: Sort the items in decreasing order of score and construct the aggregated ranking R.

3.7. Self-adaptive differential evolution algorithm

Let the objective function as $f : \mathbb{R}^n \rightarrow \mathbb{R}$
where n is the number of dimensions in the search space.

The Self-Adaptive Differential Evolution Algorithm (SA-DEO) uses a separate evolution strategy to adapt the scaling factor for each candidate solution (Qin and Suganthan, 2005; Omran et al., 2005; Brest et al., 2006; Srivastava et al., 2023b). In SA-DEO, the scaling factor is updated using the rule shown in Eq. (21).

$$F_i^{(t+1)} = F_i^{(t)} + \eta \cdot (\bar{F}^{(t)} - F_i^{(t)}) \quad (21)$$

where F_i is the scaling factor for candidate solution i, $\bar{F}^{(t)}$ is the mean scaling factor for the current population, and η is a scaling factor for the adaptation rate.

Algorithm 3: Self-Adaptive Differential Evolution Optimization to optimize assigned weights of a class in weighted class training

---

Input: Training data $x^{(i)}, y^{(i)} i = 1^m$, loss function $l(y, f(x, \theta))$, machine learning model $f(x, \theta)$, population size N, maximum number of iterations $Tmax$, scaling factor F, adaptation rate η, and stopping criterion.

Output: Optimized weights for the training examples.

1: Initialize the population with N weight vectors
$w_i = (w_{i,1}, w_{i,2}, ..., w_{i,m})$ with random values between 0 and 1.

2: Evaluate the objective function $J(w_i)$ for each weight vector w_i.

3: Set the iteration counter $t = 0$.

4: while *the stopping criterion is not met* **do**

  a: Generate trial vectors v_i for each weight vector w_i using the formula $v_i = w_{r_1} + F_i \cdot (w_{r_2} - w_{r_3})$, where r_1, r_2, r_3 are randomly selected indices from the population, and F_i is the scaling factor for the weight vector w_i.

  b: Apply the projection operation $w_i = \min(\max(w_i, 0), 1)$ to each weight vector w_i to ensure that they stay within the range of 0 and 1.

  c: Evaluate the objective function $J(v_i)$ for each trial vector v_i.

  d: for *each weight vector w_i* **do**

    i: Generate a new weight vector u_i using the formula
$$u_{i,j} = \begin{cases} v_{i,j} & \text{if } rand(0, 1) \leq CR \text{ or } j = j_{rand} \\ w_{i,j} & \text{otherwise} \end{cases},$$
where CR is the crossover probability, j_{rand} is a randomly selected index, and $rand(0, 1)$ is a random number between 0 and 1.

    ii: Evaluate the objective function $J(u_i)$ for the new weight vector u_i.

    iii: If $J(u_i) < J(w_i)$, set $w_i = u_i$.

  end

  e: Calculate the mean weight vector $\bar{w}^{(t)}$ for the current population.

  f: Update the adaptation rate η and the scaling factor $F_{i,j}$ for each weight vector using the formulas $\eta^{(t+1)} = \eta^{(t)} + \Delta\eta$ and $F_{i,j}^{(t+1)} = F_{i,j}^{(t)} + \Delta F_{i,j}$, where the values $\Delta\eta$ and $\Delta F_{i,j}$ are derived from a normal distribution with a mean of 0 and a standard deviation of 1, and are assigned randomly.

  g: Increment the iteration counter t.

end

5: Return the optimized weights w_i.

The trial vector for each candidate solution x_i is generated using the rule shown in Eq. (22).

$$v_i = x_{r_1} + F_i \cdot (x_{r_2} - x_{r_3}) \tag{22}$$

where r_1, r_2, r_3 are randomly selected indices from the population, and F_i is the scaling factor for candidate solution i.

The comparison between the trial solution and the current solution x_i is carried out according to the rule depicted in Eq. (23).

$$x_i^{(t+1)} = \begin{cases} v_i & \text{if } f(v_i) < f(x_i) \\ x_i & \text{otherwise} \end{cases} \tag{23}$$

where t is number of iterations.

The population is typically initialized with random candidate solutions, and each candidate solution is represented as a vector in the search space $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,n})$
where n is the number of dimensions in the search space.

The adaptation rate η can be either fixed or adaptively adjusted. In the latter case, a value of η is randomly generated for each candidate
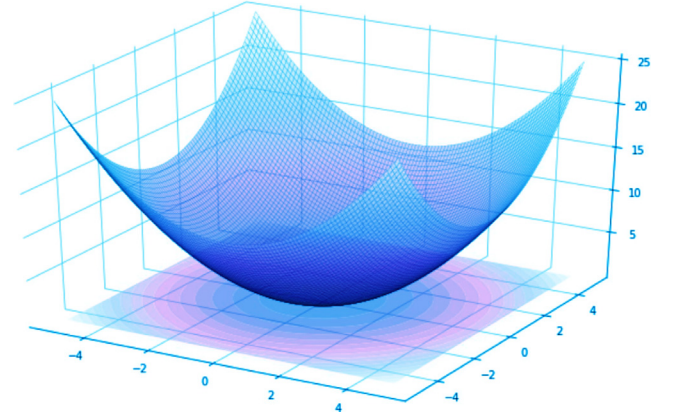

Fig. 5. Representation of $f(x) = \sum_i^n x_i^2/n$.

solution in each iteration. One popular strategy for adaptive scaling is to use the update rule shown in Eq. (24).

$$\eta^{(t+1)} = \eta^{(t)} + \Delta\eta \tag{24}$$

where η is the adaptation rate, and $\Delta\eta$ is a random perturbation.

Let $f(x) = \sum_i^n x_i^2/n$, for n = 32 dimensions. $f(x) = \sum_i^n x_i^2/n$ 2D view is shown in Fig. 5.

3.8. Proposed method-II: Self-Adaptive Differential Evolution Optimization to optimize weights of a class in weighted class training

In the proposed method-II, we assigned a separate weightage for each class during weighted class training, with a higher weight given to the minority class and a lower weight given to the majority class. To optimize these weights, we employed the Self-Adaptive Differential Evolution Optimization algorithm, which takes input parameters including a training dataset, loss function, machine learning model, population size, maximum number of iterations, scaling factor, adaptation rate, and stopping criterion.

The algorithm works by generating a population of weight vectors and iteratively creating new candidate solutions through mutation and recombination. At each iteration, trial vectors are created by combining different weight vectors from the population using a scaling factor and crossover probability, and the objective function is evaluated for each trial vector. The best trial vectors are then selected to replace the original weight vectors in the population.

To adapt the algorithm over time, a self-adaptive mechanism is included to adjust the scaling factor and adaptation rate of the population based on the mean weight vector and random values. The goal of the algorithm is to optimize the weights of a class in weighted class training by iteratively adjusting the weight vector population to minimize the loss function for the given machine learning model.

We start with the objective function that we want to optimize, which is typically the loss function of a machine learning model. In the case of weighted class training, we want to minimize the objective function shown in Eq. (25).

$$J(\theta) = \sum_{i=1}^m w_i l(y^{(i)}, f(x^{(i)}, \theta)) \tag{25}$$

where θ are the parameters of the machine learning model, m is the number of training examples, $x^{(i)}$ and $y^{(i)}$ are the input and output of the ith training example, respectively, $f(x^{(i)}, \theta)$ is the predicted output of the model for the ith training example, l is the loss function, and w_i is the weight for the ith training example.

We can use the Self-Adaptive Differential Evolution Algorithm to optimize the weights w_i of the training examples. We represent the
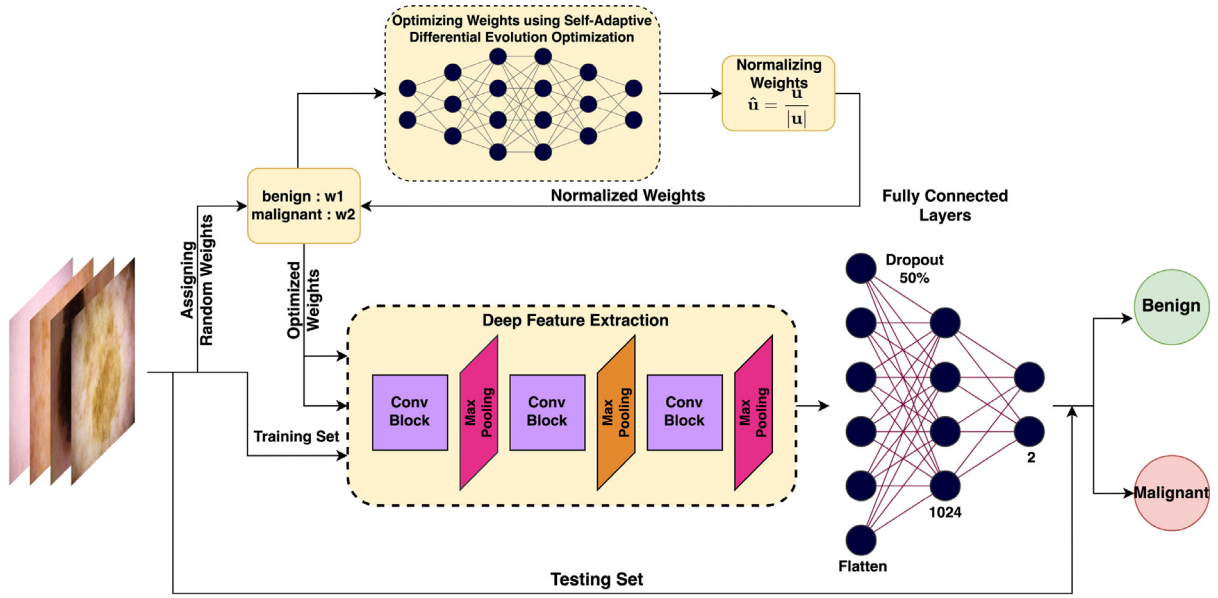
Fig. 6. Graphical abstract of the proposed self-adaptive DEO for weighted class training.

weights as a vector in the search space $w = (w_1, w_2, \ldots, w_m)$. The objective function becomes as shown in Eq. (26).

$$J(w) = \sum_{i=1}^{m} w_i l(y^{(i)}, f(x^{(i)}, \theta)) \qquad (26)$$

The Self-Adaptive Differential Evolution Algorithm updates the weights using the rule shown in Eq. (27).

$$w_i^{(t+1)} = w_i^{(t)} + \eta \cdot (\bar{w}^{(t)} - w_i^{(t)}) \qquad (27)$$

where w_i is the weight for the ith training example, $\bar{w}^{(t)}$ is the mean weight for the current population, and η is a scaling factor for the adaptation rate.

The trial vector for the weight vector w is generated using the rule shown in Eq. (28).

$$v_i = w_{r_1} + F_i \cdot (w_{r_2} - w_{r_3}) \qquad (28)$$

where r_1, r_2, r_3 are randomly selected indices from the population, and F_i is the scaling factor for the weight vector w_i.

The trial solution is then compared with the current solution w_i using the rule shown in Eq. (29).

$$w_i^{(t+1)} = \begin{cases} v_i & \text{if } J(v_i) < J(w_i) \\ w_i & \text{otherwise} \end{cases} \qquad (29)$$

where t is the iteration number.

The population is typically initialized with random weight vectors, and each weight vector is represented as a vector in the search space $w_i = (w_{i,1}, w_{i,2}, \ldots, w_{i,m})$
where m is the number of training examples.

The adaptation rate η and the scaling factor F_i can be either fixed or adaptively adjusted. In the latter case, a value of η and F_i is randomly generated for each weight vector in each iteration. One popular strategy for adaptive scaling is to use the update rule shown in Eqs. (30) and (31) respectively.

$$\eta^{(t+1)} = \eta^{(t)} + \Delta\eta \qquad (30)$$

$$F_{i,j}^{(t+1)} = F_{i,j}^{(t)} + \Delta F_{i,j} \qquad (31)$$

The adaptation rate, denoted by η, is involved in the calculation along with two random values: $\Delta\eta$ and $\Delta F_{i,j}$. These values are drawn from a normal distribution with a mean of 0 and a standard deviation of 1.

The weights w_i are typically constrained to be between 0 and 1, which can be enforced using a projection operation shown in Eq. (32).

$$w_i = \min(\max(w_i, 0), 1) \qquad (32)$$

The Self-Adaptive Differential Evolution Algorithm is designed to repeatedly generate and select trial solutions until a predetermined stopping criterion is met, which may include a maximum number of iterations or a minimum change in the objective function. Upon conducting this iterative process for a total of 1000 times, the researchers obtained the most optimal weights for their dataset and subsequently trained their model using these weights. This approach is depicted in a Fig. 6 and its training procedure is outlined in detail in Algorithm 3.

In this algorithm, CR is typically set to a value between 0.1 and 0.9, and j_{rand} is randomly selected from the indices of the weight vector w_i. The adaptation rate η and the scaling factor F_i are updated for each iteration using the formulas above, which allows the algorithm to adaptively adjust the exploration–exploitation balance based on the current population.

The projection operation is used to ensure that the weight vectors stay within the range of 0 and 1, which is important for weighted class training, where the weights represent the importance of each class. The objective function $J(w_i)$ evaluates the performance of the machine learning model with the given weights on the training data, and can be defined as the loss function $l(y, f(x, \theta))$ weighted by the weight vector w_i.

4. Experimental results and analysis

4.1. Dataset description : SIIM-ISIC 2020

The dataset employed in this study is the official dataset of the SIIM-ISIC Melanoma Classification Challenge, known as the "ISIC 2020 Challenge Dataset" (Rotemberg et al., 2021). This dataset was explicitly curated for the 2020 Summer Kaggle SIIM-ISIC Melanoma Classification Challenge and encompassed 33,126 dermoscopic training images depicting distinct benign and malignant skin lesions from approximately 2000 patients. The training set includes ground truth labels for all 33,126 images, whereas the test data lack such information. Consequently, our analysis focused exclusively on the training data.

The dataset used in this study comprises 32,542 images of benign lesions and 584 images of malignant melanomas. The malignant

diagnoses were validated through histopathological analysis, while the benign diagnoses were established based on expert evaluation, long-term monitoring, or histopathology. To ensure the accuracy of the ground truth labels for all malignant lesions, a retrospective assessment of histopathology reports was conducted, further supported by a dermoscopy specialist for diagnostic plausibility. Suspect cases were carefully re-examined in comparison to histopathology reports to uphold label accuracy. It is worth noting that both melanomas in situ and invasive melanoma were classified as melanoma, while other qualified images, including highly dysplastic nevi, were categorized as benign (Lott et al., 2016; Piepkorn et al., 2014).

Within the ISIC 2020 dataset, approximately 79.2% of the patients did not present melanoma, whereas about 20.8% exhibited at least one case of melanoma, averaging 16 lesions per patient. The dataset comprised 33,126 dermoscopic images, of which only 584 (1.8%) were histopathologically confirmed as melanomas, distinct from benign melanoma mimickers. The training images had an average pixel count of 12,743,090, ranging from 307,200 to 24,000,000 pixels. The dataset consists of two distinct classes: malignant and benign. For our study, we utilized a dataset consisting of 32,542 images of benign lesions and 584 images of malignant melanomas. Consequently, the imbalanced ratio of our dataset is approximately 55.7:1, accurately reflecting the real-world prevalence of melanoma cases. The detailed data description is shown in Table 5.

4.2. Experimental settings

The present investigation utilized the Tensorflow framework for coding implementation, and the Models were trained on a workstation equipped with 12th Gen Intel(R) Core(TM) i7-1265U processor and Nvidia RTX 3090 GPU. In order to ensure accurate classification accuracy during the training, validation, and testing phases, all models were trained for a duration of 200–500 epochs.

4.3. Results

The outcomes of the proposed methodology for melanoma detection are discussed and assessed in the following section. Six different DCNN classifiers were used in all of the experiments. The first section explored and assessed various experimental methods, including undersampling and oversampling. Next, the proposed techniques' results are discussed in the following section. The authors have used the proposed approach in various ways, including one that uses the usual loss function and another that alters the loss function. The subsequent sections thoroughly detail and explain each experiment.

4.3.1. Traditional class imbalance approaches

The primary and essential step in developing a deep learning model involves the specification of its architecture. For our investigation, we opted to utilize pre-existing networks to extract prominent features from histopathological images, which had previously undergone training on an extensive ImageNet dataset. This technique significantly reduces the computational power required to adjust weights to fit our ISIC 2020 dataset. To extract deep features, we utilized six DCNN models: VGG16, InceptionV3, Xception, InceptionResNetV2, ResNet152V2, and DenseNet201. To enhance the efficiency of our deep convolutional neural network (DCNN) models, the authors used the different hyperparameters as shown in Table 4. Moreover, early stopping callbacks were incorporated to save the best weights when a monitored parameter stopped improving.

To mitigate class inequality, we have adopted several strategies, with random undersampling being the first. As shown in Table 6, ResNet152V2 outperforms other pre-trained networks in experiments, with a testing accuracy of 82.26%. However, excessive information loss may lead to a lower overall accuracy. In an undersampling approach, the majority class discards most of the data, resulting in the loss of most

Table 4
Hyperparameters set for the model training.

Hyperparameters	Values
1st momentum decay rate (β_1)	0.9
2nd momentum decay rate (β_2)	0.999
Epsilon (ϵ)	1e−7
Starting learning rate (α)	0.001
Factor	0.1
Patience	10
Total epochs	250–300
Optimizer	Adam
Dropout value	0.5
Batch size	32

Table 5
Dataset description.

Dataset name	ISIC 2020 challenge dataset
Total number of images	33,126
Number of patients	Approximately 2000
Classes	Malignant, Benign
Number of images (Benign)	32,542
Number of images (Malignant)	584
Imbalanced ratio (Benign to malignant)	Approximately 55.7:1
Percentage of patients without melanoma	79.2%
Percentage of patients with melanoma	20.8%
Average lesions per patient with melanoma	16
Histopathologically confirmed melanomas	584 (1.8% of the dataset)
Average image pixel count	12,743,090 pixels
Minimum image pixel count	307,200 pixels
Maximum image pixel count	24,000,000 pixels

features from the benign class. As a result, the model has difficulty discriminating the malignant class from the benign class due to the insufficient features in the training set.

Conversely, random oversampling yields significantly improved results, as demonstrated in Table 7. The DenseNet201 classifier with the highest testing accuracy of 97.43% performs the best. In this scenario, oversampling is a much better approach than undersampling since no features are discarded. However, augmenting the minority class by around 50–60 times may lead to overfitting, whereby the model views the same data more frequently. Thus, the likelihood of this model failing in a real-world scenario is considerably higher. In light of this, our research proposes two approaches to handle class imbalance, which are expounded on in the subsequent section.

4.3.2. Proposed approach

The authors proposed two approaches to address the issue class imbalance in the skewed dataset. The first approach involves partitioning the dataset into eight folds and training eight base classifiers, using the Kemeny–Young method based majority voting. Each base classifier learns to distinguish the characteristics of one-eighth of the benign from the malignant classes, and each casts a vote when presented with test data. During testing, each base learner casts a vote, and the result with the most votes is considered. This approach subdivides the work into eight smaller tasks, allowing each model to learn and discriminate from the malignant class using just 1/8th of the necessary features. The accuracy on the majority vote increased compared to the prior 8-fold accuracy obtained as per the results in Table 8. The loss and accuracy curve for the training procedure of each fold of the VGG16 model is shown in Figs. 8 and 9, respectively.

The second proposed approach involves weighted class training with Self-Adaptive Differential Evolution Optimization. The authors used a meta-heuristic-based approach to identify the optimum weights using the Self-Adaptive differential evolution method as a minimization problem while monitoring the loss value. The optimized weights obtained after running the algorithm 1000 times are shown in Table 9. The authors also compared the optimized weights with the standard weights calculation method pre-defined in the sklearn library. The
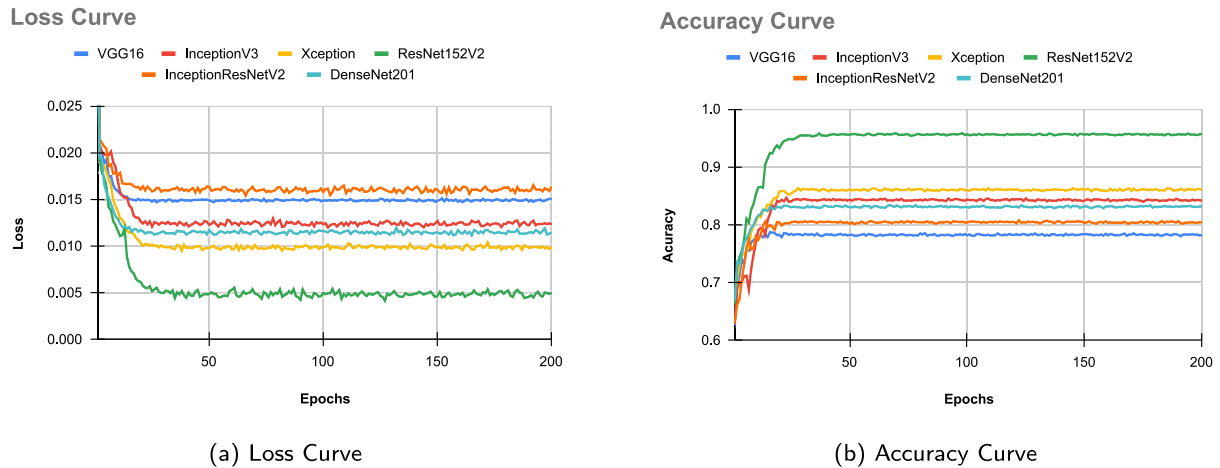
(a) Loss Curve

(b) Accuracy Curve

Fig. 7. Loss and accuracy curve for training procedure of optimized weighted class training based ensemble model.

Table 6
Results obtained with the random undersampling method.

Classifiers	Training error (%)	Validation error (%)	Testing error (%)	Sensitivity (%)	Specificity (%)	F1-score (%)
VGG16	13.72	18.62	18.09	34.78	93.99	43.96
Xception	0.05	22.41	19.69	31.30	92.87	39.34
InceptionV3	0.35	18.62	19.33	30.43	93.54	39.11
ResNet152V2	0.05	16.90	17.74	38.26	93.54	46.81
Inception ResNetV2	0.25	23.10	20.22	27.83	93.10	35.96
DenseNet201	0.45	20.69	19.69	41.74	90.20	46.38

Table 7
Results obtained with the random oversampling method.

Classifiers	Training error (%)	Validation error (%)	Testing error (%)	Sensitivity (%)	Specificity (%)	F1-score (%)
VGG16	2.51	4.08	4.29	13.79	97.17	10.13
Xception	0.84	3.02	3.02	15.52	98.43	15.25
InceptionV3	6.10	5.71	5.80	22.41	95.48	11.93
ResNet152V2	0.84	3.89	3.59	13.79	97.88	11.85
Inception ResNetV2	5.86	5.16	4.96	20.69	96.37	12.77
DenseNet201	0.68	2.66	2.57	6.90	99.05	8.60

Table 8
Results obtained with the Kemeny–Young based majority voting method on each fold and the final fold.

Classifier	Fold 1 Error (%)	Fold 2 Error (%)	Fold 3 Error (%)	Fold 4 Error (%)	Fold 5 Error (%)	Fold 6 Error (%)	Fold 7 Error (%)	Fold 8 Error (%)	Majority Voting Error (%)
VGG16	10.70	5.89	10.46	13.12	8.92	5.35	13.72	6.41	**2.44**
Xception	15.69	12.58	10.04	12.64	16.90	15.90	14.03	16.26	**11.79**
InceptionV3	18.65	15.30	15.18	19.98	14.30	16.63	16.26	13.97	**9.94**
ResNet152V2	12.88	15.36	12.52	14.60	14.42	16.23	12.58	14.12	**7.64**
Inception ResNetV2	19.29	15.87	19.44	19.41	17.90	14.48	17.59	18.74	**8.27**
DenseNet201	11.55	10.46	9.28	10.31	10.01	9.76	11.46	9.25	**3.77**

Table 9
Optimized weights with Self-Adaptive Differential Evolution Optimization.

Method	Benign : $w1$	Malignant : $w2$
Random weights	1	1
Traditional method	0.0179	1
Optimized weights	0.0729	1

model was trained using the optimized weights using a process known as cost-sensitive learning, which involves giving weights to each class. The loss and accuracy curves for the training procedure of all six DCNN classifiers are shown in Fig. 7. Table 10 contrasts the outcomes of five alternative strategies, including the proposed approach with and without weights, and with and without altering the loss function. These experiments were conducted to examine the robustness of the proposed approach.

4.4. Comparison with the state-of-the-art

The International Skin Imaging Collaboration (ISIC) releases a yearly dataset for melanoma detection. The challenge named SIIM-ISIC is conducted every year. Various articles that we evaluated employed different datasets. The ISIC 2020 dataset, the most recent dataset made available by the ISIC, has been considered in this research. To address the class imbalance issue, we choose to use the highly unbalanced ISIC 2020 dataset as the main focus of our research. Table 11 demonstrates that our two proposed methods surpassed several state-of-the-art models/methodologies proposed in earlier studies.

As evidenced by the findings presented in Table 11, a significant number of the proposed methodologies employ Oversampling-based approaches or Data Augmentation techniques to address the issue of imbalanced classes by increasing the representation of the minority class. While this approach is generally regarded as effective, it is

Table 10

Results obtained with the Weighted class training method.

Method	Metric	VGG16	Xception	InceptionV3	ResNet152V2	InceptionResNetV2	DenseNet201
Normal	Error rate (%)	1.78	1.75	1.75	1.75	1.75	1.75
	Sensitivity (%)	0	0	0	0	0	0
	Specificity (%)	99.97	100	100	100	100	100
	F1-score (%)	0	0	0	0	0	0
Normal + Focal loss	Error rate (%)	1.81	1.81	1.78	1.75	1.75	1.69
	Sensitivity (%)	0	0	0	1.72	0	3.45
	Specificity (%)	99.94	99.97	99.97	99.97	100	100
	F1-score (%)	0	0	0	3.33	0	6.67
Weighted class training	Error rate (%)	21.20	14.73	16.06	5.68	19.66	11.35
	Sensitivity (%)	68.97	62.07	63.79	39.66	72.41	70.69
	Specificity (%)	78.98	85.68	84.30	95.30	80.49	88.97
	F1-score (%)	10.23	12.86	12.21	19.66	11.43	17.90
Focal loss + Weighted class training	Error rate (%)	15.16	8.70	10.02	2.66	13.62	5.31
	Sensitivity (%)	68.97	62.07	63.79	39.66	72.41	70.69
	Specificity (%)	85.13	91.83	90.44	98.37	86.63	95.11
	F1-score (%)	13.75	20	18.23	34.33	15.70	31.78
Focal loss + Optimized weighted class training	Error rate (%)	8.82	2.36	3.99	2.36	7.28	1.99
	Sensitivity (%)	86.21	79.31	63.79	56.90	89.66	87.93
	Specificity (%)	91.27	97.97	96.59	98.37	92.78	98.19
	F1-score (%)	25.51	54.12	35.92	45.83	30.14	60.71

Table 11

Comparitive study between proposed methods and the state-of-art.

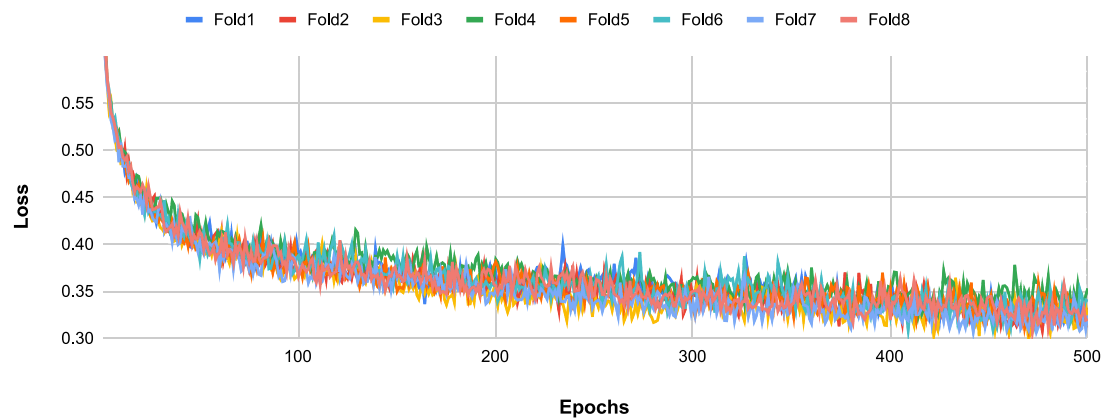| Paper	Method used	Dataset	Classification error (%)
Kaur et al. (2022)	DCNN using random oversampling	ISIC 2016	18.59
Oliveira et al. (2019)	Correlation-based feature selection method	ISIC 2016	7.70
Kaur et al. (2022)	DCNN using random oversampling	ISIC 2017	11.77
Li and Shen (2018)	Lesion Indexing Network and Lesion Feature Network	ISIC 2017	8.80
Adegun and Viriri (2019)	Encoder Decoder Network	ISIC 2017	5.01
Kassem et al. (2020)	Transfer Learning using pre-trained GoogleNet	ISIC 2019	5.08
Yu et al. (2016)	Deep Residual Networks	ISIC 2020	15.60
Pham et al. (2018)	Deep CNN and Data Augmentation	ISIC 2020	12.80
Kaur et al. (2022)	DCNN using random oversampling	ISIC 2020	9.58
Al-Masni et al. (2018)	Deep full resolution convolutional networks	ISIC 2020	5.97
Hosny et al. (2019)	Transfer learning and pre-trained Deep neural network	ISIC 2020	4.09
İlkin et al. (2021)	Bacterial colony optimization based SVM	ISIC 2020	2.50
Proposed method	**Kemeny–Young rule based Majority Voting**	ISIC 2020	**2.46**
Proposed method	**Self Adaptive Differential Evolution Optimization based cost-sensitive learning**	ISIC 2020	**1.99**

Fig. 8. Loss curve for training procedure of each fold of Kemeny–Young based majority voting ensemble model.

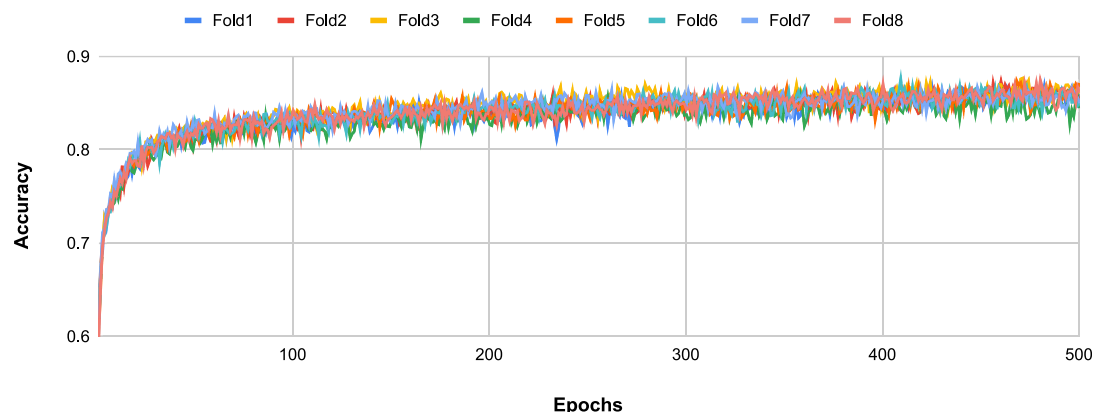## Accuracy Curve of Kemeny-Young Rule Based Ensemble Model

Fig. 9. Accuracy curve for training procedure of each fold of Kemeny–Young based majority voting ensemble model.

important to note that these models often exhibit overfitting tendencies when applied to test data, resulting in suboptimal predictions for the minority class, which ultimately undermines our primary objective.

5. Discussion and future directions

In this section, we first discuss the flexibility of the proposed methods, and the essential advantages of proposed methods in solving class imbalance problems. Finally, limitations of the proposed work and future directions are considered.

5.1. Flexibility of the proposed methods

The strength of the proposed methods for addressing class imbalance in melanoma detection lies in their inherent flexibility and adaptability to diverse scenarios and datasets. Firstly, the method based on the Kemeny–Young rule for rank aggregation takes into consideration the unique characteristics of each class and aims to minimize bias between them, irrespective of the specific data distribution. This flexibility renders the method suitable for datasets exhibiting varying degrees of class imbalance, ranging from mild to severe. Its ability to adapt to different data distributions ensures its applicability in diverse melanoma detection scenarios.

Likewise, the cost-sensitive learning approach employing Self-Adaptive Differential Evolution Optimization offers flexibility in assigning optimal weights to each class. In contrast to traditional methods that rely on predefined criteria for weight assignment, this approach dynamically determines the weights based on the inherent characteristics of the dataset under consideration. This adaptability allows the method to adjust to the specific challenges posed by imbalanced data, thereby enabling effective handling of varying degrees of class imbalance and consequent improvement in classification performance.

Moreover, the proposed methods can be seamlessly integrated into existing deep learning frameworks and workflows. Deep learning has emerged as a potent tool in the domain of biomedical image analysis, and the proposed methods seamlessly align with this paradigm. The compatibility of these methods with deep learning architectures, such as CNNs, facilitates their smooth integration into established pipelines and enhances their adoption by researchers and practitioners in the field.

It is important to note that the proposed methods are not limited solely to melanoma detection but can also be adapted and extended to address class imbalance in other biomedical imaging applications. While skin cancer classification exemplifies a scenario where imbalanced class distributions present challenges, numerous other medical imaging tasks, such as the detection of rare diseases or abnormalities, encounter similar issues. The proposed methods serve as a foundation

for addressing class imbalance across diverse medical imaging domains. By adapting these methods to different datasets and application contexts, researchers can effectively tackle class imbalance and enhance the accuracy and reliability of various computer-aided diagnosis systems.

5.2. Limitations and future directions

Although the proposed method for addressing class imbalance in melanoma detection has demonstrated promising results and outperformed state-of-the-art approaches, it is important to acknowledge its limitations and identify potential avenues for future research.

One limitation of the current approach pertains to its reliance on the Kemeny–Young rule for rank aggregation. While this rule has proven effective in reducing bias between classes and improving classification accuracy, its suitability may vary across different datasets or classification tasks. Another aspect to consider is the computational complexity associated with the proposed method, particularly concerning the utilization of the Self-Adaptive Differential Evolution Optimization algorithm for determining optimal weights. While this algorithm has exhibited promising results in identifying optimal class weights, it may present computational challenges when applied to larger datasets or real-time applications.

In terms of future directions, one avenue of exploration involves the integration of multi-modal data for melanoma detection. The combination of information from multiple imaging modalities, such as dermoscopy, clinical images, and histopathology, has the potential to enhance the accuracy and reliability of melanoma diagnosis. Investigating the fusion of different data sources and developing robust algorithms capable of handling heterogeneous data (Zhang et al., 2022) can further augment the performance of melanoma detection systems. Moreover, the proposed method can be extended to address other challenges in dermatological image analysis beyond melanoma detection. Exploring the application of the proposed method to other skin disease classification tasks, such as detecting non-melanoma skin cancers or differentiating various dermatological conditions, would broaden the utility and impact of the research.

6. Conclusion

The application of deep learning in biomedical imaging remains an active area of research, with the challenge of addressing the problem of imbalanced class distributions receiving significant attention. This issue is particularly evident in the context of identifying diseases such as COVID-19 during the recent pandemic, where datasets comprising chest X-ray or CT scans initially exhibit severe class imbalance. Similarly, in

our study, the Melanoma dataset (ISIC 2020) also presented a highly imbalanced class distribution.

In response to these challenges, this manuscript proposes a novel method to effectively handle class imbalance. The authors extensively explore and compare alternative deep learning techniques employed by researchers to address this issue. The first proposed approach involves the application of a Kemeny–Young-based majority voting ensemble model, where the dataset is partitioned into chunks and the majority vote determines the final prediction. This method, referred to as Proposed Method 1, achieves an classification error of 2.44%. Additionally, the authors propose a weighted class training approach as the second method. This approach utilizes the Self-Adaptive Differential Evolution algorithm, a meta-heuristic-based optimization technique, to determine optimal weights for the dataset. The second proposed method achieves an impressive overall classification error of 1.99%.

The findings of this research provide valuable insights for future scholars in their endeavors to tackle class imbalance and explore alternative approaches. Moreover, the proposed methodologies have wider implications beyond the specific context of this study, particularly in the field of biomedical imaging. The proposed method can greatly assist in distinguishing between benign and malignant melanoma images. The authors anticipate that the effective detection of melanoma enabled by this research will have practical benefits for hospitals and physicians, leading to improved diagnostic capabilities through the integration of these methodologies into implantable devices and graphical user interfaces (GUI).

CRediT authorship contribution statement

Gaurav Srivastava: Conceptualization, Formal analysis, Methodology, Software, Code implementations, Writing – original draft, Writing – review & editing, Validation, Investigation. **Nitesh Pradhan:** Validation, Resources, Investigation, Writing – review & editing, Supervision, Research administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data utilized in this study can be found at the following publicly available URL: https://challenge2020.isic-archive.com/.

Appendix A

Abbreviation	Meaning
AI	Artificial Intelligence
CAD	Computer-Aided Diagnosis
DL	Deep Learning
ML	Machine Learning
DCNN	Deep Convolutional Neural Network
CNN	Convolutional Neural Network
SIIM	Society for Imaging Informatics in Medicine
ISIC	International Skin Imaging Collaboration

ELM	Extreme Learning Machine
W-ELM	Weighted Extreme Learning Machines
FCRN	Fully Convolutional Residual Networks
LIN	Lesion Indexing Network
LFN	Lesion Feature Network
LICU	Lesion Indexing Calculation Unit
BES	Bald Eagle Search
RBF	Radial Basis Function
SVM	Support Vector Machine
HHO	Harris Hawk Optimization
MLP	Multi-layer Perceptron Network
SA-DEO	Self-Adaptive Differential Evolution Optimization
SMOTE	Synthetic Minority Oversampling Technique

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.engappai.2023.106738.

References

Adegun, A.A., Viriri, S., 2019. Deep learning-based system for automatic melanoma detection. IEEE Access 8, 7160–7172.

Adepu, A.K., Sahayam, S., Jayaraman, U., Arramraju, R., 2023. Melanoma classification from dermatoscopy images using knowledge distillation for highly imbalanced data. Comput. Biol. Med. 154, 106571.

Ahmed, M.T., Hussain, M.M., Chennam, K.K., 2017. Designing a consensus ranking algorithm for same domain entities. In: 2017 2nd International Conference on Communication and Electronics Systems (ICCES). IEEE, pp. 12–16.

Al-Masni, M.A., Al-Antari, M.A., Choi, M.-T., Han, S.-M., Kim, T.-S., 2018. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. Comput. Methods Programs Biomed. 162, 221–231.

American Cancer Society, 2022. Melanoma skin cancer, https://www.cancer.org/cancer/melanoma-skin-cancer.html.

Aspinwall, L.G., Leaf, S.L., Dola, E.R., Kohlmann, W., Leachman, S.A., 2008. CDKN2A/p16 genetic test reporting improves early detection intentions and practices in high-risk melanoma families. Cancer Epidemiol. Biomarkers Prevent. 17 (6), 1510–1519.

Bansal, P., Vanjani, A., Mehta, A., Kavitha, J., Kumar, S., 2022. Improving the classification accuracy of melanoma detection by performing feature selection using binary Harris hawks optimization algorithm. Soft Comput. 26 (17), 8163–8181.

Bozorg-Haddad, O., Solgi, M., Loáiciga, H.A., 2017. Meta-Heuristic and Evolutionary Algorithms for Engineering Optimization. John Wiley & Sons.

Brest, J., Zumer, V., Maucec, M.S., 2006. Self-adaptive differential evolution algorithm in constrained real-parameter optimization. In: 2006 IEEE International Conference on Evolutionary Computation. IEEE, pp. 215–222.

Chakraborty, A., Ghosh, K.K., De, R., Cuevas, E., Sarkar, R., 2021. Learning automata based particle swarm optimization for solving class imbalance problem. Appl. Soft Comput. 113, 107959.

Chamseddine, E., Mansouri, N., Soui, M., Abed, M., 2022. Handling class imbalance in COVID-19 chest X-ray images classification: Using SMOTE and weighted loss. Appl. Soft Comput. 129, 109588.

Chen, Y., Wang, Y., Gu, Y., He, X., Ghamisi, P., Jia, X., 2019. Deep learning ensemble for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12 (6), 1882–1897.

Chopard, B., Tomassini, M., 2018. Particle swarm optimization. In: An Introduction To Metaheuristics for Optimization. Springer, pp. 97–102.

Codella, N.C., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., et al., 2018. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, pp. 168–172.

Codella, N.C., Nguyen, Q.-B., Pankanti, S., Gutman, D.A., Helba, B., Halpern, A.C., Smith, J.R., 2017. Deep learning ensembles for melanoma recognition in dermoscopy images. IBM J. Res. Dev. 61 (4/5), 5–51.

Dong, Q., Gong, S., Zhu, X., 2018. Imbalanced deep learning by minority class incremental rectification. IEEE Trans. Pattern Anal. Mach. Intell. 41 (6), 1367–1381.

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542 (7639), 115–118.

Fargnoli, M.C., Kostaki, D., Piccioni, A., Micantonio, T., Peris, K., 2012. Dermoscopy in the diagnosis and management of non-melanoma skin cancers. Eur. J. Dermatol. 22 (4), 456–463.

Hosny, K.M., Kassem, M.A., Foaud, M.M., 2019. Classification of skin lesions using transfer learning and augmentation with Alex-net. PLoS One 14 (5), e0217293.

Huang, G.-B., Wang, D.H., Lan, Y., 2011. Extreme learning machines: a survey. Int. J. Mach. Learn. Cybern. 2, 107–122.

İlkin, S., Gençtürk, T.H., Gülağız, F.K., Özcan, H., Altuncu, M.A., Şahin, S., 2021. hyb-SVM: Bacterial colony optimization algorithm based SVM for malignant melanoma detection. Eng. Sci. Technol., Int. J. 24 (5), 1059–1071.

Jain, S., Pise, N., et al., 2015. Computer aided melanoma skin cancer detection using image processing. Procedia Comput. Sci. 48, 735–740.

Johnson, J.M., Khoshgoftaar, T.M., 2019. Survey on deep learning with class imbalance. J. Big Data 6 (1), 1–54.

Kassem, M.A., Hosny, K.M., Fouad, M.M., 2020. Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning. IEEE Access 8, 114822–114832.

Kato, J., Horimoto, K., Sato, S., Minowa, T., Uhara, H., 2019. Dermoscopy of melanoma and non-melanoma skin cancers. Front. Med. 6, 180.

Kaur, R., GholamHosseini, H., Sinha, R., Lindén, M., 2022. Melanoma classification using a novel deep convolutional neural network with dermoscopic images. Sensors 22 (3), 1134.

Khan, M.S., Alam, K.N., Dhruba, A.R., Zunair, H., Mohammed, N., 2022. Knowledge distillation approach towards melanoma detection. Comput. Biol. Med. 105581.

Knackstedt, T., Knackstedt, R.W., Couto, R., Gastman, B., 2018. Malignant melanoma: diagnostic and management update. Plast. Reconstr. Surg. 142 (2), 202e–216e.

Lallas, A., Argenziano, G., Zendri, E., Moscarella, E., Longo, C., Grenzi, L., Pellacani, G., Zaludek, I., 2013. Update on non-melanoma skin cancer and the value of dermoscopy in its diagnosis and treatment monitoring. Expert Rev. Anticancer Ther. 13 (5), 541–558.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.

Levin, J., Nalebuff, B., 1995. An introduction to vote-counting schemes. J. Econ. Perspect. 9 (1), 3–26.

Li, D.-C., Liu, C.-W., Hu, S.C., 2010. A learning method for the class imbalance problem with medical data sets. Comput. Biol. Med. 40 (5), 509–518.

Li, Y., Shen, L., 2018. Skin lesion analysis towards melanoma detection using deep learning network. Sensors 18 (2), 556.

Li, Y., Zhang, J., Zhang, S., Xiao, W., Zhang, Z., 2022. Multi-objective optimization-based adaptive class-specific cost extreme learning machine for imbalanced classification. Neurocomputing 496, 107–120.

Ling, C.X., Sheng, V.S., 2008. Cost-sensitive learning and the class imbalance problem. Encycl. Mach. Learn. 2011, 231–235.

Lott, J.P., Elmore, J.G., Zhao, G.A., Knezevich, S.R., Frederick, P.D., Reisch, L.M., Chu, E.Y., Cook, M.G., Duncan, L.M., Elenitsas, R., et al., 2016. Evaluation of the melanocytic pathology assessment tool and hierarchy for diagnosis (MPATH-dx) classification scheme for diagnosis of cutaneous melanocytic neoplasms: Results from the international melanoma pathology study group. J. Am. Acad. Dermatol. 75 (2), 356–363.

Maier, H.R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L.S., Cunha, M.C., Dandy, G.C., Gibbs, M.S., Keedwell, E., Marchi, A., et al., 2014. Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions. Environ. Model. Softw. 62, 271–299.

Maier, H.R., Razavi, S., Kapelan, Z., Matott, L.S., Kasprzyk, J., Tolson, B.A., 2019. Introductory overview: Optimization using evolutionary algorithms and other metaheuristics. Environ. Model. Softw. 114, 195–213.

Mayo Clinic, 2022. Melanoma, https://www.mayoclinic.org/diseases-conditions/melanoma/symptoms-causes/syc-20374884.

NHS, 2022. Skin cancer melanoma, https://www.nhs.uk/conditions/melanoma-skin-cancer.

Nikolouzakis, T.K., Falzone, L., Lasithiotakis, K., Krüger-Krasagakis, S., Kalogeraki, A., Sifaki, M., Spandidos, D.A., Chrysos, E., Tsatsakis, A., Tsiaoussis, J., 2020. Current and future trends in molecular biomarkers for diagnostic, prognostic, and predictive purposes in non-melanoma skin cancer. J. Clin. Med. 9 (9), 2868.

Oh, Y., Park, S., Ye, J.C., 2020. Deep learning COVID-19 features on CXR using limited training data sets. IEEE Trans. Med. Imaging 39 (8), 2688–2700.

Oliveira, R.B., Pereira, A.S., Tavares, J.M.R., 2019. Computational diagnosis of skin lesions from dermoscopic images using combined features. Neural Comput. Appl. 31 (10), 6091–6111.

Omran, M.G., Salman, A., Engelbrecht, A.P., 2005. Self-adaptive differential evolution. In: Computational Intelligence and Security: International Conference, CIS 2005, Xi'an, China, December 15-19, 2005, Proceedings Part I. Springer, pp. 192–199.

Pham, T.-C., Luong, C.-M., Visani, M., Hoang, V.-D., 2018. Deep CNN and data augmentation for skin lesion classification. In: Intelligent Information and Database Systems: 10th Asian Conference, ACIIDS 2018, Dong Hoi City, Vietnam, March 19-21, 2018, Proceedings, Part II 10. Springer, pp. 573–582.

Piepkorn, M.W., Barnhill, R.L., Elder, D.E., Knezevich, S.R., Carney, P.A., Reisch, L.M., Elmore, J.G., 2014. The MPATH-Dx reporting schema for melanocytic proliferations and melanoma. J. Am. Acad. Dermatol. 70 (1), 131–141.

Pradhan, N., Dhaka, V.S., Rani, G., Chaudhary, H., 2020. Transforming view of medical images using deep learning. Neural Comput. Appl. 32, 15043–15054.

Pradhan, N., Singh, V., Kumar, V., Goel, P., Dhaka, V.S., 2021. Conversion of two dimensional images into multi-view images of bone using deep learning. Comput. Methods Biomech. Biomed. Eng.: Imaging Vis. 9 (1), 106–113.

Qin, A.K., Suganthan, P.N., 2005. Self-adaptive differential evolution algorithm for numerical optimization. In: 2005 IEEE Congress on Evolutionary Computation, Vol. 2. IEEE, pp. 1785–1791.

Rajaram, N., Reichenberg, J.S., Migden, M.R., Nguyen, T.H., Tunnell, J.W., 2010. Pilot clinical study for quantitative spectral diagnosis of non-melanoma skin cancer. Lasers Surg. Med. 42 (10), 876–887.

Ren, M., Zeng, W., Yang, B., Urtasun, R., 2018. Learning to reweight examples for robust deep learning. In: International Conference on Machine Learning. PMLR, pp. 4334–4343.

Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., et al., 2021. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci. Data 8 (1), 1–8.

Saari, D.G., Merlin, V.R., 2000. A geometric examination of Kemeny's rule. Soc. Choice Welf. 17 (3), 403–438.

Saini, M., Susan, S., 2020. Deep transfer with minority data augmentation for imbalanced breast cancer dataset. Appl. Soft Comput. 97, 106759.

Salcedo-Sanz, S., Del Ser, J., Landa-Torres, I., Gil-López, S., Portilla-Figueras, J., 2014. The coral reefs optimization algorithm: a novel metaheuristic for efficiently solving optimization problems. Sci. World J. 2014.

Sánchez-Monedero, J., Pérez-Ortiz, M., Saez, A., Gutiérrez, P.A., Hervás-Martínez, C., 2018. Partial order label decomposition approaches for melanoma diagnosis. Appl. Soft Comput. 64, 341–355.

Sayed, G.I., Soliman, M.M., Hassanien, A.E., 2021. A novel melanoma prediction model for imbalanced data using optimized SqueezeNet by bald eagle search optimization. Comput. Biol. Med. 136, 104712.

Skin Cancer Foundation, 2022. Melanoma overview, https://www.skincancer.org/skin-cancer-information/melanoma.

Smaoui, N., Bessassi, S., 2013. A developed system for melanoma diagnosis. Int. J. Comput. Vis. Signal Process. 3 (1), 10–17.

Srivastava, G., Chauhan, A., Jangid, M., Chaurasia, S., 2022a. CoviXNet: A novel and efficient deep learning model for detection of COVID-19 using chest X-Ray images. Biomed. Signal Process. Control 103848.

Srivastava, G., Chauhan, A., Kargeti, N., Pradhan, N., Dhaka, V.S., 2023a. ApneaNet: A hybrid 1DCNN-LSTM architecture for detection of Obstructive Sleep Apnea using digitized ECG signals. Biomed. Signal Process. Control 84, 104754.

Srivastava, G., Chauhan, A., Pradhan, N., 2023b. CJT-DEO: Condorcet's Jury Theorem and Differential Evolution Optimization based ensemble of deep neural networks for pulmonary and Colorectal cancer classification. Appl. Soft Comput. 132, 109872.

Srivastava, G., Pradhan, N., Saini, Y., 2022b. Ensemble of Deep Neural Networks based on Condorcet's Jury Theorem for screening Covid-19 and Pneumonia from radiograph images. Comput. Biol. Med. 149, 105979.

Sun, J., Li, J., Fujita, H., 2022. Multi-class imbalanced enterprise credit evaluation based on asymmetric bagging combined with light gradient boosting machine. Appl. Soft Comput. 130, 109637.

Tan, T.Y., Zhang, L., Lim, C.P., 2019. Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models. Appl. Soft Comput. 84, 105725.

Wang, Y., Cai, J., Louie, D.C., Wang, Z.J., Lee, T.K., 2021. Incorporating clinical knowledge with constrained classifier chain into a multimodal deep network for melanoma detection. Comput. Biol. Med. 137, 104812.

Xiao, W., Zhang, J., Li, Y., Zhang, S., Yang, W., 2017. Class-specific cost regulation extreme learning machine for imbalanced classification. Neurocomputing 261, 70–82.

Xie, F., Fan, H., Li, Y., Jiang, Z., Meng, R., Bovik, A., 2016. Melanoma classification on dermoscopy images using a neural network ensemble model. IEEE Trans. Med. Imaging 36 (3), 849–858.

Yu, L., Chen, H., Dou, Q., Qin, J., Heng, P.-A., 2016. Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans. Med. Imaging 36 (4), 994–1004.

Yu, Z., Jiang, F., Zhou, F., He, X., Ni, D., Chen, S., Wang, T., Lei, B., 2020. Convolutional descriptors aggregation via cross-net for skin lesion recognition. Appl. Soft Comput. 92, 106281.

Zhang, J., Zhao, Y., Shone, F., Li, Z., Frangi, A.F., Xie, S.Q., Zhang, Z.-Q., 2022. Physics-informed deep learning for musculoskeletal modelling: Predicting muscle forces and joint kinematics from surface EMG. IEEE Trans. Neural Syst. Rehabil. Eng..