



# **Automatika**

**Journal for Control, Measurement, Electronics, Computing and Communications**

**ISSN: (Print) (Online) Journal homepage: [www.tandfonline.com/journals/taut20](https://www.tandfonline.com/journals/taut20?src=pdf)**

# **Segmenting and classifying skin lesions using a fruit fly optimization algorithm with a machine learning framework**

**R. Sonia, Jesla Joseph, D. Kalaiyarasi, N. Kalyani, Amara S. A. L. G. Gopala Gupta, G. Ramkumar, Hesham S. Almoallim, Sulaiman Ali Alharbi & S.S. Raghavan**

**To cite this article:** R. Sonia, Jesla Joseph, D. Kalaiyarasi, N. Kalyani, Amara S. A. L. G. Gopala Gupta, G. Ramkumar, Hesham S. Almoallim, Sulaiman Ali Alharbi & S.S. Raghavan (2024) Segmenting and classifying skin lesions using a fruit fly optimization algorithm with a machine learning framework, Automatika, 65:1, 217-231, DOI: <u>[10.1080/00051144.2023.2293515](https://www.tandfonline.com/action/showCitFormats?doi=10.1080/00051144.2023.2293515)</u>

**To link to this article:** <https://doi.org/10.1080/00051144.2023.2293515>



# **Segmenting and classifying skin lesions using a fruit fly optimization algorithm with a machine learning framework**

R. Sonia<sup>a</sup>, Jesla Joseph<sup>b</sup>, D. Kalaiyarasi<sup>c</sup>, N. Kalyani<sup>d</sup>, Amara S. A. L. G. Gopala Gupta<sup>e</sup>, G. Ramkumar<sup>f</sup>, Hesham S. Almoallim<sup>g</sup>, Sulaiman Ali Alharbi<sup>h</sup> and S.S. Raghavan<sup>i</sup>

<span id="page-1-3"></span><span id="page-1-2"></span><span id="page-1-0"></span>aDepartment of Computer Applications, B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India; <sup>b</sup>School of CSA, REVA University, Bangalore, India; <sup>c</sup>Department of Electronics and Communication Engineering, Panimalar Engineering College, Chennai, India; <sup>d</sup>Department of Computer Science and Engineering, R. M. K College of Engineering and Technology, Thiruvallur, India; <sup>e</sup>Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India; <sup>f</sup>Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India; gDepartment of Oral and Maxillofacial Surgery, College of Dentistry, King Saud University, Riyadh, Saudi Arabia; <sup>h</sup>Department of Botany and Microbiology, College of Science, King Saud University, Riyadh, Saudi Arabia; <sup>i</sup> Department of Health Sciences, University of Texas Health Science Center, Tyler, USA

#### <span id="page-1-6"></span>**ABSTRACT**

The deadliest forms of skin cancer, melanomas have a large fatality rate. In the United States of America, 196,060 new cases of melanoma are anticipated in 2020. In the past, many automated methods for diagnosing skin lesions have been proposed, but they have not yet proven to be very accurate. Based on skin cells' exposure to sunlight, aberrant skin cell development frequently results in skin cancer. Ultraviolet radiation, viruses, bacteria, chemicals, and fungi are the main contributors to skin conditions. The creation of a precise computer-aided system for diagnosing breast cancer is of tremendous clinical importance. An improved machine learning framework has been developed in this research to detect skin lesions or skin cancer. Hence it is important to segment and classify the skin lesion. The research utilizes the fruit fly optimization algorithm and machine learning framework to segment and classifies skin disease and cancer. This platform's central idea is to use the fruit fly optimization algorithm (FOA) to improve two crucial SVM variables and create an FOA-based SVM (FOA-SVM) for the diagnosis of skin cancer. The integrative approach not only improves accuracy but also provides important data for more accurate classification.

# **1. Introduction**

Given that melanoma is one of the deadliest different cancers of all those that have been identified in mankind, it is a fascinating area for research in both diagnostic imaging and machine learning. The two primary types of skin cancer are mild and malignant, with melanoma being the most dangerous and deadly type if detected later. Cancer is brought on by repeated exposures to ultraviolet (UV) rays on the epidermis. Melanoma and benign skin cancers are typically the two major categories. Nevertheless, melanoma is a very invasive and fatal form of cancer with no typical signs and accounts for around 75% of all skin cancer cases [\[1\]](#page-14-0). The benign spread very slowly and is treatable if discovered in its initial phases. Because of this, if a diagnosis is made too late, cancer spreads further and develops quicker under the skin, and prescribed medications challenge and lower the odds of survival. Worldwide, there are around 133,000 new instances of melanoma diagnosed each year, and the number is

expected to rise. In the USA, 207,390 skin cases of cancer are expected to be discovered in 2021. This demonstrates that the anticipated mortality rate for 2021 is 4.9%. In 2010, melanoma-related fatalities were estimated to have claimed 49,110 lives globally. As per a report by the International Institute of Research on Cancer (IARC), there will be 17 million new cases of melanoma and around 9.6 million fatalities worldwide in 2018 [\[2\]](#page-14-1).

<span id="page-1-10"></span>Enlarged and lit imaging of a patch of skin is acquired using a process that starts with a simple, noninvasive skin scanning procedure, to provide a clearer view of the skin's patches. The detection of melanoma is frequently made through thermoscopic image assessment, which has substantially greater average accuracy than assessment with the unaided eye. However, a dermatologist's subjective review of microscopic images is typically laborious, prone to mistakes, and arbitrary (even a dermatologist with extensive training could come up with widely disparate diagnostic procedures).

<span id="page-1-9"></span>**CONTACT** G. Ramkumar **C** [pgrvlsi@gmail.com](mailto:pgrvlsi@gmail.com) **D** Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai 602105, Tamil Nadu, India

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License [\(http://creativecommons.org/licenses/by/4.0/\)](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

#### <span id="page-1-8"></span><span id="page-1-7"></span><span id="page-1-5"></span><span id="page-1-4"></span><span id="page-1-1"></span>**ARTICLE HISTORY**

Received 13 September 2023 Accepted 6 December 2023

**KEYWORDS** Skin lesion; machine learning; segmentation



**a** OPEN ACCESS **a** Check for updates

<span id="page-2-0"></span>Therefore, automatic recognition systems are extremely popular [\[3\]](#page-14-2). Due to inadequate information and a concentration on standardized activities like microscopic examination, which is the inspection of the skin employing an epidermal layer microscope, earlier work in dermatology computer-aided categorization has lacked the generalization performance of medical experts. With computer-aided diagnostics, skin disorders can be swiftly and accurately classified to provide treatments depending on medical complaints. This research offers a reliable method that, using supervisory techniques that cut down on diagnostic expenses, can accurately describe skin illnesses. The progression of sick growth is evaluated using a grey-level co-occurrence grid. The reliability of the diagnosis is important for a thorough evaluation of the abnormalities for improved therapy and lower pharmaceutical expenses [\[4\]](#page-14-3).

<span id="page-2-2"></span><span id="page-2-1"></span>To enhance melanoma diagnostic accuracy, the procedure of the dermoscopic image was created. Dermoscopy is a non-invasive skin scanning approach that enlarges and illuminates a skin region to provide a clearer picture of the spots. Reducing surface reflection improves the skin lesion's visual impact. However, due to some obstacles, automatically detecting cancer from dermoscopy images is a challenging process. First, it is challenging to identify precise lesion locations due to the limited contrast between skin conditions and actual skin region. Secondly, this might be challenging to discern between melanoma and non-melanoma lesions due to the large chance of visual similarities between the two [\[5\]](#page-14-4). UV radiation-induced skin tissue destruction is the primary cause of these skin cancer types. Visual inspection by a dermatologist is a typical clinical approach for skin cancer diagnosis. The accuracy of the diagnostic techniques may be misleading. Dermoscopy is a non-invasive diagnostic technique that bridges the gap between clinical dermatologists and dermatology by allowing the presentation of morphological characters that are invisible to the unaided eye. With various methods, including solar scans, epiluminescence (XLM), pubescence microscopy (ELM), cross-polarization, and side transillumination, the morphological features could be more clearly seen. As a result, additional clinical diagnoses are provided to the dermatologist. Comparing dermoscopy to a nondiscrete eye, diagnostic performance can be increased by 10% to 30% [\[6\]](#page-14-5).

<span id="page-2-3"></span>For several decades already, skin malignancies have been examined using the biopsy technique. Although it is the easiest technique accessible, its accuracy is in doubt. The ABCDE principle and the seven-point criteria are two further testing techniques. These procedures, nevertheless, call for skilled dermatology. Medical experts now diagnose skin cancer by using microscopic pictures and dermoscopy. Portable webcams could be used to see the low-resolution microscopic <span id="page-2-4"></span>examination. Dermoscopy, on either side, is a modern imaging technique that increases diagnosis precision and may help lower the risk of human extinction. Diagnosis produces high-resolution scans that reveal deeper skin features. Doctors with expertise check these photos visually [\[7\]](#page-14-6). Rules like the ABCDrule, 3-point checklist, 7-point checklist, and Menzies-rule are frequently used in procedures that depend on the rule for identifying the kind of skin conditions. Dermatologists have used these guidelines as their starting point for diagnosis and monitoring. In the ABCDrule, the letters ABCD stand in for asymmetrical, boundary structuring, colour variation, and diameter, correspondingly. Asymmetry denotes an imbalance between the two edges, whereas symmetrical denotes a match. This helps to differentiate between benign or malignant skin conditions. For instance, the colour composition for benign will always be one, however for cancerous, it could be two or more. Malignant tumors have a larger and wider diameter than benign tumors, which are always really microscopic, like a fraction of an inch in circumference [\[8\]](#page-14-7).

<span id="page-2-6"></span><span id="page-2-5"></span>Systems for computer-aided diagnostics (CAD) get around these issues, take up little time, and offer significantly higher accuracy than traditional conventional methods. Because CAD systems operate in some interconnected processes, including preprocessing, manual feature extraction, and finally categorization, which combines to limit the total classification performance, their history is not particularly outstanding (OA). These procedures first get rid of bubbles and artifacts before carrying out the subsequent task. In comparison, deep learning approaches outperform these algorithms and manual inspection. In deep learning feature extraction, convolutional neural networks (CNNs) with many convolutional layers are frequently utilized. On either side, several studies have attempted to recognize melanoma immediately using hand-crafted characteristics without the need for a segmentation process [\[9\]](#page-14-8). Deep learning networks utilize hierarchies to continuously features extracted, in contrast to methods that use hand-crafted characteristics. Some researchers have started to use deep learning algorithms for melanoma detection as a result of the advancements made by deep learning in an enormous array of image-processing jobs. In applied to measure melanoma, Research suggested a hybrid technique that used support vector machines (SVM), sparse coding, and convolutional neural networks (CNN). In a recent study, Codella and his coworkers developed a system integrating cuttingedge data mining and machine learning techniques for segmenting and classifying skin lesions. A convolutional network was used to extract multi-scale characteristics for melanoma identification. Different boundary recognition, selection, extraction, and classification methods have been used to identify numerous CAD systems. In addition to analysing and studying image

processing methods for skin cancer diagnosis, some researchers have also evaluated the diagnostic efficacy of CAD and Artificial Intelligence (AI) systems to that of the skilled dermatologist. To improve the diagnostic accuracy of automatic decision support, additional effort is needed to define and remove ambiguity [\[10\]](#page-14-9).

Data-driven diagnosis is necessary since skin disorders tend to take many different forms, there aren't enough skilled dermatologists, and they're dispersed incorrectly. The development of lasers and photonicsbased medical technologies has made it considerably easier and faster to identify skin conditions. The price of such a diagnosis is currently high and restricted. In terms of efficiency, deep learning models are relatively good at classifying data and images. In the realm of clinical diagnosis, there has been a need for exact anomaly detection and illness classification using CT, X-ray, Magnetic Resonance, PET, and signal data such as electrocardiograms and electroencephalograms. Quality therapeutic care would be possible because of the accurate classification of the diseases [\[11\]](#page-14-10). Deep neural networks can tackle complex issues by automatically identifying the features from the input data, and they are flexible enough to change as the issue under consideration does. Even with negligible computational models, deep neural networks will obtain the assumed knowledge to find and examine the features in the unexposed patterns in data, yielding a significant amount of performance. This inspired the authors to think of using a deep learning model to determine the skin disease category from an image of the affected area. Inside the field of medical image analysis, a variety of machine learning techniques have been widely applied to the categorization and extraction of features of skin lesions. These techniques have previously been used to evaluate a range of skin lesions before studying the variation of the lesions. This research uses some classification techniques to categorize skin lesions according to colour, skin type, and texture. As a result, the categorization and recognition of skin lesions are important for the skin diagnosis of diseases. Adaptive techniques have performance improvements in image processing applications recently [\[12\]](#page-14-11).

<span id="page-3-2"></span>The actual phenomena that these investigations were typically performed on the reduced characteristics on picture pixels and the high-level ones that were discarded can be discovered in all of the preceding works, which indicates that these researches may well not reflect prior clinical qualifications. As a result, researchers suggested in this work that the diagnosis of breast cancer is made utilizing high-level features that were specified based on previous medical expertise. This description is based on the work of two highly qualified pathologists. Doctors with medical knowledge have a higher ability to distinguish among breast tumors and cancer in over-all and have greater comprehension since these characteristics also include

<span id="page-3-3"></span><span id="page-3-0"></span>knowledge of the doctor [\[13\]](#page-14-12). Then, for separating malignant breast cancers from normal skin, we suggested a unique learning architecture support vector Regression. Penalty ratio and function width are the two critical criteria in traditional SVM, and they are typically handled via search algorithm and stochastic gradient. Nevertheless, it is simple to incorporate these techniques into local optimal solutions. Finding the global perfect option has lately become simpler thanks to several bio-inspired metaheuristic search algorithmic, including genetic algorithms (GA), the fruit fly optimization (FOA),particle swarm optimization algorithms (PSO), and moth-flame optimization (MFO). FOA a new method in the swarm intelligence family is motivated by the foraging habits of actual fruit flies. The FOA has some exceptional qualities, including a straightforward mathematical formulation, straightforward construction, and straightforward comprehension with just a few system parameters. Due to its beneficial characteristics, FOA has developed into a helpful tool for numerous issues in the real world [\[14\]](#page-14-13).

<span id="page-3-4"></span><span id="page-3-1"></span>Traditional dermatological diagnostic methods, including manual microscopic image assessment, suffer from labour-intensive processes, potential human errors, and subjectivity. The research introduces automatic recognition systems leveraging machine learning, aiming to mitigate these challenges and enhance diagnostic accuracy by implementing a more objective and efficient approach through computational analysis of skin lesions.

<span id="page-3-5"></span>Like other swarm intelligence techniques, FOA is an optimization algorithm approach that, when contrasted to gradient descent and grid searching, finds the global optimum or an approximation best decision more quickly. Nevertheless, the convergence speed of the conventional FOA technique just isn't very good, and it may result in the local optimum for complicated optimization algorithms. To accelerate convergence while lowering the likelihood that FOA will enter the local optimum, this research work presents a mechanism for updating fruit fly placements. This methodology has been extensively employed to improve numerous heuristic algorithms. The two important variable pairs in this article's SVM approach, the penalty component and the width of the kernel function, were optimized using the modified FOA method to provide the best model (FOA-SVM) [\[15\]](#page-14-14). Based on a high-level characteristics database, this system would also be tested for the detection of cancer. According to our knowledge, this research is the first to resolve the SVM with FOA variable optimization process. An in-depth contrast of LOA – SVM, back propagation neural network (BPNN), FOA-SVM (system based on the primitive fruit fly optimization technique),random forest (RF), GA-SVM (prototype based on the genetic methodologies), PSO-SVM (prototype due to the particle swarm optimization algorithms), and SVM was

done in the test using a 10-fold cross-validation method on the data. The empirical findings showed that in classification performance, sensitivity, Mathew's correlation coefficient (MCC), and specificity, the suggested FOA-SVM was preferable to previous approaches. The research proposes an integrative approach using the fruit fly optimization algorithm and a machine learning framework to segment and classify skin lesions. The focus is on optimizing support vector machine (SVM) variables through FOA, leading to improved accuracy and more reliable classification compared to previous methods.

# **2. Related works**

Inside the fusion of handcrafted features connected to the clinical algorithmic rules, a new Computer-Aided Detection (CAD) approach for the identification and tracking of harmful skin conditions (cancer subtype) is provided in this study. Deep learning characteristics using Mutual Information assessments include asymmetrical boundaries, colours, dermatoscopic features, and characteristics. To identify the Region of Interest, a lesion picture is improved, processed, and segmented during the pre-processing phase. The feature extraction method is then carried out. Shape, colour, and texture are examples of handcrafted features that are utilized to symbolize the ABCD rule, while DL features are retrieved using a CNN model structure that has been pre-trained on the ABCD rule. Utilizing MI assessment as a fusion rule, the most crucial data from both kinds of characteristics are gathered. Furthermore, numerous techniques are used during the classification process, including Support Vector Machines (SVMs), Linear Regression (LR), and Relevant Vector Machines (RVMs). The available database from ISIC 2018 was used to evaluate the developed system. In terms of classification reliability, precison, and sensitivity attained in the training and validation phases, the conceptual approach seems to perform better than other cuttingedge approaches. Furthermore, researchers suggest and support a unique method for modifying evaluation measures for unbalanced information that are typical for various types of skin lesions. To enable the detection of many diseases present in the ISIC challenge dataset, the subsequent work includes developing a system for classification tasks and the use of both characteristics and features [\[16\]](#page-14-15).

<span id="page-4-0"></span>Early diagnosis and delineation of the lesion boundaries are crucial for precise malignant region identification and treatment options for skin lesions. Skin cancer incidence is greater than average, particularly melanoma, which is more dangerous due to the high probability of metastases. Consequently, early detection is essential for treating it before malignancy develops. The assessment and separation of lesion borders from dermoscopy images are done to solve this issue. A

variety of techniques have been utilized, from texture examination of the photographs to visual assessment of the photos. Nevertheless, due to the sensitivity required in surgical interventions or drug applications, the reliability of these techniques is poor for standard clinical therapy. This offers a chance to create an automatic system that is accurate enough to be applied in a clinical environment. The automated approach for fragmenting lesion borders proposed in this research, known as Res-Unet, integrates the U-Net and ResNet frameworks. Additionally, researchers removed hair using picture in painting, which considerably enhanced the classification outcomes. Researchers used the ISIC 2017 database to train our system and the PH2 database and the ISIC 2017 test set to verify it. On the ISIC 2017 test set and the PH2 dataset, our suggested model achieved Jaccard Index values of 0.772 and 0.854, respectively, which are competitive with the best state-of-the-art methods currently in use. To avoid the system from overfitting, the small amount of training data utilized must be heavily supplemented. Consequently, a huge database is required for the model to be more accurate and comprehensive [\[17\]](#page-14-16).

<span id="page-4-1"></span>Convolutional neural networks (CNNs) can classify photos of melanoma with accuracy on pace with that of a dermatologist, according to recent research. It has not yet been documented how well a Convolutional neural network training with solely medical photos of a pigmentation skin infection performed in a medical image analysis challenge when pitted against doctors. Researchers collected 5847 clinical photos of pigmented skin lesions from 3551 individuals for this research. Basal cell carcinoma and malignant melanoma were two examples of benign and pigmented skin lesions (nevus, seborrhoeic keratosis, senile lentigo, and hematoma). Researchers chose one image from each of the 666 individuals researchers selected randomly for the testing dataset, and we annotated the remaining images with region proposals to produce the training data (4732 images, 2885 patients). Then, using the training information to train a faster, region-based CNN (FRCNN), researchers evaluated the effectiveness of the algorithm using the testing dataset. Furthermore, researchers evaluated the diagnostic value of 10 board-certified dermatologists (BCDs) and ten dermatology trainees (TRNs) with FRCNN. The reliability of the FRCNN for six-class categorization was 86.2 percent, compared to 79.5 and 75.1 percent for the BCDs according to FRCNN, the accuracy, sensitivity, and specificity for two-class categorization were 91.6, 83.4 percent, and 94.percent; by BCD, 86.7, 86.4, and 86.8 percent; and by TRN, 85.6, 83.9, and 85.7 percent false positive rates and significant predictive scores for the FRCNN, BCD, and TRN, correspondingly, were 5.7 and 84.8, 13.5 and 70.6, and 14.2 and 68.6 percent. As an outcome, FRCNN's classification performance was superior to dermatologists. To

boost the forecasting of skin cancer, researchers intend to introduce this technology into society and encourage widespread usage of it [\[18\]](#page-14-17).

<span id="page-5-0"></span>Automated Computer-Aided Identification is necessary due to the propensity of skin disorders to present in a variety of ways, the scarcity and uneven distribution of trained dermatologists, and the requirement for prompt and correct diagnosis (CAD). By investigating Deep Learning's capacity to categorize thousands of skin illnesses, enhancing classification efficiency, and employing condition nomenclature, this study hopes to advance earlier work on CAD for dermatological. On two of the biggest publicly viewable skin large datasets, DermNet and ISIC Archive, researchers trained cutting-edge Deep Neural Networks. Researchers additionally used disease taxonomy, where available, to enhance the classification efficiency of these systems. With 80% correctness and a 98 percent Area Under the Curve for the classification of 25 disorders on DermNet, we create new stateof-the-art. Additionally, researchers established standards for categorizing all 622 distinct sub-classes in this dataset, achieving 67 percent accuracy and 98 percent AUC. Researchers diagnosed each of the seven disorders on the ISIC Registry with an accuracy level of 93% and an Area under the curve of 99%. This work demonstrates the huge potential of deep learning to identify a wide range of skin conditions with near-human accuracy and significantly higher consistency. Supporting medical professionals in mass screening utilizing medical or dermoscopic pictures may play a potential role in the practical real-time detection of skin diseases. Furthermore, the research on leveraging disease taxonomies to considerably increase the overall accuracy of classification demonstrates that, if such supplementary information is acquired, Deep networks can utilize it and enhance their classification efficiency [\[19\]](#page-14-18).

<span id="page-5-1"></span>When carried out manually, the investigation and identification of skin cancer disorders from skin lesions was always laborious. One of the main causes of this is the complexity of skin lesion images. Oil, hair and blood vessels, bubbles, and skin lines are among the noise and artifacts present in the photographs of skin lesions. Additionally, they have irregular borders, low contrast, and variable hues. Different computational methods have been developed in the past to help with skin lesion images in the identification and diagnosis of skin cancer disorders. Due to the disturbance of the aforesaid skin lesion traits, the methods now in use have been restricted. Skin cancer has recently been detected using machine learning approaches, particularly deep learning approaches. Nevertheless, these are currently restricted to skin lesion images with hazy and erratic boundaries and little difference between the sick lesion and healthy components. To assess and separate skin lesion photos, researchers

used a probability model for the improvement of a fully convolutional network-based deep learning programme in this research. The probabilistic model includes a fully linked conditional random variable with a Gaussian kernel and an effective mean-field approximation stochastic reasoning technique. The margins of skin lesions are further refined by the probabilistic model. The entire system is examined and tested using databases of skin lesions from ISBI 2017 and PH2 that are available to the general audience. The system performed better, with a 98 percent average accuracy [\[20\]](#page-14-19).

# <span id="page-5-2"></span>**3. Methodology**

The proposed method is based on the segmentation and classification of skin lesions. In the first stage, the dataset was collected and processed using the preprocessing technique and then passed to the segmentation process and the feature extraction is processed by using fruit fly optimization and the feature ranking and selection process is carried out and finally, the image classification procedure is done using SVM technique and skin lesion is classified. The techniques of preprocessing, segment, extraction of features, and image analysis used in this work to identify and classify images of skin lesions are shown in Figure [1.](#page-6-0) Additionally, the focus has been placed on accurately identifying and classifying three skin conditions, including three type of disease mentioned in the Figure [1](#page-6-0) [\[21\]](#page-14-20).

<span id="page-5-3"></span>Feature ranking and selection are crucial for precise classification and reducing dimensionality. The Eigenvector Centrality feature ranking and selection (ECFS) method is employed in this research. ECFS assigns grades to each feature based on their importance, allowing for the identification of essential components that contribute to more accurate skin lesion classification. This process aids in improving classification accuracy and reducing errors.

## *3.1. Dataset*

Database arrangement is essential for classifying various skin problems. An image database is compiled from several sources for investigations. Online medical data archives, research issues, and experts in the field are all resources. The online information sources include DermQuest, DermIS, PH2 dataset, and DermNZ. Images of healthy people are taken from the publicly accessible database source "11k hands." Investigators who specialize in the identification of skin lesions have provided some photographs of eczema and normal categories. Since 2016, the IEEE International Symposium on Biomedical Imaging (ISBI) has hosted an annual global skin lesion categorization competition. For this effort, a uniform database has been produced after gathering all the information from diverse sources [\[1\]](#page-14-0). The



<span id="page-6-0"></span>**Figure 1.** Schematic Work of FOA-SVM method.

**Table 1.** Statistic data.

<span id="page-6-1"></span>

Group	No of Image
Malignant	300
Benign	300
Psoriasis	300
Eczema	300
Acne	300
Healthy	300
Overall	1800

classification method may lean towards categories with more photos; hence it is crucial to handle data imbalance when training the network. In light of this, the database was balanced using a layered sampling strategy. A randomized down-sampling strategy is used after the set of data, which was obtained from the resources indicated above, is arranged according to the disease characteristics. The database in the other groups is down-sampled to make the database balanced because the psoriasis group includes the fewest photos (N = 300). 1800 photos with a combined size of  $227 \times$  $227 \times 3$  were also utilized to train and test the classification method after down-sampling. Table [1](#page-6-1) presents the detailed database breakdown used here study project [\[8\]](#page-14-7).

## *3.2. Preprocessing*

The hair and poor light levels are the main causes of the skin images' regular susceptibility to disturbance. The preprocessing procedure has been used to enhance generalization skills while also removing image noise and artifacts. The three stages of the preprocessing

procedure are colour constancy, normalizing, noise reduction, and hair removal.

The preprocessing stage plays a vital role in improving generalization skills by addressing issues such as image noise, artifacts, and variations in lighting conditions. Techniques like colour constancy, normalization, and noise reduction ensure that the skin lesion images are standardized, leading to more reliable feature extraction and subsequent classification.

# *3.3. Colour constancy*

The medical photos utilized in this work were produced utilizing several dermoscopy techniques under various lighting conditions. The grey world consistency approach is used to correct and normalize the colour alteration that uneven light intensity caused in the image. The procedure measured the average grey scale of the red  $(R)$ , green  $(G)$ , and blue  $(B)$  coordinates and assessed the colour image underneath an unknown source of light.

# *3.4. Normalization*

<span id="page-6-2"></span>The range of pixel intensity levels is changed by the normalization technique. For this investigation, the medical photos were gathered from various sources. Researchers utilize the normalizing method to improve the dynamic range and pixel intensities value just after the colour constancy phase. The complete training and test image database has now been removed from the average measured RGB values [\[22\]](#page-14-21).

#### *3.5. Noise and hair removal process*

<span id="page-7-0"></span>And then using segmented or extraction of features approaches, the noise reduction, and hair artifacts reduction process is crucial for precise prevention of skin lesions. Because of various in-image distortions, classification findings are not always precise. During the procurement procedure, the noise was removed using a Gaussian filter. There just aren't many distinct hair hues in the photograph, removing thick hair is challenging. The persistence of hair colour results in numerous unintended lesion area expansions. The image segmentation structure is changed as a result. To find and remove hair from the skin disease region, some segmentation methods were used. For the goal of hair removal, different thresholding and filtering techniques have been applied. However, this method is difficult due to inaccurate identification, and data is lost in the lesion area [\[23\]](#page-14-22). The elimination of the right object from the skin lesion region is ensured by the careful selection of the structuring element (SE), such as element structure and dimensions. The segmentation algorithms have eliminated the hair and other items from the lesion region. If the smoother is darker than their other sections, structural bottom hat filtration is used. To eliminate hair from the lesion region with a width based on hair thickness, the circular structure factor is important. For the hair colour to be removed from the image, the kernel radius is crucial. The ratio of the lengths of the major and minor axes has been used to gauge hair thicknesses. The SE means removing thin hairs at a range of four pixels, and it does not leave any hair shading in the main image when thick hairs are removed at a range of six or eight pixels. The thin or thick hairs were efficiently eliminated following the size of the SE [\[24\]](#page-14-23).

# <span id="page-7-1"></span>**4. Segmentation and detection**

The skin lesion has been segmented using the segmentation technique following preprocessing. Due to the complexity and modification of skin disease, segmenting the region of a skin lesion is typically a interesting task. The size of SE was chosen for the current work depending on classification results. For structural closure procedures, the SE employs a circular kernel since the preponderance of skin disorders has a circular shape. After removing the closed structural picture and the complements of the initial grey image, the precise skin lesion region was obtained. This approach has produced a minor distinction between the lesion site and background. After performing the morphology closure procedure, the segmentation skin lesion region was obtained. The range of the criterion was then determined by lowering the image variability [\[25\]](#page-14-24). Once the likelihood and histogram for the equilibrium quantity of the image's minimum to maximum <span id="page-7-3"></span>intensity have indeed been acquired, the variation and average have indeed been computed. To remove the shade corners, a circular quantum mask was applied to the segmentation process. The segmentation portion of the lesion's dimension lengths has caused the circle's size to change. Additionally, the circle's radius is bigger than its primary axis' length. This binary cylindrical mask, whose position and circumference had to be altered, has expanded by the segmentation portion since its development. The border of a skin lesion is crucial for identifying illnesses. The morphological gradient method has been used for border identification, although this approach increases the lesion edge. The extracted features in this research begin to deteriorate due to the SE's circular kernel, which has a two-pixel diameter. The differential between the initial segmentation process and the degraded version is then calculated. This method locates the pixel location with the greatest intensity changes [\[26\]](#page-14-25). The segmentation process, employing circular structuring elements, plays a crucial role in achieving precise identification and separation of skin lesion regions from the background. Circular kernels are specifically chosen due to their compatibility with the circular shapes commonly found in skin disorders. Through structural closure procedures using circular elements, the segmentation process effectively captures and isolates the distinctive features of skin lesions. The circular structuring elements aid in highlighting boundaries, ensuring a nuanced definition between the lesion and the surrounding background. This approach enhances accuracy by aligning with the typical shapes of skin disorders. By tailoring the segmentation technique to the characteristics of skin disorders, particularly their circular shapes, the method ensures that the identified regions align closely with the actual boundaries of the lesions. Consequently, this precision in segmentation is pivotal for subsequent analysis and classification, forming a foundation for reliable dermatological diagnostics.

## **5. Feature extraction**

<span id="page-7-2"></span>Given that several conditions may share traits, feature extraction for multi-disease categorization is a very hard and difficult process. The variety of skin lesions makes it an essential task as well. For instance, skin cancer images make it simple to retrieve shape information because they have a distinct boundary and a specific size, but images of acne, eczema, and psoriasis make it more difficult because these conditions may completely cover the body in the image being recorded and lack a distinct shape. This research project proposes a set of features that may be retrieved from any photograph of a skin lesion. 35 colour-texture characteristics are collected from the skin lesion photos in the feature extraction process to classify the images into many categories. For the feature extraction process fruit fly

algorithm is used and the following section provides a brief section about the fruit fly algorithm [\[27\]](#page-15-0).

## *5.1. Fruit-Fly algorithm (FOA)*

A meta-heuristic algorithm is introduced the fruit fly optimization algorithm (FOA) was developed as a result of fruit fly foraging performance. Fruit flies arrange food when hunting by using their clear perception of a smell. When addressing optimization issues, FOA utilizes the fruit fly's flight pattern to find the optimum solution. In FOA, the colony of fruit flies (possible resolution) is initially produced at random in the solution space, and then each fruit flier updates its site following its mode of flight. Throughout the iteratively, the fruit fly population steadily increases its efficiency (grade of solutions) [\[28\]](#page-15-1). One of the clever improvement computations is FOA. It is very easy to set up, simple to carry out, and quick to improve. In either regard, it also has a few drawbacks. To determine the beginning reaches a point during the variable implement new, FOA uses a randomized process. Dazzle kinds of equipment are used during the time of natural product fly individual position update. It takes time to sign up and is simple to get remarkable qualities. Currently, there are many evaluation standards for classifier performance. Estimates for classifier streamlining always use classification accuracy and error rate as the fundamental measures.

FOA is employed to optimize SVM hyperparameters by creating a population of fruit flies that represents possible solutions in the solution space. Through iterations, each fruit fly updates its position based on the SVM's fitness, gradually improving the overall efficiency of the population. This process continues until the global optimum, representing the best set of SVM hyperparameters, is achieved, enhancing the model's performance.

In skin lesion classification, the Fruit-Fly optimization aids feature extraction by adapting its foraging behaviour to identify relevant information. This bio-inspired metaheuristic, coupled with Eigen-vector Centrality feature ranking, ensures essential features are identified. The algorithm's ability to navigate complex parameter spaces allows for the extraction of discriminative features crucial for accurate classification.

In this paper, an unique evolution Support Vector Machine that uses the FOA approach is proposed, along with the resulting FOA-SVM. The two crucial hyperparameters for SVM can be determined by the BMC framework flexibly. Figure [2](#page-9-0) illustrates the suggested technique's integrated framework. The inner process to improve and the assessment of the outer classifier performance make up the bulk of the suggested model [\[29\]](#page-15-2). The FOA technology utilizes the 5-fold cross-validation (CV) research to continuously alter the SVM variables during the internal variable optimization problem.

<span id="page-8-0"></span>The Fruit-Fly Optimization algorithm optimizes SVM hyperparameters by mimicking fruit fly foraging behaviour. This bio-inspired approach offers advantages in complex parameter spaces, providing a more efficient and faster convergence to global optima. The randomized generation of a fruit fly population in solution space enables effective exploration, making it particularly beneficial for optimizing SVM variables.

The 10-fold CV assessment is then used to accomplish the categorization job for diagnosis of skin cancer in the outer loop of the SVM forecasting model utilizing the discovered optimal process parameters. The optimization algorithm was based on classification results presented in equation (1):

$$
fitness = \left(\sum_{u=1}^{N} ACC_u\right) / n \lim_{x \to \infty} \tag{1}
$$

<span id="page-8-1"></span>In the above equ (1) the average accuracy is represented by  $ACC_u$  and it is attained by the SVM classifier through 5-fold cross-validation.

The following steps are directed by the FOA-SVM defined as follows:

- Stage 1: Set the input variable for FOA and comprise the maximum number of iterations, population size, upper and lower bound variables, and the dimensions of the problems.
- Stage 2: Depending on the lower and upper constraints of the factors, the placement of the fruit fly swarm was created at random.
- Stage 3: Depending on the position of the fruit-fly the primary population for FOA is generated.
- Stage 4: Utilizing SVM and the positioning of the fruit fly as a variable, assess the fitness of the overall population of fruit flies.
- Stage 5: Represent a general optimum, where the best fruit fly is located, to be the position of the fruit-fly.
- Stage 6: Modify each fruit fly's location in the swarming process, and assess each fly's fitness.
- Stage 7: If indeed the fitness of the top fly in the community is higher than the global optimal, upgrade the global optimum.
- Stage 8: Modify the iteration v, v-v + 1, v is greater than a maximum number of iterations then go back to stage 6.
- Stage 9: Give the ideal SVM parameter pair  $(F, \mu)$  as the global optimum.

## *5.2. Feature ranking and selection*

<span id="page-8-2"></span>A strategy for choosing the essential components to be used in the classification stage from among the many retrieved characteristics is called feature ordering. This approach provides more precise categorization and



<span id="page-9-0"></span>**Figure 2.** FOA-SVM flow diagram.

reduces errors. Additionally, reducing the dimensionality of the data would reduce the categorization challenge, the timing of extracting features, testing and training duration, as well as memory required space by the classification algorithm. This research has used the Eigen-vector Centrality feature ranking and selection (ECFS) method to assign grades to each characteristic [\[30\]](#page-15-3).

<span id="page-9-1"></span>Stage 1: Every node and vertex of the unfocused diagram indicates a certain feature. The boundaries among the characteristics that precede would be assessed. h stands for the whole feature, and the collection of features  $C = \{C^{(1)}, \cdots C^{(2)}\}$  . Here, researchers first assess the ordering quality and then construct an unfocused network  $I = \{X, Y\}$ , W, where X stands for the vertex and Y for the weighted edges between the characteristics. The neighboring matrix A that is associated with I is a  $\{d \times d\}$  a matrix that includes all types that can be used in pairs.  $f_{bn}$  describes pairwise weighted connections and bilateral provides an additional level among any characteristic that may be expressed in equation (2):

$$
f_{bn} = \omega(c^{(b)}, c^{(n)})
$$
 (2)

The feature grade is denoted as the  $\omega$  which is depending on the differential range.

Stage 2: Compute the feature estimate metrics, such as the similarity measure and fisher criterion  $(c_i)$ The kernel K satisfies the supervised or unsupervised parameters listed below by using  $(c_f)$  and  $(a_u)$  mentioned in the equation (3):

$$
c_f = \sum_{n=1}^{n} \frac{(\alpha_{b,n} - \alpha_b)^2}{\sigma_b^2}, \sigma_b^2 = \sum_{n=1}^{n} (\sigma_{b,n})
$$
 (3)

where  $a_{b,n}$  and  $a_{b,n}$  are the mean and SD that indicates set n in the b<sup>th</sup> feature whereas the entire b<sup>th</sup>characteristics dataset has  $\alpha_b$ *ando<sub>b</sub>*.

The important stage to estimate every feature of the mutual information  $a_u$ . The sample is specified as i and the t denotes the related class labels.

$$
a_u = \sum_{t \in t} \sum_{I \in v^{(b)}} L(i, t) \log \left( \frac{L(i, t)}{L(i) \cdot L(t)} \right) \tag{4}
$$

In the above equation(4),  $L(i,t)$  denotes the probability distribution. To compute the fisher criteria and mutual information each feature is computed.

Stage 3: The kernel Q is a  $d \times d$  matrix computed by using the dot product mentioned in the equation (5):

$$
Q = (c_f a_u^v) \tag{5}
$$

Stage 4: The  $\sigma$  for the entire feature pair  $(b, n)$  has given by  $\sigma(b, n) = max(\sigma^{(b)} \alpha^{(n)})$ , the following equation (6)



<span id="page-10-0"></span>**Figure 3.** Skin disease classification.

represents the adjacency matrix S is connected to the graph I have constructed as follows:

$$
S = \mathbf{J}Q + (1 - \mathbf{J})\sigma \tag{6}
$$

Here, I loading coefficient with a value between 0 and 1 is present. Finally, after computing S's Eigen-values and vectors, the ranking vector would be the Eigen-vector paired with the highest Eigen-values.

## *5.3. Classification process based on SVM*

SVM, or support vector machines, is a crucial image categorization technique. Finding the hyper-plane that optimizes the maximum margin is the main goal of the SVM classification. In comparison to other classification techniques, this classifier performs more accurately and produces high categorization accuracy. Additionally, the SVM avoids the issue of generalization that the neural networks faced.

The three-stage SVM classification process involves distinguishing between melanoma, seborrheic keratosis, and lupus erythematosus. SVM-RFE feature selection refines the classification by improving the accuracy of images that were incorrectly identified. This

comprehensive approach ensures accurate categorization and contributes to the overall effectiveness of the skin lesion classification system.

<span id="page-10-1"></span>The classifier was employed in the suggested study to recognize and categorize the skin lesions. To ensure accurate categorization at every stage of the classifier, classification performance should be improved. Here, researchers categorize the three skin conditions lupus erythematosus, seborrheic keratosis, and melanoma shown in Figure [3.](#page-10-0) To distinguish between photos of melanoma as well as other skin disorders, the first algorithm is applied. Melanoma is considered the positive category, and other disorders are considered the negative class. Before classifying skin lesions using the SVM search technique, researchers significantly distinguish the characteristics of melanoma from other skin abnormalities [\[31\]](#page-15-4). The seborrheic keratoses are indeed being separated from lupus erythematosus using the second classification. Classify the photos as seborrheic keratoses and other conditions in the binary classification stage, though. In this instance, others stand in for the incorrectly labelled pictures of melanoma, seborrheic keratoses, and lupus erythematosus. To improve the classification performance of images that were incorrectly identified and seborrheic keratoses from

Lupus Erythematosus, another SVM-RFE feature selection method has been provided. Finally, Lupus Erythematosus was distinguished from the other incorrectly classified photos using the third predictor. The SVM classification scheme is used in this work to categorize the three skin disorders [\[32\]](#page-15-5).

### <span id="page-11-1"></span>**6. Result and discussion**

This development progress has been made using the MATLAB R2017a program and an Intel Core i5- 2410M 2.3 GHz processor. Think about skin lesion images from several databases that range in size from  $540 \times 722 - 4499 \times 6748$  pixels. The entire frame should then be resized to 512 by 512 pixels. 3 types of skin diseases have been trained using instance photographs in the classification model. Assume that for the training stage, 80% of the training images were used to create 5264 photos, 72% of the training images created 4605 images, 64% and 56% of the training images created 3496, and 3290 images were utilized to create 3290 testing samples. The problem of overfitting and dataset disparity has been reduced by the use of the data augmentation technique. The data augmentation technique has been used in conjunction with the picture rotations and shift methods. Eight pixels are moved to the right and left in addition to moving the images by 900, 1800, and 2700 degrees. The skin lesion class image has indeed been expanded by using the data augmentation approach to ensure the data balance throughout the training step of the classification algorithm.

# *6.1. Performance measure*

Calculating the accuracy (AC), sensitivity (SN), computing image and specificity (SF), and similarity indices like the Jaccard similarity index (JSI) and the Dice similarity coefficient have all been used to test the efficiency of segmented (DSC). The formulas that follow describe how well segmentation is performed in the equations (7) to (9):

$$
Accuracy = \frac{T_{Po}^* + T_{Ne}^*}{T_{Po}^* + T_{Ne}^* + F_{Po}^* + F_{Ne}^*}
$$
 (7)

$$
Specificity = \frac{T_{Po}^*}{T_{PO}^* + F_{PO}^*}
$$
\n(8)

$$
Sensitivity = \frac{T_{Ne}^{*}}{T_{Ne}^{*} + F_{Po}^{*}} \tag{9}
$$

$$
JSI(U, V) = \left| \frac{U \cap V}{U \cup V} \right| \tag{10}
$$

Here, the letters  $T_{Po}^*, T_{Ne}^*, F_{Po}^*, F_{Ne}^*$ stand for false negative, true positive, false positive, and true negative pixels, respectively. According to the two sets of U and V Jaccard Similarity Index (JSI) which is presented in the equation (10):

<span id="page-11-0"></span>**Table 2.** Segmentation image efficiency.

			Segmentation					
Database	<b>Standards</b>	Acc	Spe	Sen	JSI	<b>DSC</b>		
A	Min	0.953	0.948	0.968	0.959	0.950		
	Max	0.993	0.985	0.985	0.991	0.986		
	Total	0.967	0.960	0.968	0.975	0.975		
B	Min	0.960	0.951	0.980	0.969	0.966		
	Max	0.992	0.985	0.990	0.985	0.983		
	Total	0.969	0.966	0.975	0.973	0.982		
C	Min	0.968	0.966	0.976	0.974	0.988		
	Max	0.994	0.988	0.985	0.990	0.993		
	Total	0.973	0.963	0.976	0.973	0.988		

The cardinal set U is denoted as  $|U|$ . Depending on  $F_{P_0}^*$ ,  $F_{Ne}^*$  and  $T_{P_0}^*$ , the JSI has indicated the following equation (11):

$$
JSI(U, V) = \frac{T_{Po}^*}{F_{Po}^* + F_{Ne}^* + T_{Po}^*}
$$
(11)

Similarly, the two groups M and N's Dice Similarity Coefficient (DSC) results are as follows in equation (12):

$$
DSC(U, V) = 2 * \frac{|U \cap V|}{|U| + |V|}
$$
 (12)

Here, U stands for the cardinal set. Depending on,  $F_{p_0}^*, F_{N_e}^*, T_{p_0}^*$ the DSC has indicated as follows in equation (13):

$$
DSC(U, V) = 2 * \left(\frac{T_{Po}^*}{2 * F_{Po}^* + F_{Ne}^* + T_{Po}^*}\right) \tag{13}
$$

The classification objective functions are listed in Table [2.](#page-11-0) The table below lists three parameters, including the Ham 10,000 information, the Xiangya derm dataset, and the Dermnet NZ database. The segment efficiency indices for minimum, maximum, and average values are then displayed in this table. Specificity, accuracy, sensitivity, JSI, and DSC have all been used to evaluate segmentation performance. The features extracted approach are shown below with estimates for the mean, entropy, and contrast. Three datasets were used in segmentation efficiency they are Dermnet Nz Database (A), Xiangya derm dataset (B), and Ham 10,000 dataset (C).

The following equations (14) to (16) represents the mean, entropy, and contrast:

$$
Mean = \frac{1}{V} \sum_{u} U.P_V(U) \tag{14}
$$

$$
Entropy = -\sum_{u} P_V(U) \cdot \log(P_V(U)) \qquad (15)
$$

$$
Contrast = \sum_{u} u^2. p_v(\mathbf{U})
$$
 (16)

The likelihood that  $X_Y(m, n)$  had the quantity one is represented by the c component of the probability density function  $P_V(U)$  for a pair of pixels separated by a distance V. The technique being suggested



**Original Image** 

<span id="page-12-0"></span>**Figure 4.** Skin hair removal process.



<span id="page-12-1"></span>**Figure 5.** Outcome of skin lesion segmentation.

includes preprocessing, segment and, feature extraction, feature ranking and selection, boundary detection and an image analysis technique, all of which were utilized to enhance the accuracy of the identification of various skin lesion disorders. Figure [4](#page-12-0) summarizes the preprocessing stage that involved removing hair and noise.

The segmentation approach is used to divide up the skin lesions after the preprocessing phase. In the segmentation procedure, the morphology closure procedure has been applied to separate the skin condition from the dermoscopic images. The segment is typically the more challenging technique due to its uneven architecture. Figure [5](#page-12-1) shows the original photos as well as segmentation skin lesion visuals.

The results of skin cancer detection using various databases are shown in Table [3.](#page-12-2) The Dermnet NZ dataset, the Xiangya derm database, and the Ham 10,000 databases containing training, and testing samples with specificity, accuracy, and sensitivity metrics for skin cancer detection, are the databases researchers



Preprocessing

**Table 3.** Outcome of skin lesion detection.

<span id="page-12-2"></span>

Dataset	Training Dataset	Testing Dataset	Accuracy	Sensitivity	Specificity
A	50	50	99.08	99.68%	99.80
B	90	10	99.55	98.99	98.13
C	80	20	98.92	98.10	99.15

used in this work, and the graphical representation is presented in Figure [6.](#page-13-0) A greater number of observations are being used for training, while a smaller number of these samples are being used for assessment. With 98.92 percent accuracy, 99.15 percent specificity, and 98.10 percent sensitivity, the Ham 10,000 database uses 90% training dataset and 10% testing samples. Additionally, parameters are validated for the Dermnet Nz dataset for 60% of training samples and 40% of testing samples with 99.80% specificity, 99.08% accuracy, and 99.69% sensitivity. With findings of 99.59 percent accuracy, 98.15 percent specificity, and 98.99 percent sensitivity, the Xiangya derm dataset produces 86% training samples and 17% testing samples.

Additionally, according to rankings, the FOA is the top approach overall, followed by the FPA, BA, SCA, MFO, DA, and Particle swarm optimization in that order are shown in Figure [7.](#page-13-1) Figure [8](#page-13-2) shows the convergence patterns of LFOA and other approaches for various test functions  $V_1$ ,  $V_2$ ,  $V_3$ ,  $V_4$ 

The proposed methodology of categorization has been evaluated with many existing approaches, including edge detection-based identification, gray-level thresholding, the k-clustering approach, and water shed-based categorization. The suggested segment offers the best segmentation accuracy when the proposed approach is contrasted with other current methods. The presented segmentation's accuracy, sensitivities, specific, JSI, and DSC efficiency ranges are 0.988, 0.985, 0.960, 0.980, and 0.975, respectively, as shown in Table [4](#page-13-3) and the graphical representation is presented in Figure [8.](#page-13-2)



**Outcome Detection** 

 $\blacksquare$  A  $\blacksquare$  B  $\blacksquare$  C



<span id="page-13-0"></span>

<span id="page-13-1"></span>**Figure 7.** FOA convergence curve.



<span id="page-13-2"></span> $\blacksquare$  Acc  $\blacksquare$  Sen  $\blacksquare$  Spec  $\blacksquare$  JSI  $\blacksquare$  DSC

**Figure 8.** Performance evaluation.

# **7. Conclusion**

An efficient FOA-SVM approach has been created in this research, this could accurately diagnose breast cancer in clinical diagnostic and give clinicians clinical decision-making information. Each dataset uses a

<span id="page-13-3"></span>**Table 4.** Performance evaluation.

	Efficiency				
Technique	Acc	Sen	Spec	JSI	<b>DSC</b>
<b>Grey Level</b>	0.745	0.852	0.935	0.955	0.965
K-clustering	0.875	0.905	0.950	0.899	0.950
Watershed	0.925	0.920	0.875	0.915	0.858
<b>Edge Detection</b>	0.955	0.935	0.870	0.925	0.920
FOA-SVM	0.988	0.985	0.960	0.980	0.975

different collection of testing and training samples. By evaluating the accuracy, sensitivity, specificity, JSI, and DSC, the efficiency has been calculated. To demonstrate its exceptional performance, the suggested technique has been compared to various existing methods that includes the different classifier. The Xiangya derm dataset, the Dermnet NZ database, and the Ham 10,000 dataset were used to choose the three skin lesion classes. The accuracy is 98 percent, specificity is 99 percent, sensitivity is 96 percent, JSI is 95 percent, and DSC is 99 percent for the suggested SVM approach, which offers improved classification efficiency. Even if the proposed technique performs better, it has a few limitations. Future research should focus on processes like the mutation approach and the opposition-based training approach that may be used to enhance the FOA approach. Additionally, we intend to use the technique to address other associated issues with a diagnosis of diseases. Future work aims to enhance the skin lesion classification system by exploring avenues to address potential limitations. Specific areas may include refining preprocessing techniques, optimizing feature extraction algorithms, and incorporating advanced machine learning models. The focus could also extend to incorporating additional datasets for increased diversity and robustness.

# **Disclosure statement**

No potential conflict of interest was reported by the author(s).

#### **Funding**

This project was supported by Researchers Supporting Project number [RSP2023R5], King Saud University, Riyadh, Saudi Arabia.

#### **References**

- <span id="page-14-0"></span>[1] Huang H, et al. A new fruit fly optimization algorithm enhanced support vector machine for diagnosis of breast cancer based on high-level features. BMC Bioinformatics. [2019;20\(S8\):290. doi:10.1186/s12859-019-](https://doi.org/10.1186/s12859-019-2771-z) 2771-z
- <span id="page-14-1"></span>[2] Afza F, Sharif M, Mittal M, et al. A hierarchical three-step superpixels and deep learning framework for skin lesion classification. Methods. [2022;](#page-1-10)202:88–102. doi[:10.1016/j.ymeth.2021.02.013](https://doi.org/10.1016/j.ymeth.2021.02.013)
- <span id="page-14-2"></span>[3] Khamparia A, Singh PK, Rani P, et al. An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. Trans. Emerg. Telecommun. Technol. [2021;](#page-2-0)32(7). doi[:10.1002/ett.3963](https://doi.org/10.1002/ett.3963)
- <span id="page-14-3"></span>[4] Srinivasu PN, SivaSai JG, Ijaz MF, et al. Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. Sensors. [2021;](#page-2-1)21(8):2852. doi[:10.3390/s21082852](https://doi.org/10.3390/s21082852)
- <span id="page-14-4"></span>[5] S RD, S A. Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM). Asian Pac J Cancer Prev [2019;20\(5\):1555–1561. doi:10.31557/APJCP.2019.20.](https://doi.org/10.31557/APJCP.2019.20.5.1555) 5.1555
- <span id="page-14-5"></span>[6] Adegun AA, Viriri S. FCN-Based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images. IEEE Access. [2020;8:150377–150396. doi:10.1109/ACCESS.2020.](https://doi.org/10.1109/ACCESS.2020.3016651) 3016651
- <span id="page-14-6"></span>[7] Kassem MA, Hosny KM, Damaševičius R, et al. Machine learning and deep learning methods for skin lesion classification and diagnosis: A systematic review. Diagnostics. [2021;11\(8\):1390. doi:10.3390/diagnostics](https://doi.org/10.3390/diagnostics11081390) 11081390
- <span id="page-14-7"></span>[8] Melbin K, Raj YJV. Integration of modified ABCD features and support vector machine for skin lesion types

classification. Multimed. Tools Appl. [2021;](#page-2-5)80(6):8909– 8929. doi[:10.1007/s11042-020-10056-8](https://doi.org/10.1007/s11042-020-10056-8)

- <span id="page-14-8"></span>[9] Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. Sensors. [2018;](#page-2-6) 18(2):556. doi[:10.3390/s18020556](https://doi.org/10.3390/s18020556)
- <span id="page-14-9"></span>[10] Goyal M, Oakley A, Bansal P, et al. Skin lesion segmentation in dermoscopic images With ensemble deep learning methods. IEEE Access. [2020;](#page-3-0)8:4171–4181. doi[:10.1109/ACCESS.2019.2960504](https://doi.org/10.1109/ACCESS.2019.2960504)
- <span id="page-14-10"></span>[11] El-Khatib H, Popescu D, Ichim L. Deep learning–based methods for automatic diagnosis of skin lesions. Sensors. [2020;](#page-3-1)20(6):1753. doi[:10.3390/s20061753](https://doi.org/10.3390/s20061753)
- <span id="page-14-11"></span>[12] Afza F, Sharif M, Khan MA, et al. Multiclass skin lesion classification using hybrid deep features selection and extreme learning machine. Sensors. [2022;](#page-3-2)22(3):799. doi[:10.3390/s22030799](https://doi.org/10.3390/s22030799)
- <span id="page-14-12"></span>[13] Ünver HM, Ayan E. Skin lesion segmentation in dermoscopic images with combination of YOLO and GrabCut algorithm. Diagnostics. [2019;9\(3\):72. doi:10.3390/diag](https://doi.org/10.3390/diagnostics9030072) nostics9030072
- <span id="page-14-13"></span>[14] Ikotun AM, Almutari MS, Ezugwu AE. K-Means-Based nature-inspired metaheuristic algorithms for automatic data clustering problems: recent advances and future directions. Appl Sci|3192 [2021;](#page-3-4)11(23):11246. doi[:10.3390/app112311246](https://doi.org/10.3390/app112311246)
- <span id="page-14-14"></span>[15] Brezočnik L, Fister I, Podgorelec V. Swarm intelligence algorithms for feature selection: A review. Appl Sci|3192 [2018;](#page-3-5)8(9):1521. doi[:10.3390/app8091521](https://doi.org/10.3390/app8091521)
- <span id="page-14-15"></span>[16] Almaraz-Damian J-A, Ponomaryov V, Sadovnychiy S, et al. Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures. Entropy. [2020;](#page-4-0)22(4):484. doi[:10.3390/e22040484](https://doi.org/10.3390/e22040484)
- <span id="page-14-16"></span>[17] Zafar K, et al. Skin lesion segmentation from dermoscopic images using convolutional neural network. Sensors. [2020;](#page-4-1)20(6):1601. doi[:10.3390/s20061601](https://doi.org/10.3390/s20061601)
- <span id="page-14-17"></span>[18] Jinnai S, Yamazaki N, Hirano Y, et al. The development of a skin cancer classification system for pigmented skin lesions using deep learning. Biomolecules. [2020;](#page-5-0)10(8):1123. doi[:10.3390/biom10081123](https://doi.org/10.3390/biom10081123)
- <span id="page-14-18"></span>[19] Bajwa MN, et al. Computer-Aided diagnosis of skin diseases using deep neural networks. Appl Sci|3192 [2020;](#page-5-1)10(7):2488. doi[:10.3390/app10072488](https://doi.org/10.3390/app10072488)
- <span id="page-14-19"></span>[20] Adegun AA, Viriri S, Yousaf MH. A probabilistic-based deep learning model for skin lesion segmentation. Appl Sci|3192 [2021;](#page-5-2)11(7):3025. doi[:10.3390/app11073025](https://doi.org/10.3390/app11073025)
- <span id="page-14-20"></span>[21] Hoang L, Lee S-H, Lee E-J, et al. Multiclass skin lesion classification using a novel lightweight deep learning framework for smart healthcare. Appl Sci|3192 [2022;](#page-5-3)12(5):2677. doi[:10.3390/app12052677](https://doi.org/10.3390/app12052677)
- <span id="page-14-21"></span>[22] Nawaz M, et al. An efficient deep learning approach to automatic glaucoma detection using optic disc and optic Cup localization. Sensors. [2022;](#page-6-2)22(2):434. doi[:10.3390/s22020434](https://doi.org/10.3390/s22020434)
- <span id="page-14-22"></span>[23] Kousis I, Perikos I, Hatzilygeroudis I, et al. Deep learning methods for accurate skin cancer recognition and mobile application. Electronics (Basel). [2022;](#page-7-0)11(9):1294. doi[:10.3390/electronics11091294](https://doi.org/10.3390/electronics11091294)
- <span id="page-14-23"></span>[24] Jain S, Singhania U, Tripathy B, et al. Deep learningbased transfer learning for classification of skin cancer. Sensors. [2021;21\(23\):8142. doi:10.3390/](https://doi.org/10.3390/s21238142) s21238142
- <span id="page-14-24"></span>[25] Banerjee S, Singh SK, Chakraborty A, et al. Melanoma diagnosis using deep learning and fuzzy logic. Diagnostics. [2020;](#page-7-2)10(8):577. doi[:10.3390/diagnostics10080577](https://doi.org/10.3390/diagnostics10080577)
- <span id="page-14-25"></span>[26] Kausar N, et al. Multiclass skin cancer classification using ensemble of fine-tuned deep learning models.

Appl Sci|3192 [2021;11\(22\):10593. doi:10.3390/app112](https://doi.org/10.3390/app112210593) 210593

- <span id="page-15-0"></span>[27] Courtenay LA, et al. Deep convolutional neural support vector machines for the classification of basal cell carcinoma hyperspectral signatures. J Clin Med [2022;](#page-8-0)11(9):2315. doi[:10.3390/jcm11092315](https://doi.org/10.3390/jcm11092315)
- <span id="page-15-1"></span>[28] Rey-Barroso L, Peña-Gutiérrez S, Yáñez C, et al. Optical technologies for the improvement of skin cancer diagnosis: A review. Sensors. [2021;](#page-8-1)21(1):252. doi[:10.3390/s21010252](https://doi.org/10.3390/s21010252)
- <span id="page-15-2"></span>[29] Rashid J, et al. Skin cancer disease detection using transfer learning technique. Appl Sci|3192 [2022;](#page-8-2)12(11):5714. doi[:10.3390/app12115714](https://doi.org/10.3390/app12115714)
- <span id="page-15-3"></span>[30] Fraiwan M, Faouri E. On the automatic detection and classification of skin cancer using deep transfer learning. Sensors. [2022;22\(13\):4963. doi:10.3390/](https://doi.org/10.3390/s22134963) s22134963
- <span id="page-15-4"></span>[31] Bakheet S. An SVM framework for malignant mela noma detection based on optimized HOG features. Computation. [2017;5\(1\):4. doi:10.3390/computation](https://doi.org/10.3390/computation5010004) 5010004
- <span id="page-15-5"></span>[32] Cesati M, et al. Investigating serum and tissue expression identified a cytokine/chemokine signature as a highly effective melanoma marker. Cancers (Basel). [2020;](#page-11-1)12(12):3680. doi[:10.3390/cancers12123680](https://doi.org/10.3390/cancers12123680)