International Research Journal of Multidisciplinary Scope (IRJMS), 2025; 6(1):1239-1252 IRJMS

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2025.v06i01.02938

Advanced Melanoma Detection Using Deep Transfer Learning

Shubhendu Banerjee[1], Soumya Bhattacharyya[2]*, Shambhu Nath Saha[2]

[1]Department of CSE, Narula Institute of Technology, Kolkata, West Bengal, India, Department of IT, [2]Narula Institute of Technology, Kolkata, West Bengal, India. *Corresponding Author's Email: sat.soumya@gmail.com

Abstract

Anywhere on the body might develop melanoma, a very serious form of skin cancer. Early detection of melanoma lesions significantly increases the likelihood of effective therapy. In recent times, learning-based segmentation techniques have outperformed conventional algorithms in the segmentation of images. This work presents a novel approach to improve the identification and classification of skin lesions with cancer. We propose a two-stage procedure based on deep learning models. To evaluate our approaches, we used the popular ISIC 2018 dataset, which is well-known for its Skin Lesion Analysis Towards Melanoma Detection Challenge. There are two primary parts to the suggested methods for segmenting and identifying lesions in real time. First, we use an enhanced version of You Only Look Once edition 8 (F-YOLOv8) to accurately localise skin lesions. We next use the updated Segmentation Network (F-SegNet) to handle the F-YOLOv8 data further. We conducted experiments on 20,250 photos from three publicly available datasets: The International Skin Imaging Collaboration (ISIC) 2019, International Symposium on Biomedical Imaging (ISBI) 2017, and PH2. The findings were promising. The suggested approach obtained accuracy of 98.50% and 98.50% on the PH2 and ISBI 2017 datasets, respectively, and a Jac score of 93.22% on the ISIC 2019 dataset. In most situations, our technique showed somewhat better performance compared to current efforts in this field utilising predetermined parameters.

Keywords: Melanoma, Skin Cancer, Skin Lesion Segmentation, YOLO.

Introduction

With its fast development and dissemination, skin cancer is one of the most common diseases detected globally. The unchecked proliferation of melanocyte cells is the cause of malignant melanoma skin tumours. Because of its greater death rate and propensity to spread to neighbouring tissues, this kind of melanoma is more hazardous than other skin tumours (1). There were 11,650 skin cancer-related fatalities in the US in 2022 due to 104,450 occurrences of skin cancer, including squamous cell and basal cell tumours. Melanoma is the primary cause of these instances, accounting for 92.5% of cases and 62.1% of fatalities from skin cancer (2, 3). Early skin lesion discovery may significantly lower mortality from melanoma. Skin malignancies are usually evaluated visually, which often results in inaccurate judgements. The most trustworthy imaging technique for analysing skin lesions is dermoscopy (4, 5). Dermatologists may improve the quality of their diagnosis and patient assessments by using a reliable, automated Computer-Aided Diagnosis (CAD) system. The goal of improving diagnostic performance and accuracy for melanoma diagnosis is to build an automated system for melanoma lesion segmentation (6-9). There are two types of automated melanoma segmentation techniques: deep learning algorithms and classical approaches (10-12). Conventional techniques include adaptive thresholding, Otsu's thresholding, level set segmentation, iterative stochastic area merging, and iterative selection threshold (13). These methods often suffer from natural or clinical artefacts that reduce their efficacy. Although Otsu's thresholding works effectively, it often results in asymmetric segment boundaries that lower the quality of the picture (14-16). The implementation of intensity averaging across pixels solved this problem. Dermoscopic image-based object identification algorithms have been used to identify melanoma-affected regions and improve segmentation results. For example, researchers used superpixel data to map the relevance of melanoma spots using a hyper-graph. Deep learning algorithms have recently surpassed conv-

(Received 23rd October 2024; Accepted 23rd January 2025; Published 13st January 2025)

entional techniques in melanoma segmentation, making substantial progress (17). This work suggests an integrated, two-phase diagnostic method that accurately identifies and segments skin lesions while lowering computational expenses by using a Convolutional Neural Network (CNN) (18). The primary objective is to increase the efficacy of diagnosis by accurately detecting and categorising melanoma lesions. There are two primary stages to the procedure. First, we identify melanoma lesions as target locations in our improved You Only Look Once version 8 network (F-YOLOv8). We then use these regions as inputs into our improved segmentation system (F-SegNet). With an emphasis on identification and segmentation, this work makes use of the PH2, ISBI 2017 and ISIC 2019 datasets, which comprises 20,250 photos and matching ground truth masks. Deep learning architectures for skin lesion diagnosis may benefit from the techniques and findings of this study. There is a lot of potential for integrating the suggested F-YOLOv8 and F-SegNet-based melanoma detection technique into clinical workflows. It can improve dermatologists' diagnostic capabilities by providing a dependable, automated solution for lesion localisation and segmentation, which lessens the need for manual interpretation and increases diagnostic accuracy. Effective screening is made possible by the system's real-time processing capabilities, which may lead to earlier melanoma identification and prompter actions. It may also be easily integrated into clinical settings, simplifying processes and lessening the workload for medical professionals, given its interoperability with current imaging technologies like dermoscopy and its capacity to manage huge datasets. Ultimately, by facilitating quicker, more precise diagnoses and individualised treatment regimens, this invention may enhance patient care.

The study's main contributions are:

- Multiple skin lesions associated with melanoma may be accurately detected in one picture using F-YOLOv8.
- F-YOLOv8 uses a deep neural network to precisely locate lesion regions in pictures. We provide a real-time lesion localisation technique that smoothly incorporates extracted regions into F-SegNet.
- Using the PH2, ISBI 2017 and ISIC 2019 datasets for assessment, F-SegNet demonstrates accurate melanoma lesion segmentation.

This is the format for the remainder of the paper: Details of datasets, segmentation, classification, and pre-processing are covered in Section II. Section III contrasts the computer-generated findings with new methods. An examination of the problems brought up and upcoming difficulties rounds off Section IV.

Methodology

This work offers a real-time approach to skin lesion segmentation for melanoma. There are two primary processes in the process: first, locating the lesion in the picture, and second, segmenting the lesion according to its geographic characteristics. Figure 1 shows the complete process, including segmentation and detection of lesions. All of the training set's data were scaled to 512 by 512 pixels prior to YOLO version 8 training. In order to specify the item to be recognised, the bounding box's height (h), width (w), and centre coordinates (x, y) are added to the picture data during the training phase. Using the coordinates of the upper-left (x1, y1) and lower-right (x2, y2) corners of the bounding box makes it easier to calculate these values, as Figure 2 illustrates.

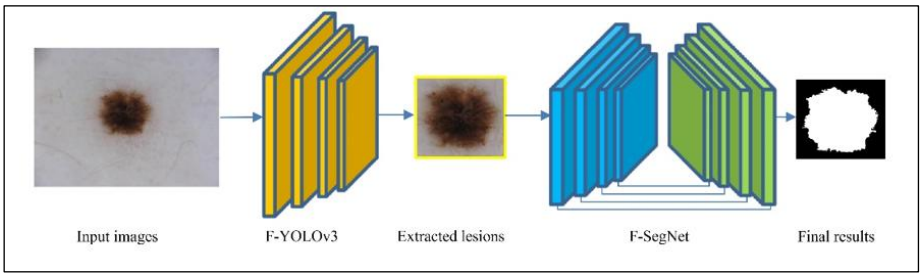

Figure 1: Localization and Segmentation of Melanoma Lesions

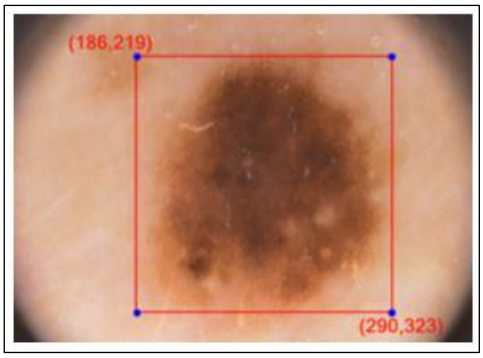

Figure 2: Skin Lesion Labelling For F-Yolov8 Training

Pre-processing

Errors and misinterpretations may happen while diagnosing malignant melanoma with the unaided eye. Dermatoscopy is often considered a more dependable option by medical professionals. Despite the high cost of dermatoscopy, research has made great strides in lowering prices without sacrificing picture quality. Although the suggested method shows encouraging results in detecting melanoma, further research is necessary to fully understand its limits when it comes to processing photos that include artefacts, lighting fluctuations, or a range of skin tones. The model's capacity to precisely locate and segment lesions may be hampered by images including artefacts like hair, shadows, or reflections, which might raise mistake rates. Similar to this, changes in illumination between datasets or in real-world environments may impact picture quality consistency, which may result in inaccurate segmentation or misclassification. Concerns about the model's generalisability in populations with different pigmentation are further raised by the fact that its performance over a range of skin tones is yet poorly understood. This study suggests a sophisticated preprocessing method to overcome these drawbacks and enhance the calibre of the images we were able to get for our investigation. Through a region-based identification approach, the initial step is cutting off the hair in the lesion area. In order to properly finish the process and improve the quality of the picture, we employ histogram equalisation in the second step. Below is a description of the suggested methods.

Let $X(i,j)$ for $(i,j) \in$ A $\equiv$ {1, 2, 3,....., M} $\times$ {1, 2, 3,....., N} be the intensity of the pixel at location (i,j) in a corrupted $M \times N$ image X. Simultaneously, a binary flag image F of the same size has been generated, where $X(i,j)$ takes values for all pixels (i,j) $f_{i,j} = 1$.

A 5× 5 window (W) is created centered on a pixel in the image X in the image X.

From this window, we form a set $\{S\}$ that includes all the components of the window. The elements of this set are then arranged in increasing order, resulting in:

$S = \{s1, s2, s3, s4, \ldots\ldots, s25\}$ where $s1 \leq s2 \leq s3 \leq \ldots\ldots \leq s25$ and $i = 1,2,3,4, \ldots\ldots,25$

Four subsets S1, S2, S3, S4 are also formed like below.

$$S1 = \{ si : 0 \leq si \leq 64 \}$$
$$S2 = \{ si : 65 \leq si \leq 129 \}$$
$$S3 = \{ si : 130 \leq si \leq 194 \}$$
$$S4 = \{ si : 195 \leq si \leq 255 \}$$

To proceed with calculating $n(S1), n(S2), n(S3), n(S4)$ are calculated and examining the maximum $n(S_i)$ for $i = 1,2,3,4$ the following steps are followed.

$$if \ \ n(S1) is \ Max \ \ m = \frac{S_5 + S_6 + S_7}{3}$$

$$if \ \ n(S2) is \ Max \ \ m = \frac{S_8 + S_9 + S_{10}}{3}$$

$$if \ \ n(S3) is \ Max \ m = \frac{S_{11} + S_{12} + S_{13}}{3}$$

$$else \qquad p = \frac{S_{15} + S_{16} + S_{17}}{3}$$

Therefore, using all of W's components and relying on the following equation, the standard deviation (σ) is calculated.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \qquad [1]$$

$if \ \ |x_{i,j} - p| \leq \sigma \ \ then \ \ x_{i,j}$ $= \ undisturbed \ pixel \ else \ noisy \ and \ F \ is \ replaced \ by$ $f_{i,j} = 0$.

$if\ (\ f_{i,j}\ ==\ 0)\ then\ x_{i,j}\ =\ 0$ and apply SCMMF to remove noise (19).

Figure 3 highlights the advancements made by showing the results of this step-by-step method. By putting these preprocessing methods into practice, we want to improve the pictures' quality and dependability and prepare them better for analysis and diagnostic operations down the road. This also lessens the need for pricey dermatoscopy techniques.

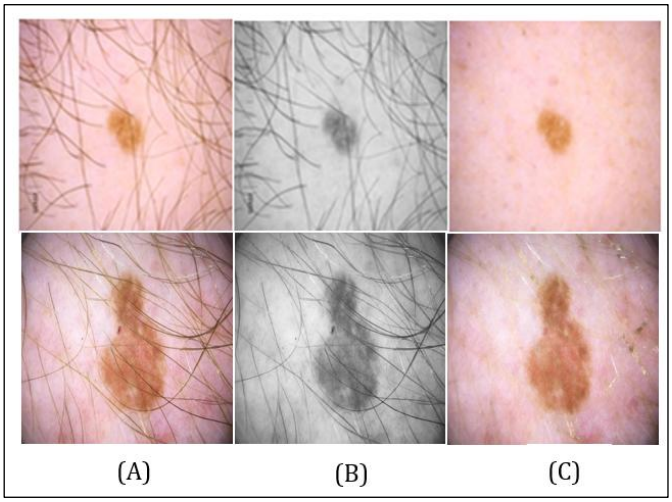

Figure 3: Skin Lesion Pre-Processing Method (A) Input Images ISIC_0087243 and ISIC_0097719 Respectively (B) Grayscale Images (C) Modified Picture

Using the PH2, ISBI 2017, And ISIC 2019 Datasets to Train Yolov8

Training algorithms with the right datasets of relevant photos has proven to be a challenging undertaking, and skin cancer detection has become a major study field in medical imaging. A total of 20,250 photos of non-melanoma and melanoma lesions from three publically available datasets—PH2, ISBI 2017, and ISIC 2019—were used to train the classifier using a holdout dataset. There were 2,530 pictures of non-melanoma and melanoma lesions in the testing data alone. There are 200 photos in the PH2 dataset, including 40 cases of melanoma, 80 atypical nevi, and 80 normal nevi. There are 2,000 training photos, 600 testing images, and 150 validation images in the ISBI 2017 dataset. There were 25,331 photos in the initial ISIC 2019 collection, of which 4,522 were melanoma and 20,809 were not. Table 1 provides details about the PH2 and ISBI 2017 datasets, while Table 2 outlines the characteristics of the ISIC 2019 dataset. We picked all 4,522 melanoma photos from the ISIC 2019 dataset and 12,778 non-melanoma images at random, for a total of 17,300 images, as we previously have 1,626 non-melanoma and 374 melanoma images from the ISBI 2017 dataset. We divided the small number of photos into three categories: 80% for training, 10% for testing, and 10% for validation. Thus, 13,840 training pictures, 1,730 testing images, and 1,730 validation images from the ISIC 2019 dataset were used to train the classifier. Table 3 shows the distribution of the chosen melanoma and non-melanoma photos for testing, validation, and training among the three datasets. The resolutions of these 24-bit RGB dermoscopic pictures range from 4499 × 6748 to 540 × 722.

Table 1: Distribution of PH2 and ISBI 2017 Datasets

Datasets	**Training Data**		**Validation Data**		**Test Data**		**Total**
Label	**M**	**NM**	**M**	**NM**	**M**	**NM**	
PH2	*	*	*	*	40	160	200
ISBI 2017	374	1626	30	120	117	483	2750

Note: M—melanoma, NM—non-melanoma, * there are no data in this field

Table 2: Distribution of ISIC 2019 Datasets

Label	MV	M	BK	BCC	SCC	VL	DF	AK	Total
ISIC 2019	12,875	4522	2624	3323	628	253	239	867	25,331

Note: MV—melanocytic nevus, M—melanoma, BK—benign keratosis, BCC—basal cell carcinoma, SCC—squamous cell carcinoma, VL—vascular lesion, DF—dermatofibroma, AK—actinic keratosis

Table 3: Used Datasets in the Proposed Work

Datasets	Training Data		Validation Data		Test Data		Total
Label	M	NM	M	NM	M	NM	
PH2	*	*	*	*	40	160	200
ISBI 2017	374	1626	30	120	117	483	2750
ISIC 2019	3622	10,218	450	1280	450	1280	17,300
Total	3996	11,844	480	1400	607	1923	20,250

Note: M—melanoma, NM—non-melanoma, * there are no data in this field

A modified CSPDarknet53 backbone serves as the convolutional neural network (CNN) for feature extraction in YOLOv8. Cross-stage partial connections are a feature of this design that allow for improved information flow across layers and increase accuracy. By merging feature maps from different backbone stages, the network's "neck," or feature extractor, enables the model to gather data at several sizes. YOLOv8 makes use of a unique C2f module in place of a conventional Feature Pyramid Network (FPN). This module improves identification accuracy, especially for tiny items, by integrating low-level spatial features with high-level semantic information. The ultimate forecasts come from the "head" of the network. For each grid cell in the feature map, YOLOv8 predicts bounding boxes, objectness scores, and class probabilities using a variety of detection modules (Figure 4). These individual forecasts are then integrated to yield the final object detections.

Using a convolutional neural network, the F-YOLOv8 approach analyses pictures and predicts bounding boxes and related class probabilities. In tensor representation, these forecasts have dimensions (S × S) * B * (5 + C). After splitting the picture into a grid of S × S non-overlapping cells, F-YOLOv8 processes each one to provide confidence ratings and identify bounding box attributes. Its confidence score tells us how likely it is that an item will be within the projected bounding box. The computation involves multiplying the probability of an object's presence (Pr(Object)) by the intersection over union (IOU) between the ground truth and the anticipated bounding box. The confidence score of a cell that is empty is 0. In the event that an object is present, the IOU—which is the ratio of the area that overlaps the actual and predicted bounding boxes to the entire area that both cover—represents the confidence score. It is possible to assess the spatial accuracy of the bounding box prediction by computing IOU, which is the result of calculating the area of intersection (numerator) and the total area of both boxes (denominator).

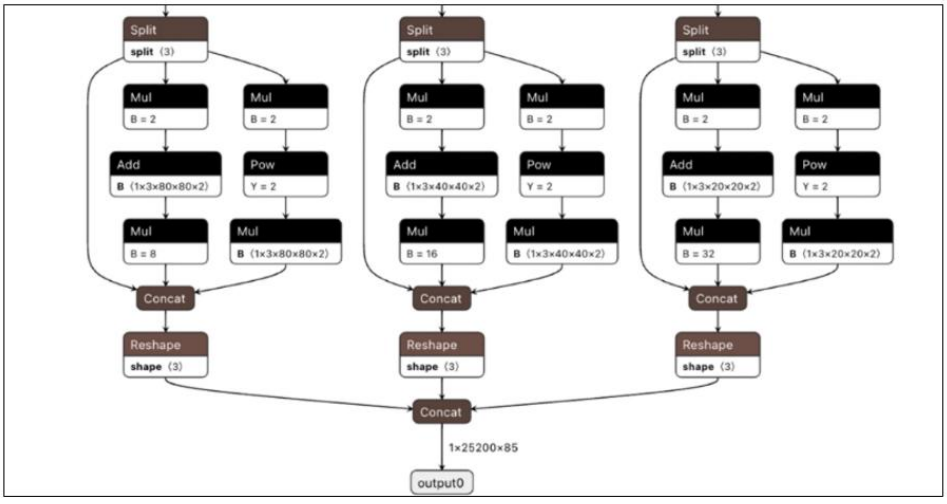

Figure 4: YOLOv8 Framework

Figure 5 shows the locations of skin lesions at three different magnification levels. With training from the ISIC 2018 dataset, we propose the F-YOLOv8 architecture (Figure 6) to identify skin lesions associated with melanoma. Resize all photos in the dataset to a consistent 512 × 512 resolution before using them to train the YOLOv8 model, even if their starting resolutions may vary. The following settings are used in the training procedure: 32 batch sizes, 0.9 momentum, 0.0005 decay, and 0.001 learning rate. The process is performed for 10,000 epochs.

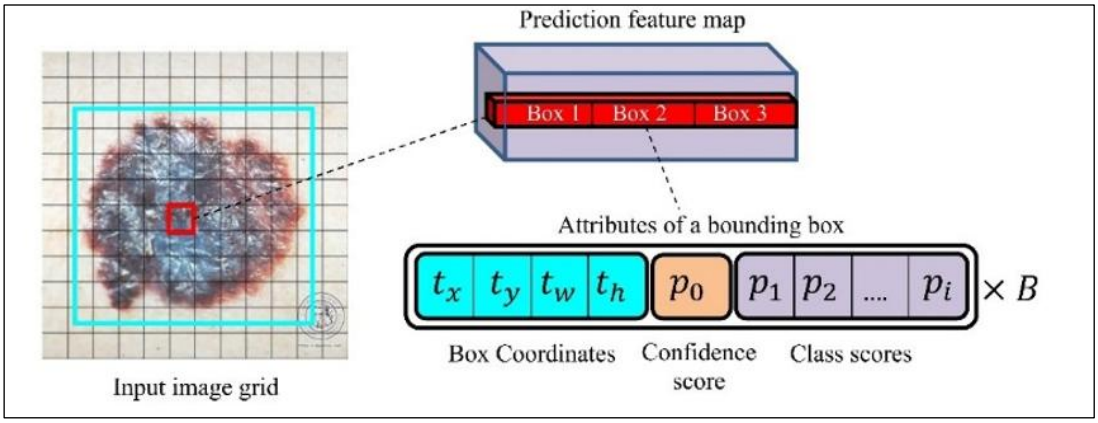

Figure 5: F-Yolov8 Reliably Predicted Melanoma. A Bounding Box Containing The Tx, Ty, Tw and Th Coordinates Established By The Red Grid Is The Blue Frame. Pi Is the Class of Probability Scores, and P0 is the Confidence Score

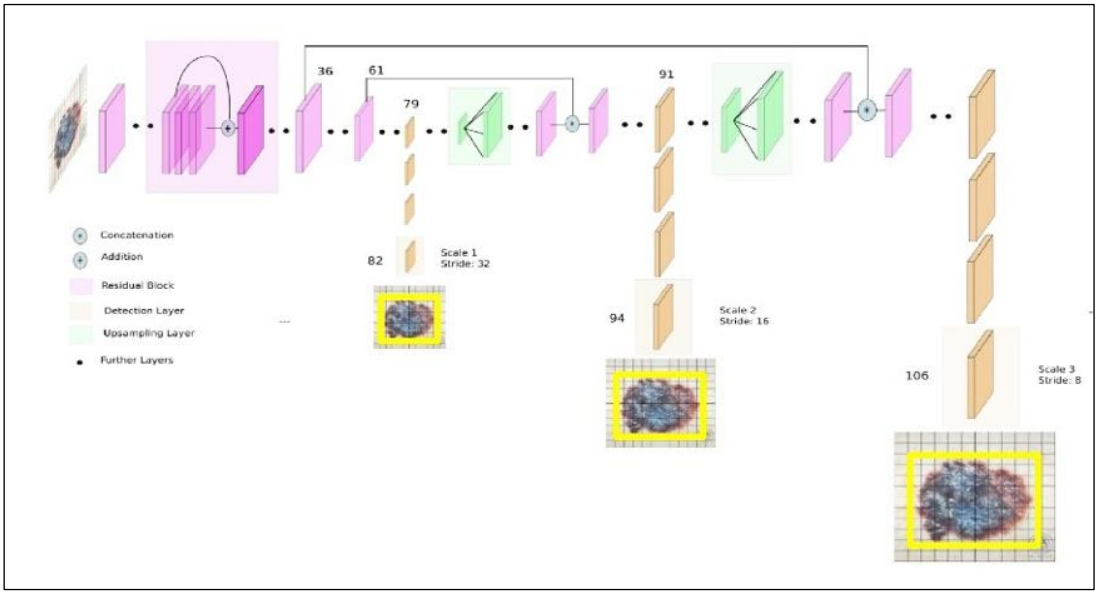

Figure 6: Detection of Melanoma Skin Lesions Using the F-Yolov8 Network

To optimize YOLOv8 for the particular features of our dataset and increase its performance in melanoma lesion identification, we fine-tuned the model's pre-trained weights. In addition, we changed the last layer of the YOLOv8 architecture to identify and categorise melanoma lesions in the output. These improvements were critical for customising YOLOv8 to the intricacies of our data and boosting its capacity to detect the target lesions. Figure 7 shows our tailored deep learning pipeline for the target domain. This process of modification and improvement guarantees that our approach successfully satisfies the particular needs of melanoma lesion identification in our target area.

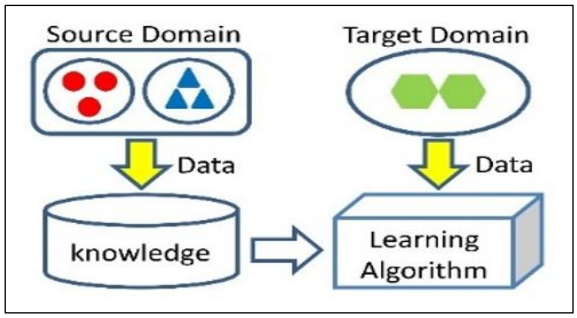

Figure 7: Shows the Transfer Learning Process Using the Target Domain from the Melanoma Dataset and the Source Domain from the Initial Weights

The example picture in Figure 8(A) shows that F-YOLOv8 has effectively recognised a melanoma lesion location. Consequently, the bounding box computation procedure is shown graphically in Figure 8(B). The portion of the bounding box where the lesion is completely contained is indicated in Figure 8(C). The last step involves getting all of the detected melanoma lesion pictures ready for F-SegNet stage input. This technique makes sure that the melanoma lesion pictures are effectively isolated and ready for the next processing stages. Using the datasets, F-YOLOv8 was trained using pictures reduced to 512 × 512 pixels and dispersed in accordance with Table 5's criteria. A total of 20,250 pictures and the related ground truth annotations were used in the training procedure. When tested on the datasets, the 53 convolutional layers and 55 million parameters that make up the F-YOLOv8 architecture produced the following outcomes. The parameter count is obtained by multiplying a filter's width and height by the number of filters in the previous layer, adding one, and then multiplying this total by the number of filters in the current layer. The assessment showed that the detection speed was 45 frames per second and the average accuracy was 99%.

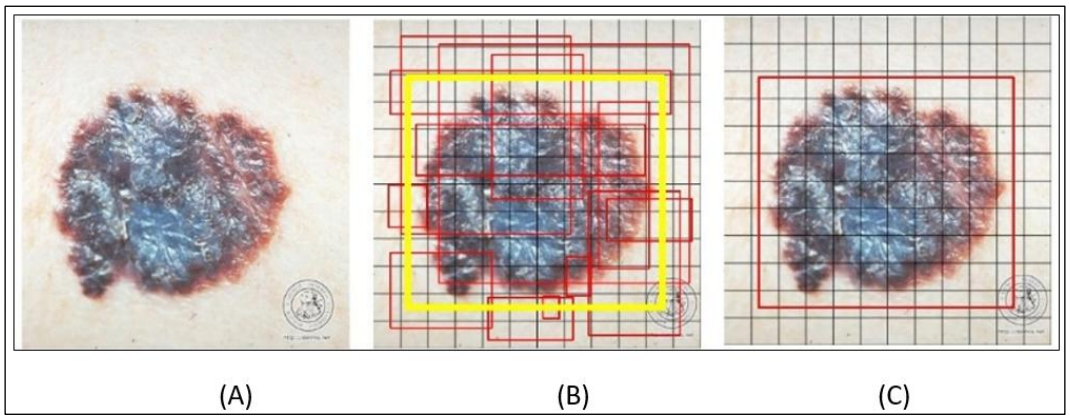

Figure 8: The Suggested F-Yolov8 is Making Progress in Extracting a Melanoma Lesion from a Clinical Picture. (A) Insert Pictures (B) Confidence and Boundary Boxes (C) Final Identification

F-Segnet for Melanoma Segmentation

For additional training, F-YOLOv8 processes images from the PH2, ISBI2017 and ISIC 2019 datasets before feeding them into F-SegNet. Figure 9 shows the configuration of the modified SegNet architecture that is used for melanoma identification in the training. Because of this modification, the SegNet model is able to efficiently complete segmentation tasks that are intended to locate melanoma lesions in the pictures. The last step in our procedure is to use the photos produced by the F-YOLOv8 process to train the F-SegNet algorithm. With regard to the SegNet model, we especially used pictures of tiny melanoma tumours to refine it. For this kind of learning and training, the symmetrical encoder-decoder structure of the SegNet architecture makes it extremely suitable, especially in situations involving complicated picture processing. The SegNet model's decoder network converts the low-resolution output into pixel-level predictions for the target class by using upsampling layers.

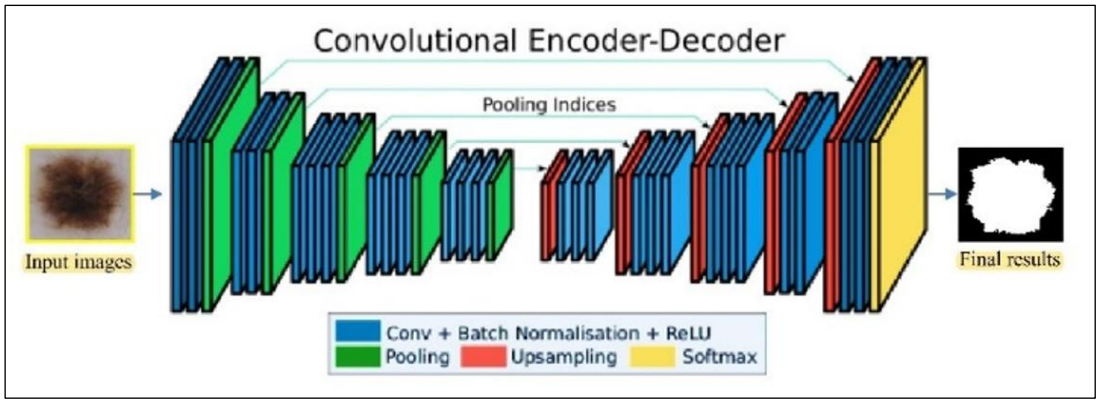

Figure 9: The Input Picture and Segmented Result Melanoma Type in F-Segnet Layers

Subsequent processing uses a low-resolution representation that the encoder network concurrently generates in addition to the initial high-resolution input. With its symmetrical encoder-decoder design, the SegNet model is especially well-suited to the problem of accurate melanoma lesion segmentation, which is the subject of this study. For optimal performance, the proposed classifier's hyperparameters include a learning rate of 0.045, a decay rate of 0.5, an epsilon value of 1.0, and a batch size of 32, with a subdivision of 16. The categorisation results are kept at the 50th epoch and are created throughout 5000 epochs. We experimented with adjustments to the SegNet model throughout the segmentation phase of our learning-based method. In particular, we presented a loss function that includes Median Frequency Balancing (MFB). Although the original training for F-SegNet (our modified SegNet) used cross-entropy loss, we included MFB in this loss function by using a particular formula, which we will discuss later.

$$K(\theta) = -\frac{1}{n}\left[\sum_{i=1}^{n} \emptyset_x(i)\left[x^i \log \hat{x}^i + (1 - x^i)\log(1 - \hat{x}^i)\right]\right] \quad [2]$$

This novel function boosts performance by weighting the average ratio of class frequencies by the melanoma class and putting it into the loss function, where $x(i)$ represents the real label, m indicates the total number of labeled pixels, and $\hat{x}^i$ symbolises the predicted output. Across the training dataset, this weighting approach is used, using class frequencies to differentiate across classes. Figure 10 shows our suggested model's chronological outputs. Crucially, our model performs very well on the unknown test dataset, proving its greater generalisation capacity and resilience. The observed discrepancy between the expected production before and after fine-tuning is a crucial finding.

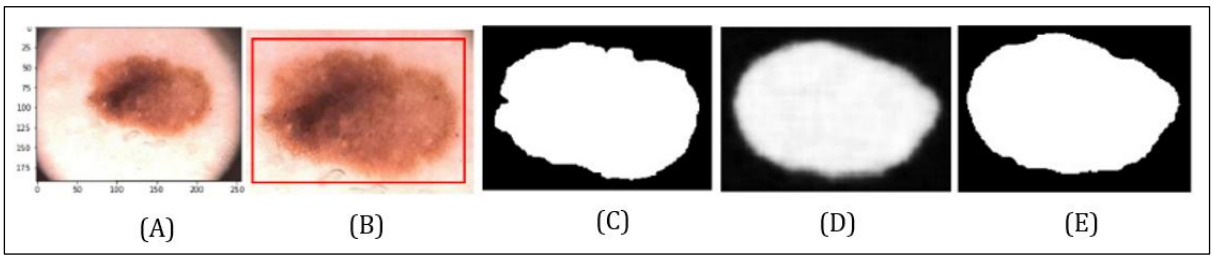

Figure 10: Chronological Progress of the Proposed Model (A) Original Image (B) Location Detection (C) Ground Truth Image (D) Predicted Segmentation Output (Before Fine-Tune The Model) (E) Final Output

Results and Discussion

This section provides a comprehensive analysis of the performance of the proposed approach. All testing were conducted on a computer with a Core-i7 CPU and 32 GB of RAM, and the recommended methodology was put into practice. The acquired images were analysed and categorised using the OpenCV framework and the Python programming language. Among the several criteria considered at different classification stages were the evaluation of the classification model, segmentation analysis, lesion localisation performance analysis, and skin refinement efficiency. To comprehend their possible ramifications in a clinical setting, the research would benefit from a thorough

examination of the many mistake kinds produced by the suggested model, such as false positives and false negatives. False positive results may cause patients needless worry and need further diagnostic testing, which would put more strain on healthcare systems. False negatives, on the other hand, can cause melanoma to go undetected, postponing necessary treatment and perhaps worsening patient outcomes. We used the PH2, ISBI 2017, ISIC 2018, and ISIC 2019 publically available datasets to assess the efficacy of the proposed methodology. Accurate localisation is necessary for the subsequent, more complex segmentation. The effectiveness of localisation is assessed using the Intersection over Union (IoU) criterion, a widely used computer vision metric. A lesion is considered accurately localised if it's IoU score is more than 80%.

$$IOU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad [3]$$

$$Sen = \frac{TP}{TP + FN} \quad [4]$$

$$Spe = \frac{TN}{TN + FP} \quad [5]$$

$$Dic = \frac{2 \times TP}{(2 \times TP) + FP + FN} \quad [6]$$

$$Jac = \frac{TP}{TP + FN + FP} \quad [7]$$

$$Acc = \frac{TP + TN}{TP + TN + FN + FP} \quad [8]$$

Additionally, the object localisation model is assessed using Mean Average Precision (mAP), which compares predicted bounding boxes to ground truth and gives higher scores to detections that are more trustworthy. These indicators are computed using exact formulae to assess the suggested method with rigour. Three metrics—sensitivity, specificity, and IOU (Intersection over Union) to precisely identify lesions—were used to assess the recognition performance. During the detection phase, the PH2 dataset produced an IOU of 95, a sensitivity of 97.5%, and a specificity of 98.5%. Using the ISBI 2017 dataset, the suggested method obtained an IOU of 92, a sensitivity of 98.47%, and a specificity of 97.51%. The results for the ISIC 2019 dataset were 90 for IOU, 97.77% for sensitivity, and 97.65% for specificity. Figure 11 shows the detailed output that the suggested model produced. The segmentation performance of our method was assessed on two datasets based on accuracy, sensitivity, specificity, Jac, and Dic metrics once the lesion site was determined. We used datasets from PH2, ISBI 2017, and ISIC 2019 to compare our work to many well-established techniques in medical imaging segmentation, which has seen significant progress. Our suggested segmentation approach performed better than several deep-learning and conventional segmentation frameworks. It outperformed the majority of approaches with the exception of one, achieving high sensitivity, specificity, Jac, and Dice scores for the PH2 dataset. With remarkable accuracy and Jac scores, it outperformed everyone on the ISBI 2017 dataset, including the best ISIC 2017 Challenge participants. Based on datasets from PH2 and ISBI 2017, respectively, Tables 4 and 5 compare the previously indicated segmentation techniques. Table 6 shows the segmentation performance results of our suggested approach on a subset of the ISIC 2019 dataset. Figure 12 displays the segmentation analysis together with the True Positive (TP) and True Negative (TN) values.

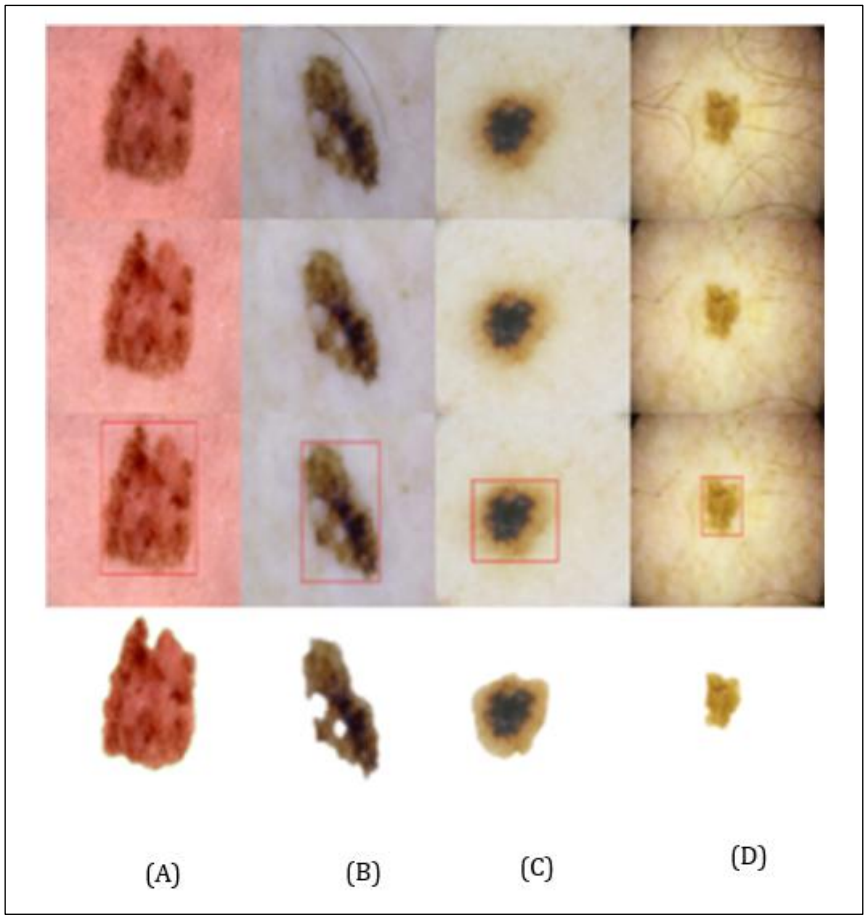

Figure 11: (A) ISIC_0014800, ISIC_0034813, ISIC_0061442, and ISIC_0599605 Input Pictures, (B) Pre-Processing Results, (C) Yolov8 Localisation, and (D) Segmentation Results

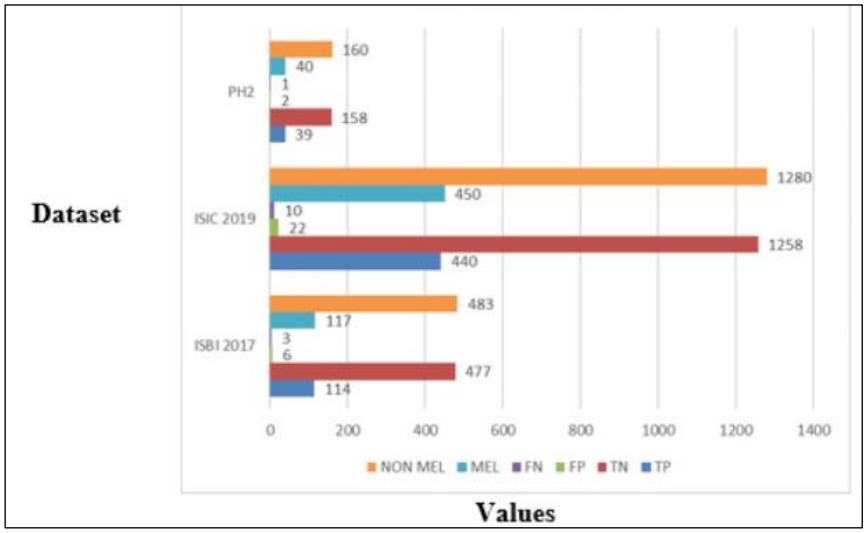

Figure 12: Segmentation Analysis Together With the True TP and True Negative TN Values

Table 4: PH2 Dataset Comparison of the Suggested Segmentation Performance (%)

References	Acc	Sen	Spe	Jac	Dic
(2)	97.50	97.50	97.50	88.64	93.97
(20)	98.70	92.90	96.90	-	-
(21)	96.50	96.70	94.60	89.40	94.20
(22)	92.99	83.63	94.02	79.54	88.13
(23)	94.24	94.89	93.98	83.99	90.66

(24)	95.03	96.23	94.52	85.90	92.10
Proposed Method	**98.50**	**97.50**	**98.75**	**92.86**	**96.30**

Table 5: ISBI 2017 Dataset Comparison of the Suggested Segmentation Performance (%)

References	Acc	Sen	Spe	Jac	Dic
(2)	97.33	91.45	98.76	86.99	93.04
(20)	95.30	87.50	85.50	-	-
(22)	93.39	90.82	92.68	74.81	84.26
(24)	95.06	86.05	95.95	79.15	88.95
(25)	81.57	75.67	80.62	-	-
(26)	93.60	81.60	98.30	78.20	87.80
(27)	93.20	82.00	97.80	76.20	84.70
Proposed Method	**98.50**	**97.44**	**98.76**	**92.68**	**96.20**

Table 6: ISIC 2019 Dataset Comparison of the Suggested Segmentation Performance (%)

References	Acc	Sen	Spe	Jac	Dic
(2)	97.33	91.45	98.76	86.99	93.04
(28)	97.86	97.56	97.97	92.23	95.96
Proposed Method	**98.15**	**97.78**	**98.28**	**93.22**	**96.49**

We evaluated the suggested recognition results with those of other classifiers, including Tree, SVM, KNN, and YOLOv8. We used You Only Look Once (YOLO) to compare our suggested approach with current deep-learning models by setting various parameters. Sensitivity, specificity, precision, accuracy, and AUC were the basis for the comparison, and time (in seconds) was also used as a statistic to confirm how quickly our strategy worked. These similarities are shown for photos from the PH2, ISBI 2017, and ISIC 2019 datasets in Tables 7, 8 and 9 respectively.

Table 7: Comparison of the Suggested Classifier with Popular Classifiers Using the PH2 Dataset

Classifier	Method	Acc (%)	Sen (%)	Spec (%)	Prec	AUC
TREE	CT	98.26	96.11	98.99	96.98	97.55
	ST	98.18	96.71	98.68	96.13	97.69
SVM	LSVM	97.95	96.71	98.38	95.28	97.54
	CSVM	98.41	96.11	99.19	97.57	97.65
	QSVM	98.18	96.41	98.78	96.41	97.60
	MGSVM	98.26	97.31	98.58	95.87	97.94
KNN	FKNN	98.64	96.41	99.39	98.17	97.90
	MKNN	98.41	96.11	99.19	97.57	97.65
	Cosine	97.88	96.41	98.38	95.27	97.39
	Cubic	98.33	95.81	99.19	97.56	97.50
	WKNN	98.26	97.01	98.68	96.14	97.84
YOLO	**Proposed Method**	**99.24**	**98.20**	**99.59**	**98.80**	**98.90**

Table 8: Comparison of the Suggested Classifier with Popular Classifiers Using the ISBI 2017 Dataset

Classifier	Method	Acc (%)	Sen (%)	Spec (%)	Prec	AUC
TREE	CT	93.83	92.39	94.54	89.22	93.46
	ST	94.83	93.91	95.29	90.69	94.60
SVM	LSVM	96.17	96.45	96.03	92.23	96.24
	CSVM	96.00	95.94	96.03	92.20	95.98
	QSVM	95.33	92.89	96.53	92.89	94.71
	MGSVM	95.00	91.37	96.77	93.26	94.07
KNN	FKNN	95.83	93.91	96.77	93.43	95.34
	MKNN	93.83	91.88	94.79	89.60	93.33
	Cosine	94.17	93.40	94.54	89.32	93.97
	Cubic	94.67	95.43	94.29	89.10	94.86

	WKNN	95.00	95.43	94.79	89.95	95.11
YOLO	**Proposed Method**	**99.17**	**99.49**	**99.01**	**98.00**	**99.25**

Table 9: Comparison of the Suggested Classifier with Popular Classifiers Using the ISIC 2019 Dataset

Classifier	Method	Acc (%)	Sen (%)	Spec (%)	Prec	AUC
TREE	CT	89.32	93.65	87.21	78.15	89.32
	ST	95.31	98.41	93.80	88.57	95.31
SVM	LSVM	91.67	96.03	89.53	81.76	91.67
	CSVM	88.28	93.65	85.66	76.13	88.28
	QSVM	92.71	96.03	91.09	84.03	92.71
	MGSVM	94.01	96.83	92.64	86.52	94.01
KNN	FKNN	92.71	96.83	90.70	83.56	92.71
	MKNN	90.10	93.65	88.37	79.73	90.10
	Cosine	94.01	96.03	93.02	87.05	94.01
	Cubic	94.01	94.44	93.80	88.15	94.01
	WKNN	88.80	96.03	85.27	76.10	88.80
YOLO	**Proposed Method**	**98.44**	**98.41**	**98.45**	**96.88**	**98.44**

Selecting YOLO as a classifier improves the effectiveness of skin lesion detection while cutting down on detection time. Preprocessing models, which automatically remove hair before applying picture enhancement and appropriate segmentation techniques, have helped to increase the accuracy of the suggested approach.

Conclusion

This research provides a unique F-YOLOv8 and F-SegNet model-based method for melanoma skin lesion identification and segmentation. To set our method apart from previous deep learning-based segmentation techniques, we use a two-step process: first, we localise the skin lesions and then we segment the areas that we have discovered. Although this approach requires more processing steps, it lessens the impact of the model including unnecessary picture regions. Tested on the PH2, ISIC 2017 and ISIC 2019 datasets, our approaches outperform state-of-the-art models in terms of detection and segmentation performance. It is possible to modify the suggested approach to handle other medical image analysis jobs, such as segmenting and detecting MRI images. In order to improve the model's robustness and clinical reliability, this study may also help guide ideas for further data incorporation or pre-processing method improvement. Future prospects for study include exploiting current improvements in the discipline to incorporate more advanced approaches within both phases.

Abbreviations

ISIC: International Skin Imaging Collaboration, ISBI: International Symposium on Biomedical Imaging, PH2: Pedro Hispano, YOLO: You Only Look Once, CAD: Computer-Aided Diagnosis, CNN: Convolutional Neural Network.

Acknowledgment

We wish to express our deep appreciation to Shubhendu Banerjee from the Department of CSE at Narula Institute of Technology, Kolkata, West Bengal, India, and Soumya Bhattacharyya and Shambhu Nath Saha from the Department of IT at Narula Institute of Technology, Kolkata, West Bengal, India, for their significant contributions to the study titled "Advanced Melanoma Detection Using Deep Transfer Learning."

Author Contributions

The work being executed under multiple authorship derived its conceptualization from the rigorous research carried out by Dr. Shubhendu Banerjee under the supervision of Dr. Shambhu Nath Saha. The much-needed software assistance, data curation, and availability of resources were ensured by Mr. Soumya Bhattacharyya. Based on the validation of data from all co-authors, Dr. Banerjee penned the original draft, which he further reviewed and edited for worthy visualization. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

The authors of this work state that they have no conflicts of interest about its publication.

Ethics Approval

Not applicable.

Funding

No specific grant from a public, private, or nonprofit organization was obtained for this study.

References

1. Busam KJ, Gerami P, Scolyer RA. Perspectives and Strategies to Minimize Harm From Melanoma Diagnosis. JAMA dermatology. 2023 Dec 1;159(12):1307-9.
2. Banerjee S, Singh SK, Chakraborty A, Das A, Bag R. Melanoma diagnosis using deep learning and fuzzy logic. Diagnostics. 2020 Aug 9;10(8):577.
3. Silveira M, Celebi ME, Mendes J, Marques JS, Marcal ARS, Mendonca T. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. IEEE J Sel Top Signal Process. 2009;3(1):35-45.
4. Masood A and Al-Jumaily AA. Fuzzy C mean thresholding based level set for automated segmentation of skin lesions. J Signal Inf Process. 2013;4(3):66.
5. Wong A, Scharcanski J, Fieguth P. Automatic skin lesion segmentation via iterative stochastic region merging. IEEE Trans Inf Technol Biomed. 2011;15(6):929-36.
6. Rajinikanth V, Satapathy SC, Fernandes SL, Dey N. Otsu's multi-thresholding and active contour snake model to segment dermoscopy images. J Med Imaging Health Inform. 2017;7(8):1837-40.
7. Alcón JF, Ciuhu C, Ten Kate W, Heinrich A, Uzunbajakava N, Krekels G, Siem D, de Haan G. Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis. IEEE journal of selected topics in signal processing. 2009 Feb 20;3(1):14-25.
8. Banerjee S, Singh SK, Chakraborty A, Basu S, Das A, Bag R. Diagnosis of Melanoma Lesion Using Neutrosophic and Deep Learning. Traitement du Signal. 2021 Oct 1;38(5):1327-1338.
9. Tong N, Lu H, Ruan X, Yang MH. Salient object detection via bootstrap learning. InProceedings of the IEEE conference on computer vision and pattern recognition. 2015:1884-1892. http://openaccess.thecvf.com/content_cvpr_2015/html/Tong_Salient_Object_Detection_2015_CVPR_paper.html
10. Yu L, Chen H, Dou Q, Qin J, Heng PA. Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging. 2016;36(4):994-1004.
11. Singh SK, Banerjee S, Chakraborty A, Bandyopadhyay A. Classification of Melanoma Skin Cancer Using Inception-ResNet. InFrontiers of ICT in Healthcare: Proceedings of EAIT 2022. 2023:65-74.http://dx.doi.org/10.1007/978-981-19-5191-6_6
12. Yuan Y, Chao M, Lo YC. Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. IEEE Trans Med Imaging. 2017;36(9):1876-86.
13. Al-Masni MA, Al-Antari MA, Choi MT, Han SM, Kim TS. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. Computer methods and programs in biomedicine. 2018 Aug 1;162:221-31.
14. Banerjee S, Singh SK, Das A, Bag R. Diagnoses of Melanoma lesion using YOLOv3. InComputational Advancement in Communication, Circuits and Systems: Proceedings of 3rd ICCACCS 2020. 2022 :291-302. http://dx.doi.org/10.1007/978-981-16-4035-3_26
15. Banerjee S, Bandyopadhyay A, Mukherjee A, Das A, Bag R. Random Valued Impulse Noise Removal Using Region Based Detection Approach. Engineering, Technology & Applied Science Research. 2017 Nov 1; 7(6):2288-2292.
16. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL. Microsoft coco: Common objects in context. InComputer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland. 2014 September 6-12. Proceedings, Part V 13. 2014:740-755. https://link.springer.com/chapter/10.1007/978-3-319-10602-1_48
17. Shapiro L, Basra M, Patel H, Payne C, Brazen B, Biglione A. Utilization of Imaging Modalities in the Diagnosis of Melanoma: A Scoping Review. Cureus. 2024 Feb;16(2). http://dx.doi.org/10.7759/cureus.54058
18. Redmon J and Farhadi A. YOLO9000: better, faster, stronger. InProceedings of the IEEE conference on computer vision and pattern recognition 2017: 7263-7271. http://dx.doi.org/10.1109/CVPR.2017.690
19. Banerjee S, Bandyopadhyay A, Bag R, Das A. Sequentially combined mean-median filter for high density salt and pepper noise removal. In 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). 2015 Nov 20:21-26. http://dx.doi.org/10.1109/ICRCICN.2015.7434203
20. Hasan MK, Reza A, Ahmed F, Sulaiman R, Mokji M, Iqbal SM. DSNet: Automatic dermoscopic skin lesion segmentation. Comput Biol Med. 2020;120:103738.
21. Xie Y, Zhang J, Xia Y, Shen C, Yuille A. A mutual bootstrapping model for automated skin lesion segmentation and classification. IEEE Trans Med Imaging. 2020;39(7):2482-93.
22. Ünver HM and Ayan E. Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. Diagnostics (Basel). 2019;9(3):72.
23. Bi L, Kim J, Ahn E, Feng D, Fulham M. Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks. arXiv preprint arXiv:1703.04197. 2017. http://dx.doi.org/10.48550/arXiv.1703.04197
24. Bi L, Kim J, Ahn E, Feng D, Fulham M. Step-wise integration of deep class-specific learning for dermoscopic image segmentation. Pattern Recognit. 2019;85:78-89.
25. Al-Masni MA, Kim DH, Kim TS. Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Comput Methods Programs Biomed. 2020;190:105351.

26. Sarker MM, Rashwan HA, Akram F, Banu SF, Saleh A, Singh VK, Chowdhury FU, Abdulwahab S, Romani S, Radeva P, Puig D. SLSDeep: Skin lesion segmentation based on dilated residual and pyramid pooling networks. InMedical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain.2018 September 16-20. Proceedings. Part II .2018:21-29. https://link.springer.com/chapter/10.1007/978-3-030-00934-2_3
27. Li Y and Shen L. Skin lesion analysis towards melanoma detection using deep learning network. Sensors (Basel). 2018;18(2):556.
28. Singh SK, Abolghasemi V, Anisi MH. Skin cancer diagnosis based on neutrosophic features with a deep neural network. Sensors (Basel). 2022;22(16):6261.