A novel artificial intelligence-based predictive analytics technique to detect skin cancer

Prasanalakshmi Balaji [1], **Bui Thanh Hung** [1], **Prasun Chakrabarti** [2], **Tulika Chakrabarti** [3], **Ahmed Elngar** [4], **Rajanikanth Aluvalu** Corresp. [5]

[1] Artificial Intelligence Laboratory, Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[2] ITM (SLS) Baroda University, Vadodara, Gujarat, India

[3] Sir Padamapat Singhania University, Udaipur, Rajasthan, India

[4] Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef, Egypt

[5] Department of IT, Chaitanya Bharathi Institute of Technology, Hyderabad, India

Corresponding Author: Rajanikanth Aluvalu
Email address: rajanikanth.aluvalu@gmail.com

One of the leading causes of death among people around the world is skin cancer. It is critical to identify and classify skin cancer early to assist patients in taking the right course of action. Additionally, melanoma, one of the main skin cancer illnesses, is curable when detected and treated at an early stage. More than 75% of fatalities worldwide are related to skin cancer. A novel Artificial Golden Eagle-based Random Forest (AGEbRF) is created in this study to predict skin cancer cells at an early stage. Dermoscopic images are used in this instance as the dataset for the system's training. Additionally, the dermoscopic image information is processed using the established AGEbRF function to identify and segment the skin cancer-affected area. Additionally, this approach is simulated using a Python program, and the current research's parameters are assessed against those of earlier studies. The results demonstrate that, compared to other models, the new research model produces better accuracy for predicting skin cancer by segmentation.

1 # A Novel Artificial Intelligence-Based Predictive
2 # Analytics Technique to Detect Skin Cancer

4 **Prasanalakshmi Balaji[1], Bui Thanh Hung[1], Prasun Chakrabarti[2], Tulika**
5 **Chakrabarti[3], Ahmed Elngar[4] , Rajanikanth Aluvalu[5]**

7 [1]Artificial Intelligence Laboratory, Faculty of Information Technology, Ton Duc Thang
8 University, Vietnam.
9 [2]ITM SLS Baroda University, Vadodara Gujarat, India.
10 [3] Sir Padampat Singhania University, Udaipur, Rajasthan, India
11 [4]Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef City, Egypt.
12 [5]Department of IT, Chaitanya Bharathi Institute of Technology, Hyderabad, India
13 Corresponding Author: Rajanikanth Aluvalu[5]
14 Email address: rajanikanth.aluvalu@gmail.com

23 # A Novel Artificial Intelligence-Based Predictive
24 # Analytics Technique to Detect Skin Cancer

26 **Prasanalakshmi Balaji[1], Bui Thanh Hung[1], Prasun Chakrabarti[2], Tulika**
27 **Chakrabarti[3], Ahmed Elngar[4] , Rajanikanth Aluvalu[5]**

29 [1]Artificial Intelligence Laboratory, Faculty of Information Technology, Ton Duc Thang
30 University, Vietnam.
31 [2]ITM SLS Baroda University, Vadodara Gujarat, India.
32 [3] Sir Padampat Singhania University, Udaipur, Rajasthan, India
33 [4]Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef City, Egypt.
34 [5]Department of IT, Chaitanya Bharathi Institute of Technology, Hyderabad, India
35 Corresponding Author: Rajanikanth Aluvalu[5]
36 Email address: rajanikanth.aluvalu@gmail.com

38 Abstract

39 One of the leading causes of death among people around the world is skin cancer.
40 It is critical to identify and classify skin cancer early to assist patients in taking the
41 right course of action. Additionally, melanoma, one of the main skin cancer
42 illnesses, is curable when detected and treated at an early stage. More than 75% of
43 fatalities worldwide are related to skin cancer. A novel Artificial Golden Eagle-
44 based Random Forest (AGEbRF) is created in this study to predict skin cancer
45 cells at an early stage. Dermoscopic images are used in this instance as the dataset
46 for the system's training. Additionally, the dermoscopic image information is
47 processed using the established AGEbRF function to identify and segment the skin
48 cancer-affected area. Additionally, this approach is simulated using a Python
49 program, and the current research's parameters are assessed against those of earlier
50 studies. The results demonstrate that, compared to other models, the new research
51 model produces better accuracy for predicting skin cancer by segmentation.

52

53 Introduction

54 Skin cancer is brought on by malignant tumors, which develop when the proliferation of
55 abnormal skin cells is left unchecked [1]. The majority of cancer cases are caused by unprotected
56 skin being exposed to UV radiation [2]. Nowadays, 95 percent of skin malignancies diagnosed
57 are squamous or carcinoma and 5 percent are melanomas [3]. Due to numerous factors like
58 infections, allergies, and exposure to the sun, skin alterations happen making it the main cause of
59 skin cancer. Skin cancer is one of the cancers that have diagonised and led to death of millions of
60 people in recent years [4]. Melanoma and eczematoid are two examples of skin malignancies.
61 Typically, skin tumors that develop on the epidermis are a common condition [45]. Apart from

62 ~~the fact that most skin tumors in the globe are benign and hardly ever develop into skin cancer,~~
63 ~~there are also dangerous malignant skin tumors that, if left untreated, can kill the host by~~
64 ~~spreading throughout the body [5, 6]. Melanoma and eczematoid are two examples of skin~~
65 ~~malignancies. Skin cancer is one of the cancers that have killed millions of people in recent years~~
66 ~~[7].~~ Melanoma and eczematoid are two examples of skin malignancies. Typically, skin tumors
67 that develop on the epidermis are a common condition [5]. Apart from the fact that most skin
68 tumors in the globe are benign and hardly ever develop into skin cancer, there are also dangerous
69 malignant skin tumors that, if left untreated, can kill the host by spreading throughout the body
70 [6,7].To lower the fatality rate, early detection and timely treatment are hence more crucial [8].
71 More than 125,000 new cases of melanoma are diagnosed each year, while there are 3 million
72 non- melanoma cases.[9].
73 More often, both melanoma and non-melanoma skin tumors ~~kill~~ may lead to death of people [10].
74 As a result, squamous cell carcinoma and basal cell carcinoma attack non-melanoma skin [11,
75 12]. Melanoma causes aberrant melanocyte proliferation, and the cells that produce pigment also
76 give skin its color. A larger probability of curing melanoma exists with early identification.
77 However, it expanded to other body areas and had permanent effects [13]. Overfitting, low
78 detection accuracy and error rate are further issues with a skin cancer diagnosis. [14]. Large-scale
79 data training for disease prediction is the main advantage of utilizing a deep learning model in
80 medical picture analysis [15]. Fig.1 defines the steps of investigation and detection of skin
81 lesions.
82   Machine learning methods are chosen in the proposed approach since~~Additionally~~, compared
83 to the Deep Learning (DL) model, the Machine Learning (ML) technique achieves greater
84 disease prediction accuracy due to the hyper parameter tuning [16]. Additionally, by preventing
85 data overfitting during training, labeled data can lower training errors [17, 18]. Additionally,
86 various rule-based techniques make it simple to solve the machine learning model [19]. Since
87 several decision tree layers are used to process the ML model's function. The memory and
88 control mechanism in the contemporary ML allows it to carry out the procedure better [20]. To
89 segment the damaged section and control the degree of disease by early-stage prediction, filtering
90 technique [22], automatic and optimal pipeline approach [24], etc. The suggested approach is
91 capable of identifying and classifying the area affected by skin cancer. The main aim of the
92 designed model is to enhance the prediction results of skin cancer by accurate segmentation and
93 classification.
94 Below are a few current literature reviews that focus on skin cancer prediction. A rectification
95 and uncovering of ML for the diagnosis of skin cancer were developed by Meike Nauta et al.
96 [21]. Dermoscopic images are initially trained using a typical VGG16 classifier on the public
97 ISIC dataset, which consists of colored patches and elliptical charts. As a result, the created
98 technique automatically eliminated the patches that were present in the trained images and
99 measured changes as they occurred during prediction; however, prediction accuracy was low
100 when compared to other strategies.

101 It is more difficult to detect similarities and weak contrast between lesions and skin. Mahbubur et
102 al. [22] suggested a method for filtering the dataset's noise using dermoscopic images like
103 anisotropic diffusion. Additionally, the affected region of the skin cancer is segmented using the
104 quick bounding box technique. Additionally, the created technique performs better; however,
105 segmenting the skin cancer-affected region took longer.
106 An innovative DL method for identifying the type of cancer depending on its phases was put
107 forth by Maad M. Mijwil [23]. Three architectures were employed to determine the precise
108 classification of images. Additionally, the evaluated high-resolution images achieved sensitivity
109 and specificity rates of 86.14 percent and 87.66 percent, respectively, although the error rate is
110 significant due to vast data. Early identification and detection of skin cancer result in a high level
111 of effort in disease management.
112 Using dermoscopic images, Liu Wei et al. [24] suggested an automatic and ideal pipeline
113 technique for the diagnosis of skin cancer. It uses the threshold technique to characterize the
114 Region of Interest (RoI) and noise reduction. The issue of offloading arises when comparing
115 existing ways to test reliability.
116 A unique ML framework for the diagnosis of skin cancer was created by Murugan et al. [25]. In
117 addition, skin cancer is the most prevalent disease in people and is determined by skin tone.
118 First, apply a median filter to the skin image, then use a mean shift to segment the affected area.
119 After that, features are retrieved and then classified using a Support Vector Machine (SVM).
120 Additionally, the segmentation of the lesion area can be done with high precision, although the
121 false detection rate is significant.
122   The key roles of this proposed model are summarized as;
123 • Initially, the dermoscopic image dataset was collected and trained in the system using the
124 python tool.
125 • Subsequently, a novel AGEbRF Model was developed to predict skin cancer by segmenting
126 the affected region.
127 • Hereafter, the fitness of the golden eagle is updated in the Random Forest (RF) that is used to
128 segment the affected part images.
129 • Hence, the developed AGEbRF segments the affected part of skin cancer with high
130 performance.
131 • Finally, the segmented images are utilized to track the dataset for predicting the possibility of
132 skin cancer.
133 • Moreover, the developed approach is validated using recent prevailing approaches in terms of
134 detection accuracy, specificity, sensitivity, F-measure, and precision.
135

Materials & Methods

137 There are 200 different forms of cancer around the globe, and they all disrupt tissue, disrupt
138 organ development, and penetrate cells. Tumors are frequently produced cells and spread to
139 other tissues. Additionally, melanoma has pigment cells that result in a broad, dark lesion. There
140 are five phases in skin cancer detection, including preprocessing, lesion identification, feature
141 extraction, and classification.

142   Fig-ure 2 depicts the fundamental model of the system and research problem. First, capture the
143   dermoscopic images and then do preprocessing to reduce noise and enhance image quality. The
144   original image is then masked using the threshold value, and the border, color, asymmetry, and
145   diameter of each feature are processed. The classification layer, which is the last to detect skin
146   cancer, has a few drawbacks, including high complexity, misclassification, poor resolution, and
147   lower detection accuracy. However, as skin cancer segmentation is handled manually, there is a
148   high possibility of incorrect classification and errors as well as erroneous predictions. This
149   creates a revolutionary machine learning approach to identify and divide the skin cancer-affected
150   area.
151   Identification of the damaged area is more important for the treatment of skin cancer because it is
152   impossible to finalize treatment measures without the knowledge of the affected area. The
153   current study has proposed an Artificial Golden Eagle-based Random Forest (AGEbRF) for the
154   exact prediction of the damaged skin areas to overcome this issue. This method updates the
155   golden eagle fitness in the classification layer to improve segmentation effectiveness.
156   Consequently, it uses dermoscopic images to segment and predicts the area affected by skin
157   cancer. Finally, the results of the intended strategy are contrasted with those of other strategies.
158   Fig-ure 3 displays the proposed methodology.

159   **A. DATA ACQUISITION**
160
161   Multiple dermoscopic images are gathered since dermoscopic skin scans are necessary for
162   identifying illnesses. For training and testing purposes, a dermoscopic image dataset obtained
163   from the internet is used in this study. The skin cancer dataset [26,27] was utilized, which also
164   has more dermoscopic images of patients. The dataset used in the study includes 10015
165   dermatoscopic images for training from Task 3 of the dataset and 1512 images from the same
166   task for testing purpose contributing to a total of 11527 images. As a result, both benign and
167   malignant skin lesions are mixed in the dataset. It consisted of around 53.98% of male patients,
168   45.45% of female patients and unknown category 0.57%. The images are also used to teach the
169   system as well as a Table.1 elaborates on the testing and training dataset. The dataset was found
170   to be biased in terms of the types of lesion, eventhen since the article concentrates on just the
171   prediction process, we donot concentrate on the bias since the bias in category of lesion doesnot
172   affect the lesion prediction results.
173   ~~As a result, both benign and malignant skin lesions are mixed in the dataset. The images are also~~
174   ~~used to teach the system as well as a table.1 elaborates on the testing and training dataset.~~

175   **B. PROCESS OF AGEbRF TO SEGMENT THE AFFECTED PART**
176
177   In this approach, the AGEbRF mechanism is developed for segmenting and predicting skin
178   cancer. The developed GEbRF is the combined form of the Golden Eagle Optimization (GEO)
179   algorithm [30] and Random Forest (RF) [31].
180   The behavior of the GEO function is initiated in the RF detection layer which can enhance
181   disease prediction. In this update, the fitness of the golden eagle enhances the tracking results of

182   the affected parts of skin cancer. It searches the stack by spiral trajectory and indicates the
183   propensity to cruise in the primary stage, which identifies the attack and provides prey
184   information to the other eagles it enhances the prediction results. In the current study, the GEO
185   function is used to pinpoint the exact location of the disease in the skin. Initially, the system is
186   trained on the dermoscopic image dataset. The Random forest has been accomplished by
187   numerous sub-trees that contain preprocessing, feature extraction, segmentation, and prediction.

188   1) PRE-PROCESSING AND FEATURE EXTRACTION
189

190   The dermoscopic image dataset is fed into the input layer, and the noise and errors in the dataset
191   are removed in the following layer. Furthermore, the input layer performs pre-processing and
192   feature extraction processes on dermoscopic skin images. Then, using a non-linear operation,
193   features are extracted from dermoscopic images based on shape, color, size, and texture. In Eqn.
194   (4) the extraction process non-linear operations are mentioned,

195   $S_i d = f\left(L(n) * S_i s\right)$ (4)

196   Let, $S_i s$ is denoted as the input image, $S_i d$ is represented as the extracted output, and non-linear

197   operation is represented as $L(n)$. The process of AGEbRF is represented in Ffig. 4.

198   2) SEGMENTATION AND PREDICTION
199   The extracted output is passed to the classification layer, which performs the classification
200   process. The golden eagle fitness function is used to improve classification in this layer. It
201   reduced the spatial resolution to achieve spatial invariance to falsification and input translations.
202   Furthermore, this layer classifies patients with skin cancer and determines whether it is benign or
203   malignant. After that, the classified output is sent to the prediction layer, which can segment the
204   affected part. As a result, the segmentation and classification processes are carried out using
205   AGEbRF, which is represented in Eqn (5),
206

207   $$S_i d^{c_l} = f\left[\sum_{k \in A_l} S_i d^{c_l - 1} * KN_{k_l}^{c_l} Srb_i + A_l^{c_l}\right] dr_{iter+1}$$ (5)

208   Let, c_l, represent classification layer, $c_l - 1$ is denoted as the sub-tree segmentation, $S_i d^{c_l - 1}$ is

209   considered as the input characteristics of this layer. Moreover, kernel maps and additive bias are

210   denoted as $KN_{k_l}^{c_l}$ and $A_l^{c_l}$. Furthermore, dr_{iter} is considered as the iteration value, Srb_i which is

211   represented by the fitness of the golden eagle. Additionally, A_l is denoted as selection input

212   maps. k, and l are represented as input and output. Finally, the output layer presents the

213   segmented output and uses a decision tree built from numerous samples to predict the skin illness
214   as benign or malignant. Here, the affected part is divided into categories using the AGEbRF
215   method . Hence, the dermoscopic image dataset was segmented, and predict the damaged portion
216   of skin cancer using the created AGEbRF method. A flowchart of the created AGEbRF is also
217   provided in fFig. 5.
218   Additionally, the dataset for determining the probability of skin cancer, whether it is benign or
219   malignant, is predicted using the segmented images.

A. CASE STUDY

220
221 Skin cancer is typically regarded as the most dangerous type of cancer. Periodically, early
222 detection is required, and a cancer diagnosis must be precise and effective if people are to be
223 saved. The suggested AGEbRF model is first applied to the dataset once the dermoscopic dataset
224 has been trained to the system. Additionally, the pre-processing and feature extraction functions
225 are used in the dataset to eliminate unintentional image flaws. The feature extraction technique,
226 which uses the suggested model, extracts the pertinent features from the filtered input images.
227 These attributes include shape, size, color, and texture. Figure 6 provides specifics on the sample
228 dermoscopic images and the expected outcomes.
229 The classification layer of the suggested model is then used to segment the impacted portion. The
230 suggested AGEbRF technique has segmented the affected area of skin cancer. The generated
231 model was then used to assess the possibility of skin cancer.
232 This uses the segmented output as the test to predict potential skin cancer cases in a given
233 dataset. Finally, the probability of skin cancer is determined using the segmented images. Here,
234 the dataset images are regarded as the testing images while the segmented images are used as
235 input images. In these images, the created AGEbRF model is used to assess the possibility of the
236 disease, whether benign or malignant.
237 The probability of skin cancer is 50 percent benign and 50 percent malignant when utilizing the
238 proposed AGEbRF model. Here, the input images are taken into consideration as 1000 numbers,
239 of which 500 are malignant and 500 are benign. As a result, the suggested model pinpoints the
240 skin cancer's affected area and Fig. 7 detailed the comparison of accuracy and loss of established
241 framework.
242 The designed AGEbRF technique is implemented using a Python program, and several metrics
243 like accuracy, sensitivity, specificity, F1-measure, and precision are measured. Additionally, the
244 developed approach is verified using established techniques such as Hybrid Feature Fusion and
245 ML (HFF) [22], Skin Cancer Classification (SCC) [23], Modified Thermal Exchange (MTE)
246 [24], and Diagnosis of skin cancer by ML (DSC) [25], Region extraction and Classification of
247 Skin Cancer (RCSC) [28] and Ensemble Lightweight Technique (ELT) [29].

248

249 1) ACCURACY MEASUREMENT
250 The degree of measurement of the proposed model's working efficiency is known as accuracy.
251 Additionally, Eqn. (6) expresses the proportion of correctly expected observance to all
252 observations.

253
$$A_c = \frac{t_p + t_n}{t_p + t_n + f_p + f_n}$$
(6)

254 Where t_p is denoted as the true positive of accurate prediction and precise segmentation, t_n is
255 represented as the true negative of accurate prediction and incorrect segmentation, f_p is denoted
256 as the false positive of inaccurate prediction and exact segmentation, and f_n is considered as the
257 false negative of inaccurate prediction and incorrect segmentation of affected skin part.

258 Using established techniques including HFF, SCC, MTE, DSC, and RCSC approaches that
259 resulted in values as listed in Table 2, the correctness of the proposed AGEbRF model is
260 estimated and validated.
261 The new strategy has a higher accuracy value of 99.96% than other ways, but the existing
262 approaches have achieved lesser accuracy, demonstrating the usefulness of the proposed
263 approach. Figure 8 displays a comparison of accuracy measurement results.
264

265 2) SENSITIVITY MEASUREMENT
266 The amount of accurately anticipated true positives is determined using sensitivity. Additionally,
267 Eqn. (7) measures the probability of predicting the skin's affected area.

$$S_{en} = \frac{t_p}{t_p + t_n} \tag{7}$$

269 The sensitivity of the developed AGEbRF technique is calculated and validated using popular
270 methods such as the HFF, MTE, SCC, ELT, DSC, and RCSC approaches described in the Table.
271 3.
272 Existing approaches have only achieved a sensitivity of nearly 95 percent. Furthermore, this
273 AGEbRF method achieved 99.34 percent higher sensitivity than other methods, demonstrating
274 the effectiveness of the developed model. Figure 9 depicts the sensitivity comparison
275 graphically.
276

277 3) SPECIFICITY MEASUREMENT
278
279 Specificity is defined as the degree used to determine the number of true negatives that are
280 correctly recognized. Additionally, specificity is used to calculate the efficiency of predicting
281 skin cancer using segmented images, which is calculated using Eqn. (8),

$$S_{pe} = \frac{t_n}{t_p + t_n} \tag{8}$$

283 The specificity of the developed AGEbRF technique is calculated and validated using commonly
284 used methods such as HFF, SCC, MTE, ELT, DSC, and RCSC approaches, which yielded values
285 as shown in Table 4.
286 The existing approaches achieved a lower specificity of nearly 98 percent. As a result, the
287 developed approach achieved 99.55 percent higher specificity than other methods, demonstrating
288 the effectiveness of the developed model. In addition, Fig. 10 depicts the comparison of
289 specificity.

290 4) PRECISION MEASUREMENT
291 This procedure has been tested for determining the number of correct positive estimates alienated
292 by the total positive estimates. Furthermore, precision is the proportion of precise cancer-affected
293 region diagnosis computed using Eqn. (9),

294
$$P_{re} = \frac{t_p}{t_p + f_p}$$
(9)

295 Using established techniques including HFF, SCC, MTE, ELT, and RCSC approaches, the
296 suggested AGEbRF model's precision is estimated and validated, and the results are listed in
297 Table 5. Here, the current methods have only managed a lower precision of roughly 96 percent.
298 Additionally, the created AGEbRF methodology obtained a 99.29% higher precision value than
299 other methods and the comparison of precision is shown in Fig. 11.

300 5) F1-SCORE MEASUREMENT
301 The computation is based on measurements of sensitivity and precision used to determine how
302 effectively the images are traced using an equation. (10),

303 $$F1-score = 2\left(\frac{P_{re} * S_{en}}{P_{re} + S_{en}}\right)$$
(10)

304 Where P_{re} denotes the calculated precision value and S_{en} represents the calculated sensitivity
305 value.
306 The suggested AGEbRF model's F1-score is determined and validated using widely used
307 techniques like as Hybrid Feature Fusion and ML (HFF) [22], Skin Cancer Classification (SCC)
308 [23], Modified Thermal Exchange (MTE) [24], and Diagnosis of skin cancer by ML (DSC) [25],
309 Region extraction and Classification of Skin Cancer (RCSC) [28] approaches. The results are
310 listed in the Table. 6.
311 Here, the suggested model has outperformed previous methods with a high F1-score value of
312 99.12% compared to the existing approaches, as shown in Fig. 12.
313

Results and Discussion

315 Python is used to process the created AGEbRF replica, and the predicted model's success rate is
316 evaluated using the mechanisms that are now in place in terms of accuracy, sensitivity, F-
317 measure, precision, and specificity. Dermoscopic images are used in this method to find skin
318 cancer. The suggested AGEbRF technique identifies skin cancer; in this case, using dermoscopic
319 images divides the affected skin cancer portions. As a result, the generated model performed well
320 in segmentation and prediction.

321 A. CASE STUDY
322 Skin cancer is typically regarded as the most dangerous type of cancer. Periodically, early
323 detection is required, and a cancer diagnosis must be precise and effective if people are to be
324 saved. The suggested AGEbRF model is first applied to the dataset once the dermoscopic dataset
325 has been trained to the system. Additionally, the pre-processing and feature extraction functions
326 are used in the dataset to eliminate unintentional image flaws. The feature extraction technique,
327 which uses the suggested model, extracts the pertinent features from the filtered input images.
328 These attributes include shape, size, color, and texture. Fig. 6 provides specifics on the sample
329 dermoscopic images and the expected outcomes.

330 The classification layer of the suggested model is then used to segment the impacted portion. The
331 suggested AGEbRF technique has segmented the affected area of skin cancer. The generated
332 model was then used to assess the possibility of skin cancer.
333 This uses the segmented output as the test to predict potential skin cancer cases in a given
334 dataset. Finally, the probability of skin cancer is determined using the segmented images. Here,
335 the dataset images are regarded as the testing images while the segmented images are used as
336 input images. In these images, the created AGEbRF model is used to assess the possibility of the
337 disease, whether benign or malignant.
338 The probability of skin cancer is 50 percent benign and 50 percent malignant when utilizing the
339 proposed AGEbRF model. Here, the input images are taken into consideration as 1000 numbers,
340 of which 500 are malignant and 500 are benign. As a result, the suggested model pinpoints the
341 skin cancer's affected area and fig. 7 detailed the comparison of accuracy and loss of established
342 framework.
343 The designed AGEbRF technique is implemented using a Python program, and several metrics
344 like accuracy, sensitivity, specificity, F1-measure, and precision are measured. Additionally, the
345 developed approach is verified using established techniques such as Hybrid Feature Fusion and
346 ML (HFF) [22], Skin Cancer Classification (SCC) [23], Modified Thermal Exchange (MTE)
347 [24], and Diagnosis of skin cancer by ML (DSC) [25], Region extraction and Classification of
348 Skin Cancer (RCSC) [28] and Ensemble Lightweight Technique (ELT) [29].
349

350 1) ACCURACY MEASUREMENT
351 The degree of measurement of the proposed model's working efficiency is known as accuracy.
352 Additionally, Eqn. (6) expresses the proportion of correctly expected observance to all
353 observations.

354
$$A_c = \frac{t_p + t_n}{t_p + t_n + f_p + f_n}$$
(6)

355 Where t_p is denoted as the true positive of accurate prediction and precise segmentation, t_n is
356 represented as the true negative of accurate prediction and incorrect segmentation, f_p is denoted
357 as the false positive of inaccurate prediction and exact segmentation, and f_n is considered as the
358 false negative of inaccurate prediction and incorrect segmentation of affected skin part.
359 Using established techniques including HFF, SCC, MTE, DSC, and RCSC approaches that
360 resulted in values as listed in table 2, the correctness of the proposed AGEbRF model is
361 estimated and validated.
362 The new strategy has a higher accuracy value of 99.96% than other ways, but the existing
363 approaches have achieved lesser accuracy, demonstrating the usefulness of the proposed
364 approach. Fig. 8 displays a comparison of accuracy measurement results.
365

366 2) SENSITIVITY MEASUREMENT

367 The amount of accurately anticipated true positives is determined using sensitivity. Additionally,
368 Eqn. (7) measures the probability of predicting the skin's affected area.

369
$$S_{en} = \frac{t_p}{t_p + t_n}$$
(7)

370 The sensitivity of the developed AGEbRF technique is calculated and validated using popular
371 methods such as the HFF, MTE, SCC, ELT, DSC, and RCSC approaches described in the table.
372 3.
373 Existing approaches have only achieved a sensitivity of nearly 95 percent. Furthermore, this
374 AGEbRF method achieved 99.34 percent higher sensitivity than other methods, demonstrating
375 the effectiveness of the developed model. Fig. 9 depicts the sensitivity comparison graphically.
376

377 3) SPECIFICITY MEASUREMENT
378
379 Specificity is defined as the degree used to determine the number of true negatives that are
380 correctly recognized. Additionally, specificity is used to calculate the efficiency of predicting
381 skin cancer using segmented images, which is calculated using Eqn. (8),

382
$$S_{pe} = \frac{t_n}{t_p + t_n}$$
(8)

383 The specificity of the developed AGEbRF technique is calculated and validated using commonly
384 used methods such as HFF, SCC, MTE, ELT, DSC, and RCSC approaches, which yielded values
385 as shown in table 4.
386 The existing approaches achieved a lower specificity of nearly 98 percent. As a result, the
387 developed approach achieved 99.55 percent higher specificity than other methods, demonstrating
388 the effectiveness of the developed model. In addition, fig. 10 depicts the comparison of
389 specificity.

390 4) PRECISION MEASUREMENT
391 This procedure has been tested for determining the number of correct positive estimates alienated
392 by the total positive estimates. Furthermore, precision is the proportion of precise cancer-affected
393 region diagnosis computed using Eqn. (9),

394
$$P_{re} = \frac{t_p}{t_p + f_p}$$
(9)

395 Using established techniques including HFF, SCC, MTE, ELT, and RCSC approaches, the
396 suggested AGEbRF model's precision is estimated and validated, and the results are listed in
397 table 5. Here, the current methods have only managed a lower precision of roughly 96 percent.
398 Additionally, the created AGEbRF methodology obtained a 99.29% higher precision value than
399 other methods and the comparison of precision is shown in fig. 11.

400 5) F1-SCORE MEASUREMENT
401 The computation is based on measurements of sensitivity and precision used to determine how
402 effectively the images are traced using an equation. (10),

403 $$F1 - score = 2\left(\frac{P_{re} * S_{en}}{P_{re} + S_{en}}\right)$$ (10)

404 Where P_{re} denotes the calculated precision value and S_{en} represents the calculated sensitivity
405 value.
406 The suggested AGEbRF model's F1-score is determined and validated using widely used
407 techniques like HFF, SCC, MTE, DSC, and RCSC approaches. The results are listed in the table.
408 6.
409 Here, the suggested model has outperformed previous methods with a high F1-score value of
410 99.12% compared to the existing approaches, as shown in fig. 12.
411

Discussion

413 By achieving the best results in accuracy, sensitivity, specificity, F-measure, and precision, the
414 suggested model of AGEbRF has demonstrated good performance. As a result, the devised
415 approach fixed the training problems at the beginning. Extraction of characteristics based on
416 color, shape, size, and texture features is the next step. Additionally, segmentation is carried out
417 in the classification layer and predicts the melanoma's malignancy or benignity in the output
418 layer. As a result, the newly developed AGEbRF technology improves the performance of
419 predicting affected skin portions.
420 The outstanding metrics comparisons are listed in table.7 involves the comparison Hybrid
421 Feature Fusion and ML (HFF) [22], Skin Cancer Classification (SCC) [23], Modified Thermal
422 Exchange (MTE) [24], and Diagnosis of skin cancer by ML (DSC) [25], Region extraction and
423 Classification of Skin Cancer (RCSC) [28] approaches, and the suggested AGEbRF had the best
424 results across every parameter validation. Furthermore, the created framework improved in terms
425 of predicting accuracy, sensitivity, and specificity, each of which reached 99.55 percent. Thus,
426 the proposed AGEbRF robustness is confirmed, and it can divide the skin's affected areas and
427 identify its stages, such as benign or malignant. To precisely segment the part of the body
428 affected by skin cancer, the proposed AGEbRF technique is used. Here, utilizing dermoscopic
429 images, the suggested method has demonstrated great performance for segmenting and
430 predicting processes in skin cancer.
431

Conclusions

433 This study develops a novel AGEbRF approach for detecting and segmenting skin cancer-
434 affected areas. The dermoscopic images are used as input images for training the system. Pre-
435 processing, feature extraction, classification, segmentation, and prediction are all part of the
436 proposed AGEbRF model. Furthermore, the developed model is applied to a dermoscopic image
437 dataset, and the affected areas of the skin cancer are segmented. Furthermore, the segmented
438 images are used to detect the presence of skin cancer. Furthermore, tThe proposed model
439 achieved better accuracy, sensitivity, specificity, and precision results. When compared to
440 existing models, it achieved 99.96 percent accuracy in predicting skin cancer-affected parts. Still
441 the future works should concentrate on unbiased data and to check if these un-biasing procedures

442 would create an impact in the results obtained and to work on the updated datasets to compare
443 the performance in several versions of datasets and on different tasks.

444
445

Data Availability Statement

447 The data taken up for study included with it the ISIC challenge Datasets 2018, accessible at
448 https://challenge.isic-archive.com/data/#2018 referrred in [26,27]

449
450

References

452 [1] Nandhini, V.L., Suresh Babu, K., Roy, S.K., Sharan, P. (2021). Multichannel Biosensor for
453 Skin Type Analysis. In: Patnaik, S., Yang, XS., Sethi, I. (eds) Advances in Machine Learning
454 and Computational Intelligence. Algorithms for Intelligent Systems. Springer, Singapore.
455 https://doi.org/10.1007/978-981-15-5243-4_57
456 [2] Reichrath J. Lessons Learned from Paleolithic Models and Evolution for Human Health: A
457 Snap Shot on Beneficial Effects and Risks of Solar Radiation. Adv Exp Med Biol. 2020;1268:3-
458 15. doi: 10.1007/978-3-030-46227-7_1. PMID: 32918211.
459 [3] Daniels, Benjamin, et al. "Risk of squamous cell carcinoma of the lip and cutaneous
460 melanoma in older Australians using hydrochlorothiazide: A population-based case-control
461 study." Basic & Clinical Pharmacology & Toxicology 127.4 (2020): 320-328.
462 [4] Holte, Kristofer, and Gemma McIntyre. "The Skin (Integument): PART II—Dermatological
463 Diseases Commonly Seen in the Head and Neck." Diseases and Injuries to the Head, Face and
464 Neck. Springer, Cham, 2021. 1625-1683.
465 [4] Kassem, Mohamed A., Khalid M. Hosny, and Mohamed M. Fouad. "Skin lesions
466 classification into eight classes for ISIC 2019 using deep convolutional neural network and
467 transfer learning." IEEE Access 8 (2020): 114822-114832.
468 [5] Boutry, Justine, et al. "The evolution and ecology of benign tumors." Biochimica et
469 Biophysica Acta (BBA)-Reviews on Cancer (2021): 188643.
470 [6] Vardasca, Ricardo, and Carolina Magalhaes. "Towards an Effective Imaging-Based Decision
471 Support System for Skin Cancer." Handbook of Research on Applied Intelligence for Health and
472 Clinical Informatics. IGI Global, 2022. 354-382.
473 [7] Kassem, Mohamed A., Khalid M. Hosny, and Mohamed M. Fouad. "Skin lesions
474 classification into eight classes for ISIC 2019 using deep convolutional neural network and
475 transfer learning." IEEE Access 8 (2020): 114822-114832.
476 [8] Poggiali, Erika, et al. "Lactate dehydrogenase and C-reactive protein as predictors of
477 respiratory failure in CoVID-19 patients." Clinica chimica acta 509 (2020): 135-138.
478 [9] Zhang, Ziqiang, et al. "Deep learning-based automatic recognition network of agricultural
479 machinery images." Computers and Electronics in Agriculture 166 (2019): 104978.

480 [10] Arroyo Mühr, Laila Sara, Emilie Hultin, and Joakim Dillner. "Transcription of Human
481 Papillomaviruses in Non-Melanoma Skin Cancers of the Immunosuppressed." International
482 Journal of Cancer (2021).

483 [11] Liu-Smith, Feng, Jinjing Jia, and Yan Zheng. "UV-induced molecular signaling differences
484 in melanoma and non-melanoma skin cancer." Ultraviolet Light in Human Health, Diseases and
485 Environment (2017): 27-40.

486 [12] Alagu, S., and K. Bhoopathy Bagan. "Skin cancer classification in dermoscopy images
487 using convolutional neural network." AIP Conference Proceedings. Vol. 2336. No. 1. AIP
488 Publishing LLC, 2021.

489 [13] Antunes, Fernanda, et al. "Effective synergy of sorafenib and nutrient shortage in inducing
490 melanoma cell death through energy stress." Cells 9.3 (2020): 640.

491 [14] Nahata, Hardik, and Satya P. Singh. "Deep learning solutions for skin cancer detection and
492 diagnosis." Machine Learning with Health Care Perspective. Springer, Cham, 2020. 159-182.

493 [15] Chen, Liang, et al. "Self-supervised learning for medical image analysis using image
494 context restoration." Medical image analysis 58 (2019): 101539.

495 [16] Galdran, Adrian, Gustavo Carneiro, and Miguel A. González Ballester. "Balanced-MixUp
496 for Highly Imbalanced Medical Image Classification." International Conference on Medical
497 Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021.

498 [17] Bilbao, Imanol, and Javier Bilbao. "Overfitting problem and the over-training in the era of
499 data: Particularly for Artificial Neural Networks." 2017 eighth international conference on
500 intelligent computing and information systems (ICICIS). IEEE, 2017.

501 [18] Alom, Md Zahangir, et al. "A state-of-the-art survey on deep learning theory and
502 architectures." Electronics 8.3 (2019): 292.

503 [19] Zhou, Xiaokang, et al. "Deep-learning-enhanced human activity recognition for Internet of
504 healthcare things." IEEE Internet of Things Journal 7.7 (2020): 6429-6438.

505 [20] Botvinick, Matthew, et al. "Reinforcement learning, fast and slow." Trends in cognitive
506 sciences 23.5 (2019): 408-422.

507 [21] Nauta, Meike, et al. "Uncovering and Correcting Shortcut Learning in Machine Learning
508 Models for Skin Cancer Diagnosis." Diagnostics 12.1 (2022): 40.

509 [22] Rahman, Md Mahbubur, et al. "Hybrid Feature Fusion and Machine Learning Approaches
510 for Melanoma Skin Cancer Detection." (2022).

511 [23] Mijwil, Maad M. "Skin cancer disease images classification using deep learning solutions."
512 Multimedia Tools and Applications (2021): 1-17.

513 [24] Wei, Liu, et al. "An Optimized Method for Skin Cancer Diagnosis Using Modified Thermal
514 Exchange Optimization Algorithm." Computational and Mathematical Methods in Medicine
515 2021 (2021).

516 [25] Murugan, A., et al. "Diagnosis of skin cancer using machine learning techniques."
517 Microprocessors and Microsystems 81 (2021): 103727.

518 [26] Noel Codella, Veronica Rotemberg, Philipp Tschandl, M. Emre Celebi, Stephen Dusza,
519 David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, Harald
520 Kittler, Allan Halpern: "Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge

521 Hosted by the International Skin Imaging Collaboration (ISIC)", 2018;
522 https://arxiv.org/abs/1902.03368

523 [27] Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of
524 multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5, 180161
525 doi:10.1038/sdata.2018.161 (2018).

526 [28] Saba, Tanzila, et al. "Region extraction and classification of skin cancer: A heterogeneous
527 framework of deep CNN features fusion and reduction." Journal of medical systems 43.9 (2019):
528 1-19.

529 [29] Wei, Lisheng, Kun Ding, and Huosheng Hu. "Automatic skin cancer detection in
530 dermoscopy images based on ensemble lightweight deep learning network." IEEE Access 8
531 (2020): 99633-99647.

532 [30] Mohammadi-Balani, Abdolkarim, et al. "Golden eagle optimizer: A nature-inspired
533 metaheuristic algorithm." Computers & Industrial Engineering 152 (2021): 107050.

534 [31] Lee, H.-C.; Lin, T.-C.; Chang, C.-C.; Lu, Y.-C.A.; Lee, C.-M.; Purevdorj, B. Clinical Risk
535 Factor Prediction for Second Primary Skin Cancer: A Hospital-Based Cancer Registry
536 Study. Appl. Sci. 2022, 12, 12520. https://doi.org/10.3390/app122412520Désir, Chesner, et al.
537 "A random forest based approach for one class classification in medical imaging." International
538 Workshop on Machine Learning in Medical Imaging. Springer, Berlin, Heidelberg, 2012.

Dataset Description

2

TABLE I
DATASET DESCRIPTION

Training	Nos	Testing	Nos
Benign	1290	Benign	500
Malignant	1307	Malignant	500
Total training dataset	2597	Total testing dataset	1000

Validation Of Accuracy

TABLE II
VALIDATION OF ACCURACY

No. of samples	Accuracy (%)					
	HFF	SCC	MTE	DSC	RCSC	Proposed
10	99.85	86.9	92.79	86.12	94.8	99.96
20	98.12	85.67	90.05	84.2	93.3	99.54
30	96.7	84	89.43	82.22	91.78	99
40	94.32	83.3	87.7	80	90	98.66
50	92	81.29	85	78.29	89.12	98.12

Validation Of Sensitivity

2

TABLE III
VALIDATION OF SENSITIVITY

No. of samples	Sensitivity (%)						
	HFF	MTE	SCC	ELT	RCSC	DSC	Proposed
10	91.65	90.99	86.14	93.4	94.5	85.31	99.34
20	88.3	87.60	84.34	92	93.2	83.02	99.12
30	85.43	85	83.03	91.02	91.03	81	98.67
40	83	82.82	81.89	89.02	90	79.93	98
50	81.10	79.32	80	88	88.73	77.7	97.26

Validation Of Specificity

2

TABLE IV
VALIDATION OF SPECIFICITY

No. of samples	Specificity (%)						
	HFF	SCC	MTE	ELT	RCSC	DSC	Proposed
10	95.70	87.66	89.19	97.4	98	84.5	99.55
20	93	85.6	87	96.5	97.45	82.53	99.12
30	91.15	84.89	84.83	94.44	96.66	80	98.77
40	89	82.3	82.15	93.12	94.87	78.31	98.12
50	87.92	80	80.05	92	93	77.43	98

Validation Of Precision

2

TABLE V
VALIDATION OF PRECISION

No. of samples	Precision (%)					
	HFF	SCC	MTE	ELT	RCSC	Proposed
10	95.60	87.47	88	85.9	95	99.29
20	93.24	86.21	86.12	84.3	94.3	99
30	92.13	84.87	83	82.03	93.35	98.87
40	91.2	83.3	81.23	80.45	91.12	98.34
50	90	82	79.42	78.8	89.4	98

Validation Of F1-Score

2

TABLE VI
VALIDATION OF F1-SCORE

No. of samples	F1-score (%)					
	HFF	SCC	MTE	DSC	RCSC	Proposed
10	92	86.12	91.2	82	94.75	99.12
20	89.32	85.5	88.12	81.04	93	98.56
30	87	84.23	86.74	78.76	92.13	98
40	85.52	83	82.23	76	91.78	97.5
50	83	82.30	79.42	74.32	89.02	97

Overall Performance Metrics

2

TABLE VII
OVERALL PERFORMANCE METRICS

Methods	Performance assessment with key metrics				
	Accuracy	precision	sensitivity	specificity	F-measure
HFF	99.85	95.60	91.65	95.70	92
SCC	86.90	87.47	86.14	87.66	86.12
MTE	92.79	88	90.99	89.19	91.2
RCSC	94.8	95	94.5	98	94.75
Proposed	99.96	99.29	99.34	99.55	99.12

Figure 1

Analysis and Detection of Skin Lesion
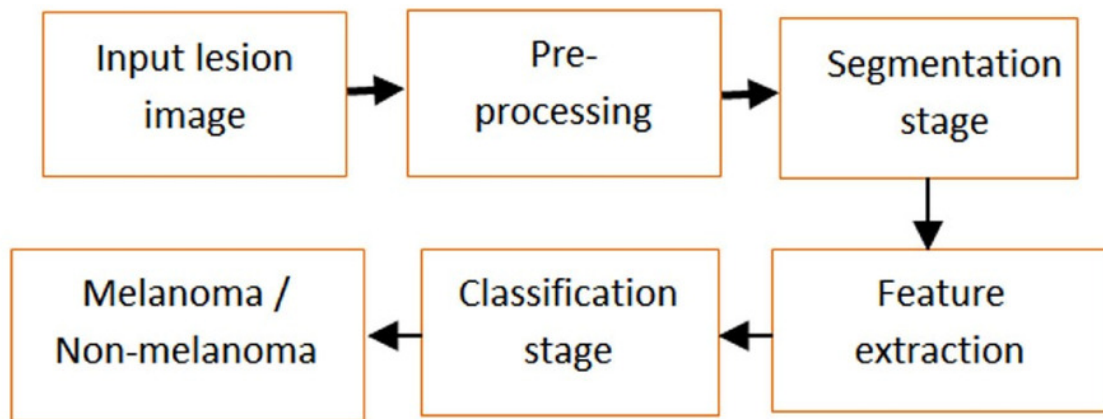
FIGURE 1.   Analysis and detection of skin lesion

Figure 2

Basic model and Problem Definition
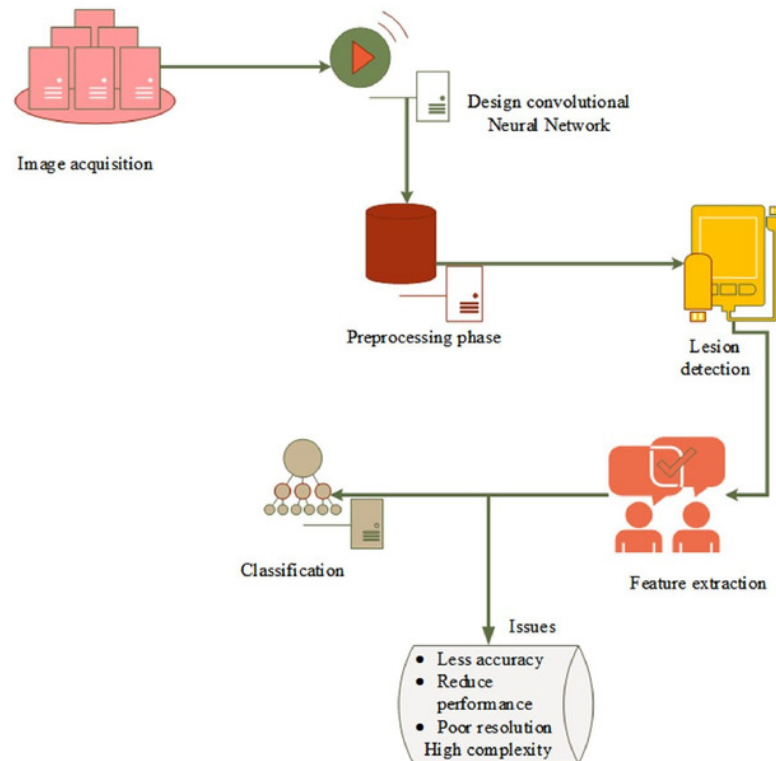
FIGURE 2. Basic model and problem definition

Figure 3

Proposed Workflow
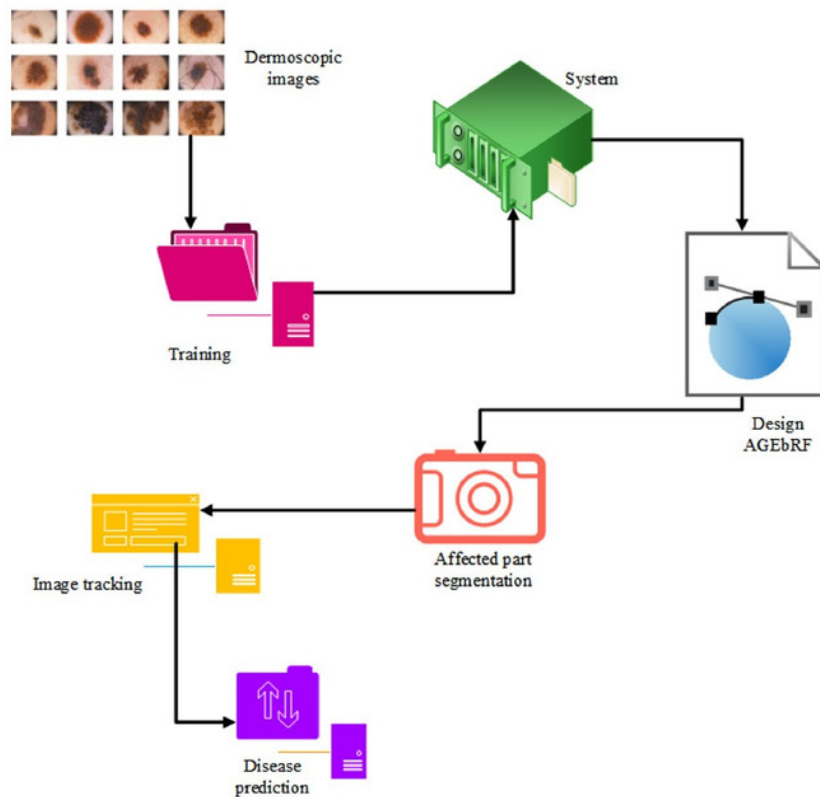
FIGURE 3. Proposed Workflow

Figure 4

Process of AGEbRF NEtwork
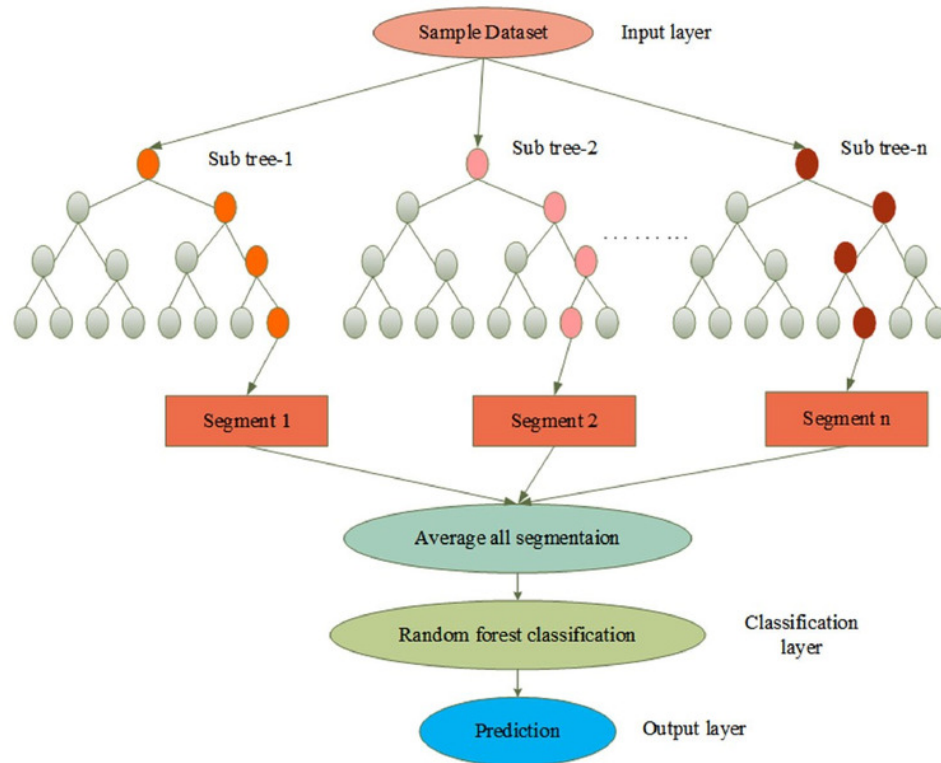
FIGURE 4. Process of AGEbRF Network

Figure 5

Flowchart of proposed AGEbRF Approach
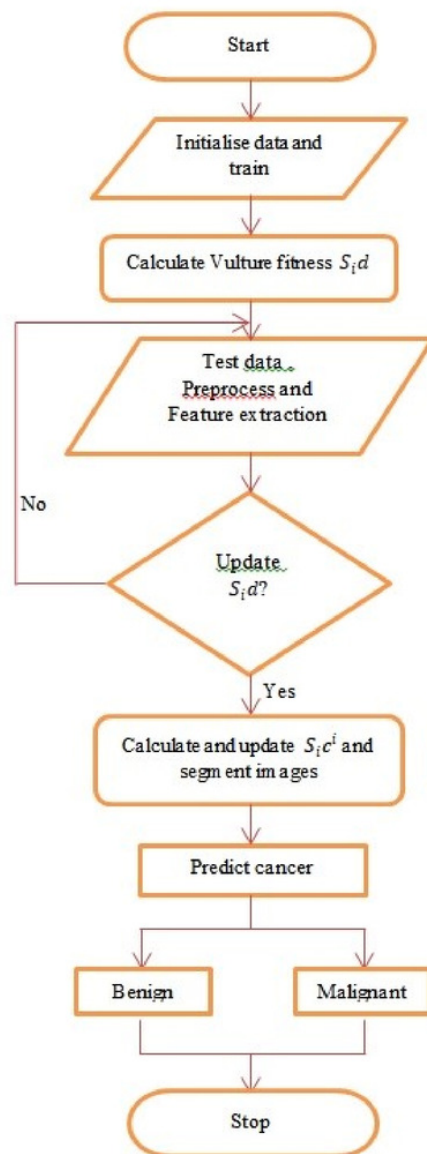
FIGURE 5. Flowchart of proposed AGEbRF Approach

Figure 6

Actual and Segmented Results
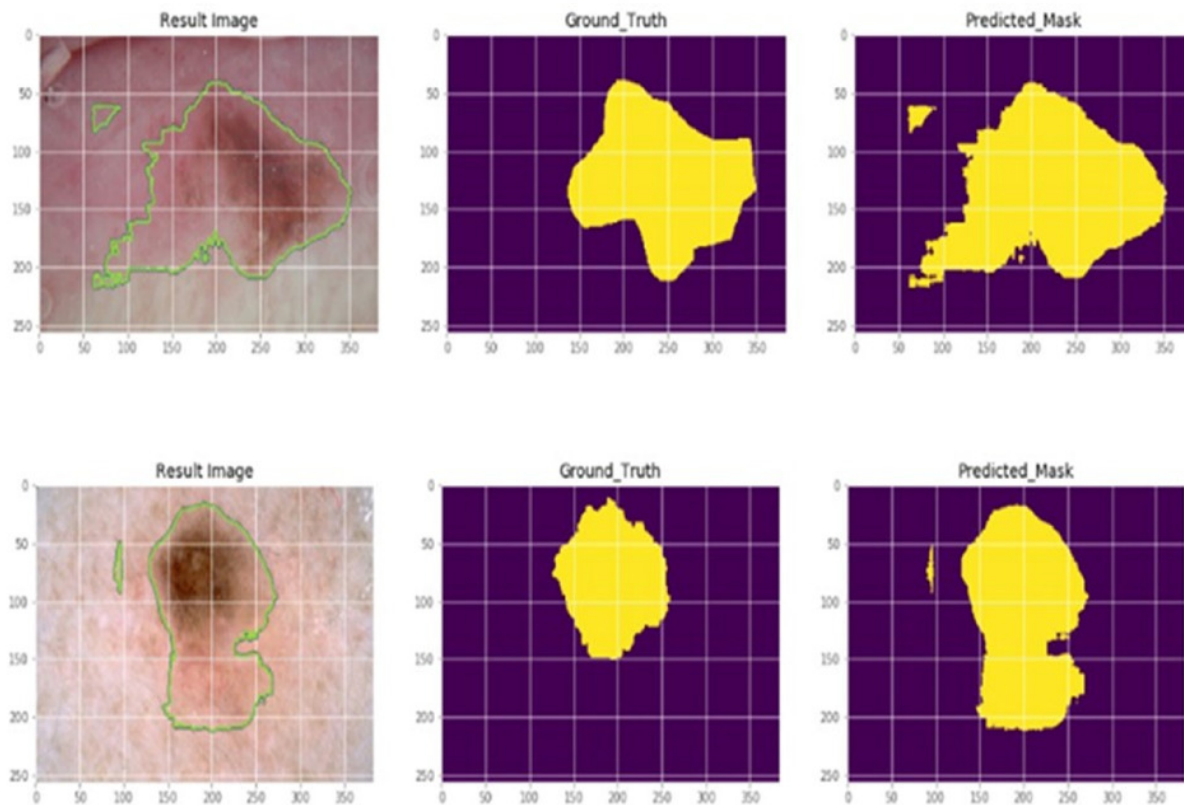
FIGURE 6. Actual and Segmented results

Figure 7

Actual and Segmented results

Accuracy and loss of actual and segmented image
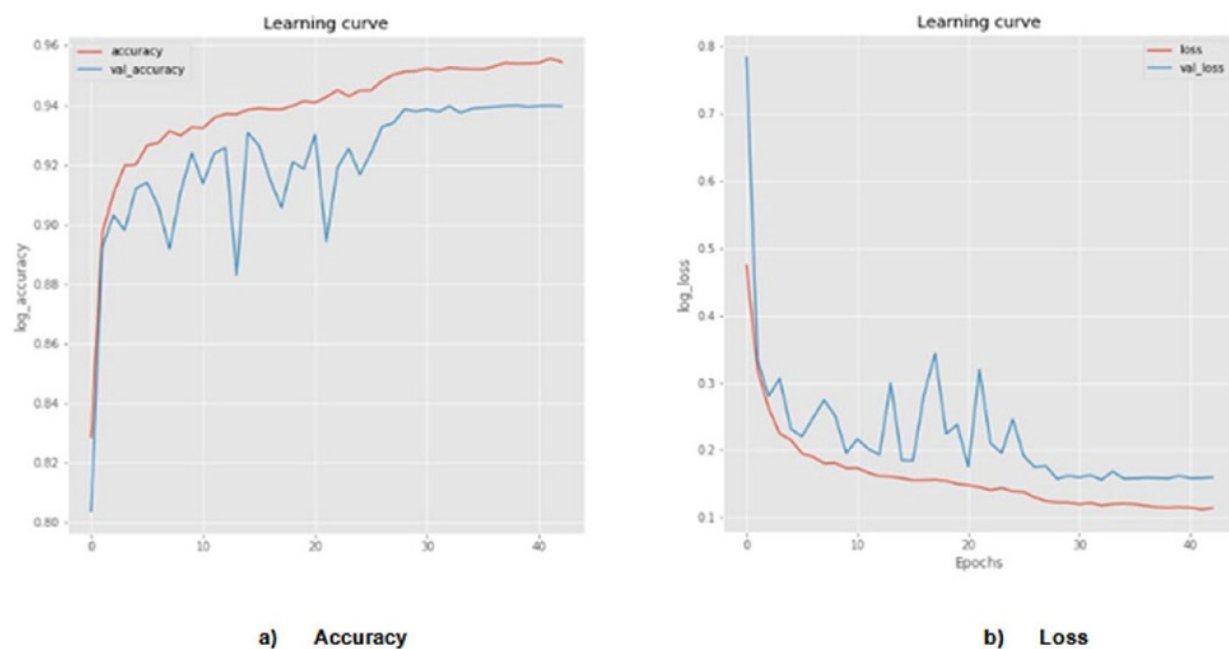
FIGURE 7. Actual and Segmented results

Figure 8

Comparison of Accuracy
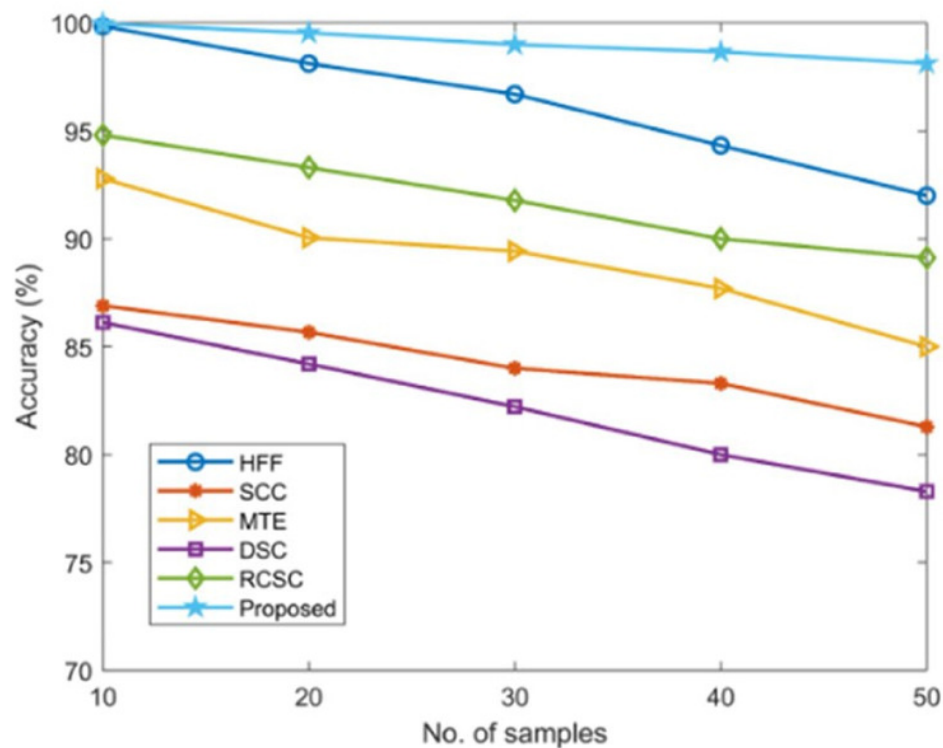
FIGURE 8. Comparison of Accuracy

Figure 9

Comparison of Sensitivity
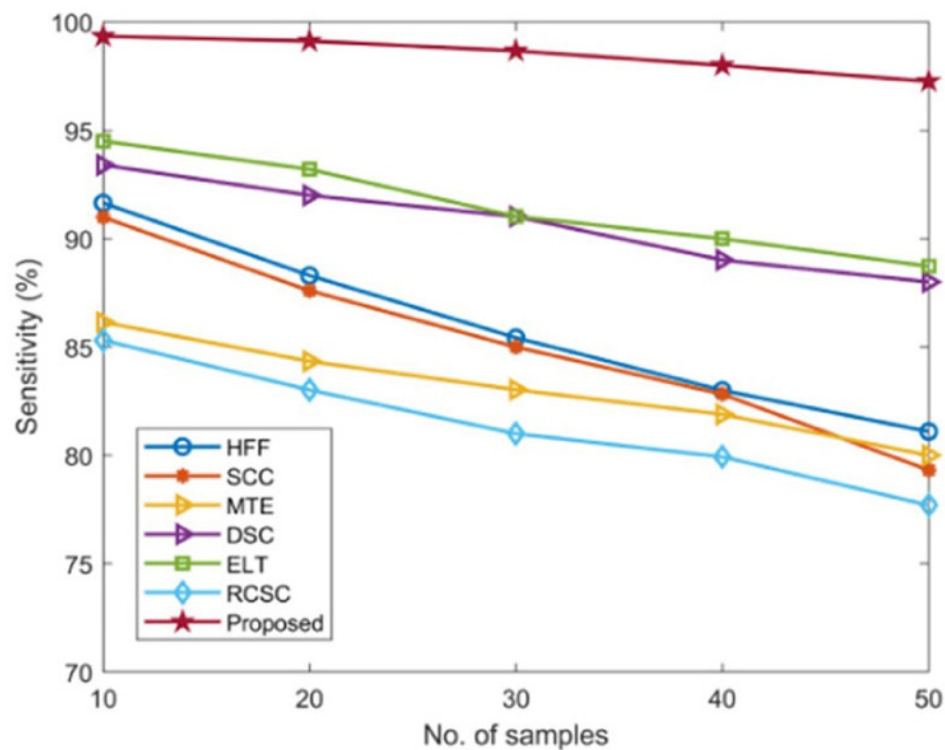
FIGURE 9. Comparison of Sensitivity

Figure 10

Comparison of Specificity
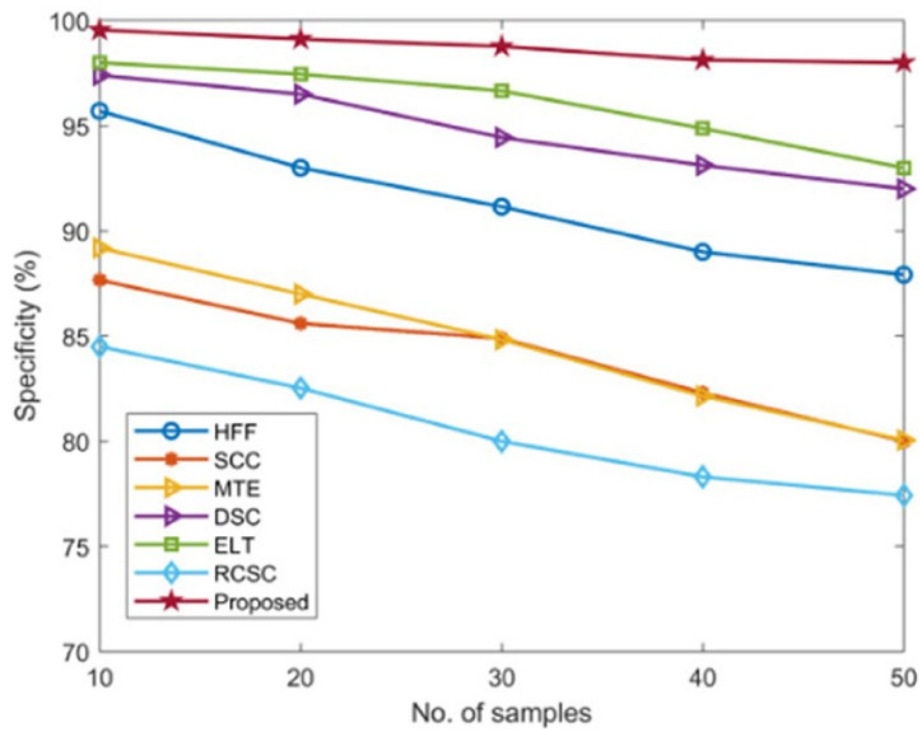
FIGURE 10. Comparison of Specificity

Figure 11

Comparison of Precision
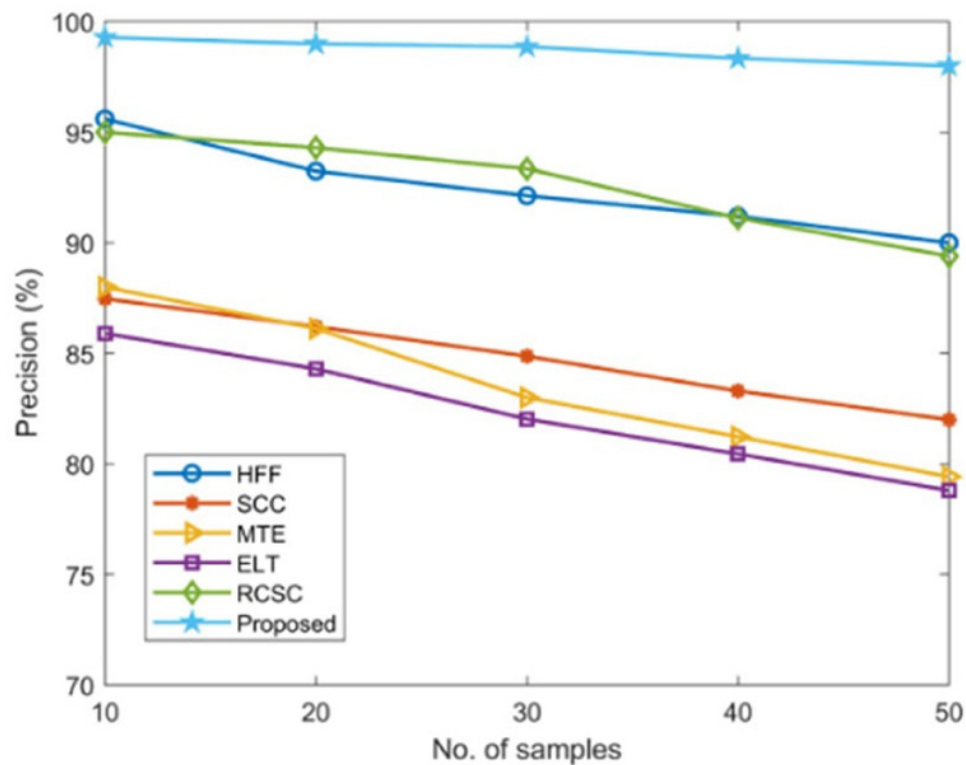
FIGURE 11. Comparison of Precision

Figure 12

Comparison of F1-Score

FIGURE 12. Comparison of F1-Score