



# A mobile augmented reality application for supporting real-time skin lesion analysis based on deep learning

Rita Francese<sup>1</sup> · Maria Frasca<sup>1</sup> · Michele Risi<sup>1</sup> · Genoveffa Tortora<sup>1</sup>

Received: 15 January 2021 / Accepted: 13 April 2021 / Published online: 3 May 2021  
© The Author(s) 2021

## Abstract

Melanoma is considered the deadliest skin cancer and when it is in an advanced state it is difficult to treat. Diagnoses are visually performed by dermatologists, by naked-eye observation. This paper proposes an augmented reality smartphone application for supporting the dermatologist in the real-time analysis of a skin lesion. The app augments the camera view with information related to the lesion features generally measured by the dermatologist for formulating the diagnosis. The lesion is also classified by a deep learning approach for identifying melanoma. The real-time process adopted for generating the augmented content is described. The real-time performances are also evaluated and a user study is also conducted. Results revealed that the real-time process may be entirely executed on the Smartphone and that the support provided is well judged by the target users.

**Keywords** Real-time mobile application · Medicine data · Deep learning · Augmented reality

## 1 Introduction

Melanoma is one of the deadliest skin cancer. It has a high rising incidence rate in the white population. Indeed, the International Agency for Research on Cancer (IARC) has estimated that 160,000 new diagnoses of skin melanoma occur annually and 62,000 of these occur in European countries [3]. When melanoma is in a metastatic stage it becomes resistant to chemo and radiotherapy: it is the cause of up to 75% of skin cancer death, while in case of early detection it is generally easy to treat by performing a simple excision.

Pigmented skin lesions are mainly due to an excessive concentration of melanin in the skin visible as darker spots on the skin. They exist in benign forms, as nevi which are melanin deposits of the epidermis. In their malignant form

melanin is produced by melanocytes at a high and abnormal rate. Their visual appearance is similar to the one of normally pigmented skin and this makes it difficult to differentiate them. This is particularly true for early-stage melanoma, whose prevention is fundamental. Dermatologists generally identify melanoma by following the "ABCDE Rule" (A = asymmetry, B = irregular edges, C = color, D = size, E = evolution), which highlights the melanoma main features for avoiding confusing it with other benign lesions.

Melanoma requires mainly a visual diagnosis by dermatologists, but this kind of recognition may be subject to errors and may take a high time for correct classification [20].

Many research work has addressed the use of Computer-Aided Diagnosis systems based on Artificial Intelligence in the early diagnosis of skin cancer. These systems propose a classification approach aiming at substituting the clinician in the diagnosis or are based on only a reduced set of parameters. Many dermatologists already use a smartphone for photographing the skin lesions of their patients. The use of augmented reality (AR) is expected to quadruplicate in the years to the visual focus dermatology and offers several opportunities and challenges [34]. Initially, the lesion measures were computed on recorded images, but as a result of the mobile devices' performance increasing the challenge is to compute them in real-time.

✉ Rita Francese  
francese@unisa.it

Maria Frasca  
mfrasca@unisa.it

Michele Risi  
mrisi@unisa.it

Genoveffa Tortora  
tortora@unisa.it

<sup>1</sup> Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 132-84084 Fisciano, SA, Italy

In [19], we proposed an approach for supporting the dermatologist in the diagnosis of melanoma which: (i) integrated a deep learning classification based on image processing and the computation of features based on the ABCD rules, (ii) applied other techniques, including a new technique (photometric stereo) never adopted for melanoma feature extraction (to the best of our knowledge) and (iii) proposed a visualization approach based on AR mobile technology supporting the skin lesion analysis.

This paper is an extension of [19] with the following improvements:

- The visualization approach has been implemented as real-time mobile application which superimposes information related to the melanoma features and the neural network classification result on the clinician camera view, while examining the framed skin lesion.
- The real-time computation process is deeply described and real-time performances are evaluated.
- A user study has been conducted to assess the app usability.
- Related work are deeply discussed and compared.

The paper is organized as follows. Section 2 discusses the background and related work concerning the diagnosis of melanoma; Sect. 3 describes the adopted real-time processing techniques and the selected features for classifying the skin lesion. Then, Sect. 4 describes the usability and real-time performance evaluation we conducted, while and Sect. 5 reports the obtained results. Finally, Sect. 6 concludes the paper and outlines future works.

## 2 Background and related work

Melanoma is a very aggressive cancer that derives from the malignant transformation of melanocytes, it can arise anew or from the modification of a pre-existing nevus. In the following, we describe the most adopted melanoma measures and the related work.

### 2.1 Melanoma measures

Typically melanoma is pigmented, with a diameter smaller than 6mm, flat or slightly raised, and of a uniform color. Due to the extreme heterogeneity of the lesions, it is very difficult to identify and diagnose melanoma early and the importance of diagnosing melanoma early is not to be underestimated, this is because the prognosis in melanoma is directly proportional to the depth of the neoplasm [33].

The ABCD rules were introduced in 1985 by Friedman et al. [22] to propose a direct and simple

interpretation tool for dermatologists. Successively, Abbasi et al. [1] improved the model by adding the letter E (evolution) for identifying a rapidly evolving lesion. Many research works validated the effectiveness of the ABCDE rule [9, 14], which evaluate the following features of a skin lesion.

(A) *Asymmetry* While a benign nevus is round, melanoma is more irregular and larger; for this reason, it is very relevant to analyze the asymmetry of the lesion. Usually, a lesion is included in a bigger spot. Asymmetry is assessed by comparing the two halves of the lesion according to its main axis. The adopted algorithm scores the symmetry of the lesion [35] as a percentage of the lesion asymmetry concerning the main axis, computed as:  $A = \frac{Dist}{\sqrt{Area}}$ , where *Dist* is the Euclidean distance between the centroid of the largest spot and the centroid of the lesion and *Area* is the area of the lesion.

(B) *Border or lesion segmentation* The process of separating the lesion from the surrounding skin to isolate the region of interest is based on edge detection or image segmentation techniques. Skin segmentation is adopted during the early phase of the CNN.

(C) *Color detection* Color is a very relevant feature because the presence of light brown, dark brown, black, red areas denotes vascularized skin region which may signal a malignant lesion [38]. Despite RGB images are largely used, often images may be represented by using an additive color composition method [41], Hue saturation brightness (HSB), also called HSV from Hue saturation value. It is an additive color composition method [41]. where *H*, *S* and *V* are defined as follows in terms of RGB color:

$$S = 1 - \frac{3}{(R + G + B)} \cdot [\min(R, G, B)]$$

$$V = \arccos \left\{ \frac{R - \frac{1}{2}(G + B)}{\sqrt{[(R - G)^2 + (R - B)(G - B)]}} \right\}$$

$$H = \begin{cases} W & \text{if } G > B \\ (2\pi - W) & \text{if } G < B \\ 0 & \text{if } G = B \end{cases}$$

(D) *Diameter* The size of a lesion is a very relevant feature because it has been shown that about 30% of lesions greater than 6 mm are invasive [37]. For detecting the diameter, we determine the center of gravity of the lesion, useful also for computing the Asymmetry [11].

(E) *Evolution* The evolution of a lesion consists of the modification of size, shape, and color in a rather short period, approximately 6–8 months. We have not yet considered this parameter in the current version of our system.

**Table 1** Technological features of related works and tools

References	Approach type	Parameters	End-users	Evaluation	Evaluation metrics	Dataset	RT	Device type	AR
[30]	C	DERM NN	C	Performance	Sensitivity, specificity	PH2, Interactive Atlas of Dermoscopy, ISIC	Unspecified	Unspecified	N
[6]	M	ABCD rules, gradient and color, hisograms	C	Performance	Sensitivity, specificity	Dermoscopic images	N	Unspecified	N
[2]	C	SVM	C	Performance	Accuracy	PH2	N	Smartphone	N
[10]	C	CNN	Unspecified	Performance	Accuracy, recall	Dermoscopic images	Unspecified	Smartphone	N
[28]	C	CNN	C	Performance	Accuracy, F-measure	Smartphone images	N	Unspecified	N
[17]	C	Lesion color, shape asymmetry, SVM, kNN	P	Performance, usability	Accuracy, sensitivity, specificity	Smartphone images	Y	Smartphone	N
[27]	M	Fractal dimension	P	Performances	Accuracy, specificity	Smartphone images	Y	Smartphone	N
[13]	C	Nearest-neighbor	P	Performances	Accuracy, sensitivity, specificity	Smartphone images	Y	Smartphone	N
This paper	C,M	ABCD, CNN	C	Performance, usability, Real-time performances	Accuracy	Smartphone images	Y	Smartphone	Y

## 2.2 Melanoma diagnosis support

Several research efforts have been devoted providing support in melanoma diagnosis. In Table 1, we classify the features of each tool depending on the following characteristics.

- *Approach type* Measure (M), classification (C). M holds when some lesion features are measured, while C corresponds to the use of a Machine Learning classification.
- *Parameters* They depend on the approach type. As an example, the considered parameters may be the ABCDE rules, in the case of M approach type, or Res-net50 in the other case.
- *End-users* The target users of the application: P (patient), C (clinician).
- *Evaluation type* Examples of evaluation types are usability evaluation or performance evaluation.
- *Evaluation metrics* They include Precision, Recall, F-measure, time, satisfaction.
- *Dataset* The dataset adopted for classification. It can be a dataset of dermoscopic images or of smartphone-taken images.
- *Real-time support* Y/N, the decision support is/is not provided in real-time.
- *Device type* Type of the device running the application. It may be a smartphone or a desktop computer.
- *Augmented reality* Y/N, the camera view is augmented by information related to the skin lesion features. Kasianos et al. [25] identified 40 smartphone apps to detect or prevent melanoma by non-specialist users. None of them was based on AR. Several works adopt artificial intelligence for classifying skin lesions [18, 30]. In the following, we detail and classify the approaches for supporting the skin lesion diagnosis following the parameters previously identified, as resumed in Table 1. Phillips et al. [30] evaluated the use of an ad-hoc developed neural network named Deep Ensemble for Recognition of Melanoma to detect malignant melanoma from dermoscopic images of pigmented skin lesions. The diagnostic accuracy achieved a ROC area under the curve (AUC) of 0.93 (95% confidence interval: 0.92–0.94), and sensitivity and specificity of 85.0 and 85.3%, respectively. Barata et al. [6] examined two systems for melanoma classification. The first system uses global methods to classify skin lesions (ABCD rules), whereas the second system uses local features, such as gradient histograms and color histograms of reduced regions. Global methods reached sensitivity of 96% and specificity of 80%, while local methods obtained sensitivity of 100% and

specificity of 75%. The dataset was composed by 176 dermoscopy images. Abuzaghlleh et al. [2] proposed a system for supporting the prevention of skin burn due to sunlight and an image analysis component that classifies the image using the PH2 dataset. The features extracted were 2-D Fast Fourier Transform, 2-D Discrete Cosine Transform, Complexity Feature Set, Color Feature Set and Pigment Network Feature Set. The experiment results revealed that the classification reached 96.3, 95.7, and 97.5% of accuracy for benign, atypical, and melanoma skin lesions, respectively. In [10], a mobile app is proposed to assist melanoma detection by using a CNN. The dataset was composed of smartphone lesion images and lesion clinical information. Due to the rarity of dermoscopic images, an evolutionary algorithm has been adopted to balance datasets. A balanced accuracy of 92% has been reached. Pacheco et al. [28] created a dataset with smartphone clinical images and patient clinical data. This information was proposed as input to deep learning models, such as CNN, to combine features from images and clinical data. Then the model performance with and without using clinical data has been compared. The accuracy was improved of 7% when additional information is considered. Hoang et al. [17] adopted a dataset of images taken by a mobile device, but the dataset is not accessible. It is devoted to self-diagnostic and the end-users are the patients. The classifier exploits numerical features to characterize a skin lesion which are a reduced set of features obtained by feature selection. We also explored the app in the stores. In particular, DermEngine [16] monitors the nevus evolution: it enables the clinician to associate a nevus picture to a specific position of the human body, in such a way as to retrieve the nevus in the future. An additional camera has to be added to the mobile device. SkinVision [27, 29] is a smartphone app that takes the lesion's picture by the phone's camera and provides the risk assessment of the lesion. The risk analysis algorithm of SkinVision is based on gray-scale image analysis and the fractal dimension of skin lesions. Recent data on a sample of proprietary smartphone images show that it reaches the sensitivity of 97% and specificity of 78%. Lūbax [13] exploits a proprietary database composed of 12,000 images of lesions certified by dermatologists. The approach creates a single high-dimensional signature for each image on the base of lesion features (i.e., size, color and shape) and a computer algorithm which compares the characteristics of new images with images in the database to identify the nearest-match diagnosis. It got sensitivity of 90.4%, specificity of 91.5%, and accuracy of 90.8%. Our mobile system supports the clinician in the diagnosis of melanoma by analyzing the images taken by a mobile phone with a Convolutional Neural Network. The ABCD

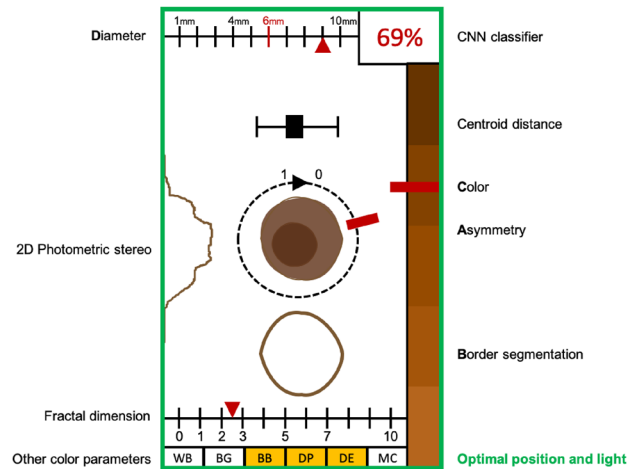


Fig. 1 A sketch of the visualization layout

rules of the nevus are also evaluated. Both these contents are shown in augmented reality modality and in Real-time while the clinician observes through the device the patient skin. The computation is conducted on the mobile device. We also conducted real-time performance and usability analysis.

### 3 The proposed approach

The system we propose aims at supporting the dermatologist during the skin lesion analysis. To this aim, we provide a lesion classification by exploiting Deep Learning techniques and adopt AR modality for visualizing in real-time the features generally evaluated for formulating a diagnosis, such as ABCDE rule and fractal dimension.

A sketch of the application layout is shown in Fig. 1, where all the considered features and the classification result are shown in real-time while exploring the patient skin. The app also enables the dermatologist to manage the patient's history with his clinical information and prescriptions. Screenshots of the app prototype are shown in Fig. 2.

#### 3.1 Real-time processing

The following fundamental steps have to be performed to provide the appropriate support:

- To analyze a nevus the dermatologist has to frame the lesion on the patient skin by using the device camera.
- To detect the distance between the patient skin and the smartphone, which should be set to 10 cm.
- To analyze the image signal if the skin lesion is in the center of the device display. When there exist several skin lesions in the center of the screen, the app high-

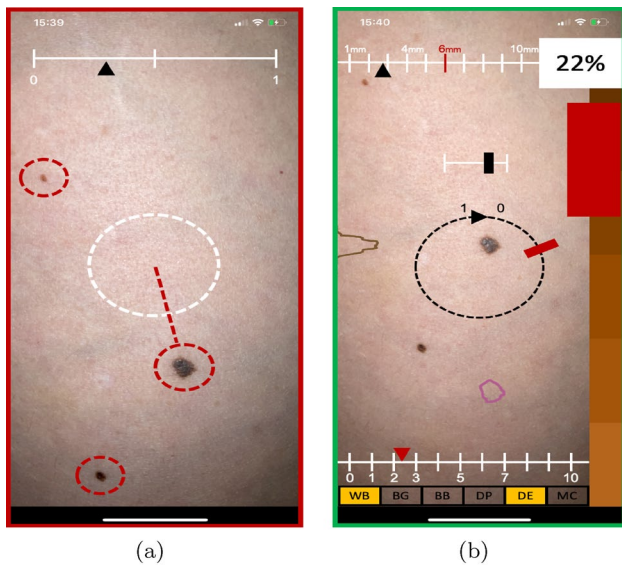


Fig. 2 Select and center a nevi (a), the AR visualization (b)

lights them, as shown in Fig. 2a and enables the clinician to select the one to be examined among them. The more central one is suggested.

- Once a nevus has been individuated the system has to keep on tracking it until it is not visible anymore.
- Classify an image and compute all the features to support the diagnosis in real-time.
- To integrate and pose the measures in AR on the mobile screen.

In this research, we focus on the collection of the skin lesion information and on the tracking of the dermatologist camera which has to overlap the support information to the

camera view in a time little enough to avoid an inappropriate waiting. When the user moves the smartphone all the information has to be recalculated and updated.

Video-based AR captures the camera video channel which is adopted as the background of the image. This image is then added it to the video of the camera. To this aim, the system has to map the position of the AR contents in real-time.

In the last years, the computational power of smartphones and tablets is largely increasing, giving them the power which was available on PC not long ago. Thus, the former may run without problems the most part of mobile applications, while some functionalities may be challenging, such as AI algorithms [23]. Thus, different client-server architectures may be adopted, depending on how we divide the computation between the client and the server. Initially, we decided to use top-level smartphones and leave all the computation on them. Then the architecture may be modified depending on the results of the real-time performance evaluation.

The real-time skin analysis process is based on the following phases, as shown in the activity diagram with object-flow depicted in Fig. 3:

1. *Real-time acquisition* A frame of the camera video is acquired.
2. *Continuous tracking* The system has to track the device’s position with respect to the patient skin. This is useful to determine the distance and the device orientation. The nevus position is also determined.
3. *Image crop* The app automatically crops the nevus.
4. *Image pre-processing* The nevus image is pre-processed for maintaining only the relevant aspects of a lesion.

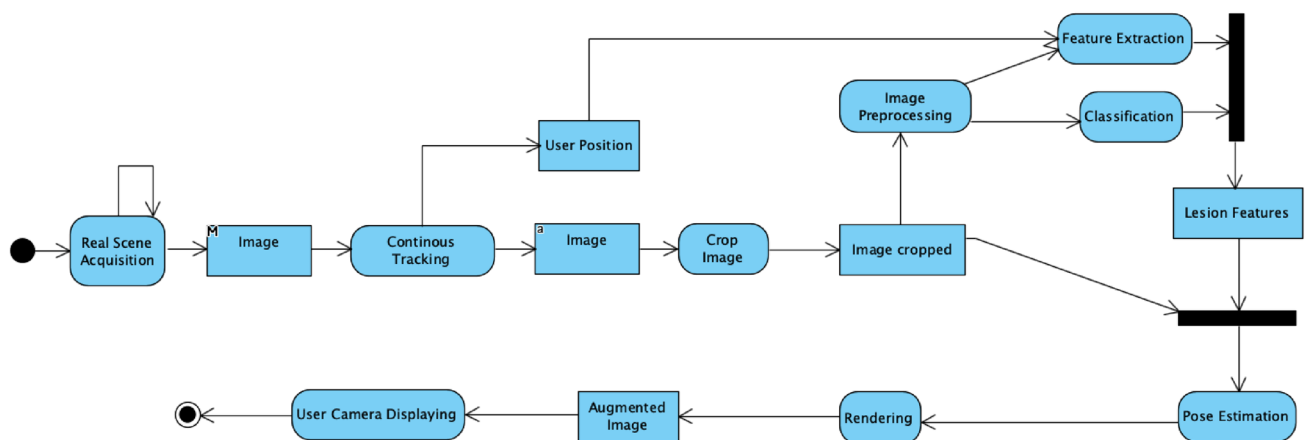


Fig. 3 The real-time process

5. *Feature extraction* ABCD rules, 2D Photometric Stereo, and Fractal Dimension are computed.
6. *Classification* CNN performs a classification of the nevus.
7. *Pose estimation* It computes the position in which the contents have to be added on the camera view with respect to the nevus position.
8. *Rendering* The evaluated features are combined with the original image.
9. *Displaying* The augmented image is shown in the dermatologist camera view.

### 3.1.1 Continuous tracking

To display information in real-time on the user camera there is the need of tracking the patient nevus while the dermatologist moves the device and determining the device's distance from the user skin. One of the main concerns to get consistent images are the lighting and the patient positioning blur [4]. The application suggests putting the point of interest in the middle of the camera view, as shown in Fig. 2b. This to avoid confusion as in the case of Fig. 2a, where the image contains two nevi.

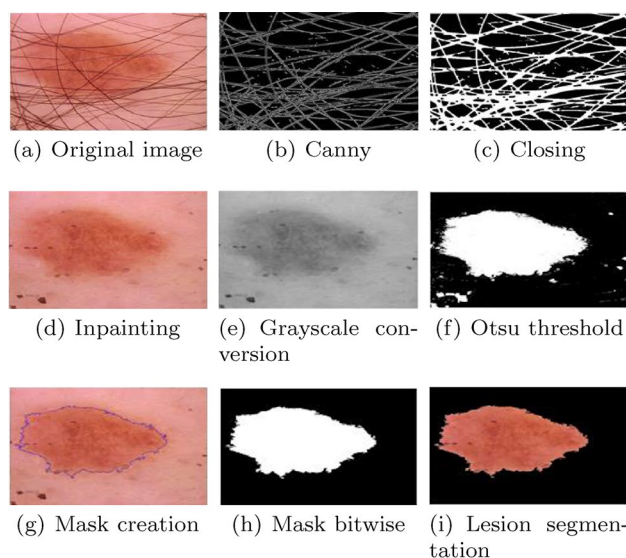
The distance between the device and the patient skin is measured by following the method proposed in [26]. In particular, this method uses the object disparity, dragging distance, and device orientation to determine the distance between the patient skin and the device. To measure the distance, users have to hold the mobile device in upright position and drag it along the vertical direction (perpendicular to the patient skin). The algorithm is based on the relation between object disparity from two images and a difference in camera positions. Acceleration signals during the dragging period are analyzed to compute the dragging distance. See [26] for further details. The application produces a red circle around the nevus when the distance is not appropriate. The circle is black when the distance is about 10cm. The distance is successively used to compute the size of the nevus.

To identify the nevus we adopted the motion tracking approach proposed in [44], where the nevus feature points taken in the first frame are tracked in the successive frames.

### 3.1.2 Image feature detection

Before using an image for extracting the relevant features of a lesion or to provide it as input to a classifier there is the need to pre-processing it. This activity is performed in the following three phases:

*Hair removal* The presence of hairs that occlude a skin lesion may interfere with its diagnosis. We used the *Canny edge detection* algorithm for the hair removal, which includes two main stages: in the former, light and dark hairs are segmented through adaptive canny edge detector and



**Fig. 4** The skin lesion image processing

refined by morphological operators. In the latter, the space left by the removed hairs is repaired on the base of the painting of the transport of coherence to several resolutions [36]. Following the Canny edge detection, the *Otsu's threshold* has been used as a mask for hair removal to obtain a black and white image. Then, the dilation operator has been used to grant greater precision in hair capture. Finally, we performed image inpainting, a technique that performs a sort of interpolation for digital image processing to reconstruct parts of damaged digital images [15]. The hair removal process is shown in Fig. 4a, b.

*Lesion segmentation* Segmentation aims at selecting specific objects or regions in an image based on a choice of properties, such as brightness, color, and texture.

The image is converted from RGB to grayscale and the lesion is separated from its background (i.e., the skin).

The Otsu's thresholding method is used to automatically perform clustering-based image thresholding or, the reduction of a gray level image to a binary image. The algorithm assumes that the image contains two classes of pixels following bimodal histogram (i.e., foreground pixels and background pixels); it then calculates the optimum threshold separating the two classes so that they are combined spread (intra-class variance) is minimal [39]. It is assumed that the edges with a gradient of intensity greater than the maximum value are real borders, while those below the minimum value are certainly not edges and therefore to be discarded. The process is shown in topmost part of Fig. 4e–i.

*Clinical feature segmentation* In this case, there is a local segmentation of the lesion, thus highlighting the clinical features of a lesion such as the texture, shape, and color. Subsequently, it was necessary to reduce the over-segmentation, a typical "problem" that occurs in the output of the previous

phase, by exploiting a technique based on temporal and spatial averaging. In this case, median filtering was used which preserves the edges by removing the over-segmentation and also the blurring effect. The median filter consists of centering a mask, sorting the pixels increasingly, and assigning the median value to the kernel pixel. Then, the *bitwise AND* binary operator was applied between the generated image and the original one. The process is shown in the bottom part of Fig. 4.

### 3.1.3 Nevus feature extraction

In the following, we describe the nevus features to be displayed in AR. We visualize information related to the ABCD features as described in Sect. 2.1 and, we also introduce the visual analysis of a peculiar characteristic of melanoma that should be considered to obtain a correct classification, the Palpable elevation of the lesion. This characteristic is evaluated by the clinician, but (for the best of our knowledge) is not considered yet in software tools. The appearance of a papule or lump in the context of a pigmented lesion can often denote the presence of a malignant lesion. A *Photometric stereo* algorithm has been adopted to measure the degree of elevation of the lesion. Photometric stereo is a method to estimate the depth and surface orientation from images of the same view taken from different directions (used in many 3D reconstruction applications) and is based on normal constructions considering the direction of light. Generally, three directions are enough to get the normals, but a larger number is needed to minimize the process noises. First, the direction of the light is computed. Lambertian surfaces are used for the generation of normal maps, where the intensity at any point on the surface can be given as  $I = \frac{N \cdot L}{\pi}$ , where  $L$  is the direction of the reflected light,  $N$  is the normal at the surface point and is computed by considering at least three light sources that are not in the same plane based on the direction of the  $x$ ,  $y$  and  $z$  axis [7]. Finally, the 3D reconstruction is mapped into a 2D image.

We also analyze the *Pigmentary reticulum and pigmentary pseudoreticle* features. The former is the most important parameter, the lines of the network correspond to the elongated epidermal crests and the spaces of the network to the dermal papillae; in benign lesions, it appears regular and nuanced in the periphery, while in suspicious lesions it appears irregular with coarse meshes, with the pigmentation of varying intensity, unshaded in the periphery and asymmetrical [24]. The Pigmentary pseudoreticle it is an interruption of pigmentation by hypopigmented patches determined by hair follicles and glandular outlets. In lentigo maligna due to the atypical melanocytes increased in number, it appears irregular and coarse [5]. In our proposal, we use the *Fractal dimension* ( $F_d$ ), which measures the repetition of each sub-structure at a specific scale applied to the

image.  $F_d$  provides the  $N_r$  distinct (non-overlapping) copies of each sub-structure scaled by a ratio  $r$  (where  $0 \leq r < 1$  when the image is scaled down) [12]. The fractal dimension  $F_d$  is given by the relation:

$$1 = N_r r^{F_d} \implies F_d = \frac{\log(N_r)}{\log\left(\frac{1}{r}\right)}$$

### 3.1.4 Melanoma classification with deep learning

To classify melanoma we adopted a Convolutional Neural Network (CNN) [40], widely used in the field of computer vision and, more generally, with data having spatial relationships. CNN uses a variable number of learnable filters. The filters in the first levels can understand simple patterns, while the following ones succeed in understanding more complex patterns [32]. Finally, the last levels, composed of fully connected neurons, are capable to make predictions.

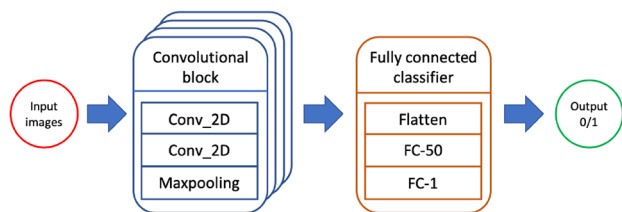
To train the CNN model we adopted a dataset composed of clinical images taken by top-level smartphones. Several kinds of lesions have been collected, including Actinic keratoses, Bowen's disease, benign keratosis-like lesions, dermatofibroma, melanoma, melanocytic nevi and vascular lesions.

We decided to consider a binary classification problem (i.e., melanoma or not melanoma). As a consequence, the data were not balanced. To solve this problem, a data augmentation process was carried out on the melanoma images. In fact, through concatenated operations composed of horizontal and vertical flips, rotations of  $180^\circ$ ,  $90^\circ$  and  $-90^\circ$ , each original image of melanoma has been used to generate seven distinct images. Thus, we have obtained a total of 8000 images.

The CNN was trained through 100 epochs with a batch size of 64 and using the Adam optimizer. We made up the neural network with 4 convolutional blocks and one last block for the classification. The training set, test set, and validation set have been formed by considering the percentages 80, 20% of the initial dataset, and 15% of the training set, respectively. The convolutional levels use a kernel size of 5 and a stride of 1 and a regularization of the L2 kernel. We used the rectified linear units (ReLU) as the activation function for each convolutional level and the Sigmoid activation function in the last fully connected level (FC) to have a binary classification of the problem. The architecture of the network is shown in Fig. 5.

### 3.1.5 Classification results

To analyze the classification results of the adopted CNN model we computed accuracy, sensitivity, and specificity



**Fig. 5** The Convolutional Neural Network architecture

**Table 2** The CNN classification results

Accuracy	Sensitivity	Specificity
78.8%	91.3%	73.0%



**Fig. 6** Nevus tracking

(see Table 2). The CNN classifier obtained an accuracy average result of 78.8%.

### 3.1.6 Pose estimation

In this phase, we determine the circle position in AR given the nevus position in the frame, as shown in Fig. 6. This is simply performed by considering the diameter dimension and the center of gravity of the lesion.

### 3.1.7 Rendering

In this phase, the app combines the augmented content and the original image on the base of the pose estimation results. In particular, the circle around the nevus has to be positioned according to the coordinate computed in the camera pose phase together with the nevus features and the classification result, as shown in Fig. 2b.

## 4 Evaluation

In this section, we describe the evaluation of the proposed approach in terms of usability and real-time performances.

**Table 3** The tasks composing the scenario of use

Task	Description
T1	Start the mobile application
T2	Logging in
T3	Enter patient data
T4	Perform the skin evaluation
T5	Examine the lesion parameters
T6	Enter the lesion evaluation
T7	End evaluation

## 4.1 Usability study

Several participatory design groups involving clinicians and researchers have been conducted before designing the application prototype analyzed in this usability evaluation. First, a mock-ups prototype has been created (a screen is shown in Fig. 1); the mock-ups were evaluated and then the running prototype has been implemented.

The *goal* of this study is to *evaluate* usability and the user satisfaction of the proposed app in a clinical setting from the *point of view* of professional dermatologists.

Six professional dermatologists voluntarily participated in the study. They were 35–60 years old and were frequent smartphone users and were accustomed to using them to take pictures of patient nevi (3 of them were male). Each one of them individually experimented in his office, adopting think-aloud protocol and under the supervision of one of the authors (the same in each evaluation). Before starting the experiment the supervisor performed a 5-min presentation showing the app's goal and operation. Participants (dermatologists and patients) provided written informed consent and were informed that their data and multimedia content were managed in an anonymous form. We asked the participants to follow the thinking aloud protocol as they were registered.

### 4.1.1 Tasks

Each dermatologist examined the skin lesions of three new patients by performing the scenario composed of the tasks reported in Table 3 for each suspect skin lesion. The scenario was derived by the requirement analysis document of the application and concerned the most relevant use cases.

The videos of the patient skin were recorded and used in the real-time evaluation of the system described in the following.

### 4.1.2 Variables and materials

We considered the following variables: the scenario completion success rates, the time (expressed in seconds) needed to



**Table 4** Post-Experiment questionnaire

ID	Question
P0	The tasks to perform were clear
P1	I think that I would like to use this system frequently
P2	I found the app unnecessarily complex
P3	I thought the app was easy to use
P4	I think that I would need the support of a technical person to be able to use this app
P5	I found the various functions in this system were well integrated
P6	I thought there was too much inconsistency in this app
P7	I would imagine that most people would learn to use this system very quickly
P8	I found the app very cumbersome to use
P9	I felt very confident using the system
P10	I needed to learn a lot of things before I could get going with this system
P11	Is the loading time of the augmented skin lesion information during the lesion analysis satisfactory
P12	The metaphors for depicting the skin lesion features during the lesion analysis are easy to understand
P13	Overall, I'm satisfied of the support offered by the tool in the skin lesion examination
P14	Open comments

accomplish the task, the number of errors while performing a task, such as navigation errors; presentation errors (e.g., failure to find and properly act upon the interface element), selection errors due to labeling ambiguities; control usage issue (wrong use of entry field). These errors were noted by the supervisor.

The dermatologist perceptions were collected at the end of the experiment through the Post-Experiment questionnaire in Table 4. In particular, the P0 question referred to the clearness of the task to perform. The participant perception of the system Usability has been collected by using the standard Italian version of the System Usability Scale (SUS) questionnaire [8], is a Likert Scale which consists of 10 questions (P1–P10 questions in Table 4). Each question is ranked from 1 (disagree vehemently) to 5 (strongly agree). We also add three questions with the same Likert scale: P11 and P12 for explicitly collect perceptions on the loading time and the visual metaphor of the augmented content during the nevus analysis, respectively; P13 (overall) resumes the participant opinion on the support offered by the tool and P14 collects open comments.

## 4.2 Real-time implementation detail and evaluation

The challenge of an AR application on a mobile device is speed [42]. In particular, the main requirements of our application are:

- performing the tracking process in a short time;
- computing the lesion features and classification in a short time;
- providing the augmented information back to the dermatologist in a short time.

The proposed approach was tested by creating video test sequences at a 2160p video at 60 and 30 fps, the video format of Samsung S10. We extracted from the videos produced during the usability evaluation 12 sequences (two for each dermatologist): 6 with melanoma, 6 with not pathological nevi, having a length ranging from 710–1320 frames. Following the direction traced in [42], we compared the proposed approach performance on a mobile device and on a desktop in such a way to be able to decide which activities shown in Fig. 3 may all be performed on the client or on the server. To this aim, there is the need of providing the video stream to both systems. Thus, we developed a frame server loading a video previously taken by the dermatologist mobile device from the file system of both the devices in substitution of the mobile camera. Because performances scale linearly with the CPU clock rate on mobile phones they do not depend on the operating system [43], we performed our benchmark by using a single device. The Samsung S10 mobile device was equipped with a Qualcomm Snapdragon 855 processor, 1785 MHz, which allows the running of trained neural networks on the device without a need for connection to the cloud [31]. The MacBookPro (2019) had a processor with 2.6GHz 6-core Intel Core i7, with 12MB shared L3 cache, 512GB SSD, 16GB of 2400MHz DDR4 onboard memory.

The CNN model was trained on a Desktop PC and then uploaded on the Android smart-phone. In particular, Qualcomm enables us to convert a pre-trained model into the deep learning container (DLC) format. The Qualcomm Neural Processing Engine (NPE) runtime then executes the neural network.

We adopted the TensorFlow<sup>1</sup> framework for the model training because it exposes C/C++ API compatible with the Android platforms.

The times which have a greater impact on the performances are related to the following activities:

- *Initialization time* This is the time intercurring between the pressing of the app icon on the mobile screen and the moment in which the app is ready to accept user input.
- *Tracking time* The tracking time is the time needed to compute the user position and the nevus feature points.
- *Pre-processing time* This is the time taken to perform the pre-processing phase described in Sect. 3.1.
- *Classification time* This is the time the server takes to classify the skin lesion, as discussed in Sect. 3.1.4.
- *Feature computation time* The time the server takes to the lesion features described in Sect. 3.1.3.
- *Pose estimation time* The time needed to compute the AR content position.
- *Rendering time* The time to superimpose the augmented content to the original frame.

## 5 Results

In this section, we report the results related to the usability evaluation and real-time performances.

### 5.1 Usability results

The scenario completion success rate was 100%. All the scenarios were completed by all the dermatologists with all the patients. The time needed to accomplish each task is reported in Fig. 7a, while the number of errors related to each task was summarized in Fig. 7b. The information related to usability perception and dermatologist satisfaction is depicted in Fig. 8.

#### 5.1.1 Time

The average time spent to perform the various tasks is summarized in the histogram in Fig. 7a, except for the time related to T5 (examine the lesion parameter), which depends on the lesion complexity and of the dermatologist need of observing it and by the augmented contents. The analysis of the time of the real-time process related to task T5 is reported in the next section. In particular, 2.3 s are needed for starting the application. The longest time is taken for filling in the form concerning patient data.

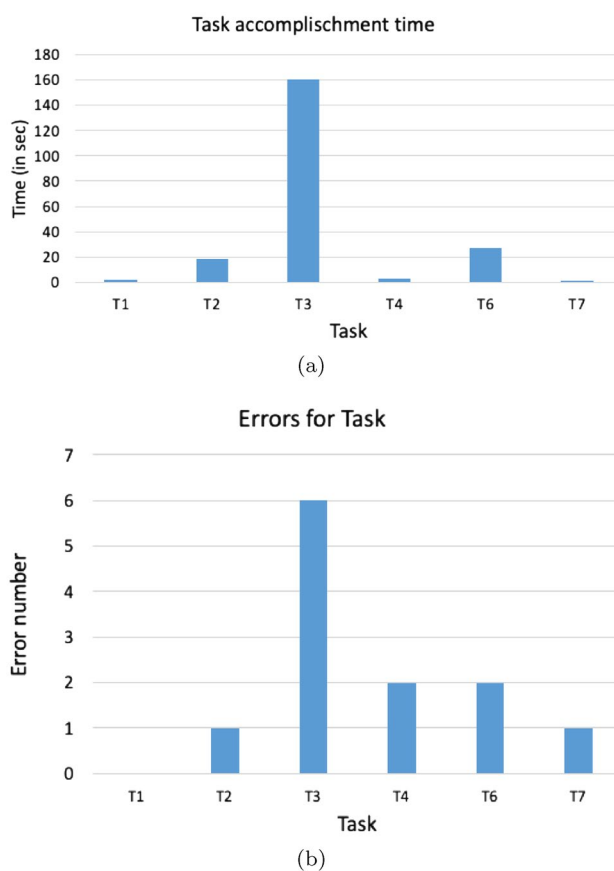


Fig. 7 Average time for task (a) and number of errors for task (b)

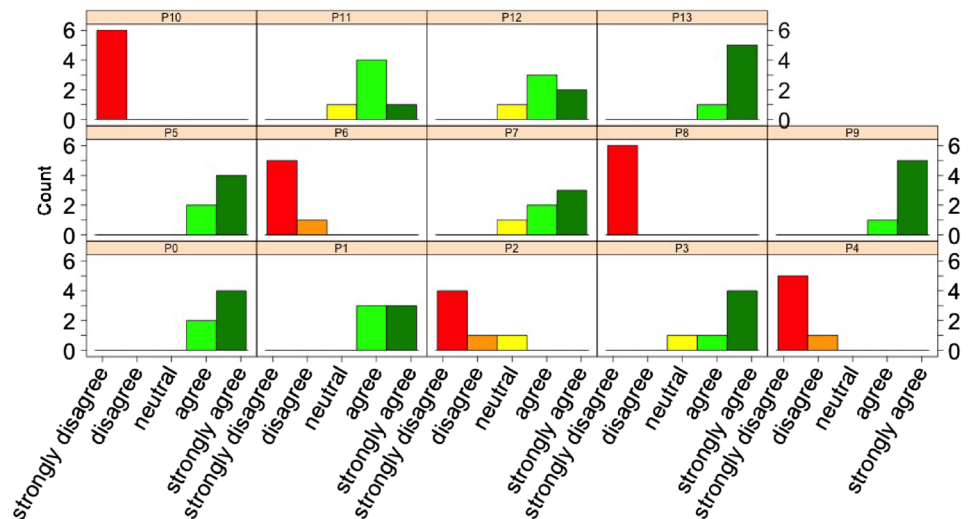
#### 5.1.2 Errors

The number of errors was very reduced. It may be because the users were all smart users. There was some navigation problem in T2 and T6, for accessing the functionalities a wrong button was pressed. The main problem was with task T3, enter the patient data. The main problems were with the data fields, as the supervisor referred.

#### 5.1.3 Usability

The user perception results collected by the Post-Test questionnaire in Table 4 are summarized by the histograms in Fig. 8, where a histogram is associated with each question. In particular, the tasks to perform were clear for all the dermatologists (P0), all of them also think to use the app frequently (P1), of which three strongly agreed. The app was not considered excessively complex (P2). The app was easy to use for five of them (P3) and the other was neutral. All of them agree that there is no need of support for using the system (P4) and the system functions were well integrate (P5). No inconsistencies were perceived (P6), while the app was perceived as easy to

<sup>1</sup> <https://www.tensorflow.org>.

**Fig. 8** The Post-Experiment questionnaire results

learn for 5 participants, one neutral (**P7**). The system was not perceived as cumbersome (**P8**), while all the participants agree that they were confident when using it (**P9**). They also considered unnecessary to learn a lot of things to be able to use the system (**P10**). Both **P11** the loading time and **P12** the adopted AR metaphors are satisfying for 5 participants. The overall satisfaction on the support provided to the skin lesion analysis is very positive for 5 participants, positive for one of them.

The same user was neutral in P3, P11, and P12. He explained in the open comments section that “The bottom ruler meaning is not clear to me. I had some difficulty in stopping the skin analysis.” Another participant lamented the difficulty in remembering the meaning of the same ruler. Thus we decided to activate an information pop-up when the user touches an element of the interface and to express the symmetry with a numeric value (1—perfectly symmetric, 0—totally asymmetric).

An appreciation in P13 was addressed to the 3D representation of the lesion, which was considered particularly useful. The same participant suggested inserting a “start analysis button” in such a way the dermatologist may observe the lesion through the device and have feedback from the system only when he needs it. We considered this suggestion very useful and decided to add it to the next prototype version.

## 5.2 Real-time performances

The results in terms of performances of the mobile device concerning the laptop are reported in Table 5. We can see that on the mobile device a cycle of the real-time process takes roughly 5 s.

## 5.3 Discussion and final remarks

Based on the real-time performance analysis we decided to leave on the mobile device all the activities. We measured the average processing time spent for each image and it is about 5 s (excluding the initialization time). We have also to consider that the algorithms do not have been executed in parallel by exploiting the device GPU. The new smartphone devices such as the Qualcomm Snapdragon Platforms of Samsung Galaxy offer the SDK for parallel development. Thus, if feature extraction and classification steps are performed in parallel the whole time may be reduced at about 4 s. These performances may be reached only on top-level smartphones, but because the application has the dermatologists as target users this requirement may be acceptable.

Concerning the accuracy level reached by the adopted CNN, it is worth mentioning that the model performs a binary classification but that the dataset images referred to different skin lesions. Moreover, smartphone images have a lower resolution than dermatoscopic ones.

**Table 5** Average process performance measures

Activity	Mobile	Desktop
Tracking time	38 ms	3.8 ms
Image pre-processing	1.51 s	1.11 s
Feature extraction	0.9 s	0.35 s
Classification	2.2 s	1.5 s
Pose estimation	340 ms	9 ms
Rendering	11 ms	1.9 ms

## 6 Conclusion

Melanoma has a large incidence and the trend is growing, in particular for white people. According to [21], many existing apps for melanoma self-diagnosis could be harmful when the recommendations are wrong, particularly in case of false-negative results, which may cause delays in medical intervention. For this reason, we provide a tool that has dermatologists as target users and only provide a decision support. In this paper, we presented an AR mobile application for supporting the computer-assisted analysis of skin lesions in real-time. To this aim the lesion classification provided by a CNN and by skin lesion features are fused in AR modality on the mobile device camera of the dermatologist. We described in detail the real-time process proposed to display the augmented nevus information and evaluated the real-time performances and the app usability. Because of the pandemic period conducting empirical work (e.g., recruiting a large number of participants) was challenging. We were able to perform a very preliminary evaluation but the obtained feedback was encouraging. Basing on the real-time performances we decide of leaving all the computation on the smartphone. We plan to perform further testing of the tool components and to parallel some process phases to further reducing the processing time.

**Acknowledgements** The authors would like to thank the volunteer participants (dermatologists and patients) who took part in the evaluation.

**Funding** Open access funding provided by Università degli Studi di Salerno within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abbasi, N.R., Shaw, H.M., Rigel, D.S., Friedman, R.J., McCarthy, W.H., Osman, I., Kopf, A.W., Polsky, D.: Early diagnosis of cutaneous melanoma: revisiting the abcd criteria. *JAMA* **292**(22), 2771–2776 (2004)
- Abuzaghlh, O., Barkana, B.D., Faezipour, M.: Noninvasive real-time automated skin lesion analysis system for melanoma early detection and prevention. *IEEE J. Transl. Eng. Health Med.* **3**, 1–12 (2015)
- Apalla, Z., Lallas, A., Sotiriou, E., Lazaridou, E., Ioannides, D.: Epidemiological trends in skin cancer. *Dermatol. Pract. Concept.* **7**(2), 1 (2017)
- Ashique, K., Kaliyadan, F., Aurangabadkar, S.J.: Clinical photography in dermatology using smartphones: An overview. *Indian Dermatol. Online J.* **6**(3), 158 (2015)
- Bär, M., Tschandl, P., Kittler, H.: Differentiation of pigmented spitz nevi and reed nevi by integration of dermatopathologic and dermatoscopic findings. *Dermatol. Pract. Concept.* **2**(1), 13 (2012)
- Barata, C., Ruela, M., Francisco, M., Mendonça, T., Marques, J.S.: Two systems for the detection of melanomas in dermoscopy images using texture and color features. *IEEE Syst. J.* **8**(3), 965–979 (2013)
- Basri, R., Jacobs, D., Kemelmacher, I.: Photometric stereo with general, unknown lighting. *Int. J. Comput. Vis.* **72**(3), 239–257 (2007)
- Brooke, J.: SUS—a quick and dirty usability scale. *Usability Eval. Ind.* **189**(194), 4–7 (1996)
- Carli, P., De Giorgi, V., Crocetti, E., Caldini, L., Ressel, C., Giannotti, B.: Diagnostic and referral accuracy of family doctors in melanoma screening: effect of a short formal training. *Eur. J. Cancer Prev.* **14**(1), 51–55 (2005)
- Castro, P.B., Krohling, B., Pacheco, A.G., Krohling, R.A.: An app to detect melanoma using deep learning: an approach to handle imbalanced data based on evolutionary algorithms. In: *International Joint Conference on Neural Networks (IJCNN)*, pp. 1–6. IEEE (2020)
- Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H.: A methodological approach to the classification of dermoscopy images. *Comput. Med. Imaging Graph.* **31**(6), 362–373 (2007)
- Chaudhuri, B.B., Sarkar, N.: Texture segmentation using fractal dimension. *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(1), 72–77 (1995)
- Chen, R.H., Snorrason, M., Enger, S.M., Mostafa, E., Ko, J.M., Aoki, V., Bowling, J.: Validation of a skin-lesion image-matching algorithm based on computer vision technology. *Telemed. e-Health* **22**(1), 45–50 (2016)
- Corbo, M.D., Wismer, J.: Agreement between dermatologists and primary care practitioners in the diagnosis of malignant melanoma: review of the literature. *J. Cutan. Med. Surg.* **16**(5), 306–310 (2012)
- Criminisi, A., Perez, P., Toyama, K.: Object removal by exemplar-based inpainting. In: *Computer Vision and Pattern Recognition*, vol. 2. IEEE (2003)
- Dermengine.: <https://www.dermengine.com>
- Do, T., Hoang, T., Pomponiu, V., Zhou, Y., Chen, Z., Cheung, N., Koh, D., Tan, A., Tan, S.: Accessible melanoma detection using smartphones and mobile image analysis. *IEEE Trans. Multimed.* **20**(10), 2849–2864 (2018)
- Fernandez Alcon, J., 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. *J. Sel. Top. Signal Process.* **3**(1), 14–25 (2009)
- Francese, R., Frasca, M., Risi, M., Tortora, G.: An augmented reality mobile application for skin lesion data visualization. In: *24th International Conference on Information Visualisation (IV)*, pp. 51–56 (2020)
- Frasca, M., Nappi, M., Risi, M., Tortora, G., Citarella, A.A.: A comparison of neural network approaches for melanoma classification. In: *25th International Conference on Pattern Recognition (ICPR)* (2020)
- Freeman, K., Dinnes, J., Chuchu, N., Takwoingi, Y., Bayliss, S.E., Matin, R.N., Jain, A., Walter, F.M., Williams, H.C., Deeks, J.J.: Algorithm based smartphone apps to assess risk of skin cancer

- in adults: systematic review of diagnostic accuracy studies. *BMJ* **368** (2020)
22. Friedman, R.J., Rigel, D.S., Kopf, A.W.: Early detection of malignant melanoma: the role of physician examination and self-examination of the skin. *CA Cancer J. Clin.* **35**(3), 130–151 (1985)
  23. Ignatov, A., Timofte, R., Chou, W., Wang, K., Wu, M., Hartley, T., Van Gool, L.: Ai benchmark: Running deep neural networks on android smartphones. In: *Procs. of the European Conf. on Computer Vision (ECCV)* (2018)
  24. Juhász, M.L., Marmur, E.S.: Reviewing challenges in the diagnosis and treatment of lentigo maligna and lentigo-maligna melanoma. *Rare Cancers Therapy* **3**(1–2), 133–145 (2015)
  25. Kassianos, A., Emery, J., Murchie, P., Walter, F.M.: Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. *Br. J. Dermatol.* **172**(6), 1507–1518 (2015)
  26. Laotrakunchai, S., Wongkaew, A., Patanukhom, K.: Measurement of size and distance of objects using mobile devices. In: *Intl. Conf. on Signal-Image Technology Internet-Based Systems*, pp. 156–161 (2013)
  27. Maier, T., Kulichova, D., Schotten, K., Astrid, R., Ruzicka, T., Berking, C., Udrea, A.: Accuracy of a smartphone application using fractal image analysis of pigmented moles compared to clinical diagnosis and histological result. *J. Eur. Acad. Dermatol. Venereol.* **29**(4), 663–667 (2015)
  28. Pacheco, A.G., Krohling, R.A.: The impact of patient clinical information on automated skin cancer detection. *Comput. Biol. Med.* **116**, 103545 (2020)
  29. Peverelli, R., Feniks, R.D.: Skinvision: Leading mobile solution to monitor, track and understand skin health. *Digital Insurance Agenda* (2017). <https://www.digitalinsuranceagenda.com/feature-insurtechs/skinvision-leading-mobile-solution-to-monitor-track-and-understand-skin-health/>
  30. Phillips, M., Greenhalgh, J., Marsden, H., Palamaras, I.: Detection of malignant melanoma using artificial intelligence: an observational study of diagnostic accuracy. *Dermatol. Pract. Concept.* **10**(1) (2020)
  31. Qualcomm: Qualcomm neural processing sdk for AIDermengine. <https://developer.qualcomm.com/software/qualcomm-neural-processing-sdk>
  32. Quang, N.H., et al.: Automatic skin lesion analysis towards melanoma detection. In: *21st Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES)*, pp. 106–111. IEEE (2017)
  33. Rigel, D.S., Russak, J., Friedman, R.: The evolution of melanoma diagnosis: 25 years beyond the abcds. *CA Cancer J. Clin.* **60**(5), 301–316 (2010)
  34. Sharma, P., Vleugels, R.A., Nambudiri, V.E.: Augmented reality in dermatology: are we ready for AR? *J. Am. Acad. Dermatol.* **81**(5), 1216–1222 (2019)
  35. Stoecker, W.V., Gupta, K., Stanley, R.J., Moss, R.H., Shrestha, B.: Detection of asymmetric blotches (asymmetric structureless areas) in dermoscopy images of malignant melanoma using relative color. *Skin Res. Technol.* **11**(3), 179–184 (2005)
  36. Toossi, M.T.B., Pourreza, H.R., Zare, H., Sigari, M.H., Layegh, P., Azimi, A.: An effective hair removal algorithm for dermoscopy images. *Skin Res. Technol.* **19**(3), 230–235 (2013)
  37. Tsao, H., Olazagasti, J.M., Cordero, K.M., Brewer, J.D., Taylor, S.C., Bordeaux, J.S., Chren, M.M., Sober, A.J., Tegeler, C., Bhushan, R., et al.: Early detection of melanoma: reviewing the abcds. *J. Am. Acad. Dermatol.* **72**(4), 717–723 (2015)
  38. Umbaugh, S.E., Moss, R.H., Stoecker, W.V.: Applying artificial intelligence to the identification of variegated coloring in skin tumors. *IEEE Eng. Med. Biol. Mag.* **10**(4), 57–62 (1991)
  39. Vala, H.J., Baxi, A.: A review on Otsu image segmentation algorithm. *Int. J. Adv. Res. Comput. Eng. Technol.* **2**(2), 387–389 (2013)
  40. Vasconcelos, C.N., Vasconcelos, B.N.: Convolutional neural network committees for melanoma classification with classical and expert knowledge based image transforms data augmentation. *arXiv preprint arXiv:1702.07025* (2017)
  41. Vocaturo, E., Zumpano, E., Veltri, P.: Features for melanoma lesions characterization in computer vision systems. In: *Information, Intelligence, Systems and Applications (IISA)*, pp. 1–8. IEEE (2018)
  42. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D.: Real-time detection and tracking for augmented reality on mobile phones. *IEEE Trans. Vis. Comput. Graph.* **16**(3), 355–368 (2010)
  43. Wagner, D., Schmalstieg, D.: First steps towards handheld augmented reality. In: *7th IEEE Intl. Symposium on Wearable Computers*, pp. 127–135 (2003)
  44. Yue, S.: Human motion tracking and positioning for augmented reality. *J. Real-Time Image Process.* 1–12 (2020)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Rita Francese** is Associate Professor at the University of Salerno. She is co-author of more than 100 papers published in scientific journals or proceedings of refereed conferences. Her research interests concern artificial intelligence, software engineering, data science, empirical evaluation, human-computer interaction, collaborative work and learning, e-learning, visual languages, and mobile application development.

**Maria Frasca** is a Ph. D student in computer science at University of Salerno, her research interests are bioinformatics on genetic data, computer vision application and pattern recognition on medical images

**Michele Risi** is Associate Professor at the University of Salerno. His research interests include artificial intelligence, data science, grammar formalisms and parsing techniques for visual languages, sketch understanding, architecture and design pattern recovery, reverse engineering of Web applications, human-computer interaction, empirical evaluation, data-warehouse and data visualization, and mobile development and applications.

**Genoveffa Tortora** is full professor in Computer Science at the University of Salerno, since 1990, where she has been Department Chair, and then Dean of the Faculty of Sciences. She has co-authored more than 300 papers published in scientific journals or proceedings of refereed conferences, and has co-edited three books. Her research interests are in the software engineering and information systems areas, and include human-computer interaction, visual languages, databases, data-warehouses, geographic information systems, image processing and biometric systems.