

Article

Deep Learning Methods for Accurate Skin Cancer Recognition and Mobile Application

Ioannis Kousis ¹, Isidoros Perikos ¹, Ioannis Hatzilygeroudis ^{1,*} and Maria Virvou ²

¹ Department of Computer Engineering and Informatics, University of Patras, 26504 Patras, Greece; kousis@ceid.upatras.gr (I.K.); perikos@ceid.upatras.gr (I.P.)

² Department of Informatics, University of Piraeus, 18534 Piraeus, Greece; mvirvou@unipi.gr

* Correspondence: ihatz@ceid.upatras.gr

Abstract: Although many efforts have been made through past years, skin cancer recognition from medical images is still an active area of research aiming at more accurate results. Many efforts have been made in recent years based on deep learning neural networks. Only a few, however, are based on a single deep learning model and targeted to create a mobile application. Contributing to both efforts, first we present a summary of the required medical knowledge on skin cancer, followed by an extensive summary of the most recent related works. Afterwards, we present 11 CNN (convolutional neural network) candidate single architectures. We train and test those 11 CNN architectures, using the HAM10000 dataset, concerning seven skin lesion classes. To face the imbalance problem and the high similarity between images of some skin lesions, we apply data augmentation (during training), transfer learning and fine-tuning. From the 11 CNN architecture configurations, DenseNet169 produced the best results. It achieved an accuracy of 92.25%, a recall (sensitivity) of 93.59% and an F1-score of 93.27%, which outperforms existing state-of-the-art efforts. We used a light version of DenseNet169 in constructing a mobile android application, which was mapped as a two-class model (benign or malignant). A picture is taken via the mobile device camera, and after manual cropping, it is classified into benign or malignant type. The application can also inform the user about the allowed sun exposition time based on the current UV radiation degree, the phototype of the user's skin and the degree of the used sunscreen. In conclusion, we achieved state-of-the-art results in skin cancer recognition based on a single, relatively light deep learning model, which we also used in a mobile application.

Keywords: skin cancer; melanoma; convolutional neural networks; deep learning; smartphone application



Citation: Kousis, I.; Perikos, I.; Hatzilygeroudis, I.; Virvou, M. Deep Learning Methods for Accurate Skin Cancer Recognition and Mobile Application. *Electronics* **2022**, *11*, 1294. <https://doi.org/10.3390/electronics11091294>

Academic Editors: Juan M. Corchado, Byung-Gyu Kim, Carlos A. Iglesias, In Lee, Fuji Ren and Rashid Mehmood

Received: 22 March 2022

Accepted: 17 April 2022

Published: 19 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Skin cancer constitutes one of the most common types of cancer, and according to the World Health Organization, it accounts for one in three diagnosed cancer types [1]. Skin cancer reports have constantly increasing incidence rates, thus making the development of more efficient and accurate diagnostic methods a true necessity. One in five people will develop skin cancer before the age of 70 [2]. More than 3.5 million new cases occur annually in the USA, and that number continues to rise [3]. The most dangerous type of skin cancer is melanoma, a type that causes 75% of skin cancer deaths [4]. It is more common in people with a sunburn history, fair skin and unnecessary exposure to UV light, and people who use tanning beds [5]. The rates of melanoma occurrence and corresponding mortality are expected to rise over the next decades [6].

The crucial point for treating skin cancer is the early and accurate detection of it. For example, if melanoma is not diagnosed in the early stages, it starts to grow and continues to spread throughout the outer skin layer, finally penetrating the deep layers, where it connects with the blood and the lymph vessels. That is why it is very important to diagnose it in the early stages, when the mortality rate is very low and a successful treatment is

possible. The estimated 5 year survival rate of diagnosed patients ranges from 15% if detected at its latest stage, to over 97% if detected at its earliest stages [7].

Skin cancer diagnosis is a difficult process, and even experienced specialist dermatologists had a success rate of only 60% until the invention of dermoscopic images, which increased success to between 75% and 84% [8]. The difficulty lies in the fact that malignant lesions are often very similar to the benign moles, and both have small diameter which does not allow for clear images with normal cameras. For example, melanoma and nevus are both melanotic types, and for that reason, the classification difficulty between them is even bigger [9]. Furthermore, most people do not regularly visit their dermatologist, and so they end up with a fatal, late diagnosis.

Thus, there is need, for those cases, to provide an easy alternative solution. The most ubiquitous digital technology typically available for individuals is that of smartphones. The imaging capability of a smartphone could be a natural way for dermatologists, general practitioners and patients to exchange information about skin lesion changes that might be worrisome [10].

Most of the existing state-of-the-art efforts use either hybrid models [11,12] or ensembles of deep learning classifiers [13–15], which are quite heavy to be used in a mobile application. To be able to build an effective mobile application, it is required to find a deep learning model that achieves state-of-the-art performance and is relatively light. Thus, the main objective of this paper is to find a single, relatively light deep learning model, which, combined with appropriate image processing methods, achieves a state-of-the-art performance. To that end, we investigate 11 CNN (convolutional neural networks) single architectures configurations to compare their ability to correctly classify skin lesions. The best model in terms of performance, required memory size and number of parameters to be specified is chosen for our mobile application to be able to classify skin lesions using a common smartphone.

The reminder of the article is structured as follows: Section 2 presents background knowledge on skin cancer types. Section 3 presents an extensive number of the most recent, related works on methods and approaches to skin cancer detection. Section 4 presents the CNN architectures and the datasets, while Section 5 presents the conducted experiments, the collected results and a discussion on them, including a comparison with existing best efforts. The development of the android application for the recognition of skin cancer is presented in Section 6. Finally, Section 7 concludes the paper and provides directions for future work.

2. Medical Knowledge

There are various types of skin cancer. In this section, we present the most significant of them, giving some medical information.

Melanoma (MEL): The most dangerous form of skin cancer, it develops when non-amendable DNA damage occurs to skin cells and causes mutations that lead the skin cells to proliferate rapidly and form malignant tumors. The cause of this phenomenon is usually exposure to ultraviolet (UV) radiation or artificial tanning devices. If it spreads (gives metastases) to the lymphatic system or internal organs, it is fatal at percentages of 38% and 86%, respectively. However, if it is diagnosed and treated at an early stage, the mortality rate is only 0.2% over the next 5 years. The difficulty of diagnosis lies in the fact that melanoma, at its early stages, is similar to other benign skin lesions, and yet, it often develops through them.

Basal Cell Carcinoma (BCC): It is the most common form of skin cancer, and at the same time, the most common form of cancer of all cancers. This type of cancer occurs in the basal cells, which are found in the deeper layers of the epidermis (the surface layer of skin). Almost all basal cell carcinomas occur on parts of the body that have been extensively exposed to sun—especially on the face, ears, neck, head, shoulders and back. In rare cases, however, tumors also develop in non-exposed areas. Further, contact with arsenic, exposure to radioactivity, open

wounds that do not heal, chronic inflammatory skin conditions and complications from burns, scars, infections, vaccines or even from tattoos are compounding factors.

Actinic Keratosis and Intraepithelial Carcinoma (AKIEC): Approximately 450,000 new cases of squamous cell carcinoma (SCC), the main type of AKIEC, are diagnosed annually. This makes this form of cancer the second most common skin cancer (after BCC). This form of cancer is found in the acanthocytes, which constitute the epidermis. SCC can occur at any area of the skin, including the mucous membranes of the mouth and genitals. Of course, it is more often observed in places that are exposed to the sun, like the ears, lower lip, face, bald part of a skull, neck, hands, arms and legs. Often, the skin at those points appears visually as if a sun damage has occurred, displaying wrinkles, changes in color and loss of elasticity.

Melanocytic nevus (NV): A nevus is the reference point (colored) characterized by the accumulation of melanocytes in different layers of the skin. They occur in the embryonic period of life, childhood and adolescence and less often in adults and the elderly. The epidermal nevus is a brown or black smooth spot and is rarely prone to enlargement and malignancy.

Benign Keratosis Lesions (BKL): This kind of benign tumor is the most common and mainly appears in middle-aged and elderly patients as seborrheic hyperkeratosis. Seborrheic hyperkeratosis is caused by the addition of keratin to the stratum corneum. In some cases, it can develop rapidly and sometimes resembles skin cancer.

Vascular lesion (VASC): Skin vascular lesions are due to the expansion of a small group of blood vessels located just below the skin's surface. They are often created on the face and feet. They are not lethal, but they can cause severe leg pain after prolonged standing, and sometimes indicate more serious venous conditions.

Dermatofibroma (DF): A common benign fibrotic skin lesion, it is caused by the non-cancerous growth of the tissue cells of the skin. It is generally a single round or oval, brownish, or sometimes yellowish, nodule of 0.5 to 1 cm in diameter.

In Table 1, the characteristics of various lesions are presented.

Table 1. Characteristics of esions.

Lesion	Characteristics			
	<i>Non-Melanocyte Lesions</i>	<i>Melanocyte Lesions</i>	<i>Benign</i>	<i>Malignant</i>
MEL		X		X
BCC	X			X
AKIEC	X			X
NV		X	X	
BKL	X		X	
DF	X		X	
VASC	X		X	

3. Related Work

There many works related to the diagnosis of skin cancer. We refer here to more recent and important works.

Afonso Menegolay et al. [16] tried to classify lesions as melanoma vs. benign, malignant vs. benign and carcinoma vs. melanoma vs. benign, and they showed that the last one gives the best results. They also proved that fine-tuning improves classification, but on the other hand, transfer learning should be implemented with care concerning the dataset they used. They used the Atlas and ISBI Challenge 2016 datasets for the lesions and ImageNet and Retinopathy datasets for fine-tuning and transfer learning. They also argued that DNNs (deep neural networks) perform better when based on VGG-M and VGG-16 architectures.

Esteva et al. [17] use a GoogleNet Inception v3 CNN architecture that was pretrained on approximately 1.28 million images from the 2014 ImageNet Large Scale Visual Recognition Challenge. They trained it end-to-end on their dataset using transfer learning. Their dataset consisted of 129,450 dermatologist-labelled clinical images, including 3374 dermoscopic images. They split it into 127,463 training and validation images and 1942 biopsy-labelled test images. Nine-fold cross-validation was used. Diseases were organized in a tree structure. For the three-class case (benign lesions, malignant lesions and non-neoplastic lesions), which concerned the first-level nodes of the tree, CNN achieved $72.1 \pm 0.9\%$ (mean $\pm$ sd) overall accuracy, which was better compared to that of two dermatologists (65.56% and 66.0%, respectively). For the nine-class case (dermal, epidermal, melanocytic-melanoma, epidermal, dermal, cutaneous lymphoma-genodermatosis and inflammatory), which are the second-level nodes, it achieved $55.4 \pm 1.7\%$ accuracy, again better than that of the same two dermatologists (53.3% and 55.9%, respectively).

Li and Shen [18] developed a lesion index calculation unit (LICU) for the lesion segmentation and classification task of ISIC 2017 Challenge, which contains 2000 skin lesion images, 374 melanoma cases, 254 seborrheic keratosis and 1372 nevus. They used the ISIC 2017 dataset, and they achieved an accuracy of 0.912. They pre-processed the data using rescaling in the high-resolution images after cropping the image in the center of the lesion, and they applied data augmentation to increase the number of the images by flipping them on the x or y axis. They additionally proposed the use of weighted softmax loss function and batch normalization.

The Recod Titans Group [11] created ensemble deep learning models to classify the skin lesion images. They used open datasets from the ISIC 2018 Challenge, the official challenge dataset with 10,015 dermoscopic images; ISIC Archive1, with over 13,000 dermoscopic images; Interactive Atlas of Extreme Learning, with 1000+ clinical cases, each with dermoscopic and close-up clinical images; Dermofit Image Library, with 1300 images; and PH2 Dataset, with 200 dermoscopic images. In addition, they used Dermatology Atlas, Derm101 and DermIS to gain an additional 631 images and reach the number of 30,726 images for their final dataset. Further, they increased that number with the use of synthetic skin lesions using GANs (generative adversarial networks) [19] and other data augmentation techniques, such as traditional color and geometric transforms, and more unusual augmentations, such as elastic transforms, random erasing and a novel augmentation, that mix two different lesions. They trained the hybrid models of Inceptionv4, ResNet-152, and DenseNet-161, which were all pretrained on the ImageNet dataset. Finally, they proposed three hybrid models: (1) the XGBoost ensemble of 43 deep learning models, (2) an average of eight best deep learning models (on the holdout set), augmented with synthetic images, and (3) an average of 15 deep learning models. Their final results were 0.732, 0.725 and 0.803, respectively, for the normalized multi-class accuracy.

The authors in [20] used a CNN (Microsoft ResNet-152 model) for classifying clinical skin images in 1 of 12 skin diseases: basal cell carcinoma, squamous cell carcinoma, intraepithelial carcinoma, actinic keratosis, seborrheic keratosis, malignant melanoma, melanocytic nevus, lentigo, pyogenic granuloma, hemangioma, dermatofibroma and warts. The CNN was trained using the training sets of the Asan and MED-NODE datasets, as well as the Atlas site images (19,398 images in total). It was tested via test sets of the Asan and Edinburgh datasets. The average values for AUC (area under curve), sensitivity and specificity were: 0.91 ± 0.01 , 86.4 ± 3.5 and 85.5 ± 3.2 (Asian test set), and 0.89 ± 0.01 , 85.1 ± 2.2 and 81.3 ± 2.9 (Edinburgh test set). Especially for malignant melanoma, they were: 0.96 ± 0.00 , 91.0 ± 4.3 and 90.4 ± 4.5 (Asian test set), and 0.88 ± 0.01 , 85.5 ± 2.3 and 80.7 ± 1.1 (Edinburgh test set). A test with 16 dermatologists showed that the results of the system were comparable to those of the dermatologists. The creation of a user-friendly, PC- and smartphone-based automatic skin disease recognition platform is also reported.

Harangi [13] investigated whether an ensemble of CNNs could increase the individual deep classifiers' performance to compensate for the shortage of large numbers of labelled melanoma images. Four CNNs of different type were ensembled: GoogLeNet, AlexNet,

ResNet and VGGNet. Only 2000 images were used for training, 150 for validation and 600 for testing, as related to three types of skin cancer (nevus, melanoma and seborrheic keratosis). Using image transformations, the images increased to 14,300. A weighted voting scheme was used for aggregating the results of individual CNNs. The ensemble achieved an average accuracy of 0.866, an average sensitivity of 0.556 and an average specificity of 0.785.

After RGB skin cancer images pre-processing, pre-trained AlexNet was used for feature extraction, and afterwards, a combined ECOC SVM algorithm was used for cancer classification in [21]. A total of 3753 images collected from the internet were used (2985 as the training set and 768 as the test set) concerning the four classes of carcinomas (actinic keratoses, basal cell carcinoma, squamous cell carcinoma and melanoma). The system achieved accuracy equal to 94.2%, sensitivity equal to 97.83 and specificity equal to 90.74 for melanoma and almost the same for the rest of the classes of carcinomas.

The authors in [22] used an attention residual learning convolutional neural network (ARL-CNN) model. It was composed of multiple ARL (attention residual learning) blocks, a global average pooling (GAP) layer and a classification layer. The ISIC-skin 2017 dataset, consisting of 2000 training, 150 validation, and 600 test images, was used for training and evaluation, enhanced with 1320 additional dermoscopic images from the ISIC archive related to melanoma, nevus and seborrheic keratosis. The model achieved an accuracy of 0.850, a sensitivity of 0.658 and specificity of 0.896 for melanoma classification, and 0.868, 0.878 and 0.867, respectively, for seborrheic keratosis classification.

Khan et al. [9] proposed an intelligent system to detect and distinguish between melanoma and nevus cases. A Gaussian filter was used for removing noise from the skin lesion and an improved k-means clustering was utilized to segment out the lesion. Textual and color features were afterwards extracted. A support vector machine (SVM) was used for the classification of skin cancer into melanoma and nevus. The DERMIS dataset, having a total number of 397 skin cancer images (146 of melanoma and 251 of nevus type), was used. The system achieved 96% accuracy, 97% sensitivity, 96% specificity and 97% precision.

Sarkar et al. [23] used what they called a deep, depthwise separable residual convolutional algorithm for binary melanoma classification. Dermoscopic skin lesion images were pre-processed by noise removal using a non-local means filter, followed by enhancement using contrast-limited adaptive histogram equalization over a discrete wavelet transform algorithm. Images were fed to the model as multi-channel image matrices, with channels chosen across multiple color spaces based on their ability to optimize the performance of the model. The proposed model achieved an accuracy of 99.50% on the international skin imaging collaboration (ISIC), 96.77% on PH2, 94.44% on DermIS and 95.23% on MED-NODE datasets.

The approach in [14] dealt with the problem of skin lesion classification in eight classes: melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AK), benign keratosis (BKL), dermatofibroma (DF), vascular lesion (VASC) and squamous cell carcinoma (SCC). The data set used was the ISIC 2019. Data augmentation was also used. They set up an ensemble of CNN architectures. The optimal ensemble comprised six EfficientNet configurations, two ResNext variants and an SENet. They achieved accuracy of 75.2% using five-fold cross-validation. The system ranked first in the ISIC 2019 Skin Lesion Classification Challenge (with a balanced accuracy of 63.6% for task 1 and 63.4% for task on official test set).

The approach in [24] dealt with the problem of skin disease classification in six classes: seborrheic keratosis (SK), actinic keratosis (AK), rosacea (ROS), lupus erythematosus (LE), basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), based on clinical images. They compared five CNN architectures: ResNet-50, Inception-v3, DenseNet121, Xception and Inception-ResNet-v2, all in two modes. In the first, the five models were trained using only clinical facial images, whereas in the second, they used transfer learning: they pretrained the models with the data of other body parts and used the parameters from the pretrained model as the initial parameters for the new model. They constructed a data set containing 4394 images (2656 of which were facial images) from a large Chinese database

of clinical images. The best result (an average precision of 70.8% and average recall of 77%) was achieved by the Inception-ResNet-v2 model.

Ameri [25] built a CNN based on a pre-trained AlexNet model that was modified for binary classification using transfer learning. All images were resized to 227×227 to match the input size of the AlexNet model. The model was trained using the SGDM algorithm, with an initial learning rate of 0.0001 and a mini-batch size of 30 and 40 epochs. A total of 3400 images (1700 benign and 1700 malignant) were used from the HAM10000 dataset, and 90% were randomly selected as a train set and the rest as a test set. A classification accuracy of 84%, a sensitivity of 81% and a specificity of 88% were achieved.

The authors in [26] investigated the use of Faster R-CNN and MobileNet v2 for skin cancer detection on mobile devices, such as smartphones. They trained and tested them using a perfectly balanced dataset of 600 images concerning two skin diseases: actinic keratosis and melanoma. The images were obtained from the ISIC archive. For training, 90% were used, while 10% were used for testing, and they were enhanced with 40 extra images taken via a smartphone. Experiments showed the superiority of Faster R-CNN both in accuracy and training time. The best accuracy result for Faster R-CNN was 87.2% (on notebook) and 86.3% (on smartphone), but for different learning rates.

Mporas et al. [27] extracted lesion features after image preprocessing and segmentation based on different color models (RGB, HSV, YIQ, etc.). They use the HAM10000 dataset to train several classifiers (SVM, NN, Bagging, AdaBoost, REP, RF, J48 and IBk). Their experiments showed better results for classifiers trained with the RGB-model-produced dataset. The winner was AdaBoost (RF), with an accuracy of 73.08%. They improved this result by combining the features of different color models. The combination of RGB plus HSV plus YIQ features and the AdaBoost (RF) achieved an accuracy of 74.26%.

In [28], the authors built a CNN for skin cancer diagnosis. The CNN consists, excepting convolution and pooling layers, of three hidden layers, using 3×3 filter sizes with 16, 32, and 64 channel outputs in sequence; a fully connected layer; and softmax activation. A total of 3000 training images and 1000 validation images concerning four carcinoma classes (dermatofibroma, nevus pigmentosus, squamous cell carcinoma and melanoma) were used from the ISIC database. Using the Adam optimizer, they achieved a 99% accuracy and an F1-score of almost 1.

The authors in [29] proposed a combination of CNN and a one-versus-all (OVA) approach to classify skin diseases. The HAM10000 dataset was used, enhanced by data augmentation techniques (e.g., rotating, scaling, translation, adding noise, etc.). No feature extraction was used. The model was trained via raw images following an end-to-end approach. The model achieved an accuracy of 92.90%, which was much better than that achieved by the CNN alone (77%).

Huang et al. [30] built two deep learning models and used two datasets for skin disease classification. First, they built a DenseNet network for the binary (malignant or benign) classification problem based on the Kaohsiung Chang Gung Memorial Hospital (KCGMH) dataset (1287 clinical images certified by biopsy) concerning five classes: BCC, SCC, MEL, benign keratosis and NV. It achieved an accuracy of 89.5%, a sensitivity of 90.8% and a specificity of 88.1%. The second model was an EfficientNet model used for multi-class skin disease classification on two datasets, the KCGMH and the HAM10000 (10,015 dermoscopic images and seven classes). The EfficientNet B4 version achieved the best overall accuracy results: 72.1% and 85.8%, respectively.

The authors in [31] used a mixture of features: handcrafted features based on shape, color and texture (after image preprocessing), and deep learning features obtained using a selected CNN architecture (VGG, Mobilenet, ResNet, DenseNet, Inception V3 and Xception), which was pre-trained in an Imagenet classification task. Then, a selected classifier (one of LR, SVM or RVM) is trained as a binary classifier (malignant or benign) on the ISIC 2018 dataset (comprised of both malignant and other benign skin lesion images), augmented by the SMOTE method, to eliminate data imbalance. The system achieved an

accuracy of 89.71% (on the test set), sensitivity of 86.41%, specificity of 90% and precision of 92.08%.

The proposed system in [12] worked in two stages. In the first stage, an end-to-end resolution-preserving deep model for segmentation, namely FrCN, was used. In the second stage, the segmented lesions were passed into a deep learning classifier (Inception-v3, ResNet-50, Inception-ResNet-v2 and DenseNet-201). Classifiers were trained and evaluated on three dermoscopic images datasets: ISIC 2016 (two output classes), ISIC 2017 (three output classes) and ISIC 2018 (seven output classes), all augmented (using rotation and flipping). The best results for the first dataset were achieved by the Inception-ResNet-v2 deep classifier (accuracy of 81.79%, sensitivity of 81.80%, specificity of 71.40% and F1-score of 82.59%). ResNet-50 was the winner of the second (accuracy of 81.57%, sensitivity of 75.33%, specificity of 80.62% and F1-score of 75.75%) and third datasets (accuracy of 89.28%, sensitivity of 81.00%, specificity of 87.16% and F1-score of 81.28%).

The authors in [32] used Deep Learning Studio (DLS) to build four different pre-trained CNN models that were based on ResNet, SqueezeNet, DenseNet and Inception v3, respectively. They trained and tested them using the HAM10000 dataset divided into training (80%), validation (10%) and test (10%) sets. The best model (Inception v3) achieved a 95.74% F1-score for the two-class prediction case.

Salian et al. [33] configured their own CNN and trained it end-to-end using two datasets of dermoscopic images, PH2 and HAM10000. They used two versions of the datasets, the original and another one produced after augmentation. Results showed that augmentation was not beneficial in all cases. Their system achieved the best accuracy values on the version with augmentation for the PH2 dataset (97.25%) and on the version without augmentation for the HAM10000 dataset (83.152%). Compared to other state-of-the-art systems (e.g., Mobilenet and VGG-16), it was much better than VGG-16 and a bit better, on average, than Mobilenet.

The authors in [15] investigated the effect of image size on skin lesion classification based on pre-trained CNNs and transfer learning. Images from the ISIC datasets were either resized to or cropped to six different sizes. Three pre-trained CNNs, namely EfficientNetB0, EfficientNetB1 and SeReNeXt-50, were used in the experiments in single and ensemble mode (called a MSM-CNN: multi-scale multi-CNN). The results showed that image cropping was a better strategy than image resizing. The best accuracy (86.2%) was achieved on the ISIC 2018 dataset by the MSM-CNN configuration.

A hair detection methodology was followed in [34] for image pre-processing using color enhancement, hair lines detection and thresholding for lesion segmentation. Afterwards, features were extracted (texture descriptors, color, convexity, circularity and irregularity index). They chose 640 images from the ISIC database and used 512 of them for training and the rest for testing. They used an ensemble voting classifier consisting of a KNN (k-nearest neighbor), an SVM (support vector machine) and a CNN classifier. The CNN took as inputs whole images, sized 124×124 pixels, and was composed of nine layers (three convolution layers with a filter size of 3×3 with a ReLU activation function, proceeded by three spatial max pooling layers, which down-sampled the feature maps). It achieved an accuracy of 88.4%, which was better than the accuracy of each individual classifier.

The authors in [35] built an architecture combining the MobileNet and LSTM architectures for skin disease classification. They used the HAM10000 dataset (seven classes) for training and evaluation. Compared to other state-of-the-art approaches, it achieved better results (accuracy of 85.34%, sensitivity of 88.24% and specificity of 92.00%). The proposed model also exhibited reduced execution time compared to other models. The authors implemented a mobile application, based on the MobileNet-LSTM model, called esmart-Health.

Acosta et al. [36] proposed a two-stage classification process: (a) create a cropped bounding box around a skin lesion, and (b) classify the cropped area using a CNN. In the first stage, Mask R_CNN was used, whereas ResNet152 was used in the second. They used 1995 dermatoscopic images for training, 149 for validation and 598 for testing, from the

ISIC 21i challenge database. There were two system output classes: benign and malignant (meaning “melanoma”). Due to the small number of training examples, they used transfer learning. They produced five different models after training. The best model, called eVida, outperformed all participants in the 2017 ISBI Challenge, achieving an average accuracy of 0.904, sensitivity of 0.820 and specificity of 0.925.

Wang and Hamian [37] used an extreme learning model (ELM) neural net model optimized by the thermal exchange optimization (TEO) algorithm. It was used to classify skin images represented by their features. Features were extracted after dermoscopy image pre-processing for noise reduction, contrast enhancement and segmentation (Otsu algorithm). The extracted features were geometric, texture and statistical. They used the SIIM-ISIC melanoma dataset with image size of 1024×1024 pixels. Compared with other methods (i.e., CNN, Genetic Algorithm, SVM, etc.), their approach achieved the best results: an accuracy of 92.65%, sensitivity of 91.18% and specificity of 89.70%.

In our approach, the objective was to find a single CNN-based model and suitable image processing methods to achieve state-of-the-art results. We chose the HAM10000 dataset, which includes seven unbalanced classes of skin cancer. We employed classical data augmentation, transfer learning and fine-tuning to improve the performance of our models. We compared 11 deep learning network models. DenseNet169 was the winner model, achieving an accuracy of 92.25% and a recall (sensitivity) of 93.59%, which outperformed deeper models. We have achieved with a single standard, deep learning, light model better results than other attempts that use hybrid or ensemble deep learning models. We also implemented a mobile application for skin cancer prevention and two-class (benign-malignant) detection.

4. Materials and Methods

4.1. Deep Learning Models

The main part of this work is the creation of the appropriate diagnostic model for skin lesions. Based on past work in this field and related issues, we decided to approach the problem with deep learning techniques, specifically with a convolutional neural network (CNN) [38]. Some of the most popular CNNs are briefly presented below and some of their characteristics in Table 2, where ‘size’ of a model represents the required memory *s* in Mbytes (MB), number of ‘parameters’ is expressed in millions (M) and ‘depth’ represents the number of levels of the model.

Table 2. Basic deep learning models and their characteristics.

Model Name	Size (MB)	Parameters	Depth
CNN	90	28.6 M	4
AlexNet	233	61 M	8
VGG16	528	138.4 M	16
VGG19	549	143.7 M	19
ResNet50	98	25.6 M	107
InceptionV3	92	23.9 M	189
MobileNetV2	14	3.5 M	105
InceptionResNetV2	215	55.9 M	449
DenseNet121	33	8.1 M	242
DenseNet169	57	14.3 M	338
DenseNet201	80	20.2 M	402

AlexNet [7] was developed in 2012 by the SuperVision group, which consisted of Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton. They created an assembly neural network which participated in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

competition. They managed to achieve a top 5 error rate (15.3%), which was 10.8 units smaller than the next in the ranking. The architecture of this network is characterized by the much larger number of levels it has in relation to earlier models. The input to this network is an image of 256×256 pixels and consists of five assembly levels and three fully connected levels. It uses the ReLU activation function after each assembly and fully connected layer.

The Visual Geometry Group (VGG) [39] presented their model showing that the depth of the network plays a decisive role in the accuracy of CNN. The model consists of two assembly levels that use the ReLU activation function. The following is a max pooling level with the ReLU activation function again. The last level of the model is a softmax level for categorization. Three models of VGG-E were proposed: VGG-11, VGG-16 and VGG-19, having 11, 16 and 19 levels, respectively. All models have three fully connected levels at the end, having the remaining 8, 13 and 16, respectively, as convolution levels.

ResNet [40] was created by He et al. as a very deep network. It succeeds in summing at the end of two levels the information they produce with the information given to the first as the input. The model continues to pass the information to the deeper levels even if some levels have zero output after ReLU (not active). ResNet was created with several different level numbers (34, 50, 101, 152 and even 1202). The most popular of those, ResNet50, consists of 49 assembly levels and a fully connected level at the end of the network. ReLU is used as an activation function.

InceptionV3 [41] is an improvement to InceptionV1 or GoogLeNet, presented by Google's Christian Szegedy, in order to reduce the computational complexity of previous models. The characteristic of GoogLeNet is that it has filters of many sizes at the same level. The network grew in width, but not in depth, with the advantage that it is not easily over-trained. InceptionV3 had extra batch normalization in axillary categorizers, a 7×7 level convolution and RMSProp optimizer.

The MobileNet [42] model is a specialized model for mobile devices and embedded systems. The architecture proposed by Google was designed to use limited computing resources. For example, compared to VGG-16, which has a size of 553 MB and 138 million parameters, MobileNet has a size of 17MB and 4.2 million parameters. Here, however, we must note that in order to reduce the requirements of the system that much, the accuracy of the model is reduced.

InceptionResNetV2 [43], introduced for the first time in 2016, is an evolution of InceptionV3 inspired by some elements of Microsoft's ResNet. This allowed the new model to have more levels than InceptionV3.

DenseNet [44] was proposed in 2017 by Gao Huang et al. The key characteristic of this model is that the output of each layer connects to all successive levels that follow in the dense block. This results in a better transfer to the classification layer of the low-level features that are exported to the first levels, compared to other models. In addition, dense connectivity between layers results in the reuse of features, which dramatically reduces network configuration. It performs three basic functions: batch normalization, the ReLU activation function, and, finally, a 3×3 level convolution.

Further, some of the most widely used methods in image recognition and classification are:

- Data Augmentation: existing augmentation methods in the task of image classification can be put into one of two very general categories: classical image transformations, such as rotating, cropping, zooming or histogram-based methods, and black-box methods based on deep neural networks, such as style transfer and generative adversarial networks [7,45,46].
- Transfer learning is a machine learning method used to solve the basic problem of insufficient training data. It tries to transfer the knowledge from the source domain to the target domain. This has positive effect on many tasks that are difficult to improve because of limited data. During transfer learning, the last few layers of the trained network are removed and retrained [47,48].

- Fine-tuning has the same concept as transfer learning and differs only in the fact that we do not only retrain the last few layers. We can retrain the whole model using a small learning rate [16].

4.2. Datasets

We used the Ham10000 [49] database, created in 2018 by the International Skin Image Collaboration, which consists of 10,015 images. On this basis, we found seven different classes of data, each belonging to one of the skin lesions we want to diagnose. This base was established by two separate sources: one of them belongs to Cliff Rosendahl and the Queensland University Medical School of Australia, and the other to the ViDIR Group of the Dermatology School at the University of Vienna in Austria.

The first problem that was observed is that of the imbalance between the number of images in different dataset classes. For example, the largest class consists of 6705 images, i.e., about 67%. In contrast, the smallest class consists of 115 images, or 1.1% of the base. Figure 1 depicts the percentages of the classes.

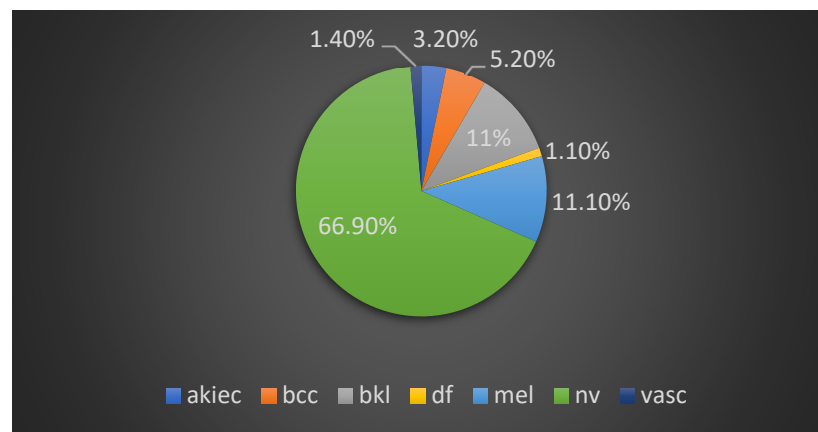


Figure 1. Distribution of images in HAM10000.

As a starting goal, we set up the creation of a model that achieves high rates in the classification average accuracy of the seven different classes and achieves melanoma detection recall [50] similar to that of specialized dermatologists at a rate of 75–84%. At the same time, we had to consider the limited resources of an average mobile phone (RAM and computing power). The first step we made was to split the dataset into training and validation (test) sets with a ratio of 4:1, or with 80% for training data and the remaining 20% for the verification process (see Figure 2).

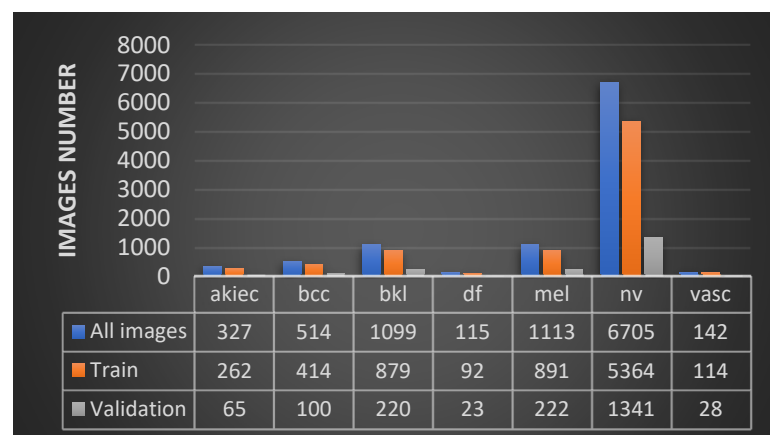


Figure 2. Distribution of images in the HAM10000 dataset of the training–validation split sets.

5. Results and Discussion

We used the TensorFlow 1.4 framework in combination with Keras API version 2.0.8 and python 3.5 to implement and run our models on a Linux system with a GTX 1060 6GB graphics card. To increase the performance of our models, we used the image techniques mentioned above. First, we used data augmentation with image random crops, rotations, zoom and horizontal and vertical flips. In addition, we employed transfer learning from the ImageNet [51] dataset while retraining the last layers, and after that, fine-tuning of the model, retraining the whole model with a smaller training rate. We also tried to train the models from the start without transfer learning, but the results were poorer, especially in the deeper models. Furthermore, we changed the color space from RGB to HSV and grayscale, but even then, the results were worse, especially for the tests with the grayscale images.

In Table 3, we present the final results based on the average (weighted) accuracy [51] of the 11 models that were trained with data augmentation, transfer learning, fine-tuning and the SGD optimizer on the original RGB images. Because of the highly unbalanced data, we used the appropriate class weights during the training process to treat the classes equally.

Table 3. Average values of metrics for the tested models.

Model Name	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	82.15	85.73	88.53	87.11
AlexNet	80.74	86.25	88.14	87.19
VGG16	90.15	91.71	92.82	92.26
VGG19	85.55	88.44	89.34	88.89
ResNet50	90.36	86.49	84.04	85.25
InceptionV3	91.13	91.34	92.56	91.95
MobileNetV2	80.69	84.98	86.65	85.81
InceptionResNetV2	90.29	91.77	92.87	92.32
DenseNet121	91.20	92.12	93.08	92.60
DenseNet169	92.25	92.95	93.59	93.27
DenseNet201	90.14	91.20	92.85	92.02

Finally, the model we chose was DenseNet169, which achieved the best values for all metrics and requires relatively less memory size and specification of relatively less parameters than other non-DenseNet candidates, e.g., InceptionV3, ResNet50 and VGG16 (see Table 2). The confusion matrix of DenseNet169 is depicted in Figure 3.

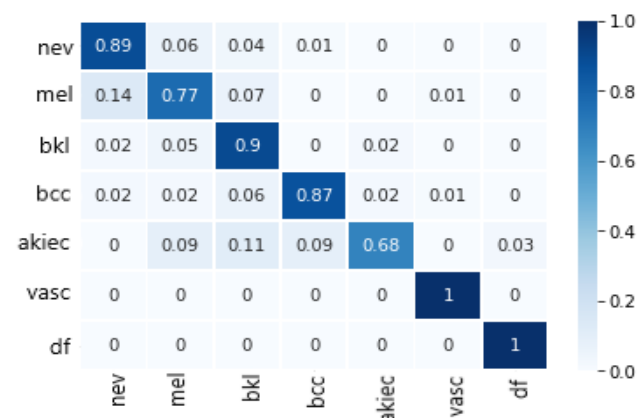


Figure 3. Confusion matrix for DenseNet169.

In order to face up with the image quality problem, we tried to insert each image during the validation phase more than one time and make a diagnosis based on their

median average. More specifically, we inserted each validation image four times into the model by making a different flip each time, and then we calculated the median average of the results in order to classify the image. This technique was used in five randomly selected models during the last 50 training epochs of DenseNet169. The results showed that there was an improvement of +1.004% on average compared to the original image case (see Table 4). From Table 4, it is also obvious that vertical flip was the main contributor to the improvement.

Table 4. Average accuracy for flip experiments.

Experiment	Flip Experiments Average Accuracy				
	Original Image	Horizontal Flip	Vertical Flip	Horizontal and Vertical Flip	Combined Result
Exp. 1	78.43%	78.37%	79.99%	80.37%	80.18%
Exp. 2	83.79%	82.92%	82.86%	80.98%	83.91%
Exp. 3	84.71%	83.48%	85.58%	83.40%	86.19%
Exp. 4	80.85%	79.84%	83.00%	79.89%	81.33%
Exp. 5	84.67%	84.11%	85.87%	84.59%	85.86%
Average	82.49%	81.74%	83.46%	81.46%	83.494

For the application, given its qualitative orientation and its use as a first diagnostic hint, we decided to use a two-class mapping of the DenseNet169 model to distinguish between benign (nv, bkl, vasc or df) and malignant (mel, bcc or akiec) cases. So, the seven-class model was used, but it gives as output to the user “benign” or “malignant”, depending on the predicted class. The metrics of the two-class mapping model are presented in Table 5.

Table 5. Metrics for the two-class DenseNet169 mapping model.

Model	Accuracy	Specificity	Sensitivity
DenseNet169—two classes	91.10	95.67	82.49

From the tested model architectures, DenseNet169 achieved the highest values for all metrics, while DenseNet121 the second highest. In contrast to our expectations, DenseNet201 did not do better than the other, less deep DenseNet models. It seems that the size of the training set was no more adequate for deeper training than that of DenseNet169. From the rest of the models, InceptionResNetV2 and VGG16 did better than the others, being very close between each other. Again, the deeper VGG19 model did not outperform the less deep VGG16 model.

The experimental results show that our model did very well in distinguishing between melanoma and nevus, which is considered a hard case given that both are melanocyte lesions and they visually look alike. Furthermore, we managed to reach a higher accuracy than specialist dermatologists.

In Table 6, a comparison with state-of-the-art approaches is attempted. In the table, we have included models that deal with comparable datasets. We also distinguish them in two groups, the two-class and seven-class groups. As is clear from Table 6, our model is the best in the seven-class category, whereas it is the second best, in terms of achieved accuracy, in the two-class category. However, compared to the first one, it has the advantage of being a single model.

Table 6. Comparison of approaches.

Research Work	Method	Mobile App	Dataset	Metrics (%)
2-class models				
[26] Hartanto and Wibowo (2020)	Faster R-CNN	✓	ISIC (600 images)—2 classes (akiec, mel)	87.2 (acc, laptop)—86.3 (acc, smartphone)
[29] Polat and Koc (2020)	CNNs + OVA (one-versus-all)		Augmented HAM10000—pairs of 2 classes	92.9 (acc)
[31] Almaraz-Damian et al. (2020)	MobileNet v2 + (LR, SVM and RVM) (handcrafted and deep learning features)		ISIC 2018—2 classes (mel, nv)	89.71 (acc)—86.41 (sen)—90.0 (spec)—92.08 (prec)
[32] Kadampur and Riyae (2020)	Inception v3		HAM10000—2 classes	98.19 (prec)—95.74 (F1)
Our model	DenseNet169	✓	HAM10000—2 classes	91.10 (acc)—82.49 (sen)—95.67 (spec)
7-class models				
[12] Al-masni et al. (2020)	FrCN + ResNet50		ISIC 2018—7 classes	89.28 (acc)—81.0 (sen)—87.16 (spec)—81.28 (F1)
[33] Salian et al. (2020)	Own CNN		HAM10000—7 classes	83.15 (acc)
[15] Mahbod et al. (2020)	MSM-CNN (multi-scale multi-CNN)		ISIC 2018—7 classes	86.2 (acc)
[35] Srinivasu et al. (2021)	MobileNet+LSTM	✓	HAM10000—7 classes	85.34 (acc)—88.24 (sen)—92.0 (spec)
Our model	DenseNet169		HAM10000—7 classes	92.25 (acc)—92.95 (prec)—93.59 (rec)—93.27 (F1)

6. Application Development

We developed a mobile application to constitute an integrated system for the prevention and diagnosis of skin lesions (Figure 4). We used Android-Studio έκδοση-3.1.3 and Android 8.0. It was designed to be a useful and easy to use tool for every smartphone user, minimizing the complex processes and making the application environment as simple and easy to understand as possible. We placed the DenseNet169 two-class mapping model inside it, after we transformed it to a TensorFlow Lite model, in order to be optimized for smartphones. The application enables us to take a photo of the skin lesion that we want to examine, manually crop the photo and keep the area that is of interest (ROI) and insert it into the classification model. Other features of the application concern the creation of a user account to provide authentication, as well as the ability to calculate the amount of time that the user can remain exposed to the sun without any burns. Studies show that the possibility of melanoma is doubled every five sunburns. For this calculation, we use the user's skin phototype, the sun UV indicator present in the user's area at the given time and the user's sunscreen index.

To make the photo quality as high as possible, we provided the option of zooming in on the area of skin damage or to open the flashlight of the mobile, if available. Then, we ask the user to manually crop the photo into a square area near the outer perimeter of the skin lesion (Figure 5).

Figure 6 presents a classification result after the use of the application for a specific case.

In addition, to reduce the likelihood of malignant skin damage of users, we developed another function in our application. Since the main factor for melanoma is the irradiated sunlight, and, more specifically, the hours exposed to the sun, according to the degree of ultraviolet radiation that we receive at that time, we inform the user of the time that he/she can safely be exposed to the sun without resulting in sunburns. We do this with the following steps:

1. Collect the user's location via the mobile phone's GPS system.
2. Send the location and the current time to OPENUV API (an open forecast and update platform for solar radiation). This returns the data we need to inform the user about sun UV index in his/her location and the time he/she can stay in the sun for every skin phototype.

3. The user selects his/her skin phototype and the degree of sunscreen (if he/she uses one) to be informed of the time that he/she can be exposed to the sun without having unpleasant side effects for his/her health.

In Figure 7, we show the corresponding snapshot of the application in which we see that in Panama City, the sun UV index is currently 5, and the user with a phototype 1 can stay in the sun for 296 min using a sunscreen with an SPF of 8.

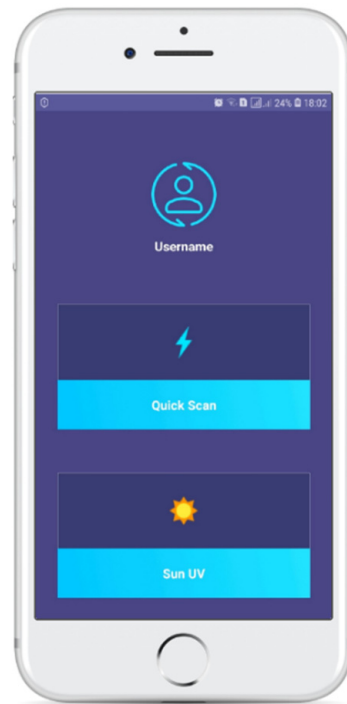


Figure 4. Application screenshot: main menu.

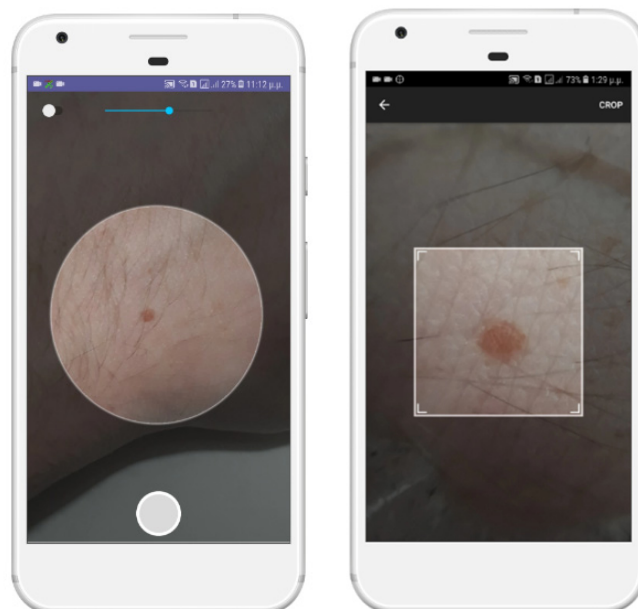


Figure 5. Application screenshots: photo taken on the left and image crop on the right.



Figure 6. Application screenshot: classification result.

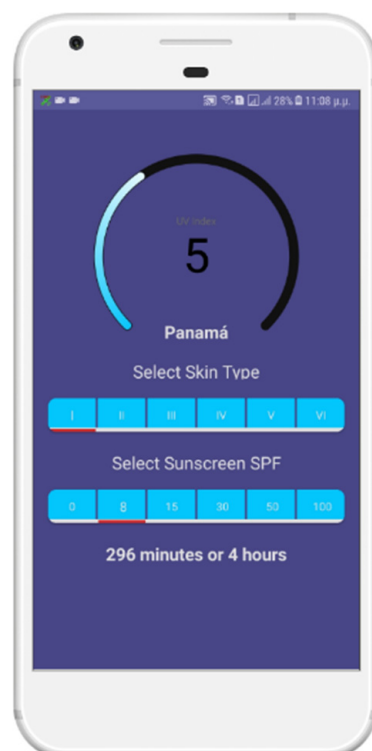


Figure 7. Application screenshots: sun UV and exposure time without burning.

Finally, we tested the app in a real environment, using at first only the phone camera lens. After that, we tried it again with an external 10× macro lens, and the result was a bit better. We also tried it with the external macro lens and a handmade stabilizer, which gave us the best result. The quality of the image is very close to the dataset images. We strongly

believe that this sector of the app is critical for the average accuracy of the model. We can see the results of the images in Figure 8.

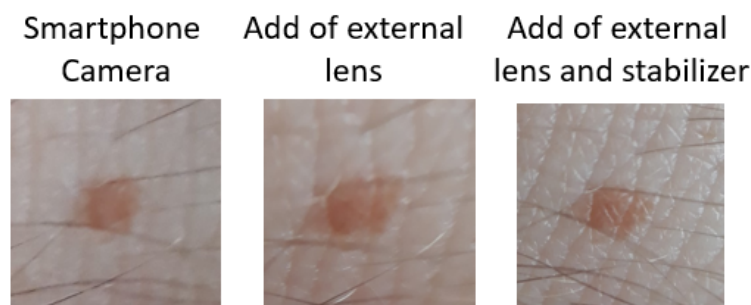


Figure 8. Images from a smartphone camera, smartphone camera with lens, and smartphone camera with lens and the handmade stabilizer.

After that, the image is automatically passed to the model to be categorized. Within a few seconds (about 2–4), depending on the capabilities of the smartphone, the classification result is displayed to the user. The user then has the option to return to the original menu or restart the process from scratch (see Figure 6).

Given the shortage of smartphone-taken skin lesion image databases, we could not test our application in a systematic way; however, small scale tests were very successful.

To use such a system may require the transfer of the image to a server, where the system could be implemented and run more efficiently. Further, to achieve better performance for the data transmission, new technologies for wireless sensor networks [52,53] could be used.

7. Conclusions

In this paper, we configured and tested 11 state-of-the-art deep learning network architectures as a means for skin cancer diagnoses. They were tested on a well-known dataset of dermoscopic images (HAM10000), which concerns seven different types of skin cancer. Our results showed that DenseNet169 performs better than the other architectures in this domain. The average accuracy of the model was 92.25%, which is higher than the accuracy of other state-of-the-art models. This also means that it does better than specialist dermatologists. We also built a two-class DenseNet169 mapping model, which achieves very good results, too (an accuracy of 91.10%).

Based on the above two-class model, we created a mobile application for helping people in having a first indication about their skin lesions. Apart from performing diagnoses, the application informs the user about how much time he/she can remain exposed to the sun without any burning, based on the user's skin phototype, the sun UV indicator present in the user's area at the given time, and the user's sunscreen index. To improve the quality of the photos taken by the mobile device, the system allows zooming and cropping.

The combination of deep learning models (ensembles) for the extraction of the result is a promising research direction for increasing the accuracy of diagnoses. However, to incorporate such a system into the mobile application may create efficiency problems due to its complicated structure and computational demands. To use such a system may require transferring the image to a server where the system could be implemented and run more efficiently.

Another issue for the mobile application is the quality of the image taken via a smartphone, which is crucial for accurate diagnoses, mostly because of the similarity and the relatively small area of the lesions. In that direction, we propose the use of a macro lens and a stabilizer product in front of the smartphone camera. Regarding the software, we could remove some noise from the image and enhance the image with AI super resolution techniques. We also propose the flipping of the taken image to be parallel with more than one photo from the lesion area.

Author Contributions: Conceptualization, I.P., I.H. and I.K.; methodology, I.K., I.P. and I.H.; software, I.K.; validation, I.H. and I.P.; formal analysis, I.K. and I.P.; investigation, I.K.; writing—original draft preparation, I.K. and I.P.; writing—review and editing, I.H. and M.V.; supervision, I.H. and M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. WHO. Available online: [https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-\(uv\)-radiation-and-skin-cancer](https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-(uv)-radiation-and-skin-cancer) (accessed on 31 January 2022).
2. Stern, R.S. Prevalence of a History of Skin Cancer in 2007: Results of an Incidence-Based Model. *Arch. Dermatol.* **2010**, *146*, 279–282. [CrossRef]
3. Li, Z.; Fang, Y.; Chen, H.; Zhang, T.; Yin, X.; Man, J.; Yang, X.; Lu, M. Spatiotemporal trends of the global burden of melanoma in 204 countries and territories from 1990 to 2019: Results from the 2019 global burden of disease study. *Neoplasia* **2022**, *24*, 12–21. [CrossRef]
4. Fornaciali, M.; Carvalho, M.; Vasques, B.F.; Avila, S.; Valle, E. Towards automated melanoma screening: Proper computer vision & reliable results. *arXiv* **2016**, arXiv:1604.04024.
5. Albahar, M.A. Skin Lesion Classification Using Convolutional Neural Network with Novel Regularizer. *IEEE Access* **2019**, *7*, 38306–38313. [CrossRef]
6. Saginala, K.; Barsouk, A.; Aluru, J.S.; Rawla, P.; Barsouk, A. Epidemiology of Melanoma. *Med. Sci.* **2021**, *9*, 63. [CrossRef]
7. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
8. Haenssle, H.A.; Fink, C.; Schneiderbauer, R.; Toberer, F.; Buhl, T.; Blum, A.; Kalloo, A.; Hadj, H.A.B.; Thomas, L.; Enk, A.; et al. Reader study level-I and level-II Groups, Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann. Oncol.* **2018**, *29*, 1836–1842. [CrossRef]
9. Khan, M.Q.; Hussain, A.; Rehman, S.U.; Khan, U.; Maqsood, M.; Mehmood, K.; Khan, M.A. Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer. *IEEE Access* **2019**, *7*, 90132–90144. [CrossRef]
10. MacKinnon, N.; Vasefi, F.; Booth, N.; Farkas, D.L. Melanoma detection using smartphone and multimode hyperspectral imaging. In Proceedings of the SPIE 9711, Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues IX, San Francisco, CA, USA, 6 April 2016; Volume 971117.
11. Bissoto, A.; Fábio, P.; Vinícius, R.; Michel, F.; Avila, S.; Valle, E. Deep-Learning Ensembles for Skin-Lesion Segmentation, Analysis, Classification: RECOD Titans at ISIC Challenge 2018. *arXiv* **2018**, arXiv:1808.08480.
12. Al-masni, M.A.; Kim, D.-H.; Kim, T.-S. Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. *Comput. Methods Prog. Biomed.* **2020**, *190*, 105351. [CrossRef]
13. Harangi, B. Skin lesion classification with ensembles of deep convolutional neural networks. *J. Biomed. Inform.* **2018**, *86*, 25–32. [CrossRef]
14. Gessert, N.; Nielsen, M.; Shaikh, M.; Werner, R.; Schlaef, A. Skin Lesion Classification Using Ensembles of Multi-Resolution Efficient Nets with Meta Data. *arXiv* **2019**, arXiv:1910.03910v1.
15. Mahbod, A.; Schaefer, G.; Wang, C.; Dorffner, G.; Ecker, R.; Ellinger, I. Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. *Comput. Methods Prog. Biomed.* **2020**, *193*, 105475. [CrossRef]
16. Menegola, A.; Fornaciali, M.; Pires, R.; Bittencourt, F.; Avila, S.; Valle, E. Knowledge Transfer for Melanoma Screening with Deep Learning. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21 April 2017. [CrossRef]
17. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [CrossRef]
18. Li, Y.; Shen, L. Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network. *Sensors* **2018**, *18*, 556. [CrossRef]
19. Bissoto, A.; Perez, F.; Valle, E.; Avila, S. Skin Lesion Synthesis with Generative Adversarial Networks. *arXiv* **2019**, arXiv:1902.03253.
20. Han, S.S.; Kim, M.S.; Lim, W.; Park, G.H.; Park, I.; Chang, S.E. Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm. *J. Investig. Dermatol.* **2018**, *138*, 1529–1538. [CrossRef]
21. Dorj, U.-O.; Lee, K.K.; Choi, J.Y.; Lee, M. The skin cancer classification using deep convolutional neural network. *Multimed. Tools Appl.* **2018**, *77*, 9909–9924. [CrossRef]
22. Zhang, J.; Xie, Y.; Xia, Y.; Shen, C. Attention Residual Learning for Skin Lesion Classification. *IEEE Trans. Med. Imaging* **2019**, *38*, 2092–2103. [CrossRef]
23. Sarkar, R.; Chatterjee, C.C.; Hazra, A. Diagnosis of melanoma from dermoscopic images using a deep depthwise separable residual convolutional network. *IET Image Process.* **2019**, *13*, 2130–2142. [CrossRef]
24. Wu, Z.; Zhao, S.; Peng, Y.; He, X.; Zhao, X.; Huang, K.; Wu, X.; Fan, W.; Li, F.; Chen, M.; et al. Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images. *IEEE Access* **2019**, *7*, 66505–66511. [CrossRef]

25. Ameri, A. A Deep Learning Approach to Skin Cancer Detection in Dermoscopy Images. *J. Biomed. Phys. Eng.* **2020**, *10*, 801–806. [\[CrossRef\]](#)
26. Hartanto, C.A.; Wibowo, A. Development of Mobile Skin Cancer Detection using Faster R-CNN and MobileNet V2 Model. In Proceedings of the 2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), Semarang, Indonesia, 24–25 September 2020; pp. 58–63.
27. Mporas, I.; Perikos, I.; Paraskevas, M. Color Models for Skin Lesion Classification from Dermatoscopic Images. In *Advances in Integrations of Intelligent Methods, Smart Innovation, Systems and Technologies*; Hatzilygeroudis, I., Isidoros, P., Foteini, G., Eds.; Springer Nature: Singapore, 2020; Volume 170, pp. 85–98. [\[CrossRef\]](#)
28. Fu'adah, Y.N.; Pratiwi, N.K.C.; Pramudito, M.A.; Ibrahim, N. Convolutional Neural Network (CNN) for Automatic Skin Cancer Classification System. *IOP Conf. Series Mater. Sci. Eng.* **2020**, *982*, 012005. [\[CrossRef\]](#)
29. Polat, K.; Koc, K.O. Detection of Skin Diseases from Dermoscopy Image Using the combination of Convolutional Neural Network and One-versus-All. *J. Artif. Intell. Syst.* **2020**, *2*, 80–97. [\[CrossRef\]](#)
30. Huang, H.W.; Hsu, B.W.-Y.; Lee, C.-H.; Tseng, V.S. Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers. *J. Dermatol.* **2021**, *48*, 310–316. [\[CrossRef\]](#)
31. Almaraz-Damian, J.-A.; Ponomaryov, V.; Sadovnychiy, S.; Castillejos-Fernandez, H. Melanoma and Nevus Skin Lesion Classification Using Handcraft and Deep Learning Feature Fusion via Mutual Information Measures. *Entropy* **2020**, *22*, 484. [\[CrossRef\]](#)
32. Kadampur, M.A.; Al Riyae, S. Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images. *Inform. Med. Unlocked* **2020**, *18*, 100282. [\[CrossRef\]](#)
33. Salian, A.C.; Vaze, S.; Singh, P.; Shaikh, G.N.; Chapaneri, S.; Jayaswal, D. Skin Lesion Classification using Deep Learning Architectures. In Proceedings of the 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA), Mumbai, India, 3–4 April 2020; pp. 168–173.
34. Daghrir, J.; Tlig, L.; Bouchouicha, M.; Sayadi, M. Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In Proceedings of the International Conference on Advanced Technologies for Signal and Image Processing, Sfax, Tunisia, 2–5 September 2020. [\[CrossRef\]](#)
35. Srinivasu, P.N.; SivaSai, J.G.; Ijaz, M.F.; Bhoi, A.K.; Kim, W.; Kang, J.J. Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. *Sensors* **2021**, *21*, 2852. [\[CrossRef\]](#)
36. Acosta, M.F.J.; Tovar, L.Y.C.; Garcia-Zapirain, M.B.; Percybrooks, W.S. Melanoma diagnosis using deep learning techniques on dermoscopic images. *BMC Med. Imaging* **2021**, *21*, 6. [\[CrossRef\]](#)
37. Wang, S.; Hamian, M. Skin Cancer Detection Based on Extreme Learning Machine and a Developed Version of Thermal Exchange Optimization. *Hindawi Omput. Intell. Neurosci.* **2021**, *2021*, 1–13. [\[CrossRef\]](#)
38. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
40. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [\[CrossRef\]](#)
41. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
42. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
43. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), Mountain View, CA, USA, 4–9 February 2017; pp. 4278–4284.
44. Huang, G.; Liu, Z.; Maaten, L.V.D.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269. [\[CrossRef\]](#)
45. Simard, P.Y.; Steinkraus, D.; Platt, J.C. Best practices for convolutional neural networks applied to visual document analysis. In Proceedings of the 7th International Conference on Document Analysis and Recognition, Edinburgh, UK, 3–6 August 2003; Volume 3, pp. 958–962.
46. Mikołajczyk, A.; Grochowski, M. Data augmentation for improving deep learning in image classification problem. In Proceedings of the 2018 International Interdisciplinary PhD Workshop (IIPhDW), Swinoujscie, Poland, 9–12 May 2018; pp. 117–122. [\[CrossRef\]](#)
47. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Volume 2, pp. 3320–3328.
48. Chuanqi, T.; Fuchun, S.; Tao, K.; Wenchang, Z.; Chao, Y.; Chunfang, L. A Survey on Deep Transfer Learning. In Proceedings of the 27th International Conference on Artificial Neural Networks, Rhodes, Greece, 4–7 October 2018. Part III. [\[CrossRef\]](#)
49. Tschandl, P.; Rosendahl, C.; Kittler, H. The HAM10000 Dataset: A Large Collection of Multi-Source Dermatoscopic Images of Common Pigmented Skin Lesions. *Sci. Data* **2018**, *5*, 180161. [\[CrossRef\]](#)

-
50. Sokolova, M.; Guy, L. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]
 51. Deng, J.; Dong, W.; Socher, R.; Li, L.; Li, K.; Li, F.-F. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255. [[CrossRef](#)]
 52. Liu, P.; Wang, X.; Hawbani, A.; Busaileh, O.; Zhao, L.; Al-Dubai, A. FRCA: A Novel Flexible Routing Computing Approach for Wireless Sensor Networks. *IEEE Trans. Mob. Comput.* **2020**, *19*, 2623–2639. [[CrossRef](#)]
 53. Hawbani, A.; Wang, X.; Zhao, L.; Al-Dubai, A.; Min, G.; Busaileh, O. Novel Architecture and Heuristic Algorithms for Software-Defined Wireless Sensor Networks. *IEEE/ACM Trans. Netw.* **2020**, *28*, 2809–2822. [[CrossRef](#)]