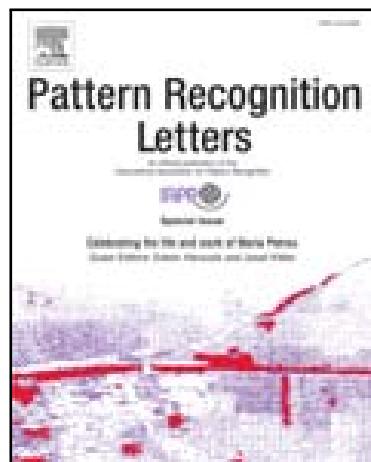


Journal Pre-proof

A new approach for classification skin lesion based on transfer learning, deep learning, and IoT system.

Douglas de A. Rodrigues, Roberto F. Ivo,
Suresh Chandra Satapathy, Shuihua Wang, Jude Hemanth,
Pedro P. Rebouças Filho

PII: S0167-8655(20)30198-7
DOI: <https://doi.org/10.1016/j.patrec.2020.05.019>
Reference: PATREC 7905



To appear in: *Pattern Recognition Letters*

Received date: 2 December 2019
Revised date: 6 May 2020
Accepted date: 17 May 2020

Please cite this article as: Douglas de A. Rodrigues, Roberto F. Ivo, Suresh Chandra Satapathy, Shuihua Wang, Jude Hemanth, Pedro P. Rebouças Filho, A new approach for classification skin lesion based on transfer learning, deep learning, and IoT system., *Pattern Recognition Letters* (2020), doi: <https://doi.org/10.1016/j.patrec.2020.05.019>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier B.V.

Research Highlights

- Twelve extractions and seven classifications for the diagnosis of skin lesions.
- Convolution neural networks as features extractors.
- The use of a consolidated IoT system to perform different experiments.



A new approach for classification skin lesion based on transfer learning, deep learning, and IoT system.

Douglas de A. Rodrigues^{a,b}, Roberto F. Ivo^{a,b}, Suresh Chandra Satapathy^c, Shuihua Wang^d, Jude Hemanth^e, Pedro P. Rebouças Filho^{a,b}

^aPrograma de Pós-Graduação em Engenharia de Teleinformática - Universidade Federal do Ceará, Fortaleza-CE, Brazil

^bLaboratório de Processamento de Imagens, Sinais e Computação Aplicada (LAPISCO), Instituto Federal de Educação, Ciência e Tecnologia do Ceará, Fortaleza-CE, Brazil

^cDepartment of CSE, KIIT Deemed to be University, Odisha, India.

^dSchool of Architecture Building and Civil Engineering, Loughborough University, Loughborough

^eDepartment of Electronics and Communication Engineering, Karunya University, India

ABSTRACT

Melanoma skin cancer is one of the most common diseases in the world. It is essential to diagnose melanoma at an early stage. Visual inspection during the medical examination of skin lesions is not a simple task, as there is a similarity between lesions. Also, medical experience and disposition can result in inaccurate diagnoses. Technologies such as the Internet of Things (IoT) have helped to create effective health systems. Doctors can use them anywhere, with the guarantee that more people can be diagnosed without prejudice to subjective factors. Transfer Learning and Deep Learning are increasingly significant in the clinical diagnosis of different diseases. This work proposes the use of Transfer Learning and Deep Learning in an IoT system to assist doctors in the diagnosis of common skin lesions, typical nevi, and melanoma. This work uses Convolutional Neural Networks (CNNs) as resource extractors. The CNN models used were: Visual Geometry Group (VGG), Inception, Residual Networks (ResNet), Inception-ResNet, Extreme Inception (Xception), MobileNet, Dense Convolutional Network (DenseNet), and Neural Architecture Search Network (NASNet). For the classification of injuries, the Bayes, Support Vector Machines (SVM), Random Forest (RF), Perceptron Multilayer (MLP), and the K-Nearest Neighbors (KNN) classifiers are used. This study used two datasets: the first provided by the International Skin Imaging Collaboration (ISIC) at the International Biomedical Imaging Symposium (ISBI); the second is *PH*². For ISBI-ISIC, this study examined lesions between nevi and melanomas. In *PH*², this work analyzed the diagnosis based on lesions of common nevus, atypical nevi, and melanomas. The DenseNet201 extraction model, combined with the KNN classifier achieved an accuracy of 96.805 % for the ISBI-ISIC dataset and 93.167 % for the *PH*². Thus, an approach focused on the IoT system is reliable and efficient for doctors who assist in the diagnosis of skin lesions.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Skin cancer is cancer, with the highest mortality rate worldwide [1]. Melanoma is the most worrying because it is more

likely to cause metastasis. According to the American Cancer Society, in 2017, in the United States of America (USA), 9730 new cases of melanoma were estimated. The mortality rate in the same year was around 1.62%, among other types of cancer [2]. If previously diagnosed, the mortality rate may be reduced by up to 90% [3].

Medical teams have difficulty making a diagnosis consistent with the patient's reality. Several factors contribute to this reality: (1) varied shapes and sizes of lesions, (2) lighting the

e-mail: douglas@lapisco.ifce.edu.br (Douglas de A. Rodrigues), robertoivo@lapisco.ifce.edu.br (Roberto F. Ivo), sureshsatapathy@gmail.com (Suresh Chandra Satapathy), shuihuawang@ieee.org (Shuihua Wang), judehemanth@karunya.edu (Jude Hemanth), pedrosarf@ifce.edu.br (Pedro P. Rebouças Filho)

environment of the medical consultation, (3) the patient's skin color, (4) the professional's experience, and (5) his tiredness and fatigue from performing multiple visual inspections of skin lesions on the same day.

The trend is that the number of cases of this disease continues to increase worldwide; consequently, the number of deaths. In the literature, the research aims to develop methods that help doctors make a quick and prior diagnosis of this cancer [4]. In this way, a patient can have his clinical health condition previously treated.

There are many attempts in the literature to overcome these challenging problems in the classification of skin lesions. Melanocytic lesions are the focus of most researches. PH^2 challenge and the challenge organized by the International Skin Imaging Collaboration (ISIC) at the International Symposium on Biomedical Imaging (ISBI) motivated the studies. The launch of these challenges reflects the concern of public health in most countries.

Therefore, in addition to developing an effective method for diagnosing skin lesions, it must be accessible to clinical use in different regions. When the goal is to revolutionize treatments and improve public health, the most viable solution are the systems that use the Internet of Things (IoT). Also, there are the Deep Learning and Transfer Learning methods that, in recent years, have shown satisfactory results for several health challenges.

1.1. Main contributions

Therefore, motivated by the concept of the Internet of Things (IoT) and Transfer Learning, this article proposes an automatic, fast, and accessible system to assist doctors in diagnosing skin lesions. Experimental results are performed using the datasets PH^2 and ISBI-ISIC 2017. The main contributions of this work are listed below:

1. It evaluates the performance of twelve convolution neural networks as resource extractions with seven classic classifiers to diagnose the type of skin lesion.
2. Uses a consolidated IoT system to carry out different experiments with the types of injuries. Therefore, making the method accessible in clinical cases worldwide.
3. It does not require the image pre-processing step, as it performs the diagnosis from the colored images acquired from the skin.

2. Literature Review

The segmentation of the injury limit, the detection of the injury attribute, and the classification of the disease are recurrent themes in the research [5]. Many studies aim to detect skin lesions. The number of research aimed at diagnosis is much smaller.

Literature studies that address the diagnosis are subdivided into manual, automated, and hybrid methods. The first method requires visual inspection of the patient's skin by a specialized and experienced professional. In computerized methods, the classification of lesions is performed with the aid of a computer,

therefore reducing subjective decisions. The hybrid techniques use the combination of the two previous methods.

Due to the wide variety of skin conditions and the quality of the images, the classification of images using an automatic method is a difficult task. Literature studies reduce the number of images or types of lesions analyzed and diagnosed. Also, the works still perform several pre-processing to remove, for example, ruler and hair, and also to make adjustments to the brightness of the colors, improving the contrast and transforming the color space. And only after that, they perform the extraction of attributes and, finally, the classification of the images.

Melanocytic lesions are classified into three categories: common nevi, atypical nevi, and melanomas. The common nevi have symmetrical shapes, with regular and well-defined edges. Also, they have a uniform color: brown or skin color. Atypical nevi have irregular borders and variations in both shape and color. Its diameter is greater than the width of the common nevi. Melanomas have characteristics similar to atypical nevi. In general, melanomas start as superficial, indolent tumors (horizontal phase). In some cases, these tumors may grow towards the depth of the dermis (vertical phase). Unlike the second phase in which there is a high probability of metastasis, if diagnosed in the horizontal phase, there is a high chance of cure through surgical excision. The atypical nevi have a pathogenic relationship with melanoma, a diagnosis consistent with the reality of the patient's clinical status and fast is essential for medical decisions.

There are some mobile connected dermatoscopic devices on the market, such as DermLite and HandyScope. However, it is not accessible to everyone because it is expensive, and training is required for its use. Some works in architecture have developed studies contributing to the development of an IoT system to assist beginning professionals and also removing subjectivity in the classification of skin lesions.

The research by Abuzaghleh *et al.* [6] has proposed a real-time alert module to help users avoid skin burns and an automated image analysis module. However, only 200 images were used. Also, the work of Abuzaghleh *et al.* [6] needed pre-processing steps and segmentation of lesions for removing noise and grounding in the image. Do *et al.* [7] address in their research the classification of images of skin lesions that are captured by smartphones under poorly controlled lighting and focal conditions. As well as the work of Abuzaghleh *et al.* [6], the work Do *et al.* [7] substantiated using few images, in a total of 184 images. In the approach proposed in this work, the 1100 images originally captured are used, with no pre-processing and segmentation steps.

Sahu *et al.* [8] proposed a portable assistant using a Raspberry Pi to classify skin lesions. Although the method does not need network connectivity, it is a method limited by the life of the battery. It is susceptible to several problems on the hardware. Like the jobs shown earlier, it requires the pre-processing step.

Through extensive research in the literature, the only study with satisfactory results that did not use these broad pre-processing steps was the research developed by Hosny *et al.* [9]. However, Hosny *et al.* [9] do not use an IoT system as a

way to facilitate diagnosis in different regions of the world. The authors used a pre-trained deep learning network and transfer of learning with AlexNet. AlexNet played the role of extractor and classifier. The authors do not mention the time required to diagnose an image, nor did they carry out other experiments among the types of injuries.

3. Methodology

In this work, the proposed approach is a system based on IoT technology for the classification of skin lesions. Figure 1 shows the diagram of the proposed approach. The doctor can upload an image or set of images of patients' skin to an online application. This system will extract data from these images and classify them into the three types of skin lesions. The system is described in detail in Subsection 3.5. The proposed approach uses CNN models trained with the ImageNet dataset for image extraction. For the classification of skin lesions among the types of melanoma, common nevi, and atypical nevi, it is using the classic classifiers.

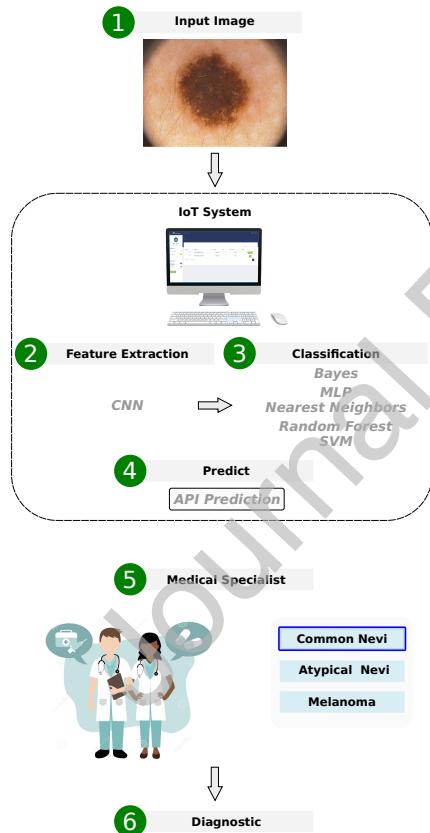


Fig. 1. Diagram of the proposed approach. Adapted from [10].

3.1. Datasets

The PH^2 ¹ and ISBI-ISIC [11] datasets total 1100 dermatoscopic images. Table 1 shows the classes of each dataset and

their respective number of images. The ISBI-ISIC dataset presents noise in your images as labels and rulers. Figure 2 shows some images of each dataset, as well as the extraction and classification steps.

Table 1. The classes of the datasets PH^2 and ISBI-ISIC, and their respective number of images.

Dataset	Class	Data augmentation	
		Before	After
PH^2	Common nevi	80	480
	Atypical nevi	80	480
	Melanomas	40	240
ISBI-ISIC	Nevi	726	4356
	Melanomas	174	1044

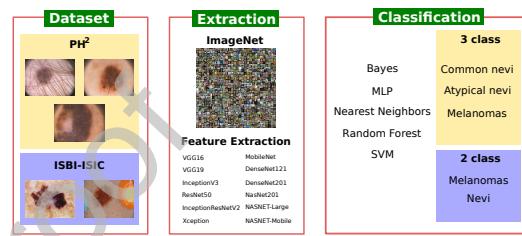


Fig. 2. Diagram showing some images from the PH^2 and ISBI-ISIC dataset, the step of extracting attributes of the skin lesion images from CNN pre-trained with the ImageNet set and the classification step with traditional classifiers.

The work Perez *et al.* [12] states that there is the best performance in the classification of skin lesions with increased data. This work used 90° rotation, horizontal and vertical rotation, gamma, logarithmic adjustment, and contrast. Figure 3 shows an example of an image from the PH^2 dataset of the six procedures used to increase the number of images. Therefore, with the data augmentation, 6600 images are totaled.

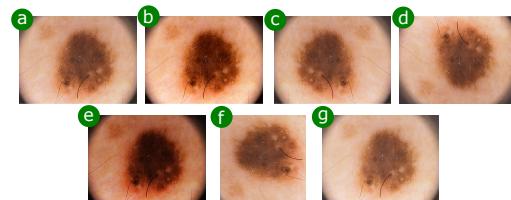


Fig. 3. Example of data augmentation of an image from the PH^2 database. Figure (a) is the original image, (b) is the result of the contrast adjustment procedure, (c) the horizontal flip, (d) the vertical flip, (e) the gamma adjustment, (f) the rotation, and (g) the logarithm adjustment.

3.2. Convolution Neural Networks as features extractors

Motivated by the concept of transfer of learning, this study made use of CNN models trained for the ImageNet dataset in the task of extracting skin lesions. The steps in using transfer learning in CNNs are: (1) Build and train a deep model on an extensive dataset; (2) Remove fully connected layers closest to the output; (3) Apply the remaining architecture to each image of a new dataset, along with the corresponding label; (4) Train a new classifier using this new dataset.

¹<https://www.fc.up.pt/addi/ph2%20database.html>

This study used the main convolutional networks in the literature: Visual Geometry Group (VGG), Inception, Residual Networks (ResNet), Inception-ResNet-V2, Extreme Inception (Xception), MobileNet, Dense Convolutional Network (DenseNet), and Neural Architecture Search Network (NASNet).

Researchers Simonyan and Zisserman [13] have investigated and evaluated the relationship of convolutional network depth with precision in large-scale image recognition. In their studies, the authors proposed the VGG network. This model makes use of convolutional filters of size 3x3. These filters provide a reduction in computational effort. The use of Factorized Convolutions allowed for increased depth without overfitting the model. In this work, we used the VGG network with 16 (VGG16) and 19 (VGG19).

The Inception [14] architecture has a small number of parameters in the network. This feature is achieved due to the use of “stacked” Inception modules, which gives it good back-propagation performance. Replacing a 5x5 convolution with two 3x3 convolutions is an example of an improvement in the Inception module since its inception. In this study, the InceptionV3 architecture was used.

ResNet consists of residual modules that are based on residual learning [15]. Equation $H(x) = f(x) + x$ displays the processing of these modules. After entering these modules, the data undergo continuous convolution operations, followed by the application of the relu function, finally convolution. The result of the $f(x)$ operation is added to the original x entry. This study used the 50-layer Resnet network (ResNet50). The Google Brain Team team combined the GoogleNet model with Resnet and created the Inception-ResNet-V2 model [16].

François Chollet has developed a new CNN model, the Xception [17]. For this purpose, his research proposed a Depthwise Separable Convolution layer. The objective is to verify the dissociation of channel correlations with spatial relations.

The MobileNet network makes use of the Depthwise Separable Convolution layer idea [18]. MobileNet differs from Xception in the use of two hyperparameters: Width Multiplier and Resolution Multiplier. Both hyperparameters allow the use on mobile platforms.

DenseNet [19] is an extension of the concept of shorter connections [15]. This network consists of the Denso Blocks. The main differences between DenseNet and ResNet are: (1) DenseNet connects each layer to all other layers with the same input resolution, while ResNet makes shorter connections between adjacent layers; (2) Every DenseNet layer performs a deep concatenation of the previous outputs, unlike ResNet that the operation is summation.

NASNet's fundament is that an efficient model can emerge from the dataset of interest [20]. NAS refers to Neural Architecture Search. NASNet search space is a new design paradigm for convolutional models.

The last fully connected layers are removed from the model to convert CNN into resource extractors. Thus, when an image is sent, the result is a vector of information. The input of classifiers to make predictions corresponds to the extracted attributes, that is, the information vectors. Table 2 shows the number of

attributes extracted by each resource extractor.

Table 2. The numbers of attributes returned for each model.

Model	Attributes	Model	Attributes
VGG16	512	MobileNet	1024
VGG19	512	DenseNet121	1024
InceptionV3	2048	DenseNet169	1664
ResNet50	2048	DenseNet201	1920
Inception-ResNet-V2	1536	NASNet-Large	4032
Xception	2048	NASNet-Mobile	1056

3.3. Classifiers

The extracted data was shuffled and randomly divided into two sets. In total, in the PH^2 dataset, there are 1080 images for training and 120 for testing. For the ISBI-ISIC dataset, there are 4830 images for training and 540 for testing. Five iterations were used to obtain the computational results. The data were trained and tested with five classifiers: Bayes, Multilayer Perceptron (MLP), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF).

The basis of the Bayes is in the Bayes Decision Theory [21]. The probability density function makes use of a gaussian distribution. This classifier makes use of a priori, a posteriori, and conditional probabilities.

The structure of the MLP neural network is consists of multiple layers [22]. The purpose of the MLP is to solve nonlinear separable problems. There are input, hidden, and output layers. Each layer has a specific activation function. The weights adjustment was performed by the descending gradient and the backpropagation of the error.

The Statistical Learning Theory is the fundamant for the SVM classifier theory. At first, SVM's idea was to solve problems with two classes, but then it was extended to multi-class problems [23]. The attributes are transformed into large vectors. Between models known as support vectors, this work used the Linear, Radial-Basis Function (RBF), and Polynomial (Poly).

Nearest Neighbors analyzes patterns that are neighboring each other in the characteristics space and concludes that they belong to the same pattern set [24]. This analysis can consider a range from one to k neighbors. For each new attribute vector not yet classified, there is a search in the set of vector images. The class of the vector more nearest vector will be assigned as the class of the new image.

The rationale of the Random Forest classifier is the decision tree method [25]. From an initial set, a group of attributes is randomly selected. Later, the data is grouped into three sets. Then, the algorithm analyzes which of the sets is most knowledgeable in solving this problem.

3.4. Evaluative metrics

From the confounding matrices, the pair of extractor-classifier combination for the diagnosis of skin lesions is evaluated. Figure 4 shows an example of a confusion matrix and the metrics equations. The values of True Positives (TP) and False Negatives (FN) correspond, respectively, to the number of lesions of a given class correctly and incorrectly classified.

True Negatives (TN) correspond to the number of lesions not belonging to a given class classified as not belonging to this class. False Positive (FP) values are the numbers of lesions incorrectly classified as belonging to a given type.

The TP, FN, TN, and FP are parameters of the equations of the accuracy (Figure 4.a), precision (Figure 4.b), recall (Figure 4.c), and F1-Score (Figure 4.d) metrics.

		Prediction	
Label	TP	FN	A
	FP	TN	B
			Acc = $\frac{TP + TN}{TP + TN + FP + FN}$
			Precision = $\frac{TP}{TP + FP}$
			Recall = $\frac{TP}{TP + FN}$
			F1-Score = $2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

Fig. 4. The Confusion Matrix structure and the equations of the accuracy, precision, recall, and F1-Score metrics.

Time is not a critical aspect of the task of classifying skin lesions. However, it is useful information if there is a want to extend the applied approach to solving tasks in real-time. Extraction time refers to the time that a CNN model needs to extract resources from an image. The time required to carry out the training of the datasets and the test time was also evaluated. The time obtained is the result of the average between all iterations.

3.5. IoT system

This work made use of the Lapisco Image Interface for Application Development (LINDA) web service. LINDA composed of a flow capable of (i) extracting, (ii) training, (iii) predicting, and (iv) storing the statistics and the results obtained for the classification of skin lesions. This IoT system performs all computational processing in the cloud. The server is located in the rectory of the Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE). In this way, computers, smartphones, and other electronic devices are allowed to send images and information to LINDA.

LINDA is a system ultimately displayed on a web system². In LINDA, a range of operations can be performed by the user: defining the number of classes, configuring the extraction and classifier parameters, and changing the extractors and classifiers used. Also, LINDA provides a graphical report with metrics that allows a coherent assessment of the extractor and classifier combination. The IoT system makes use of the PostgreSQL database, the Python programming language, and the TensorFlow and Keras libraries.

4. Results and discussions

This section presents the results obtained by combining the CNN resource extraction models with the machine learning methods applied in the classification of skin lesions.

This work evaluates 84 combinations for each dataset. Table 4 shows the average values and their respective standard deviations for the metrics accuracy, F1-Score, recall, and precision for the *PH*² and ISBI-ISIC, respectively.

For the *PH*², nine combinations achieved an accuracy greater than 90.0 %: both the Xception and DenseNet201 models with the KNN classifier, the MobileNet model with the MLP, KNN and SVM-RBF classifiers, the DenseNet168 model with the KNN, and the model DenseNet121 with standard classifiers KNN and SVM-RBF. High metrics values F1-Score, recall, and precision accompany the accuracy value. These percentage values of the metrics reaffirm the reliability of the achieved values of the accuracy of the extractor-classifier combinations.

Considering the results of ISBI-ISIC, 21 combinations out of a total of 84 achieved accuracy values higher than 90.0 %. These combinations were with the DenseNet121 model together with the MLP, KNN and SVM-RBF classifiers, the DenseNet169 model with the MLP, SVM-RBF, and KNN, the DenseNet201 with the KNN and SVM-RBF, the Inception-ResNet-V2 with the KNN and SVM-RBF, MobileNet with classifiers MLP, KNN and SVM-RBF, model VGG16 with MLP, KNN and SVM-RBF, model VGG19 with KNN and SVM-RBF, and finally, the Xception model with the MLP, KNN and SVM-RBF classifiers. These ISBI-ISIC extractor-classifier combinations achieved high values for F1-Score, Recall, and Precision.

Eight combinations achieved high results in both sets of images. In Table 4, these combinations are highlighted in green for both datasets. These combinations were the DenseNet121 model with the K-Nearest Neighbors classifiers with the SVM-RBF, the DenseNet169 model with the KNN, the DenseNet201 model with the KNN, the MobileNet model with the MLP, KNN, and SVM-RBF, and the Xception model with the KNN.

For the *PH*² dataset, the DenseNet121 extraction CNN model combined with the SVM classifier and the RBF kernel achieved the highest accuracy metric among all combinations. This set obtained 93.167 % for the accuracy metric. In addition, it reached 93.155 % F1-Score, 93.167 % for recall and 93.259 % for precision.

This combination also reached the highest values of the metrics for the ISBI-ISIC dataset. The result obtained was 96.805 % for the accuracy metric, 96.785 % for the F1-Score, 96.785 % for the recall, and 96.77 % for the precision.

Figure 5 shows the absolute confusion matrix for the CNN DenseNet201 extraction model with the K-Nearest Neighbors classifier for the *PH*² and ISBI-ISIC datasets. For *PH*², classes 0, 1, and 2 represent common nevi, atypical nevi, and melanoma, respectively. For ISBI-ISIC, class 0 represents the nevi type and class 1 the melanoma type.

For *PH*², the KNN classifier more often confuses injuries between the common nevi and atypical nevi classes. Differently with the melanoma class, in which the classifier missed only one injury about the 23 images destined for the melanoma class test.

The extraction time, training time, and test time for the *PH*², respectively, were 284.98, 629.13 ms, and 353.85 ms. For ISBI-ISIC, the extraction time per image was 449.42 ms, the training time was 4583.93 ms, and the test time was 2361.88 ms. Therefore, the time required to perform the entire procedure is reduced compared to the time necessary for a doctor.

Due to the results achieved, the proposed approach for the

²<http://lapiscoapi.ifce.edu.br>

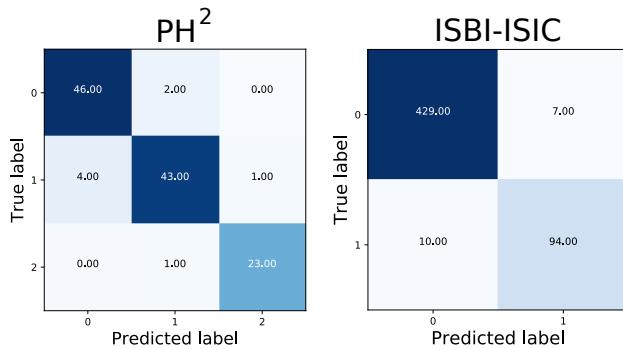


Fig. 5. Absolute confusion matrix for the combination of the DenseNet201 extraction CNN model with the K-Nearest Neighbors classifier.

diagnosis of skin lesions based on the use of transfer learning and machine learning in an IoT system is promising. It is an effective method for application in clinical routines. In this sense, one of its main characteristics is to be accessible for use in different regions and in an easy to handle manner. The limitation, therefore, is related to internet access. In which, a good connection is needed to connect to the API (Application Programming Interface) present in LINDA and send the images.

5. Conclusions

This work proposes the diagnosis of three types of skin lesions: common nevi, atypical nevi, and melanomas. In clinical routines, there are some difficulties in classifying these lesions due to similarities in physical characteristics and medical experience for their diagnosis. Images of these injuries are present in the dataset of two challenges launched in recent years: *PH²* and *ISBI-ISIC*. This work used both datasets for the experiments of this work.

This research presents an approach for the classification of skin lesions using transfer learning and convolutional neural networks. The work evaluated the combinations of twelve CNN models with seven different classifier configurations. For both datasets, the greatest results for accuracy, F1-Score, recall, and precision metrics were achieved with the DenseNet201 model combined with the KNN classifier. For *ISBI-ISIC*, the accuracy value was 96,805 %, for *PH²* the accuracy was 93.167 %.

Given the results achieved, the approach is promising and can be used by doctors to assist in the diagnosis of skin lesions, common nevi, atypical nevi, and melanoma. Professionals can easily send the images to LINDA and then obtain the classification of that image of the lesion on the patient's skin. Consequently, use this information to make a more accurate and consistent diagnosis. For future work, we intend to expand this approach to other types of skin lesions with different datasets.

6. Acknowledgment

The authors received support from the Brazilian National Council for Research and Development (CNPq).

References

- [1] Z. Ma, J. M. R. S. Tavares, A novel approach to segment skin lesions in dermoscopic images based on a deformable model, *IEEE Journal of Biomedical and Health Informatics* 20 (2) (2016) 615–623. doi:10.1109/JBHI.2015.2390032.
- [2] R. L. Siegel, K. D. Miller, A. Jemal, *Cancer statistics*, 2017, CA: A Cancer Journal for Clinicians 67 (1) (2017) 7–30. doi:10.3322/caac.21387.
- [3] R. Kasmie, K. Mokrani, Classification of malignant melanoma and benign skin lesions: implementation of automatic abcd rule, *IET Image Processing* 10 (2016) 448–455. doi:10.1109/JSEN.2020.2969414.
- [4] Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review, *Biomedical Signal Processing and Control* 39 (2018) 237 – 262. doi:10.1016/j.bspc.2017.07.010.
- [5] T. Polevaya, R. Ravodin, A. Filchenkov, Skin lesion primary morphology classification with end-to-end deep learning network, in: 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), 2019, pp. 247–250. doi:10.1109/ICAIIIC.2019.8668980.
- [6] O. Abuzaghleh, B. D. Barkana, M. Faezipour, Noninvasive real-time automated skin lesion analysis system for melanoma early detection and prevention, *IEEE Journal of Translational Engineering in Health and Medicine* 3 (2015) 1–12. doi:10.1109/JTEHM.2015.2419612.
- [7] T. Do, T. Hoang, V. Pomponiu, Y. Zhou, Z. Chen, N. Cheung, D. Koh, A. Tan, S. Tan, Accessible melanoma detection using smartphones and mobile image analysis, *IEEE Transactions on Multimedia* 20 (10) (2018) 2849–2864. doi:10.1109/TMM.2018.2814346.
- [8] P. Sahu, D. Yu, H. Qin, Apply lightweight deep learning on internet of things for low-cost and easy-to-access skin cancer detection, in: *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, Vol. 10579, International Society for Optics and Photonics, 2019, p. 1057912. doi:10.1117/12.2293350.
- [9] K. M. Hosny, M. A. Kassem, M. M. Foaud, Skin cancer classification using deep learning and transfer learning, in: 2018 9th Cairo International Biomedical Engineering Conference (CIBEC), 2018, pp. 90–93. doi:10.1109/CIBEC.2018.8641762.
- [10] C. M. D. Jr, S. P. P. da Silva, R. V. M. da Nóbrega, A. C. da S. Barros, P. P. R. Filho, V. H. C. de Albuquerque, Deep learning iot system for online stroke detection in skull computed tomography images, *Computer Networks* 152 (2019) 25 – 39. doi:10.1016/j.comnet.2019.01.019.
- [11] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, J. Rozeira, *Ph2 - a dermoscopic image database for research and benchmarking*, in: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 5437–5440. doi:10.1109/EMBC.2013.6610779.
- [12] F. Perez, C. Vasconcelos, S. Avila, E. Valle, Data augmentation for skin lesion analysis, in: *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*, Springer, 2018, pp. 303–311. doi:10.1007/978-3-030-01201-4.
- [13] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *CoRR* abs/1409.1556.
- [14] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9. doi:10.1109/CVPR.2015.7298594.
- [15] S. Wu, S. Zhong, Y. Liu, Deep residual learning for image steganalysis, *Multimedia Tools and Applications* 77 (9) (2018) 10437–10453. doi:10.1007/s11042-017-4440-4.
- [16] M. Längkvist, L. Karlsson, A. Loutfi, A review of unsupervised feature learning and deep learning for time-series modeling, *Pattern Recognition Letters* 42 (2014) 11 – 24. doi:10.1016/j.patrec.2014.01.008.
- [17] F. Chollet, Xception: Deep learning with depthwise separable convolutions, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1800–1807. doi:10.1109/CVPR.2017.195.
- [18] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient convolutional neural networks for mobile vision applications, *CoRR* abs/1704.04861.
- [19] G. Huang, Z. Liu, L. v. d. Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269. doi:10.1109/CVPR.2017.243.
- [20] B. Zoph, Q. V. Le, Neural architecture search with reinforcement learning, *CoRR* abs/1611.01578. arXiv:1611.01578.

- [21] S. Theodoridis, K. Koutroumbas, Chapter 2 - classifiers based on bayes decision theory, in: S. Theodoridis, K. Koutroumbas (Eds.), Pattern Recognition (Fourth Edition), fourth edition Edition, Academic Press, Boston, 2009, pp. 13 – 89. doi:10.1016/B978-1-59749-272-0.50004-9.
- [22] S. Haykin, Neural Networks and Learning Machines, 3/E, Pearson Education India, 2010.
- [23] K.-B. Duan, S. S. Keerthi, Which is the best multiclass svm method? an empirical study, in: N. C. Oza, R. Polikar, J. Kittler, F. Roli (Eds.), Multiple Classifier Systems, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 278–285. doi:10.1007/1149468328.
- [24] K. Fukunaga, P. M. Narendra, A branch and bound algorithm for computing k-nearest neighbors, IEEE Transactions on Computers C-24 (7) (1975) 750–753. doi:10.1109/T-C.1975.224297.
- [25] L. Breiman, Random forests, Machine Learning 45 (1) (2001) 5–32. doi:10.1023/A:1010933404324.

Table 3. Accuracy and F1-Score results of the extraction model and classifier combinations.

Model	Classifier	PH^2					I
		Accuracy	F1-Score	Recall	Precision	Accuracy	
MobileNet	Bayes	66.667±2.887	66.140±3.018	66.667±2.887	67.839±2.938	62.916±0.013	66.
	MLP	90.000±1.748	89.992±1.751	90.000±1.748	90.107±1.675	91.666±0.007	91.
	KNN	92.667±1.434	92.662±1.436	92.667±1.434	93.336±1.455	96.388±0.008	96.
	RF	87.333±3.223	87.332±3.211	87.333±3.223	87.551±3.199	86.759±0.006	84.
	SVM Linear	88.167±3.266	88.147±3.264	88.167±3.266	88.214±3.247	81.805±0.020	82.
	SVM Poly	82.500±8.898	82.671±8.560	82.500±8.898	85.163±5.498	71.527±0.030	68.
DenseNet169	SVM RBF	91.167±3.859	91.157±3.854	91.167±3.859	91.252±3.780	93.796±0.002	93.
	Bayes	54.667±4.673	52.036±5.333	54.667±4.673	62.765±3.558	80.648±0.001	72.
	MLP	88.167±2.438	88.125±2.479	88.167±2.438	88.331±2.411	92.685±0.009	92.
	KNN	92.500±1.491	92.520±1.478	92.500±1.491	92.628±1.422	96.805±0.004	96.
	RF	75.333±8.954	74.758±9.928	75.333±8.954	81.504±3.247	85.046±0.020	80.
	SVM Linear	84.500±2.718	84.136±3.079	84.500±2.718	85.565±2.133	81.481±0.009	82.
DenseNet121	SVM Poly	40.000±0.000	22.857±0.000	40.000±0.000	16.000±0.000	19.259±0.000	6.2
	SVM RBF	90.833±2.789	90.719±2.890	90.833±2.789	91.547±2.203	93.287±0.025	92.
	Bayes	53.833±1.716	50.327±2.092	53.833±1.716	62.580±1.222	80.787±0.000	72.
	MLP	89.833±1.434	89.825±1.440	89.833±1.434	90.135±1.651	92.916±0.002	92.
	KNN	91.667±1.826	91.673±1.794	91.667±1.826	92.189±1.809	80.787±0.000	72.
	RF	86.000±5.039	86.028±5.087	86.000±5.039	89.133±2.815	86.249±0.013	83.
Xception	SVM Linear	84.667±1.716	84.539±1.834	84.667±1.716	85.887±1.293	79.583±0.025	80.
	SVM Poly	40.000±0.000	22.857±0.000	40.000±0.000	16.000±0.000	19.259±0.000	6.2
	SVM RBF	93.667±2.211	93.686±2.197	93.667±2.211	94.068±2.197	94.722±0.009	94.
	Bayes	64.500±3.598	64.408±3.498	64.500±3.598	65.001±3.344	80.78±0.000	72.
	MLP	88.667±2.718	88.673±2.672	88.667±2.718	89.176±2.592	92.91±0.002	92.
	KNN	90.000±1.581	89.967±1.587	90.000±1.581	90.563±1.268	96.01±0.011	95
DenseNet201	RF	83.500±4.927	83.376±5.003	83.500±4.927	84.415±4.835	86.24±0.013	83.
	SVM Linear	85.833±3.206	85.763±3.203	85.833±3.206	86.261±3.061	79.58±0.025	80.
	SVM Poly	40.000±0.000	22.857±0.000	40.000±0.000	16.000±0.000	19.250±0.000	6.2
	SVM RBF	88.167±5.807	87.464±7.232	88.167±5.807	89.340±4.205	93.287±0.009	92.
	Bayes	56.000±4.130	51.892±4.422	56.000±4.130	61.127±9.523	80.747±0.000	72.
	MLP	89.333±1.616	89.332±1.629	89.333±1.616	89.795±1.557	92.592±0.004	92.
Inception-ResNet-V2	KNN	93.167±2.068	93.155±2.074	93.167±2.068	93.259±1.984	96.805±0.005	96.
	RF	78.333±8.182	77.627±9.667	78.333±8.182	83.582±2.203	87.546±0.010	85.
	SVM Linear	86.667±4.216	86.568±4.384	86.667±4.216	87.555±3.659	80.550±0.015	81.
	SVM Poly	40.000±0.000	22.857±0.000	40.000±0.000	16.000±0.000	19.259±0.000	6.2
	SVM RBF	89.833±4.927	89.708±5.207	89.833±4.927	90.538±3.665	96.157±0.003	96.
	Bayes	55.833±3.206	54.754±3.138	55.833±3.206	56.273±3.031	75.787±0.010	74.
NASNet-Mobile	MLP	80.000±3.801	79.976±3.806	80.000±3.801	80.315±3.843	87.685±0.012	87.
	KNN	76.000±5.307	75.830±5.249	76.000±5.307	76.551±5.494	91.481±0.009	91.
	RF	74.167±4.503	74.190±4.470	74.167±4.503	74.847±4.560	83.009±0.005	78.
	SVM Linear	78.667±2.819	78.635±2.848	78.667±2.819	78.827±2.793	77.685±0.016	78.
	SVM Poly	42.000±4.000	26.253±6.791	42.000±4.000	24.777±17.554	19.259±0.000	6.2
	SVM RBF	85.167±2.953	85.168±2.901	85.167±2.953	85.731±2.919	91.435±0.011	91.
NASNet-Large	Bayes	52.833±4.876	51.509±5.009	52.833±4.876	53.093±5.692	75.092±0.016	75.
	MLP	75.667±2.494	75.597±2.484	75.667±2.494	76.033±2.587	85.509±0.009	84.
	KNN	74.000±3.590	73.836±3.831	74.000±3.590	75.442±3.321	84.675±0.019	83.
	RF	68.833±4.491	68.773±4.552	68.833±4.491	69.805±4.003	82.453±0.003	76.
	SVM Linear	68.167±3.472	67.834±3.548	68.167±3.472	68.835±3.215	74.722±0.020	74.
	SVM Poly	40.000±0.000	22.857±0.000	40.000±0.000	16.000±0.000	19.259±0.000	6.2
Inception-ResNet-V2	SVM RBF	78.667±3.674	78.475±3.632	78.667±3.674	79.349±3.547	85.092±0.010	82.
	Bayes	52.667±4.295	52.457±4.358	52.667±4.295	54.040±3.866	71.115±0.022	72.
	MLP	76.000±3.180	75.992±3.186	76.000±3.180	76.570±3.505	86.898±0.005	86.
	KNN	75.000±3.375	74.842±3.456	75.000±3.375	77.794±2.471	88.204±0.005	87.
	RF	74.667±5.667	74.585±5.858	74.667±5.667	76.082±4.921	83.750±0.012	79.
	SVM Linear	75.333±3.145	75.208±3.062	75.333±3.145	75.714±2.957	81.296±0.017	81.
NASNet-Large	SVM Poly	40.000±0.000	22.857±0.000	40.000±0.000	16.000±0.000	19.259±0.000	6.
	SVM RBF	80.833±3.416	80.866±3.459	80.833±3.416	81.719±3.432	84.629±0.030	81.
	Bayes	61.500±5.538	61.148±5.737	61.500±5.538	61.962±5.932	80.462±0.007	74.
NASNet-Mobile	MLP	77.333±2.438	77.250±2.456	77.333±2.438	77.950±2.013	81.481±0.011	81.
	KNN	74.833±2.321	74.358±2.439	74.833±2.321	76.140±2.706	86.805±0.005	86.

Conflict of interest statement

There is NO conflict of interest among the authors