Deep learning algorithm based on YOLOV5 Neural Network for dermatoscospic classification and detection of epithelial cancer (MELANOMA)

Thiago V. Lecchi[1], Gustavo M. de Almeida[1], Rafael P. D. Vivacqua[1]

[1]*Control and Automation Engineering Program, Federal Institute of Education, Science and Technology of Espírito Santo (Ifes),*
Avenida dos Sabiás, 330 - Address of Laranjeiras, Zip-Code: 29166 - 630, Serra/Espírito Santo, Brazil
thiagolecchi@hotmail.com, gmaia@ifes.edu.br, rafsat@ifes.edu.br

Abstract. Technical concepts are introduced about epithelial cancer or melanoma, one of the most common and aggressive types of existing cancers, and about the use of applications or tools based on Artificial Intelligence techniques and the influence of the application of machine learning to perform pre-diagnoses in medicine in general, which accelerate the time of discovery of the disease and exponentially increase the chances of cure of the affected patient. The technical details of the algorithm based on Deep Learning Yolov5 are developed. The possibility of applying this tool for various purposes is discussed. The dataset used in this research is analyzed and the possibilities of characterizing tumors between benign and malignant are verified, as well as the types of epithelial formations found are classified. The possibility of using these classifications as input for algorithms based on deep learning by scanning images is discussed. It is concluded that it is possible to create logical filters that use the readings of the output data provided by the execution of the Yolov5 algorithm on a photographic basis of dermoscopic exams, obtaining relevantly high levels of accuracy that would point out with great precision the possible existence of cancerous formations, thus facilitating the pre-diagnosis and early combat of skin cancer, considerably increasing the survival of patients.

Keywords: deeplearning, melanoma, Yolov5, dermoscopy

1 Introduction

Statistically documented as the most common and frequent type of cancer in Brazil and in the world, corresponding, according to the National Cancer Institute (INCA) [1], to more than 27% of the country's malignant tumor records, skin cancer has the main root cause excessive exposure of the dermis to the sun without the use of filters or protection.

In tropical regions with a high incidence of UV rays, as in Brazil, it is very important that the population is always alert to specific symptoms of the disease, aiming to prevent, detect, diagnose and treat it, preferably at an early stage. For this, it is necessary that each of us is always attentive to warning signs, such as the appearance of new spots on the skin, which can itch, burn, flake or even bleed or wounds that do not heal within four weeks.

Melanoma, the most serious form of skin cancer, can occur in different locations throughout the epithelial tissue or the mucous membranes, and may take the form of spots, spots or signs, usually having an asymmetrical shape, irregular edges, pluricoloration and rapid increase in size. According to the Brazilian Ministry of Health [2], this is a rare type of the disease, affecting only 8,400 patients annually, however, with a high degree of aggressiveness due to the high chance of metastasis, with the “spreading” of the disease to other organs of the body, and finally, lead to death.

To prevent the appearance of possible lesions from being overlooked, and to facilitate pre-diagnosis – which is currently performed visually by physicians, specialists or not in dermatology, Mohiyuddin [3] says that numerous researches have been working towards the creation of algorithms and neural networks capable of improving diagnostic performance rates. As a solution to this problem, this research proposes to develop an algorithm based on the Yolov5 neural network – and its variations – so that it is possible to carry out the training, validation and subsequent testing of the learning of the network, under a dataset of 30,000 photos provided by the Society for Imaging Informatics in Medicine (SIIM) and the International Skin Imaging Collaboration (ISIC).

2 Epitelial cancer - Melanoma

According to the National Cancer Institute (INCA) [1], cancer is a term that encompasses more than 100 types of diseases, whose common characteristic among all is the disordered growth of cells. With their rapid division, cancer cells tend to be aggressive and uncontrollable, determining the formation of tumors, which can spread to other regions of the body, compromising the functioning of a patient's tissues or organs.

In his study, Noury [3] mentions that skin cancer is a disease that affects more Caucasian people around the world, and its difficulty has a higher incidence in an increasingly younger age group. Therefore, prevention, diagnosis and early initiation campaigns have relevant significance in public health initiatives and government policies.

2.1 Cure rate by evolutionary stage and the importance of pre-diagnosis methods

There are two distinct groups of skin cancer: non-melanoma (NMSC), more common and less aggressive, which arise in basal or squamous cells, and melanoma (COM), which originates in melanocytes, cells that produce melanin, the pigment that gives color to the skin, are rarer and have a higher mortality rate. According to Zink [4], the cure rate of squamous cell carcinoma (SCC), a type of NMSC, when it is detected and treated early, is over 95%.

Argenziano [5] describes that the analysis of nodules and the diagnosis of melanoma is performed by a dermatologist and begins initially by analyzing the clinical history of the patient and their families. After this initial verification, a physical examination is performed, known as dermoscopy, which is a method of visualizing the skin with a dermoscope, a device or apparatus that amplifies the image from 10 to 20 times, illuminates it, allowing the observation of the most depths of the skin by the optical effect of immersion with a liquid medium, gel or polarized light.

According to Caorsi [6], a study carried out with a defined and monitored group of patients with melanoma skin cancer (CPM) in hospitals in the city of Porto Alegre - RS, correlated the advance of the tumor stage with the reduction of patient survival. The survey concluded that patients who had their diagnoses made late, in which the tumor was discovered when it was at a more advanced stage of the disease, had a shorter survival time.

Table 1. Survival by staging at 1, 3, and 5 years of follow-up

Staging	%		
	1 Year	3 Years	5 Years
I	99%	98%	95%
II	99%	92%	84%
III	88%	64%	60%

The creation of tools that use a visual conference methodology under suspected epithelial lesions effectively and accurately, and which end up anticipating the detection and subsequent treatment of a cancer, results in avoiding the aggravation of a suspected lesion to a more serious level or state of tumor spread, ensuring greater chances of survival of a patient with melanoma.

2.2 Use of AI in medicine: an important tool to aid clinical diagnosis

According to Rotemberg [7], the use of artificial intelligence (AI) techniques to assist analytical tasks has evolved considerably in recent years, and, in the field of medicine, it has been a relevant factor in the search for reducing mortality, morbidity and health costs associated with cancer treatment. melanoma, improving access to expertise, diagnostic accuracy, and screening efficiency.

In the article “*Diagnostic accuracy of content-based dermatoscopic image retrieval with deep classification features*” of Tschandl [8], the ability of AI algorithms to match or even surpass the performance of dermatologists in the activity of diagnosing images of individual dermatological lesions has been demonstrated. In this same review, the reader's study did not accurately reflect clinical scenarios with all of the additional tools and methods that clinicians have access to examine, characterize, and identify all lesions in a patient.

There are other different types of techniques for *deep learning* – in addition to the spectrum of variations in Yolov5, which is the subject of study in this work – which allow for an accurate diagnosis in dermoscopic examinations. In the search "*Using Deep Learning for Dermatologist-level Detection of Suspicious Pigmented Skin Lesions from Wide-field Images*" developed by a group of MIT researchers, Soenksen mentions that a deep convolutional neural network (DCNN) was trained on a database of photos of suspicious pigmented lesions (SPLs) with satisfactory results, obtaining identification performance similar to the precision of trained eyes of medical professionals specialized in dermatology [9].

3 Object and pattern detection using Yolo

Deep learning-based object detection has been an important research point in recent years. To perform a better identification of images by AI methods, it is necessary not only to focus on the classification process of different types of images, but also to accurately estimate which objects are present in the images and their disposition/location in each figure to be processed. Zhao [10] describes this task as being known as object detection and Malta [11] describes the detector algorithms are deep neural networks, namely known as convolutional neural networks (CNN).

Among the various algorithms that exist for image detection and recognition, the R-CNN series and the YOLO series stand out. The R-CNN series have better accuracy and much lower detection speed than the YOLOv5. Jia Yao [12] cites this performance problem as the main factor that leads to the inability of the first algorithm to perform tasks that require speed, such as real-time monitoring of video records.

The YOLO series of algorithms uses the idea of regression to simply learn the general visual characteristics of a target, optimizing the time needed to solve the problem. According to the synthesis of R. Jin and Q. Niu [13], it uses single-stage neural networks to perform a complete detection of the positioning and classification of objects under analysis.

Zhao [10] cites that the main idea YOLO's main idea is to use the figure entirely as an input to the network, directly returning the positioning of a bounding box and the category to which the bounding box belongs related to the possible outputs.

In the article “*Face Mask Detection using YOLOv5 for COVID - 19*”, Lecun and Hinton [14] describes that the YOLOv5 algorithm trained on a specific database and images was able to correctly classify people who were wearing or not wearing viral protection masks. The model was tested using the YOLOv5s and YOLOv5x network versions. The first version, “s”, demonstrated a better performance than the version “x”, obtaining for the experiment in question close levels of precision with less processing time.

The object detection method can also be used to identify specific patterns that point, for example, to the existence of phenomena or anomalies, such as structural failures. This is what the article of R. Jin and Q. Niu [13], “*Automatic Fabric Defect Detection Based on na Improved YOLOv5*” demonstrates, a comparative approach about the performance of YOLOv5 in the identification process, compared to the performance of other neural networks developeds.

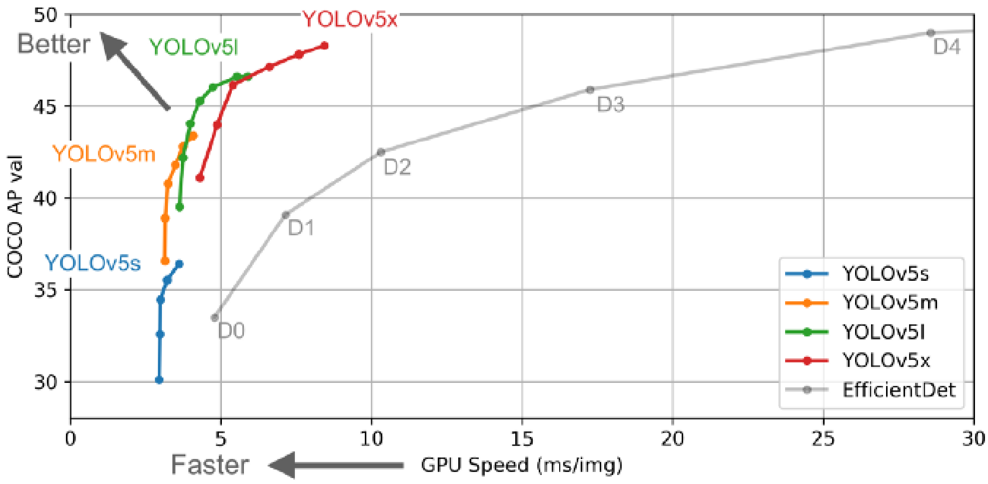

Figure 1. Speed Accuracy of each network version model
Source: GitHub.com/Ultralytics (https://github.com/ultralytics/yolov5)

4 Dataset used

For the present work, the dataset provided by the SIIM-ISIC 2020 challenge was used as a dataset. The photos of this database were collected in different hospitals, such as Memorial Sloan Kettering, Melanoma Institute Australia, University of Queensland, Medical University of Vienna , University of Athens and University of Barcelona.

In this database, the epithelial lesions represented in the dermoscopy presented were classified into two clusters, benign and malignant lesions. The only type of malignant lesion presented is the melanoma type lesion. As for the group of lesions classified as benign, seven subclassifications were made and an eighth group containing unclassified photos of benign lesions. Table 2 below represents the number of samples per dataset subgroup:

Tabel 2. Classification and number of samples from the dataset used

Tipo	Amount
Benign	**32.542**
Unknown	27.124
Nevus	5.193
Seborrheic keratosis	135
Lentigo NOS	44
Lichenoid Keratosis	37
Solar Lentigo	7
Cafe-Au-Lait Macule	1
Atypical Melanocytic Proliferation	1
Malign	**584**
Melanoma	584
Total	**33.126**

Source: Own Production

5 Performance evaluation

5.1 Metrics used

To evaluate the performance of an object detection algorithm, such as YOLOv5, it is essential to use certain metrics that can measure the effectiveness in the main task performed by a software of this type, which is to, for each photograph of a test dataset, point and isolate in a bounding box objects identified in image.

Some of the most used metrics to measure the performance of deeplearning algorithms are demonstrated in the equations (1) – (3) [14]:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{k}(R_i - R_{i-1}).P_i \tag{3}$$

The true positive (TP) is the contactor that registers the correct detection of an object that actually exists in the image. False positive (FP) records an incorrect detection of an object, i.e. the net marks an object that is not in the photo. False Negative (FN) counts the number of times an object that actually exists in the image is not detected by the network. mAP represents the average of different APs, where AP represents the area under the P-R curve.

5.2 Experiment results

5.2.1 Assumptions used in the experiment

The dataset used for this work has singularities that had to be considered prior to training the neural network. There was not, for the different classes present in the database, a proportion that would guarantee an equal amount of images per classification, thus preventing it from being possible to carry out the learning in an equal and quality way for each type of suspicious lesion exemplified in section 4.2 of this document.

From this finding, it was decided to develop the training of the neural network with the photos of the classes represented by the largest number of samples among the sets of dermoscopy with benign and malignant results. Finally, photos of benign lesions of the nevus type (5,193 photos) and malignant lesions of the melanoma type (584) were selected for the study.

5.2.2 Comparison of performance parameters

The following table demonstrates the performance comparison of the variations of the YOLOv5 algorithm for each of the defined classes and for the training/validation dataset. The “P” and “R” columns respectively show the accuracy and recall performed by machine learning. The next two columns show the mAP for the specified IoU.

According to the results presented, YOLOv5x has a slightly better performance than the other networks, but with a much higher processing time. Similarly, we have a much shorter learning time for the YOLOV5s algorithm, without necessarily having significant performance losses, when it comes to key indicators such as precision, recall or mAP. As for this study, it is not necessary to have a recurrent update of the machine learning under the dataset, we defined as the best network option to use as the YOLOv5x version, with lower speed for learning, but high performance of parameter recognition determined in the problem classes.

After executing the 100 algorithm training epochs, defined after analyzing the processing time limitations of Google Colab, the simulation platform used to test the algorithm, a satisfactory mAP 0.5 was obtained in the system, with results for versions of the YOLO with the highest performance, reaching overall levels close to 80%. The performance just did not reach rates close to 100% due to the low number of samples to perform the training of the neural network to detect cases of melanoma for all possible visual variants of this type of malignant formation. The identification of melanomas obtained results close to 70%, a considerably good performance value, with the possibility of improvement in case of a greater availability of photographic samples or a reclassification of the types of melanoma by their visual structure, identifiable with the naked eye: pigmented network, irregular brown globules, peripheral black dots and globules, structures, blue-white veil, white scar area, milky white area, blots, punctate vessels, inverted network, area of ulceration and chrysalis.

Tabel 3. Performance of YOLOv5 models (s, m, l, x) for the training and validation dataset

Class	Images	Labels	P				R				mAP@,5				mAP@,5:,95:				Training Time (h)			
			s	m	l	x	s	m	l	x	s	m	l	x	s	m	l	x	s	m	l	x
All	1.135	1.193	0,736	0,783	0,764	0,803	0,768	0,741	0,773	0,766	0,770	0,787	0,786	0,798	0,457	0,477	0,472	0,485				
nevus	1.135	1.079	0,884	0,889	0,885	0,884	0,940	0,939	0,942	0,939	0,942	0,946	0,937	0,936	0,539	0,548	0,538	0,545	2,603	6,359	10,002	19,767
melanoma	1.135	114	0,588	0,678	0,644	0,722	0,596	0,544	0,605	0,594	0,598	0,629	0,635	0,661	0,376	0,405	0,405	0,425				

Fonte: Own Production

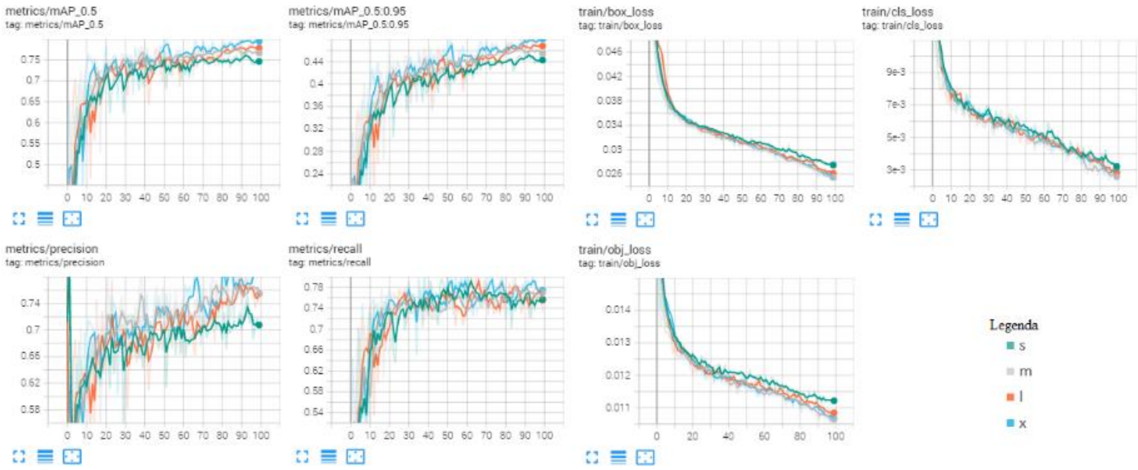

Figure 2. Performance Charts in the Training and Validation Process
Source: Author's Own Production Using the Tensorboard Tool based on YOLOv5 results

The graphs in Figure 2, which portray the performance of the learning models of each of the subtypes used in the YOLOv5 algorithm. These show that the model is progressively learning over each epoch. Performance is increasing relative to the mAP accuracy indicator until approximately epoch 120, when the system saturates and learning cannot become more efficient. For the current dataset, no more processing epochs are needed to reach the maximum average precision point of the trained network, for the dataset and defined hyperparameter settings.

The loss function shows the performance of a given predictor in classifying the input in a dataset. The lower the loss, the better the classifier is performing the identification process.

6 Final considerations

The proposed objective for this project was to develop a tool for photographic analysis of dermoscopy and selection of patients with a high probability of a positive diagnosis for melanoma, or skin cancer. The tool was built on the YOLOv5 neural network, based on deep learning techniques, and proved to be effective, accurate and fast in identifying patterns for selecting targets suspected of irregularity.

To train the network, a dataset with thousands of photos of dermoscopy with different results was used. The dataset was treated and classes of interest with a sufficient amount of photos were separated for later training and validation, creating a model by machine learning.

Several different versions of YOLOv5 were tested, comparing their performances in the main variables of yield and precision. The tests proved that the network is effective in what it is proposed to do, reaching levels of accuracy close to 70% for cases with a clinical diagnosis of the “melanoma” type. It is recommended to make improvements, such as increasing the dataset samples, mainly for the melanoma photo group, or creating a new classification, distinguishing the visual variations of the types of melanoma, so that the performance of this algorithm reaches higher levels, close to the levels obtained in the detection of cases of the “nevus” type, in which an accuracy of approximately 95% was obtained.

Acknowledgements: The authors gratefully acknowledge the support from the faculty and researchers, students of the professional master's course in control and automation engineering at Ifes, Propecaut, for their support and assistance in the preparation of this study, and to CAPES/FAPES - PDPG Cooperation, ICT+TAC project, for the financial support (TO 133/2021, Process No. 2021-CFT5C).

Authorship statement: The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] "Instituto Nacional do Cãncer (INCA)," 30 11 2020. [Online]. Available: https://www.inca.gov.br/o-que-e-cancer. [Accessed 29 05 2022].

[2] T. Saraiva, "Ministério da Saúde," Governo Federal, 04 12 2020. [Online]. Available: https://www.gov.br/saude/pt-br/assuntos/noticias/2020/dezembro/cancer-de-pele-saiba-como-prevenir-diagnosticar-e-tratar. [Accessed 28 05 2022].

[3] K. Noury, Skin Cancer, 1ª Ed. ed., Mc Gralw Hill, 2007.

[4] B. S. Zink, "Câncer de pele: a importância do seu diagnóstico, tratamento e prevenção," *Hupe,* 2014.

[5] S. H. C. S. T. R. C. R. S. F. e. a. Argenziano G, "Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet," *J Am Acad Dermatol* , vol. 670, no. 93, p. 48, 2003.

[6] F. Z. Caorsi, "Estágio Tumoral e Sobrevivência em Pacientes Diagnosticados com Melanoma Tratados na Cidade de Porto Alegre," 2015.

[7] V. K. N. B.-S. B. e. a. Rotemberg, "A patient-centric dataset of images and metadata for identifying melanomas using clinical context," *Sci Data,* vol. 8, p. 34, 28 01 2021.

[8] P. A. G. R. M. &. Y. J. Tschandl, "Diagnostic accuracy of content-based dermatoscopic image retrieval with deep classification features," *Br J Dermatol,* vol. 181, p. 155–165, 2019.

[9] T. K. S. T. C. B. M.-F. J. S. B. J. T.-S. A. N. R. R. S. C. C. K. M. M. S. J. A.-I. J. J. C. R. B. M. G. Luis R. Soenksen, "Using deep learning for dermatologist-level detection of suspicious pigmented skin lesions from wide-field images," *SCIENCE TRANSLATIONAL MEDICINE,* p. 12, 2021.

[10] Z. Zhao, P. Zheng, S. Xu and X. Wu, "Object detection with deep learning," *IEEE Trans. Neural Netw,* 2019.

[11] A. Malta, M. Mendes and T. Farinha, "Augmented Reality Maintenance Assistant Using YOLOv5," *MDPI - Applied Sciences,* 2021.

[12] 2. J. Q. J. Z. H. S. 1. ,. J. Y. a. X. L. 1. Jia Yao 1, "A Real-Time Detection Algorithm for Kiwifruit Defects Based on YOLOv5," MDPI.

[13] R. Jin and Q. Niu, "Automatic Fabric Defect Detection Based on an Improved YOLOv5," in *Mathematical Problems in Engineering*, 2021.

[14] Y. B. Y. LECUN and G. HINTON, Deep learning, 1 ed., Nature, 2015, p. 23 e 38.

[15] A. B. U. G. V. P. ,. S. A. ,. O. B. N. a. M. R. Aqsa Mohiyuddin, "Breast Tumor Detection and Classification in Mammogram Images Using Modified YOLOv5 Network," *Hindawi - Computational and Mathematical Methods in Medicine,* vol. 2022, p. 16, 2022.

[16] V. Sharma, Face Mask Detection Using YOLOv5 for COVID-19, 2020.