

ABSTRACT

The healthcare industry plays a critical role in saving lives every day. As a result, researchers, physicians, and experts are constantly working to find new ways to address illnesses and disabilities. In addition, technological advancements, especially in artificial intelligence and machine learning, have helped the scientific community design and propose advanced diagnostic tools to help physicians make crucial patient care decisions. These tools allow researchers to analyze vast amounts of data in new ways, often in real-time, for various purposes, such as detecting patterns behind illnesses, analyzing signals and detecting potential cancer from images.

In this context, this work was dedicated to the study of the Melanoma Image Binary Classification Problem (MIBCP), mainly by analyzing and proposing solutions to addressing the open issues in this field that did not allow a massive utilization of computer-aided diagnostic systems for early diagnosis. In particular, this work focuses on the resolution of the problems that may be behind high-performance automatic prediction models: the need to minimize risk situations, even by accepting lower overall performance; the opportunity to use clinical images instead of instrumental images in early diagnosis; the need able doctors to evaluate how the automatic prediction models learn and choose the results, rather than blindly relying only on the statistical values that can be calculated by analyzing the performance of the system on training, validation and test tests; the need for a scalable architecture specialized in allowing the refinement of prediction models in a fast and accessible way to non-experts.

The results reported aim to help increase trust in the automatic system that can be implemented thanks to deep learning, in particular by showing these systems' advantages, limitations and disadvantages and providing tools that show the potential to overcome these limitations. Also, this work aims to improve Melanoma early detection, which is now a limiting factor for first-line therapies in this tumour pathology.

ABSTRACT IN ITALIANO

Il settore sanitario svolge un ruolo fondamentale nel salvare vite ogni giorno. Di conseguenza, ricercatori, medici ed esperti lavorano costantemente per trovare nuovi modi per affrontare malattie e disabilità. Inoltre, i progressi tecnologici, in particolare nell'intelligenza artificiale e nell'apprendimento automatico, hanno aiutato la comunità scientifica a progettare e proporre strumenti diagnostici avanzati per aiutare i medici a prendere decisioni cruciali sulla cura del paziente. Questi strumenti consentono ai ricercatori di analizzare grandi quantità di dati in modi nuovi, spesso in tempo reale, per vari scopi, come rilevare modelli dietro malattie, analizzare segnali e rilevare potenziali tumori dalle immagini. In questo contesto, questo lavoro è stato dedicato allo studio del Melanoma Image Binary Classification Problem (MIBCP), principalmente analizzando e proponendo soluzioni per affrontare le questioni aperte in questo campo che non hanno consentito un utilizzo massiccio di sistemi diagnostici assistiti da computer per la diagnosi precoce. In particolare, questo lavoro si concentra sulla risoluzione dei problemi che possono essere alla base di modelli di previsione automatica ad alte prestazioni: la necessità di minimizzare le situazioni di rischio, anche accettando prestazioni complessive inferiori; l'opportunità di utilizzare immagini cliniche invece di immagini strumentali nella diagnosi precoce; la necessità di medici in grado di valutare come i modelli di previsione automatica apprendono e scelgono i risultati, piuttosto che affidarsi ciecamente solo ai valori statistici che possono essere calcolati analizzando le prestazioni del sistema su training, validation e test set; la necessità di un'architettura scalabile specializzata nel consentire l'affinamento dei modelli di previsione in modo rapido e accessibile ai non esperti. I risultati riportati mirano ad aumentare la fiducia nel sistema automatico che può essere implementato grazie al deep learning, in particolare mostrando vantaggi, limiti e svantaggi di questi sistemi e fornendo strumenti che mostrano il potenziale per superare questi limiti. Inoltre, questo lavoro mira a migliorare la diagno-

si precoce del melanoma, che ora è un fattore limitante per le terapie di prima linea in questa patologia tumorale.

CONTENTS

1	Introduction	1
1.1	Overview	1
1.2	Open Challenges	4
1.3	Contribution of This Thesis	6
1.4	Organization of the Thesis	7
2	Background	9
2.1	Melanoma	9
2.2	Melanoma Assessment and Visual inspection	11
2.2.1	Dermoscopic and Histological Inspection	14
2.2.2	Lymph Node Mapping	16
2.2.3	Computer Tomography	17
2.2.4	Positron Emission Tomography	18
2.2.5	Magnetic Resonance Image	20
2.2.6	Standard blood chemistry tests	21
2.3	State of the art and related works	22
2.4	Machine Learning and Deep Learning	23
2.4.1	Underfitting and Overfitting	24
2.4.2	Training methodologies	25
2.5	Evaluation of melanoma classification	33
2.5.1	Explainable AI - XIA	34
2.5.2	Local Interpretable Model-Agnostic Explanations (LIME)	35
3	Increasing trust in CAD using FNR minimization and XIA	37
3.1	Background	37
3.2	Methods	39
3.2.1	Dataset preparation	39
3.2.2	Image improvement method	41
3.2.3	Image segmentation method	41
3.2.4	CNN refactoring and evaluation	42
3.2.5	XIA analysis	44
3.3	Results and Discussion	45
3.4	Conclusion	48
4	Genetic Algorithms for Melanoma Classification	53
4.1	Idea	54

4.2	Genetic Algorithms on Clinical Image	56
4.2.1	Dataset	57
4.2.2	Materials and Method	57
4.2.3	Results and Discussion	59
4.2.4	Conclusion	61
4.3	Genetic Algorithms on Dermoscopic Images	62
4.3.1	Dataset	62
4.3.2	Pre-processing	63
4.3.3	Training, Validation and Test sets	65
4.3.4	Genetic Algorithms	65
4.3.5	GA definition	66
4.3.6	Experiments Setup	69
4.3.7	Results and Discussion	71
4.3.8	Comparison with the literature	74
4.3.9	Conclusion	75
5	A Cloud Approach for Melanoma Detection	77
5.1	Background and considered issues	77
5.2	Proposed method	78
5.3	Used Networks	80
5.3.1	MED-NODE dataset and Pre-Processing	81
5.4	Results	82
5.4.1	Results for the transfer learning reliability evaluation	82
5.4.2	Results for the impact of the three-layers architecture	83
5.5	Conclusion	86
6	Conclusions	87
6.1	Summary	87
6.1.1	Changing the POV regarding CAD performances	87
6.1.2	The potential contribute of the clinical images	89
6.1.3	Addressing the intra-class/extra-class issue with the continuous retraining	90
6.2	Future Works	91
	Bibliography	93

LIST OF FIGURES

Figure 2.1	Stages of Melanoma.	11	
Figure 2.2	Several examples of neavus and melanoma images.	14	
Figure 2.3	Example of simple naevus with similar colour, thickness and diameter.	15	
Figure 2.4	Example of clinical and dermoscopic image.	15	
Figure 2.5	Sentinel lymph node.	17	
Figure 2.6	Melanoma metastasi in the liver.	19	
Figure 2.7	Melanoma PET scan.	20	
Figure 2.8	Melanoma metastases detect by MRI [29].		21
Figure 2.9	A summary of the sierological prognostic markers for cutaneous Melanoma.	22	
Figure 3.1	Images of Naevi before and after application are shown from top to bottom. Specifically in the upper part the images before the application where you can also see the skin while in the lower part the images after the segmentation.	42	
Figure 3.2	Images of Melanoma before and after application are shown from top to bottom. Specifically in the upper part the images before the application where you can also see the skin while in the lower part the images after the segmentation.	43	
Figure 3.3	The InceptionV3 prediction explained by LIME.	45	
Figure 3.4	The global performances of the CNN on the four datasets.	48	
Figure 3.5	Comparison of SN and SP for the INA e IA datasets.	51	
Figure 3.6	Comparison of SN and SP for the NINA e NIA datasets.	51	
Figure 4.1	GACNN performance over 100 iterations.		61

Figure 4.2	The image segmentation process.	65
Figure 4.3	An overview of the experiment environment architecture.	70
Figure 4.4	Trend of death ratio over the 11 iterations.	71
Figure 4.5	Structure of the best network in the last iteration.	73
Figure 4.6	Confusion matrix of the best NN in the 11 th iteration.	73
Figure 5.1	General operation of the three layers architecture for melanoma detection.	80
Figure 5.2	Several SDs values computed for all networks.	84

LIST OF TABLES

Table 2.1	Summary of the melanoma stages	12
Table 3.1	The global performances of the CNNs on the INA and IA datasets are reported. In addition, image improvement techniques are active.	46
Table 3.2	The global performances of the CNNs on the NINA and NIA datasets are reported. In addition, image improvement techniques are not active.	47
Table 3.3	The FNR and the other metrics of the CNNs on the IA and INA datasets are reported. The image improvement techniques are active.	49
Table 3.4	The FNR and the other metrics of the CNNs on the NINA and NIA datasets are reported. The image improvement techniques are active.	50
Table 4.1	Performance of AlexNet on the MED-NODE dataset.	60

Table 4.2	Work Environment	70	
Table 4.3	Evolution during the iterations		72
Table 4.4	Experiments results	74	
Table 4.5	Reference literature	75	
Table 5.1	Performance on MED-NODE dataset for ACCs with Otsu segmentation and with and without data augmentation		82
Table 5.2	Performance on MED-NODE dataset for ACCs without Otsu segmentation and with and without data augmentation		83
Table 5.3	Performance on MED-NODE with Otsu segmentation and with and without data augmentation	83	
Table 5.4	Performance on MED-NODE without Otsu segmentation and with and without data augmentation	84	
Table 5.5	Performance drop after 100 training steps (related to Training and Validation steps)		85
Table 5.6	Clock time (in seconds) measured for both the experiments	85	

LISTINGS

ACRONYMS

1.4 ORGANIZATION OF THE THESIS

After this Chapter *Introduction*, the thesis is divided as follows:

In *Chapter 2*, there is an overview of skin cancers, Melanoma and detection issues. In particular, it reports an overview of the scientific literature relating to melanoma detection problems using clinical images (MDCI). The Chapter begins with a review of some processes required for melanoma detection and ends with a description of the classification algorithms proposed in the literature to address these issues.

Chapter 3 illustrated the experimental results related to multiple CNN architectures trained on clinical images with and without segmentation and data augmentation in order to obtain the best model of CNN and for the minimization of False Negative Rate (FNR).

Chapter 4 presents an alternative way to use an extended version of GA to address the MDCI. In particular, the Chapter presents the experimental results obtained using GA (selection, mutation, merging and crossover) to perform the design of a CNN driven by the GA scoring function; the maximization of the prediction accuracy and the minimization of the FNR was used as scoring functions.

In *Chapter 5*, the contribution of the intra-class dissimilarities (ICD) and extra-class similarity (ECS) presence in melanoma images dataset in affect classification performance is reported;

then, a hybrid architecture design on the continuous re-training approach is presented and analyzed.

Finally, conclusions and future studies are followed in *Chapter 6*. This Chapter highlights the contributions proposed by these works and any future directions.