

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي - ٠٨ ماي ١٩٤٥ بسيدي بلعباس
Ecole Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbes



MEMOIRE

En Vue de l'obtention du diplôme de Master

Filière : Informatique

Spécialité : Ingénierie des Systèmes Informatiques (ISI)

Thème

ÉTUDE COMPARATIVE DES MÉTHODES D'APPRENTISSAGE AUTOMATIQUE UTILISÉES POUR LA DÉTECTION ET LA CLASSIFICATION DES CANCER DE LA PEAU

Présenté par :

-M Khodja Moussa
-M Balbal Oussama

Soutenu le : 00/00/2022 Devant le jury composé de :

-M Badsi Hichem Président
-M Meddah Ishak Encadreur
-M Bekkouche Mohamed Examinateur

ACKNOWLEDGEMENT

“We would like to express our special thanks and gratitude to the following people for helping with this master thesis,

First of all, We thank Allah the Almighty for giving us courage, and patience needed to complete this work. We would like to thank our parents, families, and friends. We would particularly like to thank our supervisor, Mr. Meddah Ishak, for his competent help, patience and encouragement. May the members of the jury find here the expression of our sincere thanks for the honor they do us by taking the time to read and evaluate this work. We would also like to thank the teaching and administrative team of ESI SBA for their efforts that provided us with excellent training. Finally, I would like to thank everyone who has contributed directly or indirectly to the accomplishment of this work.”

Abstract

Skin cancer is one of the most common cancers in the world, and it can be fatal if not treated early, that is why its early diagnosis is considered to be the best treatment for it. And under the light of recent advancements in computational power and in the artificial intelligence field (especially its 2 subdomains machine learning and deep learning) C.A.D (computer aided diagnosis) is considered to be one of the best ways for early skin cancer diagnosis. That is why in this article we are going to do a comparative study of recent methods and algorithms applied in skin cancer analysis, detection and classification our comparison is going to be based on different types of datasets used for training, different algorithms applied, and famous performance metrics calculated by researchers such as accuracy, specificity, AUC (area under curve) ...etc. In the hopes of better understanding the problem at hand and its applied solutions, and understanding some new explored ideas and challenges faced by researchers and contributors and finally this article will help new researchers to understand what is ahead of them and get a general view on the various applied methods before engaging and contributing to this field.

الملخص

يعد سرطان الجلد من أكثر أنواع السرطانات شيوعاً في العالم ، ويمكن أن يكون قاتلاً إذا لم يتم علاجه مبكراً ، ولهذا السبب يعتبر التخسيص المبكر له هو أفضل علاج له. وفي ضوء التطورات الحديثة في القوة الحسابية وفي مجال الذكاء الاصطناعي (خاصة المجالين الفرعيين للتعلم الآلي والتعلم العميق) ، يُعد CAD (التشخيص بمساعدة الكمبيوتر) أحد أفضل الطرق للتخلص المبكر لسرطان الجلد. هذا هو السبب في أننا سنقوم في هذه المقالة بدراسة مقارنة للأساليب والخوارزميات الحديثة المطبقة في تحليل سرطان الجلد واكتشافه وتصنيفه ، وستستند مقارنتنا إلى أنواع مختلفة منمجموعات البيانات المستخدمة للتدريب ، وخوارزميات مختلفة مطبقة ، ومقاييس الأداء المشهورة التي يحسبها الباحثون مثل الدقة او (المنطقة تحت المنحنى) . على أمل فهم المشكلة المطروحة وحلولها التطبيقية بشكل أفضل ، وفهم بعض الأفكار الجديدة المستكشفة والتحديات التي يواجهها الباحثون والمساهمون ، وأخيراً ستساعد هذه المقالة الباحثين الجدد على فهم ما ينتظرون قبل الانخراط في هذا المجال والمساهمة فيه..

Resumé

Le cancer de la peau est l'un des cancers les plus répandus dans le monde et il peut être mortel s'il n'est pas traité tôt, c'est pourquoi son diagnostic précoce est considéré comme

le meilleur traitement. et à la lumière des avancées récentes en matière de puissance de calcul et dans le domaine de l'intelligence artificielle (en particulier ses 2 sous-domaines d'apprentissage automatique et d'apprentissage profond), la C.A.D (diagnostic assisté par ordinateur) est considérée comme l'un des meilleurs moyens de diagnostic précoce du cancer de la peau. c'est pourquoi, dans cet article, nous allons faire une étude comparative des méthodes et algorithmes récents appliqués à l'analyse, à la détection et à la classification du cancer de la peau. Notre comparaison va être basée sur différents types d'ensembles de données utilisés pour l'entraînement, différents algorithmes appliqués et célèbres indicateurs de performance calculées par les chercheurs telles que la précision, la spécificité, l'AUC (aire sous la courbe) ... etc dans l'espoir de mieux comprendre le problème à résoudre et ses solutions appliquées, et de comprendre certaines nouvelles idées explorées et les défis auxquels sont confrontés les chercheurs et les contributeurs et enfin cet article aidera les nouveaux chercheurs à comprendre ce qui les attend avant de s'engager et de contribuer à ce domaine.

Keywords : Skin cancer - Melanoma - Machine learning - Classification - Diagnostic

Contents

1 General Introduction	8
2 General Medecal Information	9
2.1 Skin	9
2.1.1 Skin Anatomy	9
2.1.2 Other entities also contained in the skin	9
2.1.3 Functions of the Skin	11
2.2 Cancer	11
2.2.1 Origin	13
2.2.2 Types	13
2.3 Skin Cancer	15
2.3.1 Symptoms	15
2.3.2 Types	16
2.3.3 Causes	16
2.3.4 Risk Factors	16
2.3.5 Prevention	17
2.3.6 Treatment	18
3 Artificial Intelligence	19
3.1 Artificial Intelligence	19
3.1.1 Overview	19
3.1.2 Definition	19
3.1.3 Turing Test	19
3.1.4 The 4 Types of Artificial Intelligence	19
3.1.5 Artificial Intelligence Categories	20
3.2 Machine Learning	20
3.2.1 Overview	20
3.2.2 Definition	21
3.2.3 Types of Machine Learning Algorithms	21
3.2.4 Examples and Applications	21
3.3 Deep Learning	22
3.3.1 Overview	22
3.3.2 Definition	22
3.3.3 What Is Next?	22
3.4 Ai vs Machine Learning vs Deep Learning	23
3.5 Computer Vision	23
3.5.1 overview	23
3.5.2 Using Machine Learning Methods	24

3.5.3	Using Deep Learning	24
3.5.4	Applications of Computer Vision	25
3.5.5	Some Technologies of Computer Vision	26
4	State Of The Art	27
4.1	Introduction	27
4.2	Skin Cancer Detection and Classification Using Machine Learning	27
4.3	Finding reduced Raman spectroscopy fingerprint of skin samples for melanoma diagnosis through machine learning	28
4.4	Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images	31
4.5	The impact of patient clinical information on automated skin cancer detection	32
4.6	An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features	33
4.7	Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer	34
4.8	Comprehensive Comparative Information	37
5	Results and Discussion	42
5.1	Introduction	42
5.2	Datasets	42
5.3	Extracted Features	43
5.4	Methods	43
5.5	Applicability	44
5.6	Line Of Reasoning	44
6	Our Contribution	46
6.1	Introduction	46
6.2	Proposed Convolutional Neural Network Model	46
6.3	Dataset (MNIST- HAM10000)	47
6.4	Pre-processing	47
6.5	Experimental results	49
7	Conclusion	52

List of Figures

2.1	Skin Anatomy [1]	10
2.2	Hair and Sweat Glands Anatomy [2]	10
2.3	Protective/moisture Barrier Functions [3]	12
2.4	Hair and Sweat Glands Anatomy [4]	12
2.5	DNA Mutation [5]	13
2.6	DNA Rearrangements [6]	14
2.7	Benign and Malignant tumors [7]	14
2.8	3 Most Common Types of Skin Cancer	16
3.1	AI vs ML vs DL [8]	24
3.2	Machine Learning in Computer Vision [9]	25
3.3	Deep Learning in Computer Vision [9]	25
4.1	Preprocessing: (a)Dull razor image, (b) Gray scale image, (c) Gaussian filter, (d) Median filter.	28
4.2	Segmentation: (a) Image labelled by cluster index, (b) Objects in cluster 1, (c) Objects in cluster 2, (d) Objects in cluster 3.	28
4.3	Confusion Matrix	29
4.4	ROC	30
4.5	DLS Interface	31
4.6	Model	33
4.7	ANN Structure	35
4.8	Evaluation Measures	35
4.9	Whole segmentation model input and output	40
6.1	proposed architecture which we have used for melanoma recognition . . .	47
6.2	A sample of each type of skin lesion [10]	48
6.3	This count plot helps to understand the distribution of the data. [10] . .	49
6.4	Pre-processing	49
6.5	Accuracy and Loss	50

List of Tables

2.1	Cancer Statistic [11]	12
4.1	A comparative table of the methods mentioned above in this article . . .	38
4.2	A comparative table of the latest methods used in skin lesion detection. DCNN(Deep convolutional neural network), FCRN (Fully Convolutional Residual Networks) [13]	39
4.3	Another comparative table of the latest methods used in skin lesion detec- tion and their limitations [12]	41
6.1	Evaluation Measures	51

Chapter 1

General Introduction

In this article we are going to do a recap and a comparative study of methods used for skin cancer detection and classification, but before diving into the technical stuff we need to define the background first so we don't get lost in medical/technical terms later on. That is why firstly in the general medical information chapter we defined and explained some medical and biological concepts we talked about skin anatomy, the background of cancers in general and ended it talking about skin cancer in particular, its definition, most widespread types, causes, risk factors and potential treatments with (early diagnosis + removal of tumor) being the best combination for treatment. Secondly in the general AI information chapter we explained and presented Artificial Intelligence concepts, we talked about its 2 sub categories machine learning and deep learning and their characteristics and ended it talking about Computer Vision because it was the most used branch of AI in the discussed methods, we mentioned its implementation in both machine learning and deep learning and its famous applications.

Reaching the state of the art, we did a recap and summary of used methods explaining their line of reasoning, datasets used, issues they faced, different feature extraction techniques and their algorithms with their evaluation metrics. Since there is a lot of work in this field we firstly presented articles that implement the usual method (image -> process -> train -> predict) then secondly concentrated on papers that presented new ideas such as using a different type of dataset (such as Raman spectroscopy or the combination of images and clinical information) and finally we presented a comparison table of said methods and added 2 more tables from previous researchers that did a comprehensive survey of literature of the last few years.

After the recap, we discussed the differences and presented and extracted useful information that could help future researchers and contributors to have an overview and a "map" of this field before diving into it.

Chapter 2

General Medecal Information

2.1 Skin

The skin is a complex organ [14], it is interactive, self renewing and represents the first and primary defense line against hostile environment, and it has several characteristics such as selective absorption, auto regeneration when injured, barrier to water loss, touch sensitivity ...etc. [15]. It represents the largest sensory organ (15% of total body weight and a total area of 1.86 m^2) [16], it has a highly adaptive structure that makes it vital for the survival of the human body, the balance between its static and dynamic properties makes it highly adaptive to the variations of the outer world [17].

2.1.1 Skin Anatomy

The skin is primarily composed of 3 main layers as shown in the figure 2.1, each layer has its unique properties and functions [16].

Epidermis The outermost layer which is constantly regenerating, contains the pigment melanin that determines the skin color, and it also represents a physical and biological barrier.

Dermis The middle layer, it supports the flexibility and gives strength to the epidermis, and it is mainly composed of connective tissue.

Hypodermis The last layer which is composed of subcutaneous fat which gives it its properties of being a main support of the overall structure of the skin and shock absorption.

2.1.2 Other entities also contained in the skin

Hair provides protection against minor trauma, and thermoregulation and filtering functions such as nasal hair and eyelashes

Sweat Glands it is found across the entire body, it provides lubrication, temperature regulation and salt and water balance.

their anatomies are shown in the figure 2.2

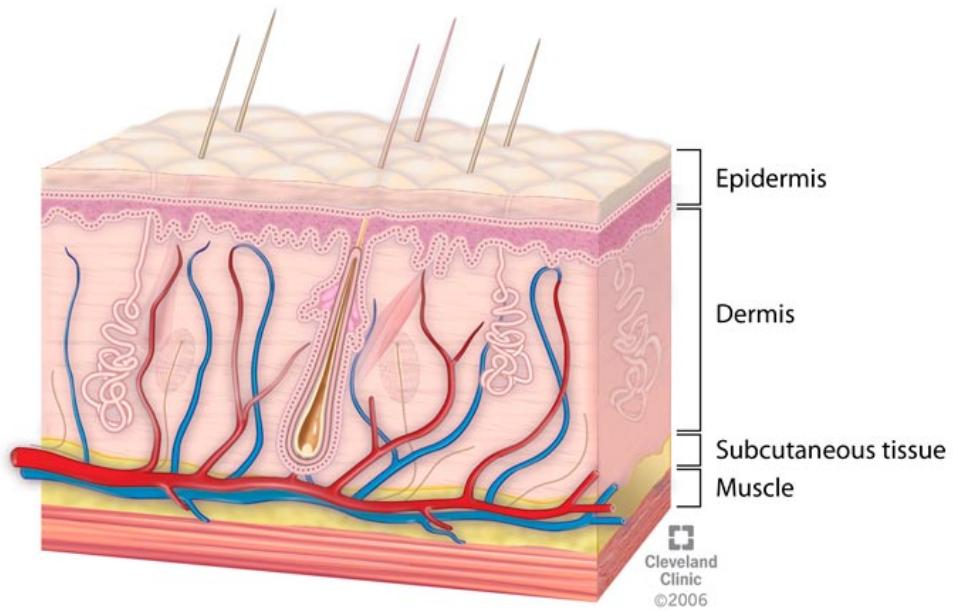


Figure 2.1: Skin Anatomy [1]

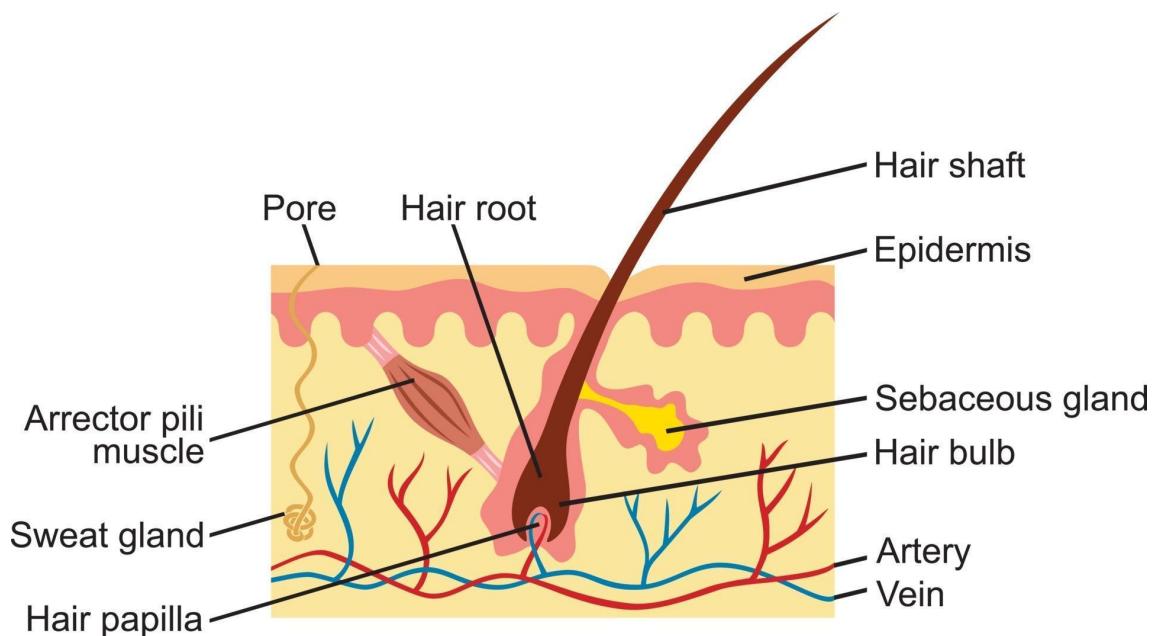


Figure 2.2: Hair and Sweat Glands Anatomy [2]

2.1.3 Functions of the Skin

The skin has 6 main functions that can be summarized as follows [16]

Protection

The skin is a direct interface between the internal organs and the environment, so it works as a protective barrier against harmful objects and pathogens (innate/adaptive immunity and ultraviolet light protection [15]) as shown in figure 2.3.

Thermostat

The skin works as a thermoregulator to keep the body at the optimal temperature of 37 C°, to achieve that it uses multiple strategies such as insensible perspiration, sweating ...etc.

Neural relay network

The skin contains a dense network of neural endings that works as receptors to various signals and provides sensations for touch, temperature and pain.

Expression and communication

A more social function is the ability for skin to enable individuals to display emotions. It acts as an indicator of one's physical state. Skin is an important component of the stress response as it acts as an immediate stress perceiver and as a target of stress responses. The skin also works as a social tool for interactions between human beings by indicating the physical state of the individual and by showing signs of stress.

Water storage

This skin works as a conservative barrier against water and body fluids' leakage as shown in figure 2.3.

Synthesis of vitamin D

The skin represents the main site of vitamin D production when exposed to the sun, it exists in the plasma membranes of basal and suprabasal keratinocytes in its inactive form then it is converted to previtamin D3 then to vitamin D in the liver and kidneys [15] as shown in figure 2.4.

2.2 Cancer

Cancer is an illness caused by the uncontrolled division and spreading of normal cells [18] unlike other diseases, cancer is caused by our own bodies and not by foreign entities, and it is one of the biggest causes of death among human beings nowadays (Table 2.1) and that is because of the ineffectiveness of traditional treatment methods such as hormones, surgery, radiation, and chemotherapy [19]. Their ineffectiveness is due to their side effects that lead the body to deteriorate more and more. But it is worth mentioning that there are some new methods and approaches being developed by researchers, a couple of those methods are the study of stem cells in relation to cancer cells and the study of the normal cells that the cancer cells came from which are called "Cancer Origin Cells", the latter approach proposes that we should study these origin cells because of their big similarities with cancer cells which will give us a roadmap to its diagnosis and therapy [20].

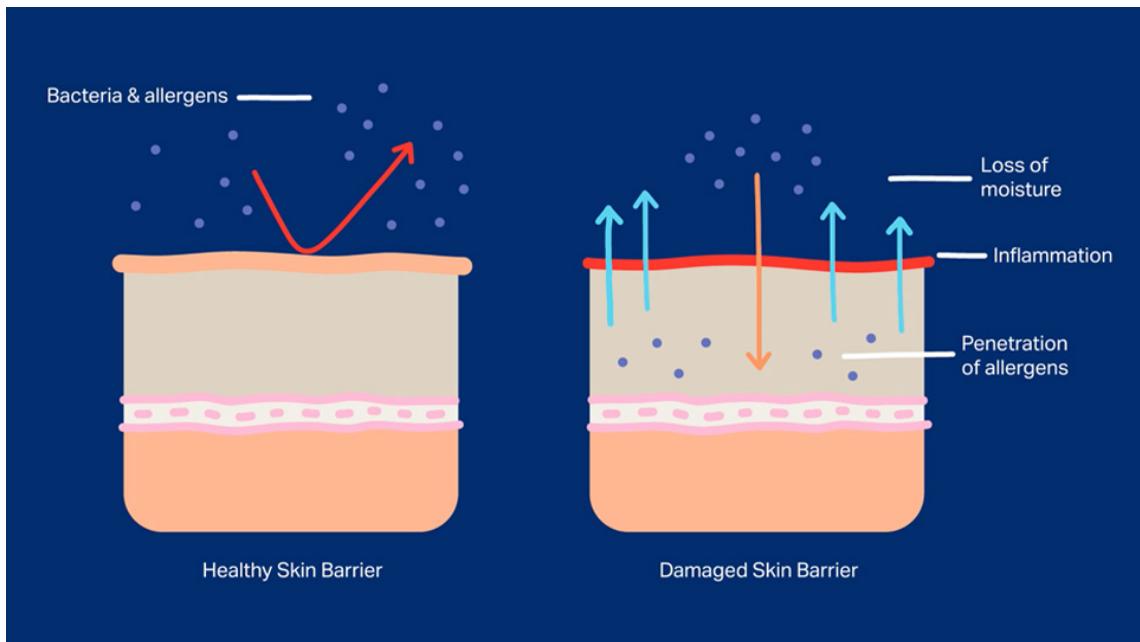


Figure 2.3: Protective/moisture Barrier Functions [3]

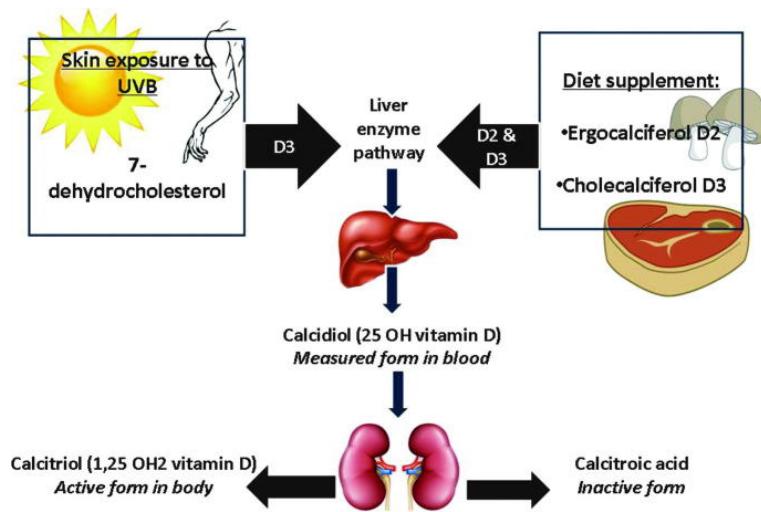


Figure 2.4: Hair and Sweat Glands Anatomy [4]

Deaths in 2020	nealy 10 million
Type	New Cases (millions) in 2020
Breast	2.26
Lung	2.21
Colon and Rectum	1.93
Prostate	1.41
Skin	1.20
Stomach	1.09

Table 2.1: Cancer Statistic [11]

2.2.1 Origin

One of the theories that discuss this is the "carcinogenesis multi-hit theory" which stipulates that for cancer to emerge there are some conditions (hits) that need to be satisfied these hits are produced by genetic mutations (figure 2.5) or rearrangements (figure 2.6) that occur over many years and the number of hits necessary is minimal ranging from 3 to 7 only [20]. But it is only fair to mention that there are some exceptions to the rule as there are some cancers caused by only one hit. And to go a step further these mutations can be caused by various elements in our environment such as chemicals in tobacco, ultraviolet rays...etc [18].

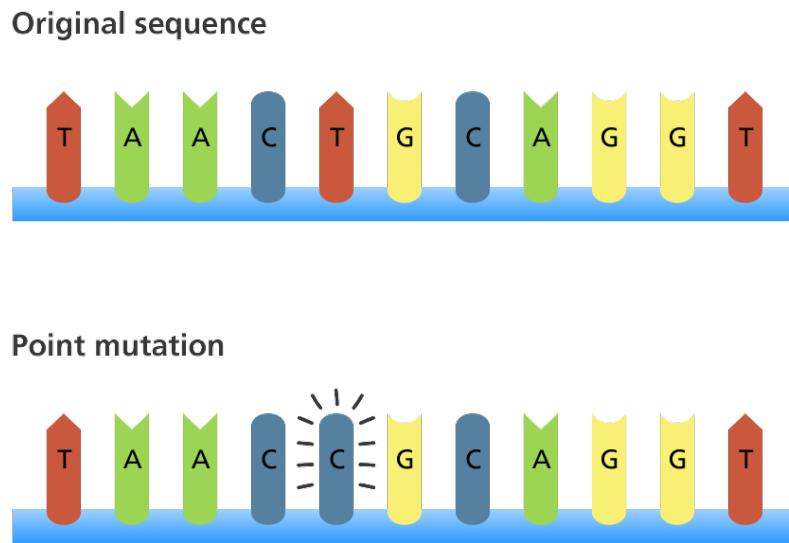


Figure 2.5: DNA Mutation [5]

2.2.2 Types

According to Fatality

Benign tumors

Are not very harmful because they do not spread to other organs and do not invade nearby tissue, and after removal, they usually don't grow back [18] as shown in figure 2.7.

Malignant tumors

Fatal if not treated, because they travel to distant places and form other tumors and invade nearby tissue [18] which makes it very hard to remove all its parts, as shown in figure 2.7.

According to Origin

Cancer is also categorized according to where it originated or its origin cells, in this category, there are over 100 types because of the different places it can appear (lung cancer, brain cancer ...) and the different origin cells that it can come from [18].

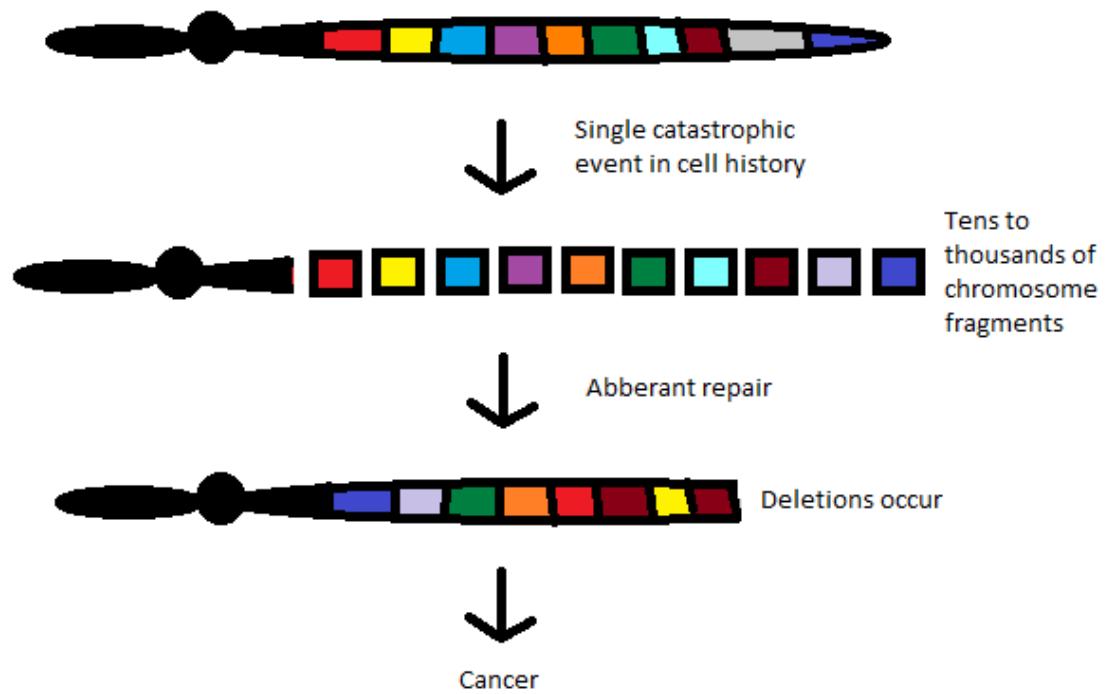


Figure 2.6: DNA Rearrangements [6]

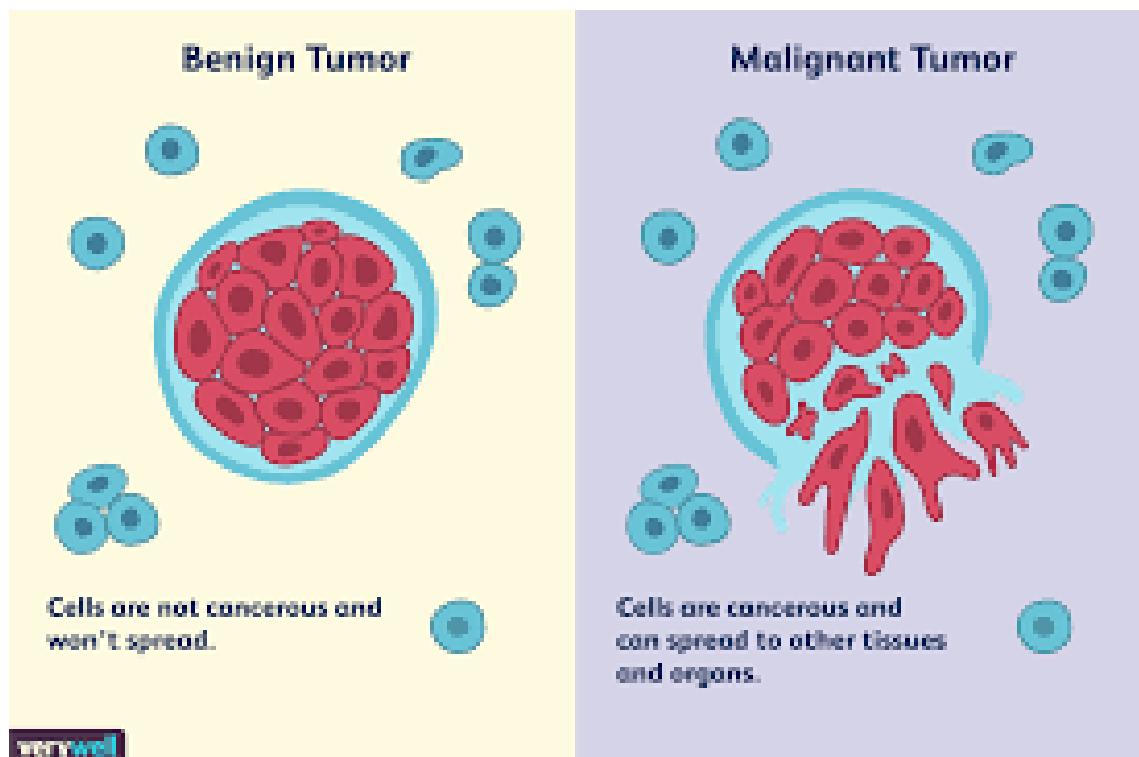


Figure 2.7: Benign and Malignant tumors [7]

Carcinoma

Most common type formed by epithelial cells.

Sarcoma

Forms in bone and soft tissue.

Leukemia

Forms in bone marrow, this type does form a tumor but travels in the blood.

Melanoma

Formed by melanocytes (cells that make melanin that gives the skin its color) ...etc.

2.3 Skin Cancer

Skin cancer is the abnormal growth of cells found in the epidermis (the outer layer of the skin) [21], it is one of the most common cancers in the world [22] and it falls under the category of a malignant tumor that is formed by fast multiplication of cells which is caused by mutations/damage in the DNA of those cells, the damage in their DNA is due to the exposure to ultraviolet rays [21] which can come from various sources but the most common are sunlight and tanning beds [21–23]. The most common types of skin cancer are basal cell carcinoma, squamous cell carcinoma , melanoma. The good news is that if it is discovered in an early stage or pre cancerous stage, it can be treated easily without leaving a scar.

2.3.1 Symptoms

Skin cancer can appear in any place on the body that is exposed to sunlight like : face, scalp, chest ...etc. But there are some cases where the cancer appeared in areas not always exposed to sunlight such as palm, soles, under the finger nails [23]. Skin cancer can happen to people of any skin color, but it is known that people with darker skins are less likely to have it because of the protection against ultraviolet rays provided by the melanin, which present in darker people in more quantities than pale people [23].

1. Basal cell carcinoma signs and symptoms Figure 2.8c

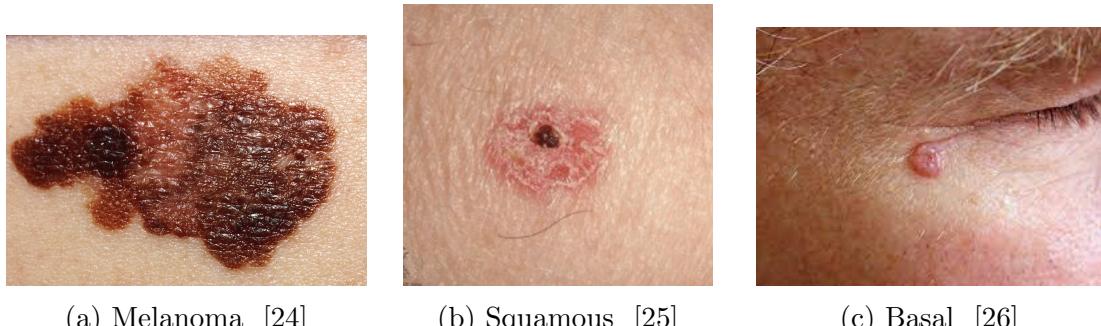
- bump
- flat brown scar
- bleeding sore that heals and returns

2. Squamous cell carcinoma signs and symptoms Figure 2.8b

- red nodule
- flat lesion with crusted surface

3. Melanoma signs and symptoms Figure 2.8a

- brownish spot
- painful lesion that itches and burns
- dark lesion



(a) Melanoma [24]

(b) Squamous [25]

(c) Basal [26]

Figure 2.8: 3 Most Common Types of Skin Cancer

2.3.2 Types

The 3 most common types are the following [21]

Basal cell carcinoma

The most common type with about 3.6 million new cases each year in the United States, if not treated early it can cause local destruction it can spread and in rare cases it is fatal.

Squamous cell carcinoma

The second most common type with about 1.8 million new cases in the United States each year, if not treated early it will spread, and it is in some cases fatal (15000 deaths/year in the United States).

Melanoma

One of the most common types, by 2022 it is estimated that, 197700 will appear in the United States, although it is treatable if detected early it is considered to be the most dangerous among common types because of its death rates (7650 deaths projected for the United States in 2022).

2.3.3 Causes

The most common and main cause of skin cancer is the exposure to ultraviolet [21–23] radiations that can primarily be found in sunlight and tanning beds, but there are some cases where the cancer appeared in areas not exposed to the sun like palms, soles, and under finger nails which indicates that other factors may contribute to the formation of skin cancer such as toxic substances, weak immune system, other types of radiation ...etc. [23]. The cells that the skin cancer originates from are squamous cells, basal cells and melanocytes. Squamous cells is just below the outer surface, basal cells is beneath squamous cells, and it produces new skin cells and melanocytes are the cells responsible for generating melanin which is the pigment responsible for the skin color. [23].

2.3.4 Risk Factors

Factors that may increase your chances of getting skin cancer are [23]

Fair skin

If you have less melanin which means your skin color is less dark than you are much

more likely to get skin cancer than a person with a darker skin because the melanin pigment is responsible for protecting the skin from ultraviolet effects.

History of sun burn

Having had sun burns before, either in childhood or adulthood, may increase your chances.

Exposure to the sun for long periods of time

Being exposed to the sun a lot or using tanning beds a lot is also one of the factors, a tan is your skin's injury response of ultraviolet rays.

High altitude climates

Living in higher places like mountains means that you are exposed to strong sunlight.

Moles

Some types of irregular moles -which are bigger than normal moles- can turn cancerous.

Precancerous skin lesions

There are some types of skin lesions -which are in themselves not cancerous - that are likely to turn cancerous, such as Bowen's disease and Actinic keratoses.

Family/Personal history of skin cancer

Weak immune system

Such as having HIV, AIDS or taking immunosuppressant drugs after an organ transplant...etc.

Exposure to radiation

Exposure to certain substances

Some harmful/unharmful substances can increase your chances, such as arsenic.

2.3.5 Prevention

As it is mentioned in [23,27]

- avoid the sun at the middle of the day
- use sunscreen to protect against sunburn with an SPF (Sun Protection Factor) over 30
- protective clothing, especially when living in the desert
- avoid tanning beds
- always check your body for abnormalities and report them to your doctor
- see a dermatologist at least once a year

2.3.6 Treatment

Before treatment, we need diagnosis first, there are two methods [28] to know that you might have skin cancer. The first method is by observing your skin frequently to see if there are some marks or abnormalities, after that you check in with a doctor who will perform further examinations which will bring us to the second method, skin biopsy - taking a part of the suspicious area of the skin and performing some laboratory tests on it to have accurate results-. After confirming that you have a skin cancer, further tests will determine what stage is it at, which is often referred to with Roman numbers (**I-** means small and limited to the area where it started / **IV-** means advanced cancer that has spread to other parts of the body). Treatment methods may vary depending on the size, type and stage of the cancer [22] but the main way to treat cancer is to remove it completely especially if it is in early or precancerous stages otherwise if additional treatment is needed, the options are as mentioned in [28]:

- freezing with liquid nitrogen
- Mohs surgery, which is for difficult cases where the surrounding healthy skin can't be removed with cancerous cells (such as the nose area)
- Curettage and electrodesiccation to eliminate remaining cancerous cells
- Radiation therapy such as X-rays
- chemotherapy with substances that contain anti cancer properties, such as lotions if the cancer is on the surface
- Photodynamic therapy, a combination of laser and chemicals
- Biological therapy using the body's own immune system

Chapter 3

Artificial Intelligence

3.1 Artificial Intelligence

3.1.1 Overview

After breaking the Enigma machine that was made by the Nazis for secure/encrypted communications in world war against the allies, Alan Turing once again changed the course of history by asking the following question "Can machines think?" in a paper he published in 1950 titled "Computing Machinery and Intelligence", this question is what gave rise to Artificial Intelligence, because all what artificial intelligence is trying to do is answer that question in the affirmative by trying to mimic human intelligence in machines [29] to do so Turing has put forward a test called "The Turing Test" which will be explained later. Because artificial intelligence is a concept that is so broad and general people don't always agree on a definition, but we found that the below definition is a good enough explanation.

3.1.2 Definition

"Artificial intelligence (AI) is a wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence." [29]

3.1.3 Turing Test

It is basically a test put forward by the mathematician Alan Turing to determine whether a machine is intelligent or not, the test goes as follows, "If a machine can engage in a conversation with a human without being detected as a machine, it has demonstrated human intelligence." [30]

3.1.4 The 4 Types of Artificial Intelligence

Reactive Machines

It is one of the most basic form of artificial intelligence because as the title suggests it only reacts to its surrounding environment, and does not use a memory to try and learn from past experience, so it is purely reactive which means that this type of artificial intelligence can only be responsible for a very narrow and specialized

set of tasks, this narrowness can be looked at as a limitation but in fact it is what makes it special in being very trustworthy and error free. A famous example of this type would be the chess playing machine Deep Blue made by IBM in the 1990s which treats each move in the game as its own separate reality and doesn't rely on past moves [29].

Limited Memory

It is a type of artificial intelligence that relies on memory and automatic training, which means learning from experience to try to make optimized decisions/predictions, the learning steps in this type can be looked at as a feedback loop (generate data, learn, make model, make predictions, accept feedback), there are 3 major models that utilize this type [29]:

- Reinforcement learning: learning from trial and error.
- Long Short Term Memory (LSTM): uses past data to make predictions, the more recent the data the more weight it has on making predictions.
- Evolutionary Generative Adversarial Networks (E-GAN): this model grows constantly by putting 2 machines against each other, and they learn by bouncing information off of each other.

Theory of Mind

This is purely theoretical and technology is still not caught up to this, and it stipulates that machines would be able to understand how humans and animals think and feel and make decisions through self reflection [29].

Self-awareness

After Theory of Mind is established, this is the next step, where machines become self-aware and comprehensive of its own existence by obtaining human level intelligence and consciousness [29].

3.1.5 Artificial Intelligence Categories

Generally speaking, there are 2 categories of artificial intelligence [29]

Narrow artificial intelligence

Also known as "Weak artificial intelligence", it operates in a limited context and is often specialized in a single task such as : Google Search, Image Recognition, Self-Driving Cars...etc.

Artificial general intelligence

Also known as "Strong artificial intelligence", it is the kind of artificial intelligence we see in Science Fiction movies implemented in robots that have human level intelligence and that can apply its intelligence to solve any problem.

3.2 Machine Learning

3.2.1 Overview

Machine learning is a subfield of artificial intelligence that has a human like ability to learn from past experience through statistics and data, and it has helped us solve difficult world

problems ranging from medical problems to environmental issues, and the special thing about machine learning is its ability to solve these problems without being explicitly programmed to do so with the usual sequence of code lines that define normal (non-artificial intelligence) algorithms, but it relies on tacit knowledge (past experience) to try and find patterns and make predictions, humans use tacit knowledge all the time for example a person can't accurately explain how he performs face recognition, but it is gained through the experience of observing that face numerous times in different angles and states [31].

3.2.2 Definition

"Machine learning is a subset of artificial intelligence that gives systems the ability to learn and optimize processes without having to be consistently programmed. Simply put, machine learning uses data, statistics and trial and error to "learn" a specific task without ever having to be specifically coded for the task." [31].

3.2.3 Types of Machine Learning Algorithms

There are 3 types [31]

Supervised Learning

Supervised machine learning algorithms provide a mathematical model that can make the connection between inputs and outputs of the training data (pre-labeled data) in the most optimized way so that when it is provided with new data, it can make very accurate predictions. Regression and classification are the most popular supervised algorithms.

Unsupervised Learning

Unsupervised algorithms take unlabeled input data and try to structure it in the form of clustering or grouping by taking into account commonalities or lack of commonalities.

Semi-Supervised Learning

This type falls in the middle, it is given labeled and unlabeled data with unlabeled being the bigger percentage than the algorithm is going to cluster the unlabeled data through the structure of the labeled data which offers a huge optimization for both sides, because supervised learning requires a huge size of labeled data which is usually done by human beings which means that it takes a lot of time and is bound to human error, and Unsupervised learning algorithms takes a lot of time also figuring out the connections in the raw unlabeled data.

3.2.4 Examples and Applications

As mentioned in [31]

Financial Services

This industry is using machine learning almost in every aspect, because of its ability to speed up the financial processes and perform tasks that used to take humans days or weeks in merely seconds. Such as handling millions of transactions, recommending personal offers ... etc.

Healthcare

This industry is also relying a lot on machine learning because of its ability to discover new treatments and detect and predict diseases, a medical professional equipped with machine learning is far more proficient because he can access a patient's relevant medical history in blink of an eye rather than digging through files or contacting other departments in the hospital. Machine learning is predicted to save the medical field billions of dollars annually.

Social Media

This industry usually uses machine learning for 2 main reasons: strengthening the feel of connection between people and eliminating bad actors, it does the former by providing individualized recommendations to friends, pages, and communities based on a user's preference or activity history, and for the latter it tries to prevent fake news before it becomes a thing, block malicious users and scams when detecting abnormalities.

...etc

3.3 Deep Learning

3.3.1 Overview

Yet again another subfield with great capabilities, although it seems to be a new concept, but it actually isn't as our professor Rahmoun Abdellatif once mentioned in a lecture talking about deep learning and neural networks, he said that the theoretical part was established a long time ago (1950's) but people back then didn't have the computational power to implement it, so it took quite some time for people to develop the necessary computational power to take on artificial neural networks and one of the scientists who made neural networks cool again is Geoffrey Hinton by demonstrating that a few of them could be trained using backpropagation for better shape recognition and word prediction and by 2012 deep learning is basically used everywhere [32].

3.3.2 Definition

"Deep learning (sometimes known as deep structured learning) is a subset of machine learning, where machines employ artificial neural networks to process information. Inspired by biological nodes in the human body, deep learning helps computers to quickly recognize and process images and speech. Computers then "learn" what these images or sounds represent and build an enormous database of stored knowledge for future tasks. In essence, deep learning enables computers to do what humans do naturally- learn by immersion and example." [32]

3.3.3 What Is Next?

Although deep learning has brought us many accomplishments, and it can be applied in various domains and when it is done right it can perform a certain task with super-human level but some scientists and researchers say it is only a small step in acquiring actual intelligent machines because it lacks the concept of abstract ideas and knowledge such

as: what objects are?, what they are for?, how to use them?...etc. And also the problem of "data" because deep learning requires a huge amount of pre labeled data to be trained which is not always available and public datasets won't cut it [32].

And there are a lot of new concepts that are presenting promising results like "deep reinforcement learning" a combination of deep learning and reinforcement learning, and we can see this implemented in a software called AlphaGo and AlphaGo Zero, another research paper suggested "Reward learning from human preferences and demonstrations" which basically means machines learn from observing humans play games which they say it works better than trial-and-error systems [32].

other ideas that are worth mentioning [32]

ONE-SHOT LEARNING and NAS (neural architecture search)

One-shot learning means we need far fewer data to learn, and NAS means an algorithm finds the best neural network architecture to solve a problem, this combination is very promising.

GANS (Generative Adversarial Networks)

A competition for deep learning which puts 2 networks against each other (a generator and a discriminator) you can think of it as a counterfeiter and a cop.

AUTOML

Learn-to-learn, which basically means machine learning algorithms do the hard work of finding the design of the network and all we need to provide is data.

3.4 Ai vs Machine Learning vs Deep Learning

After all what we have talked about it is obvious that the relationship between the three is an inclusion relationship, deep learning is a subset of machine learning which is a subset of artificial intelligence as shown in Figure 3.1 .

3.5 Computer Vision

3.5.1 overview

Yet another subfield of artificial intelligence which is used to train machines to see, and by see we mean process analyze and extract useful information from images/videos just like us human beings, although our vision is far more advanced in many aspects because our brains were trained since birth to see, analyze objects, understand the distance and relationship between objects, attribute abstract information to objects...etc. But it is safe to say that machines can surpass our vision in certain specialized tasks because of their ability to process thousands of images/frames in a short period of time due to the constant increase in computational power especially (graphical processing). Computer vision is used in a wide variety of industries, and its market is estimated to reach 48.6 billion USD by 2022 [33].

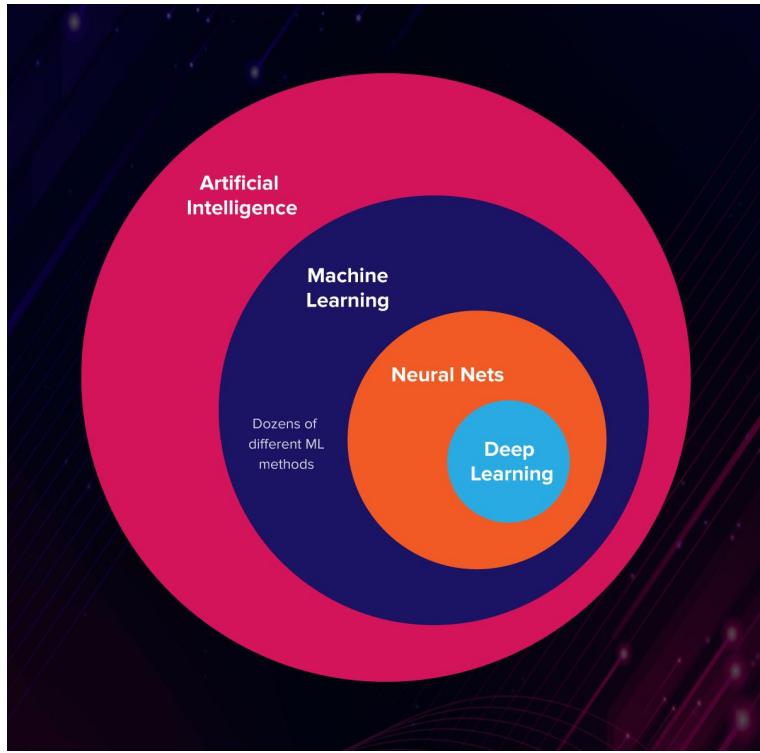


Figure 3.1: AI vs ML vs DL [8]

3.5.2 Using Machine Learning Methods

In the case of using machine learning for computer vision there are mainly 4 steps to execute, the first step is data preparation (preprocessing) in this step we need to perform some manipulations and transformations to clean the image data, some of these manipulations are cleaning noise, converting images to the same format, cropping, using gray scale instead of RGB...etc. each case requires its own set of manipulations and transformations. The second step is feature extraction which represents the hard work in most of the cases, in this step we extract a certain set of predefined features to be fed later to the algorithm, the third step is model training using the pre-labeled feature vectors, and the last step is predictions made for new image data, and for this we can choose from a variety of machine learning algorithms depending on our problem: Bayesian Nets, Decision Trees, Nearest Neighbors...etc [9].

3.5.3 Using Deep Learning

Applying deep learning in computer vision is totally different from applying classical machine learning algorithms, firstly, deep learning requires quantity (huge amounts of image data) over quality to have a robust model with accurate predictions, secondly neural networks saves us the trouble of feature extraction especially when using Convolution Neural Networks [34](Convolution: a mathematical operation on two functions to produce a third function [33]) this architecture of neural networks is specialized in processing image data and it is built on three primary layers Convolution layer, pooling layer and fully connected layer [9].

Convolution layer

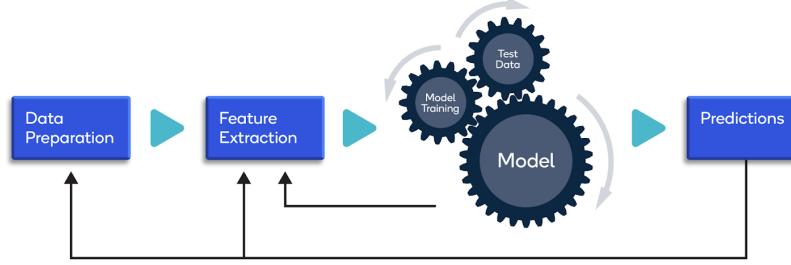


Figure 3.2: Machine Learning in Computer Vision [9]

This layer does most of the hard work by identifying and extracting the features, this is done by applying a filter of random size to blocks of the input image using the dot product between matrices.

pooling layer

After the feature extraction resulting from the Convolution layer we need to simplify (by reducing a bloc of values to a single value) the image for easy learning, there are 2 pooling operations max pooling and average pooling.

fully connected layer

It operates on a flattened input, where each input is connected to all the neurons, it is usually found at the end of the network connecting the hidden layers to the output which help in optimizing the class scores.

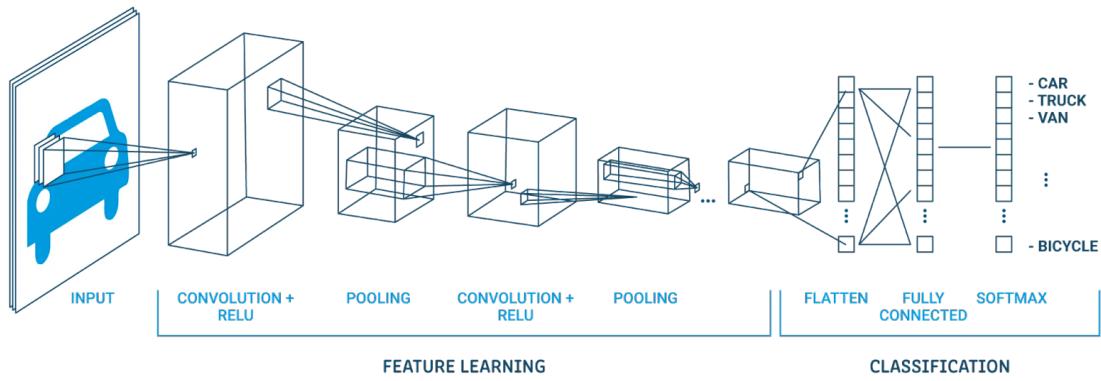


Figure 3.3: Deep Learning in Computer Vision [9]

3.5.4 Applications of Computer Vision

There are a lot of industries using computer vision and these are just a few examples [9]

medical imaging

It helps medical professionals interpret faster and diagnose abnormalities.

law enforcement and security

Like in surveillance and authentication.

self-driving machines

Like cars and robots.

gaming

Augmented reality and virtual reality.

pattern recognition

3.5.5 Some Technologies of Computer Vision

Because of the wide utility of computer vision and its benefits there are a lot of libraries and frameworks that facilitates a lot of the hard and repeated tasks, here we mention a few of them [9]

openCV

A python library for computer vision,

- super easy to use,
- a huge library of image processing algorithms,
- open source,
- works with GPUs

Tensorflow

made by Google and one of the most popular machine learning frameworks

- with a wide range of machine/deep learning algorithms,
- open source,
- GPU configured

PyTorch

made by Facebook, a neural network framework,

- used a lot by researchers,
- open source,
- works with GPUs

Caffe

a deep learning framework developed by Berkeley AI Research

- open source
- c++ based
- easy to use
- fast execution

Chapter 4

State Of The Art

4.1 Introduction

In this chapter we are going to present some recent work and research that was done regarding skin cancer detection and classification using machine learning, we are going to explore the various methods, tools, new ideas and challenges that were handled by researchers in the hope of getting a clear understanding of the problem and how to go about solving it depending on each one's conditions, requirements and goals.

4.2 Skin Cancer Detection and Classification Using Machine Learning

Proposed methodology

The proposed methodology in this article [35] uses a 6-step process (input data - preprocessing - segmentation - feature extraction - classification - output data).

Input data

Dermoscopic images from the ISIC (International Skin Imaging Collaboration) 2019 challenge containing 8 classes of skin lesions, and for simplicity reasons only 800 images out of, 25000 are used.

Preprocessing

Because of the heterogeneity of the input data, a preprocessing step is required to enhance the quality of images and remove irrelevant parts. The main techniques used here are gray scale conversion and the application of the Gaussian and median filter for noise removal and enhancement, and for the unwanted hair they applied the Dull Razor method (a preprocessing algorithm), as shown in figure 4.1.

Segmentation

Segmentation is used to extract the region of interest and for that they used a k-means clustering algorithm as shown in figure 4.2.

Feature extraction

For this they used 2 well known methods, ABCD method and GLCM. ABCD is used in dermatological applications and diagnosis for skin lesions such as melanomas, and

it is the abbreviation of Asymmetry, Border, Color and Diameter. Grey Level Co-occurrence Matrix (GLCM) is used for texture analysis, other features are also used in addition to these 2 methods for further classification such as Autocorrelation, correlation, Standard vector...etc.

Classification

For classification, they used MSVM (Multi-class Support vector machine) machine learning algorithm, they used training and testing ratios of 70:30 and obtained an accuracy of 96.25% and the confusion matrix shown in figure 4.3.

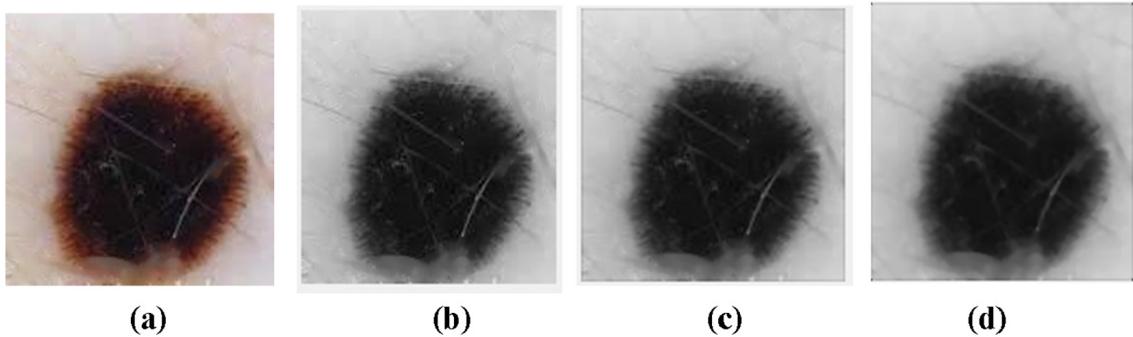


Figure 4.1: Preprocessing: (a)Dull razor image, (b) Gray scale image, (c) Gaussian filter, (d) Median filter.

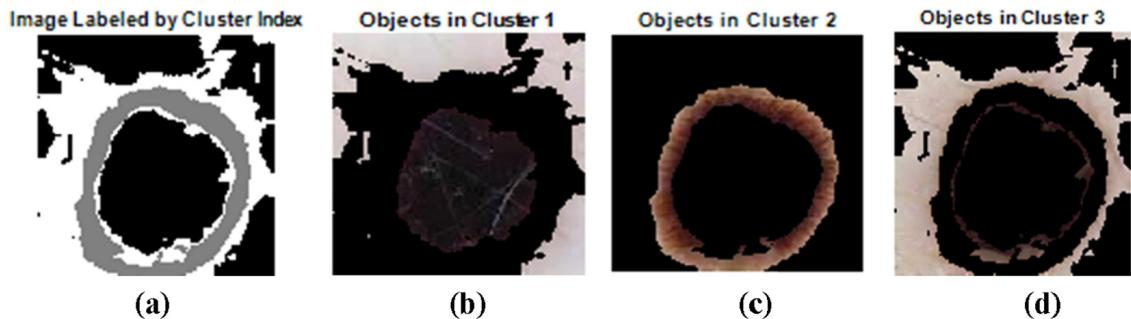


Figure 4.2: Segmentation: (a) Image labelled by cluster index, (b) Objects in cluster 1, (c) Objects in cluster 2, (d) Objects in cluster 3.

4.3 Finding reduced Raman spectroscopy fingerprint of skin samples for melanoma diagnosis through machine learning

This article [36] uses a new non-invasive approach to classify malignant and benign tumors, and that is by using Raman spectral data instead of images, Raman Spectroscopy is a way to analyze the chemical structure using light and vibrational energy modes of molecules [37].

Confusion Matrix								
True class	Actinic keratosis	2	1	1				1
	Basal cell carcinoma	58			1		1	
	Benign keratosis		60					
	Dermatofibroma	1	2		55		2	
	Melanocytic nevus				60			
	Melanoma				1	59		
	Squamous cell carcinoma		2				58	
	Vascular lesion	1			1		1	57
	Actinic keratosis	Basal cell carcinoma	Benign keratosis	Dermatofibroma	Melanocytic nevus	Melanoma	Squamous cell carcinoma	Vascular lesion
Predicted class								

Figure 4.3: Confusion Matrix

Data and method

dataset: For the dataset they brought 33 benign and 51 malignant samples and cut them into regular cuts of $2mm^3$, a laser was used to excite the samples to collect the Raman signals using special tools after this they acquired 436 Raman spectra (spectrum graphs $y=f(x)$ where x is frequency or wave number cm^{-1} and y is the intensity of scattered light). And they focused on the biological fingerprint spectral region from 800 to 1800 cm^{-1} .

Fluorescence background data pre-processing: Fluorescence is a radiation that is emitted by molecules after interacting with electromagnetic radiation and this could overshadow and disturb the study of Raman spectra, to deal with this noise they used a low frequency laser to lower the probability of fluorescence emissions and by this they could jump the preprocessing step.

Feature extraction

They divided the obtained spectrums into subsequences (local spectrums) and extracted some statistical measures from it such as arithmetic mean, standard deviation, derivative ...etc.

Results and discussion

These statistical features were then given to a machine learning classification algorithm, a complex decision tree implemented using lightGBM (open source software), other algorithms were also used such as K-nearest neighbors and XGBOOST (Extreme Gradient Boosting an open source software) but the best performance was obtained using lightGBM. Further research led them to only use the derivative as a feature and a spectral region from 896 to 1039 cm^{-1} because these two were proved to have the most discriminative information between malignant and benign tumors and by this they obtained a high performant model ($AUC \geq 0.97$) shown in figure 4.4.

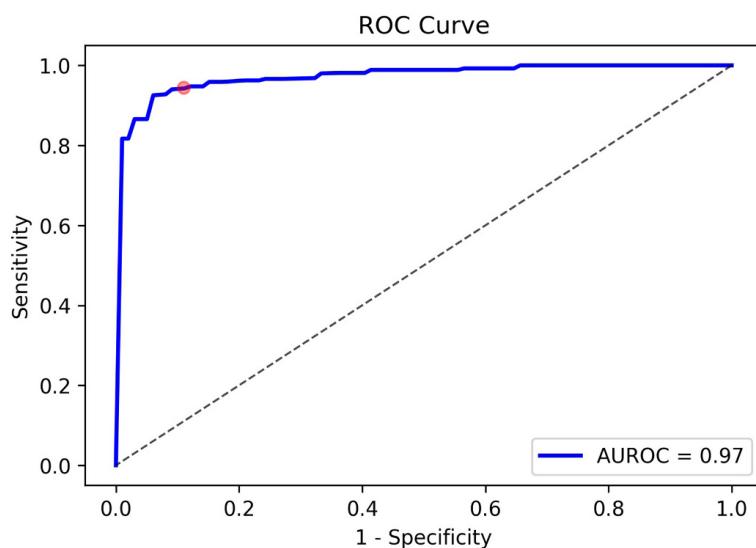


Figure 4.4: ROC

4.4 Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images

Summary

In this paper [38] the researchers are presenting a model driven approach to develop deep learning algorithms for detecting skin cancer by using a tool called DLS (deep learning studio) which is a software that allows you to build deep learning algorithms without being a specialist in programming languages, it presents a simple drag and drop interface for building models it also comes with desktop / cloud versions and community / enterprise editions with multi-GPU training and the possibility to obtain the code of the model, download the model and host it as a REST API (Representational state transfer Application programming interface), the interface dashboard is shown in figure 4.5.

Advantage

The advantage of this non-programmatic approach is for researchers and practitioners to be able to create and test their own models without the need for prior programming knowledge.

Application and Results

And then they proceed using this tool DLS to show its efficacy and ease of use, they have built and tested 5 models using famous CNN architectures squeezeNet, densenet, and inception v3. With model1 acquiring an AUC of 99.77% .

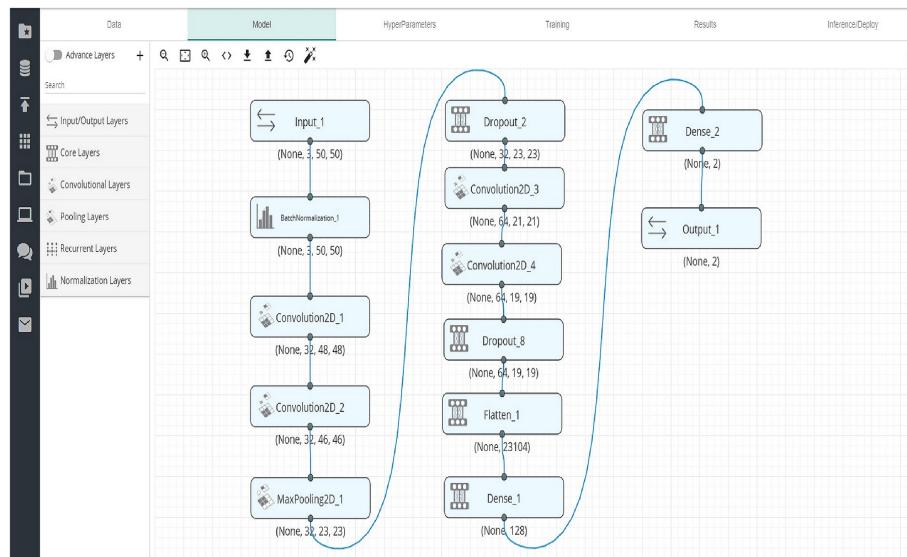


Figure 4.5: DLS Interface

4.5 The impact of patient clinical information on automated skin cancer detection

In this work [39] the researchers propose a new idea, which is the use of clinical information in addition to the image dataset and the study of this addition's effect on the deep learning model's performance.

Dataset

To build their hybrid dataset, they proposed a mobile application given to doctors and students to help collect the necessary data from Dermatological Assistance Program (PAD) dataset at the Federal University of Espírito Santo (UFES), which consists of images of the lesion, their clinical diagnosis and 8 clinical information based on common questions that dermatologists ask:

- age
- part of the body where the lesion is located,
- if the lesion itches,
- bleeds or has bled,
- hurts,
- has recently increased,
- has changed its pattern,
- and if it has an elevation

A total of 1612 images of 6 lesions.

Because the image dataset is imbalanced they used multiple strategies to overcome that such as, transfer learning (refining a pretrained model on their dataset), data augmentation, horizontal and vertical rotations, adjusting brightness...etc. and for the clinical data they used one-hot encoding (converting categorical data to augment the performance) which transformed the 8 features collected to an array of 28 values.

Training

They used 4 CNN architectures VGGNet-13/19-bn, ResNet-50/101, MobileNet, GoogleNet now a problem arose when trying to combine (by concatenation) clinical data with image features extracted by the CNN feature extractor because image features are far more great in size than clinical data, this imbalance is not good for the training and classification because the effect of image features will be greater than the clinical data, that is why they implemented an NN feature reducer on the extracted image features before combining it with the clinical data as shown in figure 4.6 and the classifier is another neural network that assigns the probabilities for each skin lesion.

Testing the effect of adding clinical data

They executed 2 scenarios for that, 1 using models trained only with images, 2 using models trained with images + clinical data then they calculated multiple performance metrics accuracy, balanced accuracy, weighted precision, weighted recall , weighted F1 score and area under the curve, and they found almost all models were

improved by 7% in almost all metrics and the best model ResNet-50 presented an $AUC \geq 95.8\%$.

Conclusion

Clinical information does make a difference when training ML models to classify skin cancer.

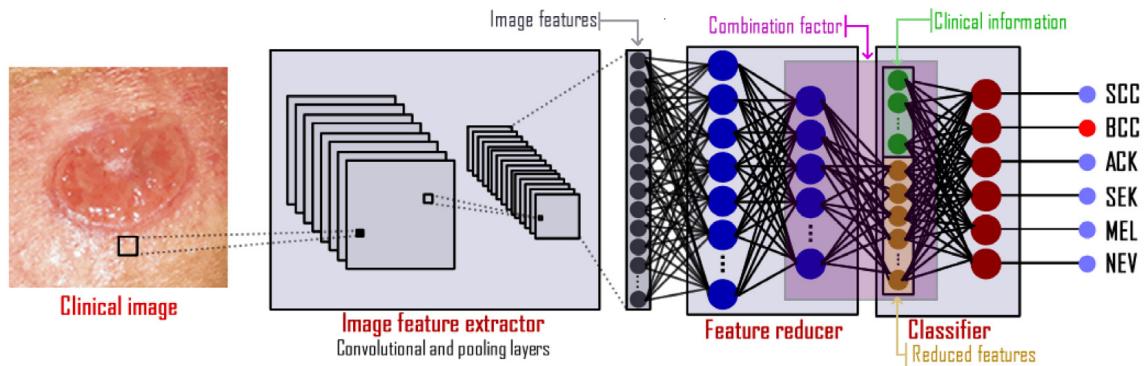


Figure 4.6: Model

4.6 An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features

In this work [40] they try to combine multiple texture features from famous methods such as ABCD, GLCM and LBP (local binary pattern) and pass all of these features to an ANN (artificial neural network) for learning.

Dataset

They preferred to use images captured using a dermatoscope because of their quality over images captured using a phone or normal camera, and they have obtained these images by combining 2 datasets: ISIC archive dataset (jpg format) and PH2 dataset (a dermoscopic image dataset in BMP format) they formed a unified dataset containing 1940 benign and 1448 malignant lesion images.

Preprocessing

Because the images are obtained from various sources, they needed to process them to standardize them in size, shape, format ...etc. And also to remove noise and enhance image quality using enhancement algorithms such as histogram equalization process that increases image contrast, and to remove body hair from the images using Maximum Gradient Intensity (MGI) algorithm .

Image segmentation

For better analysis and to remove unwanted parts they segmented the images to keep only the lesion area and for that they used a segmentation method called Otsu's Thresholding .

Feature extraction

They used ABCD (Asymmetry, Border, Color, Diameter), GLCM (energy, contrast, correlation, homogeneity) and LBP (local binary pattern used for textural analysis) as features to train their neural network.

Classification

A feed-forward neural network with backpropagation mechanism is used with the input layer receiving the extracted features and a hidden layer of 100 neurons and an output layer for the final result (1 is malignant and 0 is benign) with biased and weights initialized randomly, Levenberg-Marquardt training and optimization functions are used and while the performance function being Mean Square Error and 2 activation functions "tansig" for the hidden layer and "purelin" for the final output. The structure of the ANN is shown in figure 4.7.

Evaluation

For the evaluation of their classifier they calculated accuracy, specificity, sensitivity and precision shown in figure 4.8 where all the measures are $> 97\%$, and furthermore they also studied the effect of each feature on the discrimination process between benign and malignant lesion, and they found that the minimum sensitivity per single feature is 69%, minimum specificity per single feature is 73% and minimum accuracy per single feature is 71% which goes to show that all the used features are playing an important role in the classification process, and lastly they did a comparative evaluation between their work and previous works on the basis of extracted features which showed that more features implies higher performance rates, an example of that is the accuracy of previous works using a combination of some but not all features in (ABCD, GLCM, LBP) always presented an accuracy $< 97\%$.

In Conclusion:

The use of hybrid features provided a higher performant model in the detection and classification of benign and malignant melanoma skin cancer.

4.7 Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer

This article [41] talks about achieving interpretability in deep learning based systems, and the reason for this is that traditional machine learning models do outperform professionals in some scenarios, but we can't explain their output because they are like black boxes we don't know what is actually going on inside or why the model chose this output instead of another output, so it can not be trusted in high stake situations, for that interpretability of models became a thing in recent research articles, and there are mainly 2 ways in which we can attain interpretability, I.) we can use Model-Agnostic Methods for Interpreting any Machine Learning Model:

- like permutation feature importance,
- Partial Dependence Plots (PDPs),

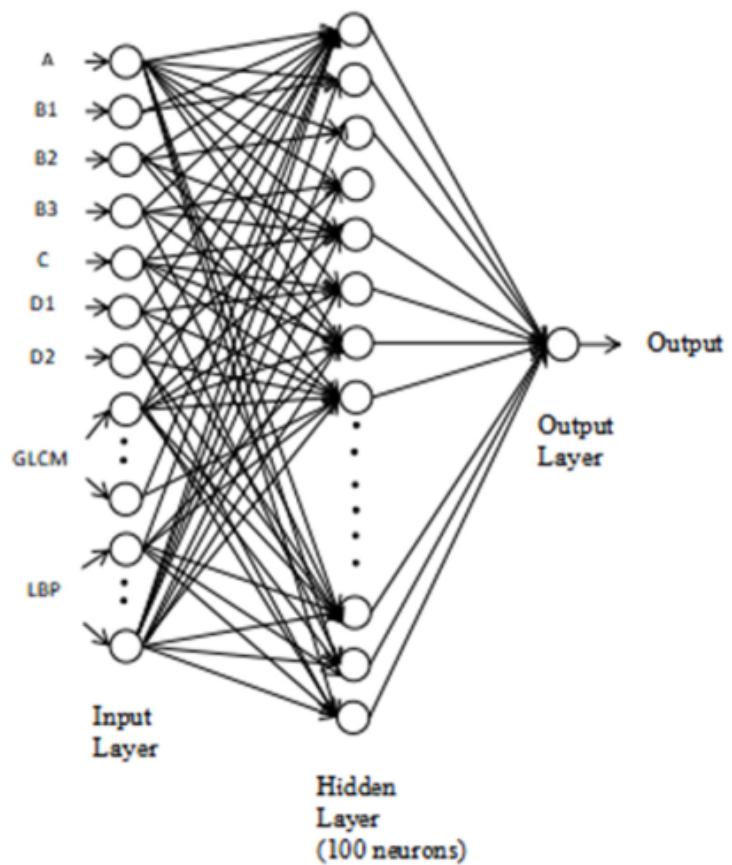


Figure 4.7: ANN Structure

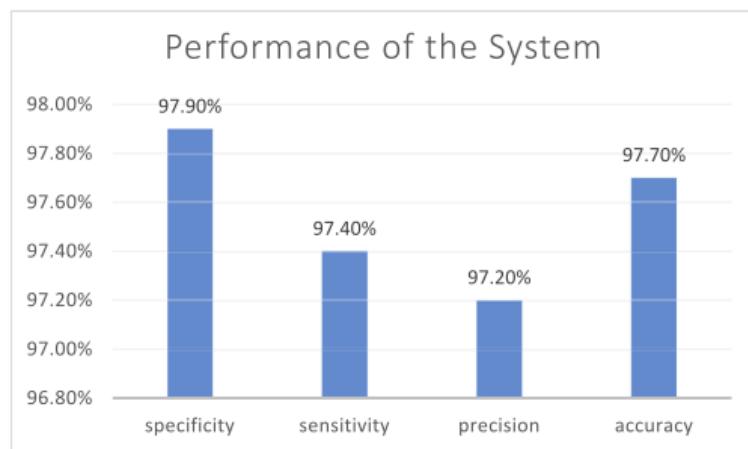


Figure 4.8: Evaluation Measures

- Individual Conditional Expectation (ICE) plots,
- global surrogate models,
- Local Interpretable Model-agnostic Explanations (LIME)
- Shapley Additive Explanations (SHAP)

To try and explain our model [42] which are statistical and visual ways used to understand a model, **II.)** There is another way which is the one used in this article and that is “naturally interpretable models”, which can be defined as models that try to solve the problem the way a human would, which means in the case of skin cancer, analyzing the whole tissue (the same way a doctor would) and not just cancerous regions of interest, and we can achieve this with semantic segmentation methods.

Dataset

MyLab Pathology provided them with their pre-existing images on non-melanoma cancers, which were taken by a microscope (one image of a cancer is the result of multiple microscopic images concatenated together) for punch, shave and excision biopsies (shave: the surface of the skin is removed with a sharp knife, punch: a round small part of the skin is removed) which meant high resolution images ($1\text{px}=0.67\mu\text{m}$) and each image was annotated by a pathologist to indicate important tissue sections that will be usefull in the classification process, any imbalance of classes was solved using augmentation (rotation and flipping...etc).

Models

Whole image segmentation

input: microscopic image see figure 4.9a.

output: $h \times w \times 12$ (12 probability distribution maps), see figure 4.9b.

the different tissue sections were colored

1. Glands (GLD)
2. Inflammation (INF)
3. Hair Follicles (FOL)
4. Hypodermis (HYP)
5. Reticular Dermis (RET)
6. Pap-illary Dermis (PAP)
7. Epidermis (EPI)
8. Keratin (KER)
9. Background (BKG)
10. BCC (Basal cell carcinoma)
11. SCC (Squamous cell carcinoma)
12. IEC (a very early treatable form of skin cancer)

To be fed to the segmentation model to train on semantic segmentation, this model was created using a combination of U-net-like architecture (U-net: a famous CNN architecture for biomedical image segmentation) and a pretrained headless ResNet50 network. Because of the high resolution of the microscopic images they were fed to the model in parts of 256×256 and 512×512 pixels.

Whole image classification

Input: output of segmentation ($h \times w \times [12 \text{ images}]$).

Output: 4 classes Healthy, BCC, SCC and IEC.

The output of whole image segmentation (which was a probability distribution for each pixel on the 12 tissue classes) was fed to a CNN to train as a classifier using Adam optimizer (used to accelerate the gradient descent algorithm) and a learning rate of 0.0001, with a ratio of 80:10:10 for training, validation and testing.

Results and discussion

The segmentation model achieved a per-pixel accuracy of 86% and overall class accuracy of 85%, they found that downgrading the image's size before training to 10 times less increases the accuracy but only by a little bit 2% which isn't much, but this information is still useful because it means that we can use low resolution images and still get a high performant model with less computational power.

The classification model achieved an accuracy of 93.6% over the 4 classes compared to other algorithms trained with the same data such as (Random Forest 87.2%, KNN 80.9%, Single-layer Perceptron 85.1%).

Conclusion

They showed that in order to build an interpretable model for skin cancer detection and classification you need to train your model the same way a real doctor would try to diagnose the skin cancer, and they did that by feeding and training the algorithm with the same data a dermatologist would use for diagnosis without ignoring anything such as hair, sweat glands ...etc. In the end they obtained a high performant model that is interpretable (which means that when a doctor sees the classification of the algorithm he can understand why it chose that classification) which will increase the possibility to use this approach in real life high stake scenarios, furthermore, because they used a diverse dataset they said their algorithm can be used for more routine work that a dermatologist would do such as assessing aggressiveness, depth, direction of growth and even calculating surgical margins (to know how much tissue to remove to guarantee that all cancerous cells are removed).

4.8 Comprehensive Comparative Information

In this section we are going to present a comprehensive comparative information for the various methods tested on benchmark datasets and their evaluation metrics that were extracted from previous literature work, we present these tables as a recap since all the methods pass through the same steps (dataset, preprocessing, feature extraction and prediction), we present a table 4.1 that recaps the methods mentioned above and 2 more tables with other methods 4.2 4.3, and it is also worth mentioning that the best methods used in the last 6 years (2016, 2021) using KPI-accuracy are SVM in machine learning and CNN in deep learning and the utilization shares are 54% for ML methods and 46% for DL methods [43].

Method	Dataset	Results
classification using MSVM	800 dermoscopic images ISIC of 8 lesion	accuracy 96.25%
classification using light-GBM (complex decision tree)	436 Raman spectra (x:frequency, y:intensity of scattered light) for benign and malignant	AUC > 97%
CNN famous architectures (implemented using deep learning studio)	Kaggle image dataset	AUC 99%
classification using multiple CNN architectures	images (1612 images of 6 lesions)+ clinical info	AUC 95.8% and clinical added a 7% to almost all performance metrics
binary classification ANN	hybrid features ABCD, GLCM, LBP from dermoscopic ISIC+PH2 datasets (1940 benign + 1448 malignant)	performance measures all > 97%, effect of each feature > 69%
segmentation to 12 tissue classes CNN (U-net, Resnet) and CNN classification (healthy and 3 lesions)	MyLab Pathology provided access to their pre-existing collection of skin cancer slides	93.6% accuracy compared to (Random Forest 87.2%, KNN 80.9%, Single-layer Perceptron 85.1%)

Table 4.1: A comparative table of the methods mentioned above in this article

Method	Dataset	Results
DCNN	PH2/ISBI 2016/ISBI 2017	98.4% on PH2 dataset, 95.1% on ISBI dataset and 94.8% on ISBI 2017 dataset
GLCM features to an SVM	ISIC	95% (Accuracy) 90% (sensitivity) 85% (specificity)
hybrid adaboost SVM	Skin Cancer and Benign Tumor Image Atlas-Contents	91.7% (Accuracy) 94.1% (sensitivity) 88.7% (specificity) 0.83% (Kappa)
ABCD features to an SVM	PH2	90.63% (Accuracy) 95% (sensitivity) 83.33% (specificity)
FCRN architecture	ISIC	0.912% (AUC) 0.857% (Accuracy) 0.490% (sensitivity) 0.961% (specificity) 0.729% (average precision)
ANN	ISIC	74.76% (Accuracy) 57.56% (validation loss)
CNN	Large collection of Multi-Source Dermatoscopic Images	75.2% (Accuracy) 0.71 (validation loss)

Table 4.2: A comparative table of the latest methods used in skin lesion detection. DCNN (Deep convolutional neural network), FCRN (Fully Convolutional Residual Networks) [13]

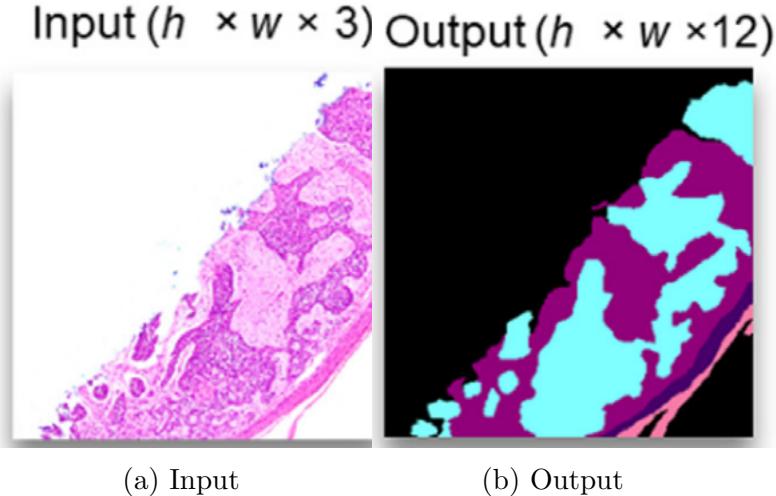


Figure 4.9: Whole segmentation model input and output

Beginning		
Method and Dataset	Results	Limitations
ResNet/ Hallym dataset 19,398 instances are used	For Basal Cell Carcinoma (BCC)/ 96% for asian dataset/ 90% for the instances of Caucasian for Melanoma –/ 96% for the instance of Asian/ 88% for Caucasian dataset	Model performance depends on patient ethnicity
Whale algorithm applied to optimize the CNN model/ Around 22,000 images of Dermquest and DermIS dataset were used	Achieved :/ Specificity 98%/ Accuracy- 94%/ Sensitivity-97%/ PPV- 90%	Results are unsatisfactory for non-Melanoma cases
ResNet/ Clinical images/ Clinical data are included	Achieved :/ 67.1% for clinical images/ 78.8% for clinical data and images	Missing values are not handled in clinical data/ Biopsy images are not included
ResNet/ 1279 dermoscopic instances are used	Achieved 89.2% for ResNet without dermatologist/ 94.3% for ResNet- with dermatologist	Clinical information to be included/ No standard evaluation criteria exist to measure classification efficacy
ResNet/ HAM10000 dataset, ISIC archive- 11,444 dermatoscopic instances are used	Mean accuracy of/ Physician 42.94/ CNN model – 81.59/ Fusion model – 82.59	Model outperformed for the trained dataset, but performance degraded for other datasets.
ResNet/ Dermnet dataset provides an instance of 23 categories of skin diseases	Achieved 97.1%	Able to identify few skin diseases like/ Acne , Rosacea, Hemangioma,/ Psoriasis, Seborrheic Dermatitis

Continuation		
Method and Dataset	Results	Limitations
Dragonfly optimized DNN model is assessed using existing techniques such as Support Vector Machine, ANN, and to display the efficiency of the system diverse evaluation criteria's are accuracy, sensitivity, and specificity are considered	Achieved/ Sensitivity – 84% / Specificity- 99.5% / Accuracy – 98.5%	Result generated for Melanoma is inadequate
CNN model used to categorized the skin cancer into Melanoma, nevi, basal cell carcinoma/ 11,444 dermoscopic images HAM10000 dataset, ISIC archive was used	89.2% of specificity and 56.5% of sensitivity attained by the dermatologists/ Whereas the CNN model achieved 98.8% of sensitivity, and specificity	Model biased and failed to classify skin lesion properly/ For melanoma images CNN model diagnosed it as nevi whereas dermatologists diagnosed it as melanoma/ model achieved the lowest specificity 94.2% for melanoma class
GoogleNet CNN model/ 4800 clinical images	Achieved/ Sensitivity 96.3% Specificity 59.5%	Model behavior depends on skin color tone, and Model biased, sensitivity and specificity dropped for Caucasian patients/ No standard evaluation criteria exist to measure classification efficacy
DenseNet, ResNet model results compared with Dermatologist/ 10,135 dermoscopy images of HAM10000: 10015, PH2: 120 data sets are used	For Melanoma and BCC – achieved 94.40% for ResNet, 99.30% for DenseNet whereas Dermatologist achieved – 82.26% and 88.82% accuracy	Model biased for the diverse dataset/ Model behavior depends on skin tone colour

Table 4.3: Another comparative table of the latest methods used in skin lesion detection and their limitations [12]

Chapter 5

Results and Discussion

5.1 Introduction

In this chapter we are going to compare the various discussed methods depending on multiple criteria and different point of views, in other terms the comparison will be based on the different datasets used, the different methods and their performance, interpretability and applicability in real life scenarios.

5.2 Datasets

These are the famous dataset used and some problems encountered using them:

- **ISIC 2019**
- **Kaggle**
- **PH2**
- **Hallym**, *problem* performance depends on ethnicity
- **Dermquest**, *problem* unsatisfactory for non-melanoma
- **DermIS**, *problem* unsatisfactory for non-melanoma
- **HAM10000 ISIC 2018**, *problem* biased for diverse dataset (but works good for binary classification)
- **Dermnet**, *problem* able to identify only few skin lesions
- **ISBI**
- **Atlas Contents**
- Compile your own dataset if you have a new idea (such as: Raman spectra, clinical data+image)
- A general problem met with all datasets is the preprocessing and cleaning (such as smoothing and hair removal)

5.3 Extracted Features

After talking about these datasets, we are going to talk about the features extracted from them, which are GLCM, ABCD, LBP, Autocorrelation, correlation, Standard vector and other statistical features (such has arithmetic mean, standard deviation for Raman spectra) all of these features are proven to play an important role in the classification process but depending on your situation and the dataset you use you could find for example a single feature that presents the same performance as the group of features (like in the case of [36] where they used the derivative only) and you can simply check that by training the model with different combination of features until you find an important difference or by using statistical methods that tells you the degree of discrimination of each feature.

5.4 Methods

These are the used algorithms:

- SVM
- hybrid adaboost SVM (Adaboost: adaptive boosting, it can be used with various algorithms to increase the performance)
- MSVM (multi-class)
- lightGBM (complex decision tree, open source software)
- ANN
- DNN (deep neural networks) with Dragonfly optimization algorithm
- CNN and its famous architectures
 - DCNN (deep CNN)
 - u-net for medical image segmentation
 - Resnet
 - FCRN (Fully Convolutional Residual Networks)
 - CNN + Whale algorithm applied for optimization
 - GoogleNet
 - DenseNet
 - MobileNet
 - Inception v3
- DLS (deep learning studio)

An overall view of literature tells us that the performance metric of both deep learning and machine learning is the same, with deep learning being slightly better. The best of both sides are: (MSVM reaching an accuracy of 96.25% and lightGBM achieving an AUC >97%) and (ANN performance metric, accuracy, specificity... all being above 97% and inceptionv3 CNN architecture reaching and AUC of 99%).

5.5 Applicability

When we talk about applicability we need to consider 4 things (place, dataset, method, interpretability)

I- where do plan on using your algorithm: in a local clinic or a national hospital or internationally?

II- Is the dataset used compatible with the place off applicability, in other words if you plan on applying your method locally your training set need to be local also, and not from another country (so famous international datasets won't be sufficient) or if you plan on using your method in a national/international level your dataset need to be comprehensive and balanced in all ways: skin color tones, ethnicity, age, sex...etc.

III- Thirdly your method need to take into consideration the available dataset and the place of applicability, do you want multiclass classification or binary classification ?, is the place you're planning on applying your method in capable of acquiring the needed technology ? (like in the case of Raman Spectroscopy, not all clinics are capable of having such a technology), is your method of diagnosis compatible with cultural specifications (invasive or non-invasive)?.

IV- Lastly, is your method interpretable? Because most of the mentioned algorithms above are not naturally interpretable (black boxes), so unless they get interpreted, they can't be trusted in high stakes scenarios.

5.6 Line Of Reasoning

Here we present a line of reasoning that could help future researchers and contributors to better navigate this field, from different point of views

Dataset

Consider the different datasets mentioned above, each with its benefits and limitations, some are underrepresenting of skin color tones or sex or age

solution: combination of multiple datasets to cover each one's limitations

challenge: pre-processing to eliminate heterogeneity and imbalances

Use a hybrid dataset by combining different data that can help in the diagnosis process,

- *bear in mind:* that not all data can be discriminative, working with the wrong feature could lead you in the opposite way

An example of such dataset is the use of images+clinical information,

- *problem:* the image features could overshadow the clinical information
- *solution:* use feature reducer NN on the image features before combining it with clinical data

Use a novel dataset you can come up with your own new idea, as we've seen with

- Raman spectra
- microscopic images
- But don't forget the labeling process, that needs expertise

Extracted Features

There are famous features used in this field which are mentioned above, You can use either one of them or a combination but take into consideration that depending on your situation and method not all features could be as helpful in the discrimination (classification) process.

So study the effect of each feature if you want to use hybrid features so you don't waste resources and time on a non-discriminative feature.

Patient Experience and Target Users

Consider the experience of the patients, is your diagnosis system going to be invasive or non-invasive?, is it affordable or not? Does it require some special tool to use which is not available to all target users?.

Methods

famous and performant algorithms

- machine learning
 - needs less data
 - MSVM
 - lightGBM using a performance boosting meta-algorithm such as Adaboost
- deep learning
 - requires a huge amount of data
 - +optimization algorithms like: Dragonfly, Whale
 - ANN
 - CNN
 - * inceptionv3
 - * Resnet
 - for multiclassification use transfer learning
 - for binary classification you can train from scratch

No programming experience

Open software made it possible for non programmers to participate in the process of building machine learning models, an example of said software is (Deep Learning Studio).

Your available resources and budget

Can you afford a server to train your algorithm?, if not are your available computational resources sufficient to build a big and performant model?.

interpretability

Lastly, if your model is intended for real world application, you really need to consider the interpretability aspect, because if the end result algorithm isn't interpretable, the probability of using your solution in high stake scenarios (and all health issues are high stake) is very low.

Chapter 6

Our Contribution

6.1 Introduction

Melanoma is a type of skin cancer, develops in the cells (melanocytes) that produce melanin — the pigment that gives your skin its color, The exact cause of all melanomas isn't clear, but exposure to ultraviolet (UV) radiation from sunlight increases your risk of developing melanoma. [44]

Melanoma is more dangerous because of its ability to spread to other organs more rapidly if it is not treated at an early stage. [45]

At present, CNN has achieved very good performance in the field of computer vision compared to other machine learning algorithms, such as object detection, image recognition, classification, etc.

”Convolutional Neural Network (CNN) is a type of deep learning model for processing data that has a grid pattern, such as images, which is designed to automatically and adaptively learn spatial hierarchies of features. CNN is a mathematical construct that is typically composed of three types of layers (or building blocks): convolution, pooling, and fully connected layers. The first two, convolution and pooling layers, perform feature extraction, whereas the third, a fully connected layer, maps the extracted features into final output, such as classification. A convolution layer plays a key role in CNN, which is composed of a stack of mathematical operations, such as convolution, a specialized type of linear operation.” [46]

Because of the difficulty of detecting melanoma cancer in an ordinary way, CNN is used to classify melanoma skin cancer.

Research on the classification and detection of melanoma cancer by various methods has been carried out. In 2016 there was a paper entitled ”Deep Residual Learning for Image Recognition” using the ResNet architecture. The paper was a winner at the 2015 ILSVRC (Imagenet competition). [47]

6.2 Proposed Convolutional Neural Network Model

The main aim of this implementation is to detect melanoma skin cancer through RGB images, to achieve this, we build a deep learning model that is capable of extracting features from the given dataset.

After delving into many articles and studies, we have found that the best convolutional neural network model we can suggest in this case is resnet50, and so we are going to implement it from scratch. As shown in figure 6.1

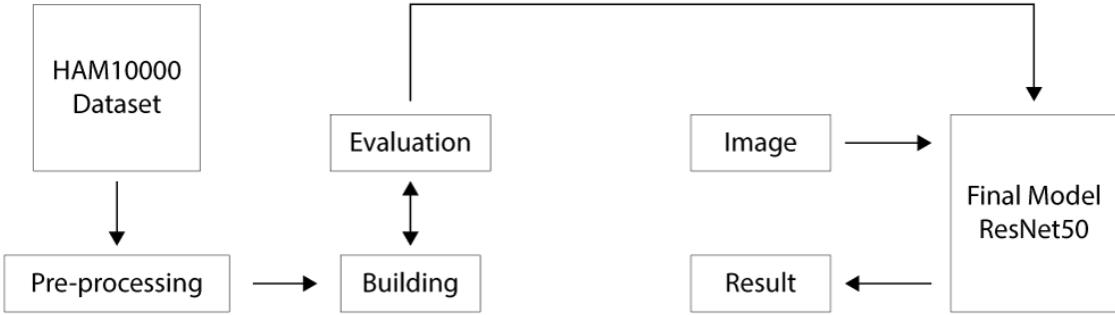


Figure 6.1: proposed architecture which we have used for melanoma recognition

6.3 Dataset (MNIST- HAM10000)

The ISIC archive is the largest public database for dermatoscopic image analysis research, and where the original HAM10000 was made available. [10]

The HAM10000 dataset is composed of 10.015 dermatoscopic images of pigmented skin lesions. The data was collected from Australian and Austrian patients. Two institutions participated in providing the images: Cliff Rosendahl in Queensland, Australia, and Medical University of Vienna, Austria. According to the authors, seven classes are defined on this dataset, where some diagnoses were unified into one class for simplicity. Information regarding patient age, sex, lesion location and diagnosis is also provided with each image. [10]

The dataset has been collated and published by Tschandl, P., Rosendahl, C. & Kittler, H. [10] A sample of each type of skin lesion present in the dataset is demonstrated in the figure 6.2. And the distribution of lesions is show in figure 6.3

6.4 Pre-processing

Before starting the model training process we need to process the dataset, as we learned earlier the dataset consists of around 10015 labeled images for 7 different types of skin lesions, but in our case, we want to get images classified on only two types of skin lesions (Melanoma and Not melanoma). We do this in several steps:

Data cleansing:

In this step, we remove unused and damaged data, also repair data that is incorrectly formatted.

Data separation:

After cleansing the data set, we separate the data set into two types of skin lesions by changing the data label for the non-melanoma types to non-melanoma, and we keep the data label for the type of melanoma as it is.

Data balancing:

When reclassifying the data set, we notice that the data set is numerically unbalanced. To solve this problem, we increase the number of images of the melanoma type by rotating,



Figure 6.2: A sample of each type of skin lesion [10]

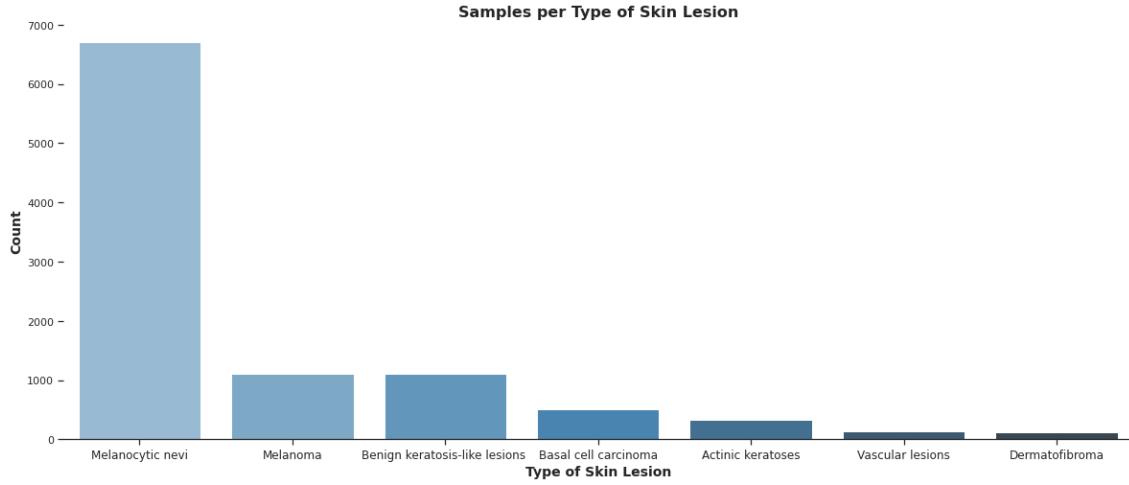


Figure 6.3: This count plot helps to understand the distribution of the data. [10]

cropping and scaling. As for the non-melanoma type, we reduce the number of images by randomly selecting a specified number of images.

Image resizing:

In this step, we reduce the image size to $75*100$ to speed up the training process of the deep learning model.

data splitting : Before the data set becomes usable, we divide it into two parts, the first part is the training set with 80 percent, and the second part is the test set with 20 percent

The diagram 6.4 helps to understand these steps

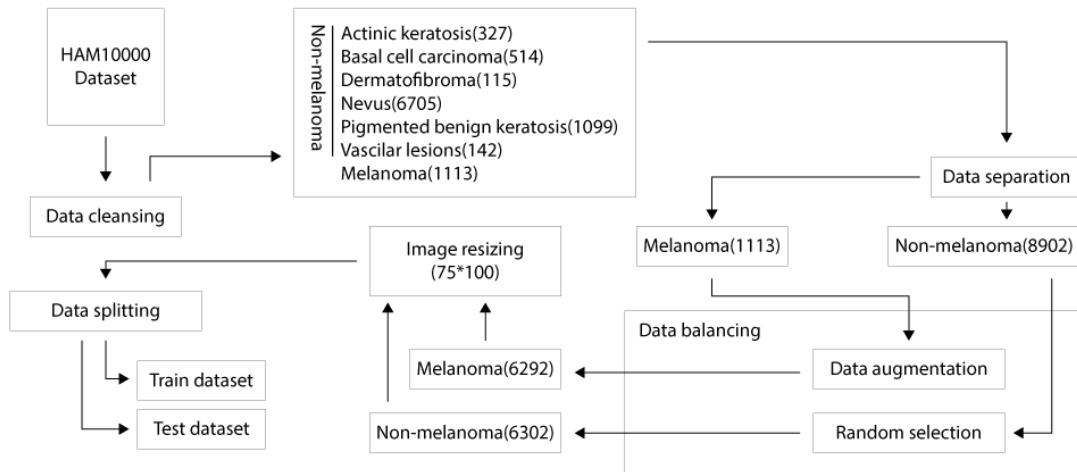


Figure 6.4: Pre-processing

6.5 Experimental results

To judge the performance of the model for the task of predicting skin lesions, we use several evaluation metrics to evaluate our model. to ensuring that our model is operating correctly and

optimally.

When the model was trained for 30 epochs, it was observed that the accuracy for both the training and test data started with rather large values and continued to increase slowly from epoch 4 until it reached its peak in epoch 30 (which goes to show that our chosen architecture is very good for this task). The test accuracy reached 93 percent and the training accuracy was 97 percent.

The plot for the accuracy and loss obtained during the training and testing process is shown in Figure 6.5

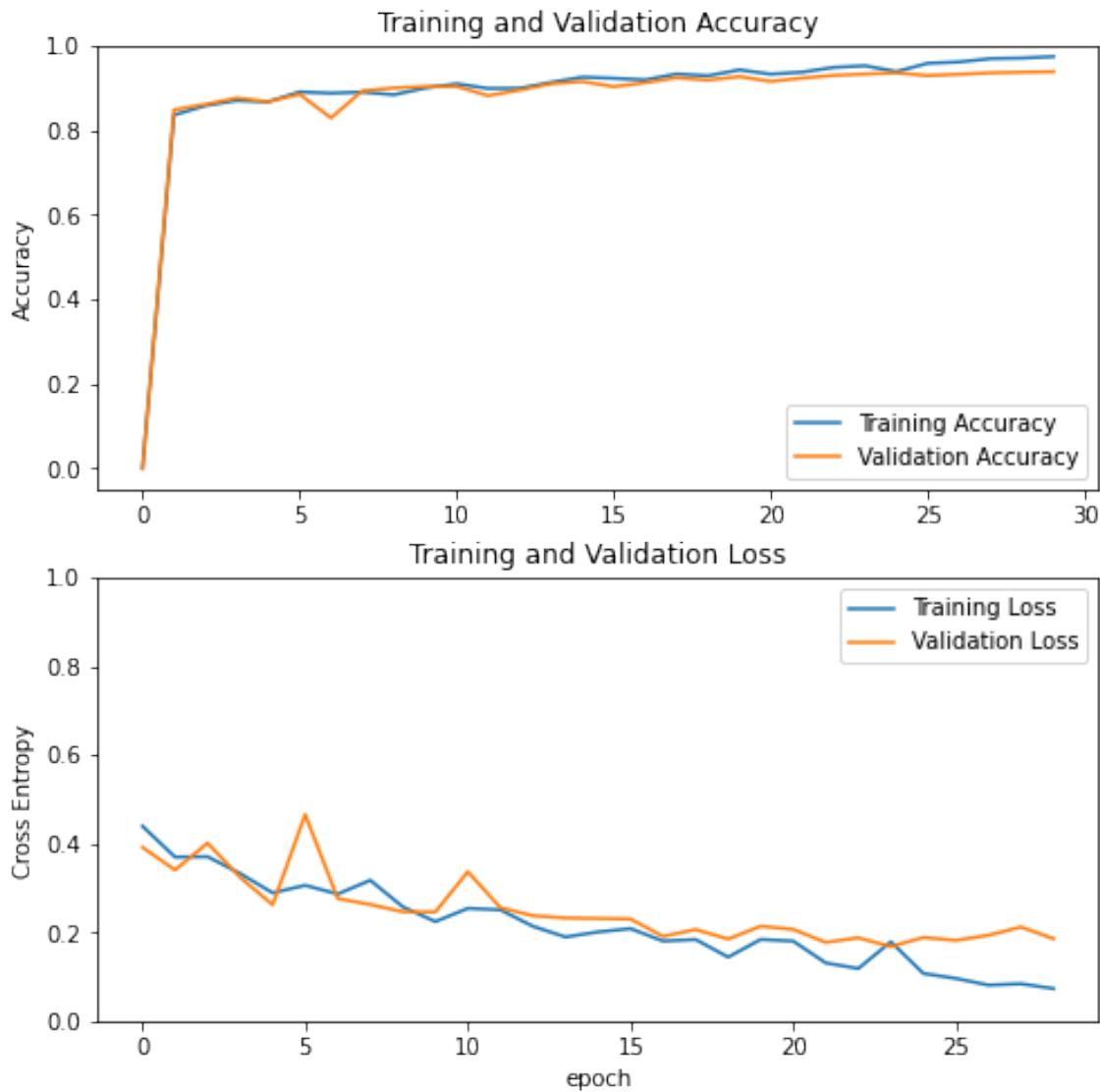


Figure 6.5: Accuracy and Loss

The table 6.1 also includes several other measurements that we used in evaluating our model

Classes	Precision	Recall	F1-score	Support
Non-melanoma	0.95	0.93	0.94	1293
Melanoma	0.93	0.95	0.94	1226

Table 6.1: Evaluation Measures

Chapter 7

Conclusion

Skin cancer is one of the most dangerous and widespread cancers in the world, it can occur to anyone of any age and of any race, just by standing in the sun for too long will increase your chances of getting it. That is why an early diagnosis will save your life. Here where the CAD's (computer aided diagnosis) systems would play an important role by implementing machine learning algorithms that could recognize, detect and classify various skin lesions.

The research in this field has brought a lot of advancements both in machine learning and deep learning and has produced a lot of models that reach an accuracy above that of an expert dermatologist, but we can't do without the expertise of doctors because of the interpretability problem since most models and research contributions don't take into consideration this aspect of algorithms.

There are a lot of datasets available online, most of them are missing some aspect or another, but that can be solved combining different datasets while taking into consideration the imbalances and heterogeneity of said datasets.

In this article we did a recap of most famous methods, algorithms and datasets used, and we did a little comparison between them based on performance metrics, dataset characteristics and applicability in real world scenarios, reading this article will help future researchers who are planning on contributing in this field.

We conclude that there are a lot of things to take into considerations when building a machine learning model, there is no right or wrong way but in the hopes of drawing a guidance map for future contributors, we made a line of reasoning that could guide them in the right direction.

Bibliography

- [1] Skin, cleveland clinic. <https://my.clevelandclinic.org/health/articles/10978-skin>.
- [2] Hair and human skin anatomy. <https://www.vecteezy.com/vector-art/3190201-hair-and-human-skin-anatomy>.
- [3] Skin barrier basics for people with eczema, national eczema association. <https://nationaleczema.org/what-is-my-skin-barrier/>.
- [4] Vitamin d and the skin: Focus on a complex relationship. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4642156/>.
- [5] What is a mutation? <https://www.yourgenome.org/facts/what-is-a-mutation>.
- [6] Chromothripsis. <https://en.wikipedia.org/wiki/Chromothripsis>.
- [7] Differences between a malignant and benign tumor, verywell health. <https://www.verywellhealth.com/what-does-malignant-and-benign-mean-514240>.
- [8] Yulia Gavrilova. Artificial intelligence vs. machine learning vs. deep learning: Essentials, serokell.io. <https://serokell.io/blog/ai-ml-dl-difference>, 2020.
- [9] Vihar Kurama. Ml-based image processing, nanonets. <https://nanonets.com/blog/machine-learning-image-processing/>, 2021.
- [10] JULIANA NEGRINI DE ARAUJO. Ham10000: Analysis and model comparison. <https://www.kaggle.com/code/jnegrini/ham10000-analysis-and-model-comparison/notebook>, 2021.
- [11] Cancer, world health organization. <https://www.who.int/news-room/fact-sheets/detail/cancer>.
- [12] S. Naresh Kumar and B. Mohammed Ismail. Systematic investigation on multi-class skin cancer categorization using machine learning approach. *Materials Today: Proceedings*, 12 2020.
- [13] Tanzila Saba. Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges, 9 2020.
- [14] Elin Jørgensen, Giulia Lazzarini, Andrea Pirone, Stine Jacobsen, and Vincenzo Miragliotta. Normal microscopic anatomy of equine body and limb skin: A morphological and immunohistochemical study. *Annals of Anatomy*, 218:205–212, 7 2018.
- [15] Joseph M. Abdo, Nikolai A. Sopko, and Stephen M. Milner. The applied anatomy of human skin: A model for regeneration, 3 2020.
- [16] Sarah A Mohamed and Rachel Hargest. Surgical anatomy of the skin, 2021.

- [17] Eliana B. Souto, Joana F. Fangueiro, Ana R. Fernandes, Amanda Cano, Elena Sanchez-Lopez, Maria L. Garcia, Patrícia Severino, Maria O. Paganelli, Marco V. Chaud, and Amélia M. Silva. Physicochemical and biopharmaceutical aspects influencing skin permeation and role of sln and nlc for skin drug delivery, 2 2022.
- [18] What is cancer?, national cancer institute. <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>, 2021.
- [19] Mahsa Alidoust Saharkhiz Lahiji and Fatemeh Safari. Potential therapeutic effects of hamscs secretome on panc1 pancreatic cancer cells through downregulation of sgk269, e-cadherin, vimentin, and snail expression. *Biologicals*, 2022.
- [20] Rachita Pandya, Kaitlyn Grace San Diego, Talha Shabbir, Arnav P. Modi, Justin Wang, Joseph Dhahbi, and Sanford H. Barsky. The cell of cancer origin provides the most reliable roadmap to its diagnosis, prognosis (biology) and therapy. *Medical Hypotheses*, 157, 12 2021.
- [21] Skin cancer 101, skin cancer foundation. <https://www.skincancer.org/skin-cancer-information/>, 2022.
- [22] Skin cancer (non-melanoma), nhs. <https://www.nhs.uk/conditions/non-melanoma-skin-cancer/>, 2020.
- [23] By Mayo Clinic Staff. Skin cancer, mayo clinic. <https://www.mayoclinic.org/diseases-conditions/skin-cancer/symptoms-causes/syc-20377605>, 2020.
- [24] Cancer, wikipidia. <https://fr.wikipedia.org/wiki/M%C3%A9lanome>, 2022.
- [25] Squamous cell carcinoma, walk-in dermatology. <https://walkidermatology.com/patient-education/squamous-cell-carcinoma/>.
- [26] First immunotherapy for advanced basal cell carcinoma approved in us, european pharmaceutical review. <https://www.europeanpharmaceuticalreview.com/news/142524/libtayo-first-immunotherapy-for-advanced-basal-cell-carcinoma-approved-in-us/>, 2021.
- [27] Skin cancer prevention, skin cancer foundation. <https://www.skincancer.org/skin-cancer-prevention/>, 2022.
- [28] By Mayo Clinic Staff. Skin cancer diagnosis, mayo clinic. <https://www.mayoclinic.org/diseases-conditions/skin-cancer/diagnosis-treatment/drc-20377608>, 2020.
- [29] Artificial intelligence, built in. <https://builtin.com/artificial-intelligence>.
- [30] JAKE FRANKENFIELD. Turing test, investopedia. <https://www.investopedia.com/terms/t/turing-test.asp>, 2022.
- [31] machine learning, built in. <https://builtin.com/machine-learning>.
- [32] Deep learning: An in-depth look into the ai-based tech, built in. <https://builtin.com/artificial-intelligence/deep-learning>.
- [33] What is computer vision?, ibm. <https://www.ibm.com/topics/computer-vision>.
- [34] machine learning in computer vision, full scale. <https://fullscale.io/blog/machine-learning-computer-vision/>, 2019.

- [35] M. Krishna Monika, N. Arun Vignesh, Ch Usha Kumari, M. N.V.S.S. Kumar, and E. Laxmi Lydia. Skin cancer detection and classification using machine learning. volume 33, pages 4266–4270. Elsevier Ltd, 2020.
- [36] Daniella Castro Araújo, Adriano Alonso Veloso, Renato Santos de Oliveira Filho, Marie Noelle Giraud, Leandro José Raniero, Lydia Masako Ferreira, and Renata Andrade Bitar. Finding reduced raman spectroscopy fingerprint of skin samples for melanoma diagnosis through machine learning. *Artificial Intelligence in Medicine*, 120, 10 2021.
- [37] What is raman spectroscopy?, edinburgh instruments ltd. <https://www.edinst.com/blog/what-is-raman-spectroscopy/>.
- [38] Mohammad Ali Kadampur and Sulaiman Al Riyae. Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images. *Informatics in Medicine Unlocked*, 18, 1 2020.
- [39] Andre G.C. Pacheco and Renato A. Krohling. The impact of patient clinical information on automated skin cancer detection. *Computers in Biology and Medicine*, 116, 1 2020.
- [40] Priyanti Paul Tumpa and Md Ahsan Kabir. An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features. *Sensors International*, 2, 1 2021.
- [41] Simon M. Thomas, James G. Lefevre, Glenn Baxter, and Nicholas A. Hamilton. Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer. *Medical Image Analysis*, 68, 2 2021.
- [42] Hennie de Harder. Model-agnostic methods for interpreting any machine learning model, towardsdatascience. <https://towardsdatascience.com/model-agnostic-methods-for-interpreting-any-machine-learning-model-4f10787ef504>, 2020.
- [43] Deepak Painuli, Suyash Bhardwaj, and Utku köse. Recent advancement in cancer diagnosis using machine learning and deep learning techniques: A comprehensive review, 7 2022.
- [44] By Mayo Clinic Staff. Melanoma, mayo clinic. <https://www.mayoclinic.org/diseases-conditions/melanoma/symptoms-causes/syc-20374884>, 2022.
- [45] Melanoma overview, skin cancer foundation. <https://www.skincancer.org/skin-cancer-information/melanoma/>, 2022.
- [46] Yamashita Rikiya, Mizuho Nishio, Do Richard Kinh Gian, and Kaori Togashi. Convolutional neural networks: an overview and application in radiology. *Insights Into Imaging*, 9, 2018.
- [47] Arief Budhiman, Suyanto Suyanto, and Anditya Arifianto. *Melanoma Cancer Classification Using ResNet with Data Augmentation*. 2019.