Understanding Skin Cancer: A Guide to Basal Cell, Squamous Cell, and Melanoma

Chopra

Contents

Chapter 1

Introduction

This chapter provides an introduction to this work. The motivation is explained in Section 1.1. The objectives this work intends to achieve are laid out in Section 1.2. Lastly, the organization of this work is shown in Section 1.3.

1.1 Motivation

Cancer is a dangerous disease with a high mortality rate when diagnosed at later stages. When it comes to skin cancer, there exist a variety of different types. Designated after the cell that originated the cancer, Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma are the first and second most common skin cancers respectively, while Melanoma is among the most dangerous skin cancers. Skin cancer has been consistently rising in these last few years, with cases of Melanoma predicted to rise from 287,723 in 2018 to 340,271 in 2025 [55]. In Portugal alone, there are over nine thousand cases of Basal Cell Carcinoma and two thousand cases of Squamous Cell Carcinoma [21], while Melanoma has one thousand cases [20].

Melanoma is given more attention due to the considerable danger it presents. Melanoma often spreads to the lymph nodes [62], being then able to spread faster to other parts of the body. While the estimated five-year survival rate from Melanoma when detected early is about 99%, the survival rate falls to 65% when the disease reaches the lymph nodes and 25% when the disease metastasizes to distant organs [26].

There has been a rise in the number of specialists (Dermatologists), however, in the United States, the number remains insufficient to provide adequate care for the general population [32]. With dermatologists also being called upon to take care of many other skin conditions, a solution is needed to address the rising number of skin cancers and insufficient number of dermatologists.

Screening has been used to attempt to identify cancerous lesions at an early stage, or before symptoms manifest, while methods based on common characteristics observed in previous cancers have been proposed and adopted. However, the screening process is time consuming, while the

classification of a lesion as cancerous remains mostly dependent on the dermatologist expertise.

With the classification process being largely visual, images are usually taken. Computer-Aided Diagnosis (CAD) systems that aid in the classification of skin cancers have been under investigation since 1987 [52]. These usually take the obtained images as input to predict a classification. Advances in technology have enabled the easier adoption of a dermatoscope, a device that enables the capture of dermoscopic images. As usually only dermatologists use dermatoscopes, clinical (macroscopic) images are still taken and kept on the medical records. Regardless, despite dermoscopic images being cleaner and having a higher level of detail than a clinical image, clinical images still contain relevant information. Even dermatologists obtain better results when they have access to various modalities such as dermoscopic images, clinical images and metadata, [35], however research on CAD systems that utilize multiple modalities is small in comparison with the larger research on CAD systems.

Recent developments have shown that deep learning is a promising technology, having surpassed human performance in visual tasks such as playing Atari games, board games like Go and object recognition [24]. This technology has been integrated in various CAD systems, obtaining comparable results to more classical.

1.2 Objectives

This thesis was developed in Fraunhofer Portugal in collaboration with Sciences Faculty of University of Porto. This work is part of project "DERM.AI: Usage of Artificial Intelligence to Power Teledermatological Screening", with reference DSAIPA/AI/0031/2018, and supported by national funds through 'FCT—Foundation for Science and Technology, I.P.'. Derm.AI project aims to contribute to processes optimization between Primary Care Units and Dermatology Services of the National Health Service, namely through the integration of a mobile application to acquire macroscopic skin lesion images and the development of AI-powered Risk Prioritization and Decision Support platform.

To support this project, this work studies the classification of skin cancer lesions using multiple modalities from the domain. The impact of using each modality, or a combination of them, to the quality of the lesion classification is investigated. Several techniques are investigated regarding their efficacy in improving the results obtained from fusing the modalities. The effects of using a simple architecture are also investigated, determining its viability. As various tests are performed, the simple architecture also facilitate a faster testing of these various subjects (modalities and techniques). Lastly, this work is heavily influenced by the work in [44], where, in addition to multitasking, several combinations of modalities are performed in the same model. This methodology is adapted and investigated. These objectives can be summarised in the following points:

1. Investigate the impact of each modality to the quality of the skin cancer classification.

2. Investigate the impact of the fusion of the modalities in the quality of the skin cancer classification.
3. Investigate the impact of several techniques to the quality of the skin cancer classification.
4. Determine the viability of a simple architecture in skin cancer classification.
5. Investigate the impact of performing several combinations of modalities on the same model.

1.3 Organization of the work

This work is made up of six chapters. This first chapter is the introduction to the work. The following two chapters provide further context, with Chapter 2 providing a full description of dermatology, skin cancer, screening and the available datasets. Chapter 3 contains a review of the state of the art regarding CAD systems. Topics such as machine learning and deep learning are introduced and explained in this chapter, followed by an overview of the methods utilized in various CAD systems with a focus works with multimodality. A set of experiments was designed to study various aspects of lesion classification using single modality and multiple modalities. The methodology and rationale for these experiments are presented in Chapter 4. Results are described, analysed and discussed in Chapter 5. Finally, Chapter 6 concludes this work and presents perspectives of future work.

Chapter 2

Skin cancer

This chapter introduces main concepts related with dermatology and skin cancer. Beginning with defining the terminology used in dermatology in Section 2.1 and teledermatology in Section 2.2. The most important skin cancers are described and the most common benign lesions are explained in Section 2.3. Screening is explained in Sub-Section 2.3.2 with the available datasets being briefly described and summarized in Section 2.4.

2.1 Dermatology

Dermatology is the field of medicine specialised in the management of skin conditions. The skin is the largest organ of the body, being composed by three layers, the epidermis, dermis and hypodermis.

The epidermis is the outermost layer of the skin, it houses cells such as squamous, basal, melanocytes and merkel cells, among others. Its job is to protect the body from the environment. The dermis stands below the epidermis layer, containing tough connective tissue, hair follicles, and sweat glands. While the deepest layer, hypodermis, is made of fat and connective tissue, functioning as an insulator and shock-absorber [61].

The scope of dermatology involves all kinds of skin conditions, from cosmetic applications to inflammatory, inherited, environmental, occupational and malignant skin diseases. Its specialists are the dermatologists. Their patients belong to all ages and sexes, and while some injuries require assistance, most of the time the patients will not stay overnight at the hospital (outpatients) [37]. Usually, the patients are redirected to a dermatologist by a general practitioner when suspicious skin lesions are found during a consultation.

Of the many conditions that dermatology oversees, skin cancer is one of the most dangerous. Depending on the situation, the treatment can range from chemotherapy to a biopsy, an operation where the affected part is removed. If the cancer is not diagnosed and treated in time, it will spread through the body (metastasis), resulting in the eventual passing of the patient.

Similar to other types of cancer, it is of great importance to confirm and accurately classify the cancer type. Early detection of cancer increases the chances of survival and reduces the morbidity and cost of the treatment [84]. When compared with other types of cancer, skin cancer has the inherent advantage of being directly visible in the body. Although this fact leads to easier detection and treatment of the lesion, it comes with its own obstacles. The skin, as the barrier between the inside and outside of the body, is home to many types of lesions. Most result from outside trauma, such as scratches, cuts or bruises. Some are benign and entirely cosmetic in nature, such as scars. Others are benign but can evolve into malignant lesions, such as moles. Malignant lesions can be divided into multiple categories such as cancers or warts, each having further specific subcategories.

The main problem in skin cancer is the confirmation of a lesion as cancerous, as the appearance of skin cancer can vary wildly. Dermatologist need to rely on their years of experience to provide a reliable diagnostic. Scoring systems and methods have been introduced that improve the diagnostic performance of less experienced clinicians [52]. However, these do not reach the desired goal of removing experience from the equation, as these do not raise the performance of a doctor to that of an experienced dermatologists, especially in rare cases. Biopsy is the test performed to obtain a reliable diagnosis, however the test is intrusive.

The usual method to distinguish a malignant lesion (cancer) from a benign lesion is to perform what has been denominated as “Pattern analysis" [68] were a dermatologist matches the traits of the lesion with previously recorded instances, attempting to classify the lesion accordingly [11]. From this practice came the discovery of reoccurring patterns, which led to the creation of several methods that streamline the classification process:

- ABCD rule [18] which consists of the analysis of four criteria, Asymmetry, Border irregularity, Color variegation and Dermoscopic (or Differential) structures, with a semi-quantitative score system. Often the D criteria is used to refer to the size of the lesion, with a lesion with Diameter larger than 6mm being an indication of cancer [1]. The ABCD score is computed as the weighted sum of the category scores. The final score varies from 1 to 8.9. Scores below 4.75 are identified as benign, from 4.75 to 5.45 as a sign of early melanoma, and above 5.45 classified as melanoma.
- ABCDE rule [1], same as above but has in consideration as well the Evolution of a lesion. A lesion is evolving when it changes its size, shape, symptoms (eg, itching, tenderness), surface (eg, bleeding), or shades of color.
- 7-point checklist [5] derived from the analysis of pigmented skin lesions, where 7 criteria, 3 major and 4 minor, are identified. The major criteria have a score value of 2 while the minor have a value of 1. A minimum of 3 total score is needed to identify malignant melanoma.

  The criteria include Atypical pigment network, Blue-white veil, Atypical vascular pattern, as the majors, and Irregular streaks, Irregular pigmentation, Irregular dots/globules and Regression structures, as the minors.

- Menzies method [11], where there are a set of features, 2 negatives and 9 positives. In this method, a lesion is not classified as melanoma if it does not contain any of the negative features and it contains at least one of the positive.

  The negative features consist of point and axial symmetry of pigmentation and presence of a single color, with the positive being blue-white veil, multiple brown dots, pseudopods, radial streaming, scar-like depigmentation, peripheral black dots-globules, multiple colors (5 or 6), Multiple blue/gray dots and broadened network.

Another attempt to streamline the classification process is teledermatology.

2.2 Teledermatology

Teledermatology [49] is the provision of dermatology services without the need to consult a dermatologist in person. This procedure exists thanks to improvements and a higher ease of access to technologies such as the internet and dermatoscopes, as well as the average cameras being able to obtain higher quality images.

With improvements to the camera and dermatoscopes, an average camera can be extended by a simple dermatoscope. This makes it able to capture images of a worrying skin lesion with dermoscopic-like quality. These images, along with a description, are sent through the internet to a queue. The queue is processed by teleconsultants (that are dermatologists). They determine if the lesion requires immediate attention, if the data is inconclusive or if it is benign. This feedback is given typically within a 24 hour period [49].

This method brings several advantages. An immediate opinion is not required, allowing specialists time to verify their conclusion. Despite usually taking 24 hours to provide a response, it is a significant improvement to the current wait time to see a dermatologist. Lastly, this procedure allows access to dermatology care in remote regions where an expert might not be available.

Despite these advantages, there has not been a consensus if teledermatology is comparable to dermatology. However, at its lowest point, it provided results slightly inferior to those of dermatology [49]. These results indicate that teledermatology is a viable way to streamline the triage of patients, ensuring the most urgent ones get treatment sooner. This improves the efficiency of the work time of the dermatologists, as the healthier patients will not necessitate a physical consultation.

A sub field of teledermatology is its mobile variant, where thanks to the improvement of mobile phone cameras, images with increased quality can be acquired by the average person. With use of a simple apparatus, a person can turn a mobile phone into an ad-hoc dermoscopic camera, further streamlining the teledermatology process and making it easier for a patients to provide dermoscopic-like images for check-ups. This is especially beneficial for patients that have

past history of skin cancer. They can better monitor a recurrent cancer [12].

Pilot study participants of this sub-procedure report that it is easy to perform and they felt motivated to monitor their skin more often. Barriers to this procedure included the difficulty to reach lesions in some places and the inadequate education of an average person about what lesions are worthy of attention [49].

A lesion can be classified into several types of skin cancer. Identifying the type of skin cancer is important as its danger is it tied to its type. Although all types of skin cancer can progress to more dangerous stages, some progress at a faster pace. A higher skin cancer stage complicates the treatment, leading to an increased cost and a larger impact on the lifestyle of the patient.

2.3 Skin Cancer

It is normal for damage to occur in cells. When such is the case and the damage can not be repaired, cells activate mechanisms to replace themselves. Cancer occurs when the mechanisms themselves are damaged. This causes the cells to not be able to terminate themselves, resulting in them starting to multiply without control [26]. This replication causes an unusual growth in the area, sometimes causing the area to be raised, as well as oozing or bleeding more easily [63]. This effect can be observed in Figure 2.1, it is a graphical representation of the effects of various cancers on the skin. The worst case occurs when the cancerous cells start spreading throughout the body (metastasis), leading to the creation of more cancerous spots and resulting in the eventual death of the patient.

Other factors that cause damage to the skin, particularly damage that can lead to cancer, is exposure to UV radiation (Sun), errors in the genetic code of the cell or usage of tanning beds [26]. A weakened immune system can also increase the risk of skin cancer [63].

2.3.1 Types of skin cancer

Skin cancer type varies from cell to cell, as such, many of their classifications are named after the cell they originate from. Due to the large variety of cells in the skin, there exists a large range of possible skin cancers, with the most common or dangerous types being the following:

- Basal Cell Carcinoma (BCC), most common. (Figure 2.2)

  It is the most common type of skin cancer, covering over 80% of skin cancer classifications, it originates from basal cells, in the lowest part of the epidermis. The basal cell is a cell responsible for producing new skin cell as old ones die off, acting as a regenerative layer for the skin [54]. This cancer tends to occur in areas where the skin is regularly exposed to the sun. Its growth is usually slow, with it being rare to metastasize or spread to nearby lymph nodes, a small bean-shaped structure that is part of the body's immune system [39], but will happen with time.

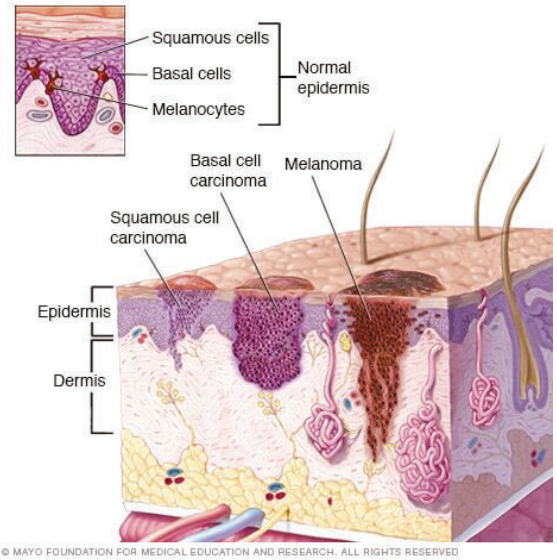

Figure 2.1: Demonstration of various types of cancerous growths on the skin and its layers [54]

BCC has a tendency to reappear/reoccur in the same location after being successfully treated/removed with a 50% chance of recurrence within five years of the first diagnosis [63]. Several features can be recognized in the cancerous region, such as open sores, red patches, pink growths, shiny bumps, scars or growths with slightly elevated, rolled edges and/or a central indentation which, at times, may ooze, crust, itch or bleed. In patients with darker skin, about half of BCCs are pigmented (meaning brown in color). BCCs can vary significantly from person to person [26].

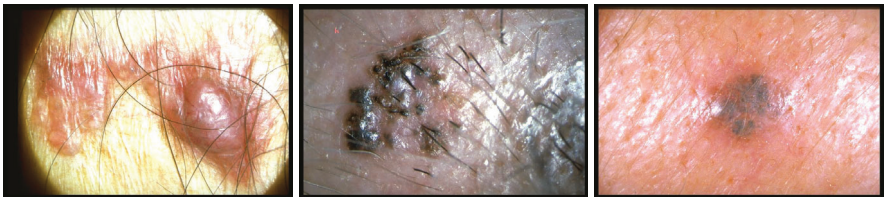

Figure 2.2: Examples of Basal Cell Carcinoma [4]

- Squamous Cell Carcinoma (SCC), second most common.

  The Second most common type of skin cancer, develops at the flat, thin squamous cells that make up much of epidermis, the outermost layer of the skin [63]. Squamous cells are found in many places of the body [54], occurring commonly in places where absorption or transportation of materials plays an important role such as in diffusion, osmosis and

filtration [10].

Like BCC, the cancer typically occurs in places that have been exposed to the sun, but may develop in other areas that contain scars or skin ulcers, as well as in the genital region. Some characteristics, like the slow growth and difficulty spreading or metastasizing, are also shared with BCC, however SCC has higher chances to invade fatty tissue beneath the skin or spread even further [63]. It occurs with features like scaly red patches, open sores, rough, thickened or wart-like skin, or raised growths with a central depression. SCCs may, at times, crust over, itch or bleed, with its features varying from person to person [26].

- Melanoma, most dangerous. (Figure 2.3)

  Melanoma is the most serious type of skin cancer, developing from melanocytes, a cell found in the epidermis. The melanocytes is a cell responsible for the production of melanin, a pigment responsible for the color of the skin [26]. These cells darken when exposed to the sun, shielding the deeper layers of the skin from the harmful effects of the ultraviolet (UV) rays from the sun [63]. Since it forms from melanocytes, it usually occurs in the skin, regardless if it is exposed to the sun, but can also appear on the eyes and, rarely, in internal organs, such as the intestines. The exact cause is not clear, but exposure to UV radiation from sunlight or tanning lamps and beds increases the risk of developing melanoma [54].

  Contrary to BCC and SCC, melanoma tends to spread to other parts of the body. The cancerous cells are often found in lymph nodes, being able to spread via the lymphatic channels, which connect other lymph nodes throughout the body [62]. It can occur with different shapes, sizes and colors. Due to this it is hard to provide an easy guide for identification to inexperienced people. Despite this, they typically either mutate from existing moles, $20 - 30\%$, or form mole-like lesions on normal skin, $70 - 80\%$ [26].

  Melanoma can be sub-categorized in several sub-types of malignant melanoma. The superficial spreading melanoma, nodular melanoma and lentigo maligna melanomas make up to 90% of all diagnosed malignant melanomas. The remaining 10% are filled by the rarer types such as acral lentiginous melanoma and acral amelanotic malignant melanoma [62]. Each sub-type of melanoma has its own features, some being extremely different from other. The superficial spreading malignant melanoma occurs as a new or a mutation from an existing mole, while nodular malignant melanoma takes a blue or red appearance, typically occurring as a new mole.

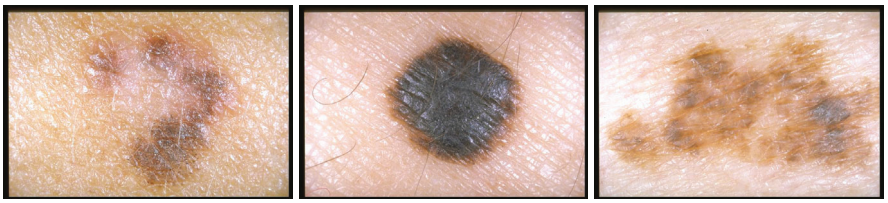

Figure 2.3: Examples of Melanoma [4]

- Merkel Cell Carcinoma (MCC), rarer but dangerous

  This is a rarer type of skin cancer but equally as dangerous as melanoma, with it forming from Merkel cells, in the epidermis layer of the skin [58, 63]. Merkel cells provide, along with nerve endings, the sense of touch from the skin [63].

  MCC appears more commonly in areas of the skin exposed to the sun, as well as in people with age over 50 and weakened immune system. Like melanoma, MCC is an aggressive form of skin cancer with a high risk of recurring and metastasizing, often within two to three years after initial diagnosis [26]. It frequently targets the brain, bones, liver and lungs when it metastasizes [63].

  MCC is harder to find than other skin cancers, as it can appear as a pearly pimple-like lump, sometimes skin-colored, red, purple or bluish-red, though they are rarely tender to the touch. Often patients and doctors only discover the cancerous formation due to its alarmingly high rate of growth [26].

  It is worth noting that Merkel Cell Carcinoma was denominated after the Merkel cell due to the similarities between the cells. However, since said cell does not self-replicate, the cancerous cells derive from the progenitor, usually epidermal precursor cells [58].

There are various medical terms that describe/encompass various types of cancer. The easiest to understand is the "non-melanoma cancer". Due to the danger and frequency of melanoma, all other types of cancer can be classified under non-melanoma cancers. Another term is the "keratinocyte carcinomas" that includes cancers originating from cells birthed from a cell called keratinocyte, BCC and SCC cancers belong in this group [64].

Cancerous cells are not the only thing a doctor should be on guard for. As cancer can develop in previously healthy lesions, some lesions should be paid attention to, either for their close resemblance to a cancer, or because they have a high likelihood to become cancerous. The Seborrheic Keratosis (SK) (Figure 2.4) is a common benign skin growth. The factors that lead to its occurrence are not well known, but they tend to occur in people older than 50 and within families, so genes are believed to play a role [54]. Generally harmless and not contagious, treatment is not necessary, however clothing can cause them to irritate or bleed. Otherwise, if signs such as sores, quick increase in size, bleeding and not healing appear, it usually indicates the development of cancer [54]. Typically appearing on the head, neck, chest or back, multiple growths are common, with their appearance being usually brown, black or light tan, with a waxy, scaly and slightly raised look [54].

Nevus (NEV) (Figure 2.5) is another benign skin lesion. Nevi (plural) is the medical term used for various skin formations such as moles, birthmarks or beauty marks. A common occurrence, nevi is a collection of harmless colored cells, typically appearing as a small brown, tan, or pink spots [65]. A person can both be born with these or develop new ones later.

The later a skin cancer is diagnosed, the harder the treatment, as such, screening has been used as an attempt to identify skin cancer early.

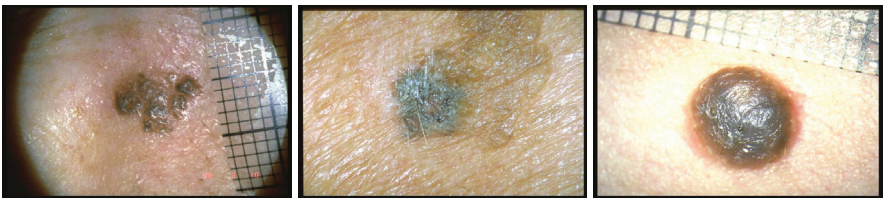

Figure 2.4: Examples of Seborrheic Keratosis [4]

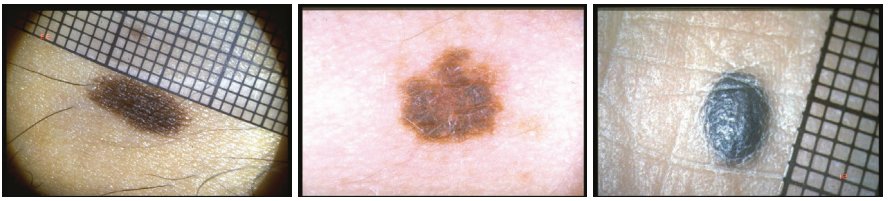

Figure 2.5: Examples of Nevus [4]

2.3.2 Screening for skin cancer

Screening is the process of performing a preemptive search for a cancer before a patient shows symptoms [67]. This results in the identification problem areas that are monitored by both the patients and doctors, quickly catching any lesion that becomes cancerous. There are attempts to implement general screenings for skin cancer however the process is not simple.

The complications arise from various factors. First there is the large variety of shapes, sizes, coloration, amongst others, that skin lesions can take. Many lesions are not malignant, even if they resemble a cancerous lesion, so a high degree of expertise is required to correctly classify a lesion. On top of that, previously benign lesions can mutate into cancer, so, even from the benign lesions, some have to be placed under surveillance. The fact that the whole perimeter of the body must be examined, coupled with some difficult decisions extends the time required to perform the screening task.

Screening for skin cancer entails the visual observing the skin in the search for nevi, or other pigmented areas that look abnormal in color, size, shape, or texture [67]. Other factors that increase the likelihood of skin cancer are also documented. Factors like sunburnt areas, especially those that are not recovering properly, fair skin, the patient using tanning beds, being older than 60 or having a weak immune system all contribute to a higher chance of skin cancer occurring [8].

If not for the extremely varied nature of the cancerous lesions, the screening process could be performed by regular medical staff, with minimal training. Attempts have been done in order to streamline the classification of cancerous lesions. The identification of common irregularities lead to the introduction of several methods that can quantify the likelihood of a lesion being cancerous.

These methods improved the diagnostic performance of less experienced dermatologists. However they fail to address the reproducibility problem of the diagnosis, even amongst experts [52].

Teledermatology aids in the screening process, as the medical staff can be educated in what constitutes a dangerous lesion, leaving the precise classification for the expert. This method expands the number of medical staff that is now qualified to perform the screening process, making it a more viable solution as well as reducing wait times for when the procedure is utilized [47].

Identifying cancerous lesions early is critical to increase the survivability of the patient, however this does not mean that the screening process is a perfect solution without drawbacks.

2.3.3 Drawbacks of screening for skin cancer

Before stating the drawbacks, it is note worthy to point out that, in addition to increasing the survivability of the patient, documenting dangerous lesions that later become cancerous provides useful data for future screenings, diagnosis and construction of automatic classifiers.

Although largely beneficial, screening has valid problems. These mainly originate from the lack of a simple and accurate diagnosis method [8, 67]. First, the screening process itself takes a considerable amount of time to be conducted, as the whole body needs to be observed. Skin cancer is more likely to develop in sun-exposed areas, but it can also occur in protected/hidden areas such as the genitalia.

Due to the complicated nature of the lesions, their large variations in shape, color and texture, they require the opinions of an expert for classification. However, obtaining the opinion of an expert takes too much time due to the inadequate number of professionals [32] and the large gap in experience [35] each expert can have.

Incorrect classification of a lesion carries problems independent of screening. However screening increases the number of lesions that are classified, leading to eventual incorrect classifications. This is a problem in screening as, in the large majority of the time, the classified skin lesions have shown no symptoms or other evidence to suggest the need for medical care. From a misclassification, two outcomes can occur. If a cancer is incorrectly classified as benign, the malignant lesion is given more time to progress, with the possibility of the patient delaying getting medical care despite having symptoms. If a benign lesion is misclassified as cancerous, the patient could suffer from anxiety, as well as receiving further treatment that would be costly and have side effects. There is also the possibility of leaving permanent damage such as scars from the excision of the lesion [67].

Lastly, the method generally used to diagnose the cancer is the biopsy. This operation requires the extraction of the lesion, either in its entirety or only portions, so that it can then be analysed by a pathologist. The operation is invasive, with the possibility of causing infection and scarring [67].

An important note is the existence of studies that show that screening for skin cancer has not led to a decrease in the chances of dying from skin cancer [67].

These set of factors have led to some organizations not having an opinion on whether to recommend or oppose the promotion of the screening of skin cancer [8]

With the recent developments of various technologies, various machine learning algorithms started to produce better results in this area. Although these algorithms will not replace the experts, they can be used to provide a reliable second opinion on a lesion.

2.4 Skin cancer data sets

Skin cancer cases have been documented in the medical record of the patients with the format of pictures and description/annotation of important structures. The datasets are made from these medical records and have some common problems. The data in the medical record of patients is confidential, so access to it is complicated. Each hospital/clinic/university hospital has different formats for their medical records, as well as the data (images), being acquired with different equipment and techniques (the images were acquired with different cameras and conditions). As such, the samples from a dataset may vary considerable from samples of another dataset, making it difficult to compare the obtained results.

Another problem is that different datasets were created with different objectives in mind. Some are created with a focus on separating Nevus from Melanoma, others to differentiate cancer from non-cancer (involving various types of skin cancer, not just Melanoma (MEL)). As such the number of classes or data types for each datasets can vary, making it more difficult to join them together without loosing information.

A big problem with the datasets from this domain is their low availability, coupled with the low number of samples. Although recently there has been an effort to congregate large numbers of samples and making them public [40], the majority of the datasets in the literature are private. Even those that are public contain a small number of samples. Some authors complimented public datasets with their privately obtained datasets. Be it due to the low number of samples from the public datasets, or intending to increase the diversity of their data, in the end the larger datasets were never released to the public.

A total of 24 datasets were considered. They are here described with as much information as was made available. For brevity, the classes that each dataset contains (when stated) are abbreviated, with only the classes that are used in this work specifically specified.

The datasets of DermIS [23], Dermnet [22] and Dermnet NZ [69] were obtained from online archives. These online archives contain a large amount of data within, however this data is spread across many specific skin cancers, with some containing only a few (1 2) samples. The data is publicly available, however contains a watermark. The original, higher quality, data can usually be obtained by requesting it from the website owners, or by paying a fee.

DermIS [23] originates from the University of Heidelburg and Erlangen, Germany. The largest reported usage mentions 397 clinical images and metadata, including melanoma and nevi skin lesions. The images have been reported to have varying resolutions (from 550×367 to 550×469) and hair. Dermnet [22] contains around 23,000 images, as of 2016, while DermNet NZ [69] contains more than 25,000 clinical or dermoscopy images, as of 2020. Their resolutions and quality vary from image to image.

DermQuest [29] would be another online medical archive but ended up losing support and being discontinued in 2019/12/31, however most of their database can be found in the SD-198 [78] dataset. The SD-198 is a dataset made from data available on the DermQuest site at the time of its creation on 2016. It contains a total of 6,584 images from 198 classes, these include different diseases from different types of eczema, acne and various cancerous conditions as well as lesions in hard-to-diagnose places.

Two of the datasets were obtained from Dermoscopy Atlases. The datasets are the Sydney Melanoma Diagnostic Centre at the Royal Prince Alfred Hospital [57] and the Interactive atlas of Dermoscopy: A tutorial (also known as EDRA) made by Argenziano et al. [4] from the university hospital of Graz, Austria, university hospital of Naples, Italy, and university hospital of Florence, Italy. Their images are true color, have a resolution of 768×512 pixels, have similar quality and some samples contained lesions that exceeded the border of the image. The first Atlas contains 168 clinical images, divided into benign and melanoma classes. The second Atlas contained 1011 lesions with clinical image, dermoscopic image, metadata and the 7-point classification for each sample. the dataset is divided into 20 classes, with the more prominent classes being BCC, Nevus, Miscellaneous (MISC), SK and MEL. Although both started as private, the Interactive atlas of Dermoscopy: A tutorial [4] was later released together with a study by Kawahara et al. [44].

The MED-NODE [31], PH2 [56], Dermofit [6], HAM10000 [81], BCN20000 [19] and ISIC archive [40] are datasets that are available to the public. Of these, only the Dermofit dataset has a one-off pay requirement. The MED-NODE [31] dataset was created with images from the Department of Dermatology of the University Medical Center Groningen, Netherlands. These images were acquired using a Nikon D3 or Nikon D1x body and a Nikkor 2.8/105 mm micro lens and lighting provided by two Multiblitz Variolite 600 flash units with a color temperature equal to 5200 Kelvin. A total of 170 clinical images, divided into MEL and NEV, was selected from the hospital. The images were manually processed to remove distracting elements, like clothes, and obtrusions, like hairs, ensuring the focus of the images were the lesion and some surrounding healthy skin.

The PH2 [56] dataset was made in a joint research collaboration between the Universidade do Porto, Técnico Lisboa, and the Dermatology service of Hospital Pedro Hispano in Matosinhos, Portugal. The images are obtained under the same conditions through Tuebinger Mole Analyzer system with a magnification of 20×, resulting in 8-bit RGB color images with a resolution of 768×560 pixels resolution. A total of 200 dermoscopic images and metadata, divided into NEV and MEL, were chosen with the best quality, resolution and features. The metadata includes