

# Recent advances in hyperspectral imaging for melanoma detection

Thomas H. Johansen\*, Kajsa Møllersen†, Samuel Ortega‡, Himar Fabelo‡,  
Aday Garcia‡, Gustavo M. Callico‡, Fred Godtlielsen\*

## Article Type:

Advanced Review

## Abstract

Skin cancer is one of the most common types of cancer. Skin cancers are classified as non-melanoma and melanoma, with the first type being the most frequent and the second type being the most deadly. The key to effective treatment of skin cancer is early detection. With the recent increase of computational power, the number of algorithms to detect and classify skin lesions has increased. The overall verdict on systems based on clinical and dermoscopic images captured with conventional RGB cameras is that they do not outperform dermatologists. Computer-based systems based on conventional RGB images seem to have reached an upper limit in their performance, while emerging technologies such as hyperspectral and multispectral imaging might possibly improve the results. These types of images can explore spectral regions beyond the human eye capabilities. Feature selection and dimensionality reduction are crucial parts of extracting salient information from this type of data. It is necessary to extend current classification methodologies to use all of the spatio-spectral information, and deep learning models should be explored since they are capable of learning robust feature detectors from data. There is a lack of large, high-quality datasets of hyperspectral skin lesion images, and there is a need for tools that can aid with monitoring the evolution of skin lesions over time. To understand the rich information contained in hyperspectral images, further research using data science and statistical methodologies, such as functional data analysis, scale-space theory, machine learning, and so on, are essential.

---

\*Department of Mathematics and Statistics, UiT The Arctic University of Norway

†Department of Community Medicine, UiT The Arctic University of Norway

‡Institute for Applied Microelectronics, University of Las Palmas de Gran Canaria

# Introduction

Skin cancer is one of the most common types of cancer in humans, and in countries with predominantly fair-skinned population the incidence trend for the past 30 years has been increasing (American Cancer Society, 2018; Ferlay et al., 2013). Skin cancers are classified as non-melanoma skin cancer (NMSC) and melanoma. NMSC is by far the most frequent, whereas melanoma is the most deadly. In 2018 the reported number of new cases of NMSC globally accounted for 5.8% of all new cancer cases, and accounting for 0.7% of all deaths. New cases of melanoma was reported to account for 1.6% of new cancer cases, but notably accounting for 0.6% of all deaths caused by cancer (Bray et al., 2018). The key to effective treatment of skin cancer is early detection, before the cancer metastasizes. Non-metastasized melanoma is reported to have a 5-year survival rate of 99%, whereas once it spreads to distant organs the survival rate drops to 20% (American Cancer Society, 2018). In dermatology one of the most commonly taught diagnostic guidelines for classifying pigmented skin lesions is the ABCD rule of dermatoscopy (Nachbar et al., 1994). The respective letters in the acronym represent different features of a skin lesion; asymmetry, border, color, and differential structures. When using the ABCD rule to diagnose a skin lesion, a score is assigned for each of the four features, and combined into a total score. The total score gives an indication of the potential for malignancy, where higher scores mean greater potential for malignancy. In clinical settings the reported sensitivity and specificity of the ABCD rule ranges from 74–91.6% and 45–67%, respectively (Annessi, Bono, Sampogna, Faraggiana, & Abeni, 2007; Unlu, Akay, & Erdem, 2014; Ahnlide, Bjellerup, Nilsson, & Nielsen, 2016). Figure 1 shows a few examples illustrating typical variation between different skin lesions. Note the differences in the shapes, borders (and lack thereof), colors, etc.

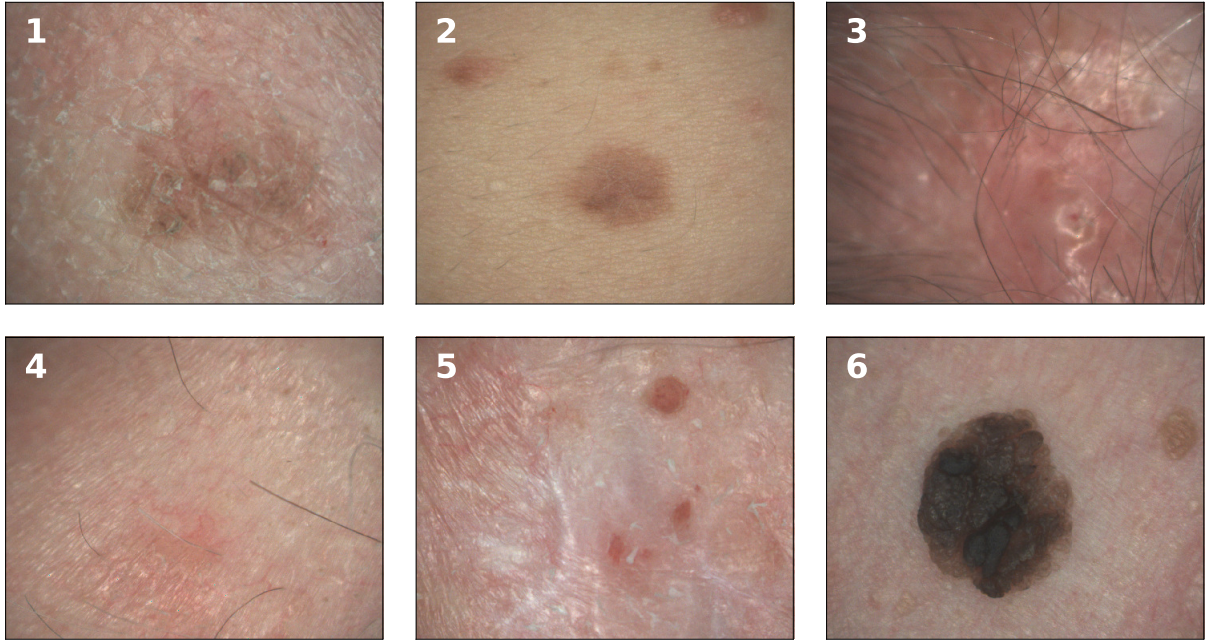


Figure 1: Examples of melanoma and non-melanoma skin cancer taken in a clinical setting from six different patients. These cases represent both non-melanoma and melanoma skin cancer. The diagnoses of the lesions based on histopathology are as follows; 1) melanoma, 2) atypical melanocytic hyperplasia, 3) squamous cell carcinoma, 4) Bowen’s disease, 5) basal cell carcinoma, 6) seborrheic keratosis.

Given the increasing trend in skin cancer prevalence, and the difficulty in detecting skin cancer at an early stage, researchers across many fields have been working to both extend and develop new diagnostic criteria and computational algorithms. For example the ABCD rule of dermoscopy has been extended to ABCDE, where the E accounts for evolution of the skin lesion over time (Abbasi et al., 2004). With the advent of machine learning and the increasing access to vast, inexpensive computational power, several research groups have been focusing on developing automated and semi-automated computational methods for detecting and classifying skin lesions. While some recent advances have been developed using conventional RGB (red, green and blue) imaging techniques (Esteva et al., 2017), other researchers have been focusing on exploring new avenues of skin cancer classification using multispectral and hyperspectral imaging techniques.

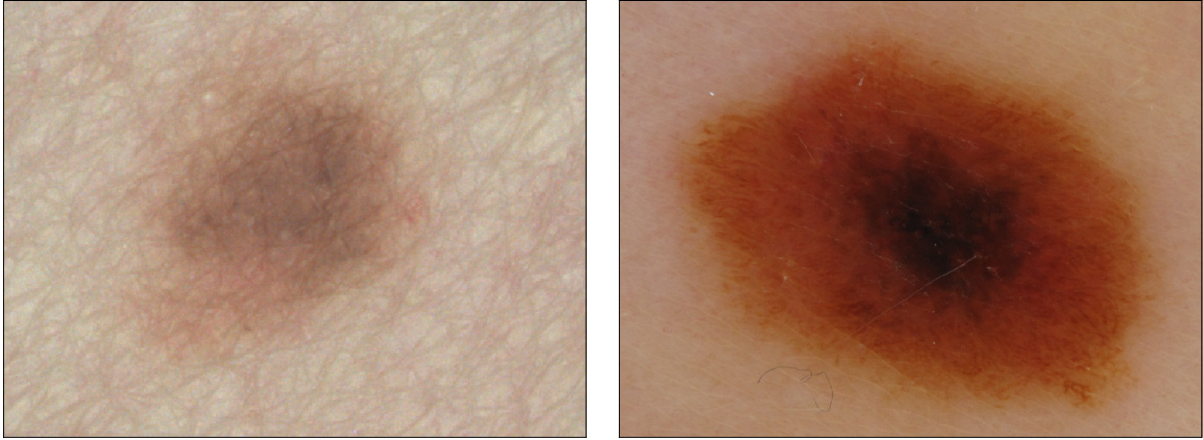


Figure 2: The image on the left is an example of a conventional, clinical image of a pigmented skin lesion, whereas the image on the right is an example of dermoscopic image.

Computer systems for classification of pigmented skin lesions has been an active research field for several decades. Early systems used conventional RGB images, but by the early 2000's almost all systems used dermoscopic images (Rosado et al., 2003). See Figure 2 for examples of both types of images. A dermoscope is a simple device consisting of a magnifying lens, a glass plate and a light source that allows the light to penetrate the uppermost layer of the skin. It is commonly used by dermatologists. The overall verdict of systems based on conventional and dermoscopic images is that they do not outperform dermatologists (Rosado et al., 2003; Vestergaard & Menzies, 2008; Korotkov & Garcia, 2012). Deep learning has been introduced to skin lesion classification, and although deep learning methods possibly outperform traditional approaches (Codella et al., 2018), it has not outperformed the dermatologist (Esteva et al., 2017). Multispectral imaging increases the amount of retrieved information and various systems has been used for skin lesion classification. Whether this increases the performance has not been established with certainty, since dermoscopic and multispectral systems have not been tested on the same set of lesions, or under strictly similar conditions. The conventional and dermoscopic systems seem to have reached an upper limit for their performance, while emerging technologies such as hyperspectral imaging can possibly increase the performance.

The main advantage of hyperspectral and multispectral imaging compared to conven-

tional imaging technologies is the possibility of exploring spectral regions beyond the human eye capabilities. Some materials present spectral features in the infrared spectral range (Lachenal & Ozaki, 1999). Besides the spectral range, the use of hyperspectral images is necessary when the material being analyzed presents narrow spectral features (Lee, Cohen, Kennedy, Maiersperger, & Gower, 2004; Jet Propulsion Laboratory, California Institute of Technology, n.d.). Such narrow spectral features cannot be detected using multispectral or RGB images, and should therefore be measured using high spectral resolution instrumentation. Figure 3 illustrates the difference in fidelity and richness of hyperspectral images in comparison to conventional RGB images.

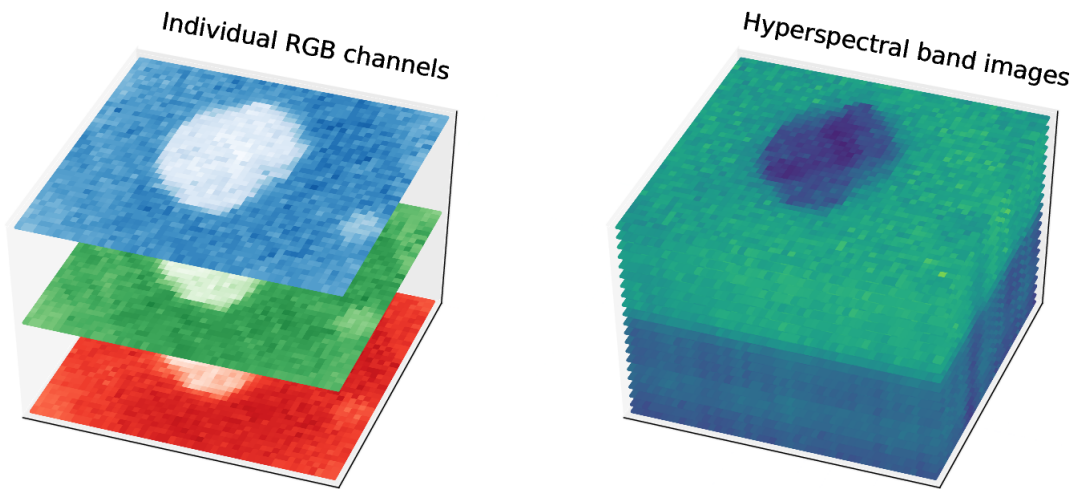


Figure 3: The conceptual difference between the information richness in a hyperspectral cube and an RGB image. In the hyperspectral cube, each horizontal slice represents spatial response for a discrete wavelength. For the RGB image each slice represents spatial information across a range of wavelengths. Each of the red, green, and blue slices are calculated based on the visual light spectrum associated with each respective color.

In this review we will report on the recent advances that specifically focus on detecting skin cancer using multi- and hyperspectral images. We will start by giving a short description of the review methodology. Then, we give a brief introduction to hyperspectral imaging and point out how this imaging technique is being used in medicine, and specifically why it is being used to classify skin cancer. Next, we focus on how feature selection is crucial for

extracting information of this type of data and thereafter we point out the need for extending current classification methodologies to include the use of spatio-spectral information. Our review finally gives some critical remarks and analysis of relevant published results before we indicate important future research directions.

## **Review methodology**

The goal of this review was to provide insight into recent advances in detection of skin cancer using hyperspectral imaging systems in order to uncover what has been achieved, and to understand what the key challenges are. Based on this we defined the following inclusion criteria with the intention of only including recent, highly relevant, peer-reviewed publications focusing on skin cancer detection using hyperspectral images;

- Peer-reviewed publication in journal or conference proceeding.
- Based on hyperspectral (or multispectral) images.
- Specifically dealing with skin lesion classification.
- Non-invasive data collection, i.e. in-vivo skin lesions.
- Published in recent years (2003–2018).

The inclusion of multispectral imaging systems was made based on preliminary searches, which uncovered that most of the relevant skin cancer research has been done with these systems. Although multispectral and hyperspectral systems are based on different concepts and technologies, from the perspective of data analysis and pattern recognition, images produced by these systems present similar benefits and challenges. Our initial threshold for “recent” was 10 years, but because some very relevant studies were published more than 10 years ago, the threshold was increased to 15 years.

In the period of August 20–23, 2018, we performed searches on Web of Science, PubMed, Scopus, and Google Scholar. Search queries were specifically adapted to each search engine, and based on the search criteria seen in Listing 1.



Listing 1: The search criteria used to construct search queries for Web of Science, PubMed, Scopus, and Google Scholar.

```
((multispectral AND classification) OR hyperspectral) AND  
(image OR camera) AND  
(skin OR melanoma) AND  
(classification OR lesion OR cancer)
```

Our searches resulted in a collection of 86 peer reviewed publications that were potential candidates for review based on their titles, keywords, and abstracts. After reading through the initial collection of candidates, we ended up with 20 publications relevant for the review, selected based upon our previously listed inclusion criteria. Figure 4 shows a breakdown of the number of publications per year, and a summary of all the reviewed publications can be seen in Table 1.

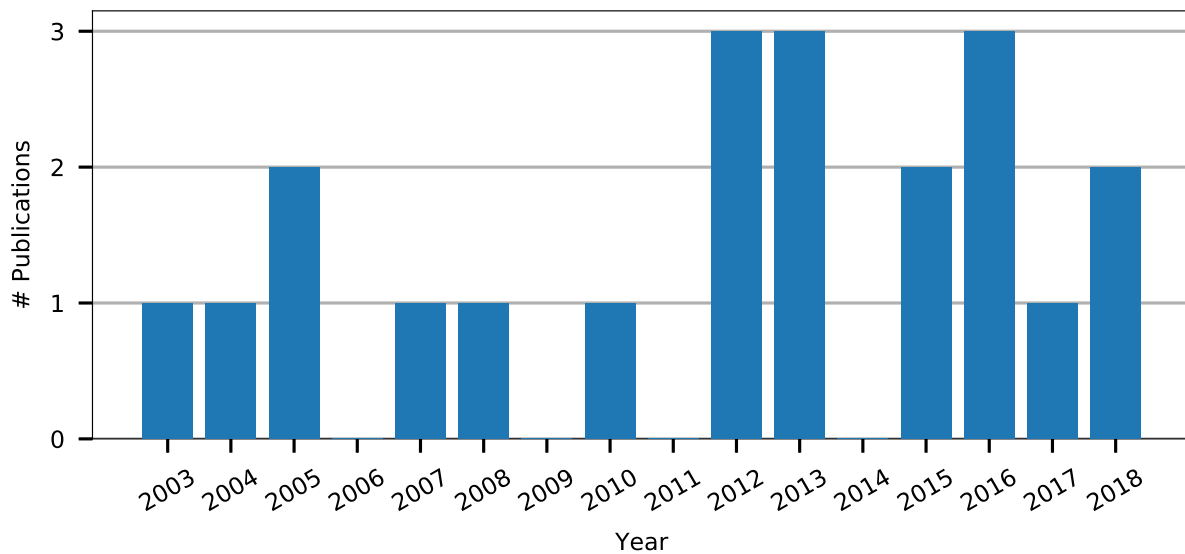


Figure 4: The number of publications per year that matched our search queries and were selected for review based on our inclusion criteria. From the plot we can see that the majority of reviewed publications were published after 2011.

Publication	Imaging System	Wavelengths (nm)	Bands	Pixel Count
Tomatis et al. (2003)	Custom	400–1040	17	—
Patwardhan et al. (2004)	Nevoscope	580, 610	2	512×512
Patwardhan et al. (2005)	Nevoscope	580, 610	2	512×512
Tomatis et al. (2005)	SpectroShade	483–950	15	640×480
Carrara et al. (2007)	SpectroShade	483–950	15	640×480
<b>Kazianka et al. (2008)</b>	Custom	—	300	640×480
Świtoński et al. (2010)	VariSpec	410–710	21	—
<b>Nagaoka et al. (2012)</b>	ImSpector V8E	380–780	124	512×512
<b>Nagaoka et al. (2012)</b>	ImSpector V8E	380–780	124	512×512
Suárez et al. (2012)	Custom	400–1100	—	—
<b>Nagaoka et al. (2013)</b>	ImSpector V8E	380–780	124	512×512
Quinzán et al. (2013)	Custom	400–1100	71	640×480
<b>Nagaoka et al. (2015)</b>	ImSpector V8E	450–750	124	1024×768
<b>Zheludev et al. (2015)</b>	VTT/Revenio	500–885	76	320×240
Lorencs et al. (2016)	Nuance EX	450–950	51	—
Song et al. (2016)	MelaFind	430–950	10	1280×1024
<b>Zherdeva et al. (2016)</b>	STC UI RAS	450–750	61	1920×1200
Stamnes et al. (2017)	Custom	365–1000	10	—
Lihacova et al. (2018)	Custom	405–964	4	—
Rey-Barroso et al. (2018)	Custom	414–1613	14	512×512

Table 1: Summary of the publications included in the review. Publications denoted in bold indicate that the research is based on hyperspectral images. The “—” entries indicate that the information is not explicitly presented in the publication.

## Hyperspectral imaging for skin cancer classification

Hyperspectral imaging has shown considerable potential as a non-invasive and non-ionizing technique, supporting rapid acquisition and analysis of diagnostic information. Unlike conventional RGB cameras, which are limited to capturing three bands in the electromagnetic spectrum, hyperspectral imaging systems are capable of capturing hundreds of narrow bands across the electromagnetic spectrum, both inside and outside the human visual spectral



range (Smith, 2012). Hyperspectral imaging has been widely used in remote sensing (Tuia, Volpi, Copa, Kanevski, & Munoz-Mari, 2011), and has been applied in-vitro, ex-vivo and in-vivo in different medical applications (Lu & Fei, 2014). For skin lesion classification, several studies have been conducted using different types of hyperspectral and multispectral acquisition systems. Based on the reviewed publications listed in Table 1, most of the research effort up until now has been based on multispectral systems. Some multispectral devices for skin lesion analysis are commercially available, such as MelaFind (Elbaum et al., 2001; Kupetsky & Ferris, 2013), and SIAscope (Moncrieff, Cotton, Claridge, & Hall, 2002), both operating in the 400–1000 nm spectral range. We are currently not aware of any commercially available hyperspectral imaging systems designed for skin lesion analysis.

Both multispectral and hyperspectral images are commonly represented as three-dimensional matrices (or data cubes), where the first two axes represent the spatial coordinates, and the third axis contains the spectral band measurements. There are two commonly used approaches to visualizing the information stored in a hyperspectral image. The first one is to pick one or more pixels (spatial coordinates) and plotting their respective spectral band measurements by wavelength. The other way is to visualize all pixels for one or more spectral bands as individual grayscale or color-mapped images. See Figure 5 for an example of both types of visualization.

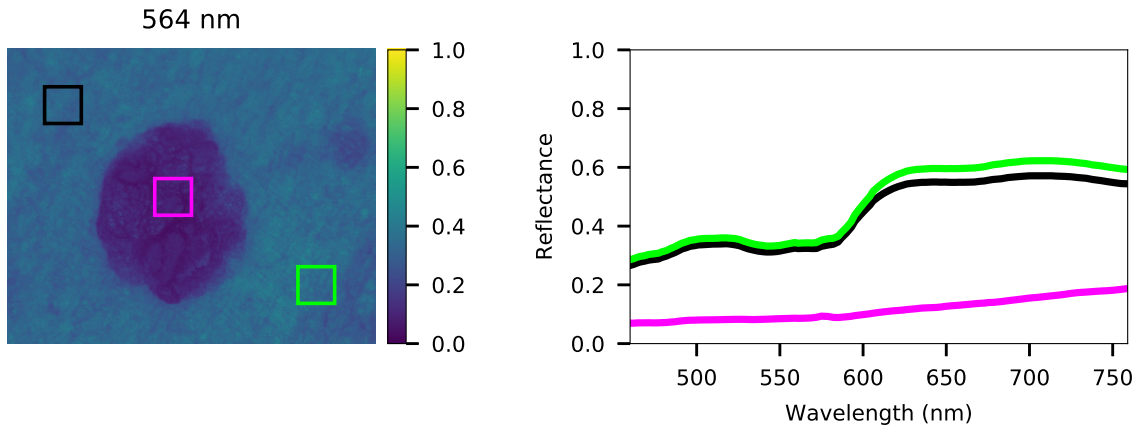


Figure 5: An example of what spectral curves for hyperspectral pixels can look like. The plot on the left shows a representation of a hyperspectral reflectance image at an arbitrarily chosen wavelength. On the right, the mean reflectance values are plotted, where the colors of the curves correspond to the colored regions in the reflectance image. The mean curves are calculated based on all pixels in each region.

There are several ways to capture both hyperspectral and multispectral images (Li et al., 2013), but from the perspective of data science applications, how images are captured is not crucial. However, what the captured image data represents is important. Both types of images contain information that represents either absorption, reflectance, or radiance at specific wavelengths across the electromagnetic spectrum. Measurements at discrete wavelengths is typically not performed, but measurements are instead performed across narrow ranges of wavelengths referred to as spectral bands. Multispectral images are often captured at specifically chosen spectral bands across the supported spectral range of the camera. In many scenarios the chosen spectral bands are picked based on prior knowledge, such as known absorption wavelengths of certain chemical compounds or similar. Other times the spectral bands are chosen somewhat arbitrarily, or at evenly spaced intervals across the entire spectral range. Commercial multispectral systems are typically capable of capturing 5–15 spectral bands across their supported spectral range. Because of spectral resolution and how wavelengths are typically chosen, captured multispectral data should be considered as consisting of discrete measurements. Hyperspectral images are captured with constant

sampling rate across the spectral range of the camera, and can have hundreds of spectral bands depending upon the resolution of camera. Therefore, measurements in hyperspectral images are often considered to be continuous, which means that each pixel in a hyperspectral image can be said to represent a continuous spectral curve.

Before multispectral or hyperspectral images can be used as input to any classifier, statistical method, or other computational algorithm where images will be compared in some sense, they need to be pre-processed. One very important pre-processing step is calibration with respect to a known reference, typically an image of certified white reference material captured. The image of the white reference is captured right before or after taking an image of a skin lesion. This ensures that both images are captured under equivalent conditions. Certified white references used with hyperspectral systems have known spectral response, e.g. 99.9% reflectance, across the entire supported spectral range, and are often intended to represent the maximum values measurable by a camera. In addition so-called “dark current” or dark reference images are usually also captured as part of the calibration process. These images can be captured by preventing light from hitting the camera sensor, and they therefore represent the minimum values measurable by a camera. An underlying assumption in this process is that the following inequality is fulfilled,

$$0 \leq I_{\text{dark}} < I_{\text{raw}} < I_{\text{white}} \quad (1)$$

where  $I_{\text{raw}}$  is the raw image before calibration,  $I_{\text{white}}$  is the white reference image, and  $I_{\text{dark}}$  is the dark reference image.

A frequently used method for calibrating hyperspectral images is relative reflectance, which in this context is performed by re-scaling spectral measurements from the skin lesion image with respect to the two reference images (K. C. Lawrence, B. Park, W. R. Windham, & C. Mao, 2003; W. Wang, Li, Tollner, Rains, & Gitaitis, 2012) captured under similar conditions. The relative reflectance image can be expressed as

$$I_{\text{reflectance}} = \frac{I_{\text{raw}} - I_{\text{dark}}}{I_{\text{white}} - I_{\text{dark}}} \quad (2)$$

Given that the inequality in (1) is fulfilled, relative reflectance images will theoretically be bounded in  $(0, 1)$ . This also implies that all calibrated images from the same camera system

are comparable in a fairly robust sense since the process reduces the effects of the camera itself and the environment in which images are captured. Furthermore, images are scaled to the same reference domain.

Another calibration technique used in some skin lesion classification research is the so-called optical density (Zherdeva et al., 2016; Lorencs, Sinica-Sinavskis, Jakovels, & Mednieks, 2016). In the context of multispectral and hyperspectral images, optical density can be defined as the logarithm of the ratio of a known reference image to the raw image,

$$I_{\text{OD}} = \log\left(\frac{I_{\text{reference}}}{I_{\text{raw}}}\right) \quad (3)$$

The reference image  $I_{\text{reference}}$  can be a white reference image, or an image of some other reference material with known spectral characteristics.

In Rey-Barroso et al. (2018) a novel hyperspectral image calibration for skin analysis is presented. The first innovation is to employ a neutral-gray color of an X-Lite ColorChecker reference instead of a conventional certified white reference. The motivation is that this reference material exhibits reflectance characteristics closer to that of human skin across the spectral range. They also perform an additional calibration step designed to account for the influence of healthy skin, reportedly boosting the effects of malignant tissue.

## Classifier input

The underlying goal of classification is to organize observations into two or more labeled classes. The classifier can be considered an algorithm that suggests a class affiliation based on the input characteristics of the observation. For early detection of skin cancer, there will typically only be two classes; malignant and benign. The classifier takes the skin lesion image, or features extracted from the image, as input and gives a binary output indicating whether the lesion is malignant or not.

The input to the classifier must contain information that makes it possible to discriminate according to the different classes. In the skin cancer situation, this means that the input must contain crucial properties of the skin lesion so that an image of a skin lesion can be assigned the correct class in a very robust manner. Since the input of the classifier plays such a crucial role, we will describe some important aspects of this for the skin cancer case.

## Feature extraction, feature selection, and dimensionality reduction

As pointed out earlier, a set of characteristics or features must be extracted from the image to construct a classifier. These features can be categorized into hand crafted features and summary statistical features. A third category, machine learned features, will not be discussed in this section since there are no deep learning classifiers yet for hyper- or multispectral skin lesions. However we will discuss some aspects related to learning features from data in later sections.

The hand crafted features aspire at mimicking some aspect that is known to be discriminatory for lesion diagnosis, often inspired by, but not limited to, the ABCD rule of dermoscopy (Nachbar et al., 1994). Several hyper- and multispectral systems apply hand crafted features, exclusively or in combination with summary statistics features (Tomatis et al., 2005; Carrara et al., 2007; Stamnes et al., 2017).

The summary statistics features are typically the mean, variance, entropy, etc., of the pixel value for each spectral band. Common for these features, and also some of the hand crafted features, is that the spatial information is not taken into account. Some systems use only the mean pixel value (Zherdeva et al., 2016; Quinzán et al., 2013), others use a different feature or a combination of summary statistics features (Lorencs et al., 2016; Lihacova et al., 2018; Rey-Barroso et al., 2018; Patwardhan, Dhawan, & Relue, 2005; Nagaoka et al., 2015).

Each feature is calculated for each spectral band, and with a combination of a large set of features and many bands, dimensionality reduction can improve the performance of the system. With many spectral bands and/or features, some of the information is probably redundant, but each feature and band adds noise. Dimensionality reduction will reduce the noise and hence improve the classifier. If the number of images is small compared to the dimensionality of the images in a dataset, which is often the case for hyperspectral image datasets, a trained classifier will be unlikely to generalize well with regards to classifying samples not seen during training. The discriminatory power of a classifier initially increases as the number of feature dimensions increases, but then begins to decrease as the number of dimensions keeps increasing. This effect is often referred to as Hughes phenomenon (Shahshahani & Landgrebe, 1994). Therefore dimensionality reduction is beneficial

even if it reduces the amount of discriminatory information. The three main strategies for dimensionality reduction are band selection (Quinzán et al., 2013; Lorencs et al., 2016), feature subset selection (Patwardhan et al., 2005; Rey-Barroso et al., 2018; Starnes et al., 2017), and principal component analysis (PCA) (Kazianka, Leitner, & Pilz, 2008; Carrara et al., 2007; Tomatis et al., 2005). In spectral band selection and feature selection, a subset of the original bands and/or features are selected. This can be done by selecting a subset of bands, then the features are calculated for this subset (Quinzán et al., 2013; Lorencs et al., 2016). It can also be done in combination (Rey-Barroso et al., 2018), where the features are calculated for all bands and then the best band-feature pairs are selected, potentially keeping all spectral bands. The advantage of the first approach is that the number of bands are reduced, which can lead to a simpler camera construction in the future. In both approaches, the interpretability is kept intact. When PCA is employed to reduce the spatial or spectral dimensions, the result is a linear combination of features and spectral bands with different positive and negative weights, and the interpretability of the resulting PCA features is to some extent lost. It can be argued that interpretability is of lesser importance if the classifier is accurate enough, but so far there are no systems with accuracies high enough to justify a “black box” approach, given the potential fatal outcome of a misclassified melanoma.

## Selecting optimal spectral bands

A promising approach for dimensionality reduction of hyperspectral images is to reduce the number of spectral bands by selecting a subset of optimal wavelengths in a given hyperspectral image. This reduction can be performed by focusing on the spectral dimension of the captured image. In hyperspectral imaging, two main approaches have been proposed to reduce dimensionality; selection of spectral features or selection of spatial features (Dai, Cheng, Sun, & Zeng, 2015). On the one hand, spectral feature selection can be based on e.g. correlation analysis of the spectral bands. Reduction of the feature set can then be achieved by selecting those bands that provide the most salient statistical information. Different search strategies have been proposed for spectral feature selection; complete, heuristic, or random search. These search strategies have been used in conjunction with many algorithms such as branch and bound (BB) (Nakariyakul & Casasent, 2007), PCA (Xing, Bravo,

Jancsó, Ramon, & De Baerdemaeker, 2005), artificial neural networks (ElMasry, Wang, & Vigneault, 2009), and competitive adaptive re-weighted sampling (CARS) (Wu & Sun, 2013). On the other hand, spatial feature selection is focused on the selection of relevant image characteristics (color, shape, etc.) to discriminate the spectral bands that contain most information about the desired features. Investigations focusing on spectral and spatial feature selection in hyperspectral images up until now have been very limited, likely due to the small number of studies that have been carried out in the field of skin cancer detection using hyperspectral images as a whole.

Practical implementation of a hyperspectral imaging system is often challenging due to the complexity and cost of a hyperspectral camera capable of capturing several hundreds of bands. Recent publications have addressed new strategies for obtaining a feasible and practical technical solution by reducing the number of spectral bands or combining different finite spectral bands. A direct consequence of the reduction of the total amount of information processed, is the reduction of the computational requirements in a given algorithm. Reducing the spectra can also enable algorithms to operate in near real-time. The common approach up until now has been to obtain a hyperspectral image composed of hundreds of bands and then analyze which bands provides more information to classify and differentiate the skin tumor.

In Zherdeva et al. (2016) an experimental setup with hyperspectral images in the 450–750 nm range is employed to discriminate between skin cancers. Based on analysis of the collected images, it was determined that the most relevant differences between healthy tissue and skin cancer are located in the spectral bands 530–570 nm and 600–700 nm. These bands correspond to the absorption wavelengths of hemoglobin and melanin, respectively (Rubins, Zaharans, Lihačova, & Spigulis, 2014). It has been reported that hemoglobin concentration and the ratio of melanin in skin lesion tissue can be important biological markers for melanoma detection (MacKinnon, Vasefi, & Farkas, 2014; Vasefi et al., 2016).

A melanoma discriminator based on few spectral channels is proposed in Lorencs et al. (2016). The spectral band selection principles are based on a correlation study of the information contained between pixel values in optical density images of each pair of bands. The triplet of spectral bands at 540 nm, 640 nm and 740 nm and at 540 nm, 640 nm and



840 nm were selected as they presented the highest correlation values.

The algorithms described represent promising approaches in achieving feasible technical implementations for dermoscopic systems. Spectral band reduction, without degrading the performance of classifying skin lesions, speeds up both training and inference of associated algorithms and this can lead to near real-time operation of the overall system. This is an important characteristic for practical applications in clinical settings.

It is worth mentioning that hyperspectral image feature selection applied to skin cancer detection has been used in very few studies. Therefore, future investigation must be carried out to demonstrate the conclusions reported in the initial studies. Furthermore, reported applications of PCA on hyperspectral skin lesion images have not been focused on selecting optimal spectral bands. By using PCA to reduce the spatial dimensionality, which means applying PCA on each individual spectral band of an image, it should be possible to study which spectral bands are most salient (Yamal et al., 2012).

## **The power of spatio-spectral information**

Treating individual hyperspectral pixels in an image as independent observations from the same patient has certain advantages and disadvantages, both in statistical methodologies and machine learning. An immediate advantage is that an approach where individual pixels are classified will yield much bigger datasets for both training and testing, even with quite few images if they have large spatial dimensions. As an example, a small dataset consisting of 10 multi- or hyperspectral images with  $1000 \times 1000$  pixels, becomes a massive dataset of 10 million observations in a pixel-wise scheme. Using a pixel-wise approach to detect skin cancer has not been widely studied, but some research has been performed; skin lesion segmentation based on a pixel-wise scheme was done in Świtoński, Michalak, Josiński, and Wojciechowski (2010). One challenge of a pixel-wise scheme is how to balance the classes in the dataset, and how to ensure the sub-division into training and test sets are distributed in a representative way with respect to the original distribution, but still performed at random. Another challenge is getting accurate labels at the pixel level, which means that for each individual pixel a corresponding individual classification or diagnosis must be known.

Acquiring accurate, fine-grained labels at the pixel level is currently not feasible, and this means that training and testing supervised models with a pixel-wise scheme will be difficult. The lack of published research using pixel-wise approaches to skin cancer detection might be indicative of these challenges, suggesting that more research is needed in this area.

Although individual pixels can be treated as independent to some extent, in reality neighboring pixels in an image are spatially dependent. By not accounting for this, models and algorithms are deprived of salient information that could otherwise be used to improve their classification performance. As an illustrative example, a dermatologist applying the ABCD rule of dermoscopy will take into account all of the spatial information visible in a dermoscope or dermoscopic image. If only presented with individual pixels, without the spatial context, classifying the skin lesion would unquestionably be much more challenging. Therefore, statistical methods and machine learning models should also be trained with the same type of spatially-dependent data, either as full images or image patches.

Just like recent successful machine learning algorithms designed for clinical and dermoscopic RGB images of skin lesions are trained on full images with all three color channels (Esteva et al., 2017), exploiting the full potential of hyperspectral images involves using all of the spatial and spectral information encapsulated within the images. While this has to some extent been done in other applications of hyperspectral imaging, such as remote sensing (Makantasis, Karantzalos, Doulamis, & Doulamis, 2015; Chen, Zhao, & Jia, 2015; Mughees, Ali, & Tao, 2017), we are not aware of any published research in the area of skin cancer detection where all of the spatio-spectral information is used in a combined, fully contextual approach. The most common practise up until now has been to use hand-crafted features, summary statistics, and other lower-dimensional features. This can work reasonably well in some cases, but most, if not all, such approaches are incapable of fully accounting for the spatial and spectral context of detected patterns and features. Many deep learning models designed for image classification, e.g. convolutional neural networks (CNN) (Krizhevsky, Sutskever, & Hinton, 2012), are specifically tailored to learn robust feature detectors from data. The learned feature detectors (sometimes referred to as feature maps) have important traits such as translation equivariance. In simplified terms, a feature detector that has learned to detect e.g. eyes, will give the same activation response regardless of the spatial

location of the pixels comprising an eye, but if the pixels are spatially translated the activation will be translated respectively. For example, two activations of an “eye” feature in an image is not enough to detect the presence of a face, but two such feature activations in close, spatial proximity is a much stronger indication of a face. This is essentially how most CNN-based models learn to detect objects by synthesizing feature detectors from one layer into increasingly complex features in the next layer (Zeiler & Fergus, 2014). In Esteva et al. (2017) they develop a deep learning model that detects and classifies skin lesions using clinical and dermoscopic RGB images. One key component of their work is employing a technique referred to as transfer learning (Pan & Yang, 2010). More specifically, they perform fine-tuning of a pre-trained CNN model using a large dataset of RGB-based clinical and dermoscopic skin lesion images. Using this type of deep learning technique is feasible when the modality of the dataset used to train the original model is equivalent to the modality of the dataset used to perform the fine-tuning. No such pre-trained models for hyperspectral images are publicly available. This is likely one of the primary reasons why there exists no published research using deep learning methods on hyperspectral images for skin cancer detection. A concrete example of the spatio-spectral information richness and variation is depicted in Figure 6. From the figure it is clear that different physical properties of the lesion are captured at different wavelengths. How to exploit this information is not immediately obvious however. One suggestion is that such knowledge should be learned from data using deep learning models as opposed to being captured by hand-crafted feature extractors, or explicitly modeled in other ways.

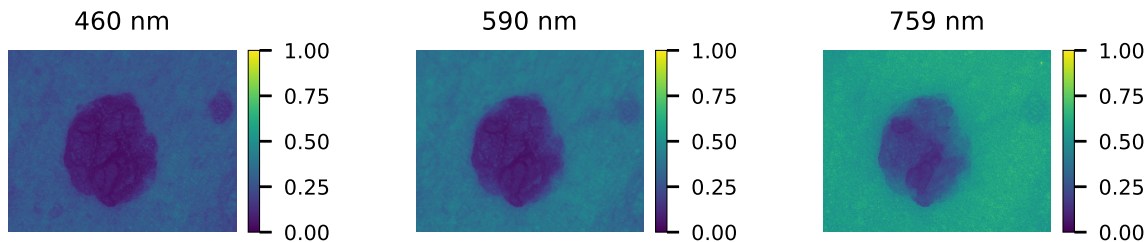


Figure 6: Examples of the different information captured in hyperspectral images at different wavelengths. Each image represents a reflectance image at a specific wavelength. Notice how certain features appear and disappear at the various wavelengths. In particular, notice how the small lesion visible near the right-most edge of the left image is almost invisible at higher wavelengths, and at higher wavelengths, smaller sub-regions and structures in the central lesion become visible.

Due to substantial differences in dataset modality, spatial dimensions, and number of channels/bands, transfer learning based on models trained on RGB images is not directly applicable to hyperspectral images. It has been shown feature detectors learned in the early layers of CNNs trained on RGB images are sensitive to colors (Zeiler & Fergus, 2014). In other words, these feature detectors have adapted to specific characteristics of RGB images, which are not trivially transferable to hyperspectral images. Therefore we believe that any model must either be trained from scratch, or novel RGB-to-hyperspectral transfer learning techniques must be developed. The number of trainable parameters in the most frequently used deep learning models designed for RGB images are in the order of  $10^7$ , sometimes as high as  $10^8$ . The number of trainable parameters in the first layer will increase substantially when the number of channels/bands in the images are increased by one or two orders of magnitude, which is the case when going from RGB to hyperspectral images. Additionally, many of the popular models are optimized for images with spatial dimensions around  $250 \times 250$  pixels and 3 color channels. Hyperspectral images used for skin cancer detection have much higher spatial resolution and many more channels. Given these differences, model architectures should be augmented for high-resolution hyperspectral images. Examples of such design adjustments might include increasing the number of learned feature detectors in each layer

of the model, and increasing the total number of layers. Increasing the complexity of the model translates into increasing the total number of trainable parameters.

Based on these observations, it is clear that training deep learning models for skin cancer detection using hyperspectral images from scratch will be challenging given current techniques and technology; it will require large amounts of computational power due to the high dimensionality of the images, and the high number of parameters being optimized during training. Training from scratch will also require sufficiently large datasets of high-quality, domain-specific, and representative observations in order for the model to generalize well at classification tasks. As far as we know, there are no publicly available datasets of hyperspectral skin lesion images that are sufficiently large to train deep learning models for skin lesion classification. The lack of publicly available datasets and pre-trained models are likely the key challenges that explain the lack of published research on deep learning methods for skin cancer detection using hyperspectral images.

## **Critical remarks and analysis of published results**

In the context of cancer detection, the ideal system provides a class label for each image in accordance with the actual pathology of the lesion in question. The gold standard for skin lesion diagnosis is histopathology for excised lesions and dermoscopic evaluation for non-excised lesions. Note that a non-biopsied lesion can only have a benign diagnosis, as suspicion of malignancy automatically leads to excision and histopathological examination. Due to the potential fatal consequences of misclassifying a melanoma as benign, even low level of suspicion leads to excision.

A system that aims at clinical relevance must either have melanoma sensitivity close to 100% combined with a reasonable specificity, or provide information that benefits the physician in the decision on whether to excise the lesion in question. Both objectives have shown to be hard to achieve, and so far no system can be said to have achieved either.

To predict the performance on future data, which do not have class labels, a system is tested on either an independent test set or by the use of cross-validation. For the outcome to be valid, the test set must be independent of all aspects of the system development,

from bandwidth selection to classifier parameter settings. In addition, the test set must be large enough for the result to be generalizable, and reflect the population from where the future data will be collected. These standards can be difficult to achieve due to the nature of the problem at hand: hyperspectral cameras are expensive, require training to operate, and melanomas are rare but fatal. This results in small datasets, and combined with high dimensionality there is often not enough data for sufficient training and adequate testing. The differences in imaging acquisition systems hinder combining different datasets.

Several publications report performance on the same set of data that were used to develop the system (Zherdeva et al., 2016; Lihacova et al., 2018; Rey-Barroso et al., 2018; Kazianka et al., 2008; Lorencs et al., 2016), which give highly optimistic results. This bias does not only apply when the test set is used to train the classification algorithm itself, but applies for all parts of system development, including bandwidth selection (Quinzán et al., 2013), feature selection (Patwardhan et al., 2005; Stamnes et al., 2017), and post hoc threshold settings for classification (Patwardhan et al., 2005; Nagaoka et al., 2015). The impact might not be obvious, but it is indisputable (Smialowski, Frishman, & Kramer, 2010).

In an effort to overcome the limitations of a small dataset, cross-validation have been used (Nagaoka et al., 2015; Quinzán et al., 2013), but when using the entire dataset for bandwidth or feature selection, or parameter setting, the results are invalid.

It can be argued that incorrect use of statistical tools not necessarily disregards the results altogether, but the drop in performance is usually dramatic. The performance of MelaFind dropped from 85% specificity (Elbaum et al., 2001) to 9% specificity for near 100% sensitivity, when tested on a proper independent test set (Monheit et al., 2011). For more examples, see Møllersen et al. (2015).

The reported classification results with independent test sets are:

Publication	Sensitivity (%)	Specificity (%)	# Melanomas	# Lesions in Total
Song et al. (2016)	50	23	4	55
Nagaoka et al. (2015)	75	97	24	132
Tomatis et al. (2005)	80	90	41	1369
Carrara et al. (2007)	95	53	76	1208

The study of Song et al. (2016) tested MelaFind in a clinical setting, but contains only 4 melanomas, and the results are therefore not generalizable. The Nagaoka et al. (2015) study consisted of 24 melanomas and 108 other skin lesions, but the lesions are from both patients and volunteers, and can therefore not be said to reflect any future population. Tomatis et al. (2005) had a large dataset with excised lesions consecutively collected, and in addition non-excised lesions that were randomly collected in a clinical setting. Ideally, both the excised and non-excised lesions should have been consecutively collected, but compared to other datasets in the field of computer-aided skin lesion classification, this dataset has high quality. The test set consisted of 41 melanomas and 306 non-melanomas, confirmed by histopathology, and 1022 lesions that were diagnosed as benign without excision. When using only the set of excised lesions, the specificity dropped to 77%, which clearly shows the enormous impact that the inclusion criteria for the dataset can have on the result. Carrara et al. (2007) reported in their study the sensitivity and specificity to whether a lesion should be excised, with the dermatologist’s decision as ground truth. The numbers reported here are according to melanoma/non-melanoma classes. Note the Tomatis et al. (2005) and Carrara et al. (2007) studies use overlapping datasets and methods.

The reported performance of a system will vary from one test set to another due to its random nature. The Clopper-Pearson confidence interval for the 95% sensitivity in the study of Carrara et al. (2007) is 87-99%, which clearly demonstrates the need for large test sets for reliable results.

The common practice of reporting of a single sensitivity-specificity pair makes comparison between systems impossible. The high specificity reported by Tomatis et al. (2005) drops when the sensitivity is increased, as shown in their receiver operating characteristic (ROC) curve, which shows the specificity as a function of sensitivity. The curve is not detailed enough to extract the exact numbers. The reported 80% sensitivity of Tomatis et al. (2005), which corresponds to missing 1 out of 5 melanomas, is not relevant for a system intended for clinical use. A 95% sensitivity corresponds to missing 1 out of 20 melanomas, and might still not be high enough. There is no consensus for a lower limit for acceptable melanoma sensitivity, and therefore, to make comparisons between systems possible, the range of corresponding specificities for sensitivities from 95% to 100% should be reported. As shown in



Møllersen, Zortea, Schopf, Kirchesch, and Godtlielsen (2017), the criterion for comparing different systems has huge impact on the resulting ranking. Summary performance measures such as the area under the ROC curve (AUC), does not distinguish between the two types of misclassifications; a system can have high AUC even if its ability to detect skin cancer is poor. This is not suitable in settings where a false negative (misclassifying a melanoma as benign) has much graver consequences than a false positive (misclassifying a benign lesion as malignant).

## **Future research directions**

### **Long-term goals**

In recent years, early detection of skin cancer using RGB images has been research focus in a large number of publications, see e.g., Oliveira, Papa, Pereira, and Tavares (2018). Although the findings presented in Codella et al. (2018) and Esteva et al. (2017) are very promising, they are still not able to outperform experienced dermatologists. Future research using RGB images will likely suffer from effects equivalent to the law of diminishing returns, and because of this additional information richness is crucial to boost classification results even further.

The ultimate goal is to obtain classification systems that can lower the number of deaths caused by skin cancer significantly. A successful classification system will benefit from research in the following two directions.

Firstly, there is a need to acquire a large quantity of high-quality data for all relevant skin cancers to be able to develop a successful classification system. Any database for clinical evaluation should be large enough to be able to provide good generalization, and hence reflecting the high-variability of data. This generalization is even more challenging in skin analysis, where the inter-patient variability across different pigmented skin lesions is also influenced by the different skin phenotypes. By acquiring RGB and hyperspectral images for all cases, it will also be possible to give a more objective answer to the proposed importance of hyperspectral information. Clearly, it will take many years before a sufficient number of datasets are available, but with such datasets available, the Common Task Framework

described by Donoho (2017) can be used to obtain the best possible classification systems. After such results are available, clinical testing needs to be carried out before the whole system can be put into use.

Secondly, patients can contribute to earlier detection of harmful skin lesions by keeping an eye on the evolution of their skin lesions. A natural first step is therefore to design a system that can be used for monitoring skin lesions. Ideally, such a system should be precise, affordable, easy to use and interpretable. By designing a system like this, early detection of skin cancer will hopefully be significantly improved since one of the reasons for skin cancer related death is the lack of early treatment. A successful monitoring system may result in earlier and more effective treatment, thereby reducing the number of deaths.

## **Short-term goals**

Although there exist several papers (Qi, Xing, Foran, & Yang, 2011; Taghizadeh, Gowen, & O'Donnell, 2011) that indicate that hyperspectral images contain information beyond RGB images, it seems natural to start with careful analyses that show how much and in what way hyperspectral information contributes in various types of classification algorithms. based in both statistical methodologies and machine learning.

Spatial and hyperspectral information gives a natural link to spatio-temporal methods and it is therefore natural to look into how such methods can be useful in the present task. In particular, there are links to image sequences in other applications of medicine. One example is fMRI where an important aim is to find areas of the brain connected to specific tasks. This may, e.g. be crucial in connection with brain surgery. Similar ideas could potentially be used to find “suspicious areas” that may be an indicator of a serious change in a skin lesion. Research in this direction should be performed in close collaboration with dermatologists, and may turn out to be well worthwhile since it could give rise to a boost in early detection of skin cancer. Another possibility is to look for particular shapes or features in the hyperspectral curves, thereby giving rise to important new features in a future classification rule.

Clustering of the hyperspectral signatures that gives rise to specific RGB values will give a potential link between RGB and hyperspectral images. This will give important knowledge

about how homogeneous such clusters are, and it may also lead to a better understanding of the extra information obtained by hyperspectral signatures.

When the research community have gathered a large number of images, these datasets may be used to learn the characteristics of each class. One important research area here would be to see if hyperspectral images could be used to better distinguish between melanoma and other types of skin cancer. This would be an extremely important result since melanomas are fatal, whereas some types of non-melanoma skin lesions are considered harmless. Dermatologists are able to distinguish these classes well, but this can be a very difficult task for general practitioners.

Preliminary results (Li, Zhou, Liu, Wang, & Guo, 2015; Q. Wang, Wang, Zhou, Li, & Wang, 2017; Ortega et al., 2018) indicate that pathology results can be improved both with respect to precision and time using hyperspectral imaging. Further investigations are needed to confirm this and to get a better understanding of how this new technology can be beneficial for this purpose.

Analysis of dermatological hyperspectral images is in our opinion the most important area for research in the near future. Monitoring the evolution of skin lesions over time is an important part of such research. In addition, it is important to analyze hyperspectral images using a large number of statistical tools, thereby gaining more knowledge about such data and be in better position to design classification systems when sufficiently large datasets become available.

For future classification systems, finding optimal data representation is a key to success. Also known as feature learning, this is the task of finding a representation of the input that will result in the best possible performance of the classification algorithm (Bengio, Courville, & Vincent, 2013). In the application at hand, the skin lesion's state is partially represented by the curves measured by the hyperspectral camera. To this end, we seek a way to represent the rich data contained in the skin lesion state that will result in a successful algorithm.

Using deep learning for reducing the dimensionality of hyperspectral images is believed to be an important field of research. Instead of using methodologies based on variance analysis, entropy, or other information measures, we suggest that learning robust lower-dimensional representations of the data using e.g. deep autoencoders (Hinton & Salakhutdinov, 2006)

could lead to better classification performance. The spatio-spectral information encoded in hyperspectral images is complex, and it is not immediately obvious that conventional methods such as PCA are sufficiently capable of capturing this. Furthermore, learning shared representations might make it feasible to combine hyperspectral skin lesion datasets (Ngiam et al., 2011). Researching the potential gains of using deep learning approaches for dimensionality reduction could yield extremely important results.

The incorporation of scale-space ideas can also be explored in obtaining an efficient state space representation. Scale-space theory is a framework for representing signals on multiple scales, developed by the computer vision, image processing and signal processing communities. Scale-space ideas could be used to select tuning parameters in the FDA approach. For a basis expansion representation of hyperspectral curves, for instance, several key parameters (e.g. bandwidth, degree of the derivative) must be selected (Chaudhuri & Marron, 2000). As the representation may be very sensitive to these parameters, scale-space methods can provide useful insight. For instance, SiZer is a visual tool to examine when the derivative of a scatterplot-smoother is significantly negative, possibly zero or significantly positive across a range of smoothing bandwidths (Chaudhuri & Marron, 1999).

## Concluding remarks

Recent advances in hyperspectral imaging for skin cancer detection show great promise, and we believe that further research can lead to a significant reduction in the number of deaths caused by skin cancer. However, there are still many open research questions that must be addressed, such as what are the benefits of training classifiers with hyperspectral skin lesion images as opposed to clinical and dermoscopic images of skin lesions captured with conventional RGB cameras. To answer this, large, high-quality datasets of skin lesion images need to be collected using both hyperspectral and conventional RGB cameras. Importantly, both types of images need to be collected from all observed skin lesions in order to make it possible to perform e.g. statistical analysis, and to compare classification algorithms trained on both types of images. Once enough data has been collected, the data can be analyzed using statistical methodologies such as functional data analysis, multivariate analysis, etc.

Furthermore, classification algorithms can be trained using conventional statistical model-based methodologies and more recent developments based on deep learning approaches. How to architect and optimize algorithms and models for skin cancer detection using hyperspectral imaging needs to be discovered by further research. Hyperspectral imaging is widely used in other fields of research such as remote sensing, and such research should provide a good foundation on which to build future research efforts towards skin cancer detection.

For reported performance results of classification systems to be valid and reliable, to ease comparison between systems, and to ensure that the clinical aspect is not ignored, we have the following recommendations for data collection and statistical analysis of the results:

1. Use an independent test set, not cross-validation.
2. Report specificities for sensitivities from 95% to 100%.
3. Collect data in a clinical-like setting, with clearly stated inclusion and exclusion criteria.

The data should be collected consecutively to reflect the underlying distribution of the population in question (e.g., hospital patients, primary care patients, etc.).

4. Report confidence intervals for the sensitivities.
5. If the available dataset is too small for independent test set, other aspects of the system such as spectral band selection or feature selection can be reported instead.

For a more detailed list that will increase the quality of a study even further, see Rosado et al. (2003).

## **Funding information**

This work has been supported in part by the Canary Islands Government through the ACIISI (Canarian Agency for Research, Innovation and the Information Society), ITHACA project “Hyperspectral Identification of Brain Tumors” under Grant Agreement ProID2017010164 and it has been partially supported also by the Spanish Government and European Union (FEDER funds) as part of support program in the context of Distributed HW/SW Platform for Intelligent Processing of Heterogeneous Sensor Data in Large Open Areas Surveillance Applications (PLATINO) project, under contract TEC2017-86722-C4-1-R. Additionally, this work was completed while Samuel Ortega was beneficiary of a pre-doctoral grant given by

the “Agencia Canaria de Investigacion, Innovacion y Sociedad de la Información (ACIISI)” of the “Conserjería de Economía, Industria, Comercio y Conocimiento” of the “Gobierno de Canarias”, which is part-financed by the European Social Fund (FSE) (POC 2014–2020, Eje 3 Tema Prioritario 74(85%)). Finally, this work has been also supported in part by the 2016 PhD Training Program for Research Staff of the University of Las Palmas de Gran Canaria.

The project has also partially been funded by grant A33020 from Tromsø Forskningsstiftelse.

## Research resources

The camera prototype used for capturing the hyperspectral images used in our example figures was provided by Revenio Research Oy.

## Acknowledgements

We would like to thank Dr. Herbert Kirchesch for his help in collecting all of the conventional RGB images and hyperspectral images used in our example figures.

## References

- Abbasi, N. R., Shaw, H. M., Rigel, D. S., Friedman, R. J., McCarthy, W. H., Osman, I., . . . Polsky, D. (2004). Early Diagnosis of Cutaneous Melanoma. *JAMA*, *292*(22).  
doi:[10.1001/jama.292.22.2771](https://doi.org/10.1001/jama.292.22.2771)
- Ahnlide, I., Bjellerup, M., Nilsson, F., & Nielsen, K. (2016). Validity of ABCD Rule of Dermoscopy in Clinical Practice. *Acta Dermato Venereologica*, *96*(3), 367–372.  
doi:[10.2340/00015555-2239](https://doi.org/10.2340/00015555-2239)
- American Cancer Society. (2018). Cancer Facts and Figures 2018. Retrieved from <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2018.html>

- Annessi, G., Bono, R., Sampogna, F., Faraggiana, T., & Abeni, D. (2007). Sensitivity, specificity, and diagnostic accuracy of three dermoscopic algorithmic methods in the diagnosis of doubtful melanocytic lesions. *Journal of the American Academy of Dermatology*, *56*(5), 759–767. doi:[10.1016/j.jaad.2007.01.014](https://doi.org/10.1016/j.jaad.2007.01.014)
- Bengio, Y., Courville, A., & Vincent, P. (2013). Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. doi:[10.1109/TPAMI.2013.50](https://doi.org/10.1109/TPAMI.2013.50)
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *00*(00), 1–31. doi:[10.3322/caac.21492](https://doi.org/10.3322/caac.21492)
- Carrara, M., Bono, A., Bartoli, C., Colombo, A., Lualdi, M., Moglia, D., . . . Marchesini, R. (2007). Multispectral imaging and artificial neural network: mimicking the management decision of the clinician facing pigmented skin lesions. *Physics in Medicine and Biology*, *52*(9), 2599–2613. doi:[10.1088/0031-9155/52/9/018](https://doi.org/10.1088/0031-9155/52/9/018)
- Chaudhuri, P. & Marron, J. S. (1999). SiZer for Exploration of Structures in Curves. *Journal of the American Statistical Association*, *94*(447), 807–823. doi:[10.1080/01621459.1999.10474186](https://doi.org/10.1080/01621459.1999.10474186)
- Chaudhuri, P. & Marron, J. S. (2000). Scale space view of curve estimation. *The Annals of Statistics*, *28*(2), 408–428. doi:[10.1214/aos/1016218224](https://doi.org/10.1214/aos/1016218224)
- Chen, Y., Zhao, X., & Jia, X. (2015). Spectral–Spatial Classification of Hyperspectral Data Based on Deep Belief Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(6), 2381–2392. doi:[10.1109/JSTARS.2015.2388577](https://doi.org/10.1109/JSTARS.2015.2388577)
- Codella, N. C. F., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., . . . Halpern, A. (2018). Skin lesion analysis toward melanoma detection: A challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 168–172). doi:[10.1109/ISBI.2018.8363547](https://doi.org/10.1109/ISBI.2018.8363547)



- Dai, Q., Cheng, J.-H., Sun, D.-W., & Zeng, X.-A. (2015). Advances in Feature Selection Methods for Hyperspectral Image Processing in Food Industry Applications: A Review. *Critical Reviews in Food Science and Nutrition*, *55*(10), 1368–1382. doi:[10.1080/10408398.2013.871692](https://doi.org/10.1080/10408398.2013.871692)
- Donoho, D. (2017). 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766. doi:[10.1080/10618600.2017.1384734](https://doi.org/10.1080/10618600.2017.1384734)
- Elbaum, M., Kopf, A. W., Rabinovitz, H. S., Langley, R. G., Kamino, H., Mihm, M. C., . . . Wang, S. (2001). Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: A feasibility study. *Journal of the American Academy of Dermatology*, *44*(2), 207–218. doi:[10.1067/mjd.2001.110395](https://doi.org/10.1067/mjd.2001.110395)
- ElMasry, G., Wang, N., & Vigneault, C. (2009). Detecting chilling injury in Red Delicious apple using hyperspectral imaging and neural networks. *Postharvest Biology and Technology*, *52*(1), 1–8. doi:[10.1016/j.postharvbio.2008.11.008](https://doi.org/10.1016/j.postharvbio.2008.11.008)
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. doi:[10.1038/nature21056](https://doi.org/10.1038/nature21056)
- Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., . . . Bray, F. (2013). *GLOBOCAN 2012 v1.0, cancer incidence and mortality worldwide: IARC cancerbase no. 11 [internet]*. International Agency for Research on Cancer. Retrieved from <http://globocan.iarc.fr>
- Hinton, G. E. & Salakhutdinov, R. R. (2006). *Reducing the Dimensionality of Data with Neural Networks* (tech. rep. No. 5786). doi:[10.1126/science.1127647](https://doi.org/10.1126/science.1127647)
- Jet Propulsion Laboratory, California Institute of Technology. (n.d.). Airborne Visible InfraRed Imaging Spectrometer (AVIRIS) - Imaging Spectroscopy. Retrieved from [https://aviris.jpl.nasa.gov/aviris/imaging%7B%5C\\_%7Dspectroscopy.html](https://aviris.jpl.nasa.gov/aviris/imaging%7B%5C_%7Dspectroscopy.html)
- K. C. Lawrence, B. Park, W. R. Windham, & C. Mao. (2003). Calibration of a pushbroom hyperspectral imaging system for agricultural inspection. *Transactions of the ASAE*, *46*(2). doi:[10.13031/2013.12940](https://doi.org/10.13031/2013.12940)

- Kazianka, H., Leitner, R., & Pilz, J. (2008). Segmentation and classification of hyper-spectral skin data. *Data Analysis, Machine Learning and Applications*, 245–252. doi:[10.1007/978-3-540-78246-9\\_29](https://doi.org/10.1007/978-3-540-78246-9_29)
- Korotkov, K. & Garcia, R. (2012). Computerized analysis of pigmented skin lesions: A review. *Artificial Intelligence in Medicine*, 56(2), 69–90. doi:[10.1016/j.artmed.2012.08.002](https://doi.org/10.1016/j.artmed.2012.08.002)
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in neural information processing systems 25 (nips 2012)*. doi:[10.1145/3065386](https://doi.org/10.1145/3065386)
- Kupetsky, E. A. & Ferris, L. K. (2013). The diagnostic evaluation of MelaFind multi-spectral objective computer vision system. *Expert Opinion on Medical Diagnostics*, 7(4), 405–411. doi:[10.1517/17530059.2013.785520](https://doi.org/10.1517/17530059.2013.785520)
- Lachenal, G. & Ozaki, Y. (1999). Advantages of near infrared spectroscopy for the analysis of polymers and composites. *Macromolecular Symposia*, 141(1), 283–292. doi:[10.1002/masy.19991410123](https://doi.org/10.1002/masy.19991410123)
- Lee, K.-S., Cohen, W. B., Kennedy, R. E., Maiersperger, T. K., & Gower, S. T. (2004). Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. *Remote Sensing of Environment*, 91(3-4), 508–520. doi:[10.1016/j.rse.2004.04.010](https://doi.org/10.1016/j.rse.2004.04.010)
- Li, Q., He, X., Wang, Y., Liu, H., Xu, D., & Guo, F. (2013). Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of Biomedical Optics*, 18(10). doi:[10.1117/1.JBO.18.10.100901](https://doi.org/10.1117/1.JBO.18.10.100901)
- Li, Q., Zhou, M., Liu, H., Wang, Y., & Guo, F. (2015). Red Blood Cell Count Automation Using Microscopic Hyperspectral Imaging Technology. *Applied Spectroscopy*, 69(12), 1372–1380. doi:[10.1366/14-07766](https://doi.org/10.1366/14-07766)
- Lihacova, I., Bolochko, K., Plorina, E. V., Lange, M., Lihachev, A., Bliznuks, D., & Derjabo, A. (2018). A method for skin malformation classification by combining multispectral and skin autofluorescence imaging. In J. Popp, V. V. Tuchin, & F. S. Pavone (Eds.), *Biophotonics: Photonic solutions for better health care vi* (Vol. 1068535, p. 113). SPIE. doi:[10.1117/12.2306203](https://doi.org/10.1117/12.2306203)

- Lorencs, A., Sinica-Sinavskis, J., Jakovels, D., & Mednieks, I. (2016). Melanoma-nevus discrimination based on image statistics in few spectral channels. *Elektronika ir Elektrotechnika*, 22(2), 66–72. doi:[10.5755/j01.eie.22.2.12173](https://doi.org/10.5755/j01.eie.22.2.12173)
- Lu, G. & Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1). doi:[10.1117/1.JBO.19.1.010901](https://doi.org/10.1117/1.JBO.19.1.010901)
- MacKinnon, N., Vasefi, F., & Farkas, D. L. (2014). Toward in vivo diagnosis of skin cancer using multimode imaging dermoscopy: (I) clinical system development and validation. *8947(1)*. doi:[10.1117/12.2041818](https://doi.org/10.1117/12.2041818)
- Makantasis, K., Karantzalos, K., Doulamis, A., & Doulamis, N. (2015). Deep supervised learning for hyperspectral data classification through convolutional neural networks. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (Vol. 2015-Novem, pp. 4959–4962). IEEE. doi:[10.1109/IGARSS.2015.7326945](https://doi.org/10.1109/IGARSS.2015.7326945)
- Møllersen, K., Kirchesch, H., Zortea, M., Schopf, T. R., Hindberg, K., & Godtlielsen, F. (2015). Computer-Aided Decision Support for Melanoma Detection Applied on Melanocytic and Nonmelanocytic Skin Lesions: A Comparison of Two Systems Based on Automatic Analysis of Dermoscopic Images. *BioMed Research International*, 2015, 1–8. doi:[10.1155/2015/579282](https://doi.org/10.1155/2015/579282)
- Møllersen, K., Zortea, M., Schopf, T. R., Kirchesch, H., & Godtlielsen, F. (2017). Comparison of computer systems and ranking criteria for automatic melanoma detection in dermoscopic images. *PLoS ONE*, 12(12). doi:[10.1371/journal.pone.0190112](https://doi.org/10.1371/journal.pone.0190112)
- Moncrieff, M., Cotton, S., Claridge, E., & Hall, P. (2002). Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. *British Journal of Dermatology*, 146(3), 448–457. doi:[10.1046/j.1365-2133.2002.04569.x](https://doi.org/10.1046/j.1365-2133.2002.04569.x)
- Monheit, G., Cognetta, A. B., Ferris, L., Rabinovitz, H., Gross, K., Martini, M., ... Gutkowitz-Krusin, D. (2011). The Performance of {MelaFind}: A Prospective Multicenter Study. *Archives of Dermatology*, 147(2), 188–194. doi:[10.1001/archdermatol.2010.302](https://doi.org/10.1001/archdermatol.2010.302)
- Mughees, A., Ali, A., & Tao, L. (2017). Hyperspectral image classification via shape-adaptive deep learning. In *2017 IEEE International Conference on Image*

- processing (icip)* (Vol. 2017-Septe, pp. 375–379). IEEE.  
doi:[10.1109/ICIP.2017.8296306](https://doi.org/10.1109/ICIP.2017.8296306)
- Nachbar, F., Stolz, W., Merkle, T., Cognetta, A. B., Vogt, T., Landthaler, M., . . . Plewig, G. (1994). The {ABCD} rule of dermatoscopy. {H}igh prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, *30*(4), 551–559. doi:[10.1016/S0190-9622\(94\)70061-3](https://doi.org/10.1016/S0190-9622(94)70061-3)
- Nagaoka, T., Kiyohara, Y., Koga, H., Nakamura, A., Saida, T., & Sota, T. (2015). Modification of a melanoma discrimination index derived from hyperspectral data: A clinical trial conducted in 2 centers between March 2011 and December 2013. *Skin Research and Technology*, *21*(3), 278–283. doi:[10.1111/srt.12188](https://doi.org/10.1111/srt.12188)
- Nagaoka, T., Nakamura, A., Kiyohara, Y., & Sota, T. (2012). Melanoma screening system using hyperspectral imager attached to imaging fiberscope. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *30*, 3728–3731. doi:[10.1109/EMBC.2012.6346777](https://doi.org/10.1109/EMBC.2012.6346777)
- Nagaoka, T., Nakamura, A., Okutani, H., Kiyohara, Y., Koga, H., Saida, T., & Sota, T. (2013). Hyperspectroscopic screening of melanoma on acral volar skin. *Skin Research and Technology*, *19*(1), 290–296. doi:[10.1111/j.1600-0846.2012.00642.x](https://doi.org/10.1111/j.1600-0846.2012.00642.x)
- Nagaoka, T., Nakamura, A., Okutani, H., Kiyohara, Y., & Sota, T. (2012). A possible melanoma discrimination index based on hyperspectral data: A pilot study. *Skin Research and Technology*, *18*(3), 301–310. doi:[10.1111/j.1600-0846.2011.00571.x](https://doi.org/10.1111/j.1600-0846.2011.00571.x)
- Nakariyakul, S. & Casasent, D. P. (2007). Adaptive branch and bound algorithm for selecting optimal features. *Pattern Recognition Letters*, *28*(12), 1415–1427.  
doi:[10.1016/j.patrec.2007.02.015](https://doi.org/10.1016/j.patrec.2007.02.015)
- Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal Deep Learning. In *Proceedings of the 28th international conference on machine learning (icml-11)* (pp. 689–696).
- Oliveira, R. B., Papa, J. P., Pereira, A. S., & Tavares, J. M. R. S. (2018). Computational methods for pigmented skin lesion classification in images: review and future trends. *Neural Computing and Applications*, *29*(3), 613–636. doi:[10.1007/s00521-016-2482-6](https://doi.org/10.1007/s00521-016-2482-6)

- Ortega, S., Fabelo, H., Camacho, R., de la Luz Plaza, M., Callico, G. M., & Sarmiento, R. (2018). Detecting brain tumor in pathological slides using hyperspectral imaging. *Biomedical Optics Express*, *9*(2), 818. doi:[10.1364/BOE.9.000818](https://doi.org/10.1364/BOE.9.000818)
- Pan, S. J. & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. doi:[10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191)
- Patwardhan, S. V. & Dhawan, A. P. (2004). Multi-spectral imaging and analysis for classification of melanoma. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *1*, 503–6. doi:[10.1109/IEMBS.2004.1403204](https://doi.org/10.1109/IEMBS.2004.1403204)
- Patwardhan, S. V., Dhawan, A. P., & Relue, P. A. (2005). Monte Carlo simulation of light-tissue interaction: Three-dimensional simulation for trans-illumination-based imaging of skin lesions. *IEEE Transactions on Biomedical Engineering*, *52*(7), 1227–1236. doi:[10.1109/TBME.2005.847546](https://doi.org/10.1109/TBME.2005.847546)
- Qi, X., Xing, F., Foran, D. J., & Yang, L. (2011). Comparative performance analysis of stained histopathology specimens using RGB and multispectral imaging. (Vol. 7963). doi:[10.1117/12.878325](https://doi.org/10.1117/12.878325)
- Quinzán, I., Sotoca, J. M., Latorre-Carmona, P., Pla, F., García-Sevilla, P., & Boldó, E. (2013). Band selection in spectral imaging for non-invasive melanoma diagnosis. *Biomedical Optics Express*, *4*(4), 514. doi:[10.1364/BOE.4.000514](https://doi.org/10.1364/BOE.4.000514)
- Rey-Barroso, L., Burgos-Fernández, F., Delpueyo, X., Ares, M., Royo, S., Malvehy, J., ... Vilaseca, M. (2018). Visible and Extended Near-Infrared Multispectral Imaging for Skin Cancer Diagnosis. *Sensors*, *18*(5), 1441. doi:[10.3390/s18051441](https://doi.org/10.3390/s18051441)
- Rosado, B., Menzies, S., Harbauer, A., Pehamberger, H., Wolff, K., Binder, M., & Kittler, H. (2003). Accuracy of Computer Diagnosis of Melanoma: A Quantitative Meta-analysis. *Archives of Dermatology*, *139*(3), 361–367. doi:[10.1001/archderm.139.3.361](https://doi.org/10.1001/archderm.139.3.361)
- Rubins, U., Zaharans, J., Ļihačova, I., & Spigulis, J. (2014). Multispectral Video-Microscope Modified for Skin Diagnostics. *Latvian Journal of Physics and Technical Sciences*, *51*(5), 65–70. doi:[10.2478/lpts-2014-0031](https://doi.org/10.2478/lpts-2014-0031)

- Shahshahani, B. M. & Landgrebe, D. A. (1994). The Effect of Unlabeled Samples in Reducing the Small Sample Size Problem and Mitigating the Hughes Phenomenon. *IEEE Transactions on Geoscience and Remote Sensing*, *32*(5), 1087–1095.  
doi:[10.1109/36.312897](https://doi.org/10.1109/36.312897)
- Smialowski, P., Frishman, D., & Kramer, S. (2010). Pitfalls of supervised feature selection. *Bioinformatics*, *26*(3), 440–443. doi:[10.1093/bioinformatics/btp621](https://doi.org/10.1093/bioinformatics/btp621)
- Smith, R. B. (2012). Introduction to Hyperspectral Imaging. Retrieved June 26, 2018, from <https://www.microimages.com/documentation/Tutorials/hyprspec.pdf>
- Song, E., Grant-Kels, J. M., Swede, H., D'Antonio, J. L., Lachance, A., Dadras, S. S., ... Rothe, M. J. (2016). Paired comparison of the sensitivity and specificity of multispectral digital skin lesion analysis and reflectance confocal microscopy in the detection of melanoma in vivo: A cross-sectional study. *Journal of the American Academy of Dermatology*, *75*(6), 1187–1192.e2. doi:[10.1016/j.jaad.2016.07.022](https://doi.org/10.1016/j.jaad.2016.07.022)
- Stamnes, J. J., Ryzhikov, G., Biryulina, M., Hamre, B., Zhao, L., & Stamnes, K. (2017). Optical detection and monitoring of pigmented skin lesions. *Biomedical Optics Express*, *8*(6), 2946. doi:[10.1364/BOE.8.002946](https://doi.org/10.1364/BOE.8.002946)
- Suárez, I. Q., Carmona, P. L., García-Sevilla, P., Boldo, E., Pla, F., Jiménez, V. G., ... de Lucía, G. P. (2012). Non-invasive Melanoma Diagnosis using Multispectral Imaging. In *Proceedings of the 1st international conference on pattern recognition applications and methods* (January, pp. 386–393). SciTePress - Science.  
doi:[10.5220/0003843803860393](https://doi.org/10.5220/0003843803860393)
- Świtoński, A., Michalak, M., Josiński, H., & Wojciechowski, K. (2010). Detection of Tumor Tissue Based on the Multispectral Imaging. (Vol. 1732, pp. 325–333).  
doi:[10.1007/978-3-642-15907-7\\_40](https://doi.org/10.1007/978-3-642-15907-7_40)
- Taghizadeh, M., Gowen, A. A., & O'Donnell, C. P. (2011). Comparison of hyperspectral imaging with conventional RGB imaging for quality evaluation of *Agaricus bisporus* mushrooms. *Biosystems Engineering*, *108*(2), 191–194.  
doi:[10.1016/j.biosystemseng.2010.10.005](https://doi.org/10.1016/j.biosystemseng.2010.10.005)

- Tomatis, S., Bono, A., Bartoli, C., Carrara, M., Lualdi, M., Tragni, G., & Marchesini, R. (2003). Automated melanoma detection: Multispectral imaging and neural network approach for classification. *Medical Physics*, *30*(2), 212–221. doi:[10.1118/1.1538230](https://doi.org/10.1118/1.1538230)
- Tomatis, S., Carrara, M., Bono, A., Bartoli, C., Lualdi, M., Tragni, G., . . . Marchesini, R. (2005). Automated melanoma detection with a novel multispectral imaging system: Results of a prospective study. *Physics in Medicine and Biology*, *50*(8), 1675–1687. doi:[10.1088/0031-9155/50/8/004](https://doi.org/10.1088/0031-9155/50/8/004)
- Tuia, D., Volpi, M., Copa, L., Kanevski, M., & Munoz-Mari, J. (2011). A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification. *IEEE Journal of Selected Topics in Signal Processing*, *5*(3), 606–617. doi:[10.1109/JSTSP.2011.2139193](https://doi.org/10.1109/JSTSP.2011.2139193)
- Unlu, E., Akay, B. N., & Erdem, C. (2014). Comparison of dermatoscopic diagnostic algorithms based on calculation: The ABCD rule of dermatoscopy, the seven-point checklist, the three-point checklist and the CASH algorithm in dermatoscopic evaluation of melanocytic lesions. *The Journal of Dermatology*, *41*(7), 598–603. doi:[10.1111/1346-8138.12491](https://doi.org/10.1111/1346-8138.12491)
- Vasefi, F., MacKinnon, N., Saager, R., Kelly, K. M., Maly, T., Booth, N., . . . Farkas, D. L. (2016). Separating melanin from hemodynamics in nevi using multimode hyperspectral dermoscopy and spatial frequency domain spectroscopy. *Journal of Biomedical Optics*, *21*(11), 114001. doi:[10.1117/1.JBO.21.11.114001](https://doi.org/10.1117/1.JBO.21.11.114001)
- Vestergaard, M. E. & Menzies, S. W. (2008). Automated diagnostic instruments for cutaneous melanoma. *Seminars in cutaneous medicine and surgery*, *27*(1), 32–36.
- Wang, Q., Wang, J., Zhou, M., Li, Q., & Wang, Y. (2017). Spectral-spatial feature-based neural network method for acute lymphoblastic leukemia cell identification via microscopic hyperspectral imaging technology. *Biomedical Optics Express*, *8*(6), 3017. doi:[10.1364/BOE.8.003017](https://doi.org/10.1364/BOE.8.003017)
- Wang, W., Li, C., Tollner, E. W., Rains, G. C., & Gitaitis, R. D. (2012). A liquid crystal tunable filter based shortwave infrared spectral imaging system: Calibration and characterization. *Computers and Electronics in Agriculture*, *80*, 135–144. doi:[10.1016/j.compag.2011.09.003](https://doi.org/10.1016/j.compag.2011.09.003)

- Wu, D. & Sun, D.-W. (2013). Application of visible and near infrared hyperspectral imaging for non-invasively measuring distribution of water-holding capacity in salmon flesh. *Talanta*, *116*, 266–276. doi:[10.1016/j.talanta.2013.05.030](https://doi.org/10.1016/j.talanta.2013.05.030)
- Xing, J., Bravo, C., Jancsó, P. T., Ramon, H., & De Baerdemaeker, J. (2005). Detecting Bruises on ‘Golden Delicious’ Apples using Hyperspectral Imaging with Multiple Wavebands. *Biosystems Engineering*, *90*(1), 27–36. doi:[10.1016/j.biosystemseng.2004.08.002](https://doi.org/10.1016/j.biosystemseng.2004.08.002)
- Yamal, J.-M., Zewdie, G. A., Cox, D. D., Neely Atkinson, E., Cantor, S. B., MacAulay, C., ... Follen, M. (2012). Accuracy of optical spectroscopy for the detection of cervical intraepithelial neoplasia without colposcopic tissue information; a step toward automation for low resource settings. *Journal of Biomedical Optics*, *17*(4). doi:[10.1117/1.JBO.17.4.047002](https://doi.org/10.1117/1.JBO.17.4.047002)
- Zeiler, M. D. & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. (pp. 818–833). doi:[10.1007/978-3-319-10590-1\\_53](https://doi.org/10.1007/978-3-319-10590-1_53)
- Zheludev, V., Pölönen, I., Neittaanmäki-Perttu, N., & Averbuch, A. (2015). Biomedical Signal Processing and Control Delineation of malignant skin tumors by hyperspectral imaging using diffusion maps dimensionality reduction. *Biomedical Signal Processing and Control*, *16*, 48–60. doi:[10.1016/j.bspc.2014.10.010](https://doi.org/10.1016/j.bspc.2014.10.010)
- Zherdeva, L. A., Bratchenko, I. A., Myakinin, O. O., Moryatov, A. A., Kozlov, S. V., & Zakharov, V. P. (2016). In vivo hyperspectral imaging and differentiation of skin cancer. *10024*, 100244G. doi:[10.1117/12.2246433](https://doi.org/10.1117/12.2246433)