Hyperspectral Imaging Reveals Spectral Differences and Can Distinguish Malignant Melanoma from Pigmented Basal Cell Carcinomas: A Pilot Study

Janne RÄSÄNEN[1,2], Mari SALMIVUORI[1–3], Ilkka PÖLÖNEN[4], Mari GRÖNROOS[1] and Noora NEITTAANMÄKI[5]

[1]Department of Dermatology, Päijät-Häme Social and Health Care Group, Lahti, [2]Department of Dermatology, Tampere University Hospital and Tampere University, Faculty of Medicine and Medical Technology, Tampere, [3]Department of Dermatology and Allergology, University of Helsinki and Helsinki University Hospital, Helsinki, [4]Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland and [5]Departments of Pathology and Dermatology, Institutes of Biomedicine and Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Pigmented basal cell carcinomas can be difficult to distinguish from melanocytic tumours. Hyperspectral imaging is a non-invasive imaging technique that measures the reflectance spectra of skin *in vivo*. The aim of this prospective pilot study was to use a convolutional neural network classifier in hyperspectral images for differential diagnosis between pigmented basal cell carcinomas and melanoma. A total of 26 pigmented lesions (10 pigmented basal cell carcinomas, 12 melanomas *in situ*, 4 invasive melanomas) were imaged with hyperspectral imaging and excised for histopathological diagnosis. For 2-class classifier (melanocytic tumours vs pigmented basal cell carcinomas) using the majority of the pixels to predict the class of the whole lesion, the results showed a sensitivity of 100% (95% confidence interval 81–100%), specificity of 90% (95% confidence interval 60–98%) and positive predictive value of 94% (95% confidence interval 73–99%). These results indicate that a convolutional neural network classifier can differentiate melanocytic tumours from pigmented basal cell carcinomas in hyperspectral images. Further studies are warranted in order to confirm these preliminary results, using larger samples and multiple tumour types, including all types of melanocytic lesions.

Key words: deep learning; neural network; basal cell carcinoma; malignant melanoma.

Accepted Jan 26, 2021; Epub ahead of print Feb 1, 2021

Acta Derm Venereol 2021; 101: adv00405.

Corr: Janne Räsänen. Department of Dermatology, Tampere University Hospital and Tampere University, Faculty of Medicine and Medical Technology, Riihipellonkatu 4 B 13, FIN-33530 Tampere, Finland. E-mail: janne.rasanen@sll.fimnet.fi

SIGNIFICANCE

This is the first study to utilize hyperspectral images and a deep-learning convolutional neural network to acquire an automated diagnosis for melanocytic tumours and pigmented basal cell carcinomas. The results of this pilot study will serve as a basis for future research. The results indicate that, with a larger sample and training dataset, the convolutional neural network could accurately classify malignant melanocytic tumours from pigmented basal cell carcinomas. This finding may be used as the basis for development of future techniques in melanoma diagnostics, which also requires a hyperspectral camera to be commercially available to clinicians.

Basal cell carcinoma (BCC) is the most common skin cancer and the most frequently occurring form of all cancers among whites in the Western countries, with an estimated lifetime risk of ≥30% (1). The most common subtypes, nodular and superficial BCC, may occasionally contain melanin, referred to as "pigmented BCC", which constitute ~7% of all BCCs (2). Dermoscopy uses a magnifying lens and polarized light, to assess different structures and colours of a lesion. In dermoscopy the pigment in BCC can be seen as blue-grey ovoid nests and globules, brown to blue-grey maple leaf-like structures, or small blue-grey dots (peppering) (2–4). Heavily pigmented BCCs may contain brown to black globules or dots, or blue-white veil-like structures, and are thus difficult to distinguish from melanocytic lesions (5). The pigmentation may cause confusion in diagnosing these lesions. Differential diagnoses for pigmented BCCs include melanoma *in situ* (MIS), invasive melanoma and benign pigmented lesions (2, 6). Usually, the clinical features suffice to establish the diagnosis of a pigmented BCC, but the diagnosis is not always clear-cut (**Fig. 1**). In uncertain cases diagnostic biopsy or excision is required.

Various non-invasive imaging modalities have been developed to aid in the diagnosis of BCC and to avoid unnecessary biopsies. These include reflectance confocal microscopy (RCM) and optical coherence tomography (OCT), as well as investigational methods still under development, such as Raman spectroscopy, high-resolution ultrasonography, and terahertz pulse imaging (4). Hyperspectral imaging (HSI) is a novel non-invasive imaging technique that measures the reflectance spectra of skin *in vivo* to identify different biological tissues (7, 8). We have previously shown that HSI can delineate the lesion borders of ill-defined BCCs and lentigo maligna (LM) more accurately than clinical examination, and can detect the invasive component in lentigo maligna melanoma (LMM) (9–11).

Convolutional neural networks (CNN) represent a form of artificial intelligence in decision-making and

doi: 10.2340/00015555-3755
Acta Derm Venereol 2021; 101: adv00405

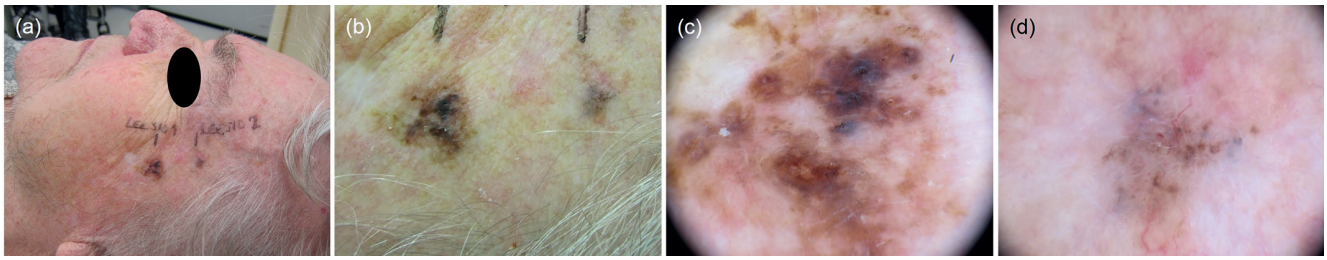

Fig. 1. Photographs of a patient included in the study with both lentigo maligna melanoma (LMM) and pigmented basal cell carcinoma (BCC). (a) Overview showing LMM (*left*) and BCC (*right*). (b) Close-up showing LMM (left) and BCC (*right*). (c) Dermoscopic image of the lesion on the left, histologically verified as LMM with Breslow depth 1.5 mm. (d) Dermoscopic image of the lesion on the right, histologically verified as a nodular pigmented BCC. The results of hyperspectral analyses are shown in Figs 2 and 3.

are a strongly emerging concept in dermatological diagnostics. CNNs can classify pigmented and non-pigmented skin tumours from clinical and dermoscopic images with an accuracy similar to that of dermatologists (12–14). In addition to 2-dimensional (2D) images, a CNN classifier has been utilized with hyperspectral images of *ex vivo* malignant tumours (15, 16).

The aim of this pilot study was to use a combination of 3-, 2- and 1-dimensional (3D, 2D and 1D) convolutional layers in a neural network classifier from hyperspectral images to distinguish between pigmented BCCs and malignant melanocytic tumours.

MATERIALS AND METHODS

Participant recruitment

The study protocol complied with the principles of the Declaration of Helsinki and was approved by the local ethics committee of Tampere University Hospital. All participants gave their written informed consent. Voluntary patients were recruited from among those referred for suspected skin malignancies to the Department of Dermatology at Päijät-Häme Central Hospital, Lahti, Finland, between May 2014 and November 2017. Inclusion criteria were clinically pigmented tumours, which were histologically confirmed as BCC, MIS, or invasive malignant melanoma (MM).

Hyperspectral image acquisition and histopathological examination

All lesions were clinically examined with a dermatoscope (Dermlite DL3, 3Gen, San Juan Capistrano, CA, USA) and photographed. For subsequent orientation purposes, a small mark was drawn on the healthy skin beside the lesion, using a black marker pen. Hyperspectral images were taken *in vivo* with an HSI camera prototype (VTT FPI VIS-VNIR Spectral Camera, VTT Technical Research Centre of Finland, Espoo, Finland). Spectral separation of the imager is based on Fabry-Pérot interferometer (FPI), which enables fast scanning in the spectral domain. The imager measures diffuse reflectance on wavebands from visible to near infrared light (500–850 nm) for every pixel of the image. The full width of half maximum (FWHM) of each waveband varies from 10 to 30 nm. The imager rapidly captures 76 wavebands within a few seconds in a 120 mm^2 field of view with a spatial resolution of 640 pixels/mm^2 (pixel = 125 × 125 µm). The imaging depth of hyperspectral imaging depends on the wavelength (17). In the wavelength range used (450–850 nm) the imaging depth varies in the range 0.5–5.0 mm. Because the whole wavelength range is used in the analysis the mean imaging depth is approximately 2 mm. The concentration of melanin affects the imaging depth, but, because the whole wavelength range is used, its effect on the mean imaging depth is minor (17). The outcome is a hyperspectral image: a 3D data cube containing a reflectance spectrum for each pixel of the 2D image. The HSI technique is described in more detail elsewhere (7, 10).

After imaging, the lesions were surgically excised and sent for routine histopathological examination. The tissue specimens were fixed in 4% formalin, embedded in paraffin, sectioned using the traditional vertical bread-loaf technique, and stained with haematoxylin and eosin (H&E).

Data pre-processing and 3-dimensional convolutional neural network

Of the 76 wavebands, 66 were used for the analyses. Thus, HSI produced a raw data cube (size 240 × 320 × 66 pixels), which was calibrated to the radiance using Saari et al.'s method (18). A white reference target was captured for each data cube, and this was used to convert imaged radiance to reflectance $R=I / I_0$, where I is the imaged region of interest and I_0 is the data cube from the white reference. To reduce the effect of vignetting and lighting irregularities, each imaged spectrum was subtracted by its average in the spectral domain (Fig. S1[1]).

A combination of 3D, 2D and 1D CNNs, operating separately and simultaneously in both spatial and spectral domains, was used to classify the lesions (19). For training the neural network, labelled data from captured hyperspectral images was used. For each image, a clinician (JER) manually annotated areas of histologically verified tumour, healthy skin and marker pen (Fig. S2[1]). From the annotated areas of each image approximately 1,000 datasets of 5×5-pixel neighbourhoods were selected for training purposes. The training set was sampled randomly from annotated points to contain a total of approximately 27,000 data-points from each class. For annotated points, data augmentation was used, such that each training cube was mirrored and flipped horizontally and vertically (20). These operations quadrupled the number of inputs in the training phase and resulted in a final training set of approximately 655,000 data-points.

Due to the limited number of lesions imaged ($n=26$), the hyperspectral images were divided into 2 parts in the vertical direction in the middle of the annotated lesion, so that one half served as a training image and the other half was used independently for classification. This ensured, firstly, that the training set did not contain data-points from the image currently being classified and, secondly, that the training set contained a sufficient variation of different lesion types (21–23).

For the implementation of CNN, the current study used Keras library, TensorFlow backend and Python 3.6. All calculations were performed on IBM PowerAI Platform, which includes 2 Nvidia

[1]https://www.medicaljournals.se/acta/content/abstract/10.2340/00015555-3755

Tesla V100-SXM2 16 GB graphics processing units (GPUs). A more detailed description of our implementation of CNN in skin cancer detection is available elsewhere (18).

Sample size

Due to the pilot study set-up, it was not possible to calculate the accurate sample size needed. The study aimed at a sample size of 10–15 pigmented BCC and 10–15 melanocytic tumours.

RESULTS

Patient demographics and lesion characteristics

A total of 32 lesions fulfilling the inclusion criteria were imaged in Päijät-Häme Central Hospital, Lahti. Hyperspectral images of 6 lesions were excluded from the study because the HSI settings were not optimal at the time of imaging, leading to imaging artefacts (under-exposure or FPI switch inadvertently off). Hence, a total of 26 lesions from 24 patients were included in the study. Patient demographics and lesion characteristics are shown in **Table I**.

Fifteen patients were men and 9 women, mean age 73.1 (range 53–91) years. Thirteen had Fitzpatrick's skin phototype II, 8 phototype III, and, in 3 patients, the phototype was not reported. Fourteen patients had a previous history of skin cancer, as follows: AK ($n=8$), Bowen's disease ($n=4$), squamous cell carcinoma ($n=3$), BCC ($n=9$), and MIS or MM ($n=4$). Five patients had a history of other malignancy (2 prostate cancers, one tongue cancer, one lip cancer, and one hepatocellular and pancreatic carcinoma). Two of the patients were immunosuppressed: one patient was an organ-transplant receiver and used sirolimus for antirejection, and one patient used azathioprine for Crohn's disease.

The imaged lesions were located on the face ($n=8$), neck ($n=3$), torso ($n=10$), upper extremity ($n=4$) and lower extremity ($n=1$). The mean ± standard deviation (SD) diameter of the lesions was 9.8 ± 2.9 mm (range 5.3–19 mm) and the mean ± SD area was 71.2 ± 46.9 mm^2 (range 11.0–247 mm^2).

Histopathological diagnoses

The histopathological diagnoses of the imaged lesions were as follows: 10 BCC (6 nodular, 2 superficial, 2 both superficial and nodular), 12 MIS (all lentigo maligna subtype), and 4 MM (3 lentigo maligna subtypes and 1 superficial spreading subtype, Breslow depths in the range 0.3–1.4 mm).

Hyperspectral analysis

The hyperspectral analyses and lesion classification with CNN were conducted by a trained mathematician (IP). Due to limited sample size, all melanocytic tumours (MIS and MM) were first combined in single group for lesion classification. Each hyperspectral image was classified

Table I. Patient demographics and baseline characteristics, histopathological diagnoses and the result of CNN classification for the lesions included in the study

Lesion number	Age, years	Sex (F/M)	Lesion location	Size, mm	Histopathological diagnosis (=true label)	Predicted label by CNN (2 classes)	Predicted label by CNN (3 classes)
1	56	F	Eyelid	2 × 7	nBCC	Pigmented BCC	Pigmented BCC
2	82	F	Temple	11 × 12	LMM (Breslow 1.5)	Melanocytic	MM
3	[a]	[a]	Temple	5 × 6	nBCC	Pigmented BCC	Pigmented BCC
4	68	M	Upper back	8 × 8	LM	Melanocytic	MIS
5	69	M	Groin	8 × 10	n/sBCC	Melanocytic[c]	Pigmented BCC
6	69	M	Shoulder	15 × 21	MM (Breslow 1.4)	Melanocytic	MM
7	57	F	Breastbone	7 × 8	sBCC	Pigmented BCC	Pigmented BCC
8	73	M	Back	10 × 13	sBCC	Pigmented BCC	Pigmented BCC
9	81	F	Neck	6 × 13	n/sBCC	Pigmented BCC	Pigmented BCC
10	72	F	Nose	10 × 10	nBCC	Pigmented BCC	Pigmented BCC
11	53	F	Eyelid	4 × 12	nBCC	Pigmented BCC	Pigmented BCC
12	86	F	Upper back	7 × 12	nBCC	Pigmented BCC	Pigmented BCC
13	66	M	Shoulder	9 × 9	LM	Melanocytic	MIS
14	75	M	Shoulder	5 × 10	nBCC	Pigmented BCC	Pigmented BCC
15	91	M	Upper back	7 × 10	LM	Melanocytic	MIS
16	84	M	Chest	4 × 8	LMM (Breslow 0.3)	Melanocytic	MM
17	83	M	Chest	8 × 12	LMM (Breslow 0.4)	Melanocytic	MM
18	69	M	Forehead	8 × 10	LM	Melanocytic	MIS
19	80	M	Neck	8 × 12	LM	Melanocytic	MIS
20	62	M	Forearm	10 × 17	LM	Melanocytic	MIS
21	75	M	Chest	11 × 13	LM	Melanocytic	MIS
22	77	F	Forearm	9 × 12	LM	Melanocytic	MIS
23	77	M	Cheek	10 × 14	LM	Melanocytic	MIS
24	71	F	Lower back	6 × 8	LM	Melanocytic	MIS
25	79	M	Temple	8 × 8	LM	Melanocytic	MIS
26	[b]	[b]	Neck	6 × 8	LM	Melanocytic	MIS

The predicted label for the whole lesion was determined by the predicted class of the majority of pixels in the lesion area. The result of classification is presented for the study sample divided into 2 classes (pigmented basal cell carcinoma (BCC) and melanocytic tumours) and 3 classes (pigmented BCC, melanoma *in situ* (MIS) and malignant melanoma (MM)).
[a]Lesions 2 and 3 belonged to the same patient. [b]Lesions 25 and 26 belonged to the same patient. [c]Convolutional neural networks (CNN) classification erroneous.
nBCC: nodular basal cell carcinoma; sBCC: superficial basal cell carcinoma; LMM: lentigo maligna melanoma.

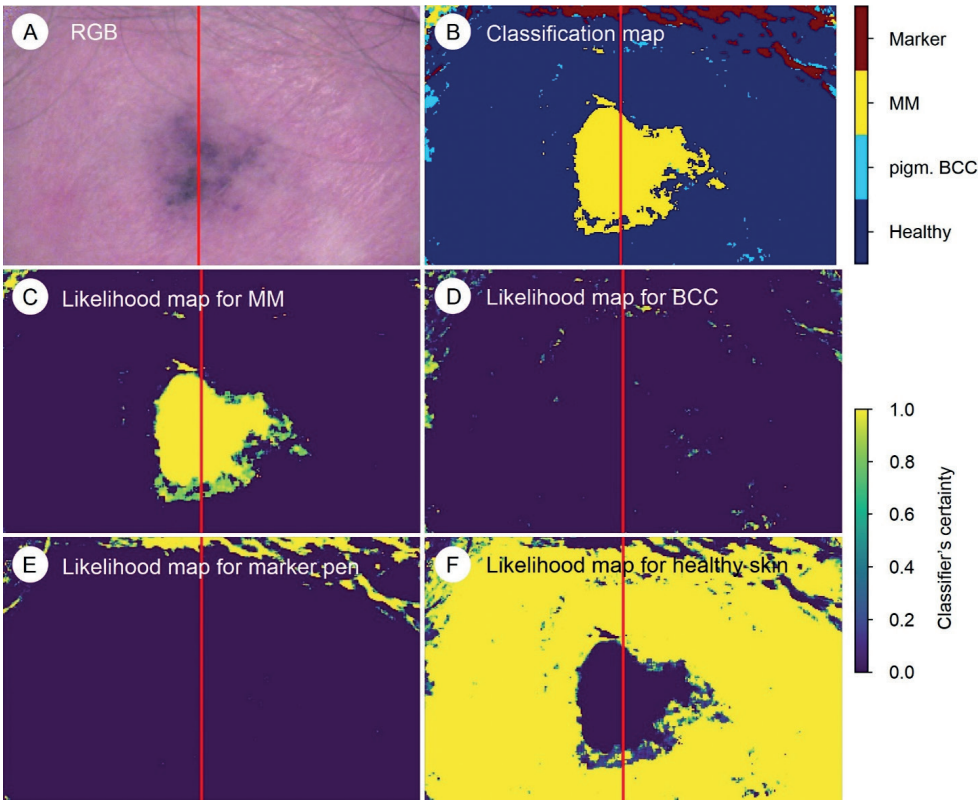

Fig. 2. Classification and likelihood maps of a histologically confirmed invasive melanoma (lentigo maligna subtype, Breslow depth 1.5 mm). The lesion was correctly classified as malignant melanoma (MM) by hyperspectral imaging using the majority of the pixels method. (A) Red, green and blue (RGB) image taken by hyperspectral camera. (B) Classification map showing most of the pixels representing invasive malignant melanoma (MM). *Upper colour bar* depicts the 4 classes in image B. (C) Likelihood map for the localization of MM. (D) Likelihood map for the localization of pigmented basal cell carcinoma (BCC). Only some scattered artefactual pixels are shown outside the lesion area. (E) Likelihood map for the localization of marker pen. (F) Likelihood map for the localization of healthy skin. *Lower colour bar* represents the certainty of the classifier in each of the likelihood maps. *Red line* shows the division of the hyperspectral image into halves, one of which was used for the training of the classifier and the other for the classifying task.

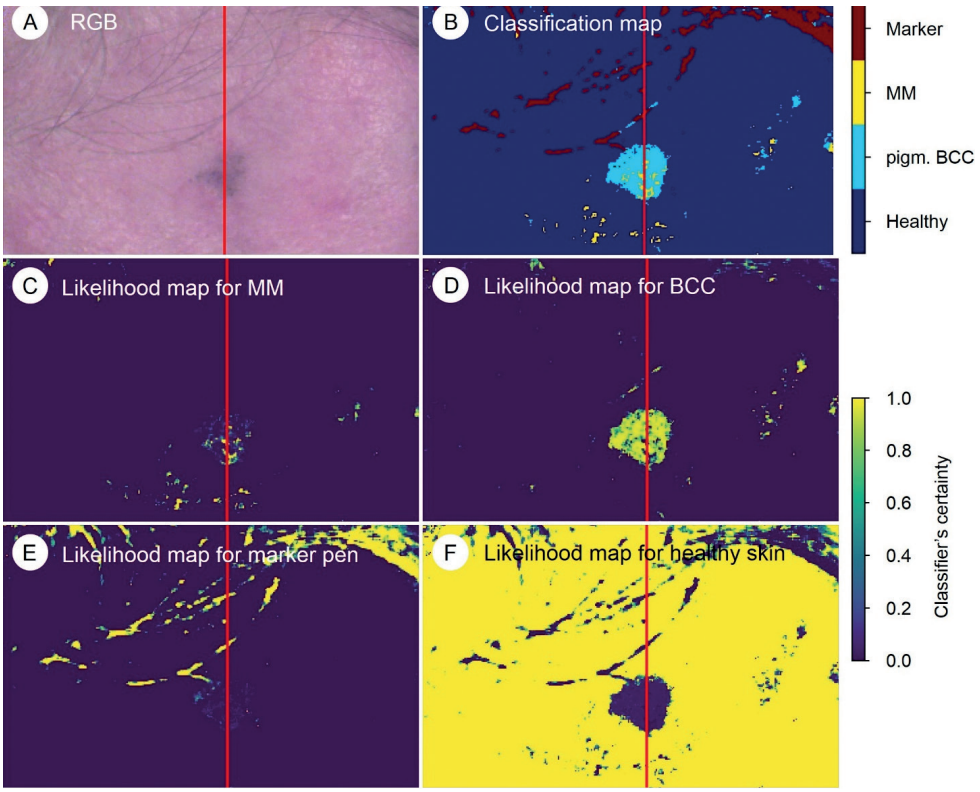

Fig. 3. Classification and likelihood maps of a histologically confirmed pigmented basal cell carcinoma (BCC). The lesion was correctly classified as BCC by hyperspectral imaging using the majority of pixels method. (A) Red, green and blue (RGB) image taken by hyperspectral camera. (B) Classification map showing most of the pixels classified as pigmented BCC. *Upper colour bar* depicts the 4 classes in image B. (C) Likelihood map for the localization of invasive melanoma (malignant melanoma; MM). (D) Likelihood map for the localization of pigmented basal cell carcinoma (BCC). (E) Likelihood map for the localization of marker pen. (F) Likelihood map for the localization of healthy skin. *Lower colour bar* represents the certainty of the classifier in each of the likelihood maps. *Red line* shows the division of the hyperspectral image into halves, one of which was used for the training of the classifier and the other for the classifying task.

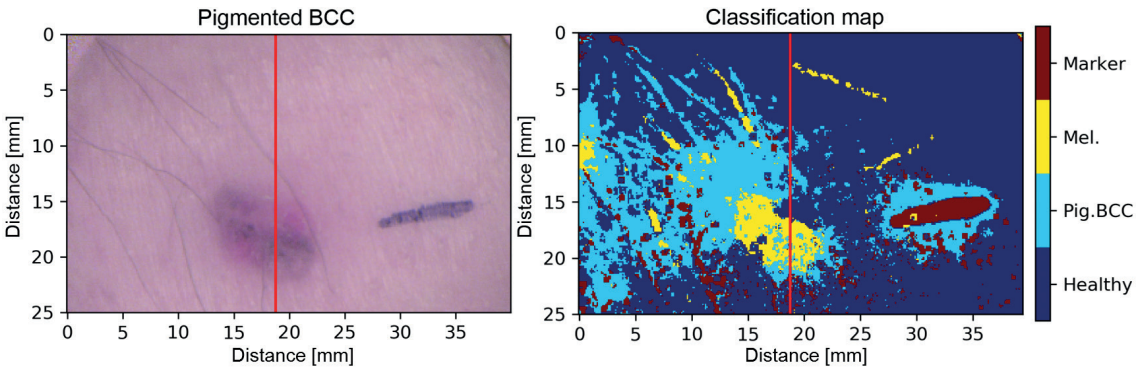

Fig. 4. Classification map of a histologically confirmed pigmented basal cell carcinoma (BCC). In this case the classifier falsely showed multiple pixels representing melanocytes in the lesional area. This may partly been explained by the fact that, histologically, there are some entrapped melanocytes in the area of pigmented BCC. Also, pigment macrophages may have been mistakenly identified as melanocytes. In this specific case the image quality was low, which also lead to misclassification of the pixels in the surrounding healthy skin.

pixel-wise by 3D CNN, which produced a classification map comprehending 4 classes ("pigmented basal cell carcinoma", "melanocytic tumour", "healthy skin" and "marker pen") (**Figs 2–4**). The last layer of the classifier also created likelihood maps indicating how certain the CNN is about the classification. To obtain the classification for each pixel over each image, the maximum likelihood (the highest value of each pixel in likelihood maps) was used. Two different approaches were used for the analyses: (*i*) a majority of pixels classification; and (*ii*) a pixel-wise classification. The histopathological diagnoses of the excision specimens represented the true classification (true label) of the lesions.

In the majority of pixels classification, the classification maps were used to acquire an automated diagnosis for the imaged lesions, i.e. the whole lesion was predicted to belong to the same class as the majority of the pixels in the lesion area. In addition, the accuracy of CNN classification per individual pixel was computed in comparison with the manual annotations in the images (4 pixel classes, as discussed above). By using the pixel-wise classification the number of classified areas increased from 26 lesions to 164,144 classified pixels. The confusion matrices for both lesion and pixel-wise classifications are shown in **Fig. 5**.

For the majority of pixel classification, sensitivity was 100% (16/16) (95% CI 81–100%), specificity 90% (9/10) (95% CI 60–98%) and PPV 94% (16/17) (95% CI 73–99%) in differentiating melanocytic tumours from pigmented BCCs. Only one true pigmented BCC was falsely predicted as a melanocytic tumour, and all melanocytic tumours were correctly predicted according to their true label. In pixel-wise classification the sensitivity for differentiating the class "melanocytic tumours" from the class "pigmented BCC" was 99.98% (95% CI 99.95–99.99%), specificity 93.2% (95% CI 92.6–93.7%) and PPV 97.8% (95% CI 97.6–98.0%).

This study also tested the accuracy of the CNN classifier, by dividing the sample into 3 groups: pigmented BCC ($n=10$), MIS ($n=12$) and MM ($n=4$), and then used the classification maps to acquire an automated diagnosis for every lesion in a manner similar to that described above. Classifier accuracy by individual pixels was also computed (5 pixel classes). In the lesion classification using the majority of pixels, calculated sensitivity was 100% (4/4) (95% CI 51–100), specificity 100% (22/22)

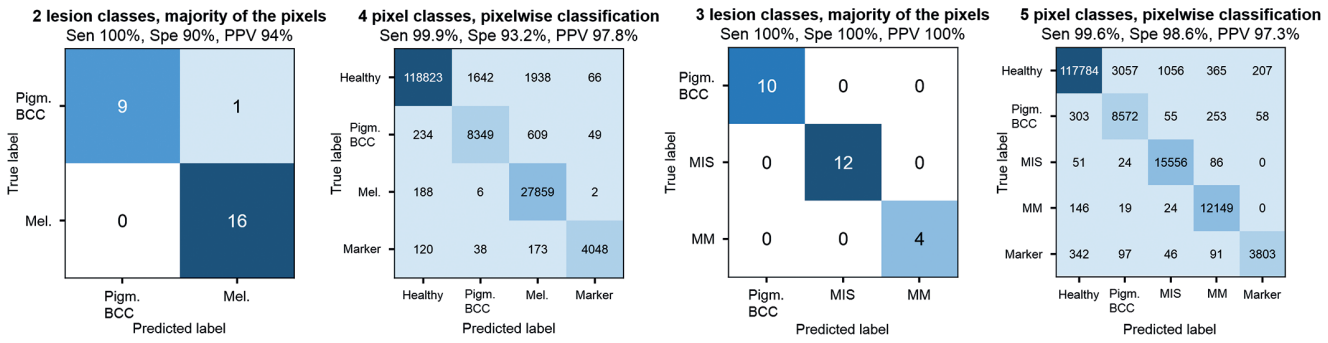

Fig. 5. Confusion matrices for the convolutional neural network (CNN) classifier. *Left*: lesion-based classification matrices for 2- and 3-class classification using the majority of the pixels to predict the label of whole lesion. *Right*: classification matrices per individual pixel compared with manual annotation in hyperspectral images (true label). In the case of 2 lesion classes the sample was divided into pigmented basal cell carcinomas (BCC) and malignant melanocytic tumours (Mel.). In the case of 3 lesion classes the sample was divided into pigmented BCCs, melanomas *in situ* (MIS) and malignant melanomas (MM). The statistical metrics were calculated for the class "Mel." in case of 2 lesion classes and 4 pixel classes and for the class "MM" in case of 3 lesion classes and 5 pixel classes. Sen: sensitivity, Spe: specificity; PPV: positive predictive value.

(95% CI 85–100) and PPV 100% (4/4) (95% CI 51–100) in distinguishing MM from other lesions (MIS and pigmented BCC). In pixel-wise classification between tumour classes only, the statistical characteristics for class "MM" were as follows: sensitivity 99.6% (95% CI 99.5–99.7), specificity 98.6% (95% CI 98.5–98.8) and PPV 97.3% (95% CI 97.0–97.6). The confusion matrices for 3-class lesion classification and 5-class pixel-wise classification are shown in Fig. 4. The result of the CNN classification for all lesions in comparison with true histopathological diagnoses is shown in Table I. The difference in the spectra between different lesion types and healthy skin is shown in Fig. S3[1].

DISCUSSION

To the best of our knowledge, this pilot study is the first to utilize *in vivo* hyperspectral imaging (HSI) and a deep-learning CNN as a tumour classifier for skin tumours. The CNN lesion classifier using 2 different approaches (majority of pixels classifier and single pixel classifier) reached high sensitivity and specificity in distinguishing melanocytic tumours from pigmented BCCs.

HSI is a novel non-invasive optical imaging method that can be used to image tissues *in vivo* and *ex vivo* tissues rapidly in various medical applications. It produces a hyperspectral data cube in which different tissue types can be identified and visualized based on their specific reflectance spectra (8). Map-like images generated by HSI can be used to delineate malignant skin lesions preoperatively and to visualize invasive parts in large lesions to target biopsy sites (9–11).

Furthermore, HSI could be useful for dermatologists as a pre-surgical diagnostic aid for planning the surgical excision margins. HSI could be a helpful diagnostic tool even for inexperienced physicians, since it is user-independent and gives an automated most likely diagnosis.

No earlier reports on HSI combined with CNN for tumour classification were found in the literature. Among the available non-invasive diagnostic modalities, one reminiscent of the current method is MelaFind®, a multispectral imaging (MSI) system that acquires 10 wavebands in the range 430–950 nm, and uses automated algorithms based on linear classifiers for pattern recognition and differential diagnosis of pigmented lesions (24). MelaFind has demonstrated high sensitivity, of 82–98%, but only moderate specificity, of 8–52%, in recognizing melanomas (25–27). A limitation of MelaFind is that it may misclassify non-melanoma skin cancers, such as BCCs, because the technology is designed to assess overall structural disorganization of lesions, rather than atypical cellular features (28). Another non-invasive MSI technique is SIAscope (spectrophotometric intracutaneous analysis), which captures 8 narrowband spectral images in the range 400–1,000 nm to visualize the quantity and microscopic architecture of chromophores (melanin, blood and collagen) in the lesion area (29). However, SIAscopy does not improve diagnostic accuracy over dermoscopy in the diagnosis of pigmented BCC (30). Tomatis et al. (31) developed an artificial neural network (ANN) classifier for 1,391 pigmented lesions imaged with MSI, which was able to distinguish melanoma from other lesions with a sensitivity of 80% and specificity of 76%. In another study, Carrara et al. (32) developed an ANN classifier for MSI data to differentiate between reassuring lesions and those requiring excision, using a dataset of 1,966 excised and 1,940 non-excised and clinically non-suspicious lesions, resulting in sensitivity of 88% and specificity of 80%. For comparison, in meta-analyses, dermoscopy has achieved sensitivity and specificity of 88% and 86%, respectively, for the diagnosis of melanoma (33), 82.4% and 83.5% for the diagnosis of LM and LMM (34), and 81.9% and 81.8% for the diagnosis of superficial BCC (35). However, the diagnostic accuracy of dermoscopy is user-dependent and depends on the experience of the clinician. Recently, deep learning CNN have been utilized to facilitate and automate diagnostic classification in clinical and dermoscopic images, reaching diagnostic accuracies similar to those achieved by expert dermatologists, with the area under the curve (AUC) of the receiver operating curve (ROC) ranging from 0.74 to 0.96 (12–14).

The results of the current study indicate that HSI might be capable of spectral differentiation at the cellular level and of recognizing the tumour cell of origin (melanocyte vs keratinocyte). The CNN classification maps in pixel-wise resolution can be thought to represent aggregates of cells in the microscopic structure of the skin. The hyperspectral image analysis appears, in most cases, to recognize the different cell types in pigmented BCC and melanocytic tumours regardless of the presence of melanin pigment. Both tumour types have increased amounts of melanin in the dermis, often histologically visualized in macrophages. Thus, the results of the current study show that HSI does not only use pigment as a chromophore, but also recognizes other characteristics beyond the pigment. Further verification of this is required in studies at the cellular level. There were some cases with pixels misclassified as melanoma (Fig. 4) within the pigmented BCCs, which could be due to the fact that, histologically, there are some entrapped melanocytes in the areas of pigmented BCCs (36).

The current study used a combination of 3D, 2D and 1D CNN for automated pixel-wise tumour classification in hyperspectral images of the skin. 3D CNN was used because the standard 2D CNN used with clinical and dermoscopic photographs may not suffice to utilize spectral data, which is 3-dimensional in nature. This method produced classification maps of imaged lesions that predict the lesion's diagnosis pixel-wise (Figs 2 and 3). The classification map not only offers identification (the most likely diagnosis), but also delineation and map-like repre-

sentation of the lesion. Figs 2 and 3 show that there are a few misclassified pixels in the tumour area, but that, by far, most pixels are classified correctly, predicting the true histopathological diagnosis. In addition, at the periphery of the image in the area of healthy skin there are some rare misclassified pixels. Furthermore, the black hair has been classified as "marker pen", since no separate class was created for hair. The few misclassified pixels could be explained by imaging errors, such as local increased noise or other perturbation in the hyperspectral image, or by inaccuracy in the CNN model. When classifying whole lesions using the majority of pixels method, the few misclassified pixels were averaged, leading to very high sensitivity and specificity for differential diagnosis of melanocytic tumours from BCCs. In the case of 3 tumour classes, the accuracy of the CNN classifier was also very high with all tumours classified correctly. The confusion matrices of individual pixel classification (Fig. 4) demonstrate the robustness of the CNN's pixel-wise classification, but it should be noted that the pixel-wise accuracy is presented only in comparison with manually annotated pixel classes in images. The pixel-wise classification is fairly accurate in distinguishing tumour classes from each other, with high sensitivity and specificity for both 2-class and 3-class lesion classification. However, if all (4 or 5) pixel classes are observed, it can be seen that there is a portion of "healthy" annotated pixels that are assigned to tumour classes "pigmented BCC", "melanocytic tumours", "MIS" and "MM", which are false positives in the whole imaged region and represent the inherent error in CNN classification. In this study setting, the majority of pixels method can be assumed to predict the diagnosis reliably, because we know *a priori* that no single BCC cells are situated inside a melanocytic tumour and vice versa. This is not necessarily the case when, for example, multiple melanocytic tumours are included in the classification task. In the development of the CNN classification method, this study strove for pixel-wise classification, because there are clinical situations in which a single skin lesion may include multiple tumour types (e.g. an invasive part of surrounding carcinoma *in situ*, or invasive melanoma arising in melanoma *in situ* or in pre-existing naevus). The CNN method used in the present study is novel and can be developed further. In addition, the classification of CNN improves with larger training datasets, which will require the collection of larger samples of different tumour types and from different locations on the body. With a larger training set the pixel-wise classification of CNN could be further improved to be more robust, thereby reducing the number of misclassified pixels.

Study limitations

This study has some limitations. Due to the pilot nature of the study, the sample size was small and halves of the images were used for the CNN training phase. Ideally, the CNN training should be conducted with a dataset separate from that used for the classification task. However, the small sample size was compensated by the pixel-wise analysis, in which the pixels were analysed separately and could be counted as separate lesions. This increases the reliability of the findings. It should be noted that benign tumours clinically mimicking pigmented BCCs (benign lentigines, seborrhoeic keratosis, benign naevi) were not included, and further studies on this topic are warranted. The CNN classification method used is novel and requires further development work. An HSI camera prototype was used in this study, which could be developed further to be more robust in order to avoid imaging noise and artefacts. There is currently no commercially available HSI device for skin imaging.

Conclusion

This is the first study to utilize hyperspectral images and deep-learning CNN to achieve automated diagnosis for melanocytic tumours and pigmented BCCs. The results will therefore serve as a basis for future studies.

The results of this study suggest that, with a larger sample and training dataset, the CNN could accurately and pixel-wise distinguish between malignant melanocytic tumours and pigmented BCCs. To confirm these preliminary results, future studies with larger samples and multiple tumour types, including all types of melanocytic lesions, are warranted.

ACKNOWLEDGEMENTS

The patients in this paper provided written informed consent for publication of their case details.

The authors would like to thank Ulla Oesch-Lääveri, nurse at the Päijät-Häme Central Hospital, for her dedication to this study, and Heikki Saari from VTT Finnish Technical Centre of Finland and statistician Martin Gillsted for their help in the study.

This study was funded by the Cancer Foundation of Finland, by Tampere University Hospital and by the State Research Funding.

IP, MG and NN have a patent (US10478071B2) licensed. MG has received a consultation fee from Revenio Group. The other authors have no conflicts of interest to declare.

REFERENCES

1. Cameron MC, Lee E, Hibler BP, Barker CA, Mori S, Cordova M, et al. Basal cell carcinoma: Epidemiology; pathophysiology; clinical and histological subtypes; and disease associations. J Am Acad Dermatol 2019; 80: 303–317.
2. Menzies SW. Dermoscopy of pigmented basal cell carcinoma. Clin Dermatol 2002; 20: 268–269.
3. Wozniak-Rito A, Zalaudek I, Rudnicka L. Dermoscopy of basal cell carcinoma. Clin Exp Dermatol 2018; 43: 241–247.
4. Cameron MC, Lee E, Hibler BP, Giordano CN, Barker CA, Mori S, et al. Basal cell carcinoma: contemporary approaches to diagnosis, treatment, and prevention. J Am Acad Dermatol 2019; 80: 321–339.
5. Altamura D, Menzies SW, Argenziano G, Zalaudek I, Soyer HP, Sera F, et al. Dermatoscopy of basal cell carcinoma:

morphologic variability of global and local features and accuracy of diagnosis. J Am Acad Dermatol 2010; 62: 67–75.
6. Bosbous MW, Dzwierzynski WW, Neuburg M. Lentigo maligna: diagnosis and treatment. Clin Plast Surg 2010; 37: 35–46.
7. Neittaanmaki-Perttu N, Gronroos M, Tani T, Polonen I, Ranki A, Saksela O, et al. Detecting field cancerization using a hyperspectral imaging system. Lasers Surg Med 2013; 45: 410–417.
8. Lu G, Fei B. Medical hyperspectral imaging: a review. J Biomed Opt 2014; 19: 10901.
9. Salmivuori M, Neittaanmaki N, Polonen I, Jeskanen L, Snellman E, Gronroos M. Hyperspectral imaging system in the delineation of Ill-defined basal cell carcinomas: a pilot study. J Eur Acad Dermatol Venereol 2019; 33: 71–78.
10. Neittaanmaki-Perttu N, Gronroos M, Jeskanen L, Polonen I, Ranki A, Saksela O, et al. Delineating margins of lentigo maligna using a hyperspectral imaging system. Acta Derm Venereol 2015; 95: 549–552.
11. Neittaanmaki N, Salmivuori M, Polonen I, Jeskanen L, Ranki A, Saksela O, et al. Hyperspectral imaging in detecting dermal invasion in lentigo maligna melanoma. Br J Dermatol 2017; 177: 1742–1744.
12. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017; 542: 115–118.
13. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 2018; 29: 1836–1842.
14. Tschandl P, Rosendahl C, Akay BN, Argenziano G, Blum A, Braun RP, et al. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. JAMA Dermatol 2019; 155: 58–65.
15. Halicek M, Little JV, Wang X, Patel M, Griffith CC, Chen AY, et al. Tumor margin classification of head and neck cancer using hyperspectral imaging and convolutional neural networks. Proc SPIE Int Soc Opt Eng 2018; 10576: 1057605.
16. Halicek M, Little JV, Wang X, Patel M, Griffith CC, El-Deiry MW, et al. Optical biopsy of head and neck cancer using hyperspectral imaging and convolutional neural networks. Proc SPIE Int Soc Opt Eng 2018; 10469: 104690X.
17. Barun, VV, Ivanov, AP, Volotovskaya, AV, Ulashchik, VS. Absorption spectra and light penetration depth of normal and pathologically altered human skin. J Appl Spectrosc 2007; 74: 430–439.
18. Saari H, Polonen I, Salo H, Honkavaara E, Hakala T, Holmlund C, et al. Miniaturized hyperspectral imager calibration and UAV flight campaigns. Sensors, Systems, and Next-Generation Satellites Xvii 2013; 8889: 88891O.
19. Polonen I, Rahkonen S, Annala L, Neittaanmaki N. Convolutional neural networks in skin cancer detection using spatial and spectral domain. Photon Dermatol Plast Surg 2019; 10851: 108510B.
20. Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning. Cambridge: MIT Press; 2016: p. 2.
21. Audebert N, Le Saux B, Lefèvre S. Deep learning for classification of hyperspectral data: a comparative review. IEEE Geosci Remote Sens Mag 2009; 7: 159–173.
22. Lunga D, Prasad S, Crawford MM, Ersoy O. Manifold-learning-based feature extraction for classification of hyperspectral data: a review of advances in manifold learning. IEEE Signal Process Mag 2013; 31: 55–66.
23. Ghamisi P, Plaza J, Chen Y, Li J, Plaza, AJ. Advanced spectral classifiers for hyperspectral images: a review. IEEE Geosci Remote Sens Mag 2017; 5: 8–32.
24. Elbaum M, Kopf AW, Rabinovitz HS, Langley RG, Kamino H, Mihm MC, Jr, et al. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: a feasibility study. J Am Acad Dermatol 2001; 44: 207–218.
25. Wells R, Gutkowicz-Krusin D, Veledar E, Toledano A, Chen SC. Comparison of diagnostic and management sensitivity to melanoma between dermatologists and MelaFind: a pilot study. Arch Dermatol 2012; 148: 1083–1084.
26. Monheit G, Cognetta AB, Ferris L, Rabinovitz H, Gross K, Martini M, et al. The performance of MelaFind: a prospective multicenter study. Arch Dermatol 2011; 147: 188–194.
27. MacLellan AN, Price EL, Publicover-Brouwer P, Matheson K, Ly TY, Pasternak S, et al. The use of non-invasive imaging techniques in the diagnosis of melanoma: a prospective diagnostic accuracy study. J Am Acad Dermatol 2020 Apr 11 (Online ahead of print).
28. March J, Hand M, Grossman D. Practical application of new technologies for melanoma diagnosis: Part I. Noninvasive approaches. J Am Acad Dermatol 2015; 72: 929–941, quiz 941–942.
29. Moncrieff M, Cotton S, Claridge E, Hall P. Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. Br J Dermatol 2002; 146: 448–457.
30. Terstappen K, Larkö O, Wennberg AM. Pigmented basal cell carcinoma – comparing the diagnostic methods of SIAscopy and dermoscopy. Acta Derm Venereol 2007; 87: 238–242.
31. Tomatis S, Carrara M, Bono A, Bartoli C, Lualdi M, Tragni G, et al. Automated melanoma detection with a novel multispectral imaging system: results of a prospective study. Phys Med Biol 2005; 50: 1675–1687.
32. Carrara M, Bono A, Bartoli C, Colombo A, Lualdi M, Moglia D, et al. Multispectral imaging and artificial neural network: mimicking the management decision of the clinician facing pigmented skin lesions. Phys Med Biol 2007; 52: 2599–2613.
33. Rajpara SM, Botello AP, Townend J, Ormerod AD. Systematic review of dermoscopy and digital dermoscopy/artificial intelligence for the diagnosis of melanoma. Br J Dermatol 2009; 161: 591–604.
34. Carapeba MOL, Alves Pineze M, Nai GA. Is dermoscopy a good tool for the diagnosis of lentigo maligna and lentigo maligna melanoma? A meta-analysis. Clin Cosmet Investig Dermatol 2019; 12: 403–414.
35. Lallas A, Tzellos T, Kyrgidis A, Apalla Z, Zalaudek I, Karatolias A, et al. Accuracy of dermoscopic criteria for discriminating superficial from other subtypes of basal cell carcinoma. J Am Acad Dermatol 2014; 70: 303–311.
36. Brankov N, Prodanovic EM, Hurley MY. Pigmented basal cell carcinoma: increased melanin or increased melanocytes? J Cutan Pathol 2016; 43: 1139–1142.