Margin Assessment of Extramammary Paget's Disease Based On Harmonic Generation Microscopy With Deep Neural Networks

Chia-I Chen, Yi-Hua Liao, and Chi-Kuang Sun ⓘ, *Fellow, IEEE*

Abstract—Surgical borders of extramammary Paget's disease (EMPD) are difficult to be identified via its clinical appearance. In this study, we propose a new diagnostic technique which combines nonlinear harmonic generation microscopy (HGM) with the deep learning method to instantaneously determine whether the imaged 3D stack is malignant EMPD or surrounding normal skin digitally. To demonstrate our proposal, in this study different locations of fresh EMPD surgical samples were 3D imaged starting from the surface up to a depth of 180 μm using stain-free HGM. With the followed histopathological examination of the same sample, we mapped the gold-standard results to 3D HGM image stacks with labels for the training of the deep learning model. With only 2095 3D image stacks as training and validation data, the results of EMPD and normal skin tissue classification achieve 98.06% sensitivity, 93.18% specificity and 95.81% accuracy. This study supports our proposed 3D convolutional-neural-network-based technique with a high potential to assist physicians to quickly map the EMPD margins by providing noninvasive instant information regarding the imaged sub-millimeter site as malignant or surrounding normal with a high accuracy.

Index Terms—Extramammary Paget's disease (EMPD), harmonic generation microscopy (HGM), deep learning, 3D convolutional neural network (3D-CNN), noninvasive, nonlinear optics.

I. INTRODUCTION

EXTRAMAMMARY Paget's disease (EMPD) is a cutaneous malignancy. It appears commonly in apocrine gland-bearing areas, particularly the anogenital regions including the vulva, perianal, perineum, and scrotum. EMPD is difficult to

visually assess due to its nonspecific clinical appearance, commonly misdiagnosed as an inflammatory or infective skin condition. Compared with nonsurgical treatments, complete surgical removal of the lesion is currently the best treatment choice for EMPD so far. However, for a lot of cases, the EMPD lesions have an ill-defined tumor border or an extended tumor spread beyond the clinically visible tumor border, which makes it difficult for physicians to define the accurate resection margin so as to remove the lesion completely in one surgical operation [1], [2]. To achieve a more accurate tumor border and avoid multiple surgical operations, several pre- or peri-operative margin assessment techniques are adopted including linear strip biopsy and mapping biopsy. The linear strip biopsy is a two-phase surgical technique which is highly invasive with skin strip resection [3]. The mapping biopsy technique utilizes punch biopsy as the tumor examination method, which is less invasive than linear strip biopsy. Statistically, it takes an average of 22 mapping biopsies at different sites to obtain an accurate EMPD border before surgery [4], [5]. These preoperational techniques assist physicians to define tumor margin more accurately; however, they consume a considerable amount of time along with pathological examination. Moreover, while these methods lower the recurrence rate, such invasive and repetitive approaches make patients suffer pain for more than once.

Recently, reflectance confocal microscopy (RCM) has been employed as a noninvasive diagnosis technique for the margin assessment of EMPD. It is a noninvasive and high-resolution 3D imaging tool for various cutaneous malignancy [6]–[8]. RCM provides quasi-histological resolution of the cellular and subcellular structures starting from the stratum corneum down to the dermoepidermal junction (DEJ) and upper papillary dermis [9]. As a diagnostic adjunct for EMPD, one study showed its sensitivity of EMPD was only 75% with a high false negative ratio [8]. This casts doubt on RCM's detection ability of EMPD margins [8]. A noninvasive diagnosis tool which expedites margin assessment of EMPD with high sensitivity and specificity is still in great demand.

In this study, we propose a new diagnostic technique which combines high-3D-resolution harmonic generation microscopy (HGM) with the 3D convolutional neural network (3D-CNN) model to instantaneously differentiate the malignant EMPD from its surrounding normal tissues digitally. Fresh EMPD surgical specimens were 3D imaged by stain-free HGM starting from the surface to the end of upper dermis. Each 3D HGM

image stack was composed of up to 100 series of 2D images, 512 × 512 pixels under 235 × 235 μm^2 field of view. Next, the surgical specimen was histopathologically examined for the EMPD cancer margins, which were treated as the result of the official clinical diagnosis for the deep learning system. We then mapped the location of the gold standard results to different HGM 3D image stacks with labels and trained the 3D-CNN model. After training, modification, and tuning the 3D-CNN model with only 2095 HGM image stacks (training and validation dataset), the result of EMPD and normal skin tissues classification achieves 98.06% sensitivity, 93.18% specificity, and 95.81% accuracy. Without subjective selection of a specific 2D image at a certain depth, our proposed 3D diagnostic technique shows high potential to assist physicians to quickly map the EMPD margins by providing noninvasive instant information regarding the imaged sub-millimeter site as malignant or surrounding normal with a high accuracy. This technique provides the basis for future clinical application in EMPD disease diagnosis, preoperational margin assessment, and intraoperational border determination in Mohs micrographic surgery.

II. MATERIALS AND METHODS

A. Harmonic Generation Microscopy

Harmonic generation microscopy, combining second harmonic generation (SHG) and third harmonic generation (THG) modalities, is a non-invasive imaging technique with an improved spatial resolution, contrast, and penetration in comparison with RCM [9]–[12]. With SHG and THG, HGM can be used to complement the single-modality confocal reflectance in order to obtain a sharp contrast between the dermis and epidermis [9]. With the virtual-transition-based nonlinear optical mechanism, HGM achieves dual contrasts without inducing nonlinear damage to the bio-samples. By utilizing longer wavelengths in the IR spectrum (1200-1300 nm), tissue linear damage can be further reduced and deeper penetration of light into the skin can be achieved. Besides, HGM has inherent confocality, and therefore, the confocal pinhole is not necessary to obtain optical sectioning. Lastly, HGM has high signal-to-noise ratio (SNR) because SHG and THG intensities scale quadratically with the number density of molecules within the focal volume [13]; therefore, SNR is only limited by integration time due to the low noise. Compared to HGM, RCM is limited by the noise from other structures, which scatter the beam before the focal spot; hence, RCM signal-to-noise is limited because the intensity scales linearly with the number density of molecules the same as the noise, and therefore, higher contrast cannot be obtained by longer integration times.

We imaged EMPD skin sample with two similar HGM systems (Fig. 1). For details, please refer to [14], [15]. The excitation source was a home-built Cr:forsterite laser, capable of producing 38 femtosecond pulses with a repetition rate of 105MHz at a central wavelength of 1262 nm (bandwidth: 91 nm). The dispersion of the optical components for lengthening the laser temporal pulsewidth in tissues was compensated by a pair of double chirp-mirrors [14],[15]. A 40x, 1.15 numerical aperture (NA) water-immersion objective (UAPON 40XW340, Tokyo,
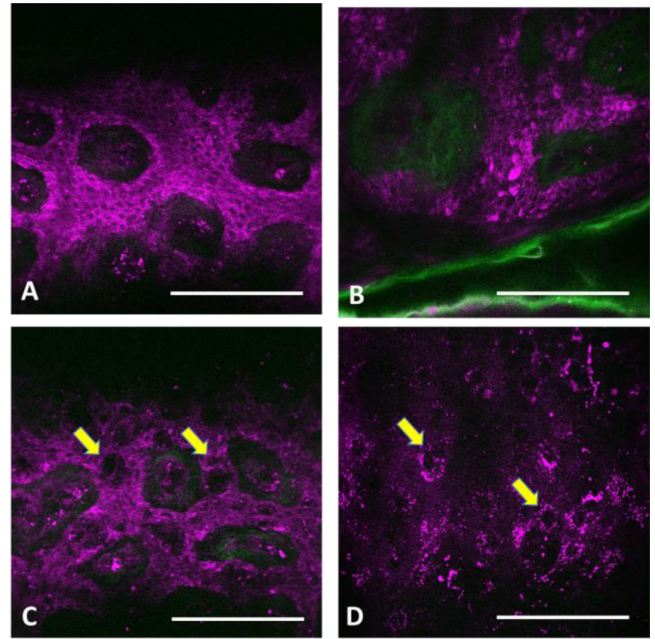

Fig. 1. Representative HGM images of EMPD lesions and normal skin. In these images, green color represents SHG signals and magenta represents THG signals. HGM image of normal skin for (A) system 1 and (B) system 2. HGM image of EMPD lesions for (C) system 1 and (D) system 2. Scale bar = 100 μm. Yellow arrows indicate Paget cells.

Olympus) focused the scanning beam onto the sample. The average power on the sample surface was around 100 mW. The above system parameters were optimized [14],[15] for penetration, contrast, and sample viability [14], [15], [16]. The noninvasiveness and photodamage concern of the systems had been reported in a previous study [16]. In this *ex vivo* study of skin samples, we observed no morphological alteration of epidermal cells or dermal structures, such as structural distortion or cellular burst. Additionally, we conducted histopathological examination on the *ex vivo* skin sample after the HGM imaging process. The H&E examination reported no photo-damage of the skin tissues.

The entire protocol was approved by the Research Ethics Committee of National Taiwan University Hospital (No. 201309036DINC), and informed consent was obtained from each subject prior to study entry. The *ex vivo* surgical samples removed from EMPD lesions were imaged immediately by HGM. A series of 3D HGM image stacks were taken at different locations across the skin sample. Each 3D HGM image stack was composed of series of 2D *en face* images, 512 × 512 pixels under 235 × 235 μm^2 field of view. Beside one sample with thin epidermis, each 3D image stack was taken from the skin surface with a depth of 180 μm. For the sample with extremely thin epidermis, the imaging depth was 138 μm, surpassing the depth of DEJ. After the HGM imaging, the skin biopsy samples were fixed in 10% formalin, embedded in paraffin, cut into 10-micron sections, and stained with hematoxylin and eosin (H&E). The histopathological features of the H&E sections were reviewed by a dermatopathologist and the borders between
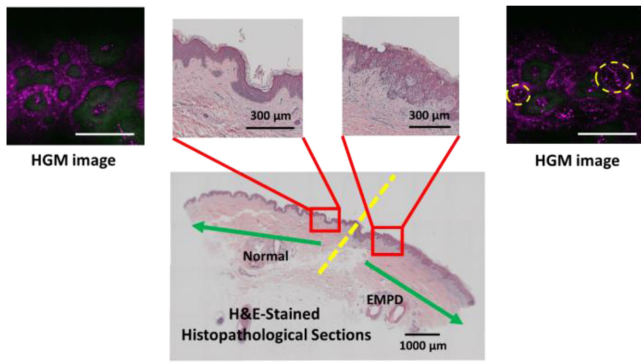
Fig. 2. Mapping clinical diagnosis result of H&E-stained histopathological sections to HGM image stacks. Scale bar = 100 μm. Yellow circles in the right panel indicate the Paget cells.
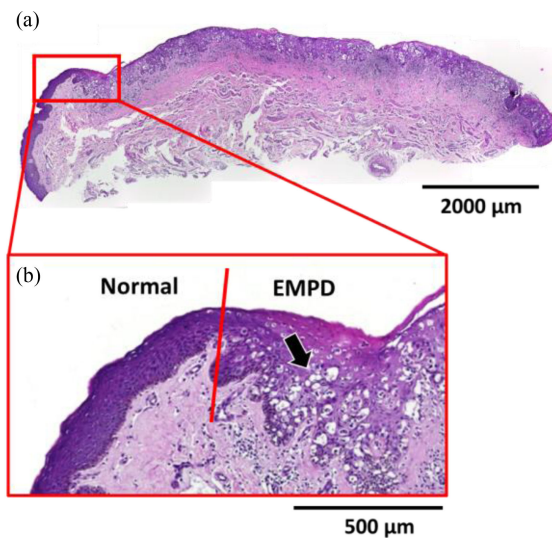

Fig. 3. Details of EMPD and perilesional normal skin in H&E-stained histopathological sections with two different scales.

normal and EMPD lesions were identified on the stitched microscopic photos. According to the locations, we then mapped the result of H&E-stained histopathological sections, which is the gold standard of EMPD examination, to the series of HGM image stacks so as to label each 3D HGM image stack as EMPD lesions or normal skin (Fig. 2). The H&E-stained histopathological sections of EMPD showed tumor cells (Paget cells) with abundant pale cytoplasm in the epidermis distributing in isolated forms or in clusters. The epidermis of EMPD lesions are often acanthotic and hyperkeratotic, comparing to the surrounding normal counterpart (Fig. 3). In 3D HGM imaging approach, SHG is the depth indicator and thus differentiates *in situ* tumor from tumor with dermal invasion. On the other hand, THG distinguishes the tumor cells from normal cells because the EMPD lesion is characterized by round-shaped, dark cells (Fig. 1) with pale THG signals in the stratum basale and stratum spinosum layers. Therefore, both SHG and THG provide essential and complimentary information for physicians to identify the condition and border in EMPD diagnostics.

B. Deep Learning Algorithm

With various medical imaging techniques nowadays, they provide physicians advanced information such as tumor locations and status for diagnosis and therapeutic decisions. Challengingly, it takes time and effort for dermatologists to analyze the information from the images [17], especially if the information is in a 3D form. To determine the lesion region and border, an approach that can accurately differentiate tissues at a specific location as cancer or normal in no time is needed. Recently, the rapid advancements in machine learning and deep learning, which is an approach of artificial intelligence, provide more accurate and faster models in medical imaging processing techniques. Several studies have conducted deep learning classification techniques based on different CNN models to distinguish normal and abnormal (cancer) states of tissues [18]–[21]. Besides, two studies applied deep learning on multiphoton images for cancer tissue classification [22], [23]. These studies demonstrated that deep learning can provide high sensitivity and high specificity diagnosis for various fields of medical imaging applications.

In this study, to reach a faster and better training result, an image pre-processing procedure was implemented to resize all HGM images by down-scaling averagely the original $512\times512\times96$ pixels (to facilitate the down-scaling process, we excluded the top 2 layers and the bottom 2 layers without affecting the results) into $64\times64\times32$ pixels. Because different skin tissue samples have different epidermal thickness, all 3D images would be padded into 96 layers if the number of 3D image layers is less than 96 layers. In our 3D HGM image dataset, we adopted 2286 3D image stacks in total as our dataset, including 1325 EMPD stacks and 961 normal stacks. Smaller images but with enough feature information could increase the efficiency of the training process within limited memory and would provide similar training performance to original 3D images without scaling. After resizing the images of each stack, the image stacks were randomly divided into 3 datasets for training, validation and test in proportion of 83.3%, 8.3% and 8.3% respectively (The proportion is 10:1:1). Although the datasets were randomly divided, the distribution of the normal and cancer tissue image stacks was the same in each dataset, which provides better training performance.

With a small number of images, fine-tuning or transfer learning on a well-trained model with own images has been reported to provide a better feature-differentiation performance than scratch from model [24]. In the medical image analysis field, transfer learning from natural image to medical image is one of the most practical method in deep learning. However, 3D medical imaging tasks are often solved in 2D models, thus losing rich 3D information and the performance is compromised. Models Genesis [25] is a powerful 3D pretrained model, significantly outperforming both in segmentation and classification of several major 3D medical image applications. With a small amount of HGM 3D image stacks, fine-tuning on Models Genesis would provide a better performance than other 2D models. To avoid the over-fitting problem, in this study we shrunk and modified the pretrained model architecture with additional layers and finetuned the model with the weights from pretrained Models Genesis.
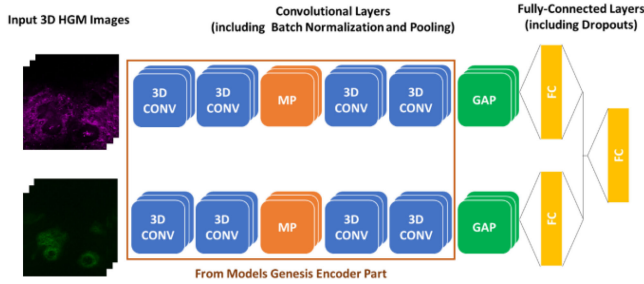
Fig. 4. EMPD classification deep learning network architecture. 3D CONV: 3D convolutional layer; MP: max pooling layer; GAP: global average pooling layer; FC: fully connected layer.

TABLE I
RESULT OF SCRATCH FROM A 3D CNN MODEL AND FINETUNE FROM A PRETRAINED MODEL

	Loss	Sensitivity (%)	Specificity (%)	Accuracy (%)
Scratch from a 3D CNN model	0.5579	97.09	37.50	74.25
Finetune from a pretrained model	0.1781	98.05	90.63	95.21

TABLE II
RESULT OF TWO-PATH VS. SINGLE-PATH MODEL

	Loss	Sensitivity (%)	Specificity (%)	Accuracy (%)
Single channel as a single-path network	0.2396	90.29	89.06	89.82
Merged channels as a single-path network	0.2302	95.15	87.50	92.22
Separate channels as a two-path network	0.1781	98.05	90.63	95.21

TABLE III
RESULT OF DIFFERENT DROPOUT SETTINGS

	Loss	Sensitivity (%)	Specificity (%)	Accuracy (%)
Without dropout	0.1903	93.20	87.50	91.02
Dropout with p=0.05	0.2371	94.17	87.50	91.62
Dropout with p=0.1	0.1781	98.05	90.63	95.21
Dropout with p=0.3	0.2496	93.20	85.94	90.42
Dropout with p=0.5	0.2689	86.41	92.20	88.62

We trained our training dataset with our designed classification network (Fig. 4). We trained the network for 300 epochs with an early stopping of 50 epochs. This technique prevents the overfitting problem by stopping training after a number of 50 epochs without a decrease in the validation loss. The part in the square (Fig. 4) was extracted from the encoder part of Models Genesis, whose initial weights were also adopted from the pretrained Models Genesis. All convolutional layers applied the rectified linear activation function (ReLU) [26] and were regularized using batch normalization. One 3D global average pooling layer was applied after the last convolutional layer. Dropout was applied to the 3D max pooling layer, the 3D global average pooling and the two fully-connected layers after the 3D global average pooling layer. Additionally, for the two fully-connected layers after the 3D global average pooling, scaled exponential linear unit (SELU) [27] activation and LeCun normal [27], [28] kernel initializer was applied to them. Sigmoid activation and L2 regularization [29] were applied to the last fully-connected layer for binary classification. Network optimization was performed using the Adam [30] with a batch size of 18 and binary cross-entropy as the loss cost.

III. RESULTS

A. Scratch 3D CNN Model vs. Finetuned Pretrain Model

To prove the concept that finetuning from a pretrained model brings better performance, we trained a simple 3D CNN model from scratch to confirm it. After training, finetuning from a pretrained model got much better sensitivity, specificity and accuracy than scratch from the 3D CNN model (Table I).

B. Two-Path Model vs. Single-Path Model

There are two imaging modalities in HGM images, SHG and THG. In this study, we conducted three different designs: two channels separately with a two-path network, two channels merged as a single-path network, and single channel (only THG signal) with a single-path network. After training from these three networks, two channels separately as a two-path network delivers the best result (Table II). We speculate that it is because SHG and THG are both important but different and need to be analyzed separately to get the correct EMPD diagnosis. For example, separate analysis can provide sharp contrast of the DEJ.

C. Dropout value

We also trained with different dropout ratio because a deep learning model contains lots of parameters, thus conducting dropout with enough parameters randomly brings better performance. After training three different settings, dropout with p = 0.1 got the best result (Table III).

D. Scaled ELU Activation and LeCun Normal Kernel Initializer

With SELU activation, when the input is in normal distribution (LeCun normal initializer), the output is also in normal distribution. SELU with LeCun normal converges the network faster, because the internal normalization (SELU) is faster than external normalization (like batch normalization). SELU with LeCun normal also suppresses the vanishing and exploding gradient

TABLE IV
RESULT OF SELU AND LECUN NORMAL SETTING

	Loss	Sensitivity (%)	Specificity (%)	Accuracy (%)
Without SELU and LeCun normal	0.1956	94.17	89.06	92.22
With SELU and LeCun normal	0.1781	98.05	90.63	95.21

TABLE V
RESULT OF L2 REGULARIZATION SETTING

	Loss	Sensitivity (%)	Specificity (%)	Accuracy (%)
Without L2 regularization	0.3910	87.38	85.94	86.83
With L2 regularization value = 0.05	0.2310	94.17	85.94	91.02
With L2 regularization value = 0.01	0.1781	98.05	90.63	95.21
With L2 regularization value = 0.001	0.2014	93.20	90.63	92.22

TABLE VI
RESULT OF BALANCED DATASET

	Loss	Sensitivity (%)	Specificity (%)	Accuracy (%)
Without balanced dataset	0.1781	98.05	90.63	95.21
With balanced dataset	0.1426	98.06	93.18	95.81

problem and brings a better training performance. After training with and without this setting, network with SELU activation and LeCun normal kernel initializer brings better performance (Table IV).

E. L2 Regularization

To keep all the image features but avoid overfitting from training data, we adopted L2 as our regularizer. After four different L2 regularization value setting trials, we found that L2 regularization value = 0.01 provide the best performance (Table V). Smaller L2 regularization value cannot avoid the overfitting problem and larger L2 regularization value might abandon too much image features.

F. Data Balance

From the first original dataset, we only adopted 3D HGM images from surgical samples of EMPD patients, therefore, EMPD images are more than normal skin tissue images. We observed that specificity was much lower than sensitivity when we utilized this dataset and we speculated that this phenomenon is from the unbalanced data. Therefore, we imaged more normal skin tissue images to make normal skin images as much as

EMPD images. As a result, training with a more balanced dataset provides better performance for both sensitivity and specificity.

IV. DISCUSSION

EMPD lesions frequently show ill-defined tumor borders, which makes it difficult for physicians to define the accurate resection line for surgery. Currently, to ensure complete removal of EMPD, physicians have to either choose wider surgical border or spend tremendous efforts to map tumor border pre-operationally. Both approaches cause large trauma, which brings problems to patients in appearance and life, and tumor may spread beyond the original border during the biopsy examination period. To avoid these problems, a noninvasive method to instantly provide accurate classification between EMPD tissues and normal skin tissues with high sensitivity, specificity and accuracy is needed. With the trauma-free mean, accurate and immediate assessment of EMPD margin before surgery would reduce the need of time-consuming and invasive strip biopsy and mapping biopsy and could lower the recurrence rate than without pre-operational border assessment. In this study, we examined if 3D CNN deep learning is capable to accurately determine whether the location of the trauma-free HGM 3D-imaged skin tissue is EMPD or surrounding normal. With 98.06% sensitivity, 93.18% specificity, and 95.81% accuracy achieved, our study confirms the performance of the proposed technology to meet the need for EMPD preoperational border assessment with *ex vivo* samples. Currently, the peripheral extent of EMPD and identification of tumor margins can be extremely difficult to define based on the clinical examination. According to the clinical examination, physicians have a predicted surgical border before the surgery, but have doubts about whether the surgical border is adequate. Our technique can provide real-time identification to assist physicians in mapping out suspected margins, which can speed up the *in vivo* margin assessment process and map the tumor pre-operatively, facilitating the planning of more rational surgery. It also has high potential to reduce the frequency of histopathologically involved margins in excision specimens and therefore reduce the need for additional surgery. The result makes us believe that the technique has great potential for the *in vivo* and noninvasive preoperational surgical border detection of EMPD. In our study, we utilized two HGM imaging systems with two slightly different spacings in the vertical direction, $1.8\mu m$ and $2\mu m$, and the result wasn't affected by this difference.

Many studies have proposed that the deep learning technique can accurately classify medical images of various diseases or can accurately determine whether it is a specific cancer [18]–[23]. For deep learning applied on nonlinear optical imaging, one multiphoton study [22] achieved over 95% sensitivity and 97% specificity on classification of the mice ovarian tissues as healthy or cancerous. Additionally, the other study [23] collected second-harmonic generation and two-photon excitation fluorescence 2D images of the DEJ in unstained fixed tissues for automatic classification of healthy and dysplastic classes using a deep learning method with 93.5% sensitivity, 95.0% specificity and 94.2% accuracy. Although these two studies demonstrated their ability on cancer tissue classification, 2D-image-based

approach needs subjective and time-consuming selection of the *en face* 2D image at a specific depth. Even though deep learning applied on 3D medical images is rare, our study supports the high accuracy of 3D deep learning on 3D HGM images to classify normal and malignant EMPD cases. With the 3D approach, instead of manually selecting 2D skin images at a specific depth for interpretation, one can now blindly perform 3D skin images starting from the surface and digitally achieve high precision automatic classification of the imaged site, which is with a sub-millimeter size. In addition, our 3D deep learning neural network can reach over 90% sensitivity, specificity and accuracy as other 2D deep learning studies.

Currently, margin assessment before surgery takes a long time. It takes an average of 35 days from the first mapping biopsy to the operation [5]. One study reported that Paget cells may spread beyond the biopsy sites during the period from the completion of the biopsy examination to the surgery. If the mapping can cover much more than 22 sites and the classification result can be obtained immediately, it can help physicians to perform the operation instantly, shorten the waiting time and prevent the spread of cancer cells. Our method, combination of 3D deep learning and 3D HGM images, shows high potential to provide such a solution in the future.

It is noted that in this study, we only recruited Asians with skin phototype III or IV as our volunteers. Even though it is unclear what the effect of other skin types on the EMPD diagnosis by deep neural networks is, pathologically it is known that the morphology of the EMPD lesions is hardly affected by different skin phototypes. Combining with our previous in vivo HGM studies on volunteers with different skin phototypes [31], [32], which showed consistent HGM capability to penetrate through the dermal-epidermal junction, we infer that with more images from patients of different skin phototypes, one needs to perform optimization of hyperparameters on the deep neural network, but there is no need to modify the architecture of the deep neural network to achieve similar results.

Even though this is only an ex vivo study, the impact of the presented development with the HGM and 3D deep learning model is to provide instant information at the pre- or intraoperation time to make the surgeons excise the lesions completely with free section margins. This technique present here will provide the basis for future ex vivo and in vivo clinical applications in EMPD disease diagnosis, preoperational margin assessment, and intraoperational border determination like in Mohs micrographic surgery.

V. CONCLUSION

We presented a novel technique for non-invasive in-vivo EMPD identification through HGM imaging analyzed by 3D-CNN models providing real-time malignancy classification to assist physicians in rapid mapping of regions with suspected malignancy. To demonstrate the technique's viability, models trained on 2095 HGM image stacks of ex-vivo specimen was benchmarked against corresponding gold-standard dermapathological examinations, investigating the influence of several deep learning approaches on its resulting performance, achieving 98.06% sensitivity, 93.18% specificity, and 95.81% accuracy.

REFERENCES

[1] T. Ito, Y. Kaku-Ito, and M. Furue, "The diagnosis and management of extramammary Paget's disease," *Expert Rev. Anticancer Ther.,* vol. 18, no. 6, pp. 543–553, Jun. 2018.

[2] A. Hendi, D. G. Brodland, and J. A. Zitelli, "Extramammary paget's disease: Surgical treatment with Mohs micrographic surgery," *J. Amer. Acad. Dermatol.,* vol. 51, no. 5, pp. 767–773, Nov. 2004.

[3] C. Gaudy-Marqueste *et al.*, "The 'spaghetti technique': An alternative to mohs surgery or staged surgery for problematic lentiginous melanoma (lentigo maligna and acral lentiginous melanoma)," *J. Amer. Acad. Dermatol.,* vol. 64, no. 1, pp. 113–118, Jan. 2011.

[4] B. J. Kim, S. K. Park, and H. Chang, "The effectiveness of mapping biopsy in patients with extramammary Paget's disease," *Arch Plast Surg.*, vol. 41, no. 6, pp. 753–758, 2014.

[5] S. O. Park, J. H. Ha, K. Y. Hong, and H. Chang, "Usefulness of mapping biopsy in the treatment of penoscrotal extramammary Paget's disease," *Ann. Surg. Oncol.,* vol. 24, no. 11, pp. 3229–3236, Oct. 2017.

[6] P. Guitera *et al.*, "Reflectance confocal microscopy for diagnosis of mammary and extramammary Paget's disease," *J. Eur. Acad. Dermatol. Venereol.,* vol. 27, no. 1, pp. e24–e29, Jan. 2013.

[7] Z.-Y. Pan, J. Liang, Q.-A. Zhang, J.-R. Lin, and Z.-Z. Zheng, "In vivo reflectance confocal microscopy of extramammary Paget disease: Diagnostic evaluation and surgical management," *J. Amer. Acad. Dermatol.,* vol. 66, no. 2, pp. e47–e53, Feb. 2012.

[8] O. Yélamos *et al.*, "Handheld reflectance confocal microscopy for the detection of recurrent extramammary Paget disease," *JAMA Dermatol.,* vol. 153, no. 7, pp. 689–693, 2017.

[9] A. P. Raphael, D. Tokarz, M. Ardigò, and T. W. Prow, "Reflectance confocal microscopy and aging," in *Textbook of Aging Skin*, M. A. Farage, K. W. Miller and H. I. Maibach, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 1–17.

[10] C. K. Sun *et al.*, "Slide-free imaging of hematoxylin-eosin stained whole-mount tissues using combined third-harmonic generation and three-photon fluorescence microscopy," *J. Biophotonics,* vol. 12, no. 5, May 2019, Art. no. e201800341.

[11] M.-R. Tsai *et al.*, "Differential diagnosis of nonmelanoma pigmented skin lesions based on harmonic generation microscopy," *J. Biomed. Opt.,* vol. 19, no. 3, 2014, Art. no. 036001.

[12] Y. H. Ho, Y. Pan, C. K. Sun, and Y. H. Liao, "Presence of intralesional melanocytes as a histopathological feature of actinic keratosis based on in vivo harmonic generation microscopy in Asians," *Photodermatol. Photoimmunol. Photomed.*, vol. 37, no. 1, pp. 20–27, Jan. 2021.

[13] R. W. Boyd, " Wave-equation description of nonlinear optical interactions," in *Nonlinear Optics*, 3rd ed., Orlando, FL, USA: Academic, 2008, pp. 69–133.

[14] S. Chakraborty, S.-T. Chen, Y.-T. Hsiao, M.-J. Chiu, and C.-K. Sun, "Additive-color multi-harmonic generation microscopy for simultaneous label-free differentiation of plaques, tangles, and neuronal axons," *Biomed. Opt. Exp.,* vol. 11, no. 2, pp. 571–585, 2020.

[15] C.-K. Sun *et al.*, "Slide-free clinical imaging of melanin with absolute quantities using label-free third-harmonic-generation enhancement-ratio microscopy," *Biomed. Opt. Exp.,* vol. 11, no. 6, pp. 3009–3024, 2020.

[16] Y.-H. Liao *et al.*, "In vivo third-harmonic generation microscopy study on vitiligo patients," *J. Biomed. Opt.,* vol. 25, no. 1, Jan. 2020, Art. no. 014504.

[17] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-net," *PLoS One,* vol. 14, no. 5, 2019, Art. no. e0217293.

[18] S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network," in *Proc. Int. Conf. Comput., Math. Eng. Technol.,* Sukkur, Pakistan, 2018, pp. 1–5.

[19] J. Hyeon, H.-J. Choi, B. D. Lee, and K. N. Lee, "Diagnosing cervical cell images using pre-trained convolutional neural network as feature extractor," in *Proc. IEEE Int. Conf. Big Data Smart Comput.,* Jeju, South Korea, 2017, pp. 390–393.

[20] K. Nazeri, A. Aminpour, and M. Ebrahimi, "Two-stage convolutional neural network for breast cancer histology image classification," Lecture Notes in Computer Science, vol. 10882, pp. 717–726, 2018.

[21] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature,* vol. 542, no. 7639, pp. 115–118, Feb. 2017.

[22] M. J. Huttunen *et al.*, "Automated classification of multiphoton microscopy images of ovarian tissue using deep learning," *J. Biomed. Opt.,* vol. 23, no. 6, Jun. 2018.

[23] M. J. Huttunen *et al.*. "Multiphoton microscopy of the dermoepidermal junction and automated identification of dysplastic tissues with deep learning," *Biomed. Opt. Exp.,* vol. 11, no. 1, pp. 186–199, Jan. 2020.

[24] Y. Bengio, "Deep learning of representations for unsupervised and transfer learning," in *Proc. ICML workshop unsupervised transfer Learn.*, Bellevue, WA, USA, 2011, pp. 17–36.

[25] Z. Zhou *et al.*, "Models genesis: Generic autodidactic models for 3D medical image analysis," in *Proc. Med. Image Comput. Comput. Assist. Intervention,* Shenzhen, China, 2019, pp. 384–393.

[26] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Int. Conf. Mach. Learn.,* Haifa, Israel, 2010, pp. 807–814.

[27] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," in *Adv. Neural Inf. Process. Syst.,* Long Beach, CA, USA, 2017, pp. 971–980.

[28] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient backprop," in *Neural Networks: Tricks of the Trade*, Berlin, Heidelberg, Springer, 2012, pp. 9–48.

[29] A. Y. Ng, "Feature selection, l 1 vs. L 2 regularization, and rotational invariance," in *Proc. 21st Int. Conf. Mach. Learn.,* New York, NY, USA, 2004, Art. no. 78.

[30] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations*, San Diego, CA, USA, May 2015.

[31] K.-H. Lin, Y.-H. Liao, M.-L. Wei, and C.-K. Sun, "Comparative analysis of intrinsic skin aging between caucasian and asian subjects by slide-free in vivo harmonic generation microscopy," *J. Biophotonics,* vol. 13, no. 4, Apr. 2020, Art. no. e201960063.

[32] S.-Y. Chen *et al.*, "In vivo virtual biopsy of human skin by using noninvasive higher harmonic generation microscopy," *IEEE J. Sel. Topics Quant.,* vol. 16, no. 3, pp. 478–492, May/Jun. 2010.

Yi-Hua Liao received the M.D. degree from the College of Medicine, National Taiwan University, and the Ph.D. degree from the Graduate Institute of Pathology, College of Medicine, National Taiwan University, Taipei, Taiwan, in 1996 and 2007, respectively. She is currently an Associate Professor with the Department of Dermatology, National Taiwan University, and a Supervisor of the Taiwanese Dermatological Association. She is responsible for dermatologic surgery special clinic, cosmetic surgery, and laser special clinic.

Chi-Kuang Sun (Fellow, IEEE) received the Ph.D. degree in applied physics from Harvard University in 1995 and was an Assistant Researcher with the UCSB QUEST Center from 1995 to 1996. In 1996, he joined National Taiwan University, where he is currently a Distinguished Professor of photonics and optoelectronics. He founded the NTU Molecular Imaging Center. His research focuses on nano-acoustics, femtosecond optics, THz optoelectronics, and biomedical imaging. He is a fellow of OSA, SPIE.

Chia-I Chen received the bachelor's degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2018. She is currently working toward the master's degree in photonics and optoelectronics from National Taiwan University, Taipei, Taiwan. Her research interests include the optical and biomedical image processing, and deep learning in image analysis.