Grenze International Journal of Engineering and Technology, June Issue

Identification and Localization of Malanoma Skin Cancer in Neuropathological Images

Rosey Chauhan[1] and Sunil Gupta[2]
[1-2]Jaipur National University/Computer Science, Jaipur, India
Email: errose.it@gmail.com, sunilg95@rediff.com

***Abstract*—Early detection and treatment of melanoma can improve survival rates. The demand for computational pathology (CPATH) systems has been emphasized by a predicted rise in skin cancer cases and a shortage of dermatopathologists. Deep learning (DL) models in CPATH systems have the capability of using underlying morphological and cellular cues to detect the existence of melanoma. This study suggests a new approach to identify melanoma and distinguish benign from malignant melanocytic tumors in whole slide images (WSI). With excellent accuracy, this technique locates lesions on a WSI to highlight possible areas of interest for pathologists. It's interesting to note that our DL technique uses a single CNN network to build localization maps first, then uses those maps to make slide level predictions to identify patients with melanoma. With a 0.992 F1 score and 0.99 sensitivity on unseen data, our top model delivers excellent patch-wise classification results.**

***Index Terms*— Computational Pathology, Deep learning, Melanoma Diagnosis, Whole Slide Images.**

I. INTRODUCTION

Skin cancer that is aggressive is called malignant melanoma [1]. Melanoma is the most common cause of death while representing just around 1% of instances of skin cancer [2]. When melanocytes begin to multiply swiftly and unchecked in the epidermis and dermis layers of the skin and form malignant lesions, melanoma skin cancer occurs [2]. If the tumor is not treated right away, it will probably advance to a new stage and could even spread to other body areas [1]. The importance of early detection and precise diagnosis cannot be overstated. A skin sample is removed via punch biopsy, processed to create a glass slide, and then examined histopathologically to detect malignancy. Afterwards, pathologists examine the glass slide under a microscope [3]. Humans frequently find it time-consuming and difficult to discern between benign tumors and early-stage cancer[17]. By using computer analysis, a digitized version of a glass slide known as a Whole Slide Image (WSI) can get beyond this limitation of standard histopathology [4, 5]. By learning morphological and cellular patterns and making predictions, computational pathology (CPATH) algorithms may automate a variety of diagnostic procedures [6, 7, 8]. By giving pathologists a second viewpoint and localizing the region of interest, CPATH systems offer a great potential for detecting melanoma and assisting them in their work (ROI). The majority of earlier studies [10, 11] on the diagnosis of melanoma concentrate on hand-crafted features selection from data such as first order texture characteristics, color intensity, entropy, eccentricity, etc. Dermoscopic pictures were employed by Jafari et al. [11] to achieve color-space transformation and extract morphological characteristics. Asymmetry, Boundary irregularity, Color patterns, and Diameter were employed in their strategy to discriminate between

Grenze ID: 01.GIJET.9.2.24

malignant and benign tumors. In a similar manner, Sheha et al. [10] employed a multilayer perceptron classifier to distinguish between malignant melanoma and melanocytic nevi using the gray-level co-occurrence matrix (GLCM) and textural characteristics. Manual feature engineering is laborious, however, and it has been shown that data-driven features outperform categorization performance in medical pictures [12]. DL methods automatically identify characteristics from data, but they need a large quantity of data and clinical labelling, which may not always be accessible. To make up for the dearth of training data, transfer learning has been a popular initialization strategy in DL [3, 13].

In order to accomplish a new classification task by automatically learning hidden features from the data, it mostly depends on transferring information to a DL model that was trained on a different task. Transfer learning was recently used by Rajitha et al. [13] to extract characteristics from unstained skin samples and categorize dermatophytosis using four Convolutional Neural Networks (CNN). Logu et al. [14] utilized histological pictures to distinguish between healthy tissue and cutaneous melanoma using a similar CNN-based method. As their model was developed for a binary objective, it was unable to distinguish between benign and malignant tumors. Wang et al[15] .'s analysis of several CNN designs allowed them to distinguish between malignant and benign nevi. In their tests, VGG16 [9] outperformed other CNN designs in patch-wise classification performance; however, their technique required the use of a different random forest classifier to make slide-level predictions.

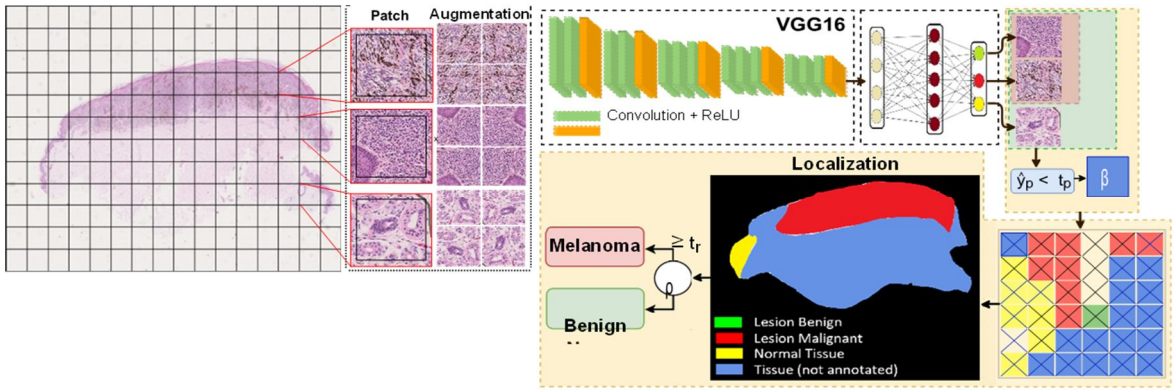

Fig. 1 shows a summary of our suggested melanoma detection technique. Patches of Whole Slide Images (WSI) are created based on labelled areas. Before to delivering the photos to the feature extractor, pre-processing is done to enhance and normalise the images. This model's main component is VGG16 [9], which is followed by a three-layer fully-connected classifier. Patch-wise classification is carried out using binary and multiclass models, which assign a class to each patch. To locate lesions, the categorised patches are joined using their coordinates. Lastly, slide-level predictions of patients with melanoma vs. those with benign nevi are made using the localization maps

This study suggests using a DL technique to distinguish between benign and cancerous lesions in a WSI, as seen in Fig. 1. It is adopted VGG16 [9] as the foundation of our model since it has been a well-liked option in literature [13, 15]. To determine the ideal parametric setting for patch-wise classification models, experiments were created. Combining the categorized patches enables pathologists to identify potential ROIs and locate lesions over the WSI. Then, using pixel-wise thresholding to identify patients with melanoma on the slide-level, we test our top models on the unseen WSIs to first forecast patches for getting the localization map.

II. DATASET

The total 81 images examined skin biopsy from ISIC dataset. These samples lacked any atypical nevi or in situ melanomas and were clearly benign or malignant. Glass slides were scanned at a 40x magnification using a Hamamatsu Nanozoomer s60 in ndpi format with 0.2199 mm x 0.2199 mm pixels. Slide level labels of benign nevus, benign lentigo, and malignant melanoma were given to all WSIs for 40, 7, and 43 WSIs, respectively. Training, validation, and testing were divided by 73/8/9, with benign nevus and benign lentigo being handled as a single label. Each slide's sections with benign lesions, malignant lesions, and various tissue types were crudely annotated by a pathologist. For a tractable calculation, annotated areas in WSIs are split into tiny sub-images (patches) [8, 12]. Sample patches from each class are shown in Fig. 2 at a 10x magnification level.

III. METHODOLOGY

Fig. 1 provides a visual summary of the melanoma detection procedure. As the feature extractor (backbone) and three fully connected (FC) layers of a bespoke classifier, our suggested approach makes use of VGG16 [9]. The

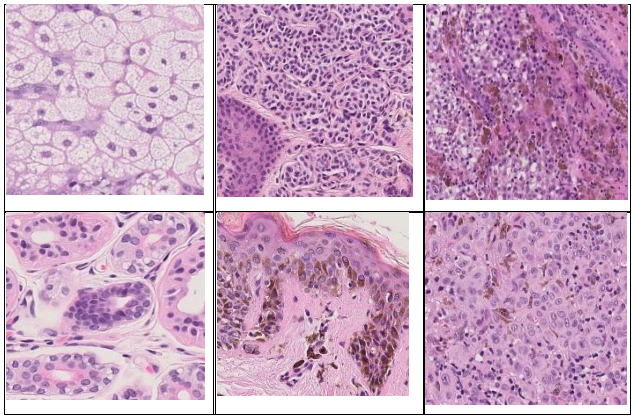

Fig. 2. Images of patches of normal tissue, benign, and malignant lesion class from the dataset at 10x magnification.

next task of melanoma detection is by employing a two-step procedure. In the first stage, we create models for binary and multiclass patch-wise classification. Malignant and benign lesion classifications are distinguished in the binary classification test. Three classifications that also include the normal tissue are used in the multiclass classification job. The slide-level result is created in the second stage by using patches categorized with a greater probability in a class to produce a localization map, as shown below.

A. Pre-processing

The annotated areas of WSIs were divided into tiny patches of 256 256 pixels at 2.5x, 10x, and 40x magnification levels since CNN cannot process an entire WSI at once. Secondly, the RGB to HSV conversion and thresholding of the Hue channel for the purple and pink tones allowed for the background-foreground separation. Afterwards, patches were extracted that had foreground and annotation masks that overlapped by 70%. For the memory-effective patching technique, then employed "patch-on-fly" [8, 9]. Geometric adjustments with 224224 random cropping were made to the underrepresented class to correct the class imbalance. Before being normalized to the mean and standard deviation of the training set, all patches were scaled to 224224 pixels to match the input size of the pre-trained model.

B. Patch-wise Classification

There are binary and multiclass models that assign a single class to a patch in the patch-wise classification stage. Using a frozen feature extractor, then trained a binary baseline model (Model baseline). The dataset that was retrieved at a 10x magnification was used to train Model baseline. Using unfrozen feature extractors, three binary models (Model2.5x, Model10x, and Model40x) are trained using datasets that were extracted at 2.5x, 10x, and 40x, respectively. A 10x magnification dataset was used to train our three-output node multiclass model (Model multiclass). Models provide a probability vector (Yp) for an input patch (x) with a ground label (Yz), as indicated in Eq. (1), where B, M, and N stand for benign lesion, malignant lesion, and normal tissue class, respectively.

$$Yp = \begin{matrix} [Y\hat{}b, Y\hat{}m] & \text{if Binary} \\ [Y\hat{}b, Y\hat{}m, Y\hat{}n\ ] & \text{if Multiclass} \end{matrix} \quad (1)$$

WSI was used for the slide-level prediction phase, therefore the inference stage will come across tissue types that haven't been seen before. A new class () and probability threshold (Tp) are developed to deal with the unseen tissue types. The predicted class (Yp) is deemed irrelevant tissue and given the label as stated in Eq if the predicted output (Yp) is less than the projected threshold (Tp) (2). In brief, the output of the model (Yp) is either binary or has three classes, but the overall forecast (Yp) has an additional class. Here, Tp aids in effectively detecting the presence of any lesion without requiring models to be trained on new tissue classes. By increasing lesion detection and reducing false positives on the validation set, used to establish the optimum tp.

$$Y\hat{}p = \begin{matrix} (argmax\ (Yp)\ if\ max\ (Yp\ ) \geq Tp \\ \beta,\ Otherwise \end{matrix} \quad (2)$$

C. Slide-level Prediction

At the slide-level prediction stage, then combine all of the previously identified patches using the top-performing binary and multiclass models to construct masks. The pixel in the down-sampled map is coloured using coordinates relating to the position of the patch. Then, use Eq to determine the ratio () of the malignant lesion class based on the localization map (3). If is higher than a ratio (tr) (see Eq. (4)), now it can forecast melanoma on a slide (ys). In other words, tr establishes the level of malignancy at which a patient with melanoma should be reported. In order to decrease the possibility of false negatives, such as the categorization of melanoma as a benign nevus, a reasonable value of tr is chosen depending on the validation set.

$$\rho = \frac{\text{No.of malignant pixels}}{\text{No.of malignant pixels} + \text{No.of benign pixels}} \quad (3)$$

$$\hat{Y}s = \textit{Malanoma if } \rho \geq Tr \quad \textit{Benign Nevus, Otherwise}$$

IV. Evaluation Matrix

The measurements for precision, F1, sensitivity, and specificity. The terms true positive, false negative, false positive, and true negative predictions, respectively, must be denoted as TP, FN, FP, and TN. The accuracy of the model's predictions is measured as (T P + T N)/(T P + F N + F P + T N). Recall = Sensitivity = T P/(T P + F N) and Precision = T P/(T P + F P), where Recall = 2 (precision recall)(precision + recall) is harmonic mean. The patches that sensitivity tests identified as malignant lesions really belonged to the malignant class. The percentage of benign lesion patches that are properly diagnosed is shown by the specificity, T N/(T N + F P).

V. Implementation

Patch extraction was carried out using the Pyvips 2 package, and the approach was implemented using the Pytorch DL framework. The classifier was started with random weights, while the VGG16 [9] feature extractor was initialised using ImageNet weights. For validation accuracy, than employed the cross- entropy loss, the Adam optimizer, the learning rate of 0.0003, the batch size of 256, and early stopping of 8 epochs. Every model was created on an Nvidia A100 40GB GPU.

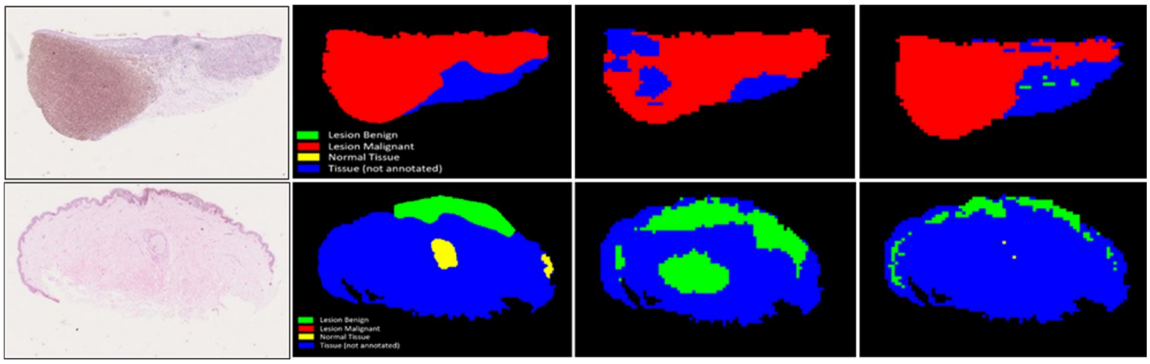

Fig. 3. A map of WSI localization the original Entire Slide Image is represented by the 1st column (WSI). Our ground truth is shown in the 2nd column. The results of the binary model are displayed in the 3rd column, while the results of the multiclass model are in 4th column

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

VI. Result

A. Patch-wise Classification

Patch extraction was carried out using the Pyvips 2 package, and the approach was implemented using the Pytorch DL framework. The classifier was started with random weights, while the VGG16 [9] feature extractor was initialized using ImageNet weights. For validation accuracy, it employed the cross-entropy loss, the Adam optimizer, the learning rate of 0.0003, the batch size of 256, and early stopping of 8 epochs. Every model was created on an Nvidia A100 40GB GPU. On the majority of assessment measures, Model2.5x surpasses all other

models. With an almost 5% F1 score, it greatly beat the baseline model. The maximum sensitivity value, however, is provided by Model10x, demonstrating that context is, at a high level, more significant than cellular specifics. Although it did not outperform binary models, Model multiclass had a greater accuracy than the baseline model and the literature. The main cause is the mislabelling of benign lesions as normal tissue. It might be because benign and malignant lesions are located close to the epidermis and may contain some normal tissue due to the rough annotations. On the unseen data, multiclass models exhibited higher F1 and sensitivity in predicting malignant lesions.

B. Slide-level Prediction

For slide-level prediction, pick the binary Model2.5x and Model multiclass from the prior experiment. Based on the validation established and fixed for the test, it is determined that the values for tp and tr are 0.99 and 0.04, respectively. According to Table 2, every patient in the test set had a correct prediction. The localization maps across a few sample WSIs from the test set are shown in Fig. 3. The binary model identifies some unannotated tissue in the first row as a malignant lesion. As shown in the annotation, the multiclass model also accurately diagnoses melanoma. It also indicates tiny benign tumours in regions without annotations. Due to its ignorance of the typical tissue architecture, the binary model interprets the normal adnexal structure in the second row as a benign lesion. Some of the tissue (on the left and right) is predicted by the multiclass model to be benign lesions. The multiclass model precisely localises lesions inside the specified border, whereas the binary model generally overestimates the size of lesions. At the slide level, both models can accurately predict the presence of benign nevi and melanoma. A diagnostic result and ROI may be obtained using the approach in clinical practice in around three minutes per WSI.

TABLE I. THE PATCH-WISE CLASSIFICATION PERFORMED ON THE VALIDATION AND TEST (UNSEEN) SET

	Models	Acc. (%)	F1	Sens.	Spec.
Validation set	Logu et al. [14]	96.5	0.965	0.957	0.977
	Wang et al. [15]	94.9	-	0.947	0.953
	Modelbaseline	93.1	0.948	0.926	0.942
	Model2.5x	99.32	0.995	0.990	0.999
	Model10x	98.13	0.986	0.994	0.954
	Model40x	96.02	0.971	0.986	0.906
	Modelmulticlass	96.63	0.979	0.986	0.882
Test	Model2.5x	99.07	0.990	0.989	0.998
	Modelmulticlass	98.27	0.992	0.990	0.973

TABLE II. SHOWS THE OUTCOMES FOR THE BEST BINARY AND MULTICLASS MODELS FOR THE SLIDE-LEVEL PREDICTION ON THE TEST SET

	Melanoma	Benign Nevus	Total (Acc.)
Model2.5x	4	5	9 (100 %)
Modelmulticlass	4	5	9 (100 %)
Ground Truth	4	5	

VII. CONCLUSION AND FUTURE WORK

Pathologists may benefit from CPATH systems automated diagnosis and region of interest. In this study, it is suggested a pixel-wise thresholding technique for classifying patches and generating a localization map to detect melanoma patients. It's interesting to note that our approach performs melanoma classification and segmentation tasks using a single Neural network. Models created at lower magnification levels successfully identify patients with melanoma using unobserved data, proving the effectiveness of our approach as a diagnostic tool. Further research should test the strategy using a bigger dataset. One additional diagnostic aspect that may be useful is the inclusion of tumour necrosis annotations.

REFERENCES

[1] C. Suephy Chen and Michelle L Pennie et al., “Diagnosing and managing cutaneous pigmented lesions: primary care physicians versus dermatologists,” Journal of general internal medicine, vol. 21, no. 7, pp. 678–682, 2006.

[2] American Cancer Society, “1.800.227.2345,” in cancer.org. https://www.cancer.org/content/dam/CRC/PDF/Public/8823.00.pdf, January 12, 2022.
[3] N. Kanwal and S. Fuster et al., “Quantifying the effect of color processing on blood and damaged tissue detection in whole slide images,” in 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). IEEE, 2022, pp. 1–5.
[4] Z. Tabatabaei and A. Colomer et al., “Residual block convolutional auto encoder in content-based medical image retrieval,” in 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). IEEE, 2022, pp. 1–5.
[5] S. Fuster and F. Khoraminia et al., “Invasive cancerous area detection in non-muscle invasive bladder cancer whole slide images,” in 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). IEEE, 2022, pp. 1–5.
[6] E. Abels and L. Pantanowitz et al., “Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the digital pathology association,” The Journal of pathology, vol. 249, no. 3, pp. 286–294, 2019.
[7] N. Wahab and I. M Miligy et al., “Semantic annotation for computational pathology: Multidisciplinary experience and best practice recommendations,” The Journal of Pathology: Clinical Research, vol. 8, no. 2, pp. 116–128, 2022.
[8] N. Kanwal and F. Perez-Buenoet al., “The devil is in the details: Whole slide image acquisition and processing for artifacts detection, color variation, and data augmentation: A review,” IEEE Access, vol. 10, pp. 58821–58844, 2022.
[9] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” CoRR, vol. abs/1409.1556, 2015.
[10] Mariam A Sheha and Mai S Mabrouk et al., “Automatic detection of melanoma skin cancer using texture analysis,” International Journal of Computer Applications, vol. 42, no. 20, pp. 22–26, 2012.
[11] M. Jafari and S. Samavi et al., “Automatic detection of melanoma using broad extraction of features from digital images,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2016, pp. 1357–1360.
[12] S. Morales, K. Engan, and V. Naranjo, “Artificial intelligence in computational pathology– challenges and future directions,” Digital Signal Processing, vol. 119, pp. 103196, 2021.
[13] K. Rajitha and S. Bhat et al., “Classification of microscopic images of unstained skin samples using deep learning approach,” in 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). IEEE, 2022, pp. 1–4.
[14] F. Logu and F. Ugolini et al., “Recognition of cutaneous melanoma on digitized histopathological slides via artificial intelligence algorithm,” Frontiers in oncology, vol. 10, pp. 1559, 2020.
[15] L. Wang and L. Ding et al., “Automated identification of malignancy in whole-slide pathological images: identification of eyelid malignant melanoma in gigapixel pathological slides using deep learning,” British Journal of Ophthalmology, vol. 104, no. 3, pp. 318–323, 2020.
[16] R. Wetteland, K. Engan, and T. Eftesol, “Parameterized extraction of tiles in multilevel gigapixel images,” in 2021 12th International Symposium on Image and Signal Processing and Analysis (ISPA). IEEE, 2021, pp. 78–83
[17] [17] R.Chauhan and B. Thankachan “A Survey: Malanoma Skin Cancer Detection Using Deep Learning Techniques”, in 2022 Design Engineering Journal.