MULTI-COLOUR SPACE CHANNEL SELECTION FOR IMPROVED CHRONIC WOUND SEGMENTATION

Christian McBride[1⋆] *Bill Cassidy*[1⋆] *Connah Kendrick*[1] *Neil D. Reeves*[2]
Joseph M. Pappachan[3] *Moi Hoon Yap*[1]

[1] Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, UK
[2] Department of Life Sciences, Manchester Metropolitan University, Manchester, UK
[3] Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

ABSTRACT

This study introduces a novel approach to chronic wound segmentation, a critical aspect of automated wound monitoring that has the potential to significantly reduce clinical workload. Addressing the challenges posed by varying wound sizes and compositions, our experiments utilise the U-Net architecture with an innovative integration of multiple colour spaces - RGB, HSV, CIE-LAB, and YCbCr. Our method involves the merging of various combinations of colour channels from these selected colour spaces. We trained and evaluated our method on the Diabetic Foot Ulcer Challenge 2022 dataset, with improved Intersection-Over-Union (+0.0187), and Dice Similarity Coefficient (+0.0183), in comparison with the baseline model. Additionally, improvements are observed on alternative test sets that include; Complex Wound DB, Advancing the Zenith of Healthcare, and Foot Ulcer Segmentation Challenge datasets. These findings highlight the importance of strategic colour channel selection in chronic wound analysis, and offer a promising direction for future research in medical image analysis. These enhancements show our method's effectiveness in capturing complex wound characteristics using colour channel selection, contributing a new research direction for medical image analysis.

Index Terms— wound segmentation, diabetic foot ulcer, deep learning, chronic wounds, colour space.

1. INTRODUCTION

Diabetic Foot Ulcers (DFUs) are a common and severe complication of diabetes mellitus, affecting approximately 6.3% of people globally [1]. These persistent wounds contribute to a significant healthcare burden [2, 3] and drastically diminish the quality of life for those affected [4, 5, 6], making timely and accurate diagnosis essential for effective treatment. Recent advances in medical imaging analysis, particularly through deep learning, have shown promise in automating the segmentation of DFUs, which is an important component in the automated analysis of these types of chronic wounds. Segmentation allows for the monitoring of wounds, including the changes in shape and size over time, which may provide important indicators of wound healing status [7]. This is a complex task due to the highly variable appearance of the wounds, influenced by factors such as lighting, skin tones, and distinct characteristics of the wounds themselves, including size, shape, or the presence of infection [8]. The conventional RGB colour model, which has been a staple in medical imaging, is often limited in capturing the complexities necessary for accurate wound analysis. This limitation has led researchers to investigate alternative colour spaces that could potentially offer more detailed information or additional insights pertinent to chronic wound assessment.

In this study, we introduce an innovative method that integrates alternative colour space data into the models input tensor to improve chronic wound segmentation accuracy. This paper presents our approach, the results of its application, and its potential impact on the field of medical image analysis. All source code used in the experiments described in this paper can be found at the following repository: to be added upon acceptance of paper.

⋆ equal contribution

2. RELATED WORK

Colour space augmentation has become a growing interest in recent deep learning research. Its application is promising in DFU segmentation to enhance deep learning model accuracy. While RGB has been the conventional colour space of choice, alternatives such as HSV, CIE-LAB, and YCbCr offer additional benefits for improved handling of varying lighting conditions and colour consistency [9]. However, these studies proved inconclusive due to the use of very small datasets in training and testing.

Prior research works have refined training workflows to incorporate additional colour representations and have shown performance improvements in segmentation tasks [10]. Expanding upon colour space channels within deep learning frameworks is an emerging research area. Ramadan and Aly

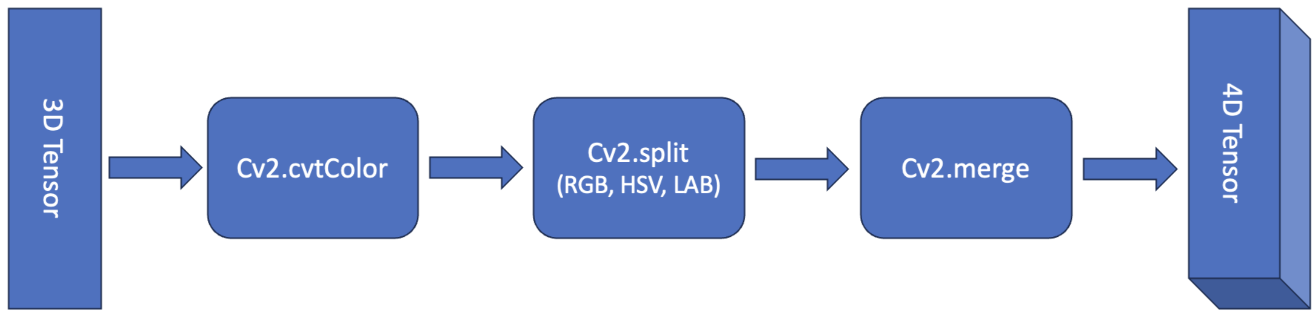

Fig. 1. Multi-colour integration strategy using OpenCV library functions (Cv2) to convert a 3-Dimensional image tensor to a 4-Dimensional image tensor. Process includes converting image colour space, extracting individual colour channels, and merging into selected number of dimensions.

(2022) [11] explored the use of a modified U-Net architecture [12] with dual and tri-colour space inputs, leading to performance increases in skin lesion segmentation. However, these methods treat colour spaces independently in the input stages of the model which incurs significant additional computational requirements.

The potential of alternate colour spaces for improved chronic wound segmentation accuracy is largely unexplored in the current literature. This research gap persists despite the promise of richer information and enhanced representation provided by these colour spaces. Our hypothesis explores the integration of additional colour channels from different colour spaces into single image tensors for enhanced segmentation performance.

3. METHODOLOGY

In this section, we define the methodology for the exploration of multi-colour space channel selection for enhancing chronic wound segmentation. We outline the preprocessing procedure that includes image normalisation, followed by the expansion of the image tensor to integrate additional colour channel information.

3.1. Dataset Description

We use 4 publicly available chronic wound segmentation datasets for our experiments: Diabetic Foot Ulcer Challenge (2022) [13], Complex Wound DB (2022) [14], Advancing the Zenith of Healthcare (2020) [15], and Foot Ulcer Segmentation Challenge (2021) [16].

The Diabetic Foot Ulcer Challenge 2022 (DFUC2022) dataset is used for both training and testing in our experiments. Analysis of the dataset reveals its composition of 4000 DFU images and corresponding binary masks, divided into training (1600), validation (400), and testing (2000) subsets. All images and corresponding masks are 640×480 pixels. The DFUC2022 dataset is the largest publicly available chronic wound segmentation dataset to date. The Complex Wound DB (CWDB) dataset comprises 27 chronic wound photographs (DFU and pressure ulcers) captured at various resolutions and includes corresponding binary masks. We use this dataset in its entirety as an exclusive test set. The Advancing the Zenith of Healthcare (AZH) dataset comprises 1109 DFU photographs at a resolution of 224×224 pixels, with 831 images in the training set and 278 images in the testing set with corresponding binary masks. We use only the testing set from this dataset in our experiments. The Foot Ulcer Segmentation Challenge (FUSC) dataset comprises 1210 DFU images, with 810 images in the training set, 200 images in the validation set, and 200 images in the testing set. All images in this dataset are 512×512 pixels. The testing set masks are not publicly available for this dataset, therefore we used the 200 validation images and corresponding masks in our experiments as an exclusive test set. All images and masks from the CWDB, AZH, and FUSC datasets are resized to match the resolution of the DFUC2022 images (640×480 pixels). The aspect ratio of the images is maintained, with black pixel padding applied where appropriate.

Data Preprocessing: To ensure correct formats for the training, validation, and testing phases of the proceeding experiments, images are preprocessed as follows:

Multi-Colour Space Tensor Merging: In line with the core purpose of this study, additional colour channels are integrated into the preprocessing pipeline (Section 3.2). This integration occurs after image loading and before image normalisation.

Image Normalisation: Pixel values within the images are normalised from the conventional range of 0-255 to a normalised range between 0 and 1. This normalisation is pivotal for the convergence and stability of deep learning models.

3.2. Model Implementation

U-Net Model: The U-Net architecture builds the foundation of our chronic wound segmentation model, selected for its robust performance in medical image segmentation [17, 18].

The implementation of this model is adapted for our specific task of chronic wound segmentation.

For our experiments, we use a standard 4-level U-Net with skip connections at each level, with each level incorporating two 3×3 convolutional layers, followed by batch normalisation and a ReLU activation function. The final output mask is produced using the sigmoid activation function for binary segmentation.

Multi-Colour Space Integration Pipeline: The colour channel integration process begins by loading the input image in its standard RGB format. We then apply the necessary colour space conversion and extract individual colour channels. These selected channels are concatenated to the original RGB channels to form a 4-6 dimensional input tensor in a process we call "tensor merging", as illustrated in Figure 1, which is then fed to the U-Net model. This higher-dimensional tensor accommodates a richer representation of image colour features.

4. EXPERIMENTAL DESIGN

Prior to completing our investigation, we determine a suitable base U-Net model which will be used in the following segmentation experiments. Our model is evaluated using a challenging testing set that comprises of 50% of the total images, ensuring performance validation and generalisability assessment.

Experimental Setup: All of our experiments were completed using the following hardware and software configuration: Intel® Core™ i7-8700 CPU @ 3.20GHz CPU, NVIDIA Quadro P5000 16GB GPU, Python (version 3.9.16), PyTorch (version 1.12.1) and CUDA (version 11.6).

U-Net Training Settings: The base U-Net training settings are summarised in Table 1. These settings were used for all experiments described in this paper. Other than the methods detailed in the augmentation experiments, no other preprocessing was used during training, except image normalisation on the training and validation sets.

Parameter	Value
Batch Size	8
Max Epochs	1000
Early Stopping	20 Epochs
Initial LR	0.001
Scheduler	Reduce on Plateau
Scheduler Factor	0.1
Scheduler Patience	6
Loss Function	Dice Loss
Evaluation	Batch IOU

Table 1. U-Net base model training parameters.

Determining the Baseline Model: In pursuit of establishing a comparable baseline for the DFU segmentation task, the U-Net model was initially trained using a standard RGB colour space. We trialed various batch sizes up to the limitations of the hardware: 2, 4, 8 and 16. The baseline model achieved its best performance at a batch size of 8, with the following validation metrics: Intersection over Union (IoU) of 0.5950 and Dice Similarity Coefficient (Dice) of 0.7341.

5. RESULTS AND DISCUSSION

Table 2 shows the top-4 best performing models in terms of training and validation metrics for the colour space experiments. The best performing model from these experiments, in terms of validation IoU and validation Dice was the RGB+Y model, with an IoU of 0.5966 and a Dice of 0.7354.

Table 2. Experiment results of multi-colour space tensor merging with training and validation metrics on the DFUC2022 dataset. Results shown are the top-4 best performing models. Ep - epoch, T - train, V - validation.

Colour Space	Ep	T-Loss	T-IoU	T-Dice	V-Loss	V-IoU	V-Dice
RGB	45	0.2371	0.7067	0.8193	0.4034	0.5950	0.7341
RGB + Y	48	0.2272	0.7164	0.8260	**0.3987**	**0.5966**	**0.7354**
RGB + A	34	0.3024	0.6501	0.7782	0.4016	0.5827	0.7239
HSV + LAB	44	0.2382	0.7150	0.8247	0.4017	0.5957	0.7331

Table 3 shows the test results of the top-4 best performing models on 4 different chronic wound test sets. For DFUC2022 and CWDB, the HSV+LAB model demonstrates the best performance in IoU (DFUC2022 = 0.4207, CWDB = 0.6698), and Dice (DFUC2022 = 0.5364, CWDB = 7791). The test results for the DFUC2022 test set indicate notable improvements[1] in IoU (+0.0187), Dice (+0.0183), FNE (-0.0206), and FPE (-0.0497).

The RGB+A model shows the best performance on the AZH dataset, with improvements in IoU (+0.0025), Dice (+0.005), FNE (-0.0236), and FPE (-0.0087). In contrast, the RGB+Y model shows best performance on the FUSC test set, with IoU (+0.0391), Dice (+0.0431), FNE (-0.0187), and FPE (-0.0642). These results highlight the importance of selecting the best colour channels for integration based on the specific characteristics of wounds in the segmentation task.

Table 4 provides detailed analysis of the performance of the models on wounds of different sizes. For small wounds (200 test samples, $< 10\%$ of wounds), the results are variable by metric. For IoU and Dice, the HSV+LAB model is the best performing model, in terms of IoU (+0.0041), and Dice (+0.0087). RGB+Y demonstrated the best FNE (-0.0118), and RGB+A the best FPE (-0.0433). For medium wounds (between 10% and 90% of wounds), HSV+LAB again demonstrated the best performance in terms of IoU (+0.0217), closely followed by RGB+Y (+0.0209), which has the highest Dice (+0.0212), and FNE (-0.0317). In the large

[1] All improvements in the form (± 0.0000) are in comparison with corresponding RGB model for the test results unless otherwise specified.

Table 3. Test results for the top-4 best performing models using four chronic wound test sets.

Dataset	Colour Space	FNE	FPE	IoU	Dice
DFUC2022	RGB	0.4313	0.3947	0.4020	0.5181
	RGB + Y	**0.4038**	0.3912	0.4176	0.5347
	RGB + A	0.4481	**0.3418**	0.4121	0.5364
	HSV + LAB	0.4107	0.3838	**0.4207**	**0.5364**
CWDB	RGB	0.2317	0.1871	0.6224	0.7376
	RGB + Y	**0.1959**	0.1813	0.6550	0.7730
	RGB + A	0.2313	0.1632	0.6362	0.7537
	HSV + LAB	0.2168	**0.1374**	**0.6698**	**0.7791**
AZH	RGB	0.2787	0.4243	0.5208	0.6048
	RGB + Y	0.2613	0.4264	0.5120	0.5958
	RGB + A	**0.2551**	**0.4156**	**0.5231**	**0.6098**
	HSV + LAB	0.3069	0.4559	0.5003	0.5802
FUSC	RGB	0.4027	0.2571	0.4855	0.5948
	RGB + Y	**0.3840**	**0.1929**	**0.5246**	**0.6379**
	RGB + A	0.4654	0.2201	0.4630	0.5742
	HSV + LAB	0.4008	0.2563	0.4982	0.6031

wound category, RGB+Y shows significant performance increases in all test metrics - IoU (+0.0391), Dice (+0.0431), FNE (-0.0187), and FPE (-0.0642).

Table 4. Test results for the multi-colour tensor merging on DFUC2022 test data with wound size splits.

Wound Size	Colour Space	FNE	FPE	IoU	Dice
Small	RGB	0.7878	0.4213	0.1956	0.2628
<10%	RGB + Y	**0.7752**	0.4264	0.1954	0.2659
	RGB + A	0.8066	**0.3780**	0.1730	0.2447
	HSV + LAB	0.7764	0.4205	**0.1997**	**0.2715**
Medium	RGB	0.4079	0.3212	0.4464	0.5633
10-90%	RGB + Y	**0.3762**	0.3170	0.4672	0.5845
	RGB + A	0.4260	**0.2592**	0.4566	0.5787
	HSV + LAB	0.3821	0.3119	**0.4681**	**0.5833**
Large	RGB	0.1108	0.3953	0.5569	0.6744
>90%	RGB + Y	**0.0970**	0.4017	0.5548	0.6694
	RGB + A	0.1133	**0.3324**	**0.6058**	**0.7196**
	HSV + LAB	0.1011	0.3593	0.5891	0.7035

These results show that careful selection of individual colour channels from diverse colour spaces can result in improved model performance in comparison with training on an individual colour space. In medical imaging of chronic wounds, characteristics of wounds can vary significantly; we show that by selection of colour channels such as ‘A’ chromaticity, significant increases are observed in larger wound types, whereas when training models with HSV and CIE-LAB colour spaces, improved predictions are observed with smaller wound types, emphasising the need for colour space and colour channel consideration in a dedicated segmentation task. As well as demonstrable performance increases, there is also the benefit of not having to feed as many images into the model during training. Additionally, the results from alternative test sets demonstrate that models trained only on DFU wounds are able to generalise to other wound types (pressure ulcers).

(a) (b)

Fig. 2. Example of a small, partially visible DFU wound prediction made by the HSV+LAB model for a case where the RGB model failed to generate a prediction. Image (a) shows the original DFU wound photograph, (b) shows the prediction mask from the HSV+LAB model. Example shown is from the DFUC2022 test set.

6. CONCLUSIONS

Our innovative method merges multiple colour channels from HSV, CIE-LAB, YCbCr, and RGB into a single tensor when training a U-Net model, addressing challenges posed by chronic wound variation. This approach offers an alternative to prior works [11], enhancing segmentation performance while reducing computational load. Our results on diverse testing sets display significant improvements in segmentation accuracy, marking notable improvements in terms of IoU, Dice, FPE, and FNE. This indicates the importance of colour channel selection in medical image analysis.

Given that the encoder for the U-Net model represents a classifier, we suggest our findings could extend to classification networks, and encourage further research in such methodologies, as well as investigation into performance in alternative model architectures, and application across other medical imaging domains. Our approach offers a step forward in developing more precise and efficient automated diagnostic tools, suggesting the potential for larger influence in medical imaging.

7. COMPLIANCE WITH ETHICAL STANDARDS

Publicly available datasets were used for this study.

8. CONFLICT OF INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

9. REFERENCES

[1] Pengzi Zhang, Jing Lu, Yali Jing, Sunyinyan Tang, Dalong Zhu, and Yan Bi, "Global epidemiology of diabetic foot ulceration: a systematic review and meta-analysis," *Annals of medicine*, vol. 49, no. 2, pp. 106–116, 2017.

[2] Brian Petersen, Walter Linde-Zwirble, Tze-Woei Tan, Gary Rothenberg, Simon Salgado, Jonathan Bloom, and David Armstrong, "Higher rates of all-cause mortality and resource utilization during episodes-of-care for diabetic foot ulceration," *Diabetes Research and Clinical Practice*, 01 2022.

[3] Elof Eriksson, Paul Liu, Gregory Schultz, Manuela Martins-Green, Rica Tanaka, Dot Weir, Lisa Gould, David Armstrong, Gary Gibbons, Randy Wolcott, Oluyinka Olutoye, Robert Kirsner, and Geoffrey Gurtner, "Chronic wounds: Treatment consensus," *Wound Repair and Regeneration*, vol. 30, 02 2022.

[4] Regina Renner and Cornelia Erfurt-Berge, "Depression and quality of life in patients with chronic wounds: ways to measure their influence and their effect on daily life," *Chronic Wound Care Management and Research*, vol. 4, pp. 143–151, 2017.

[5] Xiao-fen Xiong, Ling Wei, Ying Xiao, Ya-Chun Han, Jinfei Yang, Hao Zhao, Ming Yang, and Lin Sun, "Family history of diabetes is associated with diabetic foot complications in type 2 diabetes," *Scientific Reports*, vol. 10, pp. 17056, 10 2020.

[6] Maria Polikandrioti, Georgios Vasilopoulos, Ioannis Koutelekos, Georgios Panoutsopoulos, Georgia Gerogianni, Victoria Alikari, Evaggelos Dousis, and Afroditi Zartaloudi, "Depression in diabetic foot ulcer: Associated factors and the impact of perceived social support and anxiety on depression," *International Wound Journal*, vol. 17, no. 4, pp. 900–909, 2020.

[7] J.M. Pappachan, B. Cassidy, C.J. Fernandez, V. Chandrabalan, and M.H. Yap, "The role of artificial intelligence technology in the care of diabetic foot ulcers: the past, the present, and the future," *World Journal of Diabetes*, vol. 13, pp. 1131–1139, 2022.

[8] Moi Hoon Yap, Bill Cassidy, Joseph M. Pappachan, Claire O'Shea, David Gillespie, and Neil D. Reeves, "Analysis towards classification of infection and ischaemia of diabetic foot ulcers," in *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, 2021, pp. 1–4.

[9] V. Godeiro, J. Neto, B. Carvalho, B. Santana, J. Ferraz, and R. Gama, "Chronic wound tissue classification using convolutional networks and color space reduction," in *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, 2018, pp. 1–6.

[10] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, "Data augmentation for skin lesion analysis," in *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. CARE CLIP OR 2.0 ISIC 2018*, 2018, vol. 11041, pp. 303–311.

[11] Rania Ramadan and Saleh Aly, "Cu-net: A new improved multi-input color u-net model for skin lesion semantic segmentation," *IEEE Access*, vol. 10, pp. 15539–15564, 2022.

[12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, Eds., Cham, 2015, pp. 234–241, Springer International Publishing.

[13] Connah Kendrick, Bill Cassidy, Joseph M. Pappachan, Claire O'Shea, Cornelious J. Fernandez, Elias Chacko, Koshy Jacob, Neil D. Reeves, and Moi Hoon Yap, "Translating clinical delineation of diabetic foot ulcers into machine interpretable segmentation," 2022.

[14] Talita A. Pereira, Regina C. Popim, Leandro A. Passos, Danillo R. Pereira, Clayton R. Pereira, and João P. Papa, "Complexwounddb: A database for automatic complex wound tissue categorization," in *2022 29th International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2022, vol. CFP2255E-ART, pp. 1–4.

[15] Chuanbo Wang, D. M. Anisuzzaman, Victor Williamson, Mrinal Kanti Dhar, Behrouz Rostami, Jeffrey Niezgoda, Sandeep Gopalakrishnan, and Zeyun Yu, "Fully automatic wound segmentation with deep convolutional neural networks," *Scientific Reports*, vol. 10, no. 1, pp. 1–9, 2020.

[16] Chuanbo Wang, Behrouz Rostami, Jeffrey Niezgoda, Sandeep Gopalakrishnan, and Zeyun Yu, "Foot ulcer segmentation challenge 2021," Mar. 2021.

[17] Vatsala Anand, Sheifali Gupta, Deepika Koundal, and Karamjeet Singh, "Fusion of u-net and cnn model for segmentation and classification of skin lesion from dermoscopy images," *Expert Systems with Applications*, vol. 213, pp. 119230, 2023.

[18] Connah Kendrick, Bill Cassidy, Neil Reeves, Joseph Pappachan, Claire O'Shea, Vishnu V Chandrabalan, and Moi Hoon Yap, "Diabetic foot ulcer grand challenge 2022 summary," 02 2023.