A Deep Multi-Modal Method for Patient Wound Healing Assessment

Subba Reddy Oota [1], Vijay Rowtula [1], Shahid Mohammed[1]
Jeffrey Galitz[1], Minghsun Liu[1], Manish Gupta[2]
[1] Woundtech Innovative Healthcare Solutions, [2] Microsoft AI Research, India
{soota, vrowtula, shmohammed, jgalitz, mliu}@woundtech.net, gmanish@microsoft.com

Abstract

Hospitalization of patients is one of the major factors for high wound care costs. Most patients do not acquire a wound which needs immediate hospitalization. However, due to factors such as delay in treatment, patient's non-compliance or existing co-morbid conditions, an injury can deteriorate and ultimately lead to patient hospitalization. In this paper, we propose a deep multi-modal method to predict the patient's risk of hospitalization. Our goal is to predict the risk confidently by collectively using the wound variables and wound images of the patient. Existing works in this domain have mainly focused on healing trajectories based on distinct wound types. We developed a transfer learning-based wound assessment solution, which can predict both wound variables from wound images and their healing trajectories, which is our primary contribution. We argue that the development of a novel model can help in early detection of the complexities in the wound, which might affect the healing process and also reduce the time spent by a clinician to diagnose the wound.

1 Introduction

In recent years, the outstanding performance of deep learning in computer vision-based tasks has resulted into its extensive application for medical image analysis. The applications include detection, diagnosis, and segmentation of pathology in cancer by Janowczyk & Madabhushi (2016), retinal by Gulshan et al. (2016), brain by Nie et al. (2016), and wound images by Goyal et al. (2018); García-Zapirain et al. (2018); Khalil et al. (2019); Elmogy et al. (2018). However, fewer practical applications were built that can help clinicians to detect wound types and anticipate the wound's healing trajectory.

In current clinical practice, the evaluation of different wound ulcers consists of wound specific tasks which help in early diagnosis of wounds. Keeping track of development and treatment procedure for each particular case depends on wound type. Moreover, many of the wound findings are documented via visual assessment by clinicians. Some of the challenges for the clinicians in terms of cost of treatment include (i) frequent assessments of a patient, (ii) entry of wound attributes in the database, and (iii) applying the right diagnosis. The clinician's assessment of the wound with its surrounding skin depends on various wound variables such as wound ulcer types, location, stage, margin, yellow slough, and granulation. We are motivated by the fact that deep learning methods have proven to perform well for object classification and recognition tasks with sufficiently large data. We tried to automatically assess various factors required for clinician's review, thereby providing semi-supervised wound variable classification.

In this paper, we attempt to detect the wound type and other factors from the wound image. We further investigated the applicability of predicted wound variable to predict patient's hospitalization risk. We also show that our models comprising of deep convolutional neural networks (CNNs) provide better

Medical Imaging Meets NeurIPS Workshop, 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

Figure 1: Manually labeled wound ulcer type images

performance than a human expert. Figure 1 shows a sample of wound ulcer type images labeled by clinicians.

2 Our Method

The essential aspects of wound assessment comprise of reliable and accurate wound documentation. To provide consistent wound assessments, we implemented a method comprising a mixture of deep neural networks, a simpler machine learning model, and clinicians' expertise to overcome inherent flaws in human wound observations. Given a wound image as input, the aim is to predict the patient's hospitalization risk. To achieve this, we first predict the "Ulcer Type", "Location", "Stage", "Joint Necrosis", and "Ligament Necrosis Exposed" attributes of a wound. In practice, it is a laborious task to train a convolution neural network model from scratch with randomly initialized weights. Moreover, a large number of samples are required to create accurate neural models, as too little data would result in model overfit. Here, we employ pre-trained CNN Xception architecture weights (Chollet, 2017) to fine-tune the model on the five heuristic wound variables. These five predictions are combined with other 16 clinician-filled variables like "BMI", "Tunneling", "Age", and "Gender", etc. (by visual observation) to train a multi-modal binary hospitalization risk (or heal/no-heal) LGBM (Light Gradient Boosted Machines) classifier (Ke et al., 2017).

2.1 Dataset and Experimental Results

Wound Dataset: We experiment with the patients' wound image dataset, which was accumulated from several years of patients' wound care. To the best of our knowledge, this is the only dataset providing all types of wound variables annotated and verified by clinicians. The dataset contains images corresponding to 20 ulcer types with 80% of the images from 5 types: "Diabetic Ulcer", "Pressure Ulcer", "Surgical Wound", "Trauma Wound", and "Venous Ulcer". The dataset details are displayed in Tables 1, 2, and 3. In this work, we mainly focused on the five wound variables to build the deep learning models. We used the data provided by clinicians on other wound heuristics to train the final heal/no-heal model jointly with outputs of deep models.

Ulcer Type	#Samples
Diabetic Ulcer	19773
Pressure Ulcer	47541
Surgical Wound	12238
Trauma Wound	13667
Venous Ulcer	32492

Table 1: Images by Ulcer Type

Wound Location	#Samples
Lower Leg	31775
Sacral	20501
Foot	12753
Heel	11226
Ankle	10375
GreatToe	4547

Table 2: Images by Location

Ulcer Stage	#Samples
Full Thickness	47849
Grade 2	20501
Stage-3	12753
Stage-4	11226
Unstageable	10375

Table 3: Images by Stage

Challenges: We faced the following challenges during the training process: (i) occlusion - wound blocked by either scale or doctor's hand when taking the picture, (ii) illumination - wound images have been captured from smartphone in different lighting conditions, (iii) imbalanced data - not all ulcers are having equal number of samples, (iv) similarity - images in one of the wound ulcer looks very similar to images of the other wound ulcers, and (v) deformation - same wound image appears in different forms.

Figure 2: Attention Heatmaps generated by CNN

Experimental Setup: We trained the heal/no heal model in two-steps. We first created five different CNN models created for the wound variables such as "Wound Ulcer Type", "Wound Location", "Wound Stage", "Joint Necrosis Exposed", and "Bone Necrosis Exposed" using our wound image dataset. We created single-task models on Wound Ulcer Type, and Location, whereas multi-task models on remaining factors (Wound Stage, Joint Necrosis Exposed, and Bone Necrosis Exposed) due to their dependency on Wound Ulcer Type. The dataset was split into 70% for training, 10% for validation, and 20% for testing. We perform a 5-fold cross-validation technique for each wound variable. In the second step, we create a heal/no heal model using the five features extracted from the first step and other features like "Wound Area", "Wound Volume", "BMI", "Patient's Age", etc. filled by a clinician. Due to space constraints, we only discuss details of the wound ulcer type prediction model and the heal/no-heal model.

Wound Ulcer Type Prediction To create the wound ulcer type model on wound image dataset, we used a pre-trained Xception architecture and fine-tuned the model. We used softmax based output classifier as the prediction layer with a size equal to the number of class labels. The dataset details are provided in Table 1, where we can observe the class imbalance problem even within the five classes. We applied the dataset augmentation pre-processing step to deal with the problem. For the CNN-Xception model, we set the number of epochs to 50, batch size to 32. We use Adadelta optimizer Zeiler (2012) with a learning rate of 0.001. We illustrate the performance of our model in Table 4. To interpret the performance of our model, we generated attention heatmaps on wound images using the last CNN layer of Xception, as shown in Figure 2. Although the images used for training have no masking and wound boundary drawn, Figure 2 shows that the model can accurately locate the salient information on the wound image.

Ulcer Type	Prec	Rec	F1-score
Diabetic Ulcer	0.79	0.84	0.81
Pressure Ulcer	0.87	0.89	0.88
Surgical Wound	0.76	0.65	0.70
Trauma Wound	0.65	0.56	0.61
Venous Ulcer	0.82	0.89	0.85

Table 4: Ulcer Type Results

Ulcer Location	Prec	Rec	F1-score
Lower Leg	0.88	0.90	0.89
Sacral	0.99	0.98	0.98
Foot	0.83	0.83	0.83
Heel	0.84	0.88	0.86
Ankle	0.73	0.77	0.75
GreatToe	0.67	0.55	0.61

Table 5: Wound Location Results

Heal/No-Heal Model The main objective of this paper is to create a heal/No-heal model, where we collect five of the wound variables extracted from wound images and other remaining variables filled by the clinician. To create a heal/no-heal model, we use a recent successful state-of-the-art LightGBM Ke et al. (2017) model to classify wound into "Risk of Hospitalization", or "Treatment Complete". We follow the survival model conditions Wang et al. (2019) where we cover the "Patient Demographic details", "Procedures", "Medications", "Laboratory/Diagnosis of Wound condition" along with deep model predictions to create the final model. Our model provides a precision of 0.68, recall of 0.91, and an F1 of 0.78 for the healing class. While for the no-heal class, we obtain a precision of 0.99, recall of 0.79, and an F1 of 0.88.

Class	Prec	Rec	F1-score
Hospitalization-Wound Related	0.68	0.91	0.78
Treatment Complete (In active)	0.99	0.79	0.88

Table 6: Heal/No-Heal Model Results

3 Conclusion

In this paper, we designed a method to identify the patient hospitalization risk using both wound images and wound attributes provided by clinician. Our results show that the method leads to high accuracy for the task.

References

François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, 2017.

Mohammed Elmogy, Begoña García-Zapirain, Connor Burns, Adel Elmaghraby, and Ayman Ei-Baz. Tissues classification for pressure ulcer images based on 3d convolutional neural network. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 3139–3143. IEEE, 2018.

Begoña García-Zapirain, Mohammed Elmogy, Ayman El-Baz, and Adel S Elmaghraby. Classification of pressure ulcer tissues with 3d convolutional neural network. *Medical & biological engineering & computing*, 56(12):2245–2258, 2018.

Manu Goyal, Neil D Reeves, Adrian K Davison, Satyan Rajbhandari, Jennifer Spragg, and Moi Hoon Yap. Dfunet: Convolutional neural networks for diabetic foot ulcer classification. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2018.

Varun Gulshan, Lily Peng, Marc Coram, Martin C Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, 316(22):2402–2410, 2016.

Andrew Janowczyk and Anant Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J. Path. informatics*, 7, 2016.

Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, pp. 3146–3154, 2017.

A. Khalil, M. Elmogy, M. Ghazal, C. Burns, and A. El-Baz. Chronic wound healing assessment system based on different features modalities and non-negative matrix factorization (nmf) feature reduction. *IEEE Access*, 7:80110–80121, 2019. ISSN 2169-3536. doi: 10.1109/ACCESS.2019.2923962.

Dong Nie, Han Zhang, Ehsan Adeli, Luyan Liu, and Dinggang Shen. 3d deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients. In *MICCAI*, pp. 212–220, 2016.

Ping Wang, Yan Li, and Chandan K Reddy. Machine learning for survival analysis: A survey. *ACM Computing Surveys (CSUR)*, 51(6):110, 2019.

Matthew D Zeiler. Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*, 2012.