Generating Hyperspectral Skin Cancer Imagery using Generative Adversarial Neural Network

Leevi Annala[1,*], Noora Neittaanmäki[2], John Paoli[3], Oscar Zaar[3] and Ilkka Pölönen[1]

***Abstract*— In this study we develop a proof of concept of using generative adversarial neural networks in hyperspectral skin cancer imagery production. Generative adversarial neural network is a neural network, where two neural networks compete. The generator tries to produce data that is similar to the measured data, and the discriminator tries to correctly classify the data as fake or real. This is a reinforcement learning model, where both models get reinforcement based on their performance. In the training of the discriminator we use data measured from skin cancer patients. The aim for the study is to develop a generator for augmenting hyperspectral skin cancer imagery.**

I. INTRODUCTION

In the northern countries, melanoma incidence is increasing and the increase is predicted to continue [1]. Furthermore the clinical accuracy of the skin lesion classification is poor, leading to costly surgery that could be avoided [2]. There is a need for reliable non-invasive methods for clinical use for skin lesion diagnostics.

Hyperspectral imaging is one such potential tool, and it has been previously used successfully in classification of skin lesions [3], [4]. While in both articles, the results were promising, both of the datasets used were relatively small. Furthermore, if the models would be used in marginally different situations, either a lengthy data-gathering process would have to take place or the existing data could be reused as is or through a data transformation/augmentation process. This transformation or augmentation could be achieved by mathematically modeling the interaction of the light and the matter as well as the structure of the matter, but the uncertainties are high and the computations complex, downgrading the usability.

Another way to transform or augment data is through a GAN [5]. The basic idea of the GAN is that the generator network and the discriminator network compete in a game. The discriminator networks task is to determine whether or not the data given to it represents the training data or not. The task of the generator network is to produce data similar to the training data to trick the discriminator. The only input the generator receives during this time is yes/no confirmations from the discriminator, while the discriminator is trained with the full training data set.

GANs have been previously used in the field of medical imaging [6]. The use cases include denoising [7], image reconstruction [8], artifact removal [9], superresolution [10], image synthesis [11], classification [12] and segmentation [13]. Despite producing promising results, the field of using GANs in the medical field is still young, and there is lots of room for experimenting [6].

In the more specific field of hyperspectral (medical) imaging, GANs have been used in two different ways [14], [15]. In [14], hyperspectral images of lung samples were automatically stained using trained conditional GAN (cGAN) [16]. In [15], Wasserstein distance was used as basis for modifying GAN and the resulting networks were verified using data sets from different fields of study. There are also various different approaches to using GANs in hyperspectral classification problems (for example [17], [18]).

The objective of this study is to provide the reader with a minimal working example of a proof-of-concept level data augmentation tool for hyperspectral skin cancer image production.

II. MATERIALS AND METHODS

A. Generative Adversarial Neural Network Architecture

The used GAN was one implementation of deep convolutional GAN (DCGAN) [19]. The used generator architecture (Table I) consisted of input layer and three transposed convolution blocks. The first two convolutional blocks had two 3-dimensional convolutional layers, the first of which took into account the spectral domain, and the second the spatial. The last block had one convolutional layer, that worked in all dimensions. Next in the blocks was batch normalisation [20]. The last two blocks utilized striding, which in the case of transposed convolution increases the output size which in the end will be $40 \times 40 \times 40$ pixels. The activation function used in the blocks was leakyReLU [21] and the loss function is cross entropy between a data cube full of ones and the generated data cube. Both the generator and discriminator used the Adam-optimizer [22], with a learning rate of 10^{-4}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-7}$.

The discriminator architecture (Table II) consisted of two blocks that contained a 3-dimensional convolution and dropout of 0.3 [23] and an output block that contained flattening and output layers. The activation function used in the convolutional layers was leakyReLU and the loss function was a sum of generator loss and cross-entropy between data cube full of ones and training data.

* Corresponding author

This research was partly funded by Academy of Finland (grant: 314519).

[1] Leevi Annala and Ilkka Pölönen are with the Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland. leevi.a.annala@jyu.fi

[2] Noora Neittaanmäki is with .

[3]John Paoli and Oscar Zaar are with the Department of Dermatology and Venereology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.

TABLE I

THE ARCHITECTURE OF THE USED GENERATOR NETWORK.

Layer	**Kernel/Pool Size, Strides and Activation**	**Filters/ Units**
Dense	LeakyReLU	100
Batch Normalization		
Conv3DTranspose	(5,1,1), (1,1,1), LeakyReLU	128
Batch Normalization		
Conv3DTranspose	(1,5,5), (1,1,1), LeakyReLU	128
Batch Normalization		
Conv3DTranspose	(5,1,1), (2,1,1), LeakyReLU	64
Batch Normalization		
Conv3DTranspose	(1,5,5), (1,2,2), LeakyReLU	64
Batch Normalization		
Conv3DTranspose	(5,5,5), (2,2,2), LeakyReLU	1
Optimiser:	Adam	
Loss:	Cross entropy	

TABLE II

THE ARCHITECTURE OF THE USED DISCRIMINATOR NETWORK.

Layer	**Kernel/Pool Size, Strides and Activation**	**Filters/ Units**
Conv3D	(5,5,5), (2,2,2), LeakyReLU	64
Dropout (0.3)		
Conv3D	(5,5,5), (2,2,2), LeakyReLU	128
Dropout (0.3)		
Flatten		
Dense		1
Optimiser:	Adam	
Loss:	Cross entropy	

B. Data Capturing and Preprosessing

Raw hyperspectral measurement data was captured at Sahlgrenska University Hospital during 2018 – 2019, using Revenio Prototype 2016 hyperspectral imager with spatial resolution of 1920×1200 pixels and spectral resolution of 120 wavelengths. The camera uses Fabry-Pérot interferometer technology [24]. All patients volunteered to participate in the study. The study protocol followed the Declaration of Helsinki and it was approved by the local ethics committee. The captured dataset consists of 316 hyperspectral images of skin. The type of lesions varies with main portion of the images being of melanoma malignum, dysplastic nevus, bening nevus or melanoma in situ (Table III). A portion of the images have no label.

TABLE III

THE DISTRIBUTION OF LABELS IN THE USED HYPERSPECTRAL DATASET.

Label	**Count**
Dysplastic nevus	81
Melanoma in situ	76
Melanoma malignum	65
Benign nevus	27
Others	58
Unlabeled	9

The captured raw measurement data cubes were then processed to radiance using a fpipy Python package [25] and downsampled to $40 \times 40 \times 40$ pixels data cubes by averaging. Dataset size is increased by mirroring images by its three spatial symmetry axes, increasing the amount of images eight-fold to 2528. This data was used in training the generator and discriminator networks to produce similar data and to classify the data as real or fake. Various selected wavelength bands from the training data can be seen in Figure 1 and a typical radiance spectrum from data can be seen in Figure 2.

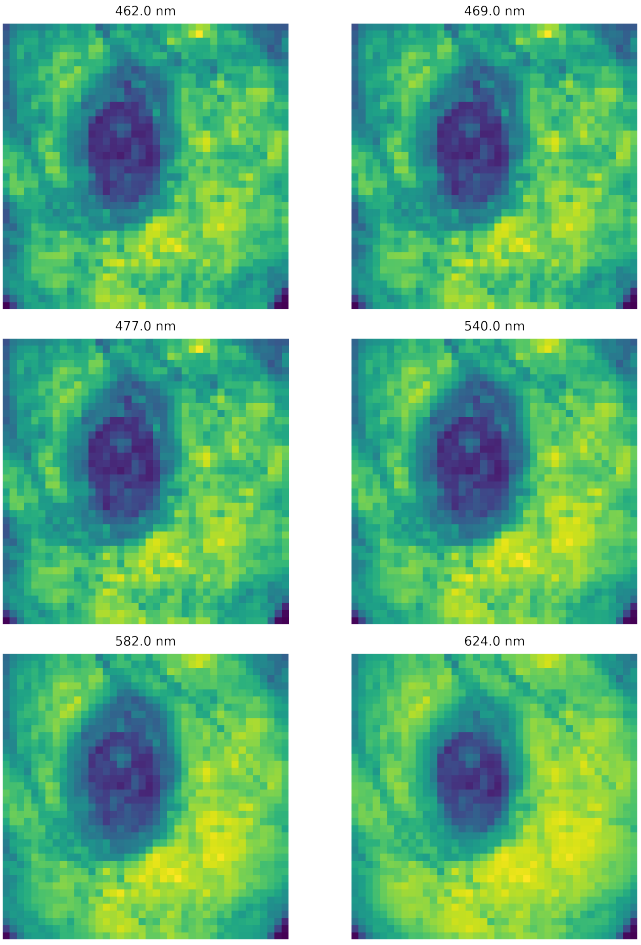

Fig. 1. Selected bands from training data hyperspectral image

III. RESULTS AND DISCUSSION

Generated hyperspectral images (Figure 3) shows positive and negative patterns. All bands, except the first band at 462 nm, look similar to the measured data in texture. They are obviously not random, and the patterns persist from band to band. However, the lesions in the generated data tend to have slight gradual shift, as can be witnessed in the Figure 3 right column. When one looks at the bottom right corner of each of the three images one sees that the lesion moves significantly to the right, in contrast to Figure 1, where the lesion becomes a little less clear, but does not move. The first band (Figure 3, 462 nm) shows some artefact.

Both of these problems likely have their roots in the convolution layers. There is no data before the first layer to convolute with. Therefore, the training affects it less. The shift of the lesions/features may have a root in the depth of the convolution. The convolution that addresses the layers 32 – 36 has little information on the layers 1 – 5.

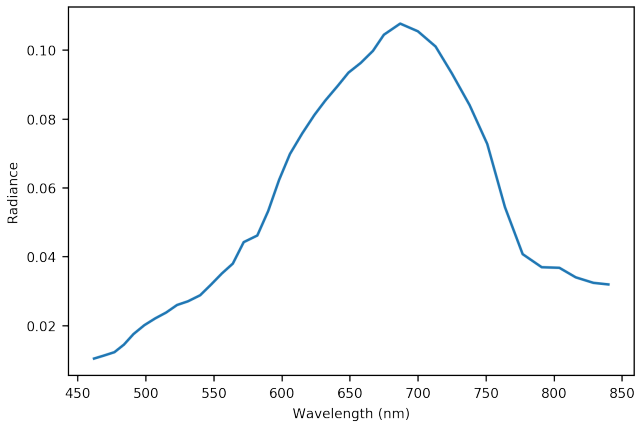

Fig. 2. Typical radiance in the training data

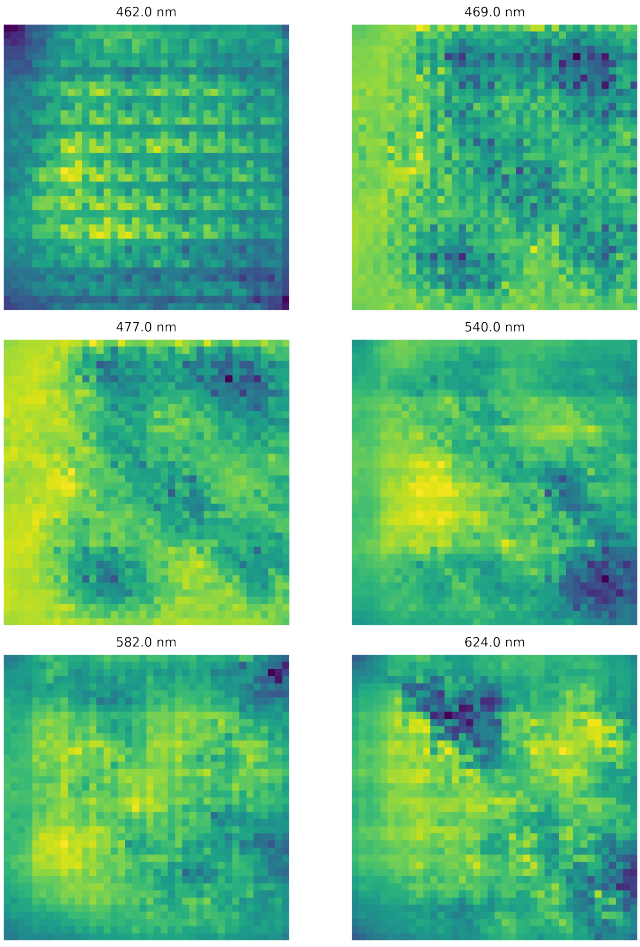

Fig. 3. Selected bands from a generated hyperspectral image

The typical radiance spectrum (Figure 4) from generated data looks similar to the typical radiance of the measured data (Figure 2), except for the artefact that causes the spectrum to fluctuate. This means that in the spectral domain the orders of magnitude are correct and the wave bands are correctly generated in that sense.

Further research of the use of GANs in the field of hyperspectral imaging are needed. The used DCGAN could be made deeper or wider, and different conditional GANs [16] could be experimented with. The future training should aim for using larger hyperspectral images to increase the accuracy of the training data. A U-net network could be tried [26]. A larger dataset should also be gathered to increase the likelihood of succesful training. In order to achieve this, one could try to experiment with a way to combine data from different studies and hyperspectral cameras in a robust manner to enlarge the size of the usable data. When the GAN produced data is judged accurate enough repression of measured data, it can be used for example in the training of other machine learning algorithms for tasks such as model inversion (such as in [27]) and hyperspectral image classification (example in the medical field [3]). There is also need to compare the method for more traditional data augmentation methods.

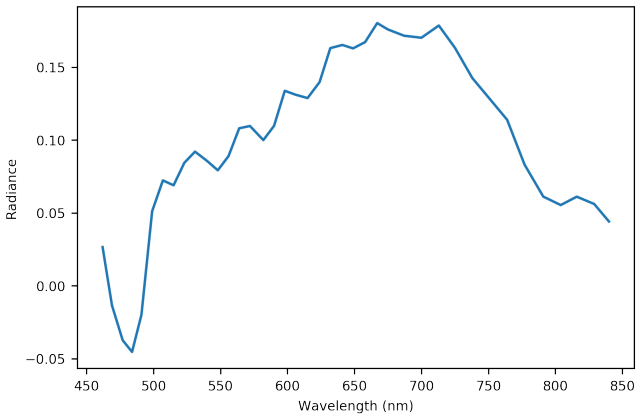

Fig. 4. Typical radiance in the generated data

IV. CONCLUSION

This study presented a proof of concept for producing augmented hyperspectral data using GAN. The used generator network was not fine-tuned to its full potential and in all likelihood there is much room for improvement in the network developement as well, as the first passable iteration was presented.

This study would have benefited from a larger hyperspectral dataset. In further research, such dataset should be gathered either by physically measuring or utilizing previously gathered different datasets in novel way.

An aim for further research could be to utilize data produced by GANs in further machine learning applications for skin cancer diagnostics by hyperspectral data to see how useful the generated data is. In further research there is a

call for using GAN-augmented data in machine learning applications to improve the quality of the algorithm training.

REFERENCES

[1] B. Møller, H. Weedon-Fekjær, T. Hakulinen, L. Tryggvadottir, H. Storm, M. Talbäck, and T. Haldorsen, "Prediction of cancer incidence in the Nordic countries up to the year 2020," *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP)*, vol. 11 Suppl 1, pp. S1–96, 2002.

[2] C. F. Heal, B. A. Raasch, P. Buettner, and D. Weedon, "Accuracy of clinical diagnosis of skin lesions," *British Journal of Dermatology*, vol. 159, no. 3, pp. 661–668, 2008.

[3] I. Pölönen, S. Rahkonen, L. Annala, and N. Neittaanmäki, "Convolutional neural networks in skin cancer detection using spatial and spectral domain," in *Photonics in Dermatology and Plastic Surgery 2019*, vol. 10851. International Society for Optics and Photonics, 2019, p. 108510B.

[4] U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9909–9924, 2018.

[5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *arXiv preprint arXiv:1406.2661*, 2014.

[6] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *Medical Image Analysis*, vol. 58, p. 101552, Dec. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1361841518308430

[7] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.

[8] S. U. H. Dar, M. Yurt, M. Shahdloo, M. E. Ildız, and T. Çukur, "Synergistic reconstruction and synthesis via generative adversarial networks for accelerated multi-contrast mri," *arXiv preprint arXiv:1805.10704*, 2018.

[9] J. Wang, Y. Zhao, J. H. Noble, and B. M. Dawant, "Conditional generative adversarial networks for metal artifact reduction in ct images of the ear," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 3–11.

[10] C. You, G. Li, Y. Zhang, X. Zhang, H. Shan, M. Li, S. Ju, Z. Zhao, Z. Zhang, W. Cong, and others, "CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE)," *IEEE Transactions on Medical Imaging*, vol. 39, no. 1, pp. 188–203, 2019.

[11] M. J. Chuquicusma, S. Hussein, J. Burt, and U. Bagci, "How to fool radiologists with generative adversarial networks? a visual turing test for lung cancer diagnosis," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 240–244.

[12] X. Yi, E. Walia, and P. Babyn, "Unsupervised and semi-supervised learning with Categorical Generative Adversarial Networks assisted by Wasserstein distance for dermoscopy image Classification," 2018.

[13] Z. Zhang, L. Yang, and Y. Zheng, "Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern Recognition*, 2018, pp. 9242–9251.

[14] N. Bayramoglu, M. Kaakinen, L. Eklund, and J. Heikkila, "Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 64–71.

[15] M. Zhang, M. Gong, Y. Mao, J. Li, and Y. Wu, "Unsupervised Feature Extraction in Hyperspectral Images Based on Wasserstein Generative Adversarial Network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2669–2688, May 2019.

[16] J. Gauthier, "Conditional generative adversarial nets for convolutional face generation," *Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition, Winter semester*, vol. 2014, no. 5, p. 2, 2014.

[17] H. Wang, C. Tao, J. Qi, H. Li, and Y. Tang, "Semi-Supervised Variational Generative Adversarial Networks for Hyperspectral Image Classification," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019, pp. 9792–9794.

[18] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Generative adversarial networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5046–5063, 2018.

[19] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," *arXiv:1511.06434 [cs]*, Jan. 2016, arXiv: 1511.06434. [Online]. Available: http://arxiv.org/abs/1511.06434

[20] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[21] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. icml*, vol. 30, 2013, p. 3.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[23] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014. [Online]. Available: http://jmlr.org/papers/v15/srivastava14a.html

[24] M. Blomberg, H. Kattelus, and A. Miranto, "Electrically tunable surface micromachined Fabry–Perot interferometer for visible light," *Sensors and Actuators A: Physical*, vol. 162, no. 2, pp. 184–188, 2010.

[25] M. Eskelinen, "Computational methods for hyperspectral imaging using Fabry–Perot interferometers and colour cameras," *JYU dissertations*, 2019.

[26] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[27] L. Annala, E. Honkavaara, S. Tuominen, and I. Pölönen, "Chlorophyll Concentration Retrieval by Training Convolutional Neural Network for Stochastic Model of Leaf Optical Properties (SLOP) Inversion," *Remote Sensing*, vol. 12, no. 2, p. 283, Jan. 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/2/283