

A Deep Learning Based Approach to Semantic Segmentation of Lung Tumour Areas in Gross Pathology Images

Matthew Gil^{1,2,4}[0000–0003–2108–3274], Craig Dick², Stephen Harrow³, Paul Murray¹[0000–0002–6980–9276], Gabriel Reines March^{1,2}[0000–0003–0761–0592], and Stephen Marshall¹[0000–0001–7079–5628]

¹ University of Strathclyde, Glasgow, UK

² NHS Greater Glasgow and Clyde, Glasgow, UK

³ NHS Lothian, Edinburgh, UK

⁴ matthew.gil@strath.ac.uk

Abstract. Gross pathology photography of surgically resected specimens is an often overlooked modality for the study of medical images that can provide and document useful information about a tumour before it is distorted by slicing. A method for the automatic segmentation of tumour areas in this modality could provide a useful tool for both pathologists and researchers. We propose the first deep learning based methodology for the automatic segmentation of tumour areas in gross pathological images of lung cancer specimens. The semantic segmentation models applied are Deeplabv3+ with both a MobileNet and Resnet50 backbone as well as UNet, all models were trained and tested with both a DICE and cross entropy loss function. Also included is a pre and post-processing pipeline for the input images and output segmentations respectively. The final model is formed of an ensemble of all the trained networks which produced a tumour pixel-wise accuracy of 69.7% (96.8% global accuracy) and tumour area IoU score of 0.616. This work on this novel application highlights the challenges with implementing a semantic segmentation model in this domain that have not been previously documented.

Keywords: NSCLC · Semantic Segmentation · Gross Pathology Photography.

1 Introduction

Pathology photography can be a useful tool for documenting ground truth anatomy before it has been distorted by the slicing processes that are used for whole slide imaging (WSI). Segmentation of regions in pathology photographs can therefore provide ground truth for the shape of an area, or volume if three dimensions are considered, of a particular anatomical region. Additionally, current pathological assessment of tumour size, which is a strong predictor of patient outcomes [14], is generally made by measuring the gross length of the tumour

across its largest dimension by hand with a ruler which often has to be reevaluated at the time of microscopic assessment [26]. Automatic segmentation of gross tumour area would provide a more reliable method of estimating the tumour volume and cellular load which are the metrics that are being estimated by gross measurements with a ruler. Additionally, if a method of automatic segmentation of singular tumours is successful it could then be expanded to be used to pick up other more subtle nodules that could be easily missed by the naked eye but may have been seen in radiology images and if used in real-time this would allow the pathologist to sample these nodules at the time of dissection. An automatic segmentation method for non-small cell lung cancer (NSCLC) tumours in gross pathology photographs, therefore, has both clinical and research applications. The work in this study aims to produce and test a methodology for the automatic semantic segmentation of lung tumours in pathology photographs of specimens that have been surgically removed from patients with NSCLC.

The procedures generally used to capture gross pathology photographs have been described in [17]. Best practices include placing the pathology specimens on a background that provides a good contrast between the specimen and background. The specimen should be well-lit with lighting located to the sides of the specimen as overhead lighting is more likely to cause reflections that may obscure anatomy. Excess moisture should also be removed from the surface of the specimen as this may obscure the underlying anatomy through the liquids opacity or the increased reflections this may cause. The specimen should also be well framed, in focus and the imaging plane should be the same as the slicing plane. The International Association for the Study of Lung Cancer (IASLC) recommends pathology photography as a standard part of pathology processing for NSCLC specimen processing after neoadjuvant therapy [27].

Gross pathology photography has been applied in some studies to provide the information necessary to transform WSI so that the geometry of the images more accurately represents what would have been observed in-vivo. This has often been for the application of registering images from the PET and pathology modalities [12, 16]. Gross pathology photography has been used as an important feature in many studies where regions are generally segmented by experienced pathologists. These studies include investigations into the mechanical properties of tissues [21], ablation treatment monitoring [29] and histologically diagnosed cardiac sarcoidosis [15]. A semi-automatic vector quantisation based pathology segmentation approach has been applied to segment regions of fibrosis in gross photographs to determine the overall prevalence of fibrotic tissue in lymph nodes [10]. Hyperspectral image based tumour segmentation has also been applied for application in real-time tissue classification during laparoscopic surgery [2].

An area where gross pathology photography has been applied more extensively than lung the lung cancer domain is skin lesion photography. There are similarities between lung lesion photography and skin lesion photography that make the greater catalogue of previous work on skin lesion segmentation relevant to this study. One such example is the work by Y. Yuan et al. who produced a fully connected (FC) convolutional neural network (CNN) based approach for

skin lesion segmentation with a Jaccard distance based loss function with their highest performing method consisting of an ensemble of six separate FC CNNs [31]. Q. Ha et. al detail their work on skin lesion segmentation that achieved 1st place in the 2020 SIIM-ISIC melanoma classification challenge [6]. Their method involved using an ensemble-based model that averages the average pixel prediction scores of models using various versions of EfficientNet, SE-ResNeXt and ResNeSt as the network backbone. Also included was a thorough image augmentation pipeline. Additionally, the ISIC skin lesion segmentation challenge [4] has run every year from 2016 until 2020 so there is a large back-catalogue of skin lesion segmentation methods all trained and tested on a standardised dataset. A detailed review of the skin lesion segmentation literature can be found in [7] which summarises 356 publications on skin lesion segmentation and 238 on skin lesion classification published between 2011 and 2022.

At the time of writing, the authors are not aware of any deep learning based semantic segmentation of gross pathology photographs of surgically removed tumours for any region of anatomy, this is supported by a google scholar search on 23/02/2023 using the key words "gross", "pathology", "deep", "learning" and "segmentation". Therefore the main contribution of this work is establishing the feasibility of this particular application of deep learning.

1.1 Loss Functions

The choice of loss function when training a deep learning based semantic segmentation model can have a large impact on the performance of the model. This is especially true for problems with unbalanced datasets where a model may greatly focus on increasing the accuracy of the class with the most instances causing the accuracy of segmentation of the underrepresented class to be low. For this particular study and more generally in many oncology based segmentation problems, the tumour class is underrepresented compared to the background class but would be considered a more important class to accurately segment. For this reason, the choice of loss function is important in this study.

Cross entropy is a distribution-based metric used for solving optimisation problems that has seen much use in deep learning applications [20]. Balanced cross-entropy (BCE) loss was introduced to improve performance when training models on unbalanced datasets that have more examples of a particular class than others by adding a weighting to the contribution each class to the loss function [30]. For binary classification problems, the balanced cross-entropy loss is expressed by equation 1 [25].

$$L_{BCE}(y, \hat{y}) = -\frac{1}{N} \sum_{n=1}^N (\beta y_n \log(p_n) + (1 - \beta)(1 - y_n) \log(1 - p_n)) \quad (1)$$

Where L_{BCE} is the balanced cross-entropy loss, N is the total number of individual pixels n . y_n is a ground truth pixel value, p_n is a predicted pixel probability outcome and β is a factor used to apply a weighting to the classes.

The DICE score was first proposed for use as a loss function for two class segmentation problems by Milletari et al. [13]. This was expanded to the generalised DICE loss function for multi-class classification problems by C. H. Sudre et al. [25]. As the problem presented in this study is a two-class classification problem, only the standard DICE loss function is required. This is described in equation 2 where L_{DICE} is the DICE loss [25].

$$L_{DICE}(y_n, p_n) = 1 - \frac{\sum_{n=1}^N 2y_n p_n}{\sum_{n=1}^N y_n + \sum_{n=1}^N p_n} \quad (2)$$

2 Methods

Unless mentioned otherwise, any computational steps applied in the following sections were applied using MATLAB ver R2022b (The MathWorks, Inc.).

2.1 Datasets

The data for this study comes from two separate datasets. Both datasets consist of photographs of lung specimens that have been surgically resected from patients with NSCLC. The specimens were all sliced as they would be for WSI and photographs were taken of the slices. Manual segmentations of the tumour boundary were produced by an expert pathologist using the software ImageJ [24].

The first dataset, referred to from now on as dataset-A, contains gross pathology photographs that were sliced and segmented using non-standard approaches. The lung specimens from this dataset consist of entire lung lobes that were inflated with agar in an attempt to create a better correlation between the shape of the ex-vivo specimens with their in-vivo shape. The lung specimens were then suspended in agar, so that the three-dimensional information was preserved through the slicing process, and sliced at 5mm intervals with photographs being taken after every slice was removed. During the collection of this dataset, care was taken in the lighting of the samples as well as partially drying the samples so that minimal reflection and maximum tissue contrast could be produced. An example of a photograph and its corresponding manual segmentation can be seen in figure 1. More information on the data gathering and pathology processing procedures for dataset-A can be found in [18].

The second dataset, which will be referred to as dataset-B, consists of pathology specimens that were sliced and photographed freehand with only the standard pathology lab lighting used to light the specimens. In this dataset, less care was taken to remove excessive moisture and reflections on the samples meaning these photographs are of poorer quality than dataset-A. The number of patients and images in each dataset is summarised in table 1.

Examples of gross pathology photographs from four separate patients are shown in figure 2. This shows the variability in the tumours and some of the different features that can be seen. For example figure 2 (b) shows a tumour

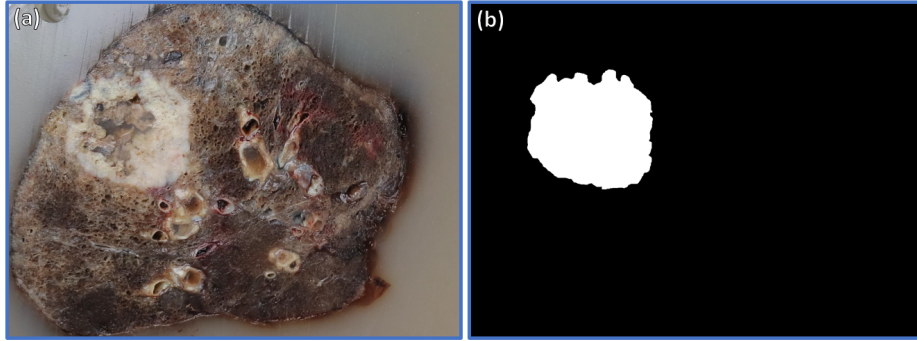


Fig. 1. An example of (a) a gross pathology photograph and (b) its corresponding manual tumour segmentation from dataset-A.

Table 1. Pathology photograph dataset information.

Dataset	Number of Patients	Number of Images
Dataset-A	9	64
Dataset-B	6	52
Total	15	116

with a large necrotic core whereas figure 2 (d) shows a tumour with no necrotic regions. Figures 2 (a),(b) and (d) all display regions with increased red or pink colour, this is due to the tumour reducing the quality of fixation in these regions. Figure 2 (c) shows an example where the tumour is displaying poorer contrast to the healthy tissue than the other examples, likely due to this example being an adenocarcinoma tumour.

2.2 Image Pre-processing and Data Augmentation

Before the pathology images and labels were used in training some pre-processing steps were applied.

Non-Tissue Background Removal Many of the images contain a large level of background area compared to the area of lung tissue. This was reduced by manually cropping the images down to a rectangular shape closely bounding the lung tissue. The aim of this step was to reduce the computational load of training the models by decreasing the image sizes and to get the CNN to focus more on areas of the lung specimen.

In all of the pathology photographs, the non-tissue background is well distinguished from the tissue regions of the image. This allows for the application of non-learning based segmentation techniques to create a mask that removes the background regions. As the pathology samples in dataset-A and dataset-B were

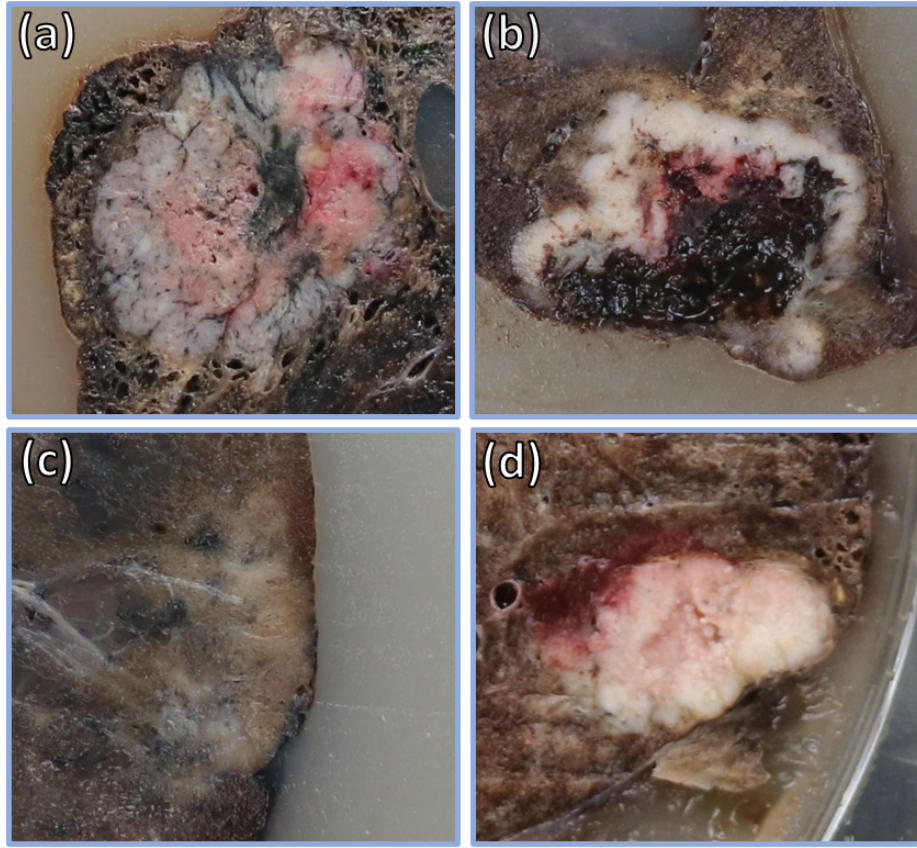


Fig. 2. Examples of tumour regions in gross photographs for four separate patients.

prepared using different methodologies, the background regions in both datasets are reasonably different. Dataset-A contains only regions of agar, these regions are of a similar colour to much of the tissue regions. The tissue has also been inflated with agar causing there to be regions of agar within the outer tissue boundary. This means a colour-based approach to background segmentation is not appropriate. Here we use spectral residual saliency detection approach as described in [11]. This was applied in MATLAB using methods adapted from [23].

Dataset-B contains samples either taken on a pathology slicing board or the perforated metal pathology workstation. There are also often separate objects such as rulers contained within the images. These images are less suited to the spectral residual approach used for dataset-A but work well with a colour-based approach due to the clear colour contrast between the tissue and non-tissue regions of the image. For dataset-B a k-means clustering approach [1] is applied

for colour segmentation where in this case $k = 2$. The parameters used in the k-means clustering algorithm are shown in table 2.

Table 2. K-means segmentation parameters

Parameter	Value
k	2
Number of clustering repetitions	3
Max Iterations	100
Accuracy Threshold	1.00e-04

Once a non-tissue background mask was produced it was used to set all the pixels of the non-tissue background to intensity values of zero.

Data Augmentation The images were then converted to patches of size 224×224 for use with the neural network. When converting the images to patches, an overlap of 50% was introduced in both the x and y image directions to conserve spatial information occurring at the borders of the patches [8]. This produces four times more patches than when the overlap is not included so it should be noted that the network is seeing the same data four separate times over one epoch of training on the patches.

A random rotation of the images between 0 and 360 degrees and a random zoom between 0.8 and 1.5 times was applied to the images after every epoch of training. The zoom here is important as the photograph height above the pathology samples is not standard so there is a variation in the zoom level within the datasets.

2.3 Segmentation Model

For the semantic segmentation task, an ensemble-based deep learning approach was applied. This involved training multiple separate deep-learning models and combining the output segmentations into a single averaged segmentation. This was followed by post-processing of the ensemble model output through background masking and morphological steps to improve the output segmentation. The full workflow of the final model is shown in figure 3 and described in the following sections.

Several different network architectures were applied to the problem. Deeplabv3+ [3] was applied with both a ResNet50 [9] and MobileNetv2 [22] backbone as well as UNet [19] with an encoder depth of 4. All of these networks were applied with a binary pixel classification output with the pixel classifications as "tumour" or "background". This means that both healthy lung tissue and the non-biological content (slicing table etc.) of the image were included in the background class.

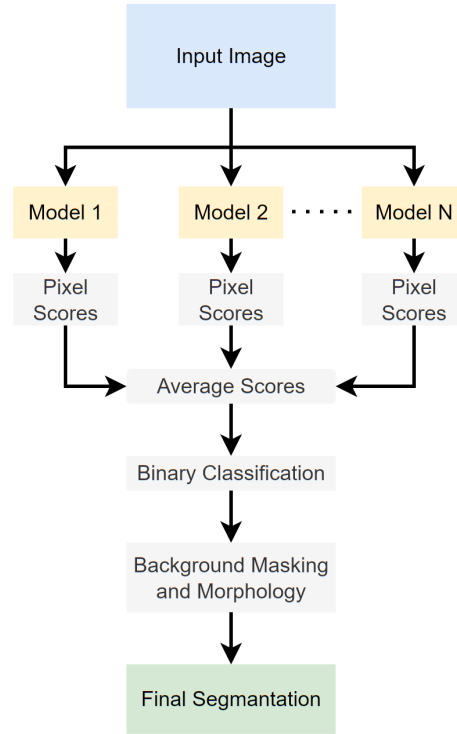


Fig. 3. The full ensemble model workflow.

All of the network architectures were trained with both weighted cross entropy and DICE loss functions. All of the models used pre-trained weights from ImageNet [5]. As the content of ImageNet is greatly different from the content of the datasets in this study, none of the layers of the models were frozen as is often done for transfer learning.

The dataset was split into training and test sets by patient so that images from one patient were only used for either training or testing. This is important as the different slices from the same patient contain similar features, such as the colour of the tumour and healthy tissue, that would bias the results if they were included in both training and test sets. Using this approach, the data was split into a training set containing 12 patients images and a test set containing 3 patients images. Using this approach causes the number of images in both the training and test sets to change depending on which patients images were used as there were more images available for some patients than others. This method generally created a split of around 96/20 images in the training/test sets. A 5-fold cross-validation approach was applied where all of the models were trained separately on different sets of 12 patients and tested on the three that were not included in the training set with the three test patients changing every fold. This

allows for the model to be tested on the full dataset. The training parameters are shown in table 3. All of the network training was performed on a single NVIDIA GeForce GTX 1080 Ti graphics card.

Table 3. Network training parameters for the individual models trained.

Model	Network Architecture	Loss Function	Initial LR	Drop Rate	Drop Factor	Epochs
1	DeepLabV3+ ResNet50	BCE	0.001	5	0.2	20
2	DeepLabV3+ ResNet50	DICE	0.001	5	0.2	20
3	DeepLabV3+ MobileNetv2	BCE	0.001	5	0.2	20
4	DeepLabV3+ MobileNetv2	DICE	0.001	5	0.2	20
5	Unet (Encoder Depth: 4)	BCE	0.001	5	0.2	20
6	Unet (Encoder Depth: 4)	DICE	0.001	5	0.2	20

In addition to applying the models individually, an ensemble-based approach was taken. To achieve this the individual pixel prediction probability outputs of each individual network in table 3 were simply averaged.

2.4 Image Post-Processing

The deep learning models often correctly segment the region of tumour in the input image but also labels some separate erroneous regions as tumour. These incorrect regions can usually be removed through some morphological operations that can be applied based on what is known about the task to improve the segmentation results. The morphology steps are detailed in the list below:

1. Small objects with a size of fewer than 5000 pixels are removed from the image (for reference, pixels are generally around 0.1x0.1mm).
2. A morphological closing operation is applied using a circular structuring element with a radius of 20 pixels.
3. Any holes in the remaining objects are filled.
4. The total number of pixels in each remaining object is calculated. Only the object made of the most pixels is kept as the final tumour segmentation.

Step 1 is applied to remove small isolated regions that were classified as tumour as these are almost always incorrect classifications, this step also improves the performance of all of the following steps. Steps 2 and 3 in the list above are required because many NSCLC tumours contain necrotic cores. These regions are pathologically and visually different from non-necrotic areas of tumour which, combined with the fact that there are few different patient examples in the datasets, causes them to be often misclassified as non-tumour. Simply closing and filling the tumour region generally fixes this problem. Step 4 can be applied as we know that the images in our dataset are from patients with one large NSCLC tumour.

3 Results

3.1 Segmentation Metrics

The results of the 5-fold cross-validation are shown in tables 4 and 5 for datasets A and B respectively. Generally for both datasets the ensemble model outperforms the individual models. Including the morphology steps improves the ensemble results across all of the metrics showing that for this particular application it is worthwhile to include them. The results on dataset-B are considerably lower than those in dataset-A. This is expected due to the lower quality of images in dataset-B.

Table 4. Results from testing on dataset-A.

Model	Global	Tumour	Tumour Background	
	Accuracy (%)	Accuracy (%)	IoU	IoU
1	93.9	69.3	0.486	0.905
2	89.9	58.7	0.332	0.854
3	91.9	57.4	0.396	0.876
4	88.1	52.3	0.262	0.832
5	90.3	56.6	0.366	0.884
6	87.2	52.3	0.296	0.826
Ensemble	95.7	63.8	0.521	0.923
Ensemble + Morphology	96.8	69.7	0.616	0.940

Table 5. Results from testing on dataset-B.

Model	Global	Tumour	Tumour Background	
	Accuracy (%)	Accuracy (%)	IoU	IoU
1	81.5	68.7	0.388	0.731
2	84.2	50.0	0.349	0.757
3	85.9	65.7	0.503	0.788
4	86.7	60.3	0.427	0.790
5	81.7	60.4	0.327	0.725
6	80.9	57.1	0.351	0.743
Ensemble	88.8	67.9	0.493	0.817
Ensemble + Morphology	89.2	68.6	0.504	0.822

3.2 Segmentation Examples

There is a large variety in the quality of the segmentation output of the ensemble model depending on the input images, some examples of this are shown in figure 4. Figure 4 (a.i) shows an example of a correct segmentation result on a

tumour with good contrast between the healthy tissue and tumour tissue. The tumour boundary in image (a.ii) aligns very closely with the ground truth mask producing an IoU of 0.956 for this image.

Figure 4 (b.i) shows an example of a partially correct segmentation where an area of necrosis has caused errors. This example has an IoU of 0.439 for the tumour class. The segmentation contour in this image outlines the region of lighter tissue which corresponds to the living tumour area. The necrotic area is not included in the segmentation output but is part of the ground truth tumour area as seen in image (b.ii). The model tends to misclassify necrotic regions as non-tumour as there are not many examples of heavily necrotic tumours in the training datasets and the coagulated blood that appears in this region also often appears in areas of healthy tissue. In other necrotic examples, this can be fixed by the post-processing morphology steps but in this case, as the living tumour area does not fully enclose the necrotic region, these steps do not solve this problem.

Figure 4 (c.i) shows an example of a failed segmentation with a tumour IoU of only 0.035. Upon analysing this image within the context of the dataset it is seen that image is from one of only two patients that had an adenocarcinoma tumour in the datasets. Adenocarcinoma has a lepidic pattern of growth causing it to be less contrasted against healthy tissue in gross images than other types of NSCLC. All other images from this patient and the other patient with adenocarcinoma have a similarly failed segmentation. It is clear from this that the dataset would need to be expanded to include more adenocarcinoma examples.

4 Discussion

The classification of the entire pathology of the lung into the two categories of tumour and non-tumour is an oversimplification that presents some problems for the segmentation model. This is most notable with adenocarcinoma tumours that are generally not recognised as tumours. Additionally, necrotic regions within the tumour are often misclassified as non-tumour regions. This can generally be fixed through the use of morphological image processing steps but it still highlights a problem with the ground truth data. This would be improved by increasing the dataset size to include more patients as the small dataset used in this study, with only 15 separate patients, included only a few examples of different pathological features such as adenocarcinomas and necrotic regions. A dataset containing a similar number of images that were all from unique patients would likely increase the performance of the trained models as this would allow the model to learn a more comprehensive array of pathological features. In the skin lesion photograph segmentation domain, large datasets such as the HAM10000 dataset [28], which contains 10000 images and ground truth segmentations of skin cancer lesions, allowing for highly accurate models to be produced. In addition to increasing the dataset size, it may be beneficial to increase the number of classes used for the segmentation to include different types of tissue though this would require

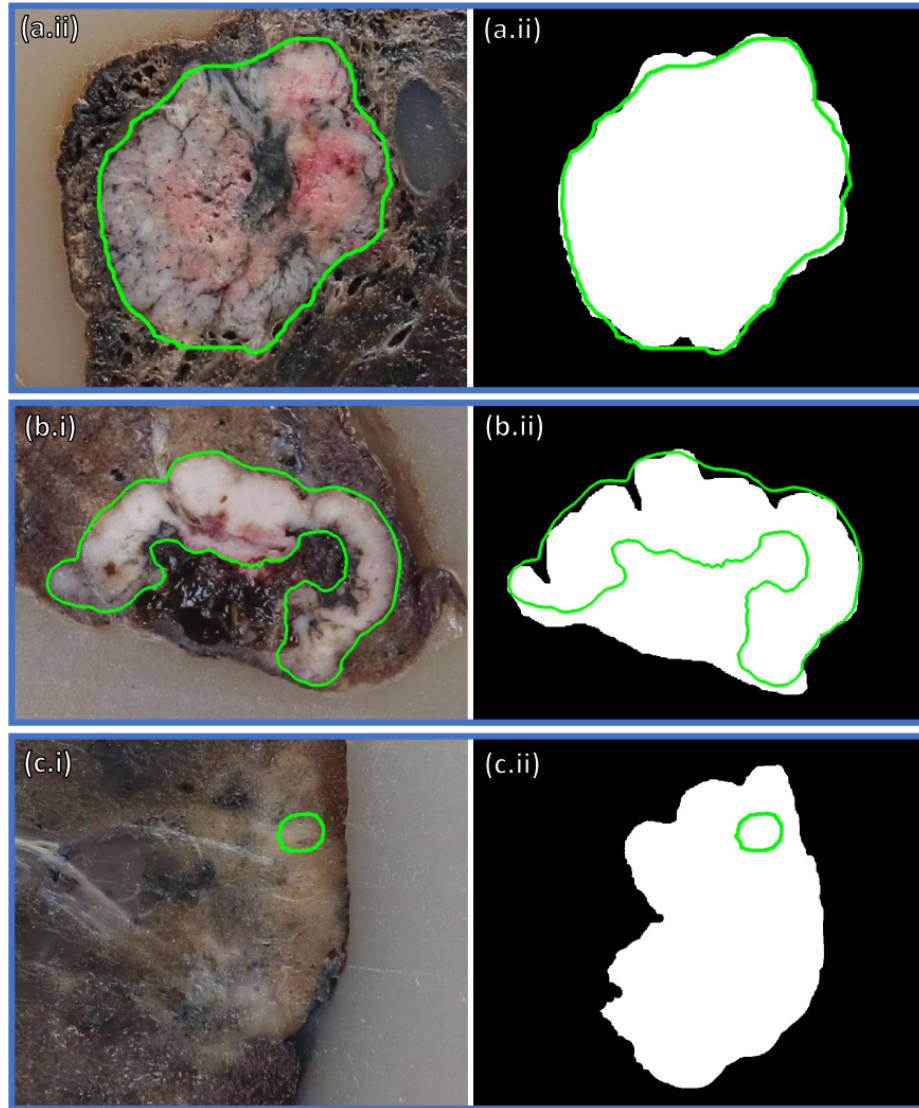


Fig. 4. Three segmentation examples from separate patients are displayed. **(a)** shows a good segmentation example, **(b)** shows a partially failed example due to a necrotic region and **(c)** shows a fully failed segmentation due to the tumour being an adenocarcinoma. Images denoted with **(i)** are the original test images zoomed in on the tumour area and images denoted with **(ii)** are the ground truth tumour segmentations. All images have the automatic segmentation contour overlaid (green line).

a time-intensive process of manual segmentation to produce the ground truth labels.

The final results for dataset-A produced better scoring metrics than those produced from dataset-B. This is unsurprising as, for the reasons described in section 2.1, the images in dataset-B are of poorer quality than those in dataset-A. This reduced image quality will increase the difficulty of segmentation first due to the image features being obscured and secondly due to there being fewer of these images of low quality in the overall training datasets. For further development and application of a system for the automatic segmentation of gross pathology photographs, care should be taken to ensure a high image quality by following the photography steps outlined in [17], though the inclusion of lower quality images in the training set may be beneficial to increase the robustness of the model.

For applications in clinical use, it may be beneficial to include some user input to produce a semi-automatic segmentation to decrease the chance of errors and improve overall accuracy. This could involve simply selecting the correct region from the output of the model to remove some of the morphology steps or marking some tumour or background pixels to be input to the network. The decision to choose a fully or semi-automatic approach would depend on the specific application and pathology workflow that the model is to be included in.

5 Conclusion

In conclusion, deep learning-based methods for semantic segmentation have been applied to the novel application of automatic segmentation of tumour areas in gross pathology photographs of specimens from patients with NSCLC. A pipeline for image preprocessing, model training and post-processing of the segmentation output has been detailed and validated. This work has proven the possibility of achieving this goal as well as highlighting some challenges for producing a fully robust system. The main barrier to improving the performance is a lack of data where the model produced in this work produced good results on more common tumour examples, tumours with less common features were poorly segmented.

Acknowledgements We would like to thank the Beatson Cancer Charity and UKRI EPSRC for funding this work as well as the CDT in Applied Photonics for facilitating this work.

References

1. Arthur, D., Vassilvitskii, S.: k-means++: The advantages of careful seeding. Tech. rep., Stanford (2006)
2. Baltussen, E.J., Kok, E.N., Brouwer de Koning, S.G., Sanders, J., Aalbers, A.G., Kok, N.F., Beets, G.L., Flohil, C.C., Bruin, S.C., Kuhlmann, K.F., et al.: Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery. *Journal of biomedical optics* **24**(1), 016002–016002 (2019)

3. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV). pp. 801–818 (2018)
4. Codella, N., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368 (2019)
5. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
6. Ha, Q., Liu, B., Liu, F.: Identifying melanoma images using efficientnet ensemble: Winning solution to the siim-isic melanoma classification challenge. arXiv preprint arXiv:2010.05351 (2020)
7. Hasan, M.K., Ahamad, M.A., Yap, C.H., Yang, G.: A survey, review, and future trends of skin lesion segmentation and classification. *Computers in Biology and Medicine* p. 106624 (2023)
8. Hashemi, S.R., Salehi, S.S.M., Erdogmus, D., Prabhu, S.P., Warfield, S.K., Gholipour, A.: Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: Application to multiple sclerosis lesion detection. *IEEE Access* **7**, 1721–1735 (2018)
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
10. Hipp, J.D., Cheng, J.Y., Toner, M., Tompkins, R.G., Balis, U.J.: Spatially invariant vector quantization: A pattern matching algorithm for multiple classes of image subject matter including pathology. *Journal of pathology informatics* **2**(1), 13 (2011)
11. Hou, X., Zhang, L.: Saliency detection: A spectral residual approach. In: 2007 IEEE Conference on computer vision and pattern recognition. pp. 1–8. Ieee (2007)
12. Meyer, C., Ma, B., Kunju, L.P., Davenport, M., Piert, M.: Challenges in accurate registration of 3-d medical imaging and histopathology in primary prostate cancer. *European journal of nuclear medicine and molecular imaging* **40**, 72–78 (2013)
13. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). pp. 565–571. Ieee (2016)
14. Okada, M., Nishio, W., Sakamoto, T., Uchino, K., Yuki, T., Nakagawa, A., Tsubota, N.: Effect of tumor size on prognosis in patients with non-small cell lung cancer: the role of segmentectomy as a type of lesser resection. *The Journal of thoracic and cardiovascular surgery* **129**(1), 87–93 (2005)
15. Okasha, O., Kazmirczak, F., Chen, K.H.A., Farzaneh-Far, A., Shenoy, C.: Myocardial involvement in patients with histologically diagnosed cardiac sarcoidosis: a systematic review and meta-analysis of gross pathological images from autopsy or cardiac transplantation cases. *Journal of the American Heart Association* **8**(10), e011253 (2019)
16. Park, H., Piert, M.R., Khan, A., Shah, R., Hussain, H., Siddiqui, J., Chenevert, T.L., Meyer, C.R.: Registration methodology for histological sections and in vivo imaging of human prostate. *Academic radiology* **15**(8), 1027–1039 (2008)
17. Rampy, B.A., Glassy, E.F.: Pathology gross photography: the beginning of digital pathology. *Surgical Pathology Clinics* **8**(2), 195–211 (2015)

18. Reines March, G.: Registration of pre-operative lung cancer PET/CT scans with post-operative histopathology images. Ph.D. thesis, University of Strathclyde (2020)
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18. pp. 234–241. Springer (2015)
20. Rubinstein, R.: The cross-entropy method for combinatorial and continuous optimization. *Methodology and computing in applied probability* **1**, 127–190 (1999)
21. Samani, A., Zubovits, J., Plewes, D.: Elastic moduli of normal and pathological human breast tissues: an inversion-technique-based investigation of 169 samples. *Physics in medicine & biology* **52**(6), 1565 (2007)
22. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4510–4520 (2018). <https://doi.org/10.1109/CVPR.2018.00474>
23. Schauerte, B., Stiefelhagen, R.: Quaternion-based spectral saliency detection for eye fixation prediction. In: Computer Vision—ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7–13, 2012, Proceedings, Part II 12. pp. 116–129. Springer (2012)
24. Schneider, C.A., Rasband, W.S., Eliceiri, K.W.: Nih image to imagej: 25 years of image analysis. *Nature methods* **9**(7), 671–675 (2012)
25. Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M.: Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Deep learning in medical image analysis and multimodal learning for clinical decision support, pp. 240–248. Springer (2017)
26. Travis, W.D., Asamura, H., Bankier, A.A., Beasley, M.B., Detterbeck, F., Flieder, D.B., Goo, J.M., MacMahon, H., Naidich, D., Nicholson, A.G., et al.: The iaslc lung cancer staging project: proposals for coding t categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the tnm classification of lung cancer. *Journal of Thoracic Oncology* **11**(8), 1204–1223 (2016)
27. Travis, W.D., Dacic, S., Wistuba, I., Sholl, L., Adusumilli, P., Bubendorf, L., Bunn, P., Cascone, T., Chaft, J., Chen, G., et al.: Iaslc multidisciplinary recommendations for pathologic assessment of lung cancer resection specimens after neoadjuvant therapy. *Journal of Thoracic Oncology* **15**(5), 709–740 (2020)
28. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data* **5**(1), 1–9 (2018)
29. Wu, P.h., Bedoya, M., White, J., Brace, C.L.: Feature-based automated segmentation of ablation zones by fuzzy c-mean clustering during low-dose computed tomography. *Medical physics* **48**(2), 703–714 (2021)
30. Xie, S., Tu, Z.: Holistically-nested edge detection. In: Proceedings of the IEEE international conference on computer vision. pp. 1395–1403 (2015)
31. Yuan, Y., Chao, M., Lo, Y.C.: Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. *IEEE transactions on medical imaging* **36**(9), 1876–1886 (2017)