Deep Learning-based methods for automated radiological detection of jaw lesions

Mateusz Grzybowski, Supervised by Dr Neda Azarmehr, Overall supervised by Dr Ali Khurram

INSIGNEO and Department of Clinical Dentistry at University of Sheffield

INSIGNEO
Institute for *in silico* Medicine

Background

The head and neck (H&N) and in particular the maxillofacial region of the human body is unique due to a wide range of pathologies that can involve the bony structures such as the maxilla, mandible, maxillary sinus and palate etc. These lesions can be asymptotic or longstanding and are usually initially identified on radiological examination such as an orthopantomogram (OPT). The radiological features although observable sometimes can cause a debate among specialists. Thus, it is harder to produce one clear objective conclusion [1].

Deep learning methods have been noted to be successful in computer-aided detection and prognosis prediction for a number of pathological conditions. Due to their ability to learn directly from data, it removes subjectivity and variation in diagnosis. On the other hand, their use is limited owing to and insufficient amount of good quality data [2].

The motivation behind this project was to provide a state-of-the-art predictive system that would effectively classify and segment a jaw lesion to help in their treatment. Moreover, this could be developed to support a learning process of a diagnostic radiographer.

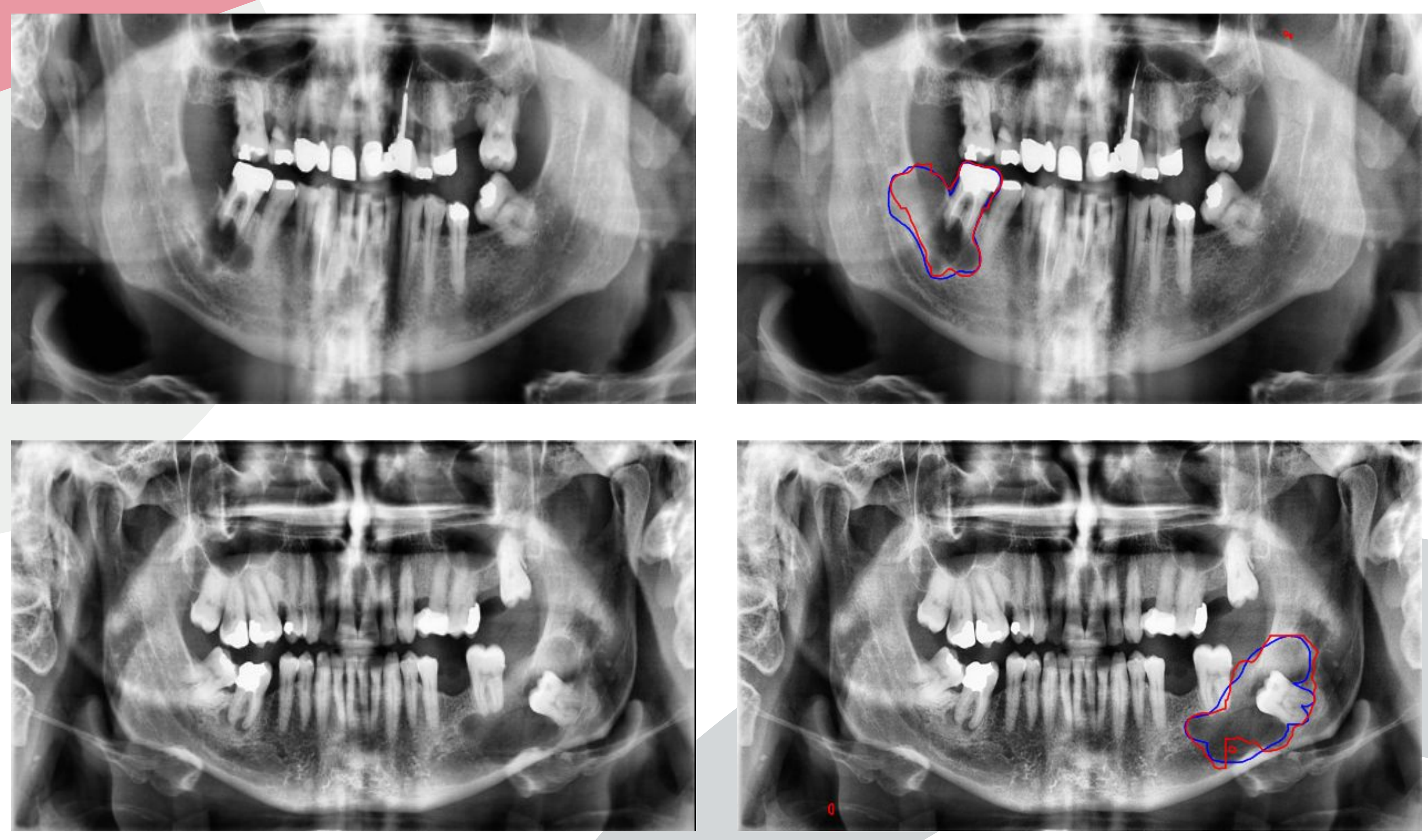

Figure 1: The examples of a very good predictions. Original images on the left and images with overlaid predictions on the right (blue curve - ground truth, red curve - prediction).

Methods

Data:

181 OPT radiology images were collected at the University of Sheffield (Ethics reference 20/WS/0017) of subjects diagnosed with one of the following jaw lesion: ameloblastomas (AM), dentigerous cysts (DC), odontogenic keratocysts (OKC), periapical granulomas (PG) or radicular cysts (RC). The resolution of the images was various hence they were all resized to 1024x1792 so that the dataset could be input to a model. Each image have been classified and manually segmented for the training purpose of a deep learning architecture.

Preprocessing:

Preprocessing stage consisted of four stages. The first was cropping the images to remove a white background that was redundant for training. Secondly, contrast limited adaptive histogram equalisation (CLAHE) applied on the radiographs had clarified the structures of jaws. Furthermore, bilateral filtering denoised and smoothed curves within each image. Lastly, the data was normalised to the range of values from 0 to 1 for more optimal computation.

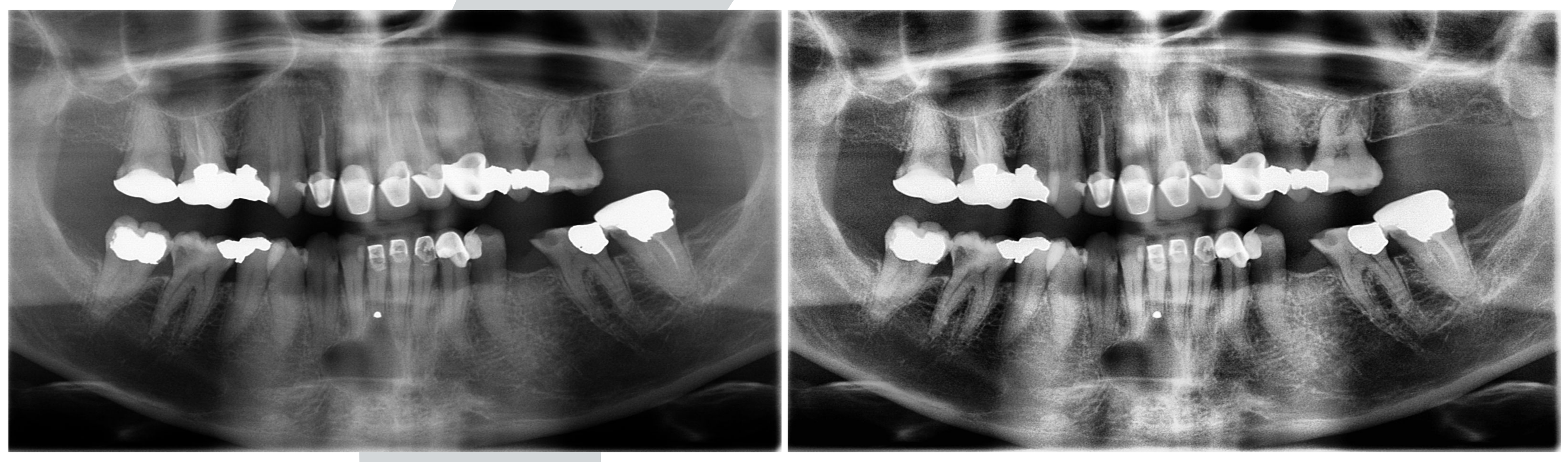

Figure 2: The impact of CLAHE on an radiography image (unprocessed on the left and processed on the right).

Data Augmentation:

Created 10 artificially created images per one original image to feed more data to network. The following transformations to the augmented data were performed: horizontal and vertical flipping, and rotation by +/- 10 degrees. As a result, there were 1260 images in total were used for training, 22 for validation and 33 for testing.

Architecture:

DeepLabV3-ResNet101 - architecture for semantic image segmentation with a ResNet101 backbone [3].

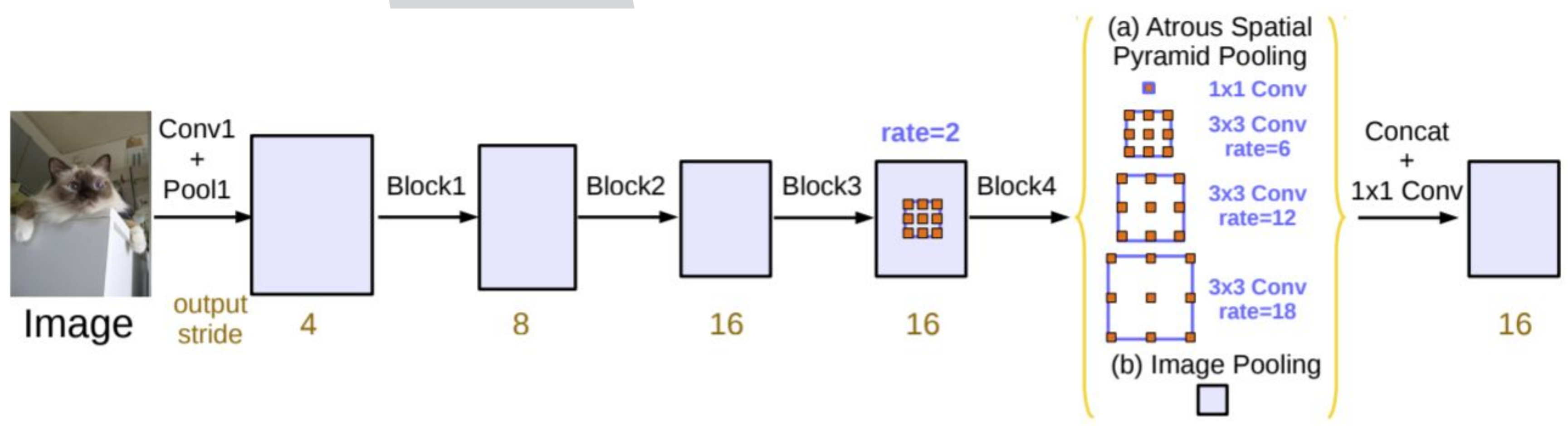

Figure 3: Head of the DeepLabV3 architecture [3].

Results

The model managed to produce some good results which are presented in Figure 1. The scores of the first to the left image were 0.934, 0.877 and 4.899 on Dice coefficient, Jaccard index and Hausdorff distance respectively. Also, for the second image, the results were the following: 0.886 on the Dice coefficient, 0.795 on the Jaccard index and 7.28 on the Hausdorff distance.

The results were obtained with the DeepLabV3-ResNet101 architecture and with the following hyperparameters:

- Learning rate = 0.001
- Epochs = 40 (due to the early stopping, originally 100)
- Batch size = 4
- Momentum = 0.9

The cross entropy function and stochastic gradient descent (SGD) were implemented to the model for this problem. Moreover, 512x256 patches were extracted from each image resulting in 21 patches per image. During testing, the predicted masks were reconstructed from the respectful predicted patches.

Discussion

The results obtained through trials suggest that the model is capable of producing good-quality segmentation masks. However, there were only a few examples with good predictions. To avoid a selection bias, only approximately 22% of images in the test set predictions had at least 70% in the Dice coefficient. Among the best predictions, 4 predictions were for subjects with OKC, 2 for subjects with DC and only one with RC. For the remaining lesion types, no good predictions were developed. This could be caused by the variety of shapes, sizes and structures of lesions which given the small dataset explains the difficulty for the model to learn.

Conclusions

Results shown in Figure 1 indicate that it is possible to make good predictions but there is room for an improvement in the model's ability to generalise and predict well on average. Although the results are currently too poor to be considered useful, solid foundations have been put into further research on the subject. The difficulties of this project were mainly an insufficient amount of data and various shapes of lesions for the semantic segmentation task which make the performance bad on average. However, there are more options to explore which could help in overcoming these issues such as:

- More extensive data augmentation
- Implementation of transfer learning
- Increasing the number of patches used for training
- Trying different models or optimizers

Additionally, as a part of our design, a classification branch was added to the model for a lesion prediction. However, it is yet to be tested due to the time limit of the placement.

Acknowledgements

INSIGNEO Summer Research Placement Programme.

References

[1] M. A. Bruno, E. A. Walker, and H. H. Abujudeh, "Understanding and confronting our mis- takes: the epidemiology of error in radiology and strategies for error reduction," Radiographics, vol. 35, no. 6, pp. 1668–1676, 2015
[2] Kim, M. et al., 2019. Deep Learning in Medical Imaging. *Neurospine*, 16(4), pp.657–668.
[3] Chen, L., Papandreou, G., Schroff, F. and Adam, H., 2017. Rethinking Atrous Convolution for Semantic Image Segmentation.

The Insigneo Institute for *in silico* Medicine is a collaborative initiative between the University of Sheffield, Sheffield Teaching Hospitals NHS Foundation Trust and Sheffield Children's Hospital.

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

Sheffield Children's NHS
NHS Foundation Trust