Implementation of a deep learning model for segmentation of multiple myeloma in CT data

Pavel Gálík
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
xgalik06@vutbr.cz

Michal Nohel
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
xnohel04@vutbr.cz

***Abstract*—This paper deals with the implementation of a deep learning model for spinal tumor segmentation of multiple myeloma patients in CT data. Deep learning is becoming an important part of developing computer-aided detection and diagnosis systems. In this study, a database of 25 patients who were imaged on spectral CT and for whom different parametric images (conventional CT, virtual monoenergetic images, calcium suppression images) were reconstructed, was used. Three convolutional neural network models based on the nnU-Net framework for lytic lesion segmentation were trained on the selected data. The results were evaluated on a test database and the trained models were compared.**

***Index Terms*—multiple myeloma, computed tomography, deep learning, nnU-Net, segmentation, monoenergetic image, calcium suppress image**

I. Introduction

The human spine forms the axis of the upright body and consists of vertebrae that are connected by ligaments, joints, and intervertebral discs. The spine consists of 33 – 34 vertebrae, which divide into 7 *vertebrae cervicales* (C1 – C7), 12 *vertebrae thoracicae* (Th1 – Th12), 5 *vertebrae lumbales* (L1 – L5), 5 vertebrae that create *os sacrum* (S1 – S5), 4 – 5 *os coccygis*. Vertebrae are made of three parts, *corpus vertebrae* (the body), *arcus vertebrae* (vertebral arch), which protect the spinal cord, and is attached to the vertebral body from behind, and *processus vertebrae* (vertebrae process) are attached to the vertebral arch and serve for vertebral mobility. The structure of the vertebrae varies according to their position in the spine, it differs in the size of the vertebral corpus, size, and count of processes [1], [2].

Multiple myeloma (MM) stands as an incurable and biologically diverse disease rooted in the plasma cells. This disease is characterized by the unregulated proliferation of monoclonal plasma cells within the bone marrow, resulting in the excessive production of nonfunctional intact immunoglobulins or their constituent chains. The microenvironment, that supports the growth of myeloma cells and their survival consists of cellular and non-cellular environment compartments. The cellular compartment consists of hematopoietic and non-hematopoietic cells. The first includes myeloid cells, T and B lymphocytes, natural killer cells, and osteoclasts. The second one includes bone marrow stromal cells, bone marrow-derived mesenchymal stromal cells, fibroblasts, osteoblasts, adipocytes, endothelial cells, and blood vessels. Multiple myeloma ranks as the second most prevalent hematologic malignancy in the medical field. It is essentially a hematological malignancy characterized by clonal expansion of plasma cells [3], [4].

One of the typical features of multiple myeloma is its extensive damage to the skeletal system, which affects 80 – 90 % of patients during the course of the disease [5]. These myeloma-induced bone lesions are exclusively osteolytic and contribute to severe and incapacitating bone pain, pathologic fractures, hypercalcemia, spinal cord compression, and heightened mortality rates. The most prevalent symptom observed in individuals with multiple myeloma is bone pain, which is experienced by more than two-thirds of patients. A significant proportion of myeloma patients fall into the elderly demographic population, with a median age at diagnosis of 69 years and a median age at the time of death of 74 years [5]. Over the past three decades, there has been notable progress in treatment, resulting in an increased 5-year survival rate. Nevertheless, it's important to note that multiple myeloma remains an incurable disease [5], [6].

According to the European Myeloma Network and the European Society for Medical Oncology, the whole body low dose CT is recommended as the initial reference standard procedure for the diagnosis of lytic bone disease in patients with MM, as it is able to detect lesions with less than 5 % trabecular bone destruction [7].

Several authors are dealing with the segmentation of tumors: Yao et al. aimed to develop computer-aided detection and diagnosis (CAD) systems, in order to help detect any suspicious abnormalities in a large amount of data. Authors introduced a system for bone metastasis detection [8]. Based on the author's observation and knowledge about bone metastasis, they devised a set of 26 quantitative features in three categories: location, shape, and density. The authors developed a CAD system to detect bone lytic metastases in the thoracolumbar spine using routine CT images [8].

Cheng et al. developed a system for the identification of the metastasis of prostate cancer in whole-body scan images by using a deep convolutional neural network [9]. The developed system exhibited satisfactory performance for a small dataset of 205 cases, 100 of which were of bone metastasis. The sensitivity and precision for bone metastasis in the chest were

0.82 ± 0.08 and 0.70 ± 0.11. The developed system has the potential to provide a prediagnostic report for physicians' final decisions [9].

J. Chmelik et al. dealt with the segmentation and classification of lytic and sclerotic metastatic lesions that are difficult to define by using spinal 3D CT images obtained from highly pathologically affected cases. The authors considered many approaches and uses of different CNN architecture, however, they would always have to overcome some unnecessary difficulties. The authors described a method for voxel-wise segmentation and classification of lytic and sclerotic bone tissues in whole-spine CT scans, based on the CNN architecture. This method can be applied to variously affected patient data acquired 29 with different CT acquisition parameters [10].

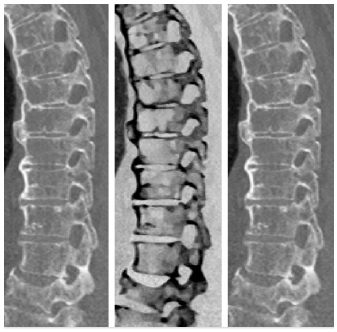
II. Dataset

For this paper, data from 15 multiple myeloma patients and 10 patients without spinal pathology were used. The data was acquired with the help of the University Hospital Brno, Department of Radiology and Nuclear Medicine, using Philips Healthcare IQon spectral CT. The acquisition parameters included a peak tube voltage of 100 kV, tube current of 10 mA, a matrix size of 512×512, slice thickness of 0.9 mm, sharp reconstruction kernel, and hybrid iterative reconstruction technique (iDose4, set to level 4). The data was obtained with the Ethics Committee´s approval under application registration number NU23J-08-00027, and all patients gave informed consent. Two separate readers, one of whom was board certified, reviewed the data utilizing a specialized workstation (Intellispace Portal version 12.1; Philips Healthcare). The diagnosis of Multiple Myeloma (MM) was confirmed through elevated levels of monoclonal immunoglobulin in the blood and an augmented plasma cell count in the bone marrow. Initial disease staging was conducted using spectral CT with a low-dose protocol, adhering to guidelines set forth by the International Myeloma Working Group (IMWG).

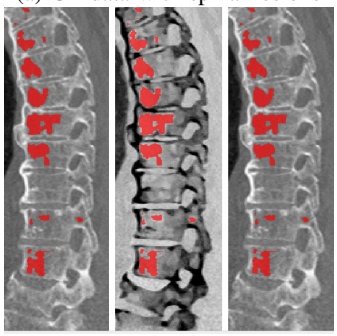
The final database consists of 25 patients: 10 without pathologies in the spine and 15 diagnosed with MM and lytic lesions in the spine. Since spectral CT was used, conventional, virtual monoenergetic images at 40 keV, and calcium-suppressed CT images with index 25 were acquired as shown in Fig. 2. For every CT scan, a mask of a segmented spine was available, which was created using the nnU-Net machine learning model presented in [11]. An example of whole-body CT data is shown in Fig. 1. Also, a manually created mask for lytic lesions was available, the mask was labeled on VMI 40 keV images using MITK software [12], as the contrast is best on these images. Images with lesion mask are shown in Fig. 2(b).

Fig. 1: On the left a whole body conventional CT scan of a patient with multiple myeloma, on the right an example of the available spinal segmentation mask.

As the CT scans are whole-body, the images are very large, thus preprocessing is needed. According to the available spinal segmentation mask, shown in Fig. 1, the images were cropped, and the size significantly decreased, resolving in much faster model training. The model was trained on the described dataset with professionally labeled MM lesions, the example of available training data is shown in Fig 2.



(a) CT data with spinal lesions



(b) CT data with lesions annotations

Fig. 2: Example of multiple myeloma CT data, from left conventional, calcium-suppressed with index 25 and monoenergetic 40 keV CT scans

III. Model description

Machine learning and its quick development have the potential to automate the work of radiologists by automatically detecting and segmenting bone lesions from obtained images. The former segmentation methods often used image segmentation techniques, such as thresholding, region-growing, edge-based segmentation, active contour models, snakes, etc [13].

The nnU-Net [14] is a robust and self-adapting framework based on 2D and 3D vanilla U-Nets. This framework is a fast and effective deep-learning method for biomedical image segmentation. The nnU-Net automatically adapts its architectures to the given image geometry. For each task, the nnU-Net automatically runs a five-fold cross-validation for three different automatically configured U-Net models and the model with the highest mean foreground dice score is chosen for final submission. This paper uses the approach of this model structure, with the code available at https://github.com/MIC-DKFZ/nnUNet [14].

One of the main advantages of the nnU-Net is the ability to handle input images of different sizes or optimize the preprocessing step. As was mentioned earlier, the dataset used for this paper consists of 25 patients, however, ground truth labels were available only for 20 of them. On those, the model was trained.

If the structure of the data for nnU-Net is met, it automatically sets the optimal architecture and learning parameters, with a learning rate set at 0.01 and gradually reduced during training. The batch size was set at 2, patch size of $112 \times 192 \times 112$, a kernel size of $3 \times 3 \times 3$, and a total of 1000 learning epochs with the ReLu activation function were performed. For model validation, 5-fold cross-validation was utilized. For the model training, the Metacenter was used, which had 12 cores, 32GB of RAM, and a dedicated graphics card.

IV. Metrics

For the evaluation of multiple myeloma segmentation, the Dice criterium was used, according to the Equation:

$$Dice(P,T) = 2\frac{|P \bigcap T|}{|P| + |T|} \tag{1}$$

Where P represents the predicted pixel value, and T represents the correct values (ground truth). The value of the Dice score is a dimensionless quantity [–] in the range of 0 and 1 (1 being the absolute match and 0 being no match).

For further evaluation, the false negative rate was used according to the Equation:

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$FNR = 1 - sensitivity \tag{3}$$

Where TP is true positive count, FN is false negative count.

V. Results

Since true labels are available only for the 10 MM and 10 healthy patients, for the 5 remaining MM patients, the evaluation was done only by visual checkup. The result of a chosen segmentation is shown in Fig 3.

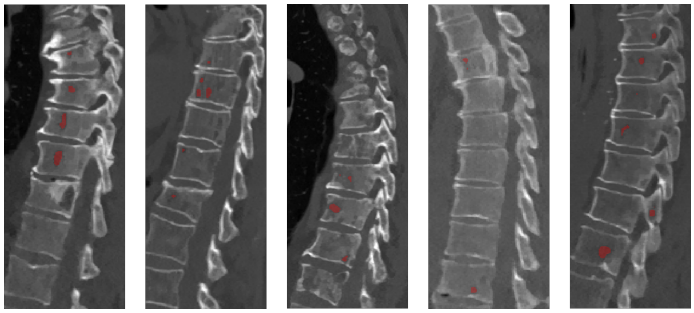


Fig. 3: Results of lesions segmentation on VMI dataset

As can be seen, detected lesions are located correctly, however, not many of the MM lesions were detected by the trained model.

The dice was computed for a validation set, and the results are shown in the following TABLE I. Model 1 represents a model trained on conventional CT data, model 2 represents a model trained on calcium-suppressed data, and model 3 represents a model trained on virtual monoenergetic images.

TABLE I: Dice coefficients for different models

Patient	Model 1	Model 2	Model 3
Myeloma_001	–	0.272	–
Myeloma_004	0.272	–	0.268
Myeloma_006	–	0.333	–
Myeloma_007	0.301	–	0.215

Also, false negative rates were computed for the validation set, and the results are shown in TABLEII. Models represent the same datasets as in the TABLE I.

TABLE II: False negative rate for different models

Patient	Model 1	Model 2	Model 3
Not_Myeloma_001	–	0	–
Not_Myeloma_004	0	–	0
Not_Myeloma_009	0	–	0
Myeloma_001	–	0.49	–
Myeloma_002	–	0.84	–
Myeloma_004	0.83	–	0.84
Myeloma_006	–	0.59	–
Myeloma_007	0.81	–	0.88

According to the results shown in the tables, the models mostly under-segmented and did not detect most of the lesions. This could be given to the small training dataset, which contains mainly patients with low MM occurrence. An example of segmented lesions of different models is shown in Fig 4.

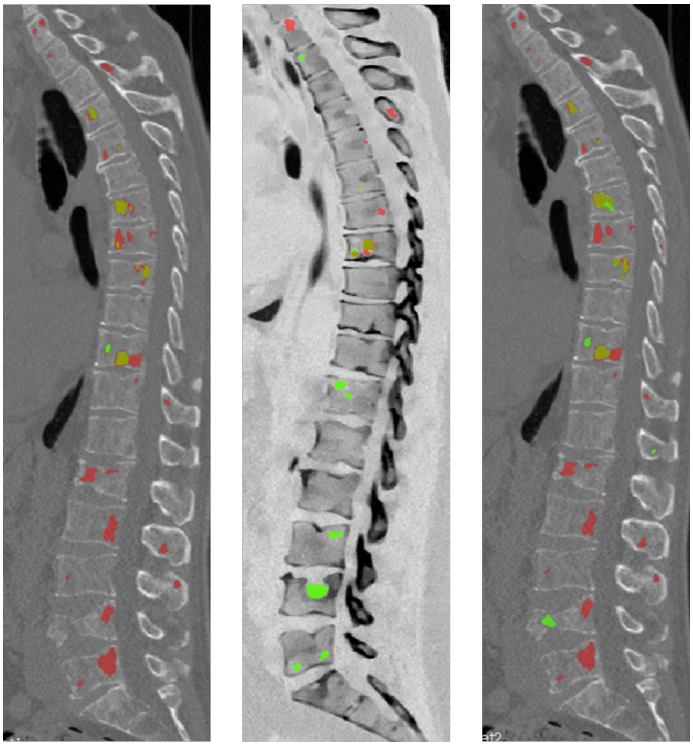


Fig. 4: CT scans with labeled lesions (green color) and with predicted lesion (mix of red and green color). From left conventional, calcium-suppressed with index 25, monoenergetic 40 keV CT scans

Fig 4 shows, that the models mostly under-segment or wrongly detect the MM lesions. A bigger training dataset would ensure better prediction results, as well as training a model on different training datasets (e.g. combined conventional with calcium-suppressed images) could improve the model performance. Also, training only on patients with MM could improve the results. Despite these shortcomings, improved models could very well serve as a checkup tool for skilled professionals.

VI. Conclusion

This paper aims to implement and evaluate the deep learning segmentation model which would be a good help, with diagnosing multiple myeloma, and simplify radiologists' and doctors' work.

Three different nnU-Net models were trained on datasets consisting of 20 patients and their conventional, calcium-suppressed CT images with index 25, and virtual monoenergetic images at 40 keV CT scans. For evaluation, the dice criterium was used with the values above 0.21. Although a relatively small training data set was available, the results shown are quite promising and provide a good basis for further research and development.

Acknowledgment

The paper and the research were supported by Philips Healthcare company and the University Hospital Brno, Department of Radiology and Nuclear Medicine. Computational resources were provided by the Ministry of Education, Youth, and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085) provided within the program Projects of Large Research, Development, and Innovations Infrastructures.

References

[1] FIALA, Pavel; VALENTA, Jiří and EBERLOVÁ, Lada, 2015. Stručná anatomie člověka. Online. Praha: Univerzita Karlova v Praze, nakladatelství Karolinum.

[2] ČIHÁK, Radomír, 2016. Anatomie. Online. Třetí, upravené a doplněné vydání. Praha: Grada.

[3] BRIGLE, Kevin and ROGERS, Barbara, 2017. Pathobiology and Diagnosis of Multiple Myeloma. Seminars in Oncology Nursing. vol. 33, no. 3, pp. 225-236. Available at: https://doi.org/10.1016/j.soncn.2017.05.012.

[4] SIMEONE, F. Joseph; HARVEY, Joel P.; YEE, Andrew J.; O'DONNELL, Elizabeth K.; RAJE, Noopur S. et al., 2019. Value of low-dose whole-body CT in the management of patients with multiple myeloma and precursor states. Online. Skeletal Radiology. vol. 48, no. 5, pp. 773-779. Available at: https://doi.org/10.1007/s00256-018-3066-6.

[5] SILBERMANN, Rebecca and ROODMAN, G. David, 2013. Myeloma bone disease: Pathophysiology and management. Online. Journal of Bone Oncology. vol. 2, no. 2, pp. 59-69. Available at: https://doi.org/10.1016/j.jbo.2013.04.001.

[6] HILLENGASS, Jens; USMANI, Saad; RAJKUMAR, S Vincent; DURIE, Brian G M; MATEOS, María-Victoria et al., 2019. International myeloma working group consensus recommendations on imaging in monoclonal plasma cell disorders. Online. The Lancet Oncology. vol. 20, issue 6, pp. e302-e312. Available at: https://doi.org/10.1016/S1470-2045(19)30309-2.

[7] ORMOND FILHO, Alípio G.; CARNEIRO, Bruno C.; PASTORE, Daniel; SILVA, Igor P.; YAMASHITA, Sâmia R. et al., 2019. Whole-Body Imaging of Multiple Myeloma: Diagnostic Criteria. Online. RadioGraphics. vol. 39, no. 4, pp. 1077-1097. Available at: https://doi.org/10.1148/rg.2019180096.

[8] YAO, Jianhua; O'CONNOR, Stacy and SUMMERS, Ronald, 2007. Computer Aided Detection of Lytic Bone Metastases in the Spine using Routine CT Images. Online. 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. pp. 512-515. Available at: https://doi.org/10.1109/ISBI.2007.356901.

[9] CHENG, Da-Chuan; LIU, Chia-Chuan; HSIEH, Te-Chun; YEN, Kuo-Yang and KAO, Chia-Hung, 2021. Bone Metastasis Detection in the Chest and Pelvis from a Whole-Body Bone Scan Using Deep Learning and a Small Dataset. Online. Electronics. vol. 10, no. 10. Available at: https://doi.org/10.3390/electronics10101201.

[10] CHMELIK, Jiri; JAKUBICEK, Roman; WALEK, Petr; JAN, Jiri; OUREDNICEK, Petr et al., 2018. Deep convolutional neural network-based segmentation and classification of difficult to define metastatic spinal lesions in 3D CT data. Online. Medical Image Analysis. vol. 49, pp. 76-88. Available at: https://doi.org/10.1016/j.media.2018.07.008.

[11] M. Nohel, R. Jakubicek, L. Blazkova, V. Valek, M. Dostal, P. Ourednicek, and J. Chmelik, "Comparison of Spine Segmentation Algorithms on Clinical Data from Spectral CT of Patients with Multiple Myeloma", in *MEDICON-23 and CMBEBIH-23*, 2024, pp. 309-317.

[12] I. Wolf, R. L. Galloway, Jr., M. Vetter, I. Wegner, M. Nolden, T. Bottger, M. Hastenteufel, M. Schobinger, T. Kunert, and H. -P. Meinzer, "The medical imaging interaction toolkit (MITK): a toolkit facilitating the creation of interactive software by extending VTK and ITK", p. 16-.

[13] RICH, Joseph M.; BHARDWAJ, Lokesh N.; SHAH, Aman; GANGAL, Krish; RAPAKA, Mohitha S. et al., 2023. Deep learning image segmentation approaches for malignant bone lesions: a systematic review and meta-analysis. Online. Frontiers in Radiology. 2023-8-8, vol. 3. Available at: https://doi.org/10.3389/fradi.2023.1241651.

[14] ISENSEE, Fabian, Paul F. JAEGER, Simon A. A. KOHL, Jens PETERSEN a Klaus H. MAIER-HEIN. NnUNet: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods. 2021, 18(2), 203-211. ISSN 1548-7091.