The Veterinary Journal 235 (2018) 90–92

Contents lists available at ScienceDirect

The Veterinary Journal

journal homepage: www.elsevier.com/locate/tvjl

Short Communication

Development of a deep convolutional neural network to predict grading of canine meningiomas from magnetic resonance images

T. Banzato[a], G.B. Cherubini[b], M. Atzori[c], A. Zotti[a,*]

[a] Department of Animal Medicine, Production and Health, Clinical Section, Radiology Unit, University of Padua, Viale dell'Università 16, Legnaro 35020, Padua, Italy
[b] Dick White Referrals (Cherubini), Six Mile Bottom, Cambridgeshire CB8 0UH, UK
[c] Information Systems Institute, University of Applied Sciences Western Switzerland (HES-SO Valais), 3960 Sierre, Switzerland

ARTICLE INFO

Article history:
Accepted 1 April 2018

Keywords:
Canine
Deep neural network
Histopathology
Magnetic resonance imaging
Meningioma

ABSTRACT

An established deep neural network (DNN) based on transfer learning and a newly designed DNN were tested to predict the grade of meningiomas from magnetic resonance (MR) images in dogs and to determine the accuracy of classification of using pre- and post-contrast T1-weighted (T1W), and T2-weighted (T2W) MR images. The images were randomly assigned to a training set, a validation set and a test set, comprising 60%, 10% and 30% of images, respectively. The combination of DNN and MR sequence displaying the highest discriminating accuracy was used to develop an image classifier to predict the grading of new cases. The algorithm based on transfer learning using the established DNN did not provide satisfactory results, whereas the newly designed DNN had high classification accuracy. On the basis of classification accuracy, an image classifier built on the newly designed DNN using post-contrast T1W images was developed. This image classifier correctly predicted the grading of 8 out of 10 images not included in the data set.

Meningiomas are the most common intracranial neoplasms in dogs (Song et al., 2013). Currently, magnetic resonance (MR) imaging is the gold standard for the clinical diagnosis of meningiomas in dogs. To date, specific MR imaging features have not been associated with the histopathological grades of canine meningiomas (Sturges et al., 2008). The aim of this study was to develop a clinically applicable test based on deep neural networks (DNNs) to predict the histopathological grading of canine meningiomas by analysis of MR images.

Two DNNs were developed using the Neural Network Toolbox in MATLAB and Statistics Toolbox Release 2017b (MathWorks). The first DNN used a transfer learning approach based on AlexNet (Krizhevsky et al., 2012), which was pre-trained on ImageNet data, and retrained and fine-tuned using MR data. The second DNN was a 'scratch' DNN (scrDNN) designed de novo. The structure of AlexNet and scrDNN and explanations of the layers comprising these DNNs are provided in Appendix: Supplementary material.

Fifty-six individual MR brain scans in dogs with a final diagnosis of intracranial meningioma were retrospectively selected from the databases of the Portoni Rossi Veterinary Hospital (PR; $n = 23$) and Dick White Referrals (DWR, $n = 33$). The MR scans were performed with a 0.22 T open permanent magnet at PR and a 0.4 T open permanent magnet at DWR. Forty-five cases included in the present study (PR, $n = 18$; DWR, $n = 27$) were also part of a previous study on texture image analysis (Banzato et al., 2017).

All images where the lesion was visible in coronal, sagittal or transverse scans were included and were reformatted to match the DNN requirements (227×227 RGB images), then manually cropped, carefully retaining the square shape of the original MR image, so that only the lesion and a small amount of the adjacent tissues were included. Images in which partial volume artefact was present were not included. Pre-contrast T1W, post-contrast T1W and T2W images were analysed separately.

The classification efficiency of the two DNNs on pre- and post-contrast T1-weighted (T1W) and T2-weighted (T2W) MR images was tested. The combination between the DNN and the MR sequence displaying the highest discriminating accuracy was used to develop an image classifier to predict the grading of new cases. The image classifier was then tested on 10 random images not included in the database order to analyse the model accuracy.

The main difference between the two DNNs is that AlexNet was pre-trained on a large-scale data set (millions of images), whereas the scrDNN was trained on a small data set. To prevent overfitting, the entire data set was divided into a training set (60%), a validation

* Corresponding author.
E-mail address: alessadro.zotti@unipd.it (A. Zotti).

https://doi.org/10.1016/j.tvjl.2018.04.001

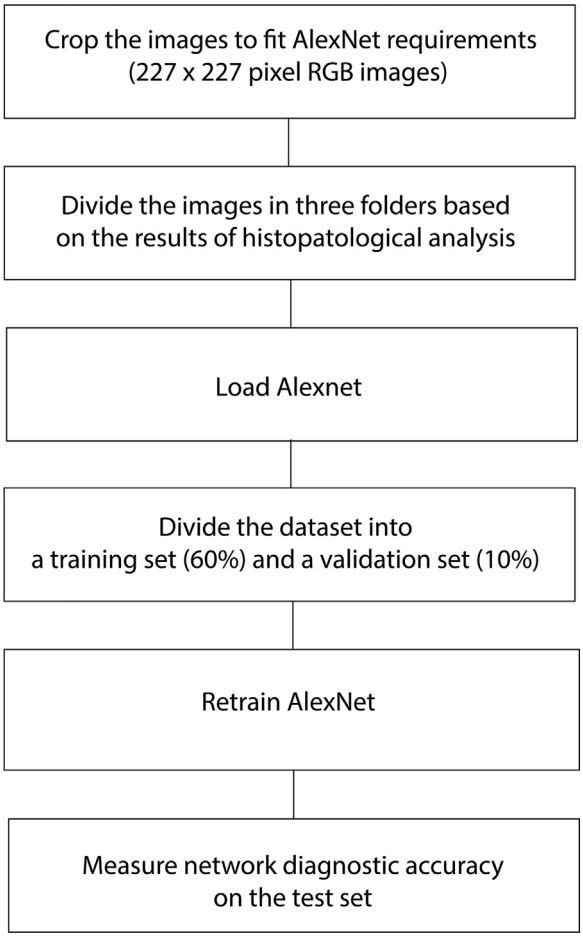

Fig. 1. Workflow used for transfer learning with AlexNet.

set (10%) and a test set (30%), maintaining the proportions of the different categories in both sets. The DNNs were retrained on the training set images, and a data augmentation algorithm that added randomly flipped images from the database. Several combinations of network hyperparameters (leanRate and miniBatchSize) were tested to fine tune the DNNs until 80% accuracy was achieved using the validation set. Classification accuracy was computed as the percentage of correctly labelled images in the test set.

The multi-class Matthew's correlation coefficient (MCMCC) was also calculated using the MATLAB Multi Class Confusion Matrix function. The accuracy of the fine-tuned network was tested on the test set of images using both the training and the validation set. To account for the random distribution of the images in the training and test sets, the procedure was repeated five times for each MR sequence. The mean, along with the limits of the overall range, of the accuracy, and of the MCMCC were calculated. A diagram of the workflow used for the transfer learning procedure is shown in Fig. 1. The combination of DNN and MR sequence displaying the highest classification accuracy was used in developing and testing an image classifier to predict the grading of 10 images that were not previously included in the data set.

Of the 56 cases of canine meningioma included in the study, 26 were classified as grade I, 22 as grade II and eight as grade III (Louis et al. 2007); six cases (two grade I, three grade II and one grade III) were discarded because the lesions were almost completely cystic and only a very small amount of tissue was available for analysis (Table 1). Complete classification results of AlexNet and of scrDNN are reported in Table 2. The combination of post-contrast T1W images and our newly developed DNN (scrDNN) was chosen in developing the image classifier. The classifier correctly classified 8/10 images; one grade III image was misclassified as grade II and one grade I image was misclassified as grade II.

The scrDNN developed in this project had good classification accuracy in prediction of the histopathological grading of canine meningiomas and outperformed the transfer learning of AlexNet. A possible explanation for the different performance of the two networks is that AlexNet includes a larger number of parameters

Table 1
Histopathological subtypes and grading score of meningiomas (Sturges et al., 2008).

Histopathological type	Grade I	Grade II	Grade III
Papillary ($n = 11$)	1	5	5
Transitional ($n = 9$)	7	2	0
Atypical ($n = 6$)	0	6	0
Meningothelial ($n = 4$)	3	1	0
Fibroblastic ($n = 4$)	2	2	0
Psammomatous ($n = 3$)	1	2	0
Syncytial ($n = 3$)	3	0	0
Lipomatous ($n = 3$)	3	0	0
Meningoendothelial ($n = 3$)	3	0	0
Chordoid ($n = 2$)	0	1	1
Anaplastic ($n = 2$)	0	1	1
Other (biphasic, cystic, malignant, microcystic, osteoid, vacuolar, vascular) ($n = 6$)	3	2	1
Total ($n = 56$)	26	22	8

Table 2
Accuracy of AlexNet and the scratch DNN (scrDNN) in the prediction of the histopathological grading of canine meningiomas from different magnetic resonance images sequences.

	Pre-contrast T1 images		Post-contrast T1 images		T2 images	
	Accuracy	MCMCC	Accuracy	MCMCC	Accuracy	MCMCC
AlexNet	65.2 (60–71.7)	0.46 (0.45–0.61)	68 (60–76.7)	0.49 (0.44–0.58)	65.6 (61.7–68.3)	0.44 (0.4–0.6)
scrDNN	78.5 (72.6–82)	0.62 (0.56–0.7)	82.2 (75–84.6)	0.68 (0.63–0.76)	75.6 (71.3–80.1)	0.61 (0.58–0.67)

Accuracy is calculated as the percentage of correctly labelled images on the total number of images in the test set. The multi-class Matthew's correlation coefficient (MCMCC) is also reported. Data are reported as mean with the limits of the overall range.

than our scrDNN and therefore the ratio between the number of parameters and the number of images in our scrDNN is more favourable. Moreover, the scrDNN proposed in this paper was developed and trained only on MR images and, most likely, features that are more specific to solving this particular classification problem were automatically selected.

The rationale beyond the use of DNNs in this study is that these algorithms are capable of determining the relationships between the extracted features at a deep level (LeCun et al., 2015). Therefore, they may be able to detect hidden patterns in the data and to accomplish classification tasks that are not appreciable by direct observation. One of the main limitations of DNNs is that they act as 'black boxes'. The structure of a DNN does not provide any useful information about the data set and, therefore, it is impossible, retrospectively, to determine which features have predictive value (Prieto et al., 2016).

Limitations of this study include the relatively low number of grade III meningiomas available in our database ($n = 7$). Also, the histopathological examinations were performed by different pathologists, thus adding inter-observer variability to the model. Furthermore, two different MR scanners were used, thus adding further sources of variability. The inclusion of a larger number of cases, including a larger number of grade III lesions, may enhance the classification accuracy of our DNN.

Conflict of interest statement

None of the authors of this paper has a financial or personal relationship with other people or organizations that could inappropriately influence or bias the content of the paper.

Acknowledgements

The authors gratefully acknowledge Dr Marco Bernardini, Department of Animal Medicine, Production and Health, University of Padua, and Portoni Rossi Veterinary Hospital, Bologna, Italy, for providing the clinical cases and discussing images. The authors would like to thank NVIDIA for donation of the GPU card used in this study. This research was supported by three research grants (CPDA124900/2012 and Junior Research Grant 2015) from the University of Padua, Italy, by a research grant (FSE 2105/201/27/1148/201) from Veneto Regional Council, Italy, and Supporting Talents in Research@University of Padua, funded for a total amount of 180.000 euros with a project entitled "Prediction of the histological grading of human meningiomas using MR texture analysis and deep learning: a traslational application of a model developed on spontaneously occurring meningiomas in dogs."

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.tvjl.2018.04.001.

References

Banzato, T., Bernardini, M., Cherubini, G.B., Zotti, A., 2017. Texture analysis of magnetic resonance images to predict histologic grade of meningiomas in dogs. American Journal Veterinary Research 78, 1156–1162.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems, vol. 1, 3–6 December 2012, Lake Tahoe, Nevada, USA, pp. 1097–1105.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

Louis, D.N., Ongaki, H., Wiestler, O.D., Cavenee, W.K., 2007. Meningiomas. In: Louis, D.N. (Ed.), WHO Classification of Tumours of the Central Nervous System. International Agency for Research on Cancer, Lyon, France, pp. 164–172.

Prieto, A., Prieto, B., Ortigosa, E.M., Ros, E., Pelayo, F., Ortega, J., Rojas, I., 2016. Neural networks: an overview of early research, current frameworks and new challenges. Neurocomputing 214, 242–268.

Song, R.B., Vite, C.H., Bradley, C.W., Cross, J.R., 2013. Postmortem evaluation of 435 cases of intracranial neoplasia in dogs and relationship of neoplasm with breed, age, and body weight. Journal of Veterinary Internal Medicine 27, 1143–1152.

Sturges, B.K., Dickinson, P.J., Bollen, A.W., Koblik, P.D., Kass, P.H., Kortz, G.D., Vernau, K.M., Knipe, M.F., LeCouteur, R.A., Higgins, R.J., 2008. Magnetic resonance imaging and histological classification of intracranial meningiomas in 112 dogs. Journal of Veterinary Internal Medicine 22, 586–595.