Deep Learning Methods in Mammography

Gustavo Carneiro

1 Review on Deep Learning Methods in Mammography

Breast cancer is one of the most common types of cancer affecting the lives of women worldwide. Recent statistical data published by the World Health Organisation (WHO) estimates that 23% of cancer related cases and 14% of cancer related deaths among women are due to breast cancer [1]. The most effective tool to reduce the burden associated with breast cancer consists of early detection in asymptomatic women via breast cancer screening programs [2], which commonly use mammography for breast imaging. Breast screening using mammography comprises several steps, which include the detection and analysis of lesions, such as masses and calcifications, that are used in order to estimate the risk that the patient is developing breast cancer. In clinical settings, this analysis is for the most part a manual process, which is susceptible to the subjective assessment of a radiologist, resulting in a potentially large variability in the final estimation. The effectiveness of this manual process can be assessed by recent studies that show that this manual analysis has a sensitivity of 84% and a specificity of 91% [3]. Other studies show evidence that a second reading of the same mammogram either from radiologists or from computer-aided diagnosis (CAD) systems can improve this performance [3]. Therefore, given the potential impact that second reading CAD systems can have in breast screening programs, there is a great deal of interest in the development of such systems.

A CAD system that can analyse breast lesions from mammograms usually comprises three steps [3]: 1) lesion detection, 2) lesion segmentation, and 3) lesion classification. The main challenges involved in these steps are related to the low signal-to-noise ratio present in the imaging of the lesion, and the lack of a consistent location, shape and appearance of lesions [4, 5]. Current methodologies for lesion detection involve the identification of a large number of candidate regions,

Gustavo Carneiro
Australian Centre for Visual Technologies, The University of Adelaide, Adelaide, Australia e-mail: gustavo.carneiro@adelaide.edu.au

usually based on the use of traditional filters, such as morphological operators or difference of Gaussians [6, 7, 8, 9, 10, 11, 12, 13]. These candidates are then processed by a second stage that aims at removing false positives using machine learning approaches (e.g., region classifier) [6, 7, 8, 9, 10, 11, 12, 13]. The main challenges faced by lesion detection methods is that they may generate a large number of false positives, while missing a good proportion of true positives [4]; in addition, another issue is the poor alignment of the detected lesion in terms of translation and scale within the candidate regions - this issue has negative consequences for the sub-sequent lesion segmentation that depends on a relatively precise alignment. Lesion segmentation is then addressed with global/local energy minimisation models on a continuous or discrete space [14, 15, 16]. The major roadblock faced by these methods is the limited availability of annotated datasets that can be used in the training of the segmentation models. This is a particularly important problem because, differently from the detection and classification of lesions, the segmentation of lesions is not a common task performed by radiologists, which imposes strong limitations in the annotation process and, as a consequence, in the availability of annotated datasets. In fact, the main reason behind the need for a lesion segmentation is the assumption that the lesion shape is an important feature in the final stage of the analysis: lesion classification. This final stage usually involves the extraction of manually or automatically designed features from the lesion image and shape and the use of those features with traditional machine learning classifiers [17, 18, 19]. In this last stage, the main limitation is with respect to the features being extracted for the classification because these features are usually hand-crafted, which cannot guarantee optimality for this classification stage.

The successful use and development of deep learning methods in computer vision problems (i.e., classification and segmentation) [20, 21, 22, 23, 24] have motivated the medical image analysis community to investigate the applicability of such methods in medical imaging segmentation and classification problems. Compared to the more traditional methods presented above (for the problem of mammogram analysis), deep learning methods offer the following clear advantages: automated learning of features estimated based on specific detection/segmentation/classification objective functions; opportunity to build complete "end-to-end" systems that take an image, detect, segment and classify visual objects (e.g., breast lesion) using a single model and a unified training process. However, the main challenge faced by deep learning methods is the need for large annotated training sets given the scale of the parameter space, usually in the order of 10^6 parameters. This problem is particularly important in medical image analysis applications, where annotated training sets rarely have more than a few thousand samples. Therefore, a great deal of research is focused on the adaptation of deep learning methods to medical image analysis applications that contain relatively small annotated training sets.

There has been an increasing interest in the development of mammogram analysis methodologies based on deep learning. For instance, the problem of breast mass segmentation has been addressed with the use of a structured output model, where several potential functions are based on deep learning models [25, 26, 27]. The assumption here is that deep learning models alone cannot produce results that are accurate enough due to the small training set size problem mentioned above, but if these models are combined with a structured output model that makes assumptions about the appearance and shape of masses, then it is possible to have a breast mass segmentation that produces accurate results - in fact this method holds the best results in the field in two publicly available datasets [19, 28]. Segmentation of breast tissue using deep learning alone has been successfully implemented [29], but it is possible that a similar structured output model could improve even more the accuracy obtained. Dhungel et al. [30] also worked on a breast mass detection methodology that consists of a cascade of classifiers based on the Region Convolutional Neural Network (R-CNN) [23] approach. The interesting part is that the candidate regions produced by the R-CNN contain too many false positives, so the authors had to include an additional stage based on a classifier to eliminate those false positives. Alternatively, Ertosun and Rubin [31] propose a deep learning based mass detection method consisting of a cascade of deep learning models trained with DDSM [28] - the main reason that explains the succesful use of deep learning models here is the size of DDSM, which contains thousands of annotated mammograms. The classification of lesions using deep learning [32, 33, 34] has also been successfully implemented in its simplest form: as a simple lesion classifier. Carneiro et al. [35] have proposed a system that can classify the unregistered two views of a mammography exam (cranial-caudal and mediolateral-oblique) and their respective segmented lesions and produce a classification of the whole exam. The importance of this work lies in its ability to process multi-modal inputs (images and segmentation maps) that are not registered, in its way of performing transfer learning from computer vision datasets to medical image analysis datasets, and also in its capability of producing high-level classification directly from mammograms. A similar high-level classification using deep learning estimates the risk of developing breast cancer by scoring breast density and texture [36, 37]. Another type of high-level classification is the method proposed by Qiu et al. [38] that assesses the short-term risk of developing breast cancer from a normal mammogram.

Based on the recent results presented above, it is clear that the use of deep learning is allowing accuracy improvements in terms of mass detection, segmentation and classification. All the studies above have been able to mitigate the training set size issue with the use of regularisation techniques or the combination of different approaches that can compensate the relatively poor generalisation of deep learning methods trained with small annotated training sets. More importantly, deep learning is also allowing the implementation of new applications that are more focused on high-level classifications that do not depend on lesion segmentation. The annotation for this higher level tasks is readily available from clinical datasets, which generally contain millions of cases that can be used to train deep learning models in a more robust manner. These new applications are introducing a paradigm shift in how the field analyses mammograms: from the classical 3-stage process (detection, segmentation and classification of lesions) trained with small annotated datasets to a 1-stage process consisting of lesion detection and classification trained with large annotated datasets.

References

1. A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, T. Murray, and M. J. Thun, "Cancer statistics, 2008," *CA: a cancer journal for clinicians*, vol. 58, no. 2, pp. 71–96, 2008.
2. B. Lauby-Secretan, C. Scoccianti, D. Loomis, L. Benbrahim-Tallaa, V. Bouvard, F. Bianchini, and K. Straif, "Breast-cancer screeningviewpoint of the iarc working group," *New England Journal of Medicine*, vol. 372, no. 24, pp. 2353–2358, 2015.
3. M. L. Giger and A. Pritzker, "Medical imaging and computers in the diagnosis of breast cancer," in *SPIE Optical Engineering+ Applications*, pp. 918908–918908, International Society for Optics and Photonics, 2014.
4. A. Oliver, J. Freixenet, J. Marti, E. Perez, J. Pont, E. R. Denton, and R. Zwiggelaar, "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis*, vol. 14, no. 2, pp. 87–110, 2010.
5. J. Tang, R. M. Rangayyan, J. Xu, I. El Naqa, and Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: recent advances," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 2, pp. 236–251, 2009.
6. E. Kozegar, M. Soryani, B. Minaei, I. Domingues, *et al.*, "Assessment of a novel mass detection algorithm in mammograms," *Journal of cancer research and therapeutics*, vol. 9, no. 4, p. 592, 2013.
7. M. Beller, R. Stotzka, T. O. Müller, and H. Gemmeke, "An example-based system to support the segmentation of stellate lesions," in *Bildverarbeitung für die Medizin 2005*, pp. 475–479, Springer, 2005.
8. G. M. te Brake, N. Karssemeijer, and J. H. Hendriks, "An automatic method to discriminate malignant masses from normal tissue in digital mammograms," *Physics in Medicine and Biology*, vol. 45, no. 10, p. 2843, 2000.
9. R. Campanini, D. Dongiovanni, E. Iampieri, N. Lanconelli, M. Masotti, G. Palermo, A. Riccardi, and M. Roffilli, "A novel featureless approach to mass detection in digital mammograms based on support vector machines," *Physics in Medicine and Biology*, vol. 49, no. 6, p. 961, 2004.
10. N. H. Eltonsy, G. D. Tourassi, and A. S. Elmaghraby, "A concentric morphology model for the detection of masses in mammography," *Medical Imaging, IEEE Transactions on*, vol. 26, no. 6, pp. 880–889, 2007.
11. M. P. Sampat, A. C. Bovik, G. J. Whitman, and M. K. Markey, "A model-based framework for the detection of spiculated masses on mammographya)," *Medical physics*, vol. 35, no. 5, pp. 2110–2123, 2008.
12. R. Bellotti, F. De Carlo, S. Tangaro, G. Gargano, G. Maggipinto, M. Castellano, R. Massafra, D. Cascio, F. Fauci, R. Magro, *et al.*, "A completely automated cad system for mass detection in a large mammographic database," *Medical physics*, vol. 33, no. 8, pp. 3066–3075, 2006.
13. J. Wei, B. Sahiner, L. M. Hadjiiski, H.-P. Chan, N. Petrick, M. A. Helvie, M. A. Roubidoux, J. Ge, and C. Zhou, "Computer-aided detection of breast masses on full field digital mammograms," *Medical physics*, vol. 32, no. 9, pp. 2827–2838, 2005.
14. J. E. Ball and L. M. Bruce, "Digital mammographic computer aided diagnosis (cad) using adaptive level set segmentation," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 4973–4978, IEEE, 2007.
15. P. Rahmati, A. Adler, and G. Hamarneh, "Mammography segmentation with maximum likelihood active contours," *Medical image analysis*, vol. 16, no. 6, pp. 1167–1186, 2012.
16. J. S. Cardoso, I. Domingues, and H. P. Oliveira, "Closed shortest path in the original coordinates with an application to breast cancer," *International Journal of Pattern Recognition and Artificial Intelligence*, 2014.
17. C. Varela, S. Timp, and N. Karssemeijer, "Use of border information in the classification of mammographic masses," *Physics in Medicine and Biology*, vol. 51, no. 2, p. 425, 2006.
18. J. Shi, B. Sahiner, H.-P. Chan, J. Ge, L. Hadjiiski, M. A. Helvie, A. Nees, Y.-T. Wu, J. Wei, C. Zhou, *et al.*, "Characterization of mammographic masses based on level set segmentation with new image features and patient information," *Medical physics*, vol. 35, no. 1, pp. 280–290, 2008.

19. I. Domingues, E. Sales, J. Cardoso, and W. Pereira, "Inbreast-database masses characterization," *XXIII CBEB*, 2012.
20. Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, 1995.
21. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks.," in *NIPS*, vol. 1, p. 4, 2012.
22. C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 8, pp. 1915–1929, 2013.
23. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pp. 580–587, IEEE, 2014.
24. Y. Zhang, K. Sohn, R. Villegas, G. Pan, and H. Lee, "Improving object detection with deep convolutional networks via bayesian optimization and structured prediction," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 249 – 258, june 2015.
25. N. Dhungel, G. Carneiro, and A. P. Bradley, "Deep learning and structured prediction for the segmentation of mass in mammograms," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, pp. 605–612, Springer, 2015.
26. N. Dhungel, G. Carneiro, and A. P. Bradley, "Tree re-weighted belief propagation using deep learning potentials for mass segmentation from mammograms," in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 760–763, April 2015.
27. N. Dhungel, G. Carneiro, and A. P. Bradley, "Deep structured learning for mass segmentation from mammograms," in *Image Processing (ICIP), 2015 IEEE International Conference on*, pp. 2950–2954, Sept 2015.
28. M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer, "The digital database for screening mammography," in *Proceedings of the 5th international workshop on digital mammography*, pp. 212–218, 2000.
29. A. Dubrovina, P. Kisilev, B. Ginsburg, S. Hashoul, and R. Kimmel, "Computational mammography using deep neural networks," in *Workshop on Deep Learning in Medical Image Analysis (DLMIA)*, 2016.
30. N. Dhungel, G. Carneiro, and A. Bradley, "Automated mass detection in mammograms using cascaded deep learning and random forests," in *Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on*, pp. 1–8, Nov 2015.
31. M. G. Ertosun and D. L. Rubin, "Probabilistic visual search for masses within mammography images using deep learning," in *Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on*, pp. 1310–1315, IEEE, 2015.
32. J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. G. Lopez, "Representation learning for mammography mass lesion classificationwith convolutional neural networks," *Computer Methods and Programs in Biomedicine*, 2016.
33. Y. Qiu, S. Yan, M. Tan, S. Cheng, H. Liu, and B. Zheng, "Computer-aided classification of mammographic masses using the deep learning technology: a preliminary study," in *SPIE Medical Imaging*, pp. 978520–978520, International Society for Optics and Photonics, 2016.
34. Z. Jiao, X. Gao, Y. Wang, and J. Li, "A deep feature based framework for breast masses classification," *Neurocomputing*, 2016.
35. G. Carneiro, J. Nascimento, and A. P. Bradley, "Unregistered multiview mammogram analysis with pre-trained deep learning models," in *Medical Image Computing and Computer-Assisted InterventionMICCAI 2015*, pp. 652–660, Springer, 2015.
36. M. Kallenberg, K. Petersen, M. Nielsen, A. Ng, P. Diao, C. Igel, C. Vachon, K. Holland, N. Karssemeijer, and M. Lillholm, "Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring," 2016.
37. K. Petersen, M. Nielsen, P. Diao, N. Karssemeijer, and M. Lillholm, "Breast tissue segmentation and mammographic risk scoring using deep learning," in *Breast Imaging*, pp. 88–94, Springer, 2014.

38. Y. Qiu, Y. Wang, S. Yan, M. Tan, S. Cheng, H. Liu, and B. Zheng, "An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology," in *SPIE Medical Imaging*, pp. 978521–978521, International Society for Optics and Photonics, 2016.