Chelonian Conservation And Biology

Vol. 17No.2 (2022) | https://www.acgpublishing.com/ | ISSN - 1071-8443

DOI:doi.org/10.18011/2022.04(1) 2036.2045

THE INTEGRATION OF RADIOLOGICAL FINDINGS WITH ELECTRONIC HEALTH RECORDS AND CLINICAL DECISION SUPPORT SYSTEMS

Bader Faize Mohammed Al Dawseri, Abdullah Mansour Akkam,Ibrahim Fahad Alhussain, Bader Hmood Alotaiabi, Ahmad Ibrahim Ibrahim Jarrah, Rashed Mohammed Alqahtani, Zaed Ali Essa Tawashi, Ibrahim Hussain Majrashi, Ali Marwi Almokhalfi, Abdulaziz Abdullah Alhazzaa, Ali Abdullah Almangoor, Mashan Muhamel Alotaibi, Naif Mohsan Almogamsi, Aldhuwalia,Yassmine Ibrahim, Raed Hamed Alkhodidi

Abstract

Deep learning techniques have the potential to significantly improve healthcare, especially in areas where medical imaging is used for diagnosis, prognosis, and treatment choices. The most advanced deep learning models available today for radiological applications only take into account pixel-value data; they do not take into account data that provides clinical context. In actuality, however, doctors are able to interpret imaging results in the proper clinical context thanks to relevant and accurate non-imaging data based on the clinical history and laboratory data. These results in increased diagnostic accuracy, informed clinical decision making, and better patient outcomes. Medical imaging pixel-based models need to be able to process contextual data from electronic health records (EHR) in addition to pixel data in order to accomplish a comparable aim utilizing deep learning. In this study, we thoroughly analyze the medical data fusion literature produced between 2012 and 2020 and discuss various data fusion strategies that can be used to merge medical imaging with EHR. We performed a thorough search for original research publications using deep learning to fuse multimodality data on PubMed and Scopus. We analyzed 985 studies in all, and we took data out of 17 publications. We present current information, highlight significant findings, and offer implementation advice through this systematic review, which can be used as a reference by researchers who are interested in using multimodal fusion in medical imaging.

Keywords: Deep learning techniques, electronic health records (EHR), medical imaging, review, clinical decision.

1. Introduction

The synthesis of information and data from several sources is crucial to the practice of modern medicine. These sources include structured laboratory data, unstructured narrative data, audio or observational data, and imaging pixel data. This is especially true when interpreting medical images, as making diagnoses frequently requires a thorough understanding of the clinical situation. For instance, it has been frequently demonstrated that a referring provider's performance and clinical utility are negatively impacted when they are unable to access clinical and laboratory data during picture interpretation1,2. The majority of radiologists (87%) who responded to a survey said that clinical information significantly affected interpretation3. Radiology is not the only imaging-based medical specialty that depends on clinical context for accurate image interpretation; several other imaging-based disciplines, including pathology, ophthalmology, and dermatology, also use this information to inform their image interpretation practices4,5,6. Physicians can interpret imaging results in the appropriate clinical context with the help of pertinent and accurate information about the patient's past medical history and current symptoms. This helps the physicians make a more relevant differential diagnosis and produces a report that is more useful to them and the patient.

There are an increasing number of radiological imaging exams in the present digital era. An average radiologist may need to interpret an image every 3.4 seconds during an 8-hour workday in order to fulfill this increasing workload demand, which raises the risk of burnout, weariness, and increased error rates7. The potential for effective automated systems to either supplement or relieve overworked physicians of cognitive effort is driving the proliferation of deep learning in the healthcare industry8,9,10. Medical images are frequently subjected to convolutional neural networks (CNNs), a type of deep learning that has shown to be highly successful in image recognition and classification tasks. Chest X-rays, skin cancer, and diabetic retinopathy are among the early medical images that CNNs were used to analyze11,12,13,14,15,16,17, and 18. Nevertheless, these models may eventually restrict practical translation because they only take into account the pixel data as a single channel for input and are unable to contextualize additional clinical information as would be done in medical practice.

For instance, numerous researchers have trained deep learning models for automated detection and categorization of diseases on chest X-rays, achieving the seemingly "simple" job of recognizing pneumonia on a chest radiograph19,20. However, these applications may ultimately have minimal influence on clinical practice in the absence of clinical context, such as patient history, principal complaint, previous diagnoses, and test data. Even while the imaging characteristics on chest X-rays consistent with pneumonia can typically distinguish it from other diseases, the findings are nonspecific, and a complete diagnosis necessitates the context of laboratory and clinical data. That is to say, a chest X-ray finding that suggests pneumonia would be accurate in a patient who has a fever and an elevated white blood cell count, but in a different patient who does not have those supporting laboratory values and clinical characteristics, the same imaging finding might instead indicate atelectasis, pulmonary edema, or even lung cancer.

Clinical context, usually in the form of organized and unstructured clinical data from the electronic health record (EHR), is essential for precise and clinically meaningful medical imaging interpretation. There are innumerable examples of this throughout various medical fields. Similar to human doctors, more clinically relevant and higher-performing models may result from automated detection and classification systems that can effectively combine clinical data from the EHR, such as patient demographics, past diagnoses, and laboratory results, with medical imaging data.

2. Multimodal deep learning models

Applications outside of medicine, such autonomous driving and video categorization, have seen success using multimodal deep learning models, which are able to assimilate pixel input along with other data types (fusion). For instance, a multimodal fusion detection system for autonomous cars is able to achieve notably higher accuracy (3.7% improvement) than a single-modal CNN detection model21 by combining visual features from cameras with data from Light Detection and Ranging (LiDAR) sensors. Comparably, a multimodal social media video classification pipeline that made use of both textual and visual characteristics raised the accuracy of the classification to 88.0%, significantly higher than single modality neural networks like Google's InceptionV3, which on the same task22 achieved an accuracy of 76.4%. Leveraging fusion methodologies for medical imaging is primarily driven by the goal to incorporate supplementary contextual information and overcome the limitations of image-only models, which is why the performance increases for these efforts not only mirror the rationale in medical applications.

Similar trends are seen in the most current work on medical imaging, where pixel and EHR data are combined in a "fusion-paradigm" to solve complicated problems that are difficult for one modality to handle alone. The terminologies and model architectures used in the new fusion paradigm encompass a broad spectrum of approaches and strategies that have not been thoroughly examined. This review study aims to define and compile pertinent terminology, provide an extensive analysis of deep learning models that utilize various modalities for medical imaging tasks, and provide an overview of the outcomes of state-of-the-art models in pertinent current literature. Our evaluation aims to provide insights for upcoming modeling frameworks and act as a guide for investigators exploring the use of multimodal fusion in medical imaging.

This review aims to compile the body of knowledge from earlier research using multimodal deep learning fusion algorithms, which fuse clinical data with medical imaging. We provide standard language for multimodal fusion methods and group previous research according to fusion tactics. In general, we discovered that compared to conventional single modality models, multimodality fusion models produced higher accuracy (1.2–27.7%) and AUROC (0.02–0.16) for the identical task. But no single fusion approach produced the best results consistently in every category. It is advised to always experiment with fusion strategies when multimodal data is available, as our literature review demonstrates that more patient data and clinical context can lead to better model performance and that fusion methods better replicate the human expert interpretation workflow.

Reviewing deep learning fusion models, we found that they cover a wide range of medical applications, from hematology29 to radiology 31. Fusion techniques, for instance, were frequently used in the diagnosis and prognosis of Alzheimer's disease25,28,33,36,41. In clinical practice, diagnosing Alzheimer's disease requires more than just imaging or clinical evidence. Utilizing deep learning fusion techniques repeatedly demonstrated gains in performance for the diagnosis of Alzheimer's disease, despite the fact that histopathological correlation42 shows that doctors still have difficulty making accurate and trustworthy diagnoses even in the presence of multimodality. This demonstrates the value and practicality of multimodal fusion approaches in medical settings.

Even in cases where single modality models have been extensively reported to achieve excellent performance, like pixel-based models for automated skin cancer diagnosis, fusion techniques have increased performance above single modality models in other, less sophisticated clinical applications43. The continuous improvement in reported performance across a wide range of clinical use cases, despite the vast variations in the fusion approach, implies that model performance based on single-modal data may not represent state of the art for a given application when multimodal data are not taken into account.

In multimodal fusion studies, the complexity of the non-imaging data was constrained, especially when considering the availability of feature-rich and time-series data in the EHR. Rather, the majority of studies concentrated mainly on basic demographic data like age and gender25,27,39; a narrow range of categorical clinical history like smoking status or hypertension32,34; or disease-specific clinical features like APOE4 for Alzheimer's25,28,33,36; or PSA blood test for prostate cancer40. Even if it is important to choose traits that are known to be linked to disease, future research may profit even more by employing vast amounts of feature-rich data, as is the case in domains other than medicine like autonomous driving44,45.

3. Guidelines for implementing fusion models

Early fusion was typically employed as an initial attempt at multimodal learning, which is a simple method that doesn't always require training multiple models. However, high-level imaging features need to be retrieved as a 1D vector before fusing with the 1D clinical data when the input modalities are not in the same dimensions, which is common when merging clinical data represented in 1D with imaging data in 2D or 3D. This was achieved through a range of techniques, such as the use of software-generated features or manually extracted imaging features25,32,33,34, 35. It is important to remember that the outputs from a CNN's linear layers are typically accurate feature representations of the original image, unless there is a strong case for doing otherwise28,29,31. This is due to the fact that learned feature representations frequently produce task-specific performance that is significantly higher than that of manually or artificially generated features46. This review supports early fusion as a starting point for fusing multimodal data, as it consistently outperformed single modality models based on the examined publications.

The same CNNs that are used to extract characteristics from imaging modalities can also be applied to joint fusion. However, neural networks are used to execute joint fusion, which might be a drawback, particularly for smaller datasets that are better suited for conventional machine learning models. Early or late fusion is preferred, for instance, if there are disproportionately few samples compared to the number of features in the dataset, or if some of the input features are sparsely represented, as these scenarios can be handled with more appropriate traditional machine learning algorithms (like Lasso and ElasticNet47).

However, joint neural networks and early fusion neural networks can both acquire shared representations, which facilitates the model's ability to learn correlations across modalities and improves performance49. Additionally, research has demonstrated that combining less correlated information from deeper layers with highly correlated features from earlier layers enhances model performance50,51. Furthermore, we believe that joint fusion models can perform better than previous fusion strategies since the method simultaneously propagates the loss to all feature extracting models, updating its feature representations iteratively to better complement each modality. Although there isn't enough data to evaluate this effect on fusion for medical imaging yet, it's a crucial field for further research.

It is preferable to use a late fusion technique when signals from various modalities do not complement one other, that is, when input modalities independently influence the final prediction and do not have inherent interdependency. This is mostly because high-dimensional vectors are produced when feature vectors from many modalities are concatenated, as in early and joint fusion. Unless a large number of input samples are available, machine learning models may find it challenging to learn without overfitting. In machine learning, this is known as the "curse of dimensionality"52,53. By using many models, each specialized on a single modality, late fusion reduces this issue by restricting the amount of the input feature vectors for each model. For instance, the pixel data from a brain MRI (e.g., Qiu et al.41) and the quantitative outcome of a Mini Mental State Examination are both essentially independent data, making them good candidates for input into late fusion models.

4. Conclusion

In an ideal world, single modality models would be constructed and refined initially in order to act as baselines and as inputs for fusion models. Multimodal fusion in medicine is a promising yet emerging topic that supports the clinical practice of medical imaging interpretation across all disciplines, according to this systematic review. In order to inform task- and modality-specific strategies, we have clarified and compiled important terminology, methods, and assessed the current state of multimodal fusion in medical imaging. New fusion techniques are anticipated as the field of multimodal fusion for deep learning in medical imaging continues to grow. Future research should concentrate on common language and metrics, and when appropriate, it should directly assess various multimodal fusion techniques. For automated medical imaging tasks, we discovered that multimodal fusion generally outperforms single modality models in terms of performance. Future research may yield more insights that might guide the best course of action.

References

1. Leslie, A., Jones, A. J. & Goddard, P. R. The influence of clinical information on the reporting of CT by radiologists. *Br. J. Radiol.* **73**, 1052–1055 (2000).
2. Cohen, M. D. Accuracy of information on imaging requisitions: does it matter? *J. Am. Coll. Radiol.* **4**, 617–621 (2007).
3. Boonn, W. W. & Langlotz, C. P. Radiologist use of and perceived need for patient data access. *J. Digit. Imaging* **22**, 357–362 (2009).
4. Comfere, N. I. et al. Provider-to-provider communication in dermatology and implications of missing clinical information in skin biopsy requisition forms: a systematic review. *Int. J. Dermatol.* **53**, 549–557 (2014).
5. Comfere, N. I. et al. Dermatopathologists' concerns and challenges with clinical information in the skin biopsy requisition form: a mixed-methods study: Clinical information in dermatopathology. *J. Cutan. Pathol.* **42**, 333–345 (2015).
6. Jonas, J. B. et al. Glaucoma. *The Lancet* **390**, 2183–2193 (2017).
7. McDonald, R. J. et al. The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Acad. Radiol.* **22**, 1191–1198 (2015).
8. Dean, N. C. et al. Impact of an electronic clinical decision support tool for emergency department patients with pneumonia. *Ann. Emerg. Med.* **66**, 511–520 (2015).
9. Banerjee, I. et al. Development and Performance of the Pulmonary Embolism Result Forecast Model (PERFORM) for Computed Tomography Clinical Decision Support. *JAMA Netw. Open* **2**, e198719 (2019).
10. Sandeep Kumar, E. & Jayadev, P. S. Deep Learning for Clinical Decision Support Systems: A Review from the Panorama of Smart Healthcare. In *Deep Learning Techniques for Biomedical and Health Informatics. Studies in Big Data* (eds. Dash, S. et al.) vol. 68, 79–99 (Springer, Cham, 2020).
11. Hinton, G. Deep learning—a technology with the potential to transform health care. *JAMA* **320**, 1101 (2018).
12. Stead, W. W. Clinical implications and challenges of artificial intelligence and deep learning. *JAMA* **320**, 1107 (2018).

13. Dunnmon, J. A. et al. Assessment of convolutional neural networks for automated classification of chest radiographs. *Radiology* **290**, 537–544 (2019).

14. Irvin, J. et al. CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. *Proc. AAAI Conf. Artif. Intell.* **33**, 590–597 (2019).

15. Jaeger, S. et al. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant. Imaging Med. Surg.* **4**, 475–477 (2014).

16. Johnson, A. E. W. et al. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci. Data* **6**, 317 (2019).

17. Kallianos, K. et al. How far have we come? Artificial intelligence for chest radiograph interpretation. *Clin. Radiol.* **74**, 338–345 (2019).

18. Gulshan, V. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316**, 2402 (2016).

19. Rajpurkar, P. et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med.* **15**, e1002686 (2018).

20. Majkowska, A. et al. Chest radiograph interpretation with deep learning models: assessment with radiologist-adjudicated reference standards and population-adjusted evaluation. *Radiology* **294**, 421–431 (2020).

21. Person, M., Jensen, M., Smith, A. O. & Gutierrez, H. Multimodal fusion object detection system for autonomous vehicles. *J. Dyn. Syst. Meas. Control* **141**, 071017 (2019).

22. Trzcinski, T. Multimodal social media video classification with deep neural networks. In *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2018* (eds. Romaniuk, R. S. & Linczuk, M.) (SPIE, 2018).

23. Ramachandram, D. & Taylor, G. W. Deep Multimodal Learning: A Survey on Recent Advances and Trends. *IEEE Signal Process. Mag.* **34**, 96–108 (2017).

24. Kiela, D., Grave, E., Joulin, A. & Mikolov, T. Efficient large-scale multi-modal classification. In *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, (2018).

25. Thung, K.-H., Yap, P.-T. & Shen, D. Multi-stage Diagnosis of Alzheimer's Disease with Incomplete Multimodal Data via Multi-task Deep Learning. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (eds. Cardoso, M. J. et al.) vol. 10553, 160–168 (Springer International Publishing, 2017).

26. Kharazmi, P., Kalia, S., Lui, H., Wang, Z. J. & Lee, T. K. A feature fusion system for basal cell carcinoma detection through data-driven feature learning and patient profile. *Skin Res. Technol.* **24**, 256–264 (2018).

27. Yap, J., Yolland, W. & Tschandl, P. Multimodal skin lesion classification using deep learning. *Exp. Dermatol.* **27**, 1261–1267 (2018).

28. Li, H. & Fan, Y. Early Prediction Of Alzheimer's Disease Dementia Based On Baseline Hippocampal MRI and 1-Year Follow-Up Cognitive Measures Using Deep Recurrent Neural Networks. In *2019 IEEE 16th International Symposium on Biomedical Imaging* (ISBI 2019) 368–371 (IEEE, 2019).

29. Purwar, S., Tripathi, R. K., Ranjan, R. & Saxena, R. Detection of microcytic hypochromia using cbc and blood film features extracted from convolution neural network by different classifiers. *Multimed. Tools Appl.* **79**, 4573–4595 (2020).

30. Liu, Mingqian, Lan, Jun, Chen, Xu, Yu, Guangjun & Yang, Xiujun Bone age assessment model based on multi-dimensional feature fusion using deep learning. *Acad. J. Second Mil. Med. Univ.* **39**, 909–916 (2018).

31. Nie, D. et al. Multi-channel 3D deep feature learning for survival time prediction of brain tumor patients using multi-modal neuroimages. *Sci. Rep.* **9**, 1103 (2019).

32. Hyun, S. H., Ahn, M. S., Koh, Y. W. & Lee, S. J. A machine-learning approach using PET-based radiomics to predict the histological subtypes of lung cancer. *Clin. Nucl. Med.* **44**, 956–960 (2019).

33. Bhagwat, N., Viviano, J. D., Voineskos, A. N., Chakravarty, M. M. & Alzheimer's Disease Neuroimaging Initiative. Modeling and prediction of clinical symptom trajectories in Alzheimer's disease using longitudinal data. *PLOS Comput. Biol.* **14**, e1006376 (2018).

34. Liu, J. et al. Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. *Eur. Radiol.* **28**, 3268–3275 (2018).

35. An, G. et al. Comparison of Machine-Learning Classification Models for Glaucoma Management. *J. Healthc. Eng.* **2018**, 1–8 (2018).

36. Spasov, S. E., Passamonti, L., Duggento, A., Lio, P. & Toschi, N. A. Multi-modal Convolutional Neural Network Framework for the Prediction of Alzheimer's Disease. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 1271–1274 (IEEE, 2018).

37. Yala, A., Lehman, C., Schuster, T., Portnoi, T. & Barzilay, R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology* **292**, 60–66 (2019).

38. Yoo, Y. et al. Deep learning of brain lesion patterns and user-defined clinical and MRI features for predicting conversion to multiple sclerosis from clinically isolated syndrome. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **7**, 250–259 (2019).

39. Kawahara, J., Daneshvar, S., Argenziano, G. & Hamarneh, G. Seven-point checklist and skin lesion classification using multitask multimodal neural nets. *IEEE J. Biomed. Health Inform.* **23**, 538–546 (2019).

40. Reda, I. et al. Deep learning role in early diagnosis of prostate cancer. *Technol. Cancer Res. Treat.* **17**, 153303461877553 (2018).

41. Qiu, S. et al. Fusion of deep learning models of MRI scans, Mini–Mental State Examination, and logical memory test enhances diagnosis of mild cognitive impairment. *Alzheimers Dement. Diagn. Assess. Dis. Monit.* **10**, 737–749 (2018).

42. Beach, T. G., Monsell, S. E., Phillips, L. E. & Kukull, W. Accuracy of the Clinical Diagnosis of Alzheimer Disease at National Institute on Aging Alzheimer Disease Centers, 2005–2010. *J. Neuropathol. Exp. Neurol.* **71**, 266–273 (2012).

43. Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).

44. Hecker, S., Dai, D. & Van Gool, L. End-to-End Learning of Driving Models with Surround-View Cameras and Route Planners. In *Computer Vision – ECCV 2018* (eds. Ferrari, V. et al.) vol. 11211, 449–468 (Springer International Publishing, 2018).

45. Jain, A., Singh, A., Koppula, H. S., Soh, S. & Saxena, A. Recurrent Neural Networks for driver activity anticipation via sensory-fusion architecture. In *2016 IEEE International Conference on Robotics and Automation (ICRA)* 3118–3125 (IEEE, 2016).

46. Goodfellow, I., Bengio, Y. & Courville, C. *Deep Learning* (MIT Press, 2017).

47. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67**, 301–320 (2005).

48. Subramanian, V., Do, M. N. & Syeda-Mahmood, T. Multimodal Fusion of Imaging and Genomics for Lung Cancer Recurrence Prediction. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* 804–808 (IEEE, 2020).

49. Ngiam, J. et al. Multimodal Deep Learning. In *Proceedings of the 28th International Conference on Machine Learning (ICML)* 689–696 (2011).

50. Karpathy, A. et al. Large-Scale Video Classification with Convolutional Neural Networks. In *2014 IEEE Conference on Computer Vision and Pattern Recognition* 1725–1732 (IEEE, 2014).

51. Neverova, N., Wolf, C., Taylor, G. & Nebout, F. ModDrop: adaptive multi-modal gesture recognition. *IEEE Trans. Pattern Anal. Mach. Intell. 38*, 1692–1706 (2016).

52. Bach, F. Breaking the curse of dimensionality with convex neural networks. *J. Mach. Learn. Res. 1*8, 1–53 (2017).

53. Mwangi, B., Tian, T. S. & Soares, J. C. A review of feature reduction techniques in neuroimaging. *Neuroinformatics 1*2, 229–244 (2014).