NERDS 2025

OCT 9-10 2025

DEEP LEARNING FOUNDATION MODELS FOR MEDICAL IMAGE ANALYSIS

LIRA PI, JAKE GAGNON, CEDRIC LIU, PHOEBE JIANG

AGENDA

- Medical imaging background
- Overview of foundation models in medical imaging
- Self-supervised learning
- Adaptation of foundation models
- Foundation model applications
- Specific application to rare cancer detection
- Deep learning models for specific imaging modalities (MRI, PET, X-rays, CT, and fundus)
- Integration of multiple imaging modalities

Medical Image Analysis

Medical imaging modalities

MRI, PET, X-rays, CT, and fundus imaging are used in clinical practice and offers unique information for diagnosis

Common DL tasks

DL can aid clinicians/technicians with image segmentation, classification, reconstruction, and registration.

Biogen

FOUNDATION MODELS

Foundation models

- Large pretrained models learning from vast unlabeled data
- Learn general features which can be applied to specific clinical tasks
- Adapt model to specific tasks with minimal labeled data
- Help address data scarcity in medical imaging
- Backbones are commonly CNN or Vision Transformers (ViT)

Backbone comparison

CNN strengths

CNN can efficiently extract local features and are computationally efficient. They also incorporate parameter sharing making them easier to train.

CNN weaknesses

Receptive fields are local so CNNs struggle with long range dependencies. They also have difficulty to detect features at various scales due to fixed kernel sizes.

ViT strengths

They can capture long range dependencies and global context. ViTs can also analyze large scale patterns.

ViT weaknesses

ViTs are computationally expensive due to their self-attention mechanisms. They may be not as effective in capturing fine grained detail in images.

Self-Supervised Learning: Learning from unlabeled data

Discriminative Learning

Learns discriminative representations (ie the differences between instances)

Generative

Learns the data structure and underlying data distributions to generate or reconstruct data

Masked image modelling is where part of the image is masked, and the model is trained to reconstruct the missing information

Multimodal

Learning shared information across multiple modalities such as medical images and text reports. These models can be used for image to text or image retrieval.

Adapting Foundation Models to Clinical Tasks

Prompt Engineering

Guiding model behavior through prompt techniques

Linear Probing

Training a linear classifier on top of a foundation model’s embeddings

Task specific heads

Adding small task specific neural networks to frozen foundation model layers

PEFT: parameter efficient fine tuning

Updating a small number of the model’s parameters.

Full fine tuning

Updating all parameters to a given task

Overview of Clinical Applications of Foundation Models

Automated Image Interpretation	Foundation models enables more efficient interpretation and report generation
Patient Communication	Simplified medical reports can improve patient communication
Diagnostic support	Foundation models can aid in diagnosis by providing lesion segmentation and disease classification
Streamlined clinical workflows	Foundation models can aid in streamlining clinical workflows
Population understanding	Foundation models can yield population insights

Foundation models applied to pathology, radiology, and ophthalmology

Pathology

- Foundation models have been proposed to model global context in whole slide images
- These models are showing promise in disease detection and classification, automated reporting, and personalized medicine applications.
- However, there are challenges in data quality and model robustness.

Radiology

- Due to the abundance of unlabeled images, SSL can be applied to train models.
- Reduces need for expensive manual annotations
- **RadClip**: image classification and image/text matching
- **RadFM**: disease diagnosis, visual question answering, and report generation

Ophthalmology

- **RETFound**: disease diagnosis and prognosis in retinal images
- **VisionFM**: covers a variety of ophthalmologic diseases; offers disease screening and diagnosis, disease prognosis, and subclassification of disease

Veldhuizen et al. 2025 https://arxiv.org/abs/2506.09095

Benefits and challenges of FM applications in medical image analysis

Benefits

Improved diagnostic accuracy

Enhanced generalizability

Reduced annotation burden

Facilitating research and clinical applications

Challenges

Data heterogeneity

Model interpretability

Computational resources

Regulatory and ethical considerations

Rare cancer modelling:
Vorontsov et al [Nature Medicine Vol 30, October 2024, 2924–2935]

Pan cancer foundation model called Virchow

- 1.5 Million H&E **unlabeled** slides for training spanning multiple tissue types (breast, skin, lymph node, lung, etc); ~100,000 patients
- Model generates data representations that can generalize well to a variety of tasks
- Helpful for rare cancer detection that may not have sufficient samples to develop a model

Technical details

- ViT architecture with 632 million parameters
- H&E slide -> tissue tiles -> embeddings per tile -> aggregation
- Trained using the DINOv2 algorithm (self-supervised learning) to determine embeddings per tile
- Train an aggregator model to go from tile level to slide level [**required labeled data of 89,417 slides**] for pan cancer detection
- Trained a different aggregator model on labeled data for each biomarker prediction task

Results

- Pan cancer detection for common and rare cancers [Training dataset efficiency: less training data but almost matches clinical grade models]
- Biomarker prediction: good performance across multiple biomarker tasks with aucs ranging from 83% to 99% depending on the task
- May reduce need for immunohistochemistry

Biogen

DEEP LEARNING FOR SPECIFIC IMAGING MODALITIES

MRI Deep Learning Models

Goals

- Improve accuracy and efficiency for tumor detection, classification, segmentation, quantification, prognosis, and response prediction
- Image enhancement: noise reduction to enhance image quality, artifact suppression, and bias field correction
- Identify neurodegenerative changes in Alzheimer's and Parkinson’s
- DL accelerates MRI scans with image reconstruction
- Image registration: align MRI with histopathology images

Common DL techniques: attention, augmentation, and multi-modal models

MONAI

- Framework for deep learning in medical imaging
- Includes various MRI DL models for 3D segmentation of brain tumor subregions, segmentation of prostate zones in 3D MRI, and others

YOLOv7: Automate brain tumor detection

Deep Learning for PET images

Goals:

- Improve image quality, reduce radiation exposure, and automate lesion detection and segmentation
- Predict patients' response to treatment
- Classify tumor type
- DL for direct reconstruction and iterative reconstruction to tomographic representation
- Improve diagnostic efficiency and reduce radiologists' workload
- Personalized treatment

Architectures include GAN and UNET:

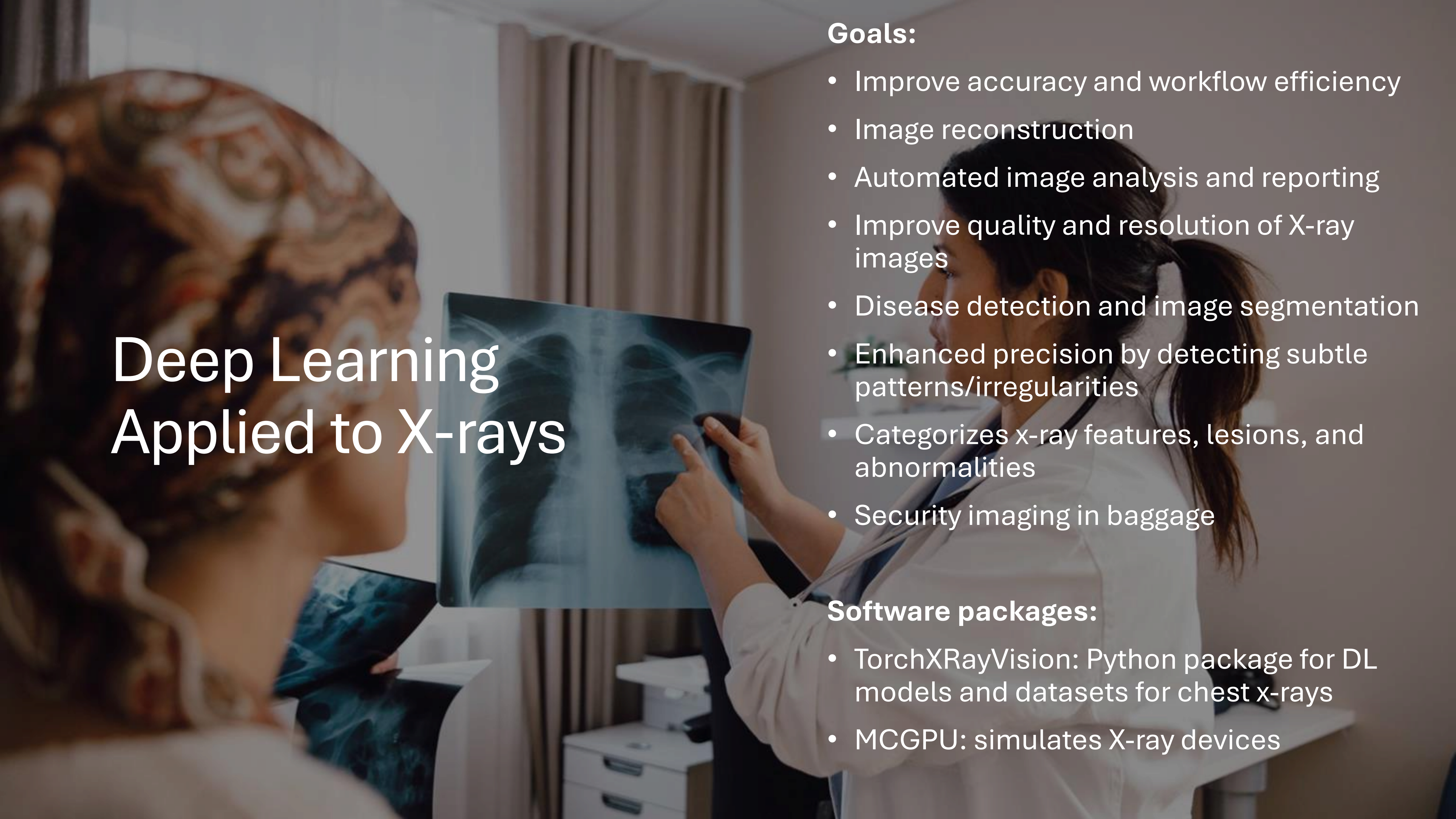
- Reduce noise of low dose PET
- Restore image quality
- Predict standard dose images from lose dose (tracer)

Packages:

- NiftyPET: python package for reconstruction, processing, and analysis of PET images
- PyTomography, CASToR, STIR, and MCGPU-PET

Recent work in multimodal integration

- PET + CT



Deep Learning Applied to X-rays

Goals:

- Improve accuracy and workflow efficiency
- Image reconstruction
- Automated image analysis and reporting
- Improve quality and resolution of X-ray images
- Disease detection and image segmentation
- Enhanced precision by detecting subtle patterns/irregularities
- Categorizes x-ray features, lesions, and abnormalities
- Security imaging in baggage

Software packages:

- TorchXRayVision: Python package for DL models and datasets for chest x-rays
- MCGPU: simulates X-ray devices

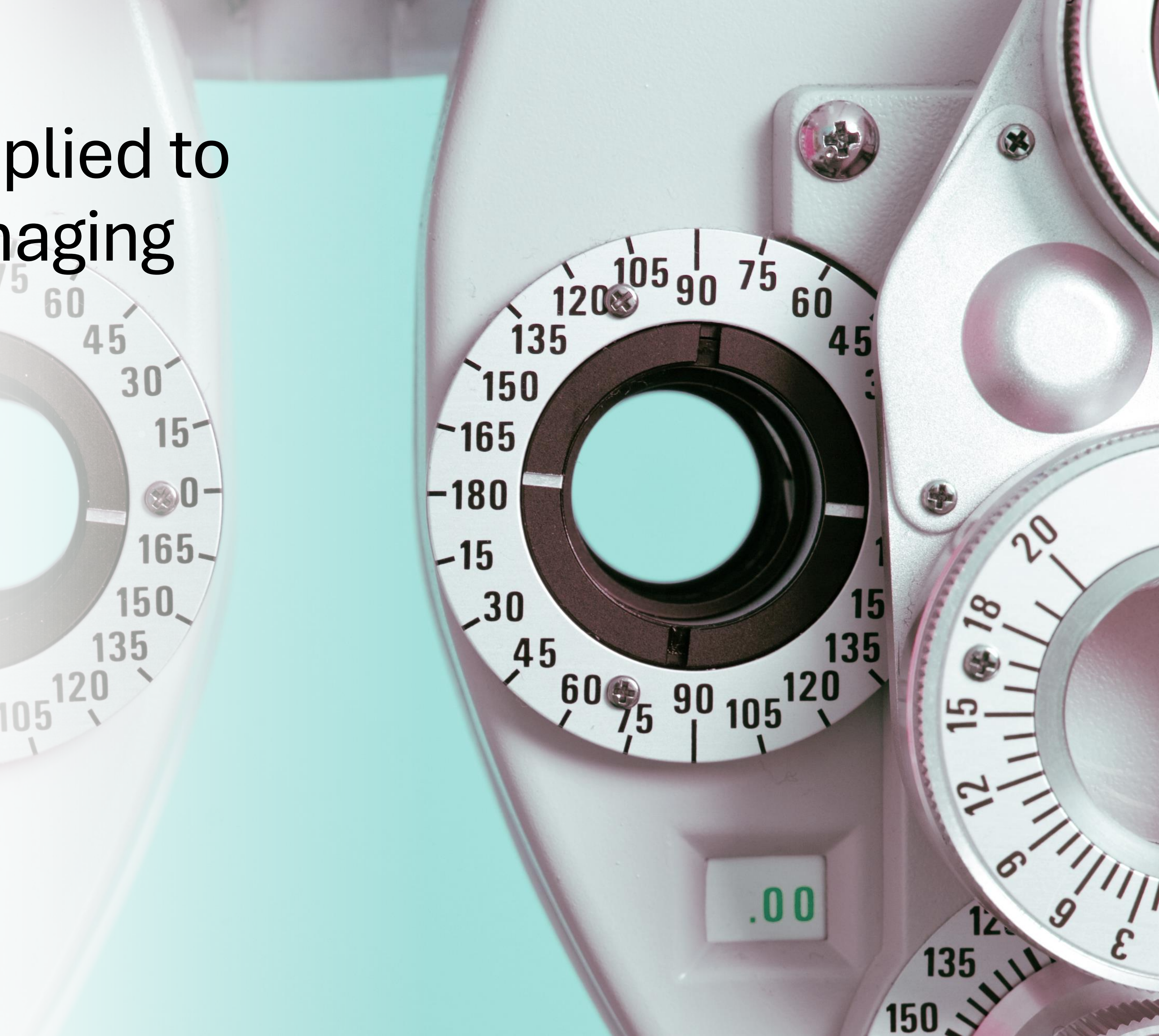
Deep Learning Applied to CT and Fundus Imaging

CT Imaging

- DL applied to tumor detection, organ segmentation, and disease classification
- Understanding of changes in CT scans over time
- Reconstruction of 3D CT scans from a smaller number of views

Fundus Imaging

- Early detection of neurodegenerative diseases
- Prediction of cardiovascular factors
- Detection of eye diseases and assignment of severity scores
- Grad-CAM to highlight fundus features



Biogen

INTEGRATION OF MULTIPLE IMAGING MODALITIES

Multi-modal Fusion

Integrating multiple imaging modalities or imaging + clinical/omics data

More comprehensive diagnosis

Aim for personalized medicine and prediction of patient outcomes

Target Diseases: cancer diagnosis and prognosis, neurodegenerative diseases, and cardiovascular diseases

Architectures: Transformers and graph convolutional networks

Software packages: MONAI, OpenMEDLab, and Molmo

Comparison of Fusion Techniques

	Early Fusion	Intermediate Fusion	Late Fusion
Concept	Multiple input modalities are combined before training a single machine learning model.	Different data modalities are first processed by individual models, before the extracted features are combined and fed into a final prediction model.	Different models are trained on separate data modalities, then the resulting predictions are merged via an aggregation function.
Advantages	Preserving the original information from each modality without losing substantial details. Simple model architecture and reduced computational complexity.	Capturing more complex interactions between modalities. Allowing separate processing pathways for different data types prior to fusion.	Easily dealing with missing data for patients.
Disadvantages	Potentially an imbalance of data richness from each modality.	Requiring a huge amount of data to precisely learn the additional interactions and combinations.	Impossible to model interactions and relationships between different modalities, with potential loss of information. Integration of results from different models will be complex.

Krones F et al, “Review of multimodal machine learning approaches in healthcare,” Information Fusion, Volume 114, 2025

Conclusions

Foundation models can aid in streamlining clinical workflows

Foundation models can aid in diagnosis by providing lesion segmentation and disease classification

Generalization of foundation models aids in rare cancer detection that may not have sufficient samples to develop a model

Challenges remain in areas of data heterogeneity, model interpretability, computational resources, and regulatory and ethical considerations

Non foundational DL models showed many clinical applications for MRI, PET, X-ray, CT, and fundus imaging

Multi modal fusion aims for a more comprehensive diagnosis and personalized medicine applications

QUESTIONS?

References

- Veldhuizen, Vivien & Botha, Vanessa & Lu, Chunyao & Erdal Cesur, Melis & Groot Lipman, Kevin & Jong, Edwin & Horlings, Hugo & Sanchez, Clárisa & Snoek, Cees & Mann, Ritse & Marcus, Eric & Teuwen, Jonas. (2025). Foundation Models in Medical Imaging -- A Review and Outlook. 10.48550/arXiv.2506.09095.
- Chaoyi Wu, Xiaoman Zhang, Ya Zhang, Yanfeng Wang, and Weidi Xie. Towards Generalist Foundation Model for Radiology by Leveraging Web-scale 2D&3D Medical Data, November 2023. URL http://arxiv.org/ abs/2308.02463.
- Zhixiu Lu, Hailong Li, and Lili He. Radclip: Enhancing radiologic image analysis through contrastive languageimage pre-training. arXiv preprint arXiv:2403.09948, 2024
- Yukun Zhou, Mark A. Chia, Siegfried K. Wagner, Murat S. Ayhan, Dominic J. Williamson, Robbert R. Struyven, Timing Liu, Moucheng Xu, Mateo G. Lozano, Peter Woodward-Court, Yuka Kihara, Andre Altmann, Aaron Y. Lee, Eric J. Topol, Alastair K. Denniston, Daniel C. Alexander, and Pearse A. Keane. A foundation model for generalizable disease detection from retinal images. Nature, 622(7981):156–163, October 2023. ISSN 1476-4687. doi:10.1038/s41586-023-06555-x.
- Jianing Qiu, Jian Wu, Hao Wei, Peilun Shi, Minqing Zhang, Yunyun Sun, Lin Li, Hanruo Liu, Hongyi Liu, Simeng Hou, Yuyang Zhao, Xuehui Shi, Junfang Xian, Xiaoxia Qu, Sirui Zhu, Lijie Pan, Xiaoniao Chen, Xiaojia Zhang, Shuai Jiang, Kebing Wang, Chenlong Yang, Mingqiang Chen, Sujie Fan, Jianhua Hu, Aiguo Lv, Hui Miao, Li Guo, Shujun Zhang, Cheng Pei, Xiaojuan Fan, Jianqin Lei, Ting Wei, Junguo Duan, Chun Liu, Xiaobo Xia, Siqi Xiong, Junhong Li, Benny Lo, Yih Chung Tham, Tien Yin Wong, Ningli Wang, and Wu Yuan. Visionfm: a multi-modal multi-task vision foundation model for generalist ophthalmic artificial intelligence, 2023.
- Dorfner, F.J., Patel, J.B., Kalpathy-Cramer, J. *et al.* A review of deep learning for brain tumor analysis in MRI. *npj Precis. Onc.* **9**, 2 (2025). https://doi.org/10.1038/s41698-024-00789-2
- Liu Z, Zhou X, Tao S, Ma J, Nickel D, Liebig P, Mostapha M, Patel V, Westerhold EM, Mojahed H, Gupta V, Middlebrooks EH. Application of Deep Learning Accelerated Image Reconstruction in T2-weighted Turbo Spin Echo Imaging of the Brain at 7T. AJNR Am J Neuroradiol. 2025 Jan 20:ajnr.A8662. doi: 10.3174/ajnr.A8662. Epub ahead of print. PMID: 39832954.
- Pulkit Khandelwal, Michael Tran Duong, Shokufeh Sadaghiani, Sydney Lim, Amanda E. Denning, Eunice Chung, Sadhana Ravikumar, Sanaz Arezoumandan, Claire Peterson, Madigan Bedard, Noah Capp, Ranjit Ittyerah, Elyse Migdal, Grace Choi, Emily Kopp, Bridget Loja, Eusha Hasan, Jiacheng Li, Alejandra Bahena, Karthik Prabhakaran, Gabor Mizsei, Marianna Gabrielyan, Theresa Schuck, Winifred Trotman, John Robinson, Daniel T. Ohm, Edward B. Lee, John Q. Trojanowski, Corey McMillan, Murray Grossman, David J. Irwin, John A. Detre, M. Dylan Tisdall, Sandhitsu R. Das, Laura E. M. Wisse, David A. Wolk, Paul A. Yushkevich; Automated deep learning segmentation of high-resolution 7 Tesla postmortem MRI for quantitative analysis of structure-pathology correlations in neurodegenerative diseases. *Imaging Neuroscience* 2024; 2 1–30. doi: https://doi.org/10.1162/imag_a_00171
- Gassenmaier S, Küstner T, Nickel D, Herrmann J, Hoffmann R, Almansour H, Afat S, Nikolaou K, Othman AE. Deep Learning Applications in Magnetic Resonance Imaging: Has the Future Become Present? Diagnostics (Basel). 2021 Nov 24;11(12):2181. doi: 10.3390/diagnostics11122181. PMID: 34943418; PMCID: PMC8700442.
- Abdusalomov AB, Mukhiddinov M, Whangbo TK. Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging. Cancers (Basel). 2023 Aug 18;15(16):4172. doi: 10.3390/cancers15164172. PMID: 37627200; PMCID: PMC10453020.
- Hashimoto, F., Onishi, Y., Ote, K. *et al.* Deep learning-based PET image denoising and reconstruction: a review. *Radiol Phys Technol* **17**, 24–46 (2024). https://doi.org/10.1007/s12194-024-00780-3
- Seyyedi, N., Ghafari, A., Seyyedi, N. *et al.* Deep learning-based techniques for estimating high-quality full-dose positron emission tomography images from low-dose scans: a systematic review. *BMC Med Imaging* **24**, 238 (2024). https://doi.org/10.1186/s12880-024-01417-y

Reference #2:

- Liu J, Ren S, Wang R, Mirian N, Tsai YJ, Kulon M, Pucar D, Chen MK, Liu C. Virtual high-count PET image generation using a deep learning method. Med Phys. 2022 Sep;49(9):5830-5840. doi: 10.1002/mp.15867. Epub 2022 Aug 13. PMID: 35880541; PMCID: PMC9474624.
- Artesani, A., Bruno, A., Gelardi, F. *et al.* Empowering PET: harnessing deep learning for improved clinical insight. *Eur Radiol Exp* **8**, 17 (2024). https://doi.org/10.1186/s41747-023-00413-1
- Illimoottil M, Ginat D. Recent Advances in Deep Learning and Medical Imaging for Head and Neck Cancer Treatment: MRI, CT, and PET Scans. Cancers (Basel). 2023 Jun 21;15(13):3267. doi: 10.3390/cancers15133267. PMID: 37444376; PMCID: PMC10339989.
- Maryam Fallahpoor, Subrata Chakraborty, Biswajeet Pradhan, Oliver Faust, Prabal Datta Barua, Hossein Chegeni, Rajendra Acharya, Deep learning techniques in PET/CT imaging: A comprehensive review from sinogram to image space, Computer Methods and Programs in Biomedicine, Volume 243, 2024, 107880, ISSN 0169-2607, https://doi.org/10.1016/j.cmpb.2023.107880.
- Lee J, Ha S, Kim REY, Lee M, Kim D, Lim HK. Development of Amyloid PET Analysis Pipeline Using Deep Learning-Based Brain MRI Segmentation-A Comparative Validation Study. Diagnostics (Basel). 2022 Mar 2;12(3):623. doi: 10.3390/diagnostics12030623. PMID: 35328176; PMCID: PMC8947654.
- Pawan Kumar Mall, Pradeep Kumar Singh, Swapnita Srivastav, Vipul Narayan, Marcin Paprzycki, Tatiana Jaworska, Maria Ganzha, A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities, Healthcare Analytics, Volume 4, 2023, 100216, ISSN 2772-4425, https://doi.org/10.1016/j.health.2023.100216.
- Anderson, P.G., Tarder-Stoll, H., Alpaslan, M. *et al.* Deep learning improves physician accuracy in the comprehensive detection of abnormalities on chest X-rays. *Sci Rep* **14**, 25151 (2024). https://doi.org/10.1038/s41598-024-76608-2
- https://www.gehealthcare.com/insights/article/deep-learning-image-reconstruction-improving-iq-and-patient-outcomes-in-radiology#:~:text=AI%2Denabled%20X%2Dray%20image,variability%2C%20increasing%20throughput%20and%20capacity.
- https://signalprocessingsociety.org/publications-resources/blog/recent-advances-deep-learning-within-x-ray-security-imaging#:~:text=The%20previous%20models%20had%20certain,the%20SIXray%20and%20GDXray%20datasets.
- https://cdrh-rst.fda.gov/mcgpu-gpu-accelerated-monte-carlo-x-ray-imaging-simulator#:~:text=MCGPU%20%5B1%2C2%5D%20is,virtual%20imaging%20trial%20project7.
- https://my.clevelandclinic.org/health/diagnostics/fundus-photography
- https://www.mdpi.com/2306-5354/12/1/57
- https://www.sciencedirect.com/science/article/abs/pii/S0161642018321857
- https://www.nature.com/articles/s41551-018-0195-0
- https://www.sciencedirect.com/science/article/pii/S0161642019318755
- https://ieeexplore.ieee.org/abstract/document/8359118

Reference #3:

- Paverd, H., Zormpas-Petridis, K., Clayton, H. *et al.* Radiology and multi-scale data integration for precision oncology. *npj Precis. Onc.* **8**, 158 (2024). https://doi.org/10.1038/s41698-024-00656-0
- Mohsen F, Ali H, El Hajj N, Shah Z. Artificial intelligence-based methods for fusion of electronic health records and imaging data. Sci Rep. 2022 Oct 26;12(1):17981. doi: 10.1038/s41598-022-22514-4. PMID: 36289266; PMCID: PMC9605975.
- Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, Vaidya AJ, Chen C, Zhuang L, Williamson DFK, Shaban M, Chen TY, Mahmood F. Artificial intelligence for multimodal data integration in oncology. Cancer Cell. 2022 Oct 10;40(10):1095-1110. doi: 10.1016/j.ccell.2022.09.012. PMID: 36220072; PMCID: PMC10655164.
- https://www.bentoml.com/blog/multimodal-ai-a-guide-to-open-source-vision-language-models
- S. Dwivedi, T. Goel, M. Tanveer, R. Murugan and R. Sharma, "Multimodal Fusion-Based Deep Learning Network for Effective Diagnosis of Alzheimer's Disease," in *IEEE MultiMedia*, vol. 29, no. 2, pp. 45-55, 1 April-June 2022, doi: 10.1109/MMUL.2022.3156471.
- Hyun, Seung Hyup MD, PhD∗; Ahn, Mi Sun MD†; Koh, Young Wha MD, PhD‡; Lee, Su Jin MD, PhD§. A Machine-Learning Approach Using PET-Based Radiomics to Predict the Histological Subtypes of Lung Cancer. Clinical Nuclear Medicine 44(12):p 956-960, December 2019. | DOI: 10.1097/RLU.0000000000002810
- Baltruschat, I.M., Nickisch, H., Grass, M. *et al.* Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. *Sci Rep* **9**, 6381 (2019). https://doi.org/10.1038/s41598-019-42294-8
- S. E. Spasov, L. Passamonti, A. Duggento, P. Liò and N. Toschi, "A Multi-modal Convolutional Neural Network Framework for the Prediction of Alzheimer's Disease," *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, USA, 2018, pp. 1271-1274, doi: 10.1109/EMBC.2018.8512468.
- Shangran Qiu, Gary H. Chang, Marcello Panagia, Deepa M. Gopal, Rhoda Au, Vijaya B. Kolachalama, Fusion of deep learning models of MRI scans, mini–mental state examination, and logical memory test enhances diagnosis of mild cognitive impairment, in: Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, vol. 10, Elsevier, 2018, pp. 737–749.
- Jinzhao Zhou, Xingming Zhang, Ziwei Zhu, Xiangyuan Lan, Lunkai Fu, Haoxiang Wang, Hanchun Wen, Cohesive multi-modality feature learning and fusion for COVID-19 patient severity prediction, IEEE Trans. Circuits Syst. Video Technol. 32 (5) (2021) 2535–2549.
- Islam Reda, Ashraf Khalil, Mohammed Elmogy, Ahmed Abou El-Fetouh, Ahmed Shalaby, Mohamed Abou El-Ghar, Adel Elmaghraby, Mohammed Ghazal, Ayman El-Baz, Deep learning role in early diagnosis of prostate cancer, Technol. Cancer Res. Treat. 17 (2018) 1533034618775530.