Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 2 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

ENHANCED DETECTION PARADIGMS: INTEGRATING AI AND DATA SCIENCE IN CT SCAN FOR EARLY DISEASE DIAGNOSIS

Syed Muhammad Ayan Naqvi
Department of Biomedical Engineering, Ziauddin University (FESTM), Karachi,74200, Pakistan.

Ansharah Hasib
Department of Biomedical Engineering, Ziauddin University (FESTM), Karachi,74200, Pakistan.

Mehwish Faiz
Department of Biomedical Engineering, Ziauddin University (FESTM), Karachi,74200, Pakistan.

Fahad Shamim
IBET, Liaquat University of Medical and Health Sciences, Jamshoro, Pakistan

Saba Yosusha
Department of Computer Information & Engineering, Mehran University of Engineering & Technology, Jamshoro, Sindh, Pakistan

Shahzad Nasim
Department of Management Sciences & Technology, The Begum Nusrat Bhutto Women University, Sukkur, Pakistan

[*]***Corresponding author:*** ***mehwish.faiz@zu.edu.pk***

Article Info

Abstract

Biomedical imaging methods such as computed tomography (CT) or magnetic resonance imaging (MRI) are important appliances for diagnostic--disease detections. However, the conventional approach of using imaging for analysis largely hinges on radiologist expertise and manual interpretation is often not able to detect early cases. In this review, we present the paradigm shift data science has introduced into biomedical imaging and how machine learning — especially in combination with deep learning algorithms— is facilitating earlier and more precise disease detection. An automatically image analysis can allow CT and MRI scans to become a higher level of codable relevance with the refinement for automating precision while detecting abnormalities that might be overlooked conventionally.

Using machine learning models trained on thousands, if not millions of medical images could potentially improve the resolution and clarity of scans so that patterns indicative diseases like cancer, cardiovascular conditions or neurological disorders can more easily be identified. Imaging techniques using data science are more sensitive and specific as compared to traditional methods making it possible for early interventions, increasing the life expectancy of patients. Still, issues regarding data privacy and ownership, algorithmic bias as well as ethical constraints need to be addressed before these technologies can safely and equitably established in the clinical routine. This narrative review enumerates significant trends in applications of data science to CT scans through a comparative examination with image analysis by traditional, semi-automated and automated methods. Central to this are advances in predictive diagnostics and AI-driven innovations that hold great promise for improving the early disease detection enabling of biomedical imaging role in overall healthcare delivery, this article likewise explores future directions including their integration.

Introduction

1.1. Overview of Biomedical Imaging and CT scans

Biomedical imaging array several techniques used for visualizing biological structures and functions, among them computed tomography (CT) is a central one. CT scans take various X-ray images from different angles to produce a detailed, 3-D image of the body. This approach is critical in managing serious diseases like trauma and cancer, where quick imaging with high accuracy required. Others, such as MRI, ultrasound and nuclear medicine are used on indications to serve specific clinical purposes in distinctive pathophysiological bases (Computed Tomography in Medical Imaging", 2022; "Biomedical Imaging", 2023; Zhou & Liu, 2024).

The medical fields with high demand of clinical CT applications include trauma assessment, stroke diagnosis and cancer management as well as many others where diagnostic information leads to important decision making. Biomedical imaging is able to expand its diagnostic potential leveraging advanced techniques like functional and molecular imaging, detecting disease in early stages with a direct impact on treatment. Nonetheless, the dangers of radiation exposure related to CT scans have encouraged further investigation for less hazardous approaches and models ("Computed Tomography in Medical Imaging", 2022; Capusten, 2022).

1.2. Role of Data Science in Modern Healthcare

Data science is revolutionizing modern healthcare by employing advanced analytical techniques to improve patient outcomes, enhance operational efficiency, and combat fraud. The application of machine learning and statistical analysis allows healthcare providers to assess treatment effectiveness and identify risk factors for various diseases. Additionally, visualization tools aid in interpreting complex datasets, fostering clearer communication of insights and trends among healthcare professionals (G. et al., 2024). Predictive analytics further streamlines operations by optimizing resource allocation and patient flow, thereby increasing productivity and reducing costs through data-driven decision-making (Russia et al., 2024).

Moreover, another edge of data science is the anomaly detection, pattern recognition that help in detecting fraudulent activities. Machine learning in real time, also embedded with data science to combat cyber threats and thus integrity of patient records is deeply ingrained into hardware including health system (Nayak et al., 2024). While the advantages are huge, potential roadblocks such as privacy concerns around data and low human capital needs to be dealt with for utilizing this transformative force of healthcare utility in conjunction with data science.

1. Constraint of traditional CT scan methods

The data science research using the traditional CT scan method faces some challenges in preparing and integrating imaging datasets for different DL applications. One of the largest confinements in the authenticity of machine learning algorithms comes from lack of public access to comprehensive CT datasets at its best or raw data, which in most cases is limited. Moreover, diversity in image orientations, types and dimensions hinder standardization that deep learning requires so as to use effectively the CT data (Fontanella et al., 2023).

In addition, these traditional deep learning frameworks even have compatibility problems. The inconsistencies of model performances were attributed by slight variations in a number of slices and the resolutions because of the different models of CT machines. Additionally, slice-wise feature learning is not possible with conventional approaches, which lowers the accuracy of classification tasks based on images ("Strong Baseline and Bag of Tricks for COVID-19 Detection from CT Scan Images", 2023).

In addition, combining data from multiple imaging platforms is a major challenge when suitable annotations required for algorithm training are not given. Moreover, the heterogeneity in high-resolution images and diversity of imaging modalities make it even harder to develop robust automated analysis methods (Yaffe 2019). Nevertheless, recent advances in deep learning and data reconstruction methods indicate potential gains for the future of biomedical imaging.

2. Data Science in CT scans

3.1. Methods and Algorithms of Automated Image Analysis in CT scans

Automated CT image analysis using machine learning tools is beneficial, especially in tracking lesions or for functional purposes which enhances diagnostic efficacy. In this regard, an algorithm is proposed by Mukherjee et al. (2024) for automated lesion correspondence. This method employs a two-step pipeline that first performs optimal lesion matching using an adaptive Hungarian algorithm and exploits the 3D rigid registration for efficient tracking over multiple time points. This substantially reduces errors and increases the accuracy of monitoring metastatic breast cancer lesion.

Deep learning has been used for CT imaging (Correa et al. 2024) and was found effective in assessing split kidney function. This work highlights the promise of automated evaluations to enhance decision-making in procedural urology and important considerations for integrating artificial intelligence into routine imaging pipelines. This progression further demonstrates how deep learning methods are gradually producing more sophisticated results that formerly required manual inspection. This development emphasizes even more the mechanization of software tools that were previously focused on manual evaluation. The more sophisticated findings we used to make under someone's careful supervision may now be found by applying a machine learning technique.

Despite its obvious potential regarding the scalability of automated methods for medical imaging, there remain concerns about standardizing these approaches across various clinical environments. As the scale and complexity of automated image analysis methods continues to grow (Brown, Irvine et al., 2022), it has been argued that an appropriate benchmark is needed for machine learning tasks aimed at clinical practice. In order to address these shortcomings and more fully leverage the potential for AI-enabled diagnosis, this work must be validated in future research.

3.2. Deep Learning and Machine learning in CT scan applications

The advances made in CT scan analysis by both machine learning (ML) and deep learning (DL), have not only greatly enhanced the diagnostic accuracy, but also significantly improved its efficiency across multiple domains of medicine. These technologies allow for automated identification, segmentation, and reconstruction of images to help facilitate quicker diagnoses with better patient results. For instance, a colonic tumor classification system by 3-phase is proposed by deep learning models with an accuracy of 95.83%, and practical value has been demonstrated in the early detection and staging of cancer (Ghosal et al.,2024).

CT images and effect of quality Innovations in thoracic CT imaging by models including Cascade-SwinUNETR and DVSR3D were also validated for the bone health assessments (Ge et al. 2024). These tools can automatically segment more complex structures such as ascites from abdominal scans extremely well, performing similarly to experts and achieving Dice scores over 0.8 compared with expert annotations

(Hou et al., 2024). Such a high level of precision reduces manual intervention and ensures consistent and objective measures.

Moreover, unsupervised learning approaches in low-dose CT imaging contribute to reducing radiation exposure without affecting image quality, essential for patient safety. Although these models have shown good performance, the main limitations coming from model interpretability as well as generalization on various populations still hinder their clinical adoption (Wang et al., 2024).

3. CT imaging improvements driven by Data science

4.1. Better Image Quality and Resolution

Machine-learning algorithms applied to CT imaging have now enhanced image quality and produced a higher resolution in low-dose circumstances with the help of data science. These include deep learning and generative models such as CycleGAN, which solved the problems presented by lower resolution images with decreased radiation exposure. Recently, cycleGAN has been employed to denoise liver CT images in an unsupervised manner which converts low-dose CTS into high dose like intensity without using any paired datasets thereby presenting optimistic results for clinical applications (Sowjanya et al., 2023). New in image reconstruction from low-dose scans, deep learning has further improved this process with specific studies using the LIDC-IDRI dataset showing that model-driven techniques combined together with DL reduced errors when compared to ground-truth images (Biguri & Mukherjee 2024).

Generative Adversarial Networks (GANs) have also boosted super-resolution, producing higher Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index over conventional techniques (Wang et al., 2024). A reference-free deep learning method has also been proposed to generate a high-resolution image from almost no low-resolution input (Ye et al., 2023).

4.2. Automated Detection and Diagnosis of Disease

Data science has improved CT imaging making automated diagnosis of disease even better, as is being seen with conditions like COVID-19. Deep learning, transfer learning, and advanced algorithms have improved diagnostic accuracy, with deep convolutional neural networks (CNNs) achieving up to 99.60% accuracy in COVID-19 detection using CycleGAN for data augmentation ("Automatic diagnosis of COVID-19 from CT images using CycleGAN and transfer learning", 2023). Similarly, a combination of U-Net and CNN models has reached 98.45% accuracy through optimized algorithms (Haennah et al., 2024). Noise reduction in low-dose CT (LDCT) imaging has also seen progress, with advanced denoising techniques enhancing image quality while preserving diagnostic details (Zhang et al., 2024). Transfer learning approaches, like those using ConvNeXt, have further streamlined anomaly detection, allowing model updates without retraining (Nizamli et al., 2024). However, challenges like overfitting and the need for ongoing model updates persist in automated diagnostic systems.

4.3. Early Detection Patterns (Biomarkers — Tumors, Cancers & Vascular Diseases)

CT imaging is one of the better-known examples where advances benefiting early detection — tumors, cancer and vascular diseases for example — have been heavily influenced by data science developments. The amalgamation of machine and deep learning in diagnostic health care makes possible the discipline on CT scans to pick out patterns more reliably. Methods have been developed to improve the consistency of quantitative imaging features by enhancing contrast-enhancement phase classification in CT scans using machine learning algorithms (Guha et al. 2024). Similarly, deep learning models specifically convolutional neural networks have improved the detection of pulmonary tumors and shows AI can process complex imaging data (Thool et al. 2023). In addition, radiomics has emerged as an essential method in cancer research providing detailed image-based features that may not be visible to the human eye and correlating CT patterns with different lung cancer types for prognostic prediction or treatment planning (Ciofiac et al., 2024). Nevertheless, obstacles such as demands for annotated datasets and risks of model overfitting exist as obstacles for further advancements in CT imaging techniques.

4. Comparative Analysis: Traditional vs. Data-Driven CT scans

Significant variations in performance indicators, including accuracy, sensitivity, and specificity, are shown by comparing data-driven and conventional CT scans. In contrast to the classical method, significant improvements have been made by data-driven approaches especially using reconstruction techniques. The learned primal-dual deep neural network (LPD-DNN) outperformed traditional filtered-back projection methods and was able to reach a normalized root mean square error of 12% and structural similarity index of 96% (Zarei et al., 2022). For pathologies of skull the traditional CT had a 72% sensitivity and specificity of 78%, whereas for similar pathologies MRI showed better performance with sensitivity rate at 87 % and high specificity to be around92% (Iqbal et al.,2023). This indicates that additional data-driven improvements could increase the utility of CT scans.

There are important clinical implications of these findings. While a classic CT scan is known for providing higher radiation doses, there are new data-specific approaches that could lead to optimal doses without compromising image quality from 25–60% lower dosage (Malik et al., 2024). While CT continues to be the gold standard for diagnosis in acute stroke with a sensitivity of 97.6% (for infarctive strokes) -data-driven methods can augment diagnostic capability, particularly when access to CT is limited (Anwar et al., 2023). Therefore, even though the traditional CT scan is an indispensable tool in clinical diagnosis and management of patients through various types of care (eg., screening, diagnostic studies), machine learning integration can help ensure imaging more accurately addresses both tradition practices will improve patient safety.

5. Challenges in Data-Driven CT Imaging: Balancing Privacy, Bias, and Security

While there are tremendous benefits for using AI with data-driven CT scans, challenges and ethical concerns also exist when blending the two — especially in terms of data privacy, security and algorithmic bias. These concerns must be resolved to ethically implement AI in medical imaging. Given the sensitivity of patient information — which is increasingly at risk for data breaches that can erode trust and expose exploitable points of entrance into a larger cyber network, privacy concerns rightly come to mind (Shukla & Taneja 2024; Acosta et al.,2023). These frameworks should aim to follow ethical standards of AI applications and assign consent by user freedom as the top priority among data protection principles (Valli et al., 2024).

There is also an unpreventable risk to vulnerable groups of the populations from algorithmic bias, which can guarantee in equitable public healthcare outcomes (Herington et al. 2023). It is key to maintain transparency in the process of model training and validation, which plays a vital role in reinforcing fairness that paves way for developing accountable AI systems [Valli et al., 2024]. Indeed, as AI-driven medical technologies become increasingly prevalent within healthcare systems worldwide, the ethical frameworks will need to evolve in order for these biases to be dealt with and ensure that they are both inclusive and fair (Herington et al. 2023). While these challenges are daunting, they also represent opportunities to advance ethical AI practices in order that protect patient rights and ensure equitable healthcare delivery.

6. Future Recommendation

The future of CT imaging driven by data science comes with great strides in both accuracy and efficiency. These include computationally efficient data augmentation, deep learning-based reconstruction techniques to enhance the quality of images and replace erroneous regions in challenging MRI scans specifically for resource-constrained settings.

New technologies including photon-counting CT and AI not only continue to improve diagnostics in terms of resolution but also at the level of decision-making speed. Nevertheless, it will be essential that healthcare disparities do not emerge as a result of this enhanced technology and this can be done by ensuring equitable access to these advancements

7. Conclusion

Biomedical imaging, especially CT and MRI scans have received tremendous benefit from data science today where early automatic image analysis can help in the detection of diseases more effectively. These techniques have improved diagnostic sensitivity and specificity over conventional methods through the use of machine learning/deep learning algorithms. Where automated systems are deployed to recognize patterns, there may be an efficient provision of timely interventions in the imaging. Data privacy, algorithmic bias and ethical considerations are some of the challenges that needs to be addressed though. AI-driven advances will continue to expand and, as they do, so too does their potential for diagnosing conditions effectively in better patient outcomes and advancing the future of healthcare.

Conflict of Interest

There is no conflict of interest to be reported by the authors in this study.

References:

- Acosta, L., Hosseini, S. E., & Pervez, S. (2023). Ethical challenges associated with security vulnerabilities and data privacy in social networking. https://doi.org/10.1109/dese60595.2023.10469464
- Automatic diagnosis of COVID-19 from CT images using CycleGAN and transfer learning. (2023). *Applied Soft Computing*. https://doi.org/10.1016/j.asoc.2023.110511
- Biomedical Imaging. (2023). https://doi.org/10.1002/9781119819028.ch2
- Brown, P. A., & Irvine, J. M. (2022). Utilization of artificial intelligence in automated image analysis. https://doi.org/10.1109/AIPR57179.2022.10092210
- Capusten, B. (2022). Biomedical imaging. *Nature Biotechnology*. https://doi.org/10.1148/radiology.163.3.644
- Ciofiac, C. M., Mamuleanu, L. L., Florescu, L. M., & Gheonea, I. A. (2024). CT imaging patterns in major histological types of lung cancer. https://doi.org/10.3390/life14040462
- Computed Tomography in Medical Imaging. (2022). https://doi.org/10.1201/978100313255
- Correa, R., Jeong, J., Patel, B. N., Banerjee, I., & Abdul-Muhsin, H. (2024). Automated analysis of split kidney function from CT scans using deep learning & delta radiomics. *Journal of Endourology*. https://doi.org/10.1089/end.2023.0488
- Fontanella, A., Li, W., Mair, G., Antoniou, A., Platt, E. J., Martin, C., Armitage, P. A., Trucco, E., Wardlaw, J. M., & Storkey, A. (2023). Challenges of building medical image datasets for the development of deep learning software in stroke. https://doi.org/10.48550/arxiv.2309.15081
- Ge, P., Li, S., Liang, Y., Zhang, S., Zhang, L., Hu, Y., Yao, L., & Wong, P. K. (2024). Enhancing trabecular CT scans based on deep learning with multi-strategy fusion. *Computerized Medical Imaging and Graphics*. https://doi.org/10.1016/j.compmedimag.2024.102410
- Ghosal, S., Das, D., Kumar, J., Pandaw, A. S., & Verma, S. (2024). iScan: Detection of colorectal cancer from CT scan images using deep learning. *ACM Transactions on Computing for Healthcare*. https://doi.org/10.1145/3676282
- Guha, S., Abdalla, I., Wu, Q., Geng, P., Chou, Y., Yang, H., Ma, J., Lu, L., Wang, D., Schwartz, L. H., Xie, C.-M., & Zhao, B. (2024). Machine learning-based identification of contrast-enhancement phase of computed tomography scans. *PLOS ONE*. https://doi.org/10.1371/journal.pone.029458
- Haennah, J., Christopher, G. R., & King, G. (2024). Combined U-Net and CNN image classification model for COVID disease detection using CXR/CT imaging. https://doi.org/10.3233/jifs-230523
- Herington, J., McCradden, M. D., Creel, K., Boellaard, R., Jones, E. C., Jha, A. K., Rahmim, A., Scott, P. J. H., Sunderland, J. J., Wahl, R. L., Zuehlsdorff, S., & Saboury, B. (2023). Ethical considerations for artificial intelligence in medical imaging: Data collection, development, and evaluation. *The Journal of Nuclear Medicine*. https://doi.org/10.2967/jnumed.123.266080
- Hou, B., Lee, S., Lee, J.-M., Koh, C., Xiao, J., Pickhardt, P. J., & Summers, R. M. (2024). Deep learning segmentation of ascites on abdominal CT scans for automatic volume quantification. *Radiology*. https://doi.org/10.1148/ryai.230601
- Iqbal, A., Ul Khursheed, S., Ibrahim, M., Ali, R., Munir, S., Ullah Khan, T., & Lodhi, K. (2023). Comparative analysis of diagnostic accuracy between MRI and CT scans in detecting skull base ENT pathologies. *CARDIOMETRY*. https://doi.org/10.18137/cardiometry.2023.29.10311
- Mukherjee, S., Coroller, T., Wang, C. K., Ravi, S., Samala, T., Hu, T., Gökçay, D., Petrick, N., Berkman, S., Sahiner, Q., & Cao, Q. (2024). Image registration-based automated lesion correspondence pipeline for longitudinal CT data. https://doi.org/10.48550/arxiv.2404.16544

- Narayanan Valli, L., Sujatha, N., Mech, M., & V, S. L. (2024). Ethical considerations in data science: Balancing privacy and utility. *International Journal of Science and Research Archive*. https://doi.org/10.30574/ijsra.2024.11.1.1098
- Nayak, R., Ghugar, U., Gupta, P., & Dash, S. (2024). Role of data science and cybersecurity in fighting healthcare fraud. In *Advances in Healthcare Information Systems and Administration* (pp. 81–97). https://doi.org/10.4018/979-8-3693-7457-3.ch004
- Ratthna, J. S., Karthikeyan, G. J., & Priyadharshini, P. (2024). Advancing healthcare through data science techniques for comprehensive analysis and visualization of healthcare data. In *Advances in Healthcare Information Systems and Administration* (pp. 1–15). https://doi.org/10.4018/979-8-3693-7457-3.ch001
- Russia, S., Murugan, K., & Vimaladevi, M. (2024). The role of data science advancements in healthcare, agriculture, and finance. https://doi.org/10.58532/v3bgct1p2ch2
- Shukla, R. P., & Taneja, S. (2024). Ethical considerations and data privacy in artificial intelligence. In *Advances in Educational Technologies and Instructional Design* (pp. 1–15). https://doi.org/10.4018/979-8-3693-2440-0.ch005
- Strong baseline and bag of tricks for COVID-19 detection of CT scans. (2023). https://doi.org/10.48550/arxiv.2303.08490
- Ud Din, M. M., Alqahtani, M., Hadadi, I., AlQhtani, A. G. M., & Alqarni, A. (2024). An analysis of computed tomography diagnostic reference levels in India compared to other countries. *Diagnostics*. https://doi.org/10.3390/diagnostics14151585
- Ul Anwar, M., Afrin, S., Hossain, S., Jahan, M. S., & Rahman, M. (2023). Sensitivity, specificity, predictive values, and accuracy of clinical diagnosis of acute stroke. *Journal of Rangpur Medical College*. https://doi.org/10.3329/jrpmc.v8i1.65038
- Umare, K. B., Thool, P. A., Wankhede, A. P., Yella, V. R., Tamijeselvan, S., Suganthi, D., & Rastogi, R. (2023). Artificial intelligence in medical imaging data analytics using CT images. https://doi.org/10.1109/icesc57686.2023.10192998
- Wang, F., Wang, R., & Qiu, H. (2024). Low-dose CT reconstruction using dataset-free learning. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0304738
- Yaffe, M. J. (2019). Emergence of "big data" and its potential and current limitations in medical imaging. *Seminars in Nuclear Medicine*. https://doi.org/10.1053/J.SEMNUCLMED.2018.11.010

- Yasser, A., Nizamli, A., Filatov, A., Fadel, W. F., & Shichkina, Y. (2024). Accurate anomaly detection in medical images using transfer learning and data optimization: MRI and CT as case studies. https://doi.org/10.1109/neuront62606.2024.10585603
- Zhang, W., Salmi, A., Yang, C., & Jiang, F. (2024). Innovative noise extraction and denoising in low-dose CT using a supervised deep learning framework. *Electronics*. https://doi.org/10.3390/electronics13163184
- Zhou, X., & Liu, Z. (2024). Computerized tomography. In X. Zhou & Z. Liu (Eds.). https://doi.org/10.1007/978-981-
- Zarei, M., Sotoudeh-Paima, S., Abadi, E., & Samei, E. (2022). A truth-based primal-dual learning approach to reconstruct CT images utilizing the virtual imaging trial platform. https://doi.org/10.1117/12.2613168