Deep Learning Model Training for Abdominal Aortic Aneurysm Detection Using CT Scans

Rasit Dinc

Rasit Dinc Digital Health & AI Research

Published: September 13, 2025 | AI in Healthcare

DOI: 10.5281/zenodo.17996409

Abstract

Explore the step-by-step deep learning training process for abdominal aortic aneurysm detection using CT scans, focusing on data handling, validation, and evaluation.

Deep Learning Model Training for Abdominal Aortic Aneurysm Detection Using CT Scans

Category: AI in Healthcare | Medical Imaging | Digital Health | Clinical Applications

Introduction

Abdominal Aortic Aneurysm (AAA) is a potentially life-threatening condition characterized by the abnormal dilation of the abdominal aorta. Early and accurate detection is crucial to prevent catastrophic rupture and improve patient outcomes. Computed Tomography (CT) imaging remains the gold standard for AAA diagnosis due to its high spatial resolution and detailed anatomical visualization. However, manual interpretation of CT scans is time-consuming and subject to interobserver variability.

Recent advances in artificial intelligence (AI), particularly deep learning, have revolutionized medical image analysis. Convolutional Neural Networks (CNNs), a subtype of deep learning architectures, have demonstrated superior ability to detect complex patterns in imaging data, making them ideal for AAA detection. This article details the comprehensive process of training a deep learning model for AAA detection using abdominal CT scans, emphasizing clinical relevance, methodological rigor, and future prospects.

Clinical Significance of AI-Driven AAA Detection

AAA affects approximately 1–2% of the general population over age 65, with a higher prevalence in men and smokers. Ruptured AAAs carry a mortality rate of up to 90%, underscoring the importance of early diagnosis and surveillance. Traditional AAA screening programs rely on ultrasound, but CT scans provide more precise assessment of aneurysm size and morphology, guiding surgical decisions.

Automated detection through AI can enhance radiologist workflow by flagging suspicious cases, reducing diagnostic delays, and standardizing interpretations. Additionally, deep learning models can quantify aneurysm dimensions and growth rates, aiding longitudinal monitoring and risk stratification. Integrating AI into clinical practice promises improved screening efficiency, cost-effectiveness, and patient safety.

Comprehensive Deep Learning Model Training Pipeline for AAA Detection

1. Data Collection and Preprocessing

- **Dataset Composition:** Assemble a large-scale, balanced dataset comprising 10,000 abdominal CT scans — evenly split between 5,000 AAA-positive and 5,000 AAA-negative cases. Data should be sourced from diverse demographics and multiple institutions to enhance generalizability. - **Annotation:** Expert radiologists must label the presence or absence of AAA and, where possible, provide segmentation masks delineating aneurysm boundaries. High-quality annotations are critical for supervised learning.

- **Preprocessing Steps:** Standardize CT images by normalizing Hounsfield units, resampling to uniform voxel spacing, and cropping to regions of interest centered on the abdominal aorta. Data augmentation techniques—such as rotation, scaling, and intensity variations—help increase dataset diversity and mitigate overfitting.

2. Dataset Splitting

To ensure unbiased evaluation, the dataset is partitioned as follows:

- **Training Set (70%):** 7,000 CT scans used to optimize model parameters. - **Validation Set (15%):** 1,500 scans employed for hyperparameter tuning and early stopping criteria. - **Test Set (15%):** 1,500 scans reserved exclusively for final model performance assessment.

Splitting is executed at the patient level to prevent data leakage and maintain independence between sets.

3. Model Architecture and Training

- **Architecture Selection:** CNNs, such as ResNet, DenseNet, or U-Net variants, are chosen for their robust feature extraction capabilities. For AAA detection, architectures may be adapted to incorporate 3D convolutions to capture volumetric context.

- **Training Procedure:** The model learns via forward propagation and backpropagation, minimizing a loss function (e.g., binary cross-entropy) using optimization algorithms like Adam or stochastic gradient descent over 100+ epochs.

- **Transfer Learning:** To leverage existing knowledge, pre-trained weights on large medical imaging datasets can be fine-tuned on the AAA dataset, accelerating convergence and improving accuracy.

4. Validation and Hyperparameter Optimization

- **Performance Monitoring:** After each epoch, evaluate metrics such as accuracy, precision, recall, F1 score, and loss on the validation set to track learning progress.

- **Hyperparameter Tuning:** Adjust learning rate, batch size, dropout rate, and network depth to optimize model generalization. Techniques such as grid search or Bayesian optimization can be employed.

- **Overfitting Prevention:** Regularization strategies—including dropout layers, weight decay, and early stopping—are implemented to prevent the model from memorizing training data at the expense of validation performance.

5. Testing and Final Evaluation

- **Unseen Data Assessment:** The finalized model is tested on the independent test set, providing unbiased estimates of clinical diagnostic performance.

- **Key Metrics:** - **Sensitivity (Recall):** Proportion of correctly identified AAA-positive cases. - **Specificity:** Proportion of correctly identified AAA-negative cases. - **Area Under the Receiver Operating Characteristic Curve (AUC):** Overall diagnostic ability.

- **Expected Outcomes:** State-of-the-art models achieve approximately 95% sensitivity, 90% specificity, and 0.96 AUC, indicating high reliability for clinical application.

Challenges in Deep Learning-Based AAA Detection

Despite promising results, several challenges remain:

- **Data Heterogeneity:** Variability in CT acquisition protocols, slice thickness, and contrast enhancement can affect model robustness.

- **Class Imbalance:** Although balanced datasets are preferred, real-world prevalence of AAA is low, complicating model deployment.

- **Annotation Quality:** Manual segmentation is labor-intensive and subject to interobserver variability, potentially impacting model training.

- **Computational Demand:** Training 3D CNNs on large datasets requires substantial computing resources and specialized hardware.

- **Interpretability:** Deep learning models are often "black boxes," limiting clinical trust. Efforts to develop explainable AI (XAI) techniques are ongoing to visualize decision rationale.

Future Directions and Research Opportunities

- **Multi-Modal Integration:** Combining CT imaging with clinical data (e.g., demographics, laboratory values) may enhance predictive accuracy.

- **Longitudinal Analysis:** Developing models to predict AAA growth trajectories and rupture risk could transform patient management.

- **Real-Time Deployment:** Incorporating AAA detection algorithms into Picture Archiving and Communication Systems (PACS) for seamless radiologist support.

- **Federated Learning:** Collaborative training across institutions without data sharing can address privacy concerns and improve generalizability.

- **Explainability Enhancements:** Developing saliency maps and attention mechanisms to elucidate model focus areas, fostering clinical acceptance.

Frequently Asked Questions (FAQs)

Q: Why is a balanced dataset critical for training AI models in medical imaging? A: Balanced datasets prevent model bias toward the majority class, ensuring equal sensitivity and specificity across positive and negative cases, which is vital for accurate clinical diagnosis. **Q: How does validation contribute to model development?** A: Validation data enable fine-tuning hyperparameters and early detection of overfitting, ensuring the model generalizes well beyond the training data. **Q: What is the significance of sensitivity and specificity in AAA detection?** A: High sensitivity ensures most cases with AAA are detected, reducing missed diagnoses, while high specificity minimizes false positives, preventing unnecessary interventions. **Q: Which techniques are effective in preventing overfitting in CNNs?** A: Dropout layers randomly deactivate neurons during training; early stopping halts training when validation performance plateaus; and hyperparameter tuning optimizes model complexity.

Conclusion

The application of deep learning models, specifically CNNs, to abdominal CT scans for AAA detection represents a significant advancement in diagnostic radiology. Through meticulous data collection, preprocessing, model training, validation, and evaluation, these AI systems can achieve high diagnostic accuracy, supporting timely clinical decisions and potentially reducing AAA-related morbidity and mortality.

Ongoing research addressing current challenges and integrating multi-modal data will further enhance model performance and clinical utility. As AI continues to mature, its integration into routine AAA screening and monitoring holds promise for transforming vascular disease management and improving patient care outcomes.

Keywords: Abdominal Aortic Aneurysm, AAA Detection, Deep Learning, Convolutional Neural Networks, CT Scans, Medical Imaging, Artificial Intelligence, Clinical Applications, Model Training, Overfitting, Validation, Diagnostic Accuracy.