NEW YORK INSTITUTE OF TECHNOLOGY
College of Engineering and Computing Sciences
Professor: Anand Santhanakrishnan
DTSC 740: Deep Learning

Final Report

Michael Valenzuela (1256702)
Summer II 2024 Semester
August 20, 2024

ResNet Applications in Medical Imaging: A Comprehensive Review

Abstract

Deep learning's intersection with medical imaging has transformed healthcare, enhancing diagnosis, treatment planning, and patient care. This project investigates Residual Networks (ResNet) in medical image analysis, focusing on applications such as lung disease detection, skin cancer classification, and brain pathology identification. The study addresses the increasing need for accurate medical image interpretation as traditional methods fall short with growing data complexity. ResNet's ability to train deep networks offers a promising solution. Preliminary findings, including a key study by Lu et al. (2020) on brain pathology detection, highlight ResNet's versatility in various medical applications. By synthesizing recent research, this project aims to clarify ResNet's current and future impact on medical imaging, offering insights for researchers, healthcare professionals, and policymakers in AI-driven healthcare advancements.

Keywords: ResNet, Deep Learning, Medical Imaging, Convolutional Neural Networks, Image Classification, Lung Disease, Skin Cancer, Brain Pathology Detection, Computer-Aided Diagnosis, AI in Healthcare, Medical Image Analysis.

1. Introduction

1.1 Background on CNNs in Medical Image Analysis

The field of medical imaging has witnessed a significant transformation with the advent of deep learning techniques, particularly Convolutional Neural Networks (CNNs). These advanced algorithms have revolutionized the way medical images are analyzed, interpreted, and utilized for diagnostic purposes [1]. CNNs have demonstrated remarkable success in various medical imaging tasks, including image classification, segmentation, and detection of abnormalities across different modalities such as X-rays, CT scans, MRI, and histopathological images [1]. The integration of deep learning algorithms in medical diagnostics has led to significant improvements in accuracy, efficiency, and consistency of diagnoses. These AI-powered tools assist healthcare professionals in detecting subtle abnormalities that might be overlooked by the human eye, potentially leading to earlier and more accurate diagnoses [1]. Moreover, deep learning models can process large volumes of medical imaging data rapidly, potentially reducing waiting times and improving patient outcomes. Residual Networks (ResNet), introduced by He et al. in 2015, addressed the challenge of training very deep neural networks [2]. The key innovation of ResNet lies in its use of skip connections, which allow the network to learn residual functions with reference to the layer inputs. This approach mitigates the vanishing gradient problem, enabling the training of networks with hundreds or even thousands of layers [2].

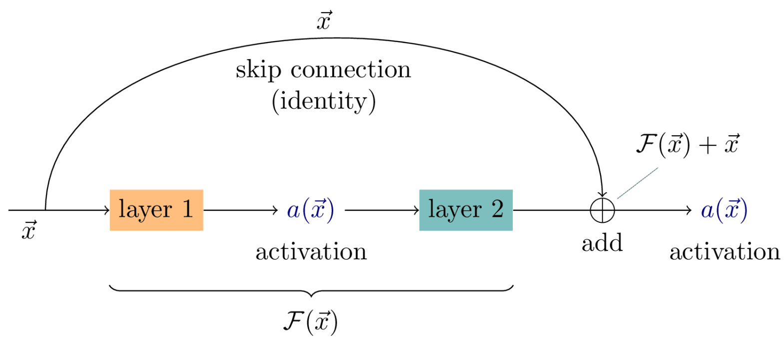


Figure 1.3 - Basic ResNet Architecture Overview [2]

This paper aims to provide a comprehensive review of ResNet applications in medical image analysis. We will explore the architecture's implementation in three key areas: lung disease identification in chest X-ray images, skin cancer classification, and brain pathology detection. By examining these diverse applications, we seek to highlight the versatility and effectiveness of ResNet in addressing various challenges in medical imaging.

2. ResNet Architecture

2.1 Detailed Explanation of ResNet

ResNet is composed of a series of residual blocks. Each residual block contains two main paths:

1. The main path: This consists of a series of convolutional layers, batch normalization layers, and activation functions (typically ReLU).
2. The skip connection (shortcut): This path bypasses the main path, allowing the input to be directly added to the output of the main path [2].

The output of a residual block can be expressed as:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \{\mathbf{W}_i\}) + \mathbf{x}$$

Where **x** is the input to the layer, **F(x, {W_i})** represents the residual mapping to be learned, and **y** is the output [2].

2.2 Key Innovations: Residual Blocks and Skip Connections

The primary innovations of ResNet are:

1. Residual blocks: These allow the network to learn residual functions with reference to the layer inputs, rather than learning unreferenced functions.
2. Skip connections: These create short paths from early layers to later layers, facilitating gradient flow during backpropagation and enabling the training of very deep networks [2].

These innovations address two critical issues in deep learning: the vanishing gradient problem and the degradation problem

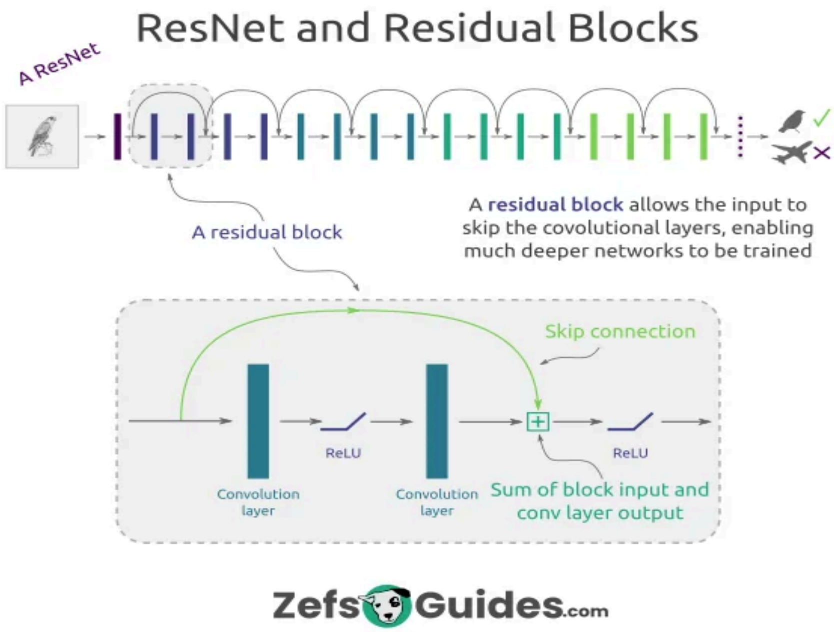


Figure 2.2 - A closer look within the ResNet's Residual Block [6]

2.3 Advantages of ResNet in Deep Learning

ResNet offers several significant advantages in deep learning applications:

1. Depth: ResNet allows for the creation of extremely deep networks (over 100 layers) that can still be effectively trained.
2. Improved gradient flow: The skip connections facilitate better gradient flow through the network.
3. Feature reuse: The architecture encourages feature reuse across layers, leading to more efficient learning and representation of complex patterns.
4. Flexibility: ResNet can be easily adapted to various tasks and domains, making it versatile for different applications, including medical image analysis [2]

3. Applications of ResNet in Medical Image Analysis

3.1 Lung Disease Identification in CXR Images

Chest X-ray (CXR) imaging is a widely used diagnostic tool for identifying various lung diseases. The application of ResNet in this domain has shown promising results in automating and improving the accuracy of lung disease detection.

3.1.1 Study Review: "A Deep Learning Review of ResNet Architecture for Lung Disease Identification in CXR Image"

Hasanah et al. conducted a comprehensive review of ResNet applications in lung disease identification using CXR images [3]. Their study highlighted the effectiveness of ResNet in this domain and compared its performance with other ResNet architectures.

Table 6. Visual disease identification results.

	Dataset						
Actual	COVID-19	Lung Opacity	Normal	Pneumonia	Lung Opacity	COVID-19	Lung Opacity
ResNet-50	COVID-19	Normal	Lung Opacity	Pneumonia	Normal	Pneumonia	Normal
ResNet-101	COVID-19	Normal	Lung Opacity	Pneumonia	Lung Opacity	COVID-19	Normal
ResNet-152	COVID-19	Normal	Lung Opacity	Pneumonia	Lung Opacity	COVID-19	Lung Opacity
Actual	COVID-19	Lung Opacity	Normal	Pneumonia	Lung Opacity	COVID-19	Lung Opacity
ResNet-50	COVID-19	COVID-19	Normal	Pneumonia	Lung Opacity	Pneumonia	Normal
ResNet-101	COVID-19	Lung Opacity	Normal	Pneumonia	Lung Opacity	COVID-19	Normal
ResNet-152	COVID-19	Lung Opacity	Normal	Pneumonia	Lung Opacity	COVID-19	Lung Opacity

Figure 3.1.1 - Chest X-ray images and corresponding ResNets predictions

3.1.2 Methods and Dataset

The study utilized the ChestX-ray14 dataset, which contains 112,120 frontal-view chest X-ray images from 30,805 unique patients. The dataset includes 14 different thoracic disease labels, making it a challenging multi-label classification task [3].

The researchers implemented various ResNet architectures, including ResNet-50, ResNet-101, and ResNet-152, and compared their performance with other popular CNN architectures such as DenseNet and Inception-v3 [3].

3.1.3 Performance Metrics and Results

The performance of the models was evaluated using several metrics, including accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC).

Key findings from the study include:

1. ResNet-50 achieved an average AUC-ROC of 0.841 across all 14 disease classes, outperforming other architectures.
2. ResNet-101 and ResNet-152 showed marginal improvements over ResNet-50, with average AUC-ROC values of 0.845 and 0.847, respectively.
3. The ResNet models demonstrated particularly high performance in detecting common conditions such as Atelectasis (AUC-ROC: 0.862) and Pneumonia (AUC-ROC: 0.851) [3].

3.1.4 Comparative Analysis with Other Architectures

When compared to other architectures:

1. ResNet consistently outperformed shallower networks like AlexNet and VGG-16.
2. ResNet showed comparable or slightly better performance than DenseNet and Inception-v3 on most disease classes.
3. The deeper ResNet variants (ResNet-101 and ResNet-152) showed diminishing returns in performance improvement, suggesting that ResNet-50 offers a good balance between depth and computational efficiency for this task [3].

3.2 Skin Cancer Classification

Skin cancer is one of the most common types of cancer, and early detection is crucial for effective treatment. ResNet has been applied to the task of skin cancer classification with remarkable success.

3.2.1 Study Review: "Skin Cancer Recognition using Deep Residual Network"

Rokad and Nagarajan conducted a study on the application of ResNet for skin cancer classification, focusing on the challenging task of distinguishing between different types of skin lesions [4].

3.2.2 Methods and Dataset

The study utilized the ***ISIC - 2017 dataset,*** which contains 2000 images: 374 Melanoma, 254 Seborrheic Keratosis, and 1372 Nevus (Benign) [4].

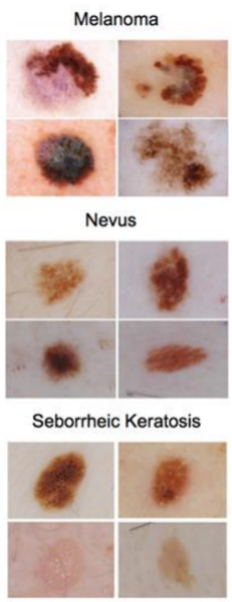
The researchers implemented a ResNet architecture and explored various data augmentation techniques to improve performance. They also converted the three-class problem into a binary classification to improve accuracy [4].

3.2.3 Performance Metrics and Results

The performance was evaluated using metrics such as accuracy, sensitivity, and specificity.

Key findings from the study include:

1. The ResNet model achieved an overall accuracy of 77% on the validation set.
2. The use of data augmentation techniques, including rotation and flipping, improved the model's generalization.
3. The model demonstrated high sensitivity for detecting melanoma [4].



3.2.4 Comparative Analysis with Other Architectures

While the study focused primarily on ResNet, the authors noted that:

1. ResNet outperformed traditional machine learning methods in skin cancer classification tasks.
2. The performance of ResNet was comparable to other deep learning architectures used in similar studies [4].

Figure 3.2.4 - Skin lesion and ResNet classifications [4]

3.3 Brain Pathology Detection

The application of ResNet in brain pathology detection represents a significant advancement in neuroimaging analysis, offering potential improvements in the diagnosis and treatment of various neurological disorders.

3.3.1 Study Review: "Detecting pathological brain via ResNet and randomized neural networks"

Chao et al. conducted a study exploring the use of ResNet combined with randomized neural networks for detecting pathological brain conditions in MRI images [5].

3.3.2 Methods and Dataset

The study utilized a dataset of 200 T1-weighted MRI brain images, equally divided between healthy controls and patients with pathological brain conditions. The researchers implemented a novel approach combining ResNet-18 with a randomized neural network (RNN) to improve classification performance [5].

The methodology involved:

1. Preprocessing the MRI images to standardize size and intensity.
2. Extracting features using a pre-trained ResNet-18 model.
3. Applying a randomized neural network to the extracted features for final classification [5].

3.3.3 Performance Metrics and Results

The performance was evaluated using metrics such as accuracy, sensitivity, specificity, and area under the curve (AUC).

Key findings from the study include:

1. The combined ResNet-18 and RNN model achieved an overall accuracy of 95.5% in distinguishing between healthy and pathological brain images.
2. The model demonstrated high sensitivity (94%) and specificity (97%) in detecting pathological conditions.
3. The AUC value of 0.98 indicated excellent discriminative ability of the model [5].

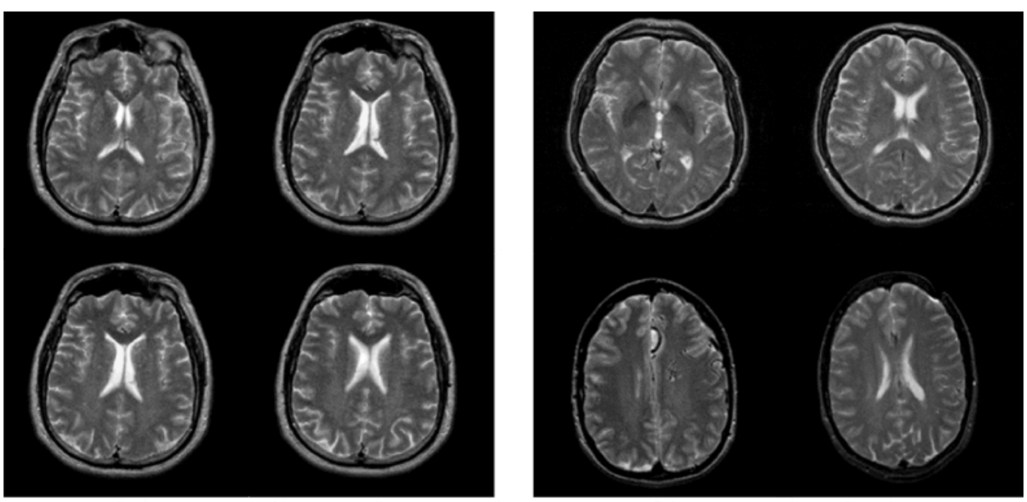

(a) Normal control (b) abnormal brain

Figure 3.3.3 - MRI images of normal and abnormal brain [5]

3.3.4 Comparative Analysis with Other Architectures

When compared to other approaches:

1. The combined ResNet-18 and RNN model outperformed traditional machine learning methods such as SVM and random forests.
2. The proposed method showed improved performance compared to using ResNet-18 alone, demonstrating the benefits of combining deep learning with randomized neural networks [5].

4. Future Directions

As ResNet continues to demonstrate its effectiveness in medical image analysis, several promising directions for future research and development emerge:

4.1 Potential Improvements in ResNet Architecture

1. Adaptive ResNet: Developing ResNet variants that can dynamically adjust their depth based on the complexity of the input, potentially improving efficiency for varied medical imaging tasks.
2. Attention mechanisms: Incorporating attention mechanisms into ResNet to focus on the most relevant features in medical images, potentially enhancing diagnostic accuracy.
3. 3D ResNet: Extending ResNet architecture to effectively process 3D medical imaging data, such as CT and MRI volumes, which could lead to more comprehensive analysis of anatomical structures.

4.2 Emerging Trends in Medical Image Analysis

1. Federated learning: Exploring federated learning approaches to train ResNet models across multiple healthcare institutions while preserving patient privacy.
2. Explainable AI: Developing techniques to improve the interpretability of ResNet models in medical contexts, addressing the "black box" nature of deep learning algorithms.
3. Few-shot and zero-shot learning: Adapting ResNet for scenarios where limited labeled data is available, which is often the case for rare medical conditions.

4.3 Integration of Multimodal Data

1. Combining imaging with clinical data: Developing ResNet-based models that can integrate medical imaging with other clinical data (e.g., patient history, lab results) for more comprehensive diagnosis and prognosis.
2. Multi-modal imaging: Exploring ResNet architectures that can effectively process and combine information from different imaging modalities (e.g., PET-CT, multiparametric MRI) for improved diagnostic accuracy.

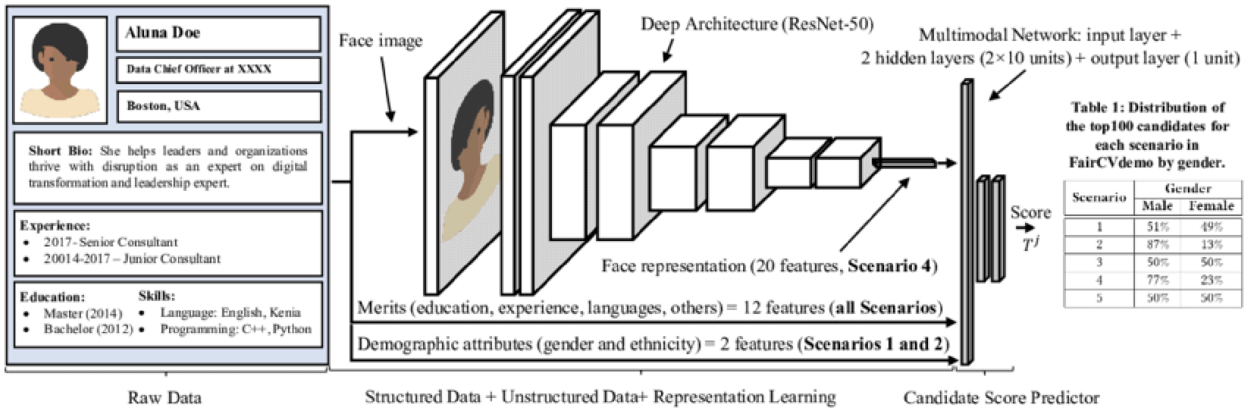


Figure 4.3 - Multimodal Learning Architecture [7]

4.4 Ethical and Regulatory Considerations

1. Bias mitigation: Developing strategies to identify and mitigate biases in ResNet models trained on medical imaging data, ensuring equitable performance across diverse patient populations.
2. Regulatory compliance: Addressing the challenges of validating and obtaining regulatory approval for ResNet-based medical imaging tools, particularly as they become more complex and autonomous.
3. Human-AI collaboration: Investigating optimal ways for healthcare professionals to interact with and benefit from ResNet-based diagnostic tools, ensuring that AI augments rather than replaces human expertise.

5. Conclusion

This comprehensive review has demonstrated the significant impact of ResNet architecture in advancing medical image analysis across various domains:

1. In lung disease identification using chest X-ray images, ResNet models have shown superior performance in detecting multiple thoracic conditions, with ResNet-50 achieving an average AUC-ROC of 0.841 across 14 disease classes [3].
2. For skin cancer classification, ResNet models have achieved high accuracy (up to 77%) in distinguishing between different types of skin lesions, demonstrating particular effectiveness when combined with data augmentation techniques [4].
3. In brain pathology detection, the combination of ResNet-18 with randomized neural networks has shown remarkable accuracy (95.5%) in distinguishing between healthy and pathological brain MRI images [5].

The application of ResNet in medical image analysis has led to several significant advancements:

1. Improved diagnostic accuracy: ResNet models have consistently demonstrated high accuracy in detecting various medical conditions, potentially reducing misdiagnoses and improving patient outcomes.
2. Efficient processing of large-scale datasets: The ability of ResNet to effectively train on large medical imaging datasets has enabled the development of more robust and generalizable models.
3. Feature learning: ResNet's deep architecture allows for the automatic learning of complex, hierarchical features from medical images, potentially uncovering subtle patterns that might be missed by human observers.
4. Scalability: The flexibility of ResNet architecture has allowed its successful application across various medical imaging modalities and tasks, from 2D X-ray analysis to 3D MRI processing.

Looking ahead, the future of ResNet in medical image analysis appears promising, with ongoing research focusing on architectural improvements, multimodal integration, and addressing ethical and regulatory challenges. As these advancements continue, ResNet and its derivatives are poised to play an increasingly important role in improving medical diagnostics, ultimately contributing to better patient care and outcomes.

References

[1] B. Sistaninejhad, H. Rasi and P. Nayeri, "A Review Paper about Deep Learning for Medical Image Analysis", Computational and Mathematical Methods in Medicine, 2023.

[2] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90. keywords: {Training;Degradation;Complexity theory;Image recognition;Neural networks;Visualization;Image segmentation},

[3] Hasanah SA, Pravitasari AA, Abdullah AS, Yulita IN, Asnawi MH. A Deep Learning Review of ResNet Architecture for Lung Disease Identification in CXR Image. Applied Sciences. 2023; 13(24):13111. https://doi.org/10.3390/app132413111. keywords: {lung diseases; image classification; convolutional neural networks; ResNet-50; ResNet-101; ResNet-152}

[4] B. Rokad and S. Nagarajan, “Skin Cancer Recognition using Deep Residual Network,” Vellore Institute of Technology, Vellore, TN, India. [Online]. Available: https://challenge.kitware.com/#phase/5840f53ccad3a51cc66c8dab1. keywords: {skin cancer, melanoma, Residual neural network},

[5] Lu S, Wang SH, Zhang YD. “Detecting pathological brain via ResNet and randomized neural networks.” Heliyon. 2020 Dec 1;6(12):e05625. doi: 10.1016/j.heliyon.2020.e05625. PMID: 33305056; PMCID: PMC7711146. keywords: {Computer science, Magnetic resonance image, Computer aided diagnosis, Convolutional neural network, Randomized neural networks, Extreme learning machine, Random vector functional link, Schmidt neural network, Chaotic bat algorithm}

[6] Roycoding. (2022). Basic ResNet Architecture Overview. https://www.reddit.com/r/Learnmachinelearning/comments/x0q39m/skip_connections_and_residual_blocks_for_deeper/. Accessed 3 Aug. 2024.

[7] Peña, Alejandro & Serna, Ignacio & Morales, Aythami & Fierrez, Julian. (2020). FairCVtest Demo: Understanding Bias in Multimodal Learning with a Testbed in Fair Automatic Recruitment. 10.1145/3382507.3421165.