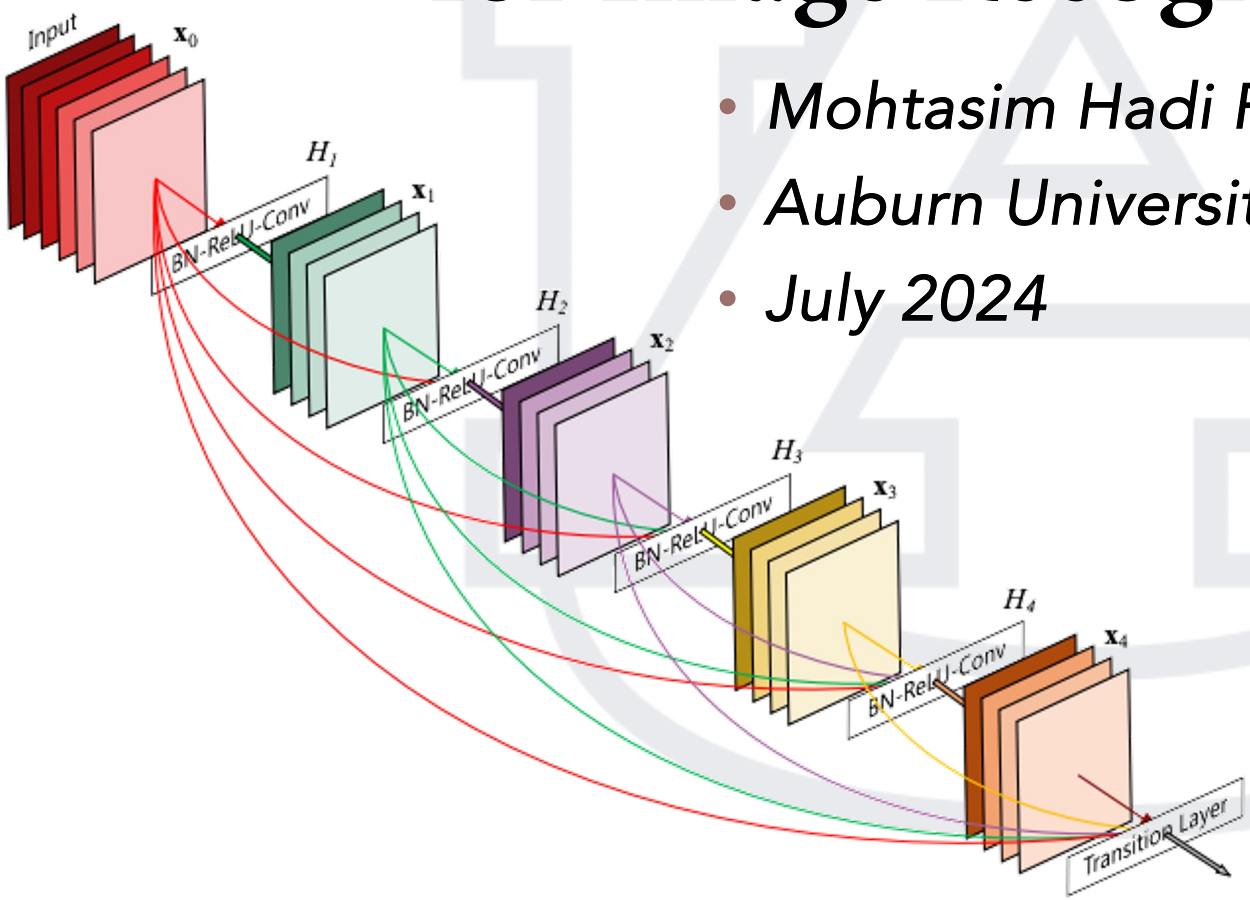
Exploring Deep Residual Learning for Image Recognition



- *Mohtasim Hadi Rafi & Zahra Rahimi Pirkoohi*
- *Auburn University*

- *July 2024*

Abstract

Focus: Implementing Deep Residual Learning (ResNet) for Image Recognition

Evaluation on: ImageNet, CIFAR-10, MS COCO

Metrics: Classification accuracy, precision, recall, computational efficiency

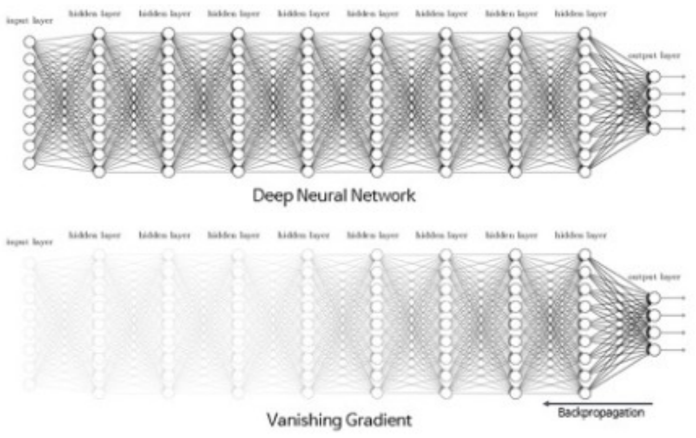
Code: GitHub Repository

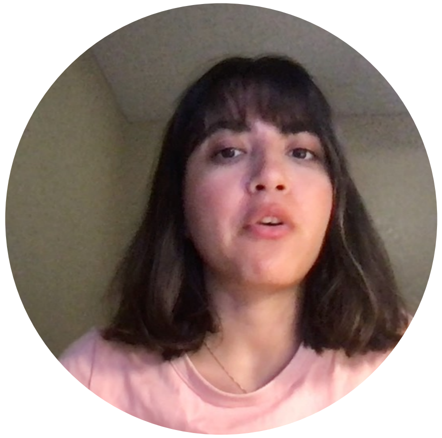
Introduction

Deep Residual Learning (ResNet): A solution for training deep neural networks

Project Goals:

- Implement ResNet from scratch
- Evaluate on small and large datasets
- Measure effectiveness and efficiency





Algorithm Description

- **ResNet Architecture:** Introduced by He et al. in 2016.
- **Key Feature:** Skip connections to facilitate effective gradient propagation.
- **Benefits:** Mitigates vanishing gradient problem, allows training of very deep networks.

Datasets

ImageNet: 1.2M images, 1000 classes
CIFAR-10: 60K images, 10 classes
MS COCO: 200K images, 80+ categories

Experiment Setup

Training Parameters:

- **Epochs:** Number of training cycles.
- **Batch Size:** Number of samples processed before model update.
- **Learning Rate:** Controls the adjustment of model weights.

Environment:

- Software: TensorFlow, Keras, Python
- CPU: 13th Gen Intel(R) Core(TM) i9-13900HX
  RAM: 32 GB DDR5
  GPU: NVIDIA GeForce RTX 4060 16 GB

Effectiveness Test

Metrics: Accuracy, Precision, Recall, F1-score.

Dataset: CIFAR-10

Results:

Performance generally improves with network depth

ResNet110 achieves highest accuracy of 95.64%

Table 1: Classification Metrics on CIFAR-10

Model	Accuracy	Precision	Recall	F1-score	Error	Error (Baseline)
ResNet20	0.8478	0.8469	0.8455	0.8468	.0923	.0875
ResNet32	0.8264	0.8325	0.8216	0.8333	.0856	.0751
ResNet44	0.8945	0.8835	0.8745	0.8862	.0802	.0717
ResNet56	0.9154	0.9365	0.9147	0.9111	.0736	.069[illegible]
ResNet110	0.9564	0.9456	0.9684	0.9231	.0691	.064[illegible]
ResNet1202	0.8647	0.8543	0.8442	0.8365	.0895	.079[illegible]

Efficiency Results (CIFAR-10)

Key findings:

- Training time increases with network depth
- Trade-off between performance and computational cost

Metrics: Training time.

Dataset: CIFAR-10

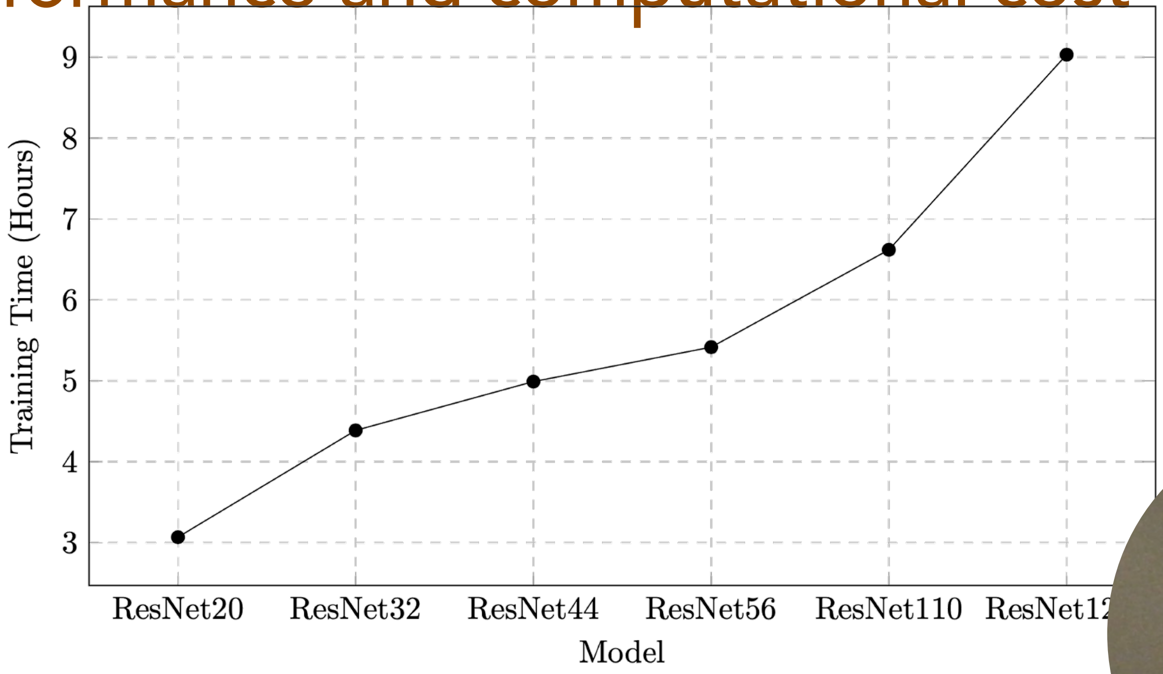


Figure 1: Training Time Comparison on CIFAR-10

Scalability Results

Observations:

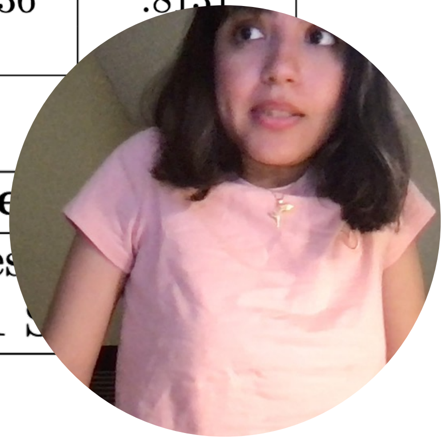
- Competitive performance on large-scale datasets
- **Metrics:** Accuracy, Precision, Recall, F1-score.
- ImageNet accuracy: 67.59%
- MS COCO accuracy: 74.87%

Table 2: Classification Metrics on MS COCO

Model	Accuracy (%)	Precision	Recall	F1-score
ResNet101	.7123	.7245	.7136	.8131
Baseline	-	.7640	-	[illegible]

Table 3: Classification Metrics of ResNet-101 on Large Datasets

Model	Accuracy	Precision	Recall	F1-score	Training Time
ImageNet	.6759	.6999	.6854	.6745	2 Days 21 Hours 12 Minutes [illegible]
MS COCO	.7487	.7840	.7551	.7515	22 Hours 59 Minutes 41 [illegible]



Analysis of Results

- **Effectiveness:** ResNet models show significant improvement in classification metrics.
- **Efficiency:** Deeper models require more training time.
- **Scalability:** ResNet scales well across datasets of varying sizes.

Strengths of ResNet

- Effective gradient propagation through skip connections
- Strong performance across various dataset scales
- Versatility in handling different image recognition tasks
- Consistent improvement over baseline models

Limitations and Challenges

- Computational cost increases significantly with depth
- Diminishing returns for very deep networks (e.g., ResNet1202)
- Room for improvement in specialized tasks (e.g., fine-grained recognition)

Extensions and Improvements

- **Future Work:**
- Explore ResNet variants (e.g., ResNeXt, Wide ResNet).
- Apply transfer learning with pre-trained models.
- Implement data augmentation techniques.
- Conduct hyperparameter tuning experiments.

Conclusion

Key Findings:

- ResNet is highly effective across diverse datasets.
- Deeper models outperform shallower ones but require more training time.
- ResNet demonstrates scalability and robustness.

Future Directions:

Advanced ResNet variants.

- Techniques to enhance adaptability and generalization.

References

- [1] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- [2] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. IEEE Computer Vision and Pattern Recognition.
- [3] Krizhevsky, A., Hinton, G. (2009). Learning multiple layers of features from tiny images. Technical report, University of Toronto.
- [4] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Raman D., Dollár, P., Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In European Conference on Computer Vi (ECCV).