# Deep Residual Learning for Image Recognition

SAI DHIRESH KILARI[1], DR. PETER WU[2]

***Abstract- Deep neural networks have potentially demonstrated dominant success in the field of image recognition and relevant tasks. However, with the significant increase of the network depth mostly leads to a specific set of optimization-based challenges, specifically in terms of vanishing as well as exploding the gradients.***

## I. INTRODUCTION

Deep neural networks have potentially demonstrated a i. dominant success in the field of image recognition and relevant tasks. However, with the significant increase of the network depth mostly leads to a specific set of optimization-based challenges, specifically in terms of vanishing as well as exploding the gradients [1]. The ii. existing traditional forms of architectures such as the iii. likes of AlexNet as well as VGG face struggle into training the very deep networks in an effective manner. This paper directly introduces the Residual Networks (ResNets), which turns out to be a novel iv. deep learning architecture explicitly addressing the v. issues by effectively utilizing the aspect of skip connections.

## II. BACKGROUND & CHALLENGES IN DEEP LEARNING

Deep learning has helped with the necessary advancements of image recognition, integrated with architectures such as VGG, AlexNet as well as GoogleNet, successfully achieving notable success. However, with the deeper growth of the network, they also face multiple challenges such as the likes of degraded accuracy, vanishing gradients as well as difficulties in terms of optimization [2]. The traditional deep neural networks also face struggle in learning about the effective forms of feature representations with the significant increase of the depth. In order to properly address such kinds of issues, researches necessarily introduced Residual Networks (ResNets), which makes use of skip connections in order to bypass all the individual layers that allows the gradients to flow in a smoother manner during the effective backpropagation. This particular innovation allows to train the ultra-deep networks, bringing improvement to generalization as well as accuracy in classification. Since, ResNets have turned out to be the marked cornerstone of computer vision as well as applications of deep learning.

Some of the difficulties faced by Deep neural networks are,

Degradation Problem: The deeper networks at times visually showcase higher error of training than all the existing shallower counterparts, notably contradicting the expectation that helped increase the depth, which should be improving the performance [3].

Difficulty in Optimization: Appropriately training the deep networks has the prime requirement of a careful weight initialization along with complex techniques of optimization.

Vanishing & Exploding Gradients: With the significant growth of the network in a deeper manner, the gradients also undergo diminishing or might also explode during the aspect of backpropagation, which makes the entire training completely unstable.

## III. RESIDUAL LEARNING – THEORETICAL FOUNDATION

3.1. Residual Mapping:

ResNets make notable introductions to residual learning aimed at tackling the problem of degradation. Instead of potentially learning the direct mapping ($H(x)$), ResNets mostly learn about the residual function:

$F(x) = H(x) - x$

Hence, the originally existing function can again be reformulated as $H(x) = F(x) + x$

Where, x mainly represents the input and the $F(x)$ refers to the residual function [4]. The identity mapping, which is ‘x’ is again propagated with the help of skip connections, which makes it easier for the network to potentially learn about the residual functions rather than having complete transformations.

3.2. Skip Connections:
The skip connections specifically allow a direct flow through the network, which prevents the vanishing gradients as well as allows an efficient training regarding deeper architectures [5]. This particular modification significantly results into an easier optimization along with a notably improved generalization.

3.3. Comparison to the Traditional Networks:
Unlike the standard deep networks that face struggle into optimization of deeper architectures, ResNets can offer training to the networks with nearly 1,000 layers while at the same time maintain proper accuracy along with reduction of errors. This particular approach is effectively different from the previously existent architectures such as GoogleNet as well as VGG, heavily relying upon the wider and the deeper structures without having an explicit form of identity mappings.

i.

## IV. RESIDUAL NETWORK ARCHITECTURE ii.

ResNet mainly consists of a specific stack of residualiii. blocks, every individual block consisting of theiv. following,

i. Two or multiple convolutional layers.
ii. ReLU activation for effective non-linearity [6].
iii. Batch normalization intended for stable training.
iv. Skip connections that effectively bypass one or multiple layers.

4.1. Basic Residual Block:
A simplified residual block follows a specific structure:
Y = F(x, W) + x.

Where:
i. x is referred to as the input.
ii. W specifically represents the learnable weights for all the convolutional layers [7].
iii. F(x,W) specifically denotes the appropriate transformation applied directly to x.

4.2. Bottleneck Residual Block:
In order to reduce the computation, various deeper ResNets such as ResNet-50, ResNet-101 make use of the bottleneck layers as mentioned below,
a. 1x1 convolutions bringing reduction to the dimensionality.
b. 3x3 convolutions carrying out specific transformations [8].
c. 1x1 convolutions restoring the dimensions.

This particular design brings notable reduction to the cost of computation while at the same time preserve the performance.

## V. EXPERIMENTAL RESULTS

ResNets had been tested on large-sized databases such as the likes of ImageNet as well as CIFAR-10, carrying out appropriate demonstration of the superior efficiency as well as accuracy.

5.1. ImageNet Performance:
ResNets had been evaluated specifically on the ImageNet (ILSVRC 2015), which turns out to be a dataset specifically consisting of nearly 1.2 images across a total of 1000 categories.

ResNet-34 had potentially outperformed VGG-19, successfully achieving an error rate of lower top-5.

ResNet-50, ResNet-152 as well as ResNet-101 specifically sets newer benchmarks in the form of classification accuracy [9].

5.2. CIFAR-10 & CIFAR-100 Results:
For CIFAR-10 (10 classes) as well as CIFAR-100 (100 classes:
a) ResNet-110 nearly 110 layers specifically surpassing the traditionally existing CNNs.
b) ResNets had been maintained with the help of a lower form of training as well as validation error directly compared to the conventional forms of existing architectures.

5.3. Efficiency of Training:
ResNets specifically train faster as well as generalise better than the existing non-residual deep networks [10]. In addition to this, Gradient propagation is much

smoother because of the skip connections, which directly leads to an efficient form of optimization.

## VI. ADVANTAGES OF RESNETS

The offered advantages of ResNets have been outlined in the following points for a better understanding,

i. Higher Accuracy: ResNets potentially outperforms the traditionally existing CNNs upon large-sized datasets.
ii. Improved Generalization: ResNets necessarily generalise properly across all the different tasks related to classification of images.
iii. Improved Gradient Flow: Skip connections specifically prevent the vanishing gradients, allow a deeper network training.
iv. Easier Optimization: Training the deeper models is much feasible without the degradation of performance.

## VII. APPLICATION OF RESIDUAL NETWORKS

ResNets have specifically transformed the applications aligned to Deep Learning:

a. Image Classification: Utilization in the aspect of medical imaging, autonomous driving as well as satellite image analysis.
b. Facial Recognition: Appropriate integration into FaceNet as well as deeper biometric systems.
c. Detection of Objects: Necessary incorporation into Faster R-CNN as well as YOLO frameworks.
d. Speech & NLP: Adapted with the aspect of speech recognition as well as natural language processing.

## VIII. CONCLUSION & FUTURE WORK

This research paper explored about the architecture of ResNet, performance as well as theoretical foundation upon the benchmark datasets. With respect to this, Residual learning allows training the extremely deep networks, overcoming the traditional limitations of deep learning. The future research might focus upon:

i. Enhancement of computational efficiency intended for the real-time applications.
ii. Exploring the unsupervised and the self-supervised learning approaches utilizing ResNets.
iii. Effectively integrating the ResNets with the emergence of AI-based technologies such as transformers.

## REFERENCES


[1] M., Shafiq, and Z., Gu. Deep residual learning for image recognition: A survey. *Applied Sciences*, *12*(18), p.8972, 2022.

[2] W., Fang, Z., Yu, Y., Chen, T., Huang, T., Masquelier, and Y., Tian. Deep residual learning in spiking neural networks. *Advances in Neural Information Processing Systems*, *34*, pp.21056-21069, 2021.

[3] I.C., Duta, L., Liu, F., Zhu, and L., Shao. Improved residual networks for image and video recognition. In *2020 25th International Conference on Pattern Recognition (ICPR)* (pp. 9415-9422). IEEE, 2021, January.

[4] L.H., Shehab, O.M., Fahmy, S.M., Gasser, and M.S., El-Mahallawy. An efficient brain tumor image segmentation based on deep residual networks (ResNets). *Journal of King Saud University-Engineering Sciences*, *33*(6), pp.404-412, 2021.

[5] C., Liu, X., Liu, D.W.K., Ng, and J., Yuan. Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications. *IEEE Transactions on Wireless Communications*, *21*(2), pp.898-912, 2021.

[6] M., Neshat, M.M., Nezhad, S., Mirjalili, D.A., Garcia, E., Dahlquist, and A.H., Gandomi. Short-term solar radiation forecasting using hybrid deep residual learning and gated LSTM recurrent network with differential covariance matrix adaptation evolution strategy. *Energy*, *278*, p.127701, 2023.

[7] Y., Hu, H., Tang, and G., Pan. Spiking deep residual networks. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(8), pp.5200-5205, 2021.

[8] D., Qiu, L., Zheng, J., Zhu, and D., Huang. Multiple improved residual networks for medical image super-resolution. *Future Generation Computer Systems*, *116*, pp.200-208, 2021.

[9] N., Shahadat, and A.S., Maida. Deep residual axial networks. *arXiv preprint arXiv:2301.04631*, 2023.

[10] C., Zhou, S., Zhou, J., Xing, and J., Song. Tomato leaf disease identification by restructured deep residual dense network. *IEEE Access*, *9*, pp.28822-28831, 2021.