Ecole Nationale
Supérieure
d'Ingénieurs
de Tunis

Deep Learning

Project :

Residual Networks (ResNet)

Realized by :Hattab hammami
M2 Data Science

Academic year : 2023/2024

Summary:

I. Introduction
II. Challenges faced by Deep Neural Networks
III. Key Components of ResNet Architecture
IV. A Look into the 34 layered Residual Neural Network
V. Comparing The Performance of Plain Networks with residual networks
VI. Conclusion

I. Introduction

Residual Network is a deep Learning model used for computer vision applications.
The ResNet (Residual Neural Network) architecture was introduced by Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian sun in their paper titled "Deep residual Learning for Image Recognition " in 2015.This article is primarily based on this paper.
One of the most significant developments in the ILVRSC (ImageNet Large Scale Visual Recognition Challenge) 2015 was the introduction of Residual network. Resnet achieved unprecedented results by addressing the challenges associated with training very deep neural networks.
Resnet outperformed other architectures by winning winning the image classification task in ILVRSC 2015 by a substantial margin with top 5 error rate of 3.57 % and achieved remarkable results .

II. Challenges faced by Deep Neural Networks

1. Vanishing/Exploding Gradient Problem:

As the number of layers in the neural network increases the gradients of the loss function with respect to the weight may become extrememly small during backpropogation. This makes it difficult for the latter layers to learn any meaningful pattern as the updates of the weights are almost negligible to have an impact.
Conversely in some cases the gradients can become very large during backpropogation leading to unstability in the training process by causing the weights to be updated by large amount, making the optimization process difficult to control.

2. Degradation Problem in Neural Network:

The degradation problem in neural networks refers to the phenomenon where, as the depth of a neural network increases ,the performance of the network on the training data saturates and then starts to degrade. The degradation problem is particularly problematic because it goes against the intuition that deeper networks are more able to extract intricate and abstract features.We can see it as twofold problem

3. **Performance Plateau:**

As the number of layers in the neural network increases the training error tends to saturate and stops improving.It means that the additional layers are not having any significant benefit on the reduction of training error.

4. **Accuracy Degradation:**

After the addition of more layers surprisingly the error on the validation set starts increasing and the performance on the unseen data becomes poor.

III. Key Components of ResNet Architecture

1. Residual Block:

Residual blocks are the main components of Residual Neural network .In a classical neural network the input is transformed by a set of convolutional layers then it is passed to the activation function. In a residual network the input to the block is added to the output of the block creating a residual connection.The output of the residual block H(x) can be represented by:

H(x) = F(x) + x

F(x) represents the residual mapping learned by the network .The presence of identity term x allows the gradient to flow more easily .

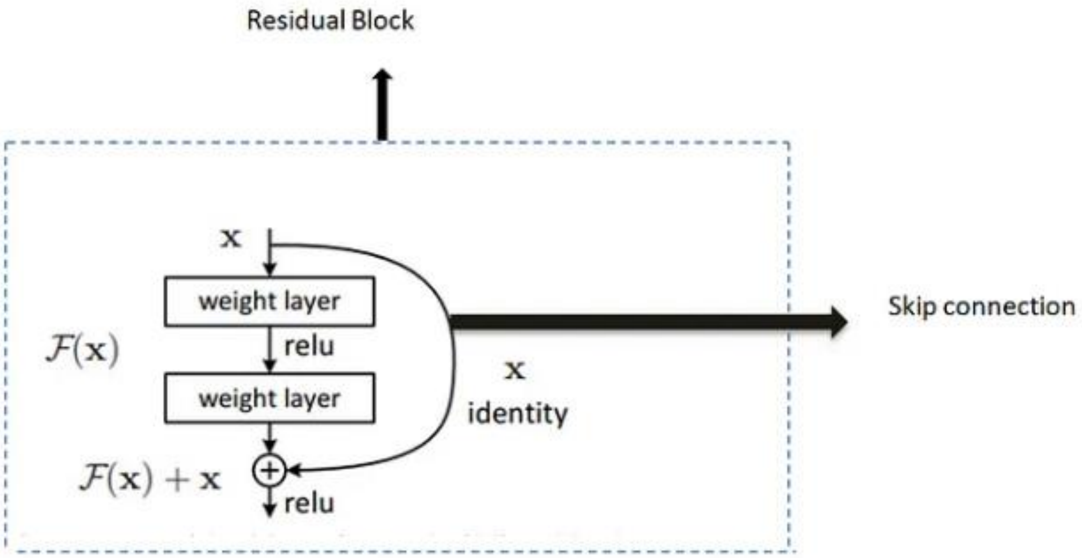


image from original paper-link-https://arxiv.org/pdf/1512.03385.pdf

2. **Skip Connection**

Skip connection helps in forming the residual blocks.Skip connection consists of the input of the residual block that is bypassed over the convolutional layer and added to the output of the residual block.

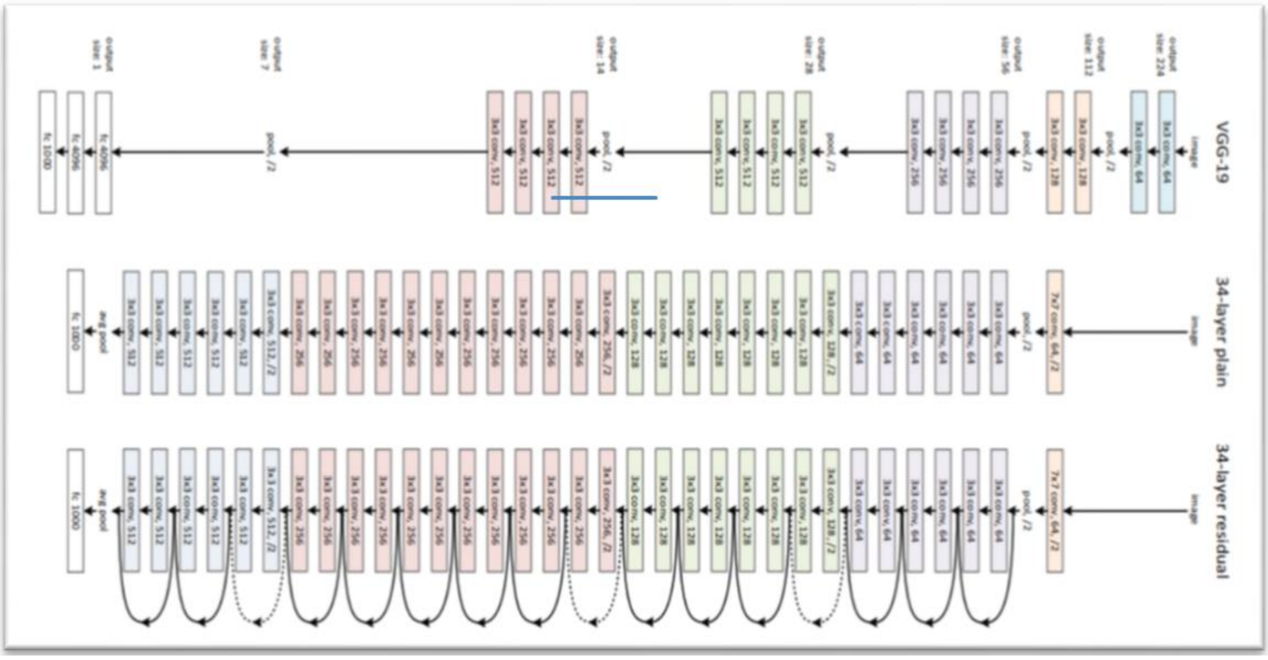
3. **Stacked layers**:

ResNet architectures are formed by stacking multiple residual blocks together.Using these multiple residual blocks together resnet architecture can be built very deep.Versions of ResNet with 50,101,152 layers were introduced.

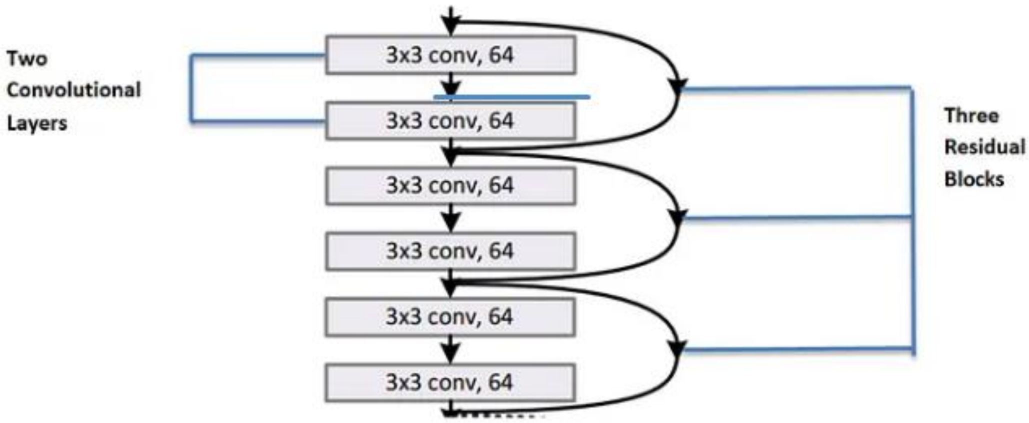
4. Global Average Pooling(GAP)

Resnet architectures typically utilises Global average Pooling as the final layer before the fully connected layer .GAP reduces spatial dimensions to a single value per feature map providing a compact representation of the entire feature map.

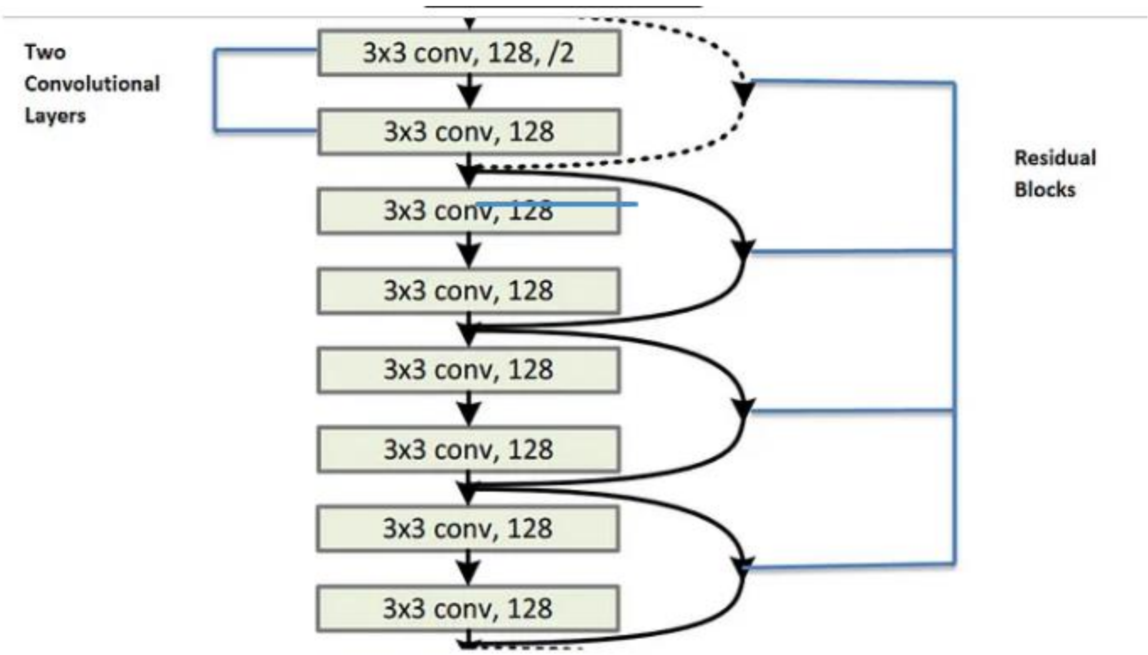
IV. A Look into the 34 layered Residual Neural Network



In this diagram we can see the VGC-19 ,34 layer plain network and 34 layered residual network.
In this residual network there are total 16 residual blocks.



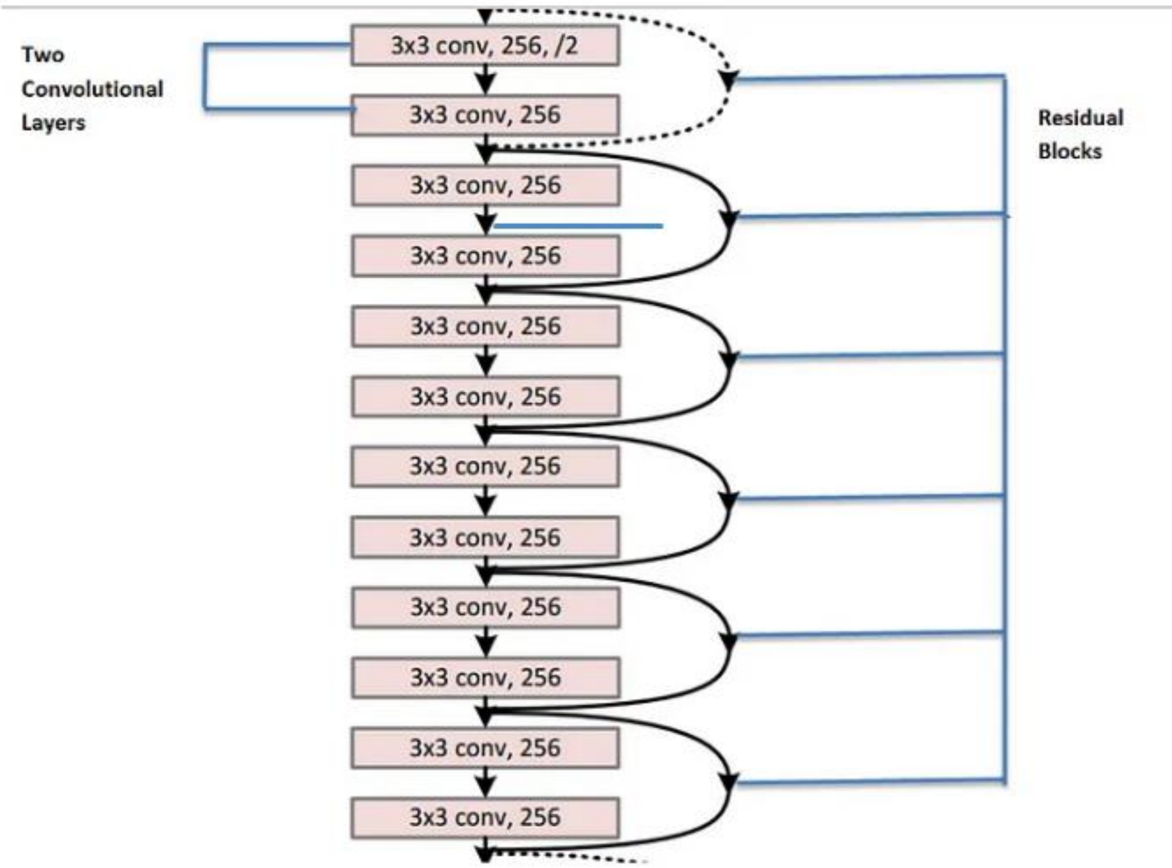
The first set consists of 3 residual blocks. Each residual block consists of 2 convolution layers where each convolution layer consists of 64 filters of size 3x3 and a skip connection which performs identity mapping.



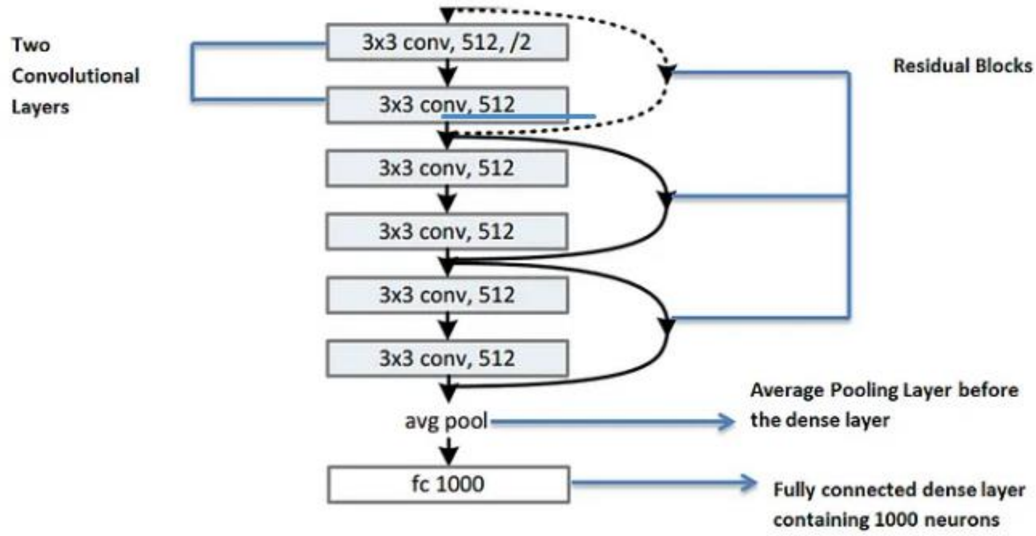
Second set consists of 4 residual blocks. Each residual block contains 2 convolutional layers where each layer consists of 128 kernels of of size 3x3 and a skip connection .

Dotted line skip connections represents the connections when the dimension are increased then it has to match the dimension of the output of convolutional layers.

When the dimensions increase (dotted line shortcuts), we consider two options:
The first is that the shortcut still performs identity mapping, with extra zero entries padded for increasing dimensions. The second is that the projection shortcut is used to match dimensions (done by 1×1 convolutions). For both options, when the shortcuts go across feature maps of two sizes, they are performed with a stride of 2.



Third set consists of 6 residual blocks where each residual blocks contains two convolutional layers . Each convolution layer has 256 filters of size 3x3.



The fourth set consists of 3 residual blocks where each residual blocks consists of 2 convolutional layers .Each convolutional layer contains 512 filters of 3x3 each.

· After that the feature map is passed through average pooling layer and then it is passes through dense layer containing 1000 neurons to classify 1000 classes.

V. Comparing The Performance of Plain Networks with residual networks

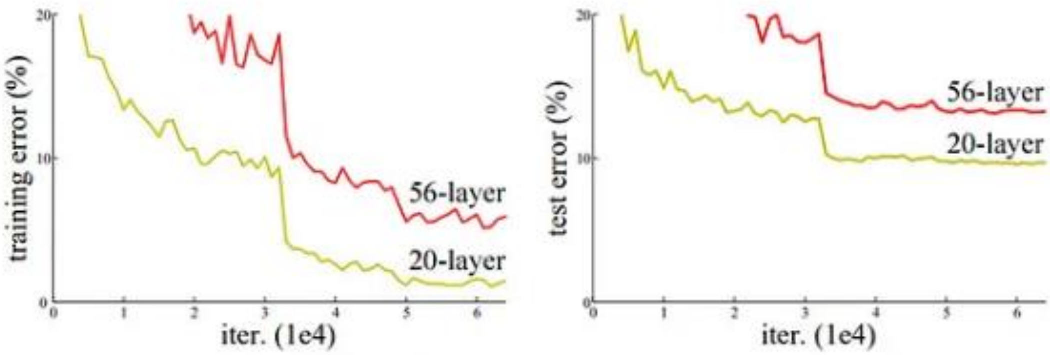


image from original paper-link-https://arxiv.org/pdf/1512.03385.pdf

Above figure shows the training error and test error on CIFAR-10 with 20 layer and 56 layer plain network. We can observe from above that as we are increasing the number of layers both training and test error of 56 layered network are higher than the 20 layered network.

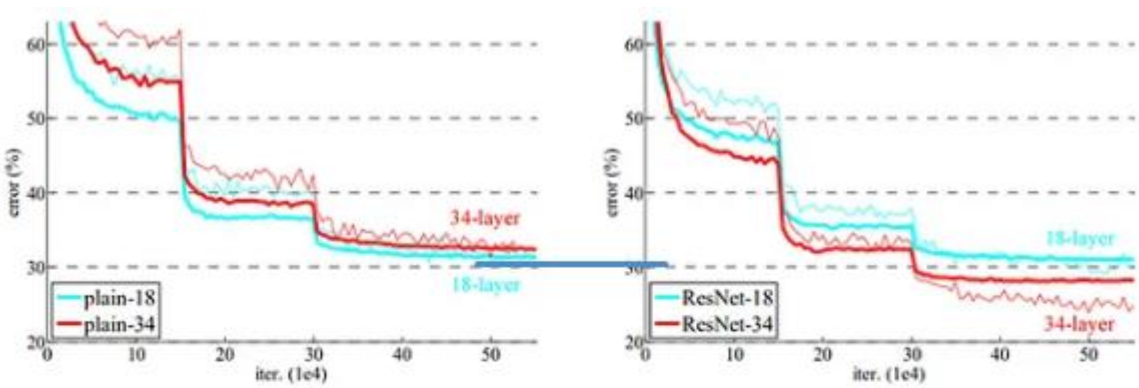


image from original paper-link-https://arxiv.org/pdf/1512.03385.pdf

VI. Conclusion

Residual Networks offer several advantages including improved training of very deep networks,faster convergence and reduced model complexity.

ResNets facilitates the training of deep neural network without suffering from degradation in performance which is the main issue in plain networks.

The skip connections in ResNets allow the network to learn identity mappings making it easier for the model to learn the identity function when needed.

Skip connections helps in optimization process which leads to faster convergence and less training time.

The skip connections in the residual networks leads to better generalization on the unseen data as the network can skip unnnecesary or irrelevant information.

REsNets have consistently demonstrated state of the art performance on the various datasets and their flexibility have made them useful in various other task including object detection,segmentation .Because of these attributes they have been widely adopted for many deep learning task.