LAND OF THE CURIOUS

Indian Institute Of Information Technology, Guwahati

IIITG

PLANT LEAF DISEASE DETECTION

Presented by **Ganesh Kavthekar (Roll No. 2101099)**

May 2, 2024

Guide: **Dr. Nilkanta Sahu**

OUTLINE

IIITG

INTRODUCTION

- India is agriculture dominated country
- Total Foodgrain production in the country is estimated at record 3296.87 Lakh tonnes in the 2022-23, click here
- These crops are threatened by wide variety of plant diseases
- These can damage the crop, lower the vegetable and fruits quality and wipe out the harvest
- About 42 percent of the total agricultural crop is destroyed yearly by diseases, click here

COMPUTER VISION AND IMAGE PROCESSING

Computer Vision and Image Processing comes into role because:

- Most plant diseases show **visible symptoms**, and the technique which is accepted today is that an experienced plant pathologist diagnoses the disease through optical observation of infected plant leaves
- **Drawbacks of pathologist over different techniques of Computer Vision:**
  - Time-consuming process
  - Subjective interpretation of symptoms
  - Limited availability of skilled pathologists
  - Costly for large-scale monitoring

RELATED WORK

IIITG

- The AlexNet is a **8 layers** CNN architecture.On PLANT VILLAGE dataset it gives around **89.33 percent** accuracy[1].
- The VGG16 is a **16 layers** CNN architecture.On PLANT VILLAGE dataset it gives around **96.26 percent** accuracy[1].
- **InceptionV3** architechitecture is **48 layers** cnn model but on the same dataset it give accuracy of **96.26 percent**.
- While increasing the depth of neural networks can potentially capture more complex patterns in data, it also introduces challenges related to **vanishing gradient**.
- **ResNet50** architecture proposed to solve the problem of multiple non-linear layers not learning identity maps and vanishing gradient problem.

RELATED WORK(CONTD.)

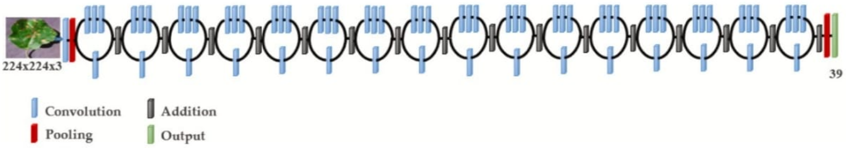

fig 1: **ResNet50**

- The ResNet50 is a **50 layers** CNN architecture.
- On PLANT VILLAGE dataset it gives around **95.44 percent** accuracy[1].
- ResNet (Residual Network) overcomes the problem of vanishing gradients by introducing **skip connections** or **residual connections**.
- Can we further improve the accuracy.....

COMPOUND SCALING

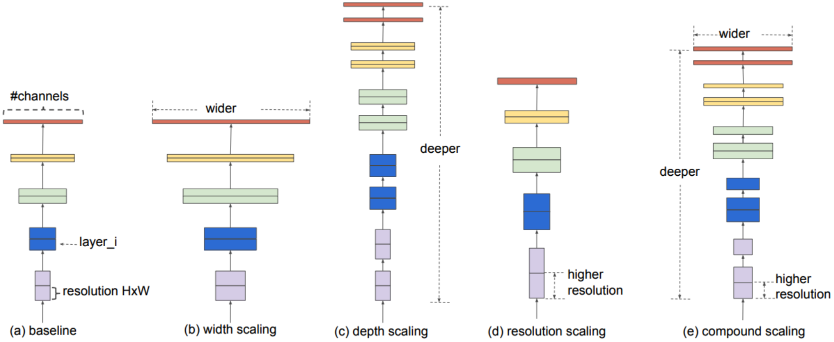

fig 2: **Scaling Model**

- Depth scaling involves increasing the number of layers in a neural network.
- Width scaling involves increasing the number of channels or filters in each layer of a neural network.
- Resolution scaling involves adjusting the input resolution of the neural network[2].

COMPOUND SCALING(CONTD.)

IIITG

- Depth scaling aims to increase the model's capacity to capture complex patterns by adding **more layers**.
  depth scaling
- Width scaling focuses on increasing the information capacity at each layer by adding **more channels**.
  width scaling
- Resolution scaling adapts the network to different input resolutions, capturing **finer details or complex features in the data**.

  resolution scaling
- Scaling up any dimension of network width,depth or resolution improves accuracy, but accuracy gain diminshes for bigger models.
- In order to pursue better accuracy and efficiency, it is critical to **balance all dimesions of network width,depth and resolution during scaling**.

EFFICIENTNET

- To scale the depth,width and resolution , a baseline model is needed which is called **"EfficientNet B0"**.
- Baseline network developed using a **Neural Architectural Search(NAS)**, then scaled up the baseline network to generate a series of models called as **"EfficientNets",B1 to B7**[2].

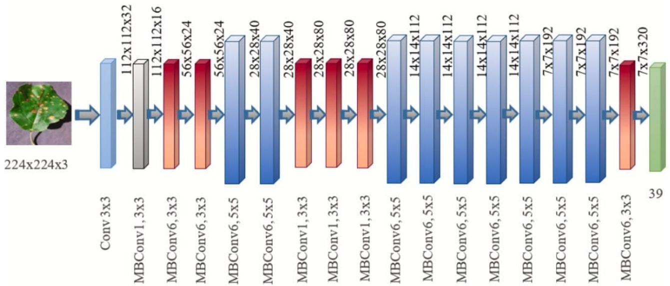

fig 3: **EfficientNet(B0)**

EFFICIENTNET(CONTD.)

IIITG

- EfficientNet introduces a novel scaling method that uniformly scales dimensions of depth, width, and resolution in a compound manner, ensuring optimal performance across various resource constraints.
- EfficientNet architectures are scalable, allowing them to be adapted to different resource constraints. Variants like EfficientNetB0 to EfficientNetB7 offer a spectrum of models with varying depths and complexities.
- It shows the large difference compared to other architectures in the accuracy and number of parameters generated while training.

CBAM

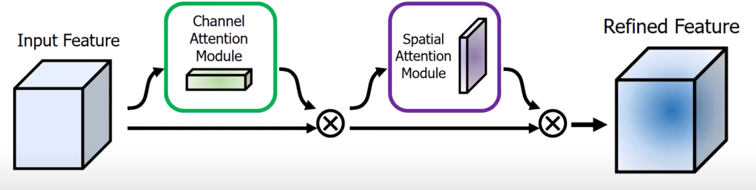

fig 4: **CBAM**

- CBAM comprises two essential modules: Channel Attention and Spatial Attention.
- The Channel Attention module recalibrates feature maps along the channel dimension.
- The Spatial Attention module refines feature maps spatially to emphasize important regions.
- CBAM helps CNN models to focus on relevant features, enhancing interpretability and robustness.

CBAM(CONTD.)

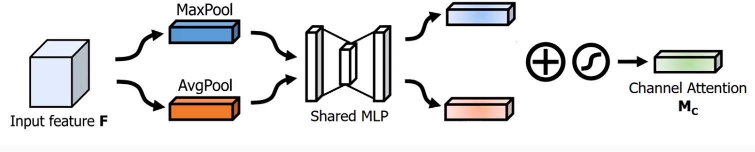

fig 5: **Channel attention**

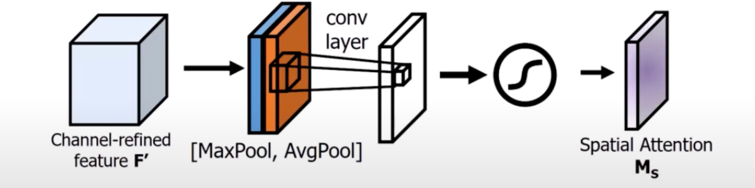

fig 6: **Spatial Attention**

DATASETS

- **PlantVillage Dataset:**

  –In this data-set, 39 different classes of plant leaf and background images are available. The data-set containing 61,486 images.

  –We used six different augmentation techniques for increasing the data-set size. The techniques are image flipping, Gamma correction, noise injection, PCA color augmentation, rotation, and Scaling.

  click here

- **Soynet:**

  –Indian Soybean Image dataset with quality images captured from the agriculture field (healthy and diseased Images).

  –Raw dataset and preprocessed dataset with a resolution of 256x256 pixels present in the dataset.

  –This dataset consists of 9000+ high-quality images of soybeans (healthy and Disease quality).

  click here

RESULTS

Architecture	No. Of parameters	test accuracy	% validation accuracy	loss
VGG16	134,420,327	0.9314	0.9567	0.7122
Alexnet	58,441,127	0.8995	0.9363	1.0201
EfficientNet	4,091,454	0.9526	0.9843	1.1230

Table: Results for differenct architectures without CBAM on Plant Village dataset

Architecture	No. Of parameters	test accuracy	% validation accuracy	loss
VGG16	134,573,398	0.8823	0.9185	0.8832
Alexnet	58,476,496	0.8401	0.8779	1.0448
EfficientNet	11,113,825	0.9137	0.9567	0.7862

Table: Results for differenct architectures with CBAM on plant village dataset

RESULTS ANALYSIS

- The base architectures (AlexNet, VGG16, EfficientNet) may already have sufficient capacity to capture the necessary features for plant disease classification. Adding CBAM may not provide significant additional benefits if the base models can already learn effective representations without attention mechanisms.

- CBAM focuses on capturing spatial and channel-wise attention maps, which might not be significantly beneficial for plant disease classification if the distinguishing features are not effectively highlighted by attention mechanisms.

FUTURE WORK

Cascaded CBAM Integration

- In future work, we plan to enhance our CNN architecture by integrating Convolutional Block Attention Module (CBAM) in a cascaded manner. This involves applying CBAM units sequentially after each convolutional layer, followed by concatenation with the original feature maps.
- By applying cascaded CBAM integration, we seek to advance the efficiancy and robustness of our CNN architecture in image analysis tasks.

TRANSFER LEARNING

- Transfer Learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.
- Integrate transfer learning by initializing pre-trained model, then **fine-tuning or feature extraction** on new plant leaf datasets, leveraging existing learned features to boost accuracy and adaptability.

REFERENCES

» [1] Atila, Ümit, Murat Uçar, Kemal Akyol, and Emine Uçar. "Plant leaf disease classification using EfficientNet deep learning model." Ecological Informatics 61 (2021): 101182.

» [2] Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." In International conference on machine learning, pp. 6105-6114. PMLR, 2019.

» [3] Lu, Jinzhu, Lijuan Tan, and Huanyu Jiang. "Review on convolutional neural network (CNN) applied to plant leaf disease classification." Agriculture 11.8 (2021): 707.