LAND OF THE CURIOUS

Indian Institute Of Information Technology, Guwahati

IIITG

PLANT LEAF DISEASE DETECTION

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

March 22, 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

IIITG

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] 'Emine Uçar'..
- **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.)

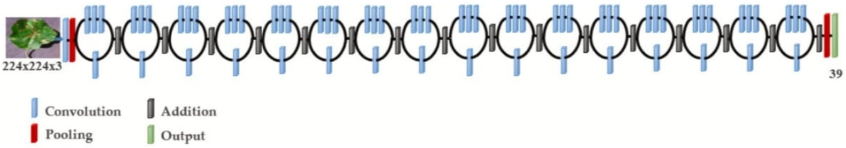

fig 1: **ResNet50**

- 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

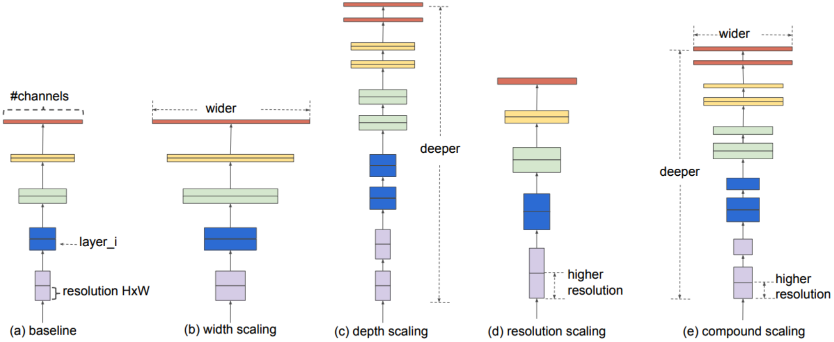

fig 2: **Scaling Model**

- 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

IIITG

- 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].

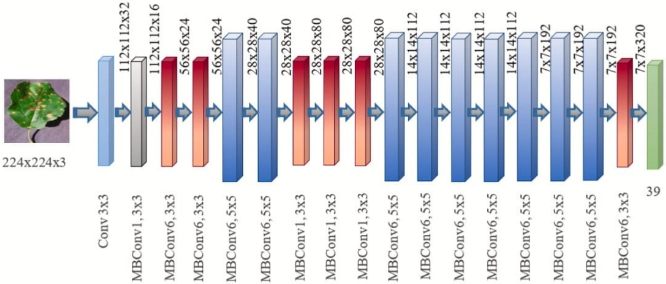

fig 3: **EfficientNet(B0)**

EFFICIENTNET(CONTD.)

- The **MBConv block**,used in EfficientNet, uses the inverted residual block and includes a squeeze and excitation(SE) module to improve its representational power.
- Expansion Phase –> Depthwise Convolution –> Batch Normalization –> Squeeze and Excitation(SE) module –> Skip Connection
- In the expansion phase, the number of channels is increased using a 1x1 convolutional layer.
- Depthwise convolution applies a separate convolutional filter for each input channel.It helps in reducing computational cost and model size by applying convolutions independently across channels.
- Batch normalization is a technique used to address the issue of internal covariate shift. The SE module is integrated into the MBConv block to enhance its representational power.It adaptively recalibrates channel-wise feature responses by explicitly modeling interdependencies between channels.

COMPARISON

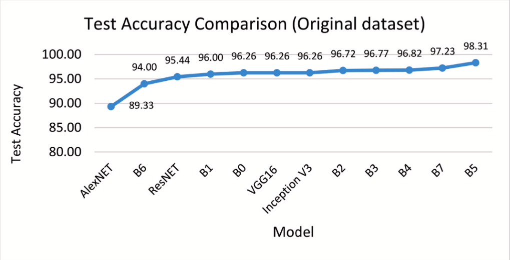

fig 4: **Test Accuracy**

Model name		Input size	Number of total parameters
AlexNet		227×227	60,954,656
ResNet50		224×224	25,636,712
VGG16		224×224	138,357,544
Inception V3		299×299	23,851,784
EfficientNet	B0	224 × 224	5,330,571
	B1	240 × 240	7,856,239
	B2	260 × 260	9,177,569
	B3	300 × 300	12,320,535
	B4	380 × 380	19,466,823
	B5	456 × 456	30,562,527
	B6	528 × 528	43,265,143
	B7	600 × 600	66,658,687

fig 5: **No.Of Parameters analysis**

- These results obtained after training all models on PLANT VILLAGE dataset[1].
- EfficientNet B5 exhibits the highest accuracy, while considerations such as overfitting and hyperparameter tuning become more prominent factors for models B6 and B7.

COMPARISON(CONTD.)

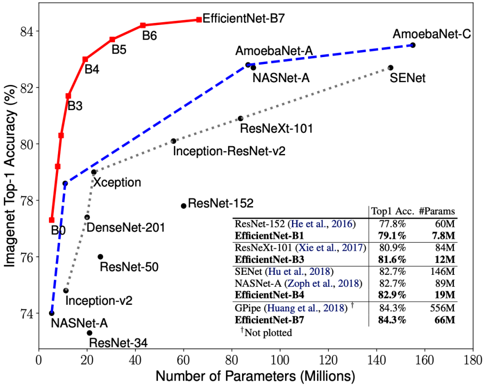

fig 6: **Different Models Analysis**

- These results obtained after training all models on **Imagenet** dataset.
- Here, B6 and B7 gives higher accuracy than B5.
- The number of parameters increases as depth, width, and resolution scale up, impacting computational complexity and memory requirements in the transition from EfficientNet B0 to B1[2]..

DATASETS

IIITG

- **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

- **Dataset for Crop Pest and Disease Detection:**

  The dataset is presented in two folds

  −The raw images which consists of 24,881 images (6,549-Cashew;7,508-Cassava;5,389-Maize and 5,435-Tomato).

  −Augmented images which is further split into train and test set consists of 102,976 images (25,811-Cashew; 26,330-Cassava;23,657-Maize and 27,178-Tomato), categorized into 22 classes

  click here

RESULTS

	no. of parameters	accuracy
basic	7,842,762	75.66
alexnet	58,437,030	78.63
vgg16	151,512,972	–
EfficientNet B0	40,91,454	–

- Vanishing Gradient implementation using the deep network.
- Difference of softmax and relu function on vanishing gradient.

FURTHER IMPROVMENT

IIITG

ATTENTION MECHANISM

- The attention mechanism enables the network to selectively emphasize important features while suppressing irrelevant ones, leading to more efficient use of computational resources.
- The attention mechanism facilitates faster convergence during training by directing the learning process towards the most informative features, leading to improved efficiency in terms of both time and computational resources.

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.