*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

# Inception-VR70: An Advanced Inception-Net Artificial Intelligence based Novel Hybrid Architecture

**[1] Pelluri Venkata Naga Balarama Murthy[1]**
Department of Computer Science and Engineering
Neil Gogte Institute Of Technology
Hyderabad,Telangana, India 500039
e-mail: balaram@kmit.in

**Sai Likhita Rage[2] and Polu V V S Sai kiran[3]**
[2] Department of Computer Science and Engineering (AI & ML)
Keshav Memorial Institute Of Technology
Hyderabad, Telangana, India 500029
e-mail: sailikhitarage.12@gmail.com
e-mail: polukiran7@gmail.com

**Abstract**—In the present day, Convolutional Neural Network (CNN) architectures are undergoing a great deal of development, which has resulted in the creation of models like VGG16, ResNet50, InceptionV3 etc., that significantly increase accuracy. Yet, a network that can deal with overfitting, a significant challenge in deep learning besides having greater accuracy and extracting useful features is required. In this paper, we propose a Deep hybrid model which is an inception of pretrained models with a different input image size, significantly leading to improved accuracy which has been tested on various datasets of different domains including health care, agriculture, and the remote sensing. The performance of this hybrid model is superior to the standalone pretrained models. It is observed that the hybrid model proposed in this paper merely has overfitting despite of having very deep layers compared to all other deep architectures. Best accuracy achieved for this model is 99.64% being train accuracy and 96.33% being test accuracy for the Satellite Images of Hurricane Damage dataset with minimum overfitting.

**Keywords**- Hybrid Deep Learning (HDL), VGG16, ResNet50, Overfitting, Faster convergence.

## I. Introduction

Deep Neural Networks (DNN), a subset of Artificial Intelligence (AI) is a milestone and is applied across varied domains such as Healthcare, sales and marketing, agriculture, military and many more has shown a remarkable achievement in all the sectors. It is a field of great interest to researches and needs much more attention. Deep Learning (DL) relies on Convnets (CNNs), a type of Artificial Neural Network (ANN) for Image Classification, Segmentation and Object Detection. In the following three years, due to various advancements in Deep Convolution Neural Networks (DCNNs) the error rate has lowered to 3.57% [1-4]. Diverse works have been conducted for Image classification using Machine Learning (ML) and DL models.

Deep neural networks are trained by tuning the network parameters in such a way that the mapping improves during the training process [5][6]. Deep networks often attempt to incorporate all of the characteristics [7] from the input picture with classifiers in a comprehensive multidimensional manner. Depending on how the layers are arranged, the proportion of valuable characteristics retrieved at different “phase” of network may be boosted. Further tests have demonstrated the importance of network depth, and the best outcomes were attained by such deep networks.

Single ML or single DL architectures were used for the tasks such as classification and is known as Solo Deep Learning (SDL). Fig.1 depicts the traditional methods of image classification using DL. Subsequent researches started focusing on combining DL architectures with ML models and then combining two SDLs. This is known as Hybrid Deep Learning (HDL) [8][9]. Many works have been proposed on the combination of DL architecture with ML model(classifier) but a combination of two DL models is very rarely seen. There are many architectures such as VGG16, VGG19, ResNet50, DenseNet, InceptionV3 which provide a considerably good accuracy. These are also available as pre-trained models where the model is trained beforehand on a different dataset that further increases the performance. Yet overfitting is a problem for all the Deep Learning architectures.

In this work we intend to propose a fusion model of two pretrained Deep Learning architectures. We propose a hybrid model that is a combination of VGG16 [10] pre-trained and ResNet50 [11] pre-trained models as shown in Fig. 2.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

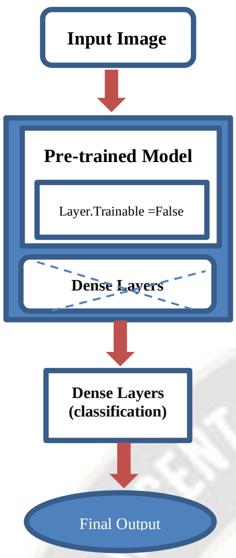


Fig. 1: Traditional Models

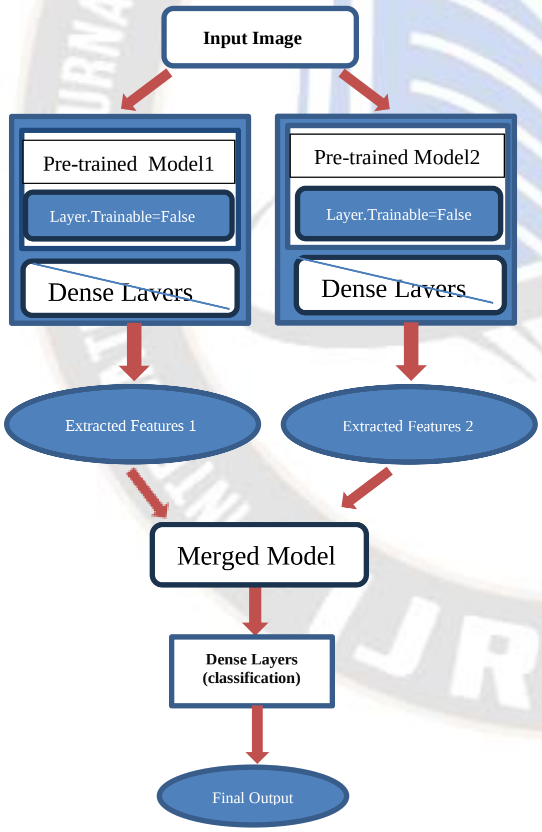


Fig. 2: Inception of two pre-trained models

This paper is further organized as follows

The main motto of this research is cited in section 2 and the contributions of this research are stated in section 3. Previous related contributions of other researchers are analysed under Previous Related Works in section 4. Section 5 comprises of Model development methods, Section 6 testing and evaluating the models. Results of the work are presented in Section 7 and Section 8 concludes the paper.

## II. Motivation

There are many models in combination of a DL architecture and an ML classifier. We aim to propose a hybrid model which is an inception of two DL architectures. Hence forth we implemented various combinations of two Deep Learning architectures so that we can minimize the overfitting as much as possible. Overfitting is the major problem in implementing DL and ML algorithms which leads to wrong predictions. It is important to avoid overfitting in real time applications. Our main intention is to propose a model which gives better accuracy with minimum overfitting and faster convergence and can be applied for all the real time applications.

## III. Contributions

- Faster Converging models with minimum overfitting.
- Greater accuracy besides faster convergence.
- Trained various Hybrid models and analysed their accuracy.
- Tested with various sizes of input images to obtain greater accuracy.

## IV. Previous Related Works

This section includes the relevant earlier efforts and investigations carried out by other scholars. Our method deals with creating a hybrid VGG16 and ResNet50 model that offers the maximum accuracy for all generalized situations with minimum overfitting and faster convergence.

Coursework by Xiaohua Zhai, Xiaohui Shen, Ziwei Liu, Jianping Shi, Chen Change Loy, and Dahua Lin examines several strategies for building residual connections in ResNet50 and assesses the effectiveness of these strategies using the ImageNet dataset. Top-1 accuracy for the paper is 76.35%, and top-5 accuracy is 93.11% [12].

Preeti Gupta and Sachin Meshram worked on Skin Lesion Detection using VGG16 and ResNet50 based hybrid Model and achieved an accuracy of 85.65% [13]. The input shape given to the model was (224,224,3) and the outputs from each model were given to each separate dense layer and then the outputs were combined. In our model we pass a different input image shape and obtained a more improved accuracy.

The contribution of Bharkad et al. (2021) in the Journal of Medical Systems includes a model for Grading Diabetic

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Retinopathy from Fundus Images. This paper used a hybrid model of VGG16 and ResNet50 with an accuracy of 89.9%. [14] Using the ImageNet dataset, Tong He, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li suggested a collection of “tricks” for convolution neural networks that include Resnet50.Top-1 accuracy for the paper is 80.4%, and top-5 accuracy is 95.1% [15].

Liu et al 2019 ’s article in IEEE utilised a Hybrid Deep Learning model for Diabetic Retinopathy Identification and achieved an accuracy of 93.54%, this study employed a hybrid model made up of VGG16 and ResNet50 [16].

## V. MODEL DEVELOPMENT AND METHODS

### 1. Proposed Model

In this study, we used two pre-trained models to create a novel architecture which is the inception of the two pre-trained models. Unlike existing inception architectures where inception of layers is performed, we perform inception of two pre-trained architectures. We performed inception of various combinations of pre-trained models by taking input images of different sizes. After analysing the accuracies of the various combinations, it was found that the inception of VGG16 and ResNet50 gave higher accuracy with least overfitting. We passed input images of sizes 224X224X3, 250X250X3 and 256X256X3 through the proposed architecture. 250X250X3 size of input image performed better compared to other input sizes. The input image of size 250X250X3 is passed parallelly through both the networks (VGG16, ResNet50) independently. *We freeze each layer and remove the dense layers of the pre-trained models*. Automatic Feature extraction from the image is done by the Convolutional layers of each model of the architecture. Since the image is passed through several layers large number of feature maps are extracted. Hence we pool the average value of each feature map using GlobalAveragePooling2D. Consequently, reducing the dimensions and overfitting. At the end of each model Global Average Pooling is applied and then the features obtained from each network are combined. Dense layers for classification are added at end of the combined model. And the features are classified into its respective category by passing through the classification(dense) layers.

In the neural network designed we used layers from the TensorFlow library, Keras Framework that includes Convolutional2D(Conv2D), MaxPooling2D, Batch Normalization (BN), GlobalAveragePooling2D, and Activation layers. 13 convolution layers from VGG16 and 50 from ResNet50(including Convolutional layers and MaxPool, AveragePool layers). Finally we add 4 dense layers. There are a total of 70 layers (including the first Conv layer of ResNet50) in the proposed model as shown in Fig. 3. With varying number of units for each layer, the filter size for each convolution layer is either 1x1 or 3x3. In some areas of the network, MaxPooling layer retains 2x2 pool size and stride 2. By including a Batch Normalization layer in our network, we optimize the complexity, number of trainable parameters, and computation of the network. ReLU activation function is used for efficient transfer of useful information from one layer to the other. *For optimizing the error rate we use Adam optimizer in our proposed model*. Adam converges faster for CNN’s and is computationally efficient with less memory requirements and is suitable for problems with large data or parameters [17]. Traditional Softmax classifier has good performance for multi classification tasks. *Softmax is often used as a loss function combined with Cross Entropy*. If we define the i-th input feature xi and label yi, then the traditional softmax classifier is as follows: [18]

$$L = \frac{e^{f_{yi}}}{\Sigma_j e^{f_j}}$$

In the model developed by Preeti Gupta and Sachin Meshram the maximum accuracy was only 85.65%. It was implemented on only a single dataset of a specific domain with the regular input image size of 224X224X3 taken in the solo pre-trained models. Our model outperforms their model since we take an input image of size 250X250X3 and the useful features are properly extracted. We also tested our model on various datasets of different domains. Our model was also being tested on images of classes from the ImageNet dataset [19] and also obtained a good accuracy of 98%.

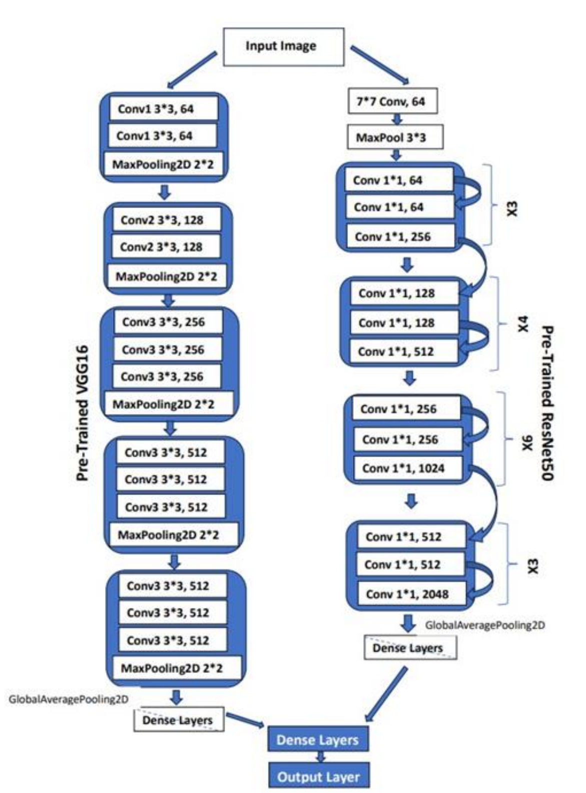


Fig.3: Proposed Model Inception VR70 - Inception of two pre-trained models.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

2. Model Implementation Environment

The datasets are being taken from the public repository “Kaggle”. These models are trained using the GPU P100 accelerator in the Kaggle environment. The models are processed on the servers in the cloud. The time taken was 5.4s for 4000 images with 40 epochs.

## VI. TESTING

The main objective of neural network testing is to confirm that the logic that was imparted to the network throughout the training phase will remain consistent regardless of how frequently we call the application. To better appreciate the adaptability of our proposed neural network we tested this network on datasets across various domains, including healthcare, agriculture, and the sciences.

1. Healthcare Domain

We tested the above network on eye disease classification and Brain Tumor detection. The eye disease classification dataset [20] consists of more than 4200 pictures that have been categorized into four classes: normal, glaucoma, diabetic retinopathy, and cataract.

Constructing such a prediction model aid in illness early detection so that the appropriate measures may be done to diagnose the condition. The test accuracy obtained for this dataset was 92.1% for the proposed model and 82.88%, 86.07% for VGG16 and ResNet50 respectively

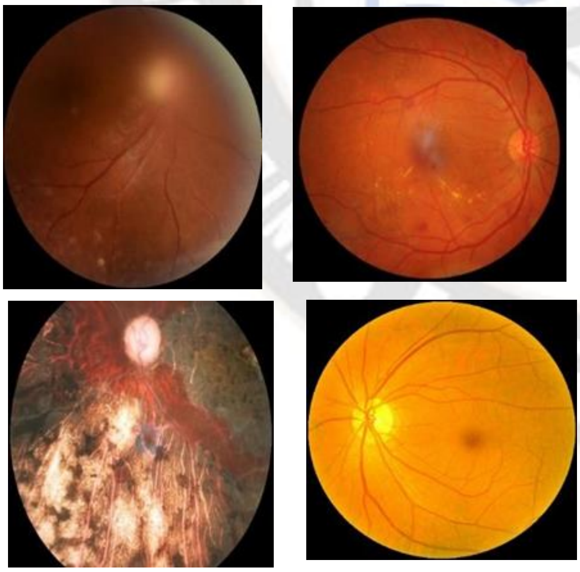

Fig.4: Images of Normal Eye, Glaucoma, Diabetic Retinopathy and Cataract respectively from eye_disease_classification dataset.

2.Agriculture Domain

Training was done using the Corn Leaf Disease Detection Dataset [21][22], which consists of more than 4200 pictures classified into four classes: Common Rust, Grey Leaf Spot, Blight, and Healthy. The model trained over such dataset will give good accuracy for real time applications and help in increasing the percentage yield of the crop. 96.77% test accuracy was observed for this problem whereas 93.15% and 91.53% for VGG16 and ResNet50

Fig.5: Common Rust, GreyLeaf Spot, Blight, Healthy leaf images respectively from Corn leaf disease detection dataset.

2. Sciences Domain

Our network is trained on damage and no-damage classes [23][24] constituting a collection of about 2000 satellite image classifications. Developing a model that can automatically determine if a certain area is likely to experience flooding damage is the aim.

Fig.6: Satellite Images showing Damage or No Damage due to hurricane from “Satellite Images of Hurricane Damage dataset”.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 9***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

## VII. RESULTS AND DISCUSSIONS

### 1. Results Table

The following Table.1 represents the model's performance on the above-mentioned datasets that includes train accuracy, test accuracy, training loss and testing loss.

Table.1: Table comprising the results i.e., training accuracy, testing accuracy, testing accuracy and testing loss for 10 epochs.

| Domain | Dataset | No.of classes | Architecture | Train accuracy | Test accuracy | Train Loss | Test Loss |
|---|---|---|---|---|---|---|---|
| Agriculture | Maize disease | 4 | VGG16 | 98.29 | 93.15 | 0.0407 | 0.2926 |
| Agriculture | Maize disease | 4 | ResNet50 | 98.86 | 91.53 | 0.0333 | 0.2559 |
| Agriculture | Maize disease | 4 | VGG16+ResNet50 | 99.57 | 96.77 | 0.0151 | 0.1547 |
| HealthCare | Eye disease | 4 | VGG16 | 92.7 | 82.88 | 0.1923 | 0.6386 |
| HealthCare | Eye disease | 4 | ResNet50 | 91.41 | 86.07 | 0.2017 | 0.385 |
| HealthCare | Eye disease | 4 | VGG16+ResNet50 | 95.7 | 92.1 | 0.0999 | 0.2248 |
| Remote Sensing | Satellite Hurricane damage | 2 | VGG16 | 99.56 | 95.8 | 0.0166 | 0.1284 |
| Remote Sensing | Satellite Hurricane damage | 2 | ResNet50 | 99.31 | 96.6 | 0.023 | 0.1076 |
| Remote Sensing | Satellite Hurricane damage | 2 | VGG16+ResNet50 | 99.64 | 96.33 | 0.0187 | 0.087 |

### 2. Graphs

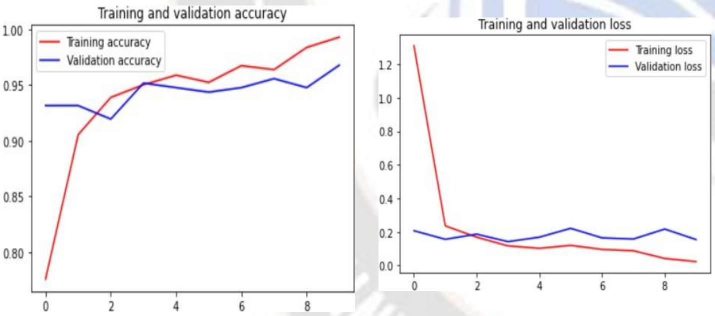


Fig.7: Train accuracy, Test accuracy and Train loss, Test loss of proposed model (Inception-VR70).

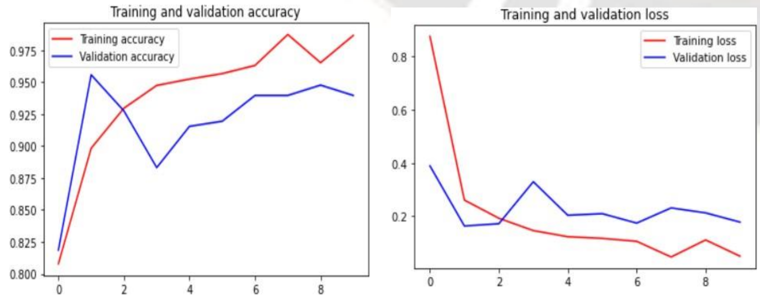


Fig.8: Train accuracy, Test accuracy and Train loss, Test loss of VGG16 model.

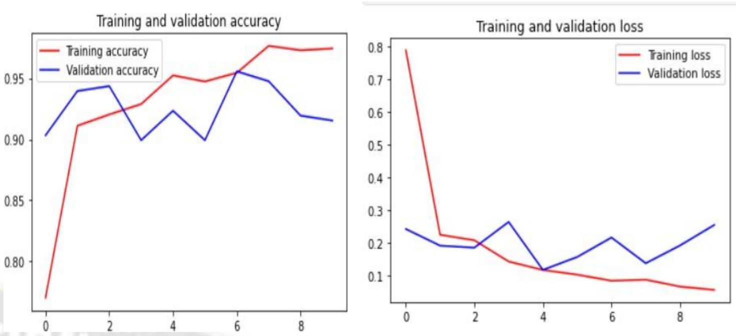


Fig.9: Train accuracy, Test accuracy and Train loss, Test loss of ResNet50 model.

### 3. Evaluation


Fig.10: Model prediction for Common Rust

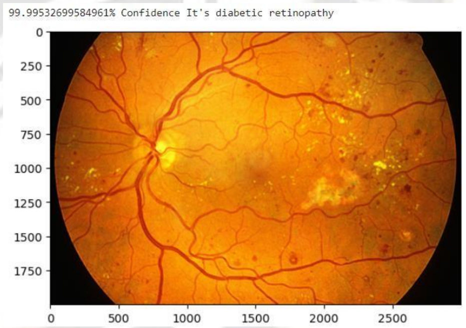


Fig.11: Model prediction for Diabetic Retinopathy

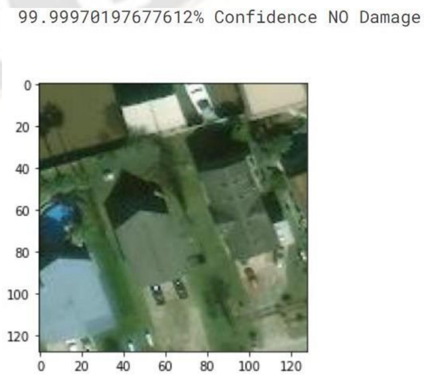


Fig.12: The Model predicted accurately as No Damage.

## VIII. Conclusion

Inception of two pretrained architectures is used for Image Classification. We performed inception on various combinations of pre-trained models such as Alexnet + VGG16, AlexNet + ResNet50. But the combination of VGG16 and ResNet50 gave best accuracy and accurate predictions on evaluation. Since the image is being passed through several layers of two deep pre-trained models, a greater number of features are extracted which enables the model to be trained on large number of varied features and allows to learn useful features. Since the images are being passed through very deep networks the input size taken (250,250,3) results in good accuracy. During Evaluation the maximum evaluation accuracy obtained for unseen images was 100%. Among all the domains we tested on our model, the highest accuracy was obtained on agriculture domain being 96.77% test accuracy. For all the domains our model resulted in least cost function.

The model presented in this study can be further used for Object detection and Object Localization tasks other than Image Classification. Further works can be done by working on various combinations of inception of pre-trained models. This advanced inception model can be further optimized computationally. It can also be extended in the fields of digital pathology and bioinformatics.

## IX. References

1. K. Simonyan and A. Zisserman. Very deep convolutional networksfor large-scale image recognition. In ICLR, 2015.
2. M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional neural networks. In ECCV, 2014.
3. K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers:Surpassing human-level performance on imagenet classification. InICCV, 2015.
4. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D.Anguelov, D. Erhan, V. Vanhoucke, and A.Rabinovich. Going deeper with convolutions. In CVPR, 2015.
5. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature 521, 436–444. doi:10.1038/nature14539
6. Schmidhuber, J. (2015). Deep learning in neural networks: an overview. Neural Netw.61, 85–117. doi: 10.1016/j.neunet.2014.09.003
7. M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional neural networks. In ECCV, 2014.
8. panelBiswajit Jena a, Sanjay Saxena a, Gopal K. Nayak a, Luca Saba b, Neeraj Sharma c, Jasjit S. Suri d Artificial intelligence-based hybrid deep learning models for image classification: The first narrative review.
9. A new hybrid deep learning model for human action recognition Author links open overlay panelNeziha Jaouedi a, Noureddine Boujnah b, Med Salim Bouhlel c
10. Karen Simonyan & Andrew Zisserman part of Visual Geometry Group, Department of Engineering Science, University of Oxford cited VGG16 model in "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION".
11. As part of Microsoft Research Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun developed the pretrained model of Deep Residual Learning for Image Recognition.
12. "Exploring Residual Connections for Image Recognition" by Xiaohua Zhai, Xiaohui Shen, Ziwei Liu, Jianping Shi, Chen Change Loy, and Dahua Lin helped improve accuracy.
13. International Journal Of Creative Research Thinking (IJCRT) published a work by Preeti Gupta and Sachin Meshram (ISSN:2320-2882).
14. "A Hybrid Deep Learning Model for Diabetic Retinopathy Grading from Fundus Images" by Bharkad et al. (2021) in the Journal of Medical Systems
15. Bag of Tricks for Image Classification with Convolutional Neural Networks" by Tong He, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li.
16. "Automated Diabetic Retinopathy Detection Using a Hybrid Deep Learning Model" by Liu et al. (2019) in the IEEE Access.
17. Quoc Dung Cao (University of Washington), Youngjun Choe (University of Washington): https://ieee-dataport.org/open-access/detecting-damaged-buildings-post-hurricane-satellite-imagery-based-customized.
18. A softmax classifier for high-precision classification of ultrasonic similar signals Author links open overlay panelFei Gao, Bing Li, Lei Chen, Zhongyu Shang, Xiang Wei, Chen He State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710000, China International Joint Laboratory for Micro/Nano Manufacturing and Measurement Technology, Xi'an Jiaotong University, Xi'an 710000, China.
19. Jia Deng; Wei Dong; Richard Socher; Li-Jia Li; Kai Li; Li Fei-Fei: ImageNet: A large-scale hierarchical image database.
20. These images are collected from various sources like IDRiD, Oculur recognition, HRF etc.
21. Singh D, Jain N, Jain P, Kayal P, Kumawat S, Batra N. PlantDoc: a dataset for visual plant disease detection. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD 2020 Jan 5 (pp. 249-253).
22. J, ARUN PANDIAN; GOPAL, GEETHARAMANI (2019), "Data for: Identification of Plant Leaf Diseases Using a 9-layer Deep Convolutional Neural Network", Mendeley Data, V1, doi: 10.17632/tywbtsjrjv.1
23.Kaggle-https://www.kaggle.com/datasets/kmader/satellite-images-of-hurricane-damage.
24. Quoc Dung Cao (University of Washington), Youngjun Choe (University of Washington): https://ieee-dataport.org/open-access/detecting-damaged-buildings-post-hurricane-satellite-imagery-based-customized.