Image classification neural networks

Department of Computer Sciences, American International University-Bangladesh

**Abstract:**

Computer vision is a collection of scientific methods that deals with training of computers to interpret the visual world. It is an AI technology. Deep learning is an artificial intelligence (AI) function that tries to follow the workings of the human brain in processing data and developing patterns for making decisions. CNNs are used for image classification and recognition because of its high accuracy. The CNN follows a hierarchical model which works on building a network, like a funnel, and finally gives out a fully-connected layer where all the neurons are connected to each other and the output is processed. Here we discussed few CNNs and showed their structure and also how do they work. We tried to figure out which CNN has advantage over others and why.

**Introduction:**

With the rapid development of deep learning technology theory and the improvement of computer performance and hardware cost, people require that the accuracy of machine vision recognition of image targets to be higher and higher. Especially in the fields of precision strike, road monitoring, product quality monitoring, and automatic driving, there is a high requirement for recognition accuracy, because misidentification may cause unpredictable consequences. Convolutional neural networks have good advantages in image classification, and have achieved excellent results in many objects recognition tasks, mainly because their network structure itself can extract multi-level features of images.

ResNet Won 1st place in the ILSVRC 2015 classification competition with a top-5 error rate of 3.57% also Won the 1st place in ILSVRC and COCO 2015 competition in ImageNet Detection, ImageNet localization, Coco detection and Coco segmentation. So, we can easily guess that ResNet models were extremely successful.

Dense Nets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters.

VGG16 is a network model based on convolutional neural network proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper “Very Deep Convolutional Networks for Large-Scale Image Recognition”.

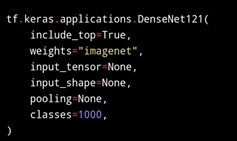
VGG19 is a variant of VGG model which in short consists of 19 layers (16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer). There are other variants of VGG like VGG11, VGG16 and others. VGG19 has 19.6 billion FLOPs.

**Literature Review:**

Image classification is a crucial step in the object detection and image analysis. Alex Net was the first convolutional network to employ the graphics processing unit (GPU) to improve performance.

The success story of image classification with CNN started with Alex net 26 27 Krichevsky et al. [2012]. To learn about thousands of objects from millions of images, we need a model with a large learning capacity. However, the immense complexity of the object recognition task means that this problem cannot be specified even by a dataset as large as ImageNet, so our model should also have lots of prior knowledge to compensate for all the data we don’t have. Convolutional neural networks (CNNs) constitute one such class of models. Their capacity can be con-trolled by varying their depth and breadth, and they also make strong and mostly correct assumptions about the nature of images (namely, stationarity of statistics and locality of pixel dependencies). Thus, compared to standard feedforward neural networks with similarly-sized layers, CNNs have much fewer connections and parameters, and so they are easier to train, while their theoretically-best performance is likely to be only slightly worse. Despite the attractive qualities of CNNs, and despite the relative efficiency of their local architecture, they have still been prohibitively expensive to apply in large scale to high-resolution images. Luckily, current GPUs, paired with a highly-optimized implementation of 2D convolution, are powerful enough to facilitate the training of interestingly-large CNNs, and recent datasets such as ImageNet contain enough labeled examples to train such models without severe overfitting. Since then, many authors have contributed on up- 27 28 dating the architectures.

**Dense Net** (Dense Convolutional Network) is reviewed. This is the paper in 2017 CVPR which got Best Paper Award with over 2000 citations. It is jointly invented by Cornwell University, Tsinghua University and Facebook AI Research (FAIR).



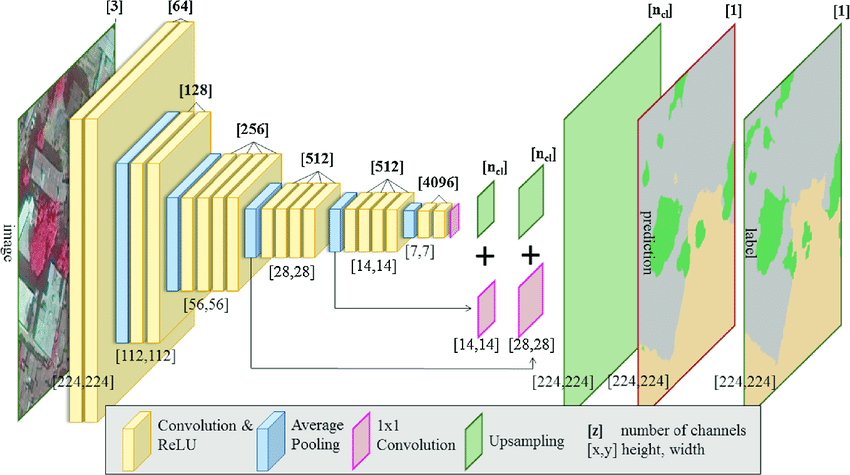
Now, talk about the **Resnet**:

It was introduced in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun in their paper “Deep Residual Learning for Image Recognition". And Won 1st place in the ILSVRC 2015 classification competition with a top-5 error rate of 3.57% also Won the 1st place in ILSVRC and COCO 2015 competition in ImageNet Detection, ImageNet localization, Coco detection and Coco segmentation. So, we can easily guess that ResNet models were extremely successful.

Some of the major papers are VGG-Net Simonyan 28 29 and Zisserman [2014], Google Net Szegedy et al. [2015], ResNet He et al. 29 30 [2016], MobileNet Howard et al. [2017] etc.

VGGNet is a Convolutional Neural Network architecture proposed by Karen Simonyan and Andrew Zisserman from the University of Oxford in 2014. ... You can find the original paper of VGGNet which is titled Very Deep Convolutional Networks for Large Scale Image Recognition.

**Proposed Method:**



**Figure: VGG19**

**Results:**

*Image Classification* is a fundamental task that attempts to comprehend an entire image as a whole. The goal is to classify the image by assigning it to a specific. Alex Net is a convolutional neural network that is 8 layers deep.

**Alex Net** is considered one of the most influential papers published in computer vision, having spurred many more papers published employing CNNs and GPUs to accelerate deep learning.

**RESNET** is a platform providing real estate solutions to banks and servicers. To gain access to this platform subscriptions are sold to real estate agents which provide exposure and potential business.

**DenseNet** is the newest convolutional-type neural network that uses dense block to build connection in layers.

***VGG16*** is a pre-trained model that takes in (224,224) RGB images and converts them into features. It comes out-of-the-box from the keras library ..

**VGG19** is a variant of VGG model which in short consists of 19 layers (16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer). There are other variants of VGG like VGG11, VGG16 and others

Compared with **VGG16**, **VGG19** is slightly better but requests more memory. **VGG16** model is composed of convolutions layers, max pooling layers, and fully connected layers. The total is 16 layers with 5 blocks and each block with a max pooling layer.

Above all the discussion and everything is subject to the fact that VGG 19 is the best models for Image Classification.

**Discussion:**

Throughout the report, the deduction and result are solely made on research papers published on image classification. As the procedure of judgment is not same for all, the result may vary. The factor we focused during the research which leads us to the decision that VGG 16, 19 must be the best convolutional neural network for image classification among the other options.

**Conclusion:**

We used Convolutional Neural Network (CNN) for image classification which contains to extract features and max pooling to decrease the size of image thus classifies the image accurately. The results can be made more accurate by increasing the number of convolution layers and hidden neurons. People can recognize the object from blurry images by using our model. Image classification is an excellent prototype problem for learning about neural networks, and it gives a great way to develop more advanced techniques of deep learning. In the future, we are planning to develop a real-time image classification system.

**Reference:**

1. <https://www.mygreatlearning.com/blog/resnet/>
2. <https://paperswithcode.com/method/dense-block>
3. [Architecture of AlexNet and its current use (opengenus.org)](https://iq.opengenus.org/architecture-and-use-of-alexnet/)
4. <https://iq.opengenus.org/vgg19-architecture/>
5. <https://www.geeksforgeeks.org/vgg-16-cnn-model/>

|  |  |  |
| --- | --- | --- |
| **Name** | **ID** | **Contribution** |
| Khan,Anamul Haque | 17-33696-1 | Abstruct,Introduction[ResNet] |
| Mery,Ummul Rubaiyat Afroz | 18-38390-2 | Literature Review,Proposed method,Result[VGG16&19] |
| Hossen, Md Shakil | 18-36689-1 | Conclusion[AlexNet] |
| Neha, Mitheela Farzana | 18-36649-1 | Dicussion[DenseNet] |