Classification of Image using Convolutional Neural Network

Ridhima Dhawan
Assistant Professor, School of Computer Science and Engineering,
Lovely Professional University.

Abstract:

Image classification, a topic of pattern recognition in pc vision, is an method of classification based totally on contextual statistics in pics. "Contextual" approach this technique is focusing on the connection of the nearby pixels, that's also referred to as neighbourhood. The intention of this approach is to classify the pics by using the contextual facts.

This paper presents the study of three different convolutional network for recognition of the objects in the images and videos. The Alex Net, Google Net, ResNet50 has been used for classification of the images. The commonly used datasets used for images is the CIFAR-100. The above table, honestly outlines that Google Net and ResNet50 has higher precision in classifying the gadgets in the picture in evaluation to Alex Net

Introduction

At the present time, the internet is jam-packed with plenty of images and videos, which has become the comprehensive subject of code driven image processing procedures [1] which is a method of meting out the images by using algorithms. The Image classification forms an integral fragment of image processing. The objective of image classification is the automatic classification of images into the thematic groups [2] therefore leading to object detection and classification problems.

The procedure of image classification includes two steps: training of the images and testing. The training of the images and testing means classifying the image into particular class pertaining through unique characteristics on the basis of training characteristics.[3].

The Convolutional Neural Network (CNN), a machine learning algorithm is being castoff for programmed classification the images. CNN's are proficiently and successfully used in many pattern and image recognition applications, such as, recognition of gestures, recognition of face, classification of object [4] and producing descriptions for the scenes. They can be found at the core of everything from tagging of photos on social media to self-driving cars. They're working effectively behind the scenes in everything from healthcare to security. The working of CNN is described in the section 2.

Working of CNN

Neural networks are built of the layers where each layer is connected to every other layer, thus creating the network. It is also called as the feed forward neural network (FFNN) where the neurons are connected with the each other through input layer and output layer, consisting hidden layers which are the intermediate between these two layers. By adjusting weights of the connection using back propagation, neurons learn. The weight adjusting procedure is called recursively till input layer is updated by the weight layer, to which it is connected.

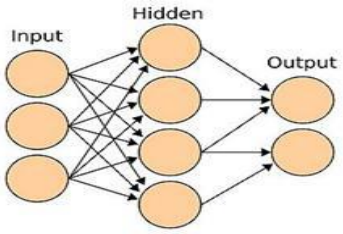

Figure 1: feed-forward neural network architecture [5]

Convolutional Neural Network (CNN) is neural-network designed artificially, inspiring from connection of neurons in the animal cortex, aimed at analysing 2D pictures proficiently and effectively [6]. CNN are neural networks which is based on the sharing of the parameters. CNN contains convolution layers, a set of learnable filters. Every filter has width, height and depth as that of input volume (3 if the input layer is image input)[i]. These CNN filters, together forms convolutional and sub-sampling layers, having sequence among them, which collectively yield an estimate data of image provided as an input. The weights of same colour are shared, thus constrained to be identical.

Neurons in layer m are connected to neurons of layer (m-1) which has adjoining receptive fields. CNN applies local connection pattern among the neurons of adjacent layers, hence exploiting correlation spatially.[3]

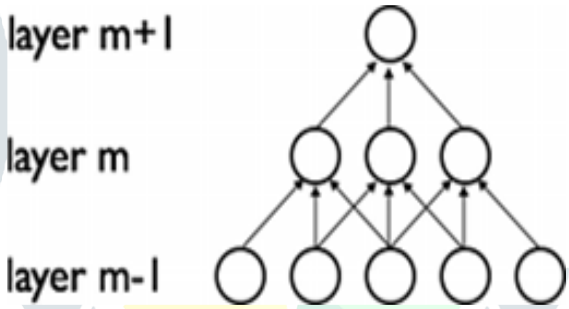

Figure 2: control flow of layers outlining linking between layers [3]

Categories of CNN layers:

1. **Input Layer:** A raw image is provided as an input containing width and height 32 each and deepness of 3.
2. **Convolution Layer:** a dot product is calculated between the filters and patch of an image. For an instance, we have 8 filters for a layer then output will 32*32*8
3. **Activation Function Layer:** An activation function such as sigmoid, RELU is applied by this layer.
4. **Poolling Layer:** This layer is intermittently interleaved in the covnets therefore reducing the size of volume thus reducing memory and preventing it from overfitting.
5. **Fully-Connected Layer:** by taking output of the previous layer as input, it yields 1D array whose size is equal to the number of classes.

CNN Models

The Alex Net, Google Net and ResNet50 are the most popular CNN neural network which are used for classification of objects from the images.

The construction of those networks' contrasts in terms of internal layers and techniques used. GoogleNet accomplish numerous sizes of convolutions which concatenate the filters for the following layer. This is carried out with the aid of Inception components [7]. In evaluation to Alex Net, it makes use of output of the preceding layer due to the fact the input to the layer [8]. ResNet is a short name for Residual Network. In residual learning, instead of seeking to learn some features, we try to learn some residual. Residual is a subtraction of function found from enter of that layer. ResNet does this with the useful resource of right away connecting enter of nth layer to a few (n+x)th layer. By using this technique, training of units to CNN has become simpler which has result in upgradation within the accuracy rate [9].

Comparative Study of CNN Models

500 pictures are separated into four hundred coaching images and eighty testing images for every category, therefore, creating a complete of 40000 numerous pictures. These eighty groups are organized into fifteen super categories. the chosen classes for coaching and testing are shirts, trousers, skirts, shoes, watches and wardrobe.

CIFAR-100		Alex Net	Google Net	ResNet50
Image classes	Shirts	0.0%	70.1%	50.2%
	Trousers	22.4%	73.9%	56.2%
	Boots	84.9%	62.8%	36.8%
	watches	00.0%	73.12%	33.5%
	skirts	23.6%	79.2%	34.5%
	jeans	31.5%	94%	92%
	jackets	32%	78%	34%
	ballerina	87%	73%	63%

Fig 3: Performance of CNN on CFIR-100 test datasets

CIFAR-100		Alex Net	Google Net	ResNet50
Image classes	Fruits	6.0%	70.1%	49.82%
	Vegetables	32.4%	74.9%	53.2%
	chairs	94.9%	68.8	36.8%
	table	10.0%	73.12%	34.5%
	computers	28.6%	79.2%	37.5%
	televisions	27.3%	28%	36.5%
	aeroplanes	35%	25%	35%
	laptops	47.56%	89%	78%
	mobiles	41.5%	95.78%	93.4%

Fig 4: Performance of CNN on CFIR-100 test datasets

Results

The prediction accuracy rate is done by testing the various images using different CNN network. Following table outline the accurate rate of the images.

	CLASSES	ALEX NET	GOOGLE NET	RESNET50
PREDICTION ACCURACY RATE	Shirts	45	35	55
	Trousers	16	17	23
	Boots	0	4	7
	watches	2	7	18
	skirts	1	17	10
	ballerina	3	35	7
	jeans	4	3	6
	televisions	0	0	0
	laptop	6	18	8

Fig 5: Accuracy rate of various images done by CNN Models

The above table, clearly outlines that Google Net and ResNet50 has better precision in classifying the objects in the image in comparison to Alex Net,

Conclusion

The toil analysed the prediction accuracy of 3 different convolutional neural networks (CNN) on most notable education and check datasets significantly CIFAR100. We have a tendency to targeted our study on fifteen categories of every dataset solely. the elemental endurance was to seek out the accuracy of the not like networks on same datasets and examination the consistency of prediction by each of those CNN. This has been obtainable thorough prediction cross-check for examination the networks' performance for varied lessons of pictures.

Bibliography

[1] Kou, F., Du, J., He, Y., & Ye, L. (2016) "Social Network Search Based on Semantic Analysis and Learning." CAAI Transactions on Intelligence Technology

[2] Lillesand, T.M. and Kiefer, R.W. and Chipman, J.W., in "Remote Sensing and Image Interpretation" 5th ed. Wiley, 2004

[3] Sowmya V., P., S. K., and Deepika, J., "Image Classification Using Convolutional Neural Networks", International Journal of Scientific & Engineering Research, vol. 5, no. 6, p. 06/2014,

[4] Sharma, N., Jain, V., & Mishra, A. (2018). An analysis of convolutional neural networks for image classification. Procedia Computer Science, 132, 377-384.

[5] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

[6] Ramprasath, M., Anand, M. V., & Hariharan, S. (2018). Image Classification using Convolutional Neural Networks. International Journal of Pure and Applied Mathematics, 119(17), 1307-1319.

[7] Ballester, P., & deAraújo, R. M. (2016, February)"On the Performance of GoogLeNet and AlexNet Applied to Sketches." in *AAAI*.

[8] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., &Fei-Fei, L. (2014)"Large-scale video classification with convolutional neural networks." in Proceedings of the IEEE Conference on *Computer Vision and Pattern Recognition*.

[9] Srinivas, S., Sarvadevabhatla, R. K., Mopuri, K. R., Prabhu, N., Kruthiventi, S. S., &Babu, R. V. (2016) "A taxonomy of deep convolutional neural nets for computer vision."

[i] https://www.geeksforgeeks.org/introduction-convolution-neural-network/