iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 Issue: X Month of publication: October 2018

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue X, Oct 2018- Available at www.ijraset.com

Classification of Images on Furniture and Household Goods by using Transfer Leaning and Fine Tuning

Kingshuk Das Bakshi[1], Dr Gagan Deep[2]
[1]Research scholar, [2]Professor Chandigarh Engineering College

Abstract: Automatic product recognition for shoppers in online shopping is a challenging task. The reason behind that is, for the same product, a picture can be taken in different intensities of light, angles, backgrounds and levels of occlusion. This causes the different fine-grained categories look very similar. Many of general-purpose recognition machines used now days, cannot perceive such subtle differences between photos.

These differences could be important for shopping decisions. In this paper, a novel approach has been proposed based on deep learning and artificial neural networks (ANN) for pattern recognition, which accurately assigns category labels for furniture and home good images.

This is done by classification of textual patterns, to help push state of art in automatic image classification. In deep learning, transfer learning is used, where two pre trained convolutional neural network (CNN) models are retrained. The CNN models used for this experiment are VGG-16 and Inception V3.

The experiment is carried on dataset taken from kaggle and classification is made among five items named bed, sofa, table, chair and swivel chair. The experimental results are measured by performance metrics, in terms of training accuracy, validation accuracy, training loss and validation loss. The results demonstrate that the accuracy of Inception V3 transfer learning model with 97.3% is more than VGG-16 and ANN with accuracy of 92% and 86%, respectively.

Index terms: Convolutional neural networks (CNN), texture classification, transfer learning, computer vision.

I. INTRODUCTION

Automatic product recognition for online shopping is a difficult task for customers. This looks much more obvious while buying furniture and household items.

While searching the product via internet manually, there is a possibility of wrong identification of product. Many products may look similar, and hence in these cases the customer may go wrong in finding the product.

For example in between ball chair and egg chair in furniture, or between dutch oven and a cook oven in cookwares, there can be a possibility of mistake as these products look very similar.

The reason behind that is for the same product, a picture can be taken in different intensities of light, angles, backgrounds and levels of occlusion.

This makes different fine-grained categories look very similar. Many of general-purpose recognition machines used now days cannot perceive such subtle differences between photos.

These differences are important for shopping decisions [1].

To this end, no such model has been prepared by any organization so that it will help the customers to buy products of absolutely the same design which he is wanting to get.

To deal with this issue a novel deep learning and artificial neural network is proposed in this paper [2]. Three automated systems are made by which the probability of making a correct assessment of products may increase.

If a buyer wants to buy an arm chair of a particular design, then if he uploads the image of that arm chair in the website of the seller, the model, belonging to a particular company, will try to identify the photo at first and then try to match the photo with any available resource which the company has, and then it uploads the details of the product to the customer.

In this process, the customer does not have to take the burden of searching for the exactly similar model of chair by himself.

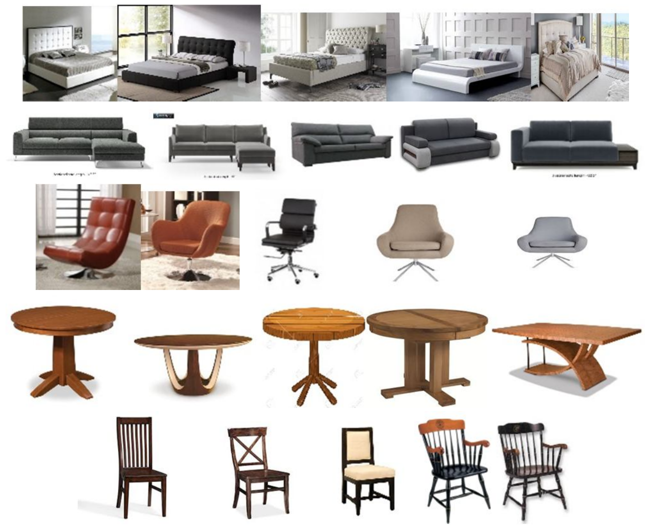


Fig 1. Typical pictures of household furniture. From top to bottom: bed, sofa, swivel chair, table, chair.

In order to assist the customer with the correct product which he has wished to buy, the model can be a big help. This is basically a computer aided system which helps in recognition of products by the model, instead of manual recognition of products by the bare eye. In deep learning, convolutional neural networks (CNN) are used [3]. Two pre trained CNN models named VGG-16 and Inception V3, are trained by transfer learning procedure [4], [5], [6]. In this paper, an experiment is done on recognition of five household image items by using the above mentioned techniques. The items are bed, sofa, table, chair and swivel chair.

II. RELATED WORK

This section is provided with a brief overview of previous efforts done in improving image classification of furniture and households, so that the customers could buy the products of their liking via the internet.

A. Autonomous Household Robots

This paper [7] proposes a mapping system machine. This machine take pictures of 3 dimensional objects, which are man-made such as kitchen, segments the images of doors, tables, drawers, and shelves, and then reconstructs geometrically. The model also acquires images of objects of daily use such glasses, plates, and similar ingredients. The model enables the recognition of the objects in cluttered scenes and the classification of newly encountered objects. The technical details behind this technology include an accurate, robust, and efficient algorithm for constructing a complete object model using three-dimensional point clouds, for constituting views of partial object. The recognition procedures are based on feature-detection and extraction for cabinets, tables, and other task-relevant furniture objects [6]. The feature extraction system also helps in the automatic inference of object instance and class signatures for objects of daily use. This enables robots to reliably recognize the objects in cluttered and real task contexts.

B. Mobile Robot

This paper [8] proposed the use of mobile robot to detect and catch the target object. The technology behind this is the using of CCD (charged coupled device) camera with an embedded system. The techniques used for image processing are Histogram, Image Spatial Resolution and Connectivity (Neighboring Pixels). These techniques are simply practical to be used with the embedded system for detecting the target object. They can be applied to mobile robots. The performance will be a small task with high flexibility and low cost. There was ease in construction and the procedure was energy saving.

C. Mobile Robot using Aforge.NET Framework

This paper [9] proposed the use of mobile robot using A forge dot NET Framework along with C# language in order to classify and recognize objects. In the detection of the target object, image processing techniques such as Blob Counter and HSL Color space are used to distinguish the object from the background and determine the position of the target object on the image. This application can be applied to enhance various capabilities on the mobile robot such as detection of the object, object following and rotation direction control of robot's camera.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue X, Oct 2018- Available at www.ijraset.com

III. MATERIALS AND METHODOLOGIES

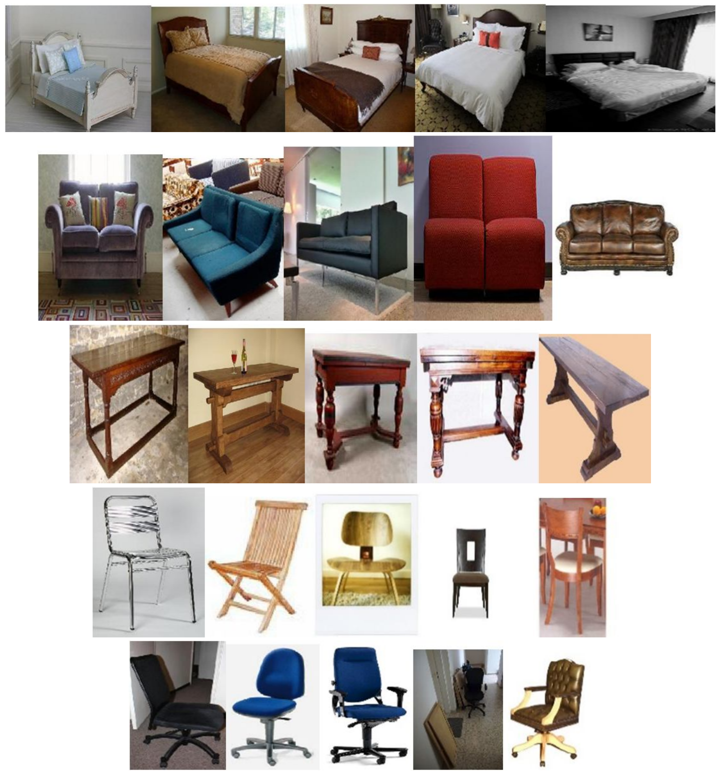


Fig 2. Typical samples from dataset for training of images. From top to bottom: bed, sofa, table, chair, swivel chair

A. Database

Dataset required in this research has been provided by kaggle. Kaggle had organized a conference on Computer Vision and Pattern Recognition (CVPR) in which it has asked highly focused and aspired data scientists to differentiate on five furniture household goods based on their shapes based on fine-grained visual categorization. It had called for a workshop which is named as the FGVC5 workshop. In this workshop, CVPR partnered with Google, Malong Technologies and Wish. They challenged the data scientists to develop algorithms which would be used

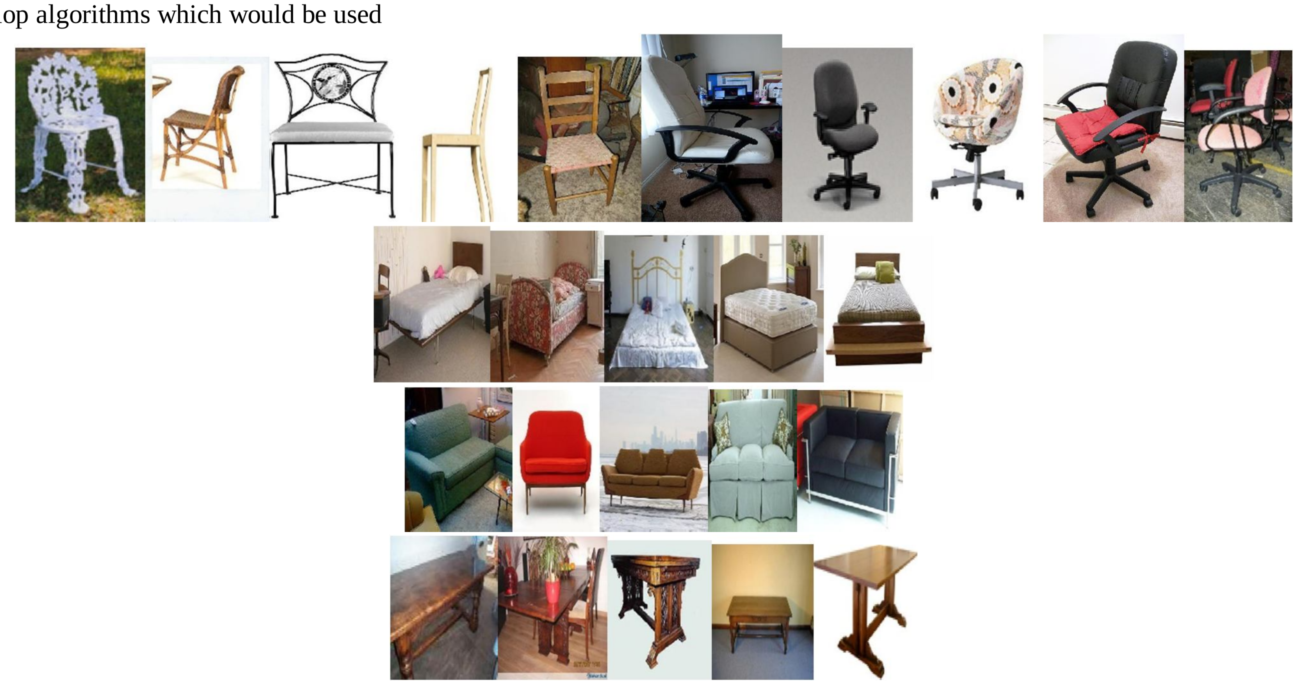


Fig 3. Typical samples from dataset for testing of images.

From top to bottom: chair, swivel chair, bed, sofa, table as a help for automatic product recognition. Kaggle partnered with research groups to expand the boundary of machine learning. Kaggle's platform is used by research competitions and research group's data science team [10]. The training dataset includes images from 5 items of furniture and home goods classes. The five types of images include that of bed, sofa, chair, swivel chair and table. It includes a total of 194,828 images for training and 6,400 images for validation and 12,800 images for testing. That makes 91% images belongs to training, 9% images for validation or testing testing. The table2 shows the distribution of objects which have been classified.

Table 1. Number of images for training and testing

Name of substance	No of images for training	No of images for testing
Bed	38947	3852
Sofa	38951	3853
Chair	38953	3851
Swivel chair	38955	3853
Table	38958	3853

Table 2. Traditional layers of CNN

Convolutional layer
ReLu layer
Average pooling layer
Flattening layer
Dense layer
Softmax layer

Table 3. Layers of an Inception V3 model

convolution layer
average pooling layer
max pooling layer
concatenation layer
dropout layer
fully connected later
softmax layer

Table 4. Number of objects for each item Methodology

Name of the object	No of objects
Bed	42799
Sofa	42804
Chair	42806
Swivel chair	42808
Table	42811

This proposed model consists of transfer learning based approach with the framework of keras and tensorflow library at the backend. Using transfer learning approach two pre trained CNN models are trained. The CNN models are VGG-16 and Inception V3 model. An artificial neural network is also developed.

Table 3. Number of images for each case

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue X, Oct 2018- Available at www.ijraset.com

B. CNN

Convolutional neural networks comes under the domain of deep learning. It is a representation learning method that can automatically discover features from a picture (raw data), which is suited for a particular task [11]. The features extractors are task specific. They are not fixed for a particular task each and every time. There are a number of layers in each network that lead hierarchical features used in the learning process.

CNNs consists a number of layers. They are convolutional layers, pooling layers, rectified linear unit layer (ReLU), pooling layer, flattening layer, dense layer and softmax layer [13]. The role of the former is to extract local features from a set of learnable filters and the role of the latter is to merge neighbouring patterns. This is done by reducing the spatial size of the previous representation and adding spatial invariance to translation. CNN is a hierarchical neural network and accuracy of its working is dependent on the design and training methods of the layers.

Some popular CNNs available in the Caffe framework [14] are Le Net [15], Alex Net [16], Google Net [17], VGG-16, and Inception V3 model. Here pretrained CNN models have been used. They are VGG-16 and Inception V3 model, which are trained in specific domain by transfer learning.

C. Transfer Learning

Transfer learning is defined as an ability of a system to utilize
the knowledge, which it has acquired from one domain (example: classification of image in medical field like cancer or identification of different animals) to another domain that shares some common characteristics [18]. In this paper, a supervised transfer learning procedure with CNNs, have been proposed. Deep CNNs have shown abilities in transferring the knowledge between apparently different image classification tasks. They have also shown abilities in transferring knowledge between imaging modalities for the same task. This is generally done by weight transferring. A network is pretrained with a particular task and then weights of some of its layers are transferred to another network that is used for another task. In some of the cases, activations of this second network are used as “off-the-shelf” features and it can be fed to any classifier. In some other cases, the non transferred weights of the network are initialized. This will be followed by a second training phase on the target task [18]. While this training is being done, the transferred weights are kept frozen or inactive at their initial values and in other cases all the weights are trained together with random weights. This process is called “fine-tuning”. When the target dataset is too small with respect to the capacity of the network, fine tuning will result into over fitting, hence the features are left frozen. The number of layers to be transferred depends on the proximity of the two tasks and on the proximity of the corresponding imaging modalities. The last layers of the network are task specific and the earlier layers are modality specific [19]. If there is no case of overfitting issues, each and every layers can be transferred and fine tuned [20]. By doing so the performance of doing all these tasks gets increased. This process is applied in training data for some tasks are limited. Transfer learning is applied on two pretrained CNN models named VGG-16 and Inception V3 model.

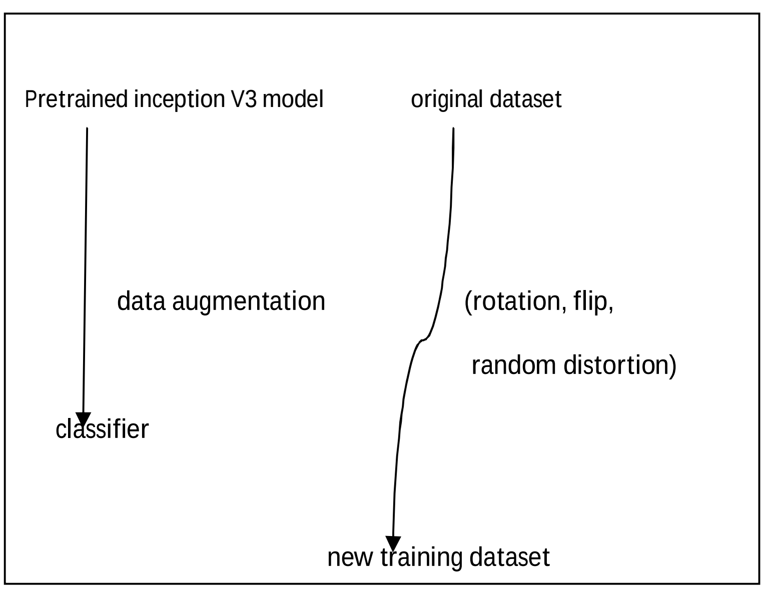


Fig 5. Workflow of proposed method of transfer learning using inception v3 model and data augmentation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue X, Oct 2018- Available at www.ijraset.com

D. VGG- 16 Model

The VGG-16 is a type of conventional neural network. It consistsof 16 main layers. Thirteen of them are convolutions, and three are fully connected layers. All the convolutional layers on VGG-16 have the same filter size. The layers include ReLU layers, max pooling layers, fully connected layers, and dropout layers.

1) *C.1 Workflow:* The procedures for building model for furniture and household classification are as follows.

a) Preprocessing of data: The steps carried out under data preprocessing are

b) Division of dataset: Dataset contains a total number of 214028 images. Dataset is divided into two parts, training and testing. Training dataset contain images of table, chair, swivel chair, bed and sofa. These images work as input for vector conversion.

c) Vector conversion: Image to vector conversion means conversion of images into array of RGB images. This is done as deep learning require numerical values as input. Size of each vector is 150*150. The reason behind it is input shape to the model is 150*150. Library used for this is OpenCV.

d) Normalization: Normalization is done in order to bring every pixels in array to the same scale. This is done in order to make faster converging of pixels and removal of any redundancy pixel.

X=X/255

Y=Y/255

2) *Splitting Between Train And Test Set:* After data preprocessing is done, dataset is divided into training and testing in the ratio 91:09 which means 91% of images are in training set and 09% in test set as shown in Table x. Test set can be considered as validation set. Hyper parameters tuning of model will be done according to validation set.

Feature Extraction/Model Training: VGG16 model is used for feature extraction as shown in Table 6. It is one of the pretrained model which is trained on larger dataset of imagenet comprising of 1000 categories. Each layer in this model is carrying out features. As for example, first and second layer are carrying out low level features such as horizontal edges or vertical edges etc.

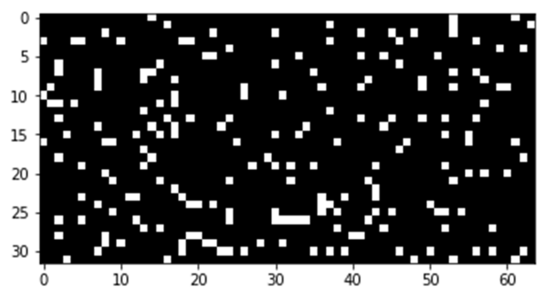


Fig 6. Features Extracted by the Model

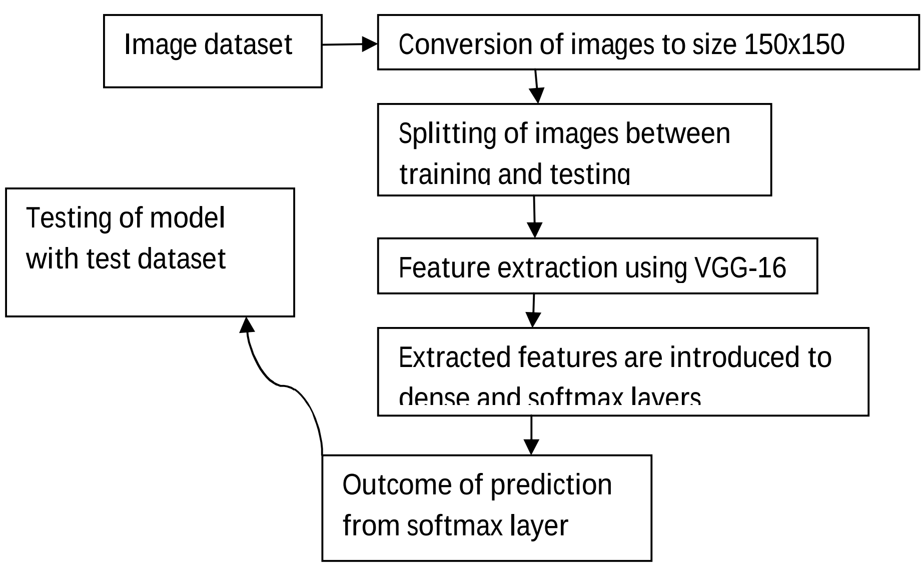


Fig 7. Workflow of VGG-16 model

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue X, Oct 2018- Available at www.ijraset.com

Table 6. Training and testing data

Training data	194828
Testing data	19200

3) *Transfer Learning:* While transferring knowledge in this model, the lower layers are freezed and higher layers are used for training. Top two layers are removed and new dense layers are added according to the number of categories to be classified into as shown in below Table 5.

Table 5. VGG-16 Feature extractor

Input Layer	(None,150,150,3)	0
Conv 2D	(None,150,150,64)	1792
Conv 2D	(None,150,150,64)	36928
MaxPooling 2D	(None,75,75,64)	0
Conv 2D	(None,75,75,128)	73856
Conv 2D	(None,75,75,128)	147584
MaxPooling 2D	(None,37,37,128)	0
Conv 2D	(None,37,37,256)	295168
Conv 2D	(None,37,37,256)	590080
Conv 2D	(None,37,37,256)	590080
MaxPooling 2D	(None,18,18,256)	0
Conv 2D	(None,18,18,512)	1180160
Conv 2D	(None,18,18,512)	2359808
Conv 2D	(None,18,18,512)	2359808
MaxPooling2D	(None,9,9,512)	0
Conv2D	(None,9,9,512)	2359808
Conv2D	(None,9,9,512)	2359808
Conv2D	(None,12,12,512)	2359808
MaxPooling2D	(None,4,4,512)	0
Dense	(None,4,4,512)	37750784
Flatten	(None,8192)	0
Output	(None,5)	1285

4) *Fine Tuning:* In this process, some of the lower layers are made inactive and top two layers are removed. Addition of new layers are done according to the dataset. The newly added layers are kept for training.

After the extraction of all the features, these features are transferred to the next layer. At first they are flattened, then they are transferred to the fully connected layers and then to softmax layers for prediction.

Model is trained for 4 epochs and for 6 iterations to get the appropriate result.

5) *Model Evaluation:* Testing of model is done in this phase, with the help of test dataset. Comparison is done between test and train accuracy. The main goal is to achieve negligible difference between test and train accuracy.

Negligible difference between train and test accuracy and both being above 80% means, the prepared model is perfect. Necessity of hyper parameters tuning is not there, otherwise there is a need.

Inception V3 model

Inception-v3 model consists of 22 neuron layers. These layers are trained on large ImageNet dataset which contains 1.2million images categorized into 5 different classes. It is a network composed of convolutional layers assembled upon one another. There are two max-pooling layers to halve the resolution of the grid.

The model accepts images of size150x150x3 as input. The output layer is a soft max layer of 5 neurons, since the classification is being done between 5 classes of image. This model shows convolution model whose kernel is larger than 3x3, can be expressed more efficiently by a series of smaller convolution. This can be done by replacing large 7x7 filters with a pair of 1x7 and 7x1

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue X, Oct 2018- Available at www.ijraset.com

convolutional layers. In order to fine-tune the inception-v3 model, the count of layers in the final classification layer is modified. This is done to correspond to the given dataset. All the weights from the pre-trained Inception-v3 model are restored except for the final classification layer. Weights at that layer will get randomly initialized. Finally, the training of the network with training images of required domain has been carried out using the tensorflow library. The machine used is of NVIDIA GeForce 830M GPU and8GB memory. The figure(x) represents an Inception-v3 model.

Artificial Neural Networks

The artificial neural network is a multilayer perceptron architecture. The network used here has one input layer, three hidden layers and one output layer. All these layers are fully connected to their previous layer. The input layer consists of M*M neurons. This corresponds to a window to be displaced over the preprocessed input image. The figure 9 represents the model of an artificial neural network. In the above figure 9, artificial neural network model is there which is classifying between five items of bed, table, chair, swivel chair and sofa. Here, three hidden layers are used for carrying out the prediction for classification between the five items. In each hidden layer, 10 units are used. The output layer has five neurons, and its output will be given by a softmax activation function, which is used for carrying out the probabilities of the items. This value will be considered as a degree of confidence.

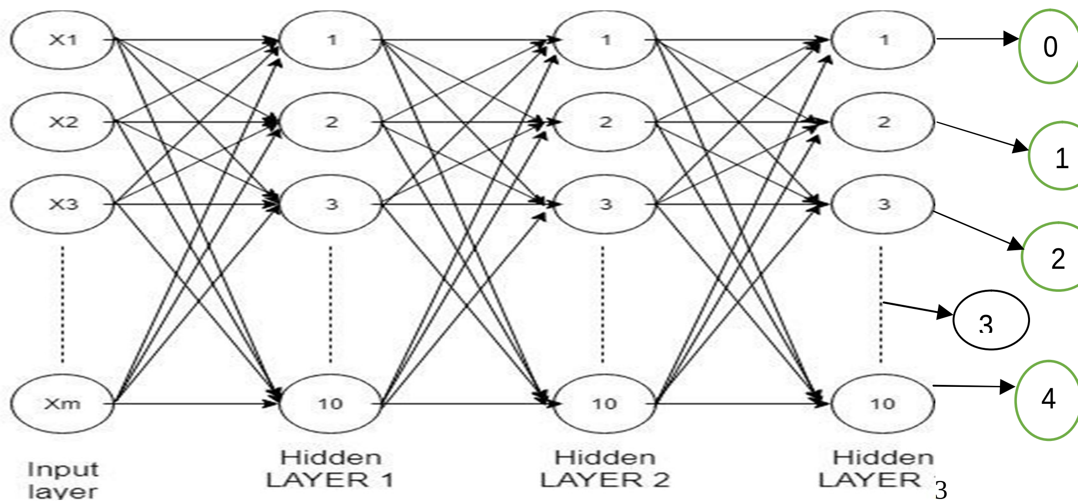


Fig 9.Artificial neural network

a) *Dataset of images:* In this step dataset is divided into five folders as table, chair, sofa, chair, swivel chair. Images are labelled as 0 for table, 1 for chair, 2 for swivel chair, 3 for sofa and 4 for bed. Each image is cropped and converted into 100*100 images. This dataset is then introduced into the model as input.
b) *Conversion of image to vector:* The dimension of each image is converted into (100*100*3, 1). For sample of m images, the dimension of data is (100*100*3, m). Array of vectors are carried out since, deep learning models require continuous value numbers as input.
c) *Splitting Of Data Between Training And Testing:* Dataset is now split into training and testing data. The size of train data is 91% of the whole data and that of validation is 2.9% for validation. 5.9% are used for testing.

Table 5. Ratio of Split Dataset

Dataset	Ratio
Training Data	91%
Testing Data	9%

6) *Data Normalization:* Training set of the data is normalized so that gets converged faster. If input training vectors do not get scaled then the ranges of distribution of features, would likely become different for each feature.
7) *Model Training:* This is the most important phase of the procedure where the preprocessed training set images are fed into the neural network model. Then each hidden layer shown in Table 3 carry out different function with features. In doing so, the model learns about those features. In this phase two processes are included.
a) *Forward Propagation*: In forward propagation, features are extracted from hidden layers. Equations 1 and 2 are used at each hidden layer. For prediction purpose, softmax activation function is applied by using the equation 3.

$z_1 = w * x + b$ (1)
$a_1 = relu(z)$ (2)
$y = softmax(a_1)$ (3)

b) *Back Propagation:* After computing the error in prediction, the error is propagated backwards to update weights along with bias, in order to reduce the overall loss. This procedure is known as back propagation. Equation 4, equation 5, equation 6 and equation 7 are used in this step.

$dZ^{[l]} = dA^{[l]} * g^{[l]'}(Z^{[l]})$ (4)
$dW^{[l]} = (dZ^{[l]} . A^{[l-1]T})/m$ (5)
$db^{[l]} = 1/m \sum dZ^{[l]}$ (6)
$dA^{[l-1]} = W^{[l]T} . dZ^{[l]}$ (7)

Table 6. Neurons in hidden layers

1st Hidden Layer	10 Neurons
2nd Hidden Layer	10 Neurons
3rd Hidden Layer	10 Neurons

8) *Model Evaluation:* In this step, different evaluation measurements are carried out. The evaluation metrics used here are accuracy and loss metrics and based on that, the overfitting state of model if known. If model is having the issue of overfitting, retraining of the model with regularization technique is required.

IV. RESULTS

We have conducted the different furniture and household recognition experiments on the single data set provided in kaggle. The whole data set is divided into two parts as train, validation or test set. Ninety one percent of the data set has been used for training and 9% has been employed for validation. Data augmentation and alignment have been applied on the training part of the data set. In the experiments, for deep CNN model training, images have been resized to 256 × 256 pixels resolution. These 256 × 256 sized images are cropped into five different images during the training phase and a single crop is taken from the center of the image during the test phase. The crop image size for VGG 16 and it is 227 × 227.

TABLE 7. Comparison in between proposed methods.

	Training accuracy	Validation accuracy	Training loss	Validation loss
VGG 16	90 %	92 %	0.30	0.225
Inception v3	98.2 %	97.3%	0.20	0.125
ANN	85 %	86 %	0.50	0.40

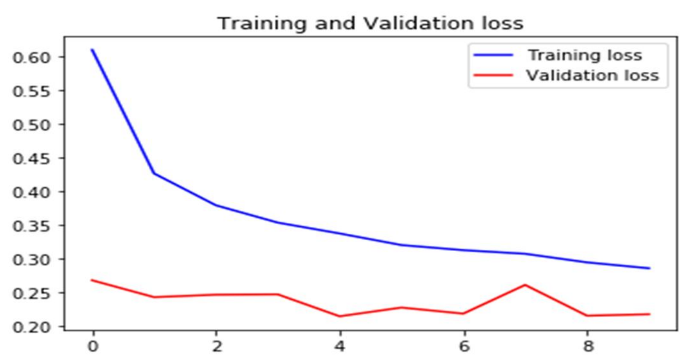


Fig 11. Graph showing training and validation loss for VGG-16.

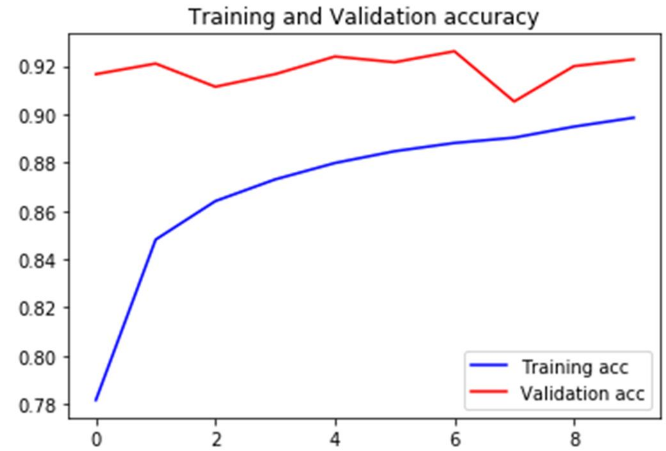


Fig 10. Graph showing training and validation accuracy for VGG-16.

The performance of the ANN models are assessed first on the collected data set from kaggle. VGG-16 [7],and Inception V3 architectures have been employed. They are fine tuned by using their pre trained models that were trained on the Image Net data set [16]. The obtained results on the test set are listed in Table x. In the table, the first column contains the name of the model, the second one contains the corresponding train loss, third column contains train accuracy, fourth column indicates test loss and fifth column indicates test accuracy. Augmentation and alignment have been applied in case of transfer learning applications. As can be seen, the achieved classification rates are quite high in case of transfer learning applications. Inception-V3 model is found to perform the best. Data augmentation has contributed around 1% to the accuracy. Alignment did not lead to an improvement. However, this point requires further investigation.

V. CONCLUSION

In this study, several aspects of recognition of furniture and households have been proposed. This work contains a comparison on performance on performance between transfer learning and artificial neural networks. For this approach, we have first taken a data set consisting of household and furniture images from kaggle. At first stage, the pre trained deep CNN models named VGG-16 and Inception V3, which were trained on the ImageNet, are fine tuned with the dataset of household and furniture items. Hence domain adaptation for the pre trained deep CNN models is done. After that artificial neural network topology is designed. The extensive experiments have been conducted on kaggle dataset. It is shown that performing two-stage fine-tuning is very beneficial for furniture and household classification and recognition. With data augmentation and without alignment, for Inception V3 model, the correct classification accuracy is of 97.3% . For VGG-16 and ANN model, the accuracy is of 92% and 86% respectively. This performance indicates the importance of transferring a pre trained CNN model from a closer domain. It is observed that the best performance is achieved by performing transfer learning of Inception V3 model and then on VGG-16 model. The performance is higher than the one obtained with ANN model. For future work in automatic furniture and household image detection, other transfer learning models like Resnet can be used for improvement in accuracy.

REFERENCES

[1] iMaterialistic challenge (Furniture) at FGVC5 image classification of furniture and home goods, kaggle. [Online].Available:https://www.kaggle.com/c/imaterialist-challenge-furniture-2018

[2] S. Christodoulidis, M. Anthimopoulos, L. Ebner, A. Christe, and S. Mougiakakou, "Multisource transfer learning with convolutional neural networks for lung pattern analysis ," IEEE Trans. Med. Imag., vol.21, no.1, pp.76-84, Jan. 2017.

[3] M. Anthimopolous, S. Christodoulidis, L. Ebner, A. Christe, and S.Mougiakakou, "lung pattern classification for interstitial lung diseases using a deep convolutional neural network," IEEE Trans. Med. Imag., vol. 21, no. 1, pp. 1207-1216, May 2016.

[4] K. Simonyan and A. Zisserman. (Sep. 2014). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: https://arxiv.org/abs/1409.1556.

[5] J. Krause, B. Sapp, A. Howard, H. Zhou, A. Toshev, T. Duerig, J. Philbin, and L. Fei-Fei. "The unreasonable effectiveness of noisy data for finegrained recognition". arXiv preprint arXiv:1511.06789, 2015.

[6] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition." in Icml, vol. 32, 2014, pp. 647–655.

[7] R. B. Rusu, N. Blowdow, Z. Marton, A. Soos, M. Beetz "Towards 3 D object maps for autonomous household robots", International Conference on Intelligent Robots and Systems, pp. 3191-3198, Dec. 2007.

[8] P. Palungsuntikul and W. Premchaiswadi, "Object detection on keeping a mobile robot by using a low cost embedded color vision system ," 2010 Eighth International conference on ICT and knowledge engineering, pp. 70-76 , Jan. 2011.

[9] K. Kungcharoen, P. Palangsantikul and W. Premchaiswadi, "Development of object detection software for a mobile robot using an AForce.Net framework ," 2011 Ninth International Conference on ICT and Knowledge Engineering, pp. 201-206, Jan. 2012.

[10] iMaterialistic challenge (Furniture) at FGVC5 image classification of furniture and home goods, kaggle. [Online].Available:https://www.kaggle.com/c/imaterialist-challenge-furniture-2018

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, pp.1097-1105, 2012,

[12] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in Information Processing in Medical Imaging, pp. 588-599, 2015.

[13] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition", Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

[14] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J.Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe; Convolutional architecture for fast feature embedding," in Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014, pp. 675-678.

[15] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition", Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

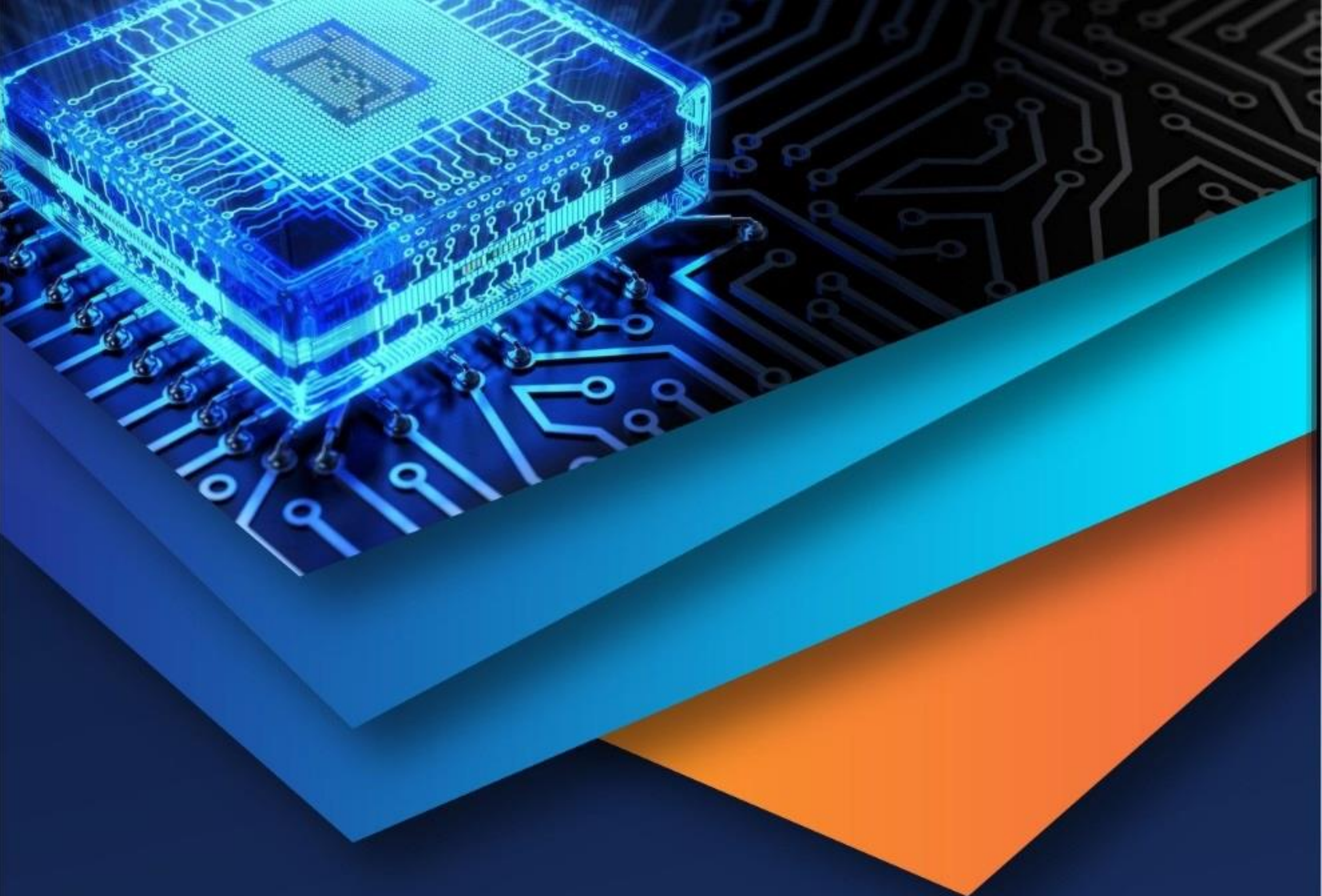
[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, pp.1097-1105, 2012,

[17] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2015, pp. 1-9.

[18] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition." in Icml, vol. 32, pp. 647–655, 2014.

[19] Yosinski, J., Clune, J., Bengio, Y., et al.: "How transferable are features in deep neural networks? ". Advances in Neural Information Processing Systems (NIPS), 2014, pp. 3320–3328, 2014.

[20] Gross, R., Matthews, I., Cohn, J.F., et al.: 'Multi-PIE'. IEEE Int. Conf. on Automatic Face and Gesture Recognition (FG), 2008.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY