iJETRM

International Journal of Engineering Technology Research & Management

IMAGE FORMATION BASED ON TRANSFER LEARNING USINGCNN

Dr.S.Kannan[1]and Immaculate Rexi Jenifer[2]
[1]Department of Computer Science &Engineering,
E.G.S Pillay Engineering college,Nagai
[2]Department of Computer Science &Engineering, AnjalaiAmmalMahalingam Engineering College,Thiruvarur

ABSTRACT

Image classification is a popular machine learning based applications of deep learning. Deep learningtechniques are very popular because they can be effectively used in performing operations on image datain large-scale. In this paper CNN model was designed to better classify images. We make use of featureextraction part of inception v3 model for feature vector calculation and retrained the classification layerwith these feature vector. By using the transfer learning mechanism the classification layer of the CNNmodel was trained with 20 classes of Caltech101 image dataset and 17 classes of Oxford 17 flower imagedataset. After training, network was evaluated with testing dataset images from Oxford 17 flower datasetandCaltech101 imagedataset.ThemeantestingprecisionoftheneuralnetworkarchitecturewithCaltech101datasetwas 98%andwithOxford17Flowerimagedatasetwas92.27%.

KEYWORDS

ImageClassification,CNN,DeepLearning,TransferLearning.

1. INTRODUCTION

In thismodern era, Computers are being powerful day by day.They have become perfectcompanion with high speed computing capabilitiesoverthe time.Fewdecadesago itwasbelieved that machines are only for arithmetic operations but not for complex tasks like speechrecognition,objectdetection,imageclassification,languagemodelingetc.Butthesedays,situati on is inverted. Machines are capable of doing these thingsmore easily with very muchhigh accuracy. Usual algorithm consisting of finite arithmetic operations cannot provide capacitytodosuchcomplextasksformachine.Forthis,ArtificialIntelligenceprovideslotsoftechniques. Learning Algorithms are used for such purpose. Huge dataset is required for trainingthe model with appropriate architecture. Testing is required to evaluate whether the model isworking properly or not. Neural Network is one of AI techniques emerged long ago in 1940s buttechnology at that time was not so advanced. It was up at time in 1980s with the development ofback-propagation[1].Lateritwasagaindiscardedduetoslowlearningandexpensivecomputation. Initially It was believed that only 2 to 3 hidden layers are sufficient for NeuralNetwork to work

iJETRM

International Journal of Engineering Technology Research & Management

properly but later on it is observed that even more layers can represent highdimensionalfeaturesoftheinputsignals[2].Imageclassificationhasreceivedextensiveattentionsincetheearlyyearsofcomputervisionresearch.Classificationremainsmainproblems in image analysis. CNNs are applicable to fields like Speech Recognition[3], textprediction[4],handwritinggeneration[5]andsoon.

In These days, RAM on a machine is cheap and is available in plenty. we need lots of labeleddata, time and GPU speed to train a CNN model from scratch[6]. With this limitation, it is notfeasible fortraining the CNN network from scratch, because itisacomputationally intensivetask and it takes several days or even weeks with high speed GPU computer[7], which is notpossible with limited resource we have. Defining appropriate model for image classificationswhich will produce good result in small training time and minimum CPU speed is the main taskthatthispaperisintendedtodo.

2. RELATEDWORKS

In[8],NovalgeneralKnearestneighborclassifierGKMNC[9]wasusedforvisualclassification. Sparse representation based method[10] for learning and deriving the weightscoefficients and FISTA[11] was used for optimization. CNN-M[12], a pretrained CNN was usedfor image features extractions then marginal PCA[13] is applied to reduce the dimension of theextracted features. In[14], Alex net[15] model, a deep neural network is used to learn sceneimage features. During the training phase, series of transformation such as convolution, maxpooling, etc are performed to obtained image features. Then two classifier SVM[16] classifierandSoftmax[17]classifieraretrainedusingextractedfeatures fromtheAlexNetmodel.

In[18], Spatial pyramid pooling was used in CNN to eliminate the fixed size input requirements.forthisnewnetworkstructureSPP-
netwasused,whichcangenerateafixedlengthrepresentation regardless of image size. Standard back propagation algorithm[1] was used fortraining, regardless of the input image. In[19], Kernalized version of Nave bayesNeabour[20]was used forimage classification and SVM[16] classifierwas trained on Bag-Of-Features[21]for visual classification. In[22], Extension of the HMAX[23], a four level NN has been used forimageclassifications.Thelocalfiltersatfirstlevelareintegratedintolastlevelcomplexfiltersto
provide aexible description of object regions. In[24], Nearest neighbor classifier[20] wasusedforvisualclassifications.SIFT[25]descriptortodescribeshape,HSV[26]valuestodescribecol orsandMRfilterstodescribetexturewere used.

3. METHODOLOGY

3.1. ImagePreprocessing

Thelearningmethodusedinthisexperimentissupervisedlearning[27].Insupervisedlearning[27], we have to label the data for training & evaluation of the model. For training andtesting the model, Caltech101[28] image dataset and Oxford 17 flower[29] image dataset wereused. In preprocessing, all images from Caltech101[28] and Oxford 17 flower[29] are resized to299x299X3 because to train aCNN using transferlearning[30],image input size to CNNmustbe same as the input sizegiven to original model.Then images from two standard imagedataset[28][29] were divided into training set, validation set and testing set. From caltech101[28]dataset,amongst70imagesperclass,60 imageswere used fortraining &validation andremaining 10 images were used for testing. For Oxford 17 flower[29] image dataset, amongt 80images per class, 64 images were used for training & validation and remaining 16 images wereusedfortesting.
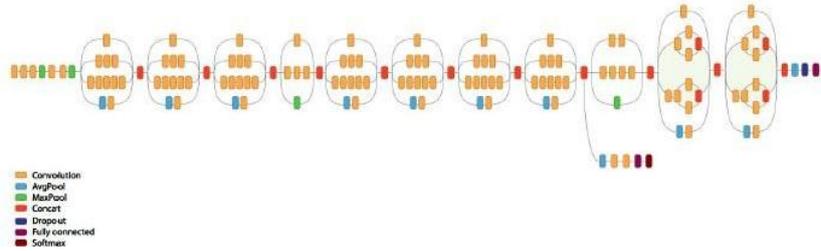
3.2. CnnModelDesign

Fig.1.ArchitectureOfInceptionV3Model[6].

TheCNNmodeldesignedhereisbasedoninceptionv3[6].Fig.1representthedetailarchitectureofIncept ionmodel.Itisa42layerDeep,pretrainedCNNmodeltrainedonImageNet[31],whichwas1strunnerupi nImageNetLargeScaleVisualRecognitionCompetition, with lower error rate i.e. top-1 error: 17.2% & top-5 error: 3.58 %. The inceptionmodel has two parts; feature extraction and classification. we make use of the feature extractionpart of inception model and retrained the classification layer with Oxford 17 flower imagedataset[29] and Caltech101 image dataset[28]. To retrain the classification layer we implementthetransferlearning[30]

3.3. TransferLearning

Inception v3[6] is a convolutional neural network model and by using GPU configured computerit takes weeks to train from scratch[6], tensorflow[32] a machine learning framework whichprovides plateform to train the classification layer with images from Caltech101[28] and Oxford17 flower[29] using transfer learning mechanism[30]. transfer learning mechanism[30], whichkeeps the weights and bias values of the feature extraction layer and removes parameters onclassification layer of inception v3[6]. First the input image of size of 299x299x3 arefeded tofeature extraction layerof CNN. After that feature extraction layer calculate the feature valuesfor each images, feature vector are 2048 float values for each images. then classification layer ofthe CNN is trained with these feature vector. The output labels in the classification layer is equaltothenumberof imageclassesonthedataset.
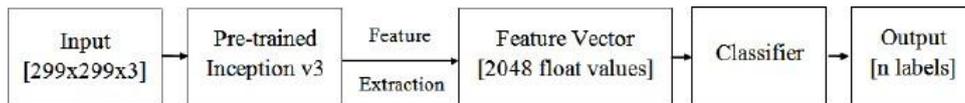

Fig.2.TransferLearningImplementationPipelines.

4. EVALUATION

The pretrained inception model[6] is used here for experimental purpose and the platform usedhere is tensorflow[32] and the hardware plateform used here is dell latitude e6410: 2.4 GHZprocessor intel i5, 4 GB RAM. For experiment Oxford 17 flower dataset[29] and Caltech101dataset[28]wereused.

4.1. Dataset

Oxford 17 flower dataset[29]: This dataset consist of 17 category flowers images with 80 imagesfor each class. The flowers chosen are some common flowers in Britain. The images has beencollectedbyMaria-ElenaNilsbackandAndrewZissermanofUniversityofOxford,UK.Caltech101imagedataset[28]:This datasethas101classesofimages.Ithas40to800imagesper class and Collected by Fei-Fei Li, Marco Andreetto, and Marc Aurelio Ranzato in September2003.

4.2. EvaluationProcedure

TotestthemodelonCaltechdataset[28],20classesofimagesformCaltech101imagedataset[28]   are used. Each category consist total of 70 images. The training and validation setconsist 60 images and testing set consists of 10 images per class. And to test the modelonOxford 17 flower dataset[29] 17 classes of flower images were used, each class consist of 80images: 64 images per class for training & validation and 16 images per class forevaluationwereused.

To train the model transfer learning mechanism[30] is implemented on pretrained inceptionmodel[6]. In transfer learning, weights and bias values of the feature extraction layer are keptsame as original CNN model and removes parameters on classification layer of a CNN. To trainthe classification layer,training setimagesfrom two dataset[29][28]were used.The Backpropagation algorithm[1] was used to train the classification layer of the CNN model and weightparameters of classification layer is adjusted by using cross entropy cost function by calculatingthe cross entropy error between classification layer output and input feature value from featureextractionlayer.

5. RESULTS

5.1. Training

The training accuracy, validation accuracy and cross entropy graph onflower dataset[29] andClatech101 dataset[28] are given in Fig. 3, Fig. 4, Fig. 5 and Fig. 6 respectively. The parametersusedforretrainingthemodelare,trainingsteps:4500,traininginterval:1andlearningrate: 0.045. The training and validation accuracy graph in Fig. 3 and Fig. 5 represents accuracy theCNN model get from training and validation images. And the cross entropy graph in Fig. 4 andFig. 6 represents the difference between the actual output and input feature value while trainingthe Classification layer of CNNmodel. In Fig. 3, 4, 5, 6 the Blue line represents accuracy &cross entropy error variation curve on training images and Orange line represents accuracy andcrossentropyerrorvariationcurveonvalidationimagesfromtwodataset[29][28].Fig.3represents the   training   &   validation   accuracy   variation   onflower   image   dataset[29]. Trainingaccuracywas91.2%atthebeginningofthetrainingprocessandstartstoincrease,aftercompleti on of (3/4) training steps it reached and remains to 100%. Validation accuracy was72.13% during initiation of training and validation process and final validation accuracy was92.7%.Fig.4representsthecrossentropyerrorfortrainingandvalidationrespectivelyon flower image[29]. The cross entropy error were 0.58 and 0.69, during initiation of training andvalidation process. As the training steps increases the cross entropy error starts to decrease. Atfinal step of training and validation the training and validation cross entropy error were 0.02 and0.09.
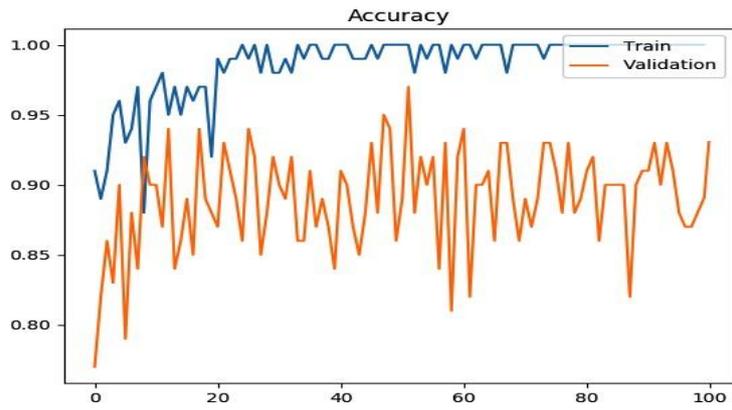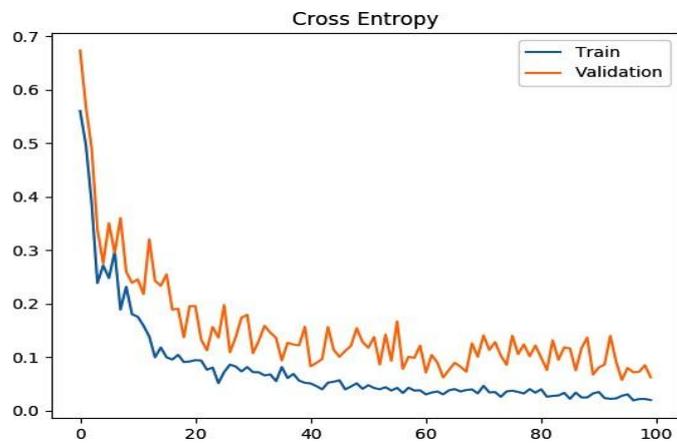
Fig.3.Accuracygraph on Oxford 17flowerdataset[29].

Fig.4.CrossentropyonOxford 17flowerdataset[29].
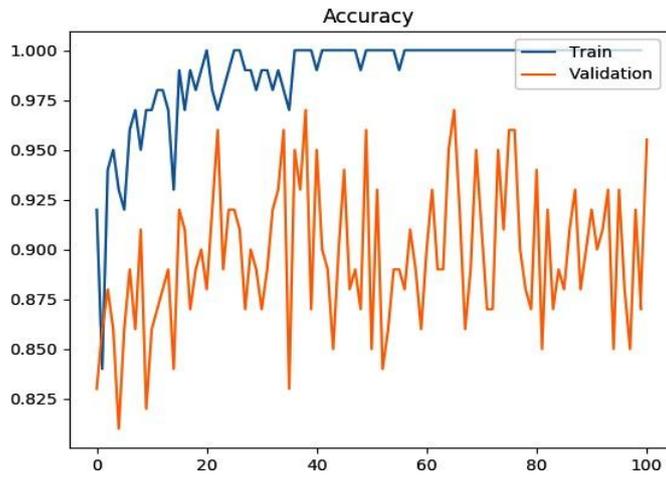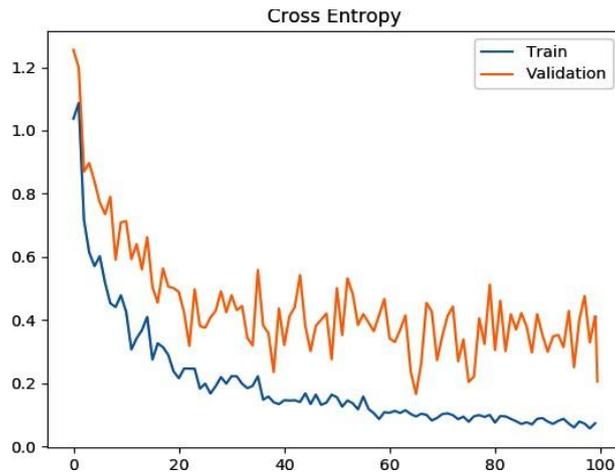

Fig.5.AccuracygraphonCaltech101dataset[28].

Fig.6.Cross entropygraphoftrainingonCaltech101dataset[28].

Fig. 5 represents the training & validation accuracy variation respectively on Caltech101 imagedataset[28]. Training accuracy was 83.6% at the beginning of the training process and starts toincrease, after completion of (3/5) training steps it reached and remains to 100%. Validationaccuracy was 82.7% during initiation of training and validation process and final validationaccuracywas94.9%.Fig.6representsthecrossentropyerrorfortrainingandvalidationrespect ivelyonCaltechimages[28].

The cross entropy were 1.0 and 1.3, during initiation of training and validation process. As thetraining steps increases the cross entropy error starts to decrease. At final step of training andvalidationprocessthetrainingandvalidationcrossentropyerrorwere0.28and0.1.

5.2 Testing andEvaluation

The classification result of the system on testing images fromflower dataset[29] and Caltechdataset[28]arerepresentedinevaluationchartgiveninFig.7andFig.8andtheresultscomparison s of our system with other experiments are represented in Table 1 and Table 2.respectively. The 'TP' in the chart is 'True Positive', which represents correctly classified imagesbythesystem.

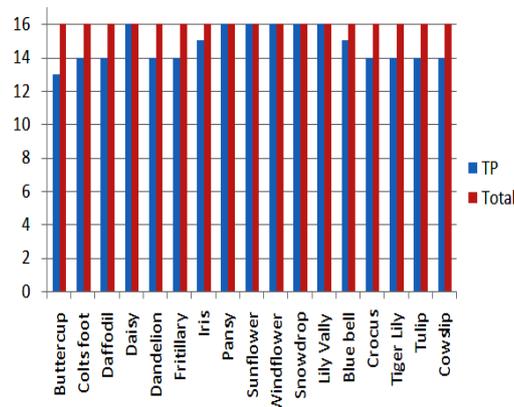Fig.7.Performanceevaluationparametersonflowerdataset[29].

For testing the modelon flowerimage[29], 16 images were used from each class, i.e. 272images from 17 classes. Amongst 272 testing images, correctly classified images by the systemwere 251 as shown in Fig. 7. For testing the model on caltech image dataset[28], 10 images wereusedfrom
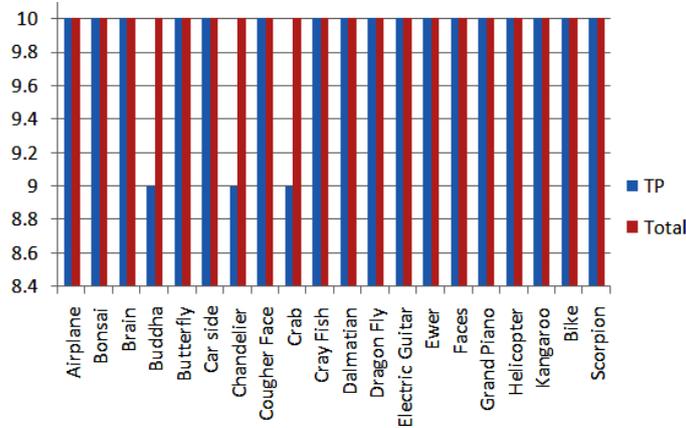

Fig.8.PerformanceevaluationmetricsonCaltechimagedataset [28].

each class, i.e. 200 images from 20 classes. Amongst 200 testing images, correctly classifiedimagesbythesystemwere196asshowninFig.8.

Table1. Performanceevaluationofoursystemon flowerdataset[29]withothermethod

S.N.	Method	MeanPrecision
1	[24]	81.3%
2	[33]	90.4%
3	**Oursystem**	**92.27%**

Table2.PerformanceevaluationofoursystemonCaltech101dataset[28]withothermethod

S.N.	Method	MeanPrecision
1	[8]	59.12%
2	[19]	75.2%
3	[22]	76.32%
4	[18]	80.4%
5	**Oursystem**	**98.0%**

6. CONCLUSIONS

Inthispaper,theclassificationlayerofpre-trainedinceptionv3modelwasretrainedsuccessfully by implementing transferlearningmechanism.Themodelyieldsprecisionof92.27 %on17classesofOxford17flowerimagedatasetand98.0%on20classesofCaltech101imagedataset.

REFERENCES

[1]Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel,Backpropagation applied to handwritten zip code recognition," Neural computation, vol. 1, no. 4, pp.541-551,1989.

[2   ]A.Krizhevsky,I.Sutskever,and   G.   E.   Hinton,Imagenet   classification   with   deep convolutionalneuralnetworks,"inAdvancesinneuralinformation   processingsystems,2012,pp.1097-1105.

[3]D.R. Reddy,Speech recognitionby machine: A review,"Proceedings of the IEEE, vol.64, no. 4,pp. 501-531, 1976.

[4A. Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, D. J. Wu, and A. Y. Ng, Textdetectionandcharacterrecognitioninsceneimageswithunsupervisedfeaturelearning,"inDocument Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE, 2011, pp.440-445.

[5]   T. Varga, D. Kilchhofer, and H. Bunke, Template-based synthetic handwriting generation for thetrainingofrecognitionsystems,"inProceedingsofthe12thConferenceoftheInternationalGraphonomic s Society,2005, pp. 206-211.

[6]   C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, Rethinking the inception architectureforcomputervision,"inProceedingsoftheIEEEconferenceoncomputervisionandpatternreco gnition,2016,pp. 2818-2826.

[7]   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 computervisionandpatternrecognition, 2015, pp.1-9.

[8]   Q. Liu, A. Puthenputhussery, and C. Liu, Novel general knn classifier and general nearest meanclassifier for visual classification," in Image Processing (ICIP), 2015 IEEE International Conferenceon.IEEE, 2015,pp. 1810-1814.

[9]   J.M.Keller,M.R.Gray,andJ.A.Givens,Afuzzyk-nearestneighboralgorithm,"IEEEtransactionsonsystems,man, andcybernetics, no. 4, pp.580-585, 1985.

[10] J. A. Tropp, Greed is good: Algorithmic results for sparse approximation," IEEE Transactions onInformationtheory,vol.50,no.10, pp.2231-2242, 2004.

[11] A.BeckandM.Teboulle,Afastiterativeshrinkage-thresholdingalgorithmforlinearinverseproblems,"SIAMjournalonimagingsciences,vol.2,no.1,pp. 183-202,2009.

[12] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, Return of the devil in the details: Delvingdeepintoconvolutionalnets,"arXivpreprintarXiv:1405.3531, 2014.

[13] I.Jolliffe,Principalcomponentanalysis,"inInternationalencyclopediaofstatisticalscience.Springer,2011, pp. 1094-1096.

[14] J. Sun, X. Cai, F. Sun, and J. Zhang, Scene image classification method based on alex-net model," inInformative and Cybernetics for Computational Social Systems (ICCSS), 2016 3rd InternationalConferenceon.IEEE, 2016, pp. 363-367.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutionalneuralnetworks,"inAdvancesinneuralinformation   processingsystems,2012,pp.1097-1105.

[16] C. Cortes and V. Vapnik, Support-vector networks," Machine learning, vol. 20, no. 3, pp. 273-297,1995.

[17] N. M. Nasrabadi, Pattern recognition and machine learning," Journal of electronic imaging, vol. 16,no.4, p.049901, 2007.

[18] K. He, X. Zhang, S. Ren, and J. Sun, Spatial pyramid pooling in deep convolutional networks forvisualrecognition,"inEuropean conferenceoncomputervision.Springer,2014,pp.346-361.

[19] T. Tuytelaars, M. Fritz, K. Saenko, and T. Darrell, The nbnn kernel," in Computer Vision (ICCV),2011IEEEInternationalConferenceon.IEEE, 2011,pp.1824-1831.

[20] K.P.Murphyetal.,Naivebayesclassifiers," UniversityofBritish Columbia,vol.18,2006.

[21] Z.S.Harris,Distributionalstructure," Word,vol.10,no.2-3,pp.146-162,1954.

[22] C. Theriault, N. Thome, and M. Cord, Extended coding and pooling in the hmax model," IEEETransactionsonImageProcessing, vol.22,no. 2,pp.764-777, 2013.

[23] M.RiesenhuberandT.Poggio,Hierarchicalmodelsofobjectrecognitionincortex,"Natureneuroscience,vol.2,no. 11, p. 1019,1999.

[24] M.-E. Nilsback and A. Zisserman, A visual vocabulary forflowerclassification," in ComputerVision and Pattern Recognition, 2006 IEEE Computer Society Conference on, vol. 2. IEEE, 2006,pp. 1447-1454.

[25] D. G. Lowe, Object recognition from local scale-invariant features," in Computer vision, 1999. TheproceedingsoftheseventhIEEEinternationalconferenceon,vol.2.Ieee,1999,pp.1150-1157.

[26] A. R. Smith, Color gamut transform pairs," ACM Siggraph Computer Graphics, vol. 12, no. 3, pp.12-19, 1978.

[27] S.B.Kotsiantis,I.Zaharakis,andP.Pintelas,Supervisedmachinelearning:Areviewofclassification techniques," Emerging artificial intelligence applications in computer engineering, vol.160,pp.3-24, 2007.

[28] L. Fei-Fei, R. Fergus, and P. Perona, Learning generative visual models from few training examples:An incremental bayesian approach tested on 101 object categories," Computer vision and Imageunderstanding,vol.106,no.1,pp.59-70,2007.

[29] M.-E. Nilsback and A. Zisserman, A visual vocabulary forflowerclassification," in ComputerVision and Pattern Recognition, 2006 IEEE Computer Society Conference on, vol. 2. IEEE, 2006,pp. 1447-1454.

[30] S. J. Pan, Q. Yang et al., A survey on transfer learning," IEEE Transactions on knowledge and dataengineering,vol.22,no. 10, pp. 1345-1359,2010.

[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, ImageNet: A Large-Scale HierarchicalImageDatabase,"inCVPR09, 2009.

[32] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M.Isard et al., Tensor flow: a system for large-scale machine learning." in OSDI, vol. 16, 2016, pp.265-283.

[33] Y. Chai, V. Lempitsky, and A. Zisserman, Bicos: A bi-level co-segmentation method for imageclassification,"2011.