FASHION APPAREL DETECTION BY MEANS OF DEEP LEARNING TECHNIQUES

[1]J Rajalekshmi , [2]Shivraj Nag , [3]Riyazul Anjum B.S
[4]Raveena S Kharvi, [5]Shivam Kumar

1, 2 ,3 ,4 ,5 Department of Computer Science and Engineering, RajaRajeswari College of Engineering, Bangalore, Karnataka-560074
[1]rajee_jegadheesan@yahoo.co.in,[2]shivrajnag@gmail.com,[3]riyazulanjum1997@gmail.com,[4]raveenaskharvi98@gmail.com,
[5]smileyshivam3042@gmail.com

ABSTRACT: Image based fashion apparel spotting is a nascent problem in computer vision that has gained a considerable research traction in the past couple of years. In this research analysis,anewcomputervisiontaskisaddressed,whichiscalledasfashionitemdetection, where the aim is to detect various fashion items a person in the images wearing or carrying.Thetypesoffashiontermswhichisconsideredinthisworkincludejeans,dress, and shirt along with their colour.The detection of fashion it em scan be an important first stepofvariouse-commerceapplicationsforfashionindustry.Theclassificationofvarious fashion apparels are used for various purposes like retrieval of similar apparels from the inventory and automate smart recommendation based on users liking to enhance the user experienceourmethodsbasedonstate-of-the-artobjectdetectionmethodpipeline,which combines object proposal methods with a Deep Convolutional Neural Network. Throughoutthework,theeffectivenessoftheproposedmethodisdemonstrated.
Keywords:Fashion,Apparel,ComputerVision,Detect,E-Commerce,DeepConvolution NeuralNetwork.

1. Introduction

The fashion industry is continuously evolving and it is required to keep up with the latest trends in this work, the method is proposed in order to detect fashion apparels a person is wearing or holding in the image. The types of fashion apparels include shirt, dress, and jeans along with their colour. Fashion apparel detection has gained significant amount of researchattentioninthepastcoupleofyear.Amainreasonisduetovarietyofapplications that a fashion apparel detection can enable. For example, predicting the details of fashion itemscanhelpindiscoveringsimilaroridenticalfashionitemsinane-commercedatabase or inventory. Similarly, classification of fashion apparel, based on the user's liking and dislikingcanbeusedtoproviderecommendationtotheuser.Fashionappareldetectionin real-time can be useful in video surveillance on retail stores where information aboutand individual's clothing choices can be guessed and similar clothes can be recommended to theuser.Unlikemostpreviousworksonthedetectionoffashionapparelthatapproachthetask asaspecializationofsemanticsegmentationtothefashiondomain,theproblemwasanobjectdetection task,wherethedetectionresultsaregivenanimagealongwiththetypeof cloth and their colour with percentage.
Detection-based spotters are more favourable as it is generally faster and have lower memoryfootprintascomparedtosemanticsegmentation,largescalepixel-perfecttraining data is extremely difficult to obtain, whereas training data as much easier to obtain instance-level identification is preferred because semantic segmentation does not distinguishdifferentinstancesbelongingtothesameclass.Althoughanycurrentmethods ofobjectdetectioncanbeimplemented,theroleoffashionappareldetectionposesitsown challenges such as the deformation of clothes is high, some classes of fashion items are extremely similar in appearance i.e. skirt and bottom of short dress and the concept of classes of fashion items may be vague i.e. pants and thighs [6]. In this work, some of these challenges are addressed by

ISSN- 2394-5125 VOL 7, ISSUE 04, 2020

incorporating state-of-the-art object detectors with various domainspecificmetricsliketheshapeandsizeoftheobject[7].Thestate-of-theartobject detectorwhichhasbeenusedinthisworkisRegionConvolutionalNeuralNetworkwhich combines object propositions with a Convolution Neural Network. The R-CNN starts by generating a set of object propositions. The images patches are then extracted from the produced boxes and resized to a fixed size. For the task of image classification, the Convolution Neural Network, pre trained on a board image database, is used to extract featuresfromeachimagepatch.Then,SVMclassifiersareappliedtoeachimagepatchto decideifthepatchbelongstoaparticularclass.TheR-CNNisidealforourtaskbecauseit can detect objects with various aspect ratios and sizes without running window search, minimizing the difficulty of the computation and the false positives [8]. Apparel classification also facilitates the automatic annotation of images with tags or descriptions related to clothing, allowing better retrieval of information and settings such as a user's photo from socialnetworks.

2. Literature Survey

Thefirstsegmentation-basedfashionapparelspottingalgorithmforgeneralfashionitems was proposed by K.Yamaguchi et al where they introduce the Fashionista Dataset and utilize a combination of local features and pose estimation to perform semantic segmentation of a fashionimage.

B.S.Hasanetalalsoproposedasegmentation-basedapproachaimedatassigningaunique labelfrom"Shirt","Tie","Jacket"and"FaceandSkin"classestoeachpixelintheimage. Their method is focused on people wearingsuits.Thereexistseveralclothingsegmentationmethodswhosemaingoalistosegmentoutthe clothing area in the image and types of clothing are not dealt with. in a clothing segmentation method based on graph-cut was proposed by A. C. Gallagher et al for the purpose of identityrecognition.

Z. Hu et al proposed a graph-cut based method to segment out upper body clothing.

N.Wangetal[15]presentedamethodforclothingsegmentationofmultiplepeople.They proposetomodelandutilizetheblockingrelationshipamongpeople J. Shen et al [16], purposed a structured learning technique for simultaneous human pose estimation and garment attribute classification is proposed. The work is focused towards detecting attributes associated with the upper body clothing, such as collar types, colour, typesofsleeves,etc.Sincelocalizationofupperbodyisessentiallysolvedbyupperbody detectors and detecting upper body is relatively easy, the focus of the above methods are mainlyonthesubsequentstageofclassification.Ontheotherhand,thefocusisonavariety offashionitemswhichareverydifficulttodetectevenwiththeperfectposeestimationdue to varioussize.

M. Yang et al [17] proposed a real-time clothing recognition method in surveillance settings. They first obtain foreground segmentation and classify lower bodies and upper bodies separately into a fashion item class.

L.Bossardetal[18]proposedaposelet-basedapproachforhumanattributeclassification. intheirwork,asetofposeletdetectorsaretrainedandforeachposeletdetection,attribute classification is done using SVM. The final results are then obtained by considering the dependencies between differentattributes.

G. W. Cottrell et al [19] proposed recognition of social styles of people in an image is addressedbyConvolutionalNeuralNetworkappliedtoeachpersonintheimageaswellas the entireimage.

3. Proposed Methodology

Theaimoftheproposedmethodistodetectfashionitemsinagivenimage,wornorcarried byasingleperson.Wedefinethefashiondetectiontaskasanobjectdetectiontaskandnot thesemanticsegmentationtask.Theproposedmethodcanbeconsideredasanextensionof theR-CNNframework,whereweutilizevariouspriorsonlocation,sizeandaspectratios of fashion apparels, which we refer to as geometricpriors.

ISSN- 2394-5125 VOL 7, ISSUE 04, 2020

4. Design andImplementation

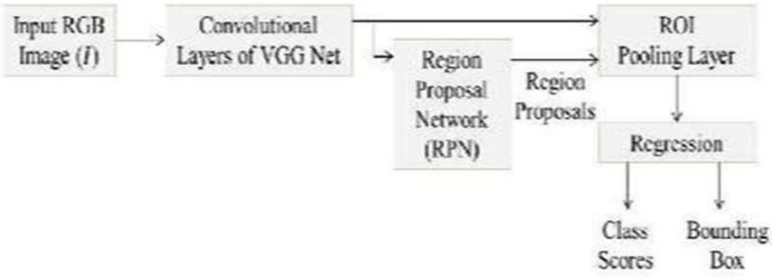

Fig1. Architectural Design

TheDesigngoesthroughtwostages:TrainingStageandDetectionStage.IntrainingStage there are training datasets which are passed through Faster RCNN and then it is used in detection stage, whereas in detection stage a yarn image and trained datasetswithFaster-RCNNarepassedinordertogetresultsafterchecksumverificationthatiswherewe decide whether the trained datasets are failed orsucceeded.

Object Proposal: Object detection based on a sliding window technique was a standard approach for exhaustive execution of object detectors at all locations and potential image scales. The most recent works detect a single object through a series of part specific detectors to accommodatematerialdeformation,andallowthepartstovaryinconfiguration Duetothe accommodation of a certain amount of deformation, the possible aspect ratiosconsidered arestilllimitedandasthenumberofcomponentdetectorsincreases,thecomputationtime increaseslinearly.Thevariationinintra-classformisbiginourmission.ACNNtrainedon ImageNet database is explicitly used as a function extractor for various related tasks and achieves impressiveperformance.

SVM Training: The training of linear SVM is done in order to classify an image patch as positive or negativeforeachclassofobjects.inordertoincreasethenumberoftrainingpatcheswerun the proposed object algorithm on the training images. The training of a series of linear SVMs have been done, each of which is trained as positive samples using instances in a givenclassasnegativesamples,andallinstancesintheremainingclass.Theparametersfor SVMs are determined from validation collection. The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes. At first approximationwhatSVMsdoistofindaseparatingline(orhyperplane)betweendataof two classes. SVM is an algorithm that takes the data as an input and outputs a line that separates those classes ifpossible.

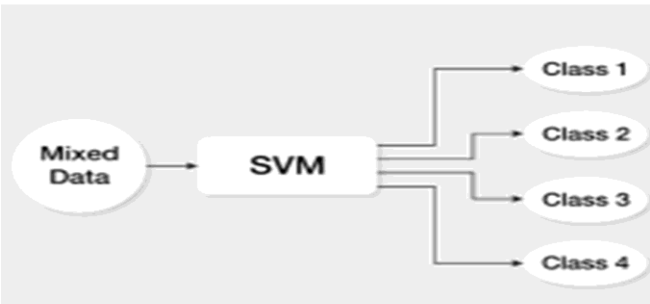

Fig2. SVM Classifier

Input / Output Design: This work consists of two Sub-Design Phases:

- High LevelDesign
- Low LevelDesign

ISSN- 2394-5125 VOL 7, ISSUE 04, 2020

IntermsofPrecisionandRetrieval,firstassessisdonetotheefficiencyofthemethodsof the item proposal. The use of the average precision i.e. AP computed from the Precision Retrieval curves to determine the efficiency of the detection methods and those are mathematically represented by the equations. This finding demonstrates the utility of geometricpriorsinthefunctionoffashionoffashionobjectdetection.Thecontentofthat identification of fashion clothing has its own special challenges. Secondly even with the different groups of fashion pieces, certain fashion items are very close to each other visually.“Tights”and“Pants”,forexamplecanlookverysimilarasbothproductscanhave arangeofcolours.Theonlydistinctivecluemaybehowclosetoit,whichisverydifficult tocapture.

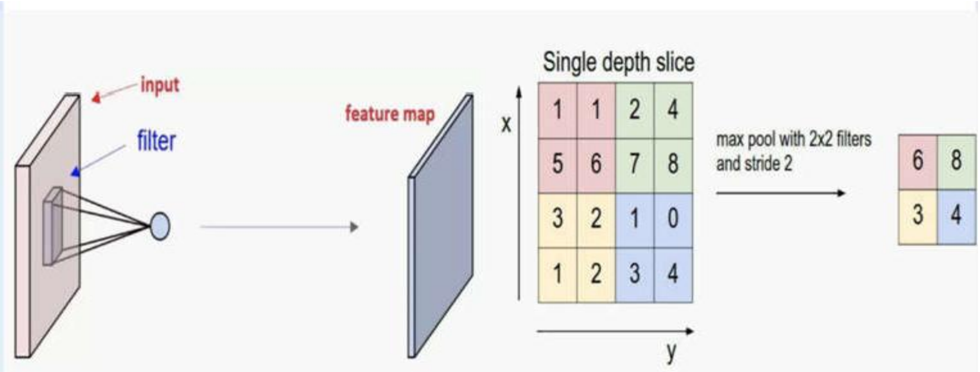

Fig3. Input/output Design

5. Results

Unit Testing: Unit testing is a software development process in which the smallest testable parts of an application, called units, are individually and independently scrutinized for proper operation.Unittestingisoftenautomatedbutitcanalsobedonemanually.Thegoalofunit testing is to isolate each part of theprogram.

Sl # Test Case	UTC-1
Name of Test	Image Capture
Items being tested	Input Image
Sample Input	Image
Expected output	Should Capture input image
Actual output Image	Captured Successful
Remarks	Pass

Test case1. Image Unit Testing

Sl # Test Case	UTC-2
Name of Test	Shirt Labelling
Items being tested	Labelling
Sample Input	Image
Expected output	Shirt portion should be labelled
Actual output	Shirt Labelled
Remarks	Test Passed

Testcase2. ShirtUnitTesting

Sl # Test Case	UTC-3
Name of Test	Pant Labelling
Items being tested	Labelling
Sample Input	Image
Expected output	Should label Pant
Actual output	Same as Expected
Remarks	Test Passed

Testcase3. PantUnitTesting

Fig 4. Keras Multi Output Dataset

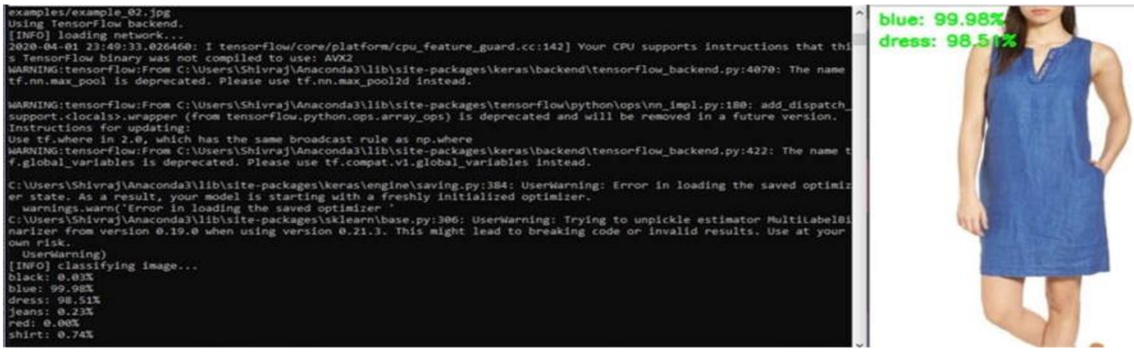

Fig 5. Blue Top Detected

6. Conclusion and Future Work

Inthisworktheuseofconvolutionalneuralnetworkstorecognizeimagesanduse them for apparel classification is been analysed. The reformulation of fashion apparel parsing, historically treated as a semantic segmentation task, as an object detectiontaskandpurposeaprobabilisticmodelwhichincorporatesstate-of-theart object detection algorithm. Online shopping for clothing can be improved by replacing the traditionally filters with the image filters through experimental evaluations,weobservetheeffectivenessofthepurposedpriorsforfashionapparel detection.

As a future work, the same approach can be extended to new domains where insufficient data and other constraints might otherwise prevent their use. This approachachievesasignificantimprovementofaverageaccuracyandstate-of-the art of image based apparel classification. A similar approach can be employed to recommend products based on user'sliking.

References

[1] L. Bossard, M. Dantone, C. Leistner, C. Wengert, T. Quack, and L.V. Gool. Apparel classification with style. ACCV,2012.

[2] H. Chen, A. Gallagher, and B. Girod. Describing clothing by semantic attributes. ECCV,2012.

ISSN- 2394-5125 VOL 7, ISSUE 04, 2020

[3] J.Donahue,Y.Jia,O.Vinyals,J.Hoffman,N.Zhang,E.Tzeng,andT.Darrell. DeCAF:ADeepConvolutionalActivationFeatureforGenericVisualRecognition. CML,2014.

[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchiesfor accurate object detectionand semantic segmentation. CVPR, 2014.

[5] B. S. Hasan and D. C. Hogg. Segmentation using Deformable Spatial Priors with Application to Clothing. BMVC, pages 83.1– 83.11,2010.

[6] A.Krizhevsky,Sutskever,andG.E.Hinton.magenetclassificationwithdeep convolutional neural networks.2012

[7] J. R. R. Uijlings, K. Van De Sande, T. Gevers, and A. Smeulders. Selective Search for Object Recognition. iJCV,2013.

[8] P.ViolaandM.Jones.Rapidobjectdetectionusingaboostedcascadeofsimple features. CVPR,2001.

[9] K.Yamaguchi,M.H.Kiapour,andT.L.Berg.PaperDollParsing:Retrieving Similar Styles to Parse Clothing items, iCCV,2013

[10] K.Fukushima.Neocognitron:ASelf-organizingNeuralNetworkModelfora Mechanism of Pattern Recognition Unaffected by Shift in Position. Biological Cybernetics, 202,1980.

[11] K.Yamaguchi,M.H.Kiapour,L.E.Ortiz,andT.L.Berg.Parsingclothingin fashion photographs. CVPR,2012.

[12] B. S. Hasan and D. C. Hogg. Segmentation using Deformable Spatial Priors with Application to Clothing. BMVC,2010

[13] A.C.GallagherandT.Chen.Clothingcosegmentationforrecognizingpeople. CVPR, June2008.

[14] Z. Hu, H. Yan, and X. Lin. Clothing segmentation using foreground and backgroundestimation.PatternRecognition,41(5):1581–1592,May2008

[15] N. Wang and H. Ai. Who Blocks Who: Simultaneous clothingsegmentation for grouping images. CCV, pages 1535–1542, Nov.2011.

[16] J.Shen,G.Liu,J.Chen,Y.Fang,J.Xie,Y.Yu,andS.Yan.UnifiedStructured Learning for Simultaneous Human Pose Estimation and Garment Attribute Classification.2014

[17] M. Yang and K. Yu. Real-time clothing recognition in surveillance videos. iCIP, 2011.

[18] L. Bourdev, S. Maji, and J. Malik. Describing people: A pose let based approach to attribute classification. iCCV,2011

[19] Y.WangandG.W.Cottrell.Bikersareliketobaccoshops,formaldressersare likesuites:RecognizingUrbanTribeswithCaffe.InWACV,2015

[20] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson. CNN Features off-the-shelf: an Astounding Baseline for Recognition. CVPR Workshop, Mar. 2014