ZKG INTERNATIONAL
ISSN: 2366-1313

Deep learning-based Attention MechanismModel for Fast Lane Detection

[1]Manda Mohan Ganesh, [2]Dr. B. Gohin

[1]MCA Student, Dept. Of MCA, Swarnandhra College of Engineering and Technology, Seetharampuram, Narsapur, Andhra Pradesh 534280,

mohanganesh9989@gmail.com

[2,]Associate Professor, Dept. Of MCA,Swarnandhra College of Engineering and Technology, Seetharampuram, Narsapur, Andhra Pradesh 534280,

Abstract:*As an important subtask in the field of unmanneddriving, lane detection has transitioned in recent years from traditionalimageprocessingtoadeeplearning-basedneuralnetworkapproach. However, since the early deep learning methods arebased on semantic segmentation at the pixel level, their largenetwork structures cannot meet the real-time requirements. Tosolve the real-time problem, a new network structure based onpredefined rows, represented by UFAST, is proposed. In thearchitecture of this network model, the network parameters aresignificantlyreduced,allowingthesystem'sreal-timeperformancetobesatisfied.Toimprovetheperformanceofrecognizinglanes in the framework of this model, we introduce attentionmechanism into the model by considering the habits of real humandriving. Finally, we not only improve the performance of themodel framework in non-ideal conditions such as poor lightingand vehicle occlusion by nearly 1.9%, but also increase the numberof model parameters by less than 0.2% of the UFAST. We addressourcecodesathttps://github.com/APPZ99/Fast-Lane-Detection-Based-on-Attention-Mechanism on*

Keywords-component; UnmannedDriving,LaneDetection,AttentionMechanisms

I INTRODUCTION

In more recent years, with the constant development ofcomputer vision and artificial intelligence, the field of un-manned driving has been rapidly developed. As an importantcomponent of unmanned driving, the development of lanedetectiondeterminesthedirectionoftheunmanneddrivingfield.With the constant

ISSN: 2366-1313

improvement and development of deeplearninginrecentyears,lanedetection basedondeeplearninghas gradually replaced the traditional lane detection methodbased on probability distribution and graphical geometry. In thefield of lane detection, various deep learning-based detectionmethods have emerged in recent years, such as Lane Net [1],SCNN [2], and VPGNet [3] etc. However, they are all based onpixel-levelsemanticsegmentationtoretrieveroad informationand categorize it, and their massive network structures do notenhancethesysteminreal-timerate.Inthiscontext,ZequnQinet al. [4] construct select locations of lanes at predefined rows ofthe image using global features to retrieve the information of thelane,whichsignificantlyimprovesthelanedetectionratewhileobtainingthe state-of-the-artperformance. Oneof the important Approaches of UFAST [4] is to artificially schedule the inputimage information of the system to the lower part of the viewinformation, the location where the lanes exist in most cases.However, real road information is complex, and if the sameapproachisusedforalltheviewinformation,itwillinevitablyyield missing or redundant information. In comparison with theclassificationaccuracyof95.87%onthe TuSimpledataset[5]under the ResNet- 18 [6] architecture, the accuracy of 68.4% ontheCULanedataset[2]initsexperimentsisslightlyinsufficient.Thereasonforthisisthat, inadditiontothemorecomplexandvariableroadconditionsintheCULanedatasetsitself,thedesignof its system framework also tends to lead to the lack of roadlaneinformationundernon-idealconditions.

Whetheritistraditional visualinformation processingmethods or deep learning based detection and classificationmethods,mostoftheeffortsarefocusedongatheringmoreandmore complex information to be utilized as a constraint fordetection.Thishasalsoledtoincreasinglymassivethewholedetection network structures and an increasing amount ofprocessed data. Although the newly proposed networks strive tosimplify the network size, it is difficult to get rid of theframework complexity of pixel-level semantic segmentation.How to abandon redundant information while retaining as muchbeneficial information as possible is the key to solve the slowspeedoflanedetectionsystem.Fromtheperspectiveofhumandriving habits, we find that drivers will observe the current laneand the lanes on both sides (or one side) from close to far duringregular driving, and pay less attention to the field of viewboundaryinformation.Thisprocessis alsotheattentionallocationmechanism.Themechanismtoimitatehumanobservation of lanes in complex situations such as low light andvehicleobstructionisexactlywhatmostlanedetectionsystemslackatpresent.

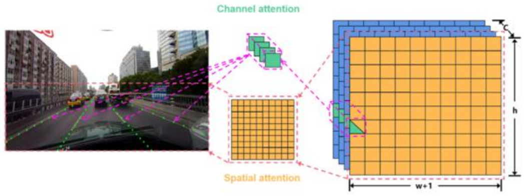

Fig.1. Rationale for the role of attention mechanisms in the UFAST architecture. Channel attention concentrates on the location and shape of each lane toenhancethedetectionofasinglelane;

So that the whole system network learns to focus on the lanedetectionpartwhilemaintainingtheprocessingspeedasfastaspossibleandachieving animprovementinthesystemprocessingperformance.Itmapsthefeaturesoftheentirenetworkandthechannels in each grid. At the same time, the network capturesthe overall features of the input image and predicts the lanes fornon-idealconditionssuchascongestedlanes,insufficientlight,etc., which are partially projected into the spatial features of theoverall network. Combining these two features, we adoptCBAM,anattentionmechanismmodulethatishighlycompatible with the UFAST framework. This module combinesspatial attention in addition to channel attention. Combined withthe CAMB module, the overall net- work architecture can betterreinforce the information of lane detection and achieve theenhancement of attention to lanes in real driving situations. Duetoitscompactsize,itonlyincreasesthenumberofparametersby 1% of the original network frame. Ultimately, we achieveperformanceimprovementsforthe UFASTframeworkincomplex situations. The contribution of this our work to theoverallfieldissummarizedintwopoints asfollows.

- Implementing the attention mechanism, which allowsthe systemtoextractlaneinformationinamoretargetedmanner?

With the attention mechanism the improvement of themodelaccuracyinroadconditionsofcomplexsituations isnearly 1.9 %on average over the UFAST[4]originalmodel.

Thispaperwillbeorganizedas follows:Sec.IIconsiders relatedwork.Sec.IIIdescribesthemethod ofmodelconstruction.Sec.IVexplainsthere levantexperimental setup as well as the results and analysis. Finally,Sec.Vconcludesourwork.

II RELATEDWORK

ISSN: 2366-1313

MethodsofLandDetection

Forthefieldoflanedetection,therearetraditionaldetection methods based on Markov random fields by Aly etal.[14]in theearlydays.Withtherapiddevelopmentofdeep learning in recent years and the increasing complexity oflaneroadconditions,deeplearning-basedlanedetectionmethodshave beenproposedandthe performance has beenimproving at thesame time.Davy Neven et al.[1]propose anend-to-endmethodtoextractthelaneinstancesthroughtheLaneNet network, and reconvert the lanes to the original mapinaleast-squaresmannerafterpassingthetransformationmatrixH outputfromHNet.Mostofthe succeedingdeeplearning-basedmethodsareinfluencedbytheLaneNetnetwork.Xingang Pan et al.[2]propose the SCNN in 2017simplifiesthesegmentationtaskinLaneNet,makingtheoverallnetwork moreconcise,andtheproposedCULaneDatasetsalsoprovideaplentifuldatatestsetforlanedetection.SeokjuLeeetal.[3]proposeTheVPGNetinthesameperiod,whichpredictstheoveralllanesbymeansofamulti-branchtask,focusingontheproblemoflanepredictionintheinvisiblepart.Althoughalltheaboveofthesemethodsachievedindustry-leadingaccuracyatthetime,itsmassivenetwork structure made the overall systemprocessingspeedstilldifficulttocopewiththereal-timerequirementsofunmanned driving. To solve this problem, Zequn Qin et al.[4]proposedtheUFASTnetworkstructurein2020greatlysimplifiedthebackboneofthe entire detection networkandcreativelybuiltapredefinedlinetoselectlaneinformationmodelbasedonthe useofglobalfeature pictures,whichsubstantiallyimprovedtheoveralldetection systemspeedwhile keepingthestate- of- art accuracy at that time.Itsreal-timeperformancehasfulfilledthelanedetectionrequirementsfordriverlesscarsinlowerspeedsituations.

MethodsofAttentionMechanisms

Humansconsciouslyfocusonthe parts that interestthemwhentheyobservethingsanddevotemoreattentionto them, which is the attention mechanism [12]. VolodymyrMnih et al.[7]propose the first RAM implemented spatialattention for convolution neural networks. Then Jie Hu et al.[8] proposeSENetimplementedanew modelofchannelattention network with adaptive,and the proposed networkalso greatly inspiredthedesignwork of attention moduleforsubsequentconvolutionneuralnetworks.DeeplyinfluencedbySENet,Sangh yunWooetal.[13]proposetheCBAMattentionmodulein2018implementsacombination of channel attention andspatialattentionforconvolution networks, while its lightweight module designfacilitatesitsinsertionbetweenanyconvolutionlayerswith

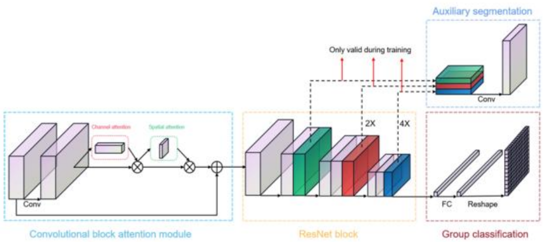

Fig.2.OverallarchitecturewithCBAMattentionmodule.Beforeinputtingintotheresidualnetwork,theCBAMmoduleextractsthechannelandspatialfeaturesof the input image, and then feeds the convolution blocks containing the features into the subsequent residual network. The second, third and fourthconvolutionblocksintheresidualnetworkareextractedandcombinedtoformtheauxiliarysegmentationchannel,whichisshowninthe upperpart.Themainchannelconnectstheconvolvednetworktothefullyconnectedlayertoobtainthefinallaneclassification.

Little impact on its network speed.In recent year, with theproposed transformer, the attention mechanism adapted to itsarchitecture structure is also gradually proposed, such as ViT[15] achieved good results in the field of computer visionthrough the transformer structure. However, this paper doesnot involve the transformer architecture, so we will not dotoomuchanalysishere.

III METHOD

In this section, we focus on the implementation of theattention mechanism in the model and the rationale of theattentionmechanismfornetworkperformanceimprovement.

ExtractionofAttentionFeatures

Themainreasonforthesignificantspeedimprovementof the UFAST [4] system is the structural design that distinguishes it from the semantic segmentation model.The feedimage only intercepts the defined grid size, and the numberofchannelsdependsonthenumberoflanesforeachdataset. Therefore, the overall network structure is much smallerthan thatofthesemanticsegmentationmethod.The channelofthe networkcorrespondstoeachlane;naturally, weimprovetheattentiontothechannelofthe network,whichisequivalent to improving the attention to each lane, thusmaking the system more sensitive to the detection of lanes.For non- ideal situations such as nearby vehicle occlusion andinsufficient light,UFASTpredictstheoccludedlanesby acquiring the image global features. To exploit the imageglobal features better and improve the feature extraction inthenetworkforthewholegridnetworkthat isadvantageousfor lane prediction, we improve the attention to the spatialfeatures of the whole convolution network. By raising theattention to features such aspavement edges and bridge

ISSN: 2366-1313

piles,we enhance the network's predictionandcorrectionoflanesas a whole.The use of the overall mechanism of attentiondetectionisshowninFig.1.

LaneRecognition

We followthe basic architecture of UFAST [4]and usethe ResNet- 18 [6] neural network to make the extraction ofoverall lanes using auxiliary channels in the training stage.The lanes are predicted by extracting the data from the second,thirdandfourthconvolutionlayerso fResNet-18,respectively, andcombining them toconvolvethem.Themainbranchclassifie sthelanesbyunfoldingtheconvolutionlaye rsto connect the classifier, resize to theoriginal predefined grid size, and calibrate the lane locationsin the image. Thus, with lane attention information is addedto the subsequent lane detection recognition system.

VI EXPERIMENTS

Experimentalsetting

Hardware: Thehardwaredeviceweuseisa differentNvidiaRTX1070GPUfromthe UFAST's NvidiaRTX-1080Ti G P U , so

Datasets: To test whether our model is improved com-pared to UFAST [4]original model,we choose the samebenchmark dataset as UFAST for lane detection: TuSimpledataset [5]and CULane dataset [2],and the details of bothdatasetsareshownintheTABLEI.

Evaluation metrics: The official evaluation criteria forthetwodatasetsaredifferent.

FortheTuSimpledataset,themainevaluationmetricsarecalculatedasshowninthefollowingequation.

$$Recall = \frac{TP}{TP+FN} \quad \text{Accuracy} = \frac{\sum_{clip} C_{clip}}{\sum_{clip} S_{clip}} \tag{2}$$

Forthe CULanedataset, theevaluationindex $F1-$ iscalculatedasfollows.

$$F1 - \text{measure} = \frac{2\times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

Control group setting:To achieve a better representationoftheroleoftheattentionmechanisminimprovingtheoverall networkperformance,wedeployedSENet[8] intotheUFAST [4]networkframeworkinordertoobtaintheamount of improvementofnetworkperformancebytheroleofchannelattentiononly.

Results

Accordingto the results shown in the TABLE III andTABLEIV,itiseasy to find that the modified UFASThasthefollowingtwocharacteristics.

- For the TuSimple dataset, the

model is similar to theoriginal model in terms of accuracy and has no significantimprovement.

• FortheCULanedataset,themodelac hievestheaccuracy improvement of the whole test domain, and the finalaverage accuracy is improved by nearly 1. 9 % compared withtheoriginalmodel.

• Withtheincorporation of the attention mechanism,our modelparametersincreasebyonlylessthan0. 2 %oftheoriginalmodel;theincreaseinFLOPsforconvolutionisabout0. 4‰oftheoriginalmodel.

TABLE III COMPARISON WITH OTHER METHODS ON TUSIMPLETEST SET.FLOPS REFERS TO FLOATING POINT OPERATIONS INTHECONVOLUTION. PARAM. REFERSTOTHETOTALNUMBEROFPARAMETERSOFTHE NETWORK

	UFAST-18	UFAST-18+SE	UFAST18+CBAM
Accuracy	**95.87**	95.46	95.85
FLOPs	**16.646G**	16.646G	16.653G
Parma.	**44.532M**	44.619M	44.620M

Analysis

For the above existed situation, we made the followinganalysis.

• For the TuSimple dataset, the accuracy of the originalmodelhasreached95.87%,whichis 2.32%differentfromtheaccuracy of the SCNN UNet ConvLSTM2 [16] model with thehighestaccuracy,andonly1.09%differentfromtheaccuracyofthePERESA[17]modelwiththesecondhighestaccuracy,which basicallyreachestheoptimalmeasurement accuracyofthis dataset, and it is difficult to require a great improvement onthe original model as it is difficult to improve on the originalmodel.Addingtheattentionmechanismmakesthenet-workasawholeover-fitandunderperformonthetestset.

• For the CULane dataset, we found that in various scenariosclosertodailylife,ourmodelhasimprovedsignificantly compared with the original model, except for theNo- Line dataset. The other datasets generally improved by1-3%. We argue that for No-Line, the attentionmechanismcannotcapturetheattentionobject better,resultinginrelatively poor model performance inthis dataset.For theother test datasets,we can find that after adding the attentionmechanism, the model can better capture the information oflanes for various environments,thus improving the overallaccuracyofthesystem.

• Combining the performance of the two datasets,theTuSimpledatasetismainlyforlanedetection in a moreidealenvironment(light,lanelineintegrity,roadvehiclecongestion) ,when the original model is also able to capturelane prediction well. However, for the CULane dataset,thecomplexandvariableroadconditionsmake theoriginalmodel'sabilitytocapturelaneinformation significantlyreduced,and the introduction of the attention mechanismimproves the problem, thus achieving an improvement in theoverall system accuracy, which is also more in line with thehumandrivinghabitsin realisticscenarios-whentheenvironment is more This is also more in line with humandriving habits in realistic scenarios - when the environmentis more complex and changing, the attention to the overalldrivingwillbemorefocused,thusensuringsafedriving.

V CONCLUSION

Inthispaper,wepropose to introduce an attentionmechanisminthelanedetectionmodeltobettersimulatethe real situation where people pay attention to lane conditions.To ensure real-time lane detection, we choose to improve theUFAST [4] model and introduce the CBAM [13] attentionmechanismmoduletoobtainthechannelandspatialfeatures of the lane detection model,while keeping the highspeed of the original model without degradation due to itscompactness.Wedidcomparisonexperimentsonourimprovedmodelin TuSimpledataset[5]andCULanedataset [2] ,and foundthat the model did not significantly improvethe lane detection intheidealenvironmentofTuSimpledataset,and generally improved the performance in multi-complexcasesinCULane dataset,withafinalaverageaccuracyimprovementofnearly1 . 9 %.Wewillalsoimprovetheattentionmodeldesignforlanedetectiontogetamoresuitableattentionmodeldesignforlanedetection.

REFERENCES

1. D. Neven, B. De Brabandere, S. Georgoulis, M.Pressman's, and L. VanGool, "Towards end-to-end lane detection: an instance segmentationapproach," in 2018 IEEE intelligent vehicles symposium (IV), pp. 286–291,IEEE,2018.
2. X.Pan,J.Shi,P.Luo,X.Wang,andX.Tang," Spatialasdeep:Spatialcnnfortrafficsceneunderstanding,"inProceedingsoftheAAAI ConferenceonArtificial Intelligence,vol.32,2018.
3. S.Lee,J.Kim,J.ShinYoon,S.Shin,O.Bailo, N.Kim,T.-H.Lee,H.SeokHong, S.-H. Han and I. So Kweon, "Vpgnet:

Vanishing point guidednetworkforlaneandroadmarkingdetectionandrecognition,"inProceedingsoftheIEEEinternationalconferenceoncomputervision,pp.1947–1955,2017.

4. Z. Qin, H. Wang, and X. Li, "Ultra fast structure-aware deep lane detection,"in European Conference on Computer Vision, pp. 276–291, Springer,2020.

5. S. Yoo, H. S. Lee, H. Myeong, S. Yun, H. Park, J. Cho, and D. H. Kim,"End-to-endlanemarkerdetectionviarow-wiseclassification,"inProceedings of the IEEE/CVF Conference on Computer Vision andPatternRecognitionWorkshops,pp.1006–1007,2020.

6. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for imagerecognition,"inProceedingsoftheIEEEconferenceoncomputervisionandpatternrecognition,pp.770–778,2016.

7. V. Mnih, N. Heess, A. Graves, et al., "Recurrent models of visual attention,"Advancesinneuralinformationprocessingsystems,vol.27,2014.

8. J.Hu,L.Shen,andG.Sun,"Squeeze-and-excitationnetworks,"inProceedings of the IEEE conference on computer vision and patternrecognition,pp.7132–7141,2018.

9. A.Vaswani,N.Shazeer,N.Parmar,J.Uszkoreit,L.Jones,A.N.Gomez,Ł. Kaiser, and I. Polosukhin, "Attention is all you need," Advances inneuralinformationprocessingsystems,vol.30,2017.

10. S. Xu, Y. Cheng, K. Gu, Y. Yang, S. Chang, and P. Zhou, "Jointly attentivespatial-temporalpoolingnetworksforvideo-basedpersonre-identification," in Proceedings of the IEEE international conference oncomputervision,pp.4733–4742,2017.

11. X. Li, W. Wang, X. Hu, and J. Yang, "Selective kernel networks," inProceedings of the IEEE/CVF Conference on Computer Vision andPatternRecognition,pp.510–519,2019.

12. M.-H. Guo, T.-X. Xu, J.-J. Liu, Z.-N. Liu, P.-T. Jiang, T.-J. Mu, S.-H.Zhang, R. R. Martin, M.-M. Cheng and S.-M. Hu, "Attention mechanismsincomputervision:Asurvey," ComputationalVisualMedia,pp.1–38,2022.

13. S.Woo,J.Park,J.-Y.Lee,andI.S.Kweon,"Cbam:Convolutionblockattentionmodule,"inProceedingsoftheEuropeanconferenceoncomputervision(ECCV),pp.3– 19,2018.

14. M. Aly, "Real time detection of lane markers in urban streets," in 2008IEEEIntelligentVehiclesSymposium,

pp.7–12,IEEE,2008.

15. Prasadu Peddi (2015) "A review of the academic achievement of students utilising large-scale data analysis", ISSN: 2057-5688, Vol 7, Issue 1, pp: 28-35

16. Prasadu Peddi (2015) "A machine learning method intended to predict a student's academic achievement", ISSN: 2366-1313, Vol 1, issue 2, pp:23-37.