1 Experiments

In this Section,we introduce our detailed implementation of building detection,and show the results of building detection.

1.1 Dataset

The overhead images used in this work comes from the ISPRS 2D Semantic Labelling Challenge Vaihigen and Potsdam.Vaihigen data consist of near infra-red(IR),red(R),green(G) imagery with corresponding digital surface models(DSMs).Potsdam data consist of near infra-red,red,green,blue imagery with corresponding normalized digital surface models(nDSMs) and row digital surface models(DSMs).Both areas cover urban scenes,while Vaihigen is a relative small village with many detached buildings and small multi story buildings,Potsdam shows a typical historic city with large building blocks,narrow streets and dense settlement structure. Since images in the data set are too large to train in the network,we cut the training and validation images into256×256 pixels patches.In addition,since the images provided by Potsdam data are too large,it may affect the final prediction result, we reduce the spatial resolution of the original resolution to 1/16.Component of the data sets are shown in Table 1.

Table 1: Composition of dataset

	Vaihigen	Potsdam
Labeled images	16	24
GSD	9cm	5cm
Bands	IR,R,G,DSM	IR,R,G,B,DSM
Training set	1,3,5,7,13, 17,21,23,26,32,37	2_10,3_10,3_11,3_12, 4_11,4_12,5_10,5_12, 6_8,6_8,6_10,6_11, 6_12,7_7,7_9,7_11,7_12
Training patches	115088	85000
validation set	11,28,34	2_11,4_10,5_11,7_10
validation patches	28376	25000
Test set	15,30	2_12,6_7,7_8

1.2 Training Settings

The implementation of our network based on the Caffe Library.Our network is initialized and fine-tuned by pre-trained HF-FCN model and trained in an end-to-end manner.HF-FCN is trained on the Massachusetts Building Dataset,which consists of 151 aerial images of the Boston area.The standard parameters used during training the networks are shown in Table 2.In order to avoid over-fitting, we set the drop-out ratio to 0.5.Because the network converged very fast,we trained only 20000 times,then chose the model which performed better both on the validation set and test set for building detection.Training parameters on two data sets are shown in Table 2.

Table 2: Parameters for Network Training

	Vaihigen	Potsdam
input size	256×256	256×256
mini-bachsize	15	15
learning rate	10^-6	10^-5
test_interval	1000	1000
type	SGD	SGD
max_iter	40000	40000
momentum	0.9	0.9
clip_gradients	10000	10000
weight_decay	0.005	0.005

1.3 Results

We applied three different information as input to compare its impact on our network performance.For the Vaihigen data set,we set up three different inputs,the 3-channel inputs are IR,R,G,the 4-channel inputs are IR,R,G,nDSMs,the 5-channel inputs are IR,R,G,DSMs,nDSMs.Table 3 shows the different prediction results.We adopted 3 different evaluation metrics,nameed precision,recall and F1-score.

Table 3: Performance comparison of the results of different inputs on Vaihigen data set

	Img	3_in:IR,R,G			4_in:IR,R,G,nDSMs			5_in:IR,R,G,DSM,nDSM		
		Pre	Rec	F1	Pre	Rec	F1	Pre	Rec	F1
Val	11	0.911	0.906	0.909	0.936	0.900	0.917	0.890	0.900	0.900
	28	0.94	0.875	0.906	0.96	0.792	0.868	0.952	0.823	0.883
	34	0.965	0.899	0.930	0.987	0.902	0.942	0.972	0.918	0.944
	Ave	0.939	0.894	**0.915**	0.961	0.865	0.909	0.939	0.880	0.907
Test	15	0.918	0.930	0.924	0.883	0.917	0.9	0.833	0.931	0.88
	30	0.921	0.929	0.926	0.931	0.827	0.876	0.875	0.877	0.876
	Ave	0.919	0.930	**0.925**	0.907	0.872	0.888	0.858	0.900	0.878

We did a similar experiment on the Potsdam data set,and compared the effects of different inputs on semantic segmentation.Unlike the Vaihigen dataset, on the Potsdam data set the 3-channel inputs are R,G,B, the 4-channel inputs are IR, R,G,B and the 5-channel inputs are IR, R,G,B, nDSMs.Table 4 shows the different results.The evaluation metrics are the same as Vaihigen data set.

Table 4: HF-FCN semantic labelling results on Potsdam data set

	Img	3_in:RGB			4_in:RGB,IR			5_in:RGB,IR,nDSM		
		Pre	Rec	F1	Pre	Rec	F1	Pre	Rec	F1
Val	2_11	0.917	0.950	0.933	0.917	0.978	0.946	0.934	0.976	0.954
	4_10	0.937	0.945	0.941	0.926	0.943	0.936	0.947	0.946	0.946
	5_11	0.930	0.972	0.950	0.959	0.975	0.966	0.956	0.977	0.967
	7_10	0.964	0.536	0.689	0.950	0.590	0.728	0.939	0.554	0.697
	Average	0.937	0.851	0.879	0.937	**0.872**	**0.894**	**0.944**	0.864	0.891
Test	2_12	0.897	0.868	0.882	0.920	0.959	0.939	0.944	0.965	0.955
	6_7	0.894	0.902	0.898	0.915	0.909	0.912	0.901	0.918	0.909
	7_8	0.975	0.929	0.951	0.977	0.950	0.957	0.976	0.946	0.960
	Average	0.922	0.900	0.910	0.937	0.935	0.936	**0.940**	**0.943**	**0.941**