Calculating Disparity Images with Denoising Joint Filtering using Deep Convolutional Networks

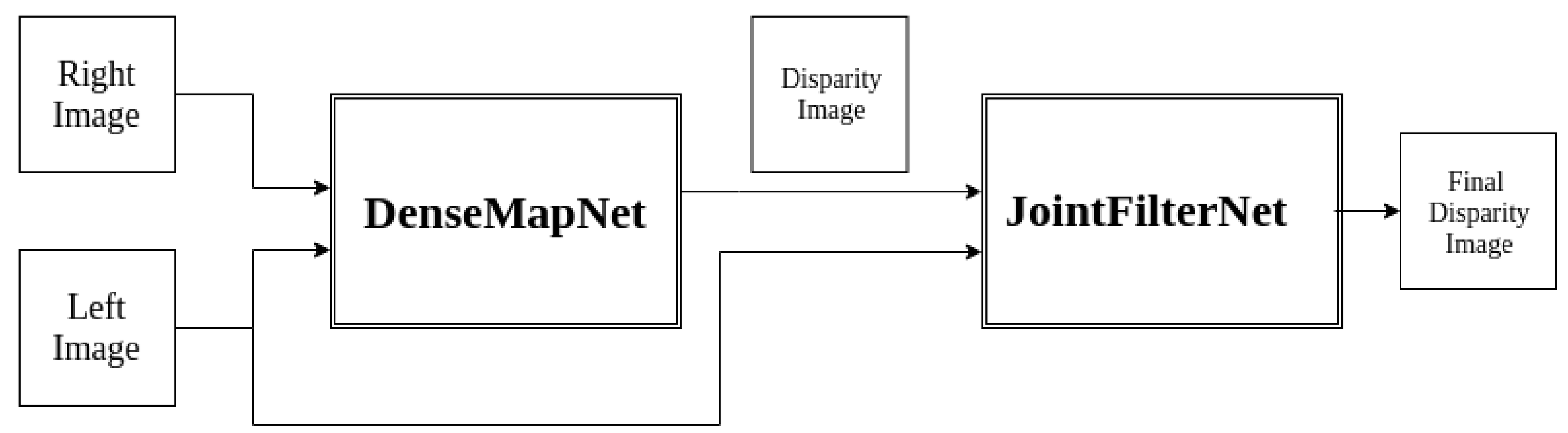
Moran Altmark & Shai Weisman

Supervised by: Dr. Raja Giryes and Dor Bank

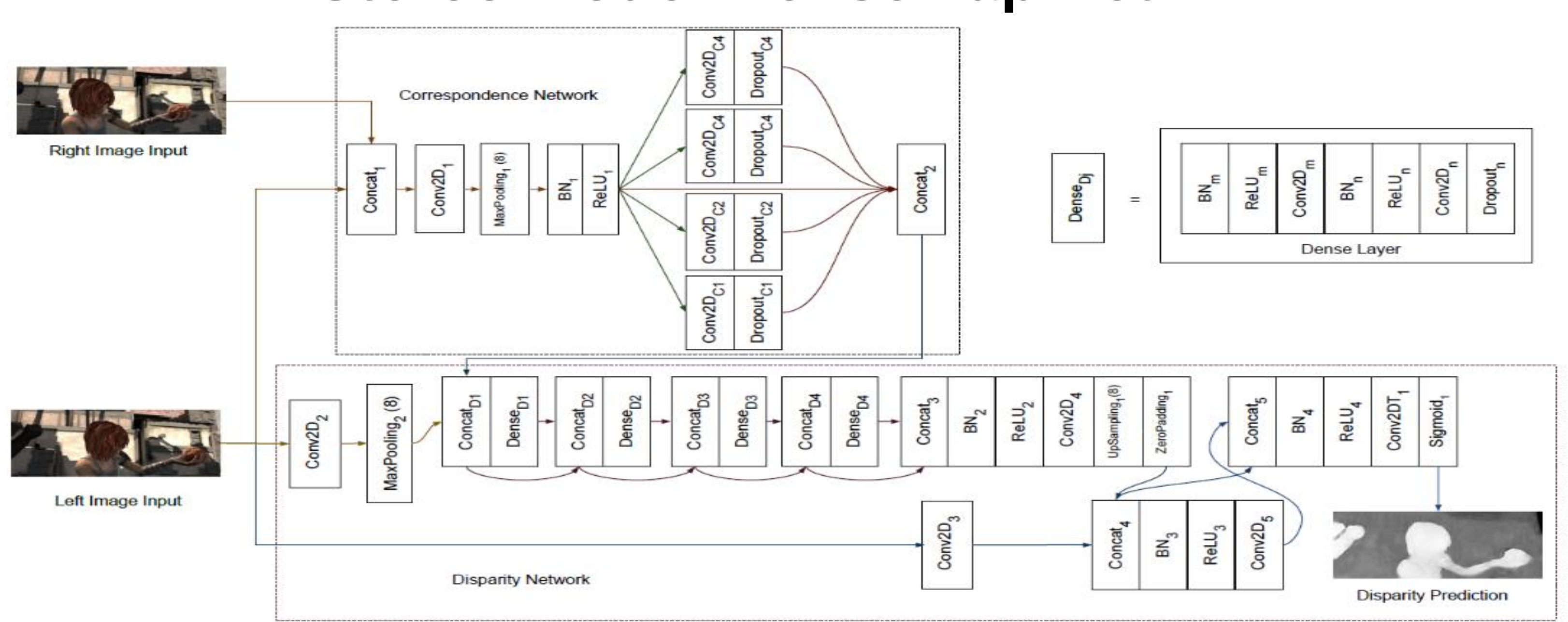
TEL AVIV UNIVERSITY אוניברסיטת תל אביב

The Network

Block Diagram



Stereo Model-DenseMap Net



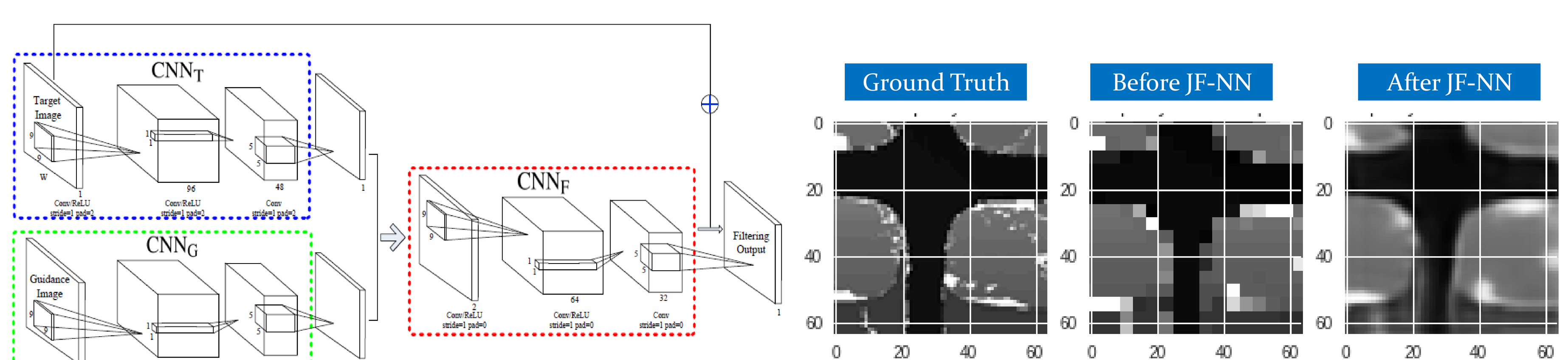
BENCHMARK ON DIFFERENT DATASETS. ALL ERRORS ARE END-POINT-ERRORS (EPE). BASELINE DATA ARE FROM [2].

Method	Sintel	Driving	FlyingThings3D	Monkaa	KITTI 2015	Parameters	Speed	GPU
DispNet	5.38	15.62	2.02	5.99	2.19	38.4M	16.67Hz	NVIDIA Titan X
SGM	19.62	40.19	8.70	20.16	7.21	-	0.91Hz	NVIDIA Titan X
MC-CNN-fast	11.94	19.58	4.09	6.71	-	0.6M	1.25Hz	NVIDIA Titan X
DenseMapNet	4.41	6.56	5.07	4.45	2.52	0.29M	>30Hz	NVIDIA GTX 1080Ti

Fast Stereo matching which reaches state of the art results

R. Atienza, "Fast Disparity Estimation using Dense Networks", 2018

Post Processing Model-Joint Filter Net



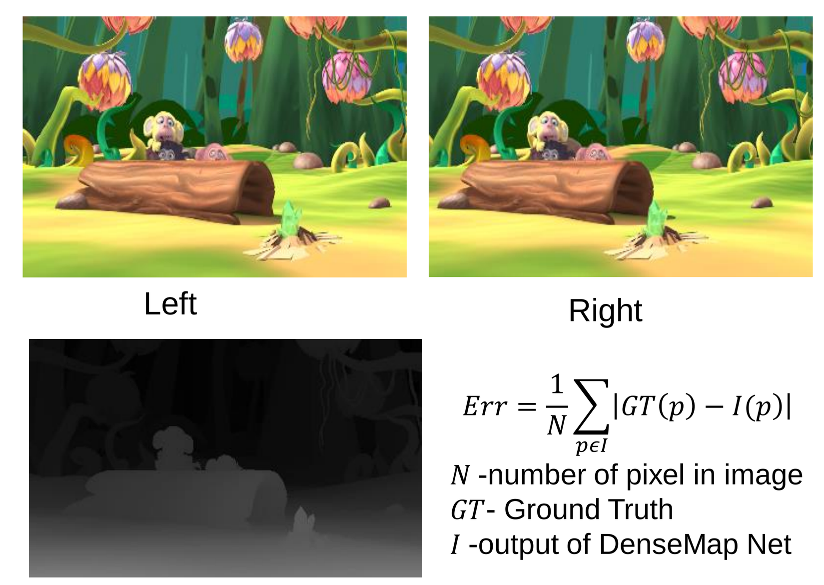
A NN approach for a joint filter, using a 3 sub-network to define the target patch and the guide patch and based on the feature maps, denoise the target image.
The NN showed great value as an edge preserving filter and smoothing quantized images (as seen in example above).

J.-B. H. N. A. a. M.-H. Y. Yijun Li, "Joint Image Filtering with Deep Convolutional Networks", 2017

Results

DenseMap Net

We trained the DenseMap net on a data set of 144 images for training and 17 images for testing and ran 200 epochs using different loss methods. We can note that after 100 epochs the MSE and LogCosh gave a minimal error of 5 (while using SGM on the all dataset, the average error was 20).



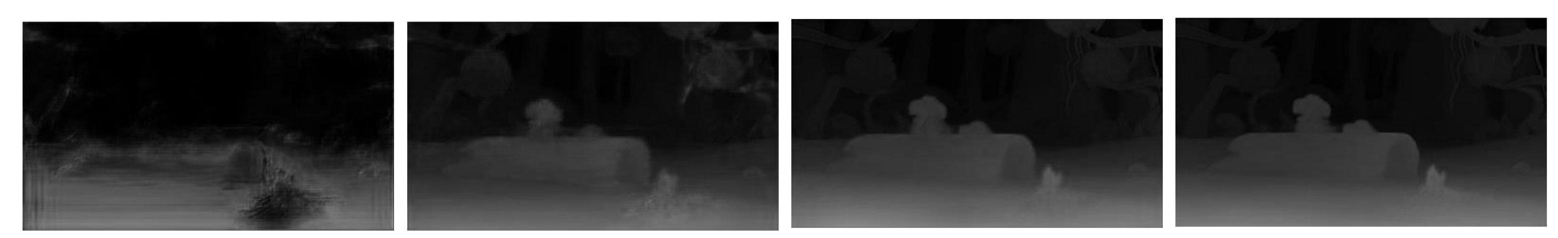
Left Right

Ground Truth

$$Err = \frac{1}{N}\sum_{p\epsilon I}|GT(p) - I(p)|$$

N -number of pixel in image
GT- Ground Truth
I -output of DenseMap Net

Binary Cross Entropy



Epoch 1 Epoch 10 Epoch 50 Epoch 100

Mean Square Error

Epoch 1 Epoch 10 Epoch 50 Epoch 100

Log Cosh

Epoch 1 Epoch 10 Epoch 50 Epoch 100

Joint Filter Net

We trained the network on 50K patches (from 13 images) and tested results on 4 images. After optimization of loss, dropout and batch normalization, the loss error decreased from 5.71 to 2.23.

Loss Function	Batch Normalization	Dropout	Error	JF Error/ DenseMap Error
MSE	X	X	6.59	1.15
MSE	**V**	**X**	**2.23**	**0.39**
MSE	V	0.2	3.24	0.57
MSE	V	0.5	24.98	4.37
Binary	X	X	3.63	0.63
Binary	V	X	8.53	1.49
Binary	V	0.2	2.99	0.52
Binary	V	0.5	7.68	1.34
Logcosh	X	X	9.97	1.74
Logcosh	V	X	7.65	1.34
Logcosh	V	0.2	4.56	0.80
Logcosh	V	0.5	27.66	4.84

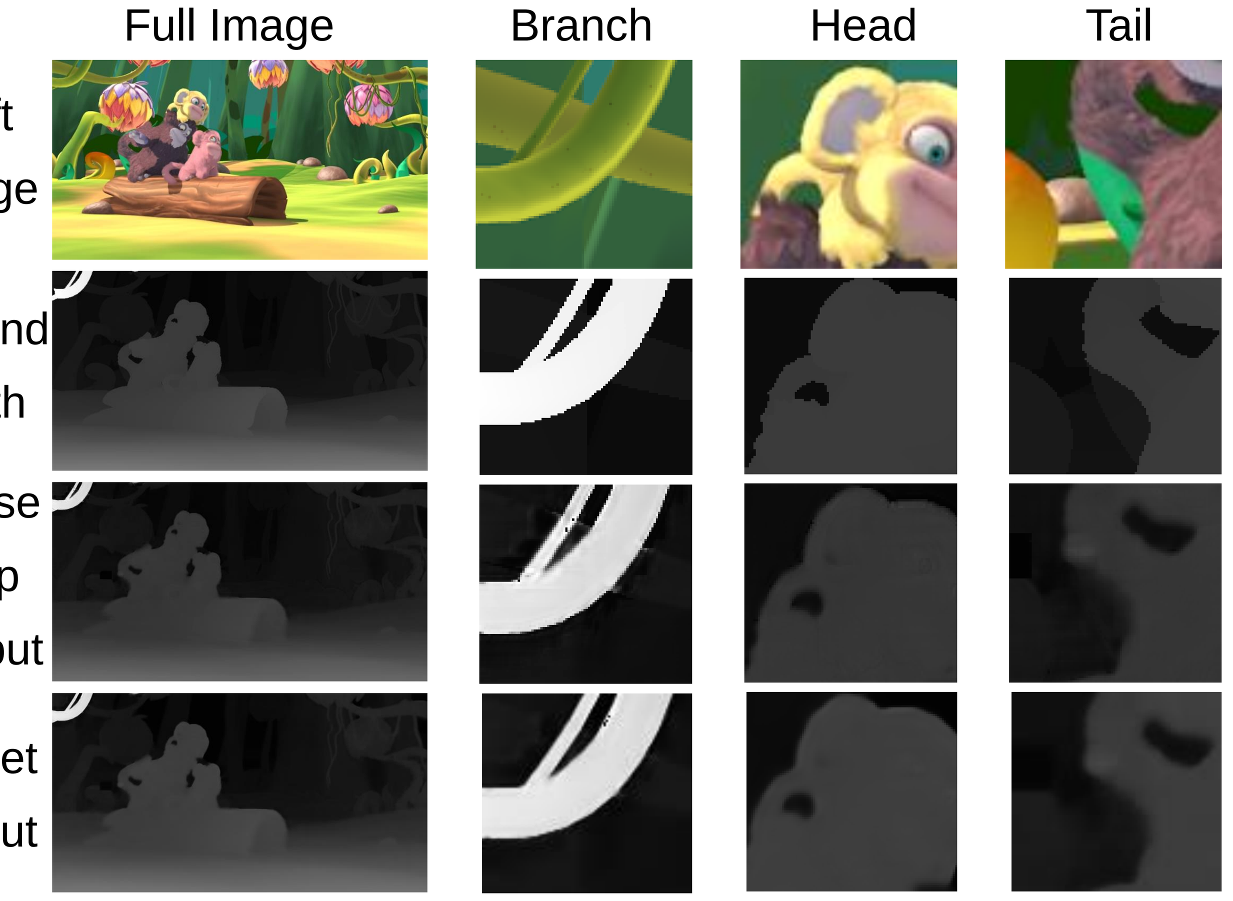
Full Image Branch Head Tail

Left Image

Ground Truth

Dense Map Output

JF-Net Output



Conclusions

The Network showed proof of concept and the feasibility of improving Disparity (Depth) images using a joint Filter Network, **Reducing the error from 5.71 to 2.23.** In addition, we can see the value of different loss function and batch Normalization.
With that said, the **use of a small data set** (163 images compared to 8000 in original DenseMapNet Model) **may cause over fitting** and may require further tuning and optimization.