Supplementary Materials for the Paper: NDDR-CNN: Layerwise Feature Fusing in Multi-Task CNN by Neural Discriminative Dimensionality Reduction

Yuan Gao[1], Jiayi Ma[2] Mingbo Zhao[3] Wei Liu[1] Alan L. Yuille[4]
[1] Tencent AI Lab [2] Wuhan University [3] City University of Hong Kong [4] Johns Hopkins University
{ethan.y.gao, jyma2010, mbzhao4}@gmail.com, wl2223@columbia.edu, alan.yuille@jhu.edu

We conduct additional experiments in this supplementary material file, including:

- Semantic Segmentation and Surface Normal Prediction using AlexNet [1] backbone.
- Additional ablation analysis for the cross-stitch network on VGG-16 backbone [4] to verify the hyperparameters for the cross-stitch network used in our main text are optimal.

1. Semantic Segmentation and Surface Normal Prediction on AlexNet

We conduct *Semantic Segmentation* and *Surface Normal Prediction* on AlexNet [1] with FCN32s [2], as those in the cross-stitch network paper [3]. We also use the same hyperparameters as the those in [3]. The results in Table S1 show that our method outperforms the cross-stitch network and the sluice network on AlexNet.

	Surface Normal Prediction					Semantic Seg.	
	Angle Dist.		Within t° (%)			(%)	
AlexNet	Mean	Med.	11.25	22.5	30	mIoU	PAcc
C.-S.	19.7	17.1	28.1	65.9	**80.0**	21.7	53.4
Sluice	19.5	16.6	29.7	66.2	79.5	21.9	53.8
Ours	**19.4**	**15.5**	**36.6**	**66.8**	79.2	**23.1**	**56.3**

Table S1. The results for *Semantic Segmentation* and *Surface Normal Prediction* on **AlexNet**.

2. Ablation Analysis for the Cross-Stitch Network on VGG-16

In this section, we verify that, in our main text, we have fair comparisons with the state-of-the-art cross-stitch network, especially regarding the hyperparameters on different network backbones. In other words, we show that the hyperparameters for the cross-stitch network, originally obtained from [3] on AlexNet, are still the best for other network backbones. This can be investigated by doing ablation analysis of the cross-stitch network on other network backbones. The ablation analysis of the cross-stitch network on VGG-16 [4] is shown in Table S2, which demonstrates that the best hyperparameters of the cross-stitch network have been used in our main text for fair comparative-evaluation.

	Surface Normal Prediction					Semantic Seg.	
	Angle Dist.		Within t° (%)			(%)	
(α, β)	Mean	Med.	11.25	22.5	30	mIoU	PAcc
(0.9, 0.1)	**15.2**	11.7	48.6	**76.0**	**86.5**	**34.8**	**65.0**
(0.7, 0.3)	15.5	**11.6**	**48.7**	75.1	85.5	34.4	64.6
(0.5, 0.5)	15.9	12.0	47.5	73.7	84.4	33.9	64.0
Scale	Mean	Med.	11.25	22.5	30	mIoU	PAcc
1	15.3	11.9	47.9	75.8	86.3	34.5	64.6
10	15.5	12.0	47.3	75.1	86.0	35.0	65.0
10^2	15.3	11.8	48.1	75.6	86.2	**35.1**	**65.2**
10^3	**15.2**	**11.7**	**48.6**	**76.0**	**86.5**	34.9	65.0

Table S2. Ablation analysis for the cross-stitch network on **VGG-16**. This is to ensure that the hyperparameters for the cross-stitch network, *i.e.*, $(\alpha, \beta) = (0.9, 0.1)$ and 1000x learning rate for fuse layers, used in our main text are the best ones for the cross-stitch network.

References

[1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *NIPS*, 2012. 1

[2] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *CVPR*, pages 3431–3440, 2015. 1

[3] Ishan Misra, Abhinav Shrivastava, Abhinav Gupta, and Martial Hebert. Cross-stitch networks for multi-task learning. In *CVPR*, 2016. 1

[4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2015. 1