AI-driven analysis and image acquisition of in-vivo neuronal network activity

Xingying Chen, Bastian Krämer
Michal Klein, René Romen

August 5, 2019

Department of Mathematics
Technical University of Munich

Table of Contents

Table of Contents

Zebrafish Larvae

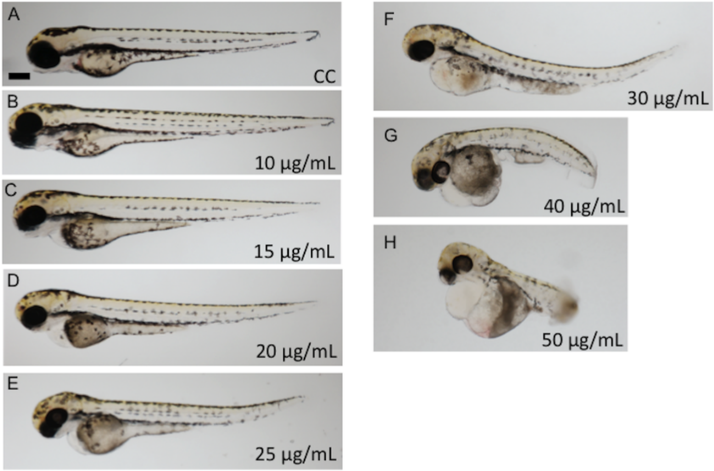


Figure 1: Zebrafish[5]

Light Field Microscope

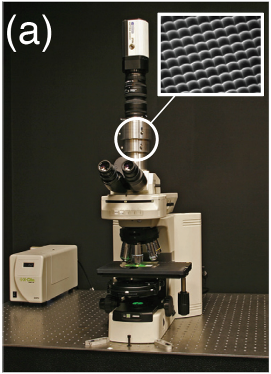


Figure 2: (a) A microscope with a microlens array[3]

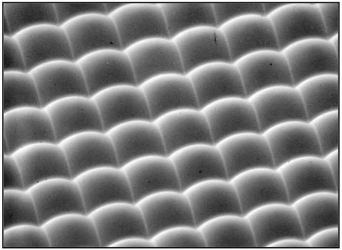


Figure 3: Microlens array[1]

Advantages of Light Field Microscopy Images

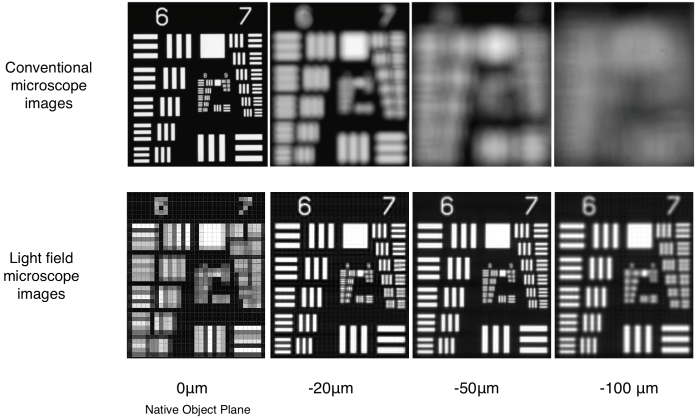


Figure 4: Comparison of images at different focal planes[3]

Light Field Microscope

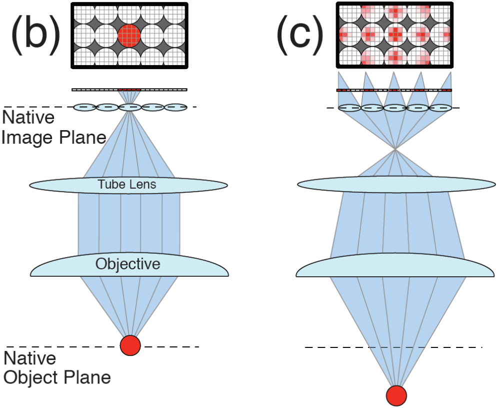


Figure 5: (b), (c) The red point stands for a point source generating illumination. Red regions on top shows the intensity of illumination arriving at the sensor plane[3].

Light Field Microscope

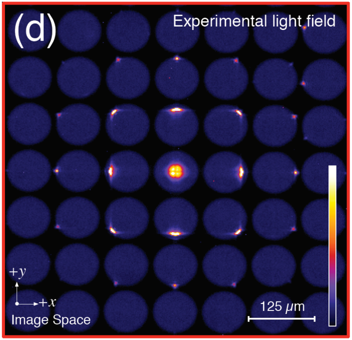
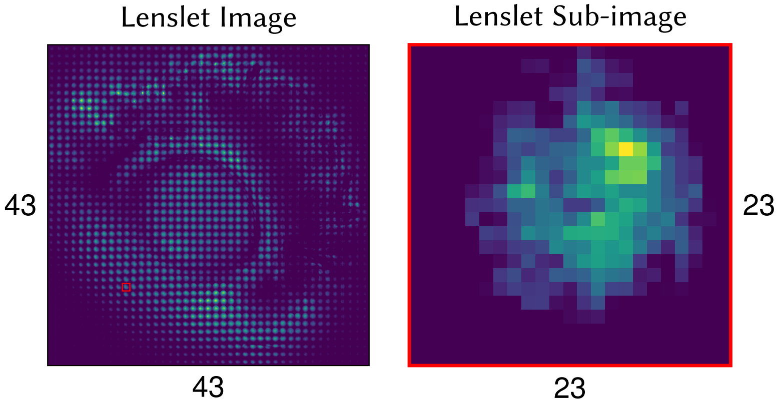


Figure 6: Example light field of a sphere[3]

Light Field Transformations



- Angular resolution: 43×43
- Spatial resolution: 23×23

Reordering Light Field

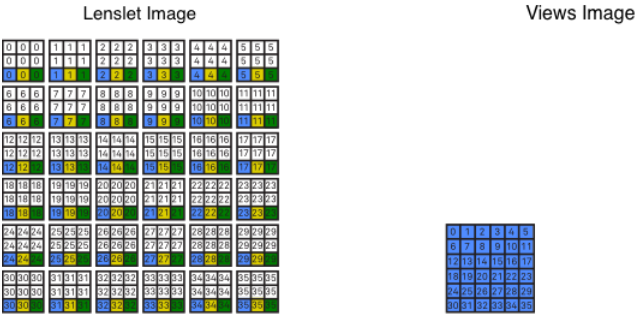


Figure 7: Reordering lenslet to views image[6]

Reordering Light Field

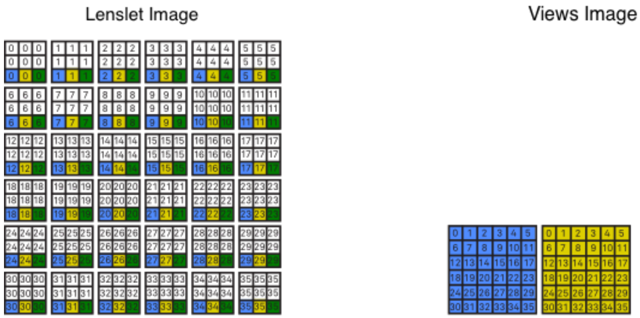


Figure 8: Reordering lenslet to views image[6]

Reordering Light Field

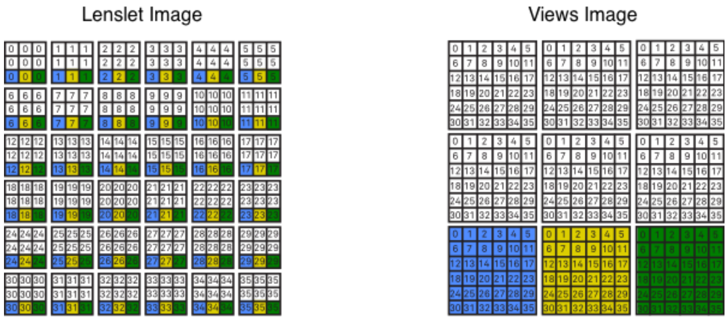
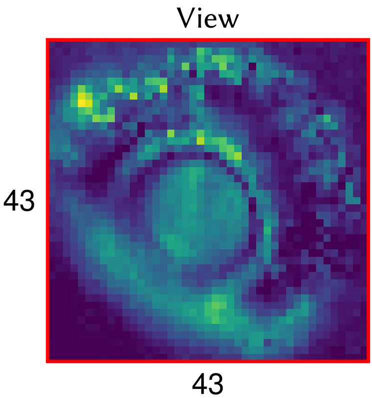
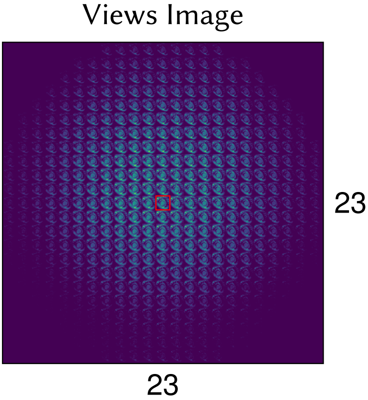


Figure 9: Reordering lenslet to views image[6]

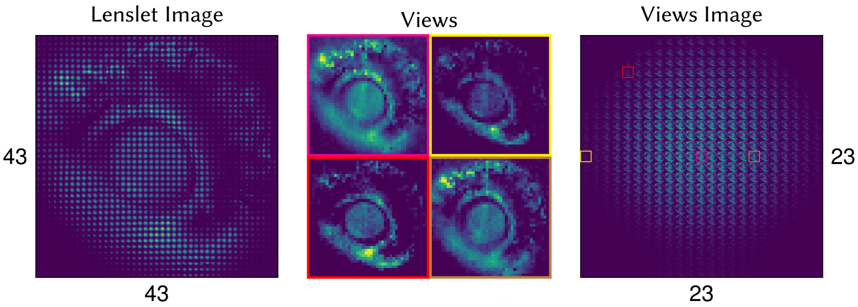
Light Field Transformations



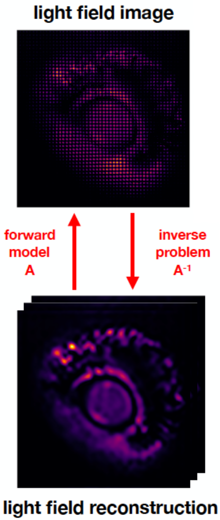


- Angular resolution: 43×43
- Spatial resolution: 23×23

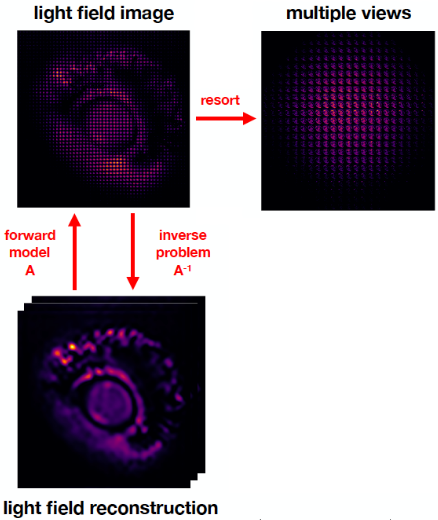
Light Field Image



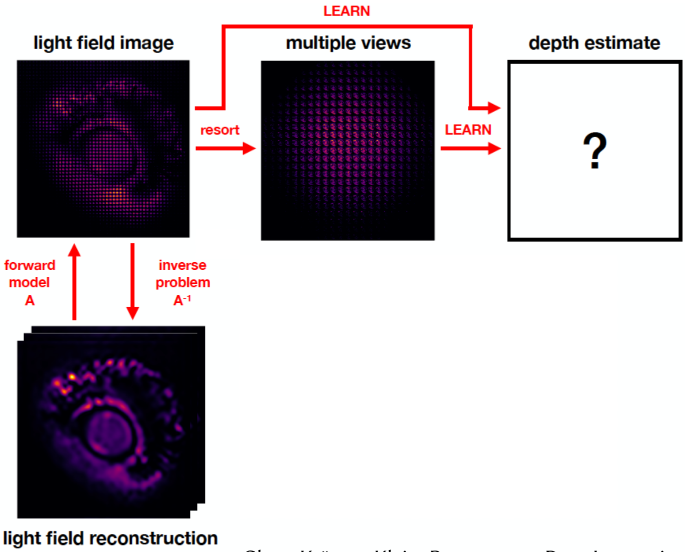
Goal


light field image
forward model
A
inverse problem
A-1
light field reconstruction

Goal


light field image
multiple views
resort
forward model A
inverse problem A^{-1}
light field reconstruction

Goal


LEARN
light field image
multiple views
depth estimate
resort
LEARN
?
forward
model
A
inverse
problem
A-1
light field reconstruction

Goal

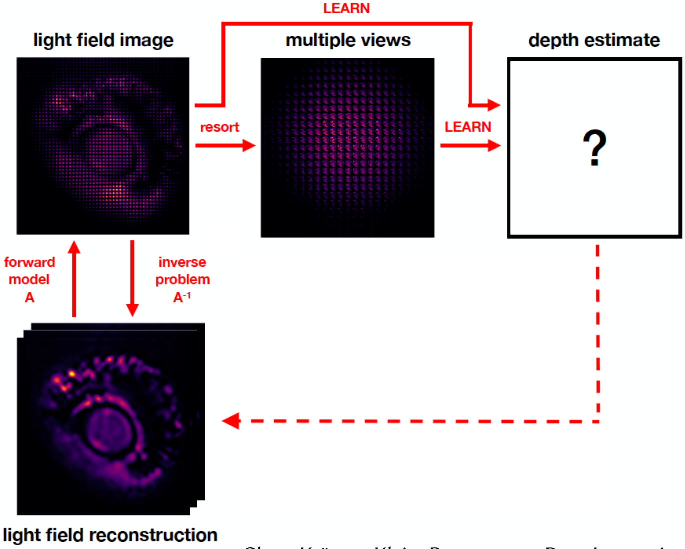

LEARN
light field image
multiple views
depth estimate
resort
LEARN
?
forward model A
inverse problem A-1
light field reconstruction

Table of Contents

Datasets - Real Images

- Fish eye

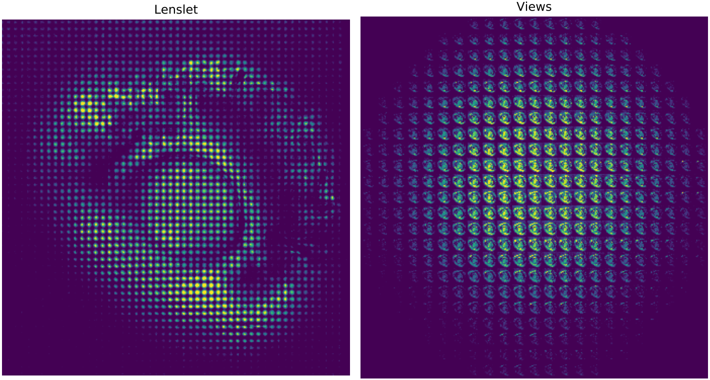


Figure 10: Fish eye

Datasets - Real Images

- Spheres

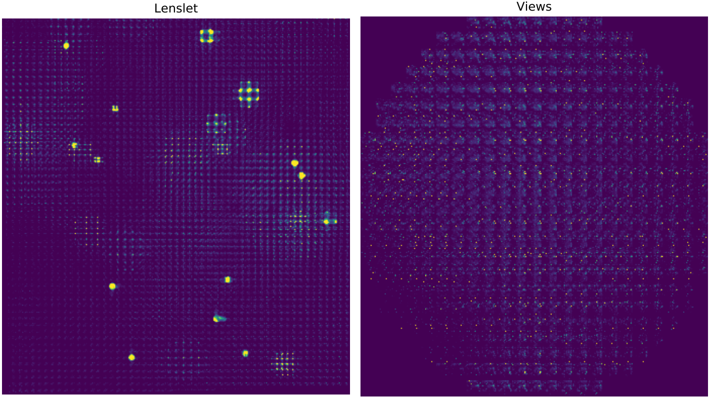


Figure 11: Spheres

Dataset - Training Data

- Created own dataset containing geometric objects

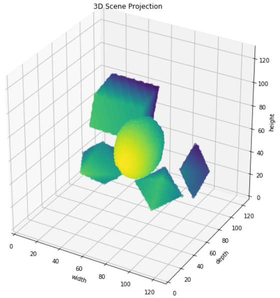


Figure 12: Scene

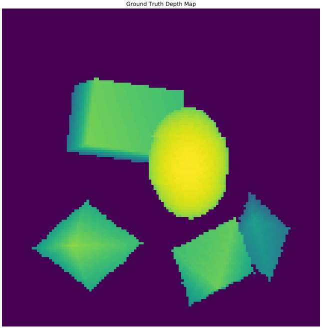


Figure 13: Resulting depth map

Dataset

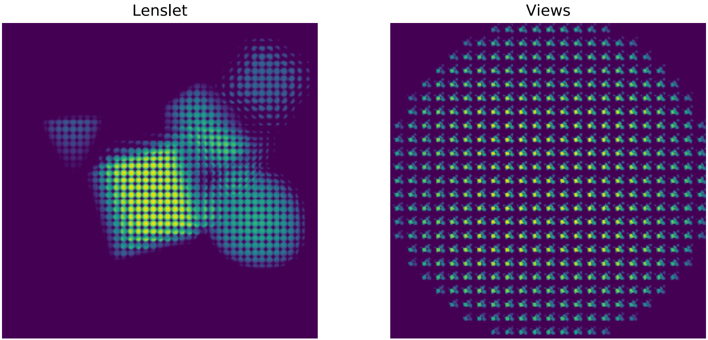


Figure 14: Resulting lenslet and views image

- 369 images

Table of Contents

Overview of Networks

- Input - views image
  - Epinet
  - Views network
- Input - lenslet image
  - Lenslet network
  - Lenslet classification network

Overview of Networks

- Input - views image
  - **Epinet**
  - Views network
- Input - lenslet image
  - Lenslet network
  - Lenslet classification network

Epinet

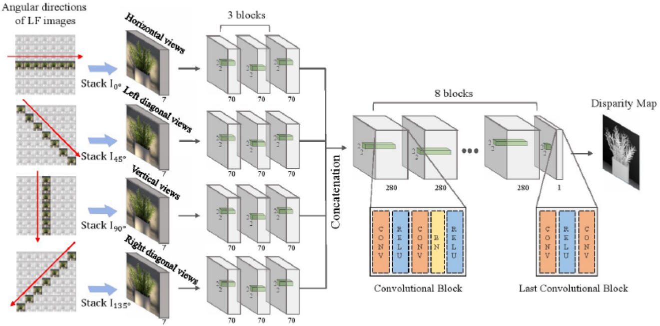


Figure 15: Epinet architecture [9]

- Multi-stream network

Epinet

- 2.2 million parameters
- Multistream convolutional encoder
- Upscaling: nearest neighbor interpolation
- Optimizer: RMSProp (no learning rate decay)
- Regularization: None
- Activation function: Leaky ReLU
- Learning rate: 10^{-4}
- Error: Mean squared error

Overview of Networks

- Input - views image
  - Epinet
  - **Views network**
- Input - lenslet image
  - Lenslet network
  - Lenslet classification network

Views Network

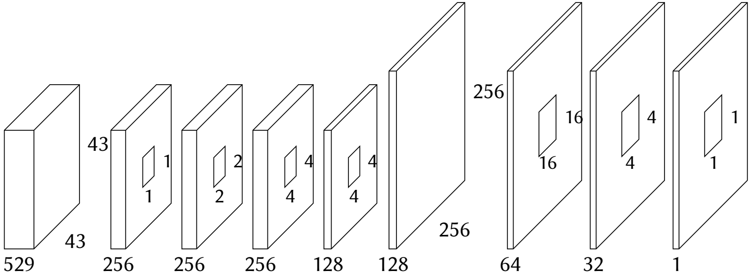


Figure 16: Left to right: input, 4 convolution layers, upscaling by nearest neighbor interpolation, 3 convolution layers.

Views Network

- 4.1 million parameters
- Convolutional encoder
- Upscaling: nearest neighbor interpolation
- Optimizer: Adam (no learning rate decay)
- Regularization: None
- Activation function: Leaky ReLU
- Learning rate: 10^{-6}
- Error: Absolute error

Overview of Networks

- Input - views image
  - Epinet
  - Views network
- Input - lenslet image
  - **Lenslet network**
  - Lenslet classification network

Lenslet Network

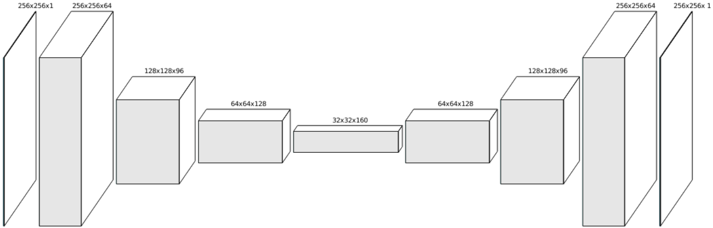


Figure 17: General structure of a convolution encoder-decoder architecture. We use Inception-ResNet-v2 [10] blocks in encoder.

Lenslet Network

- 7.6 million parameters
- Inception-ResNet-v2 blocks encoder
- Upscaling: fast up-convolution (x3)
- Optimizer: Adam (exp. decaying learning rate (400 examples))
- Regularization: L2 (scale: $5 * 10^{-6}$)
- Activation function: Leaky ReLU (after each block)
- Learning rate: $5 * 10^{-6}$
- Error: Mean squared error

Overview of Networks

- Input - views image
  - Epinet
  - Views network
- Input - lenslet image
  - Lenslet network
  - **Lenslet classification network**

Lenslet Classification Network

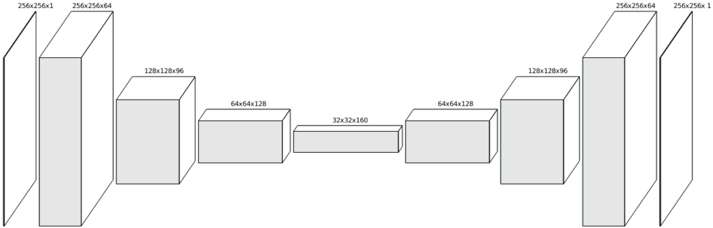


Figure 18: General structure of a convolution encoder-decoder architecture. We use Inception-ResNet-v2[10] blocks in encoder.

Lenslet Classification Network

- 5 million parameters
- Inception-ResNet-v2 encoder
- Upscaling: fast up-convolution (x3)
- Optimizer: Adam (exp. decaying learning rate (400 examples))
- Regularization: L2 (scale: $5 * 10^{-6}$)
- Activation function: Leaky ReLU (after each block)
- Learning rate: $5 * 10^{-6}$
- Error: Mean absolute error

Table of Contents

Quantitative Results

Network	MSE	BPR ($\delta = 0.15$)	Input	#params (millions)
Epinet	**0.0049958**	**0.04304**	Stacked views	2.2
Views	0.0122427	0.12022	Stacked views	4.1
Lenslet	0.00595414	0.05586	Lenslet	7.6
Lenslet (cls)	26.9925	0.99521	Lenslet	5

Qualitative Results - Test Data

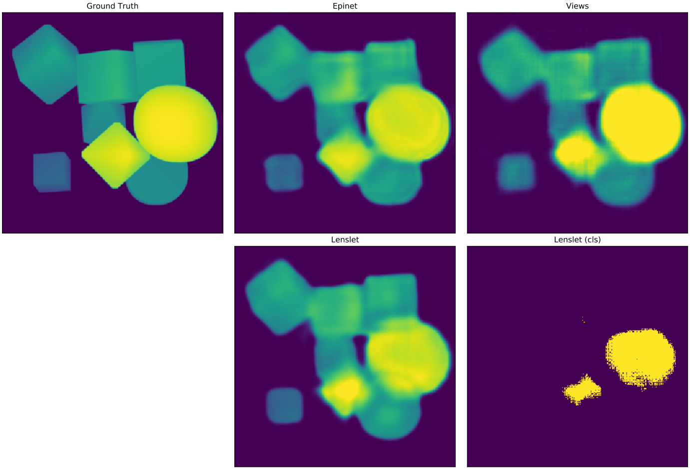


Figure 19: Depth prediction of an image from test set after 300 epochs.

Qualitative Results - Fish Eye

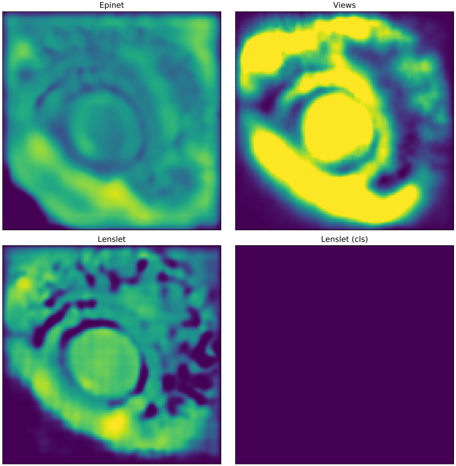


Figure 20: Depth prediction of fish eye after 300 epochs.

Qualitative Results - Spheres

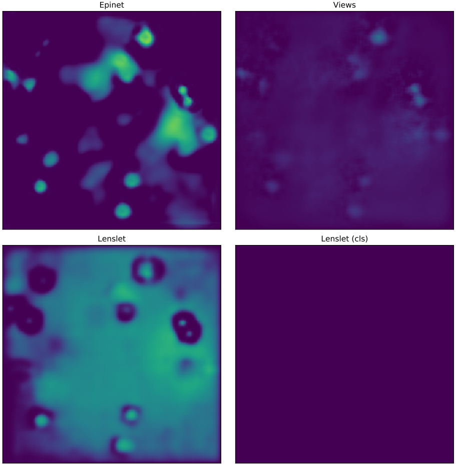


Figure 21: Depth prediction of spheres after 300 epochs.

Discussion - Real Data

- Image of fish eye and spheres taken with different configurations

Discussion - Real Data

- Image of fish eye and spheres taken with different configurations
- Better networks on the test set pick up more noise

Discussion - Real Data

- Image of fish eye and spheres taken with different configurations
- Better networks on the test set pick up more noise
- No ground truth to compare the results

Table of Contents

Discussion - Generated Dataset

- Simulation needed to overcome lack of data

Discussion - Generated Dataset

- Simulation needed to overcome lack of data
- General problem when using deep networks on light field data

Discussion - Generated Dataset

- Simulation needed to overcome lack of data
- General problem when using deep networks on light field data
- Geometric objects are biologically not plausible

Discussion - Generated Dataset

- Simulation needed to overcome lack of data
- General problem when using deep networks on light field data
- Geometric objects are biologically not plausible
- Real data contains noise

Discussion - Generated Dataset

- Simulation needed to overcome lack of data
- General problem when using deep networks on light field data
- Geometric objects are biologically not plausible
- Real data contains noise
- Realistic simulation with unrealistic scene

Discussion - Light Field Data

- Typical light field resolutions in research papers:
  - Spatial resolution of 512×512
  - Angular resolution of 3×3, 5×5 or 9×9
- e.g. see EPINET[9], VommaNet[8]

Discussion - Light Field Data

- Typical light field resolutions in research papers:
  - Spatial resolution of 512×512
  - Angular resolution of 3×3, 5×5 or 9×9
- e.g. see EPINET[9], VommaNet[8]
- Light field microscopy data in our project:
  - Spatial resolution of 43×43
  - Angular resolution of 23×23

Discussion - Light Field Data

- Depth predictions with a resolution of 43×43 too small

Discussion - Light Field Data

- Depth predictions with a resolution of 43×43 too small
- Upsampling of spatial resolution to 256×256

Discussion - Light Field Data

- Depth predictions with a resolution of 43×43 too small
- Upsampling of spatial resolution to 256×256
- Differences in baseline and depth range

Discussion - Light Field Data

- Depth predictions with a resolution of 43×43 too small
- Upsampling of spatial resolution to 256×256
- Differences in baseline and depth range
- Architectures proposed in research papers were not directly suited for our data

Outlook

- Realistic scene generation

Outlook

- Realistic scene generation
- Networks that are invariant to camera configurations

Outlook

- Realistic scene generation
- Networks that are invariant to camera configurations
- Light field reconstruction using depth prediction

Summary

Thank you for your attention

Questions?

Bibliography I

[1] https://www.rpcphotonics.com/product/mla-s100-f12/.

[2] Nearest neighbour interpolation.

[3] M. Broxton, L. Grosenick, S. Yang, N. Cohen, A. Andalman, K. Deisseroth, and M. Levoy.
Wave optics theory and 3-d deconvolution for the light field microscope.
Optics express, 21(21):25418–25439, 2013.

[4] X. Duan, X. Ye, Y. Li, and H. Li.
High quality depth estimation from monocular images based on depth prediction and enhancement sub-networks.
In *2018 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2018.

Bibliography II

[5] L. D. Ellis, E. C. Soo, J. C. Achenbach, M. G. Morash, and K. H. Soanes.
Use of the zebrafish larvae as a model to study cigarette smoke condensate toxicity.
PLoS One, 9(12):e115305, 2014.

[6] C. Hahne, A. Aggoun, V. Velisavljevic, S. Fiebig, and M. Pesch.
Baseline and triangulation geometry in a standard plenoptic camera.
International Journal of Computer Vision, 126(1):21–35, Jan 2018.

[7] I. Laina, C. Rupprecht, V. Belagiannis, F. Tombari, and N. Navab.
Deeper depth prediction with fully convolutional residual networks.
In *2016 Fourth international conference on 3D vision (3DV)*, pages 239–248. IEEE, 2016.

[8] H. Ma, H. Li, Z. Qian, S. Shi, and T. Mu.
Vommanet: an end-to-end network for disparity estimation from reflective and textureless light field images.
arXiv preprint arXiv:1811.07124, 2018.

Bibliography III

[9] C. Shin, H.-G. Jeon, Y. Yoon, I. So Kweon, and S. Joo Kim.
Epinet: A fully-convolutional neural network using epipolar geometry for depth from light field images.
In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4748–4757, 2018.

[10] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi.
Inception-v4, inception-resnet and the impact of residual connections on learning.
In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

Backup Slides

Depth Estimate

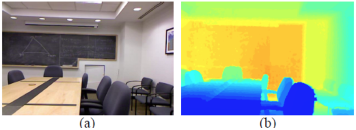


Figure 22: Example of a depth estimate.[4]

Bad Pixel Ratio

$$BPR(x, \hat{x}; \delta) = \frac{1}{NM} \sum_{i=0}^{N} \sum_{j=0}^{M} I[\delta < |x_{i,j} - \hat{x}_{i,j}|]$$

Upsampling Strategies

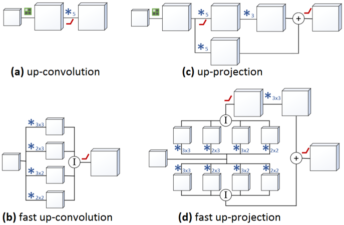


Figure 2. **From up-convolutions to up-projections. (a)** Standard up-convolution. **(b)** The equivalent but faster up-convolution. **(c)** Our novel up-projection block, following residual logic. **(d)** The faster equivalent version of (c)

Figure 23: Different up-sampling strategies[7]

Upsampling Strategies

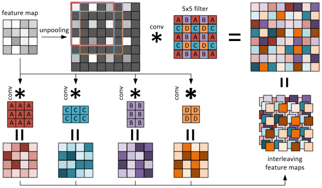


Figure 3. **Faster up-convolutions.** Top row: the common up-convolutional steps: unpooling doubles a feature map's size, filling the holes with zeros, and a 5×5 convolution filters this map. Depending on the position of the filter, only certain parts of it (A,B,C,D) are multiplied with non-zero values. This motivates convolving the original feature map with the 4 differently composed filters (bottom part) and interleaving them to obtain the same output, while avoiding zero multiplications. A,B,C,D only mark locations and the actual weight values will differ

Figure 24: Faster up-convolution[7]

Upsampling Strategies

- Deconvolution can cause checkerboard artifacts

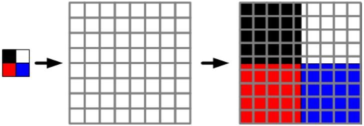


Figure 25: Nearest neighbor upsampling[2]

Lenslet Network - Fish Eye - Fewer Epochs

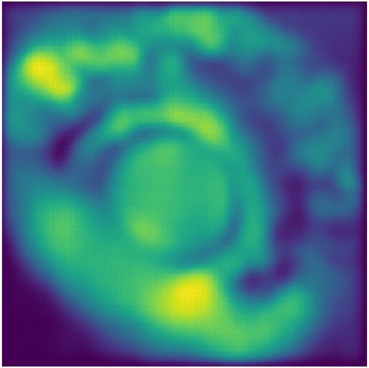


Figure 26: Spheres prediction (best epoch)

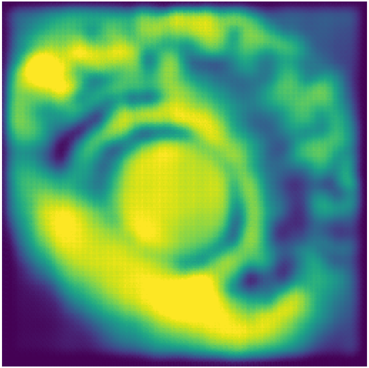


Figure 27: Spheres prediction (last epoch)

Lenslet Network - Spheres - Fewer Epochs

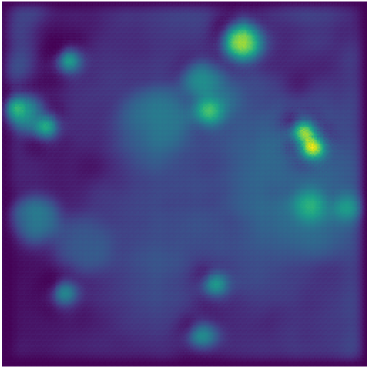


Figure 28: Spheres prediction (best epoch)

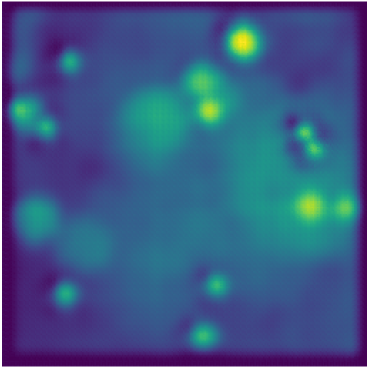


Figure 29: Spheres prediction (last epoch)

Lenslet Classification Network - Fewer Epochs

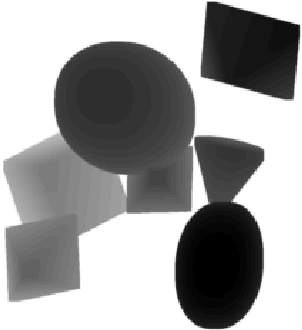


Figure 30: Ground truth.

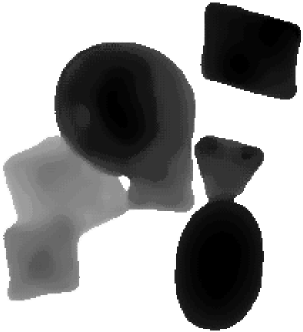


Figure 31: Predicted depth map.