

Multiobject Tracking Repetitive Patterns with Autoencoder

Antonin Sulc, Bastian Goldlücke

Computer Vision and Image Analysis University of Konstanz

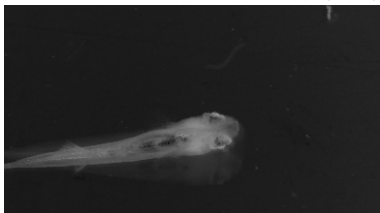
Github

<https://tinyurl.com/y5mnubtd>

Provide tools for tracking and analysis of their repetitive patterns.

Key Features

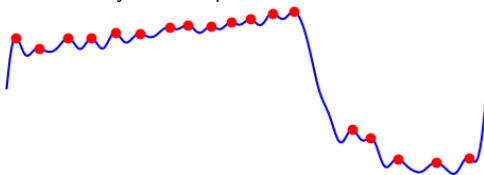
Selection of Region of Interest (ROI)



Tracking Individuals or Their Parts

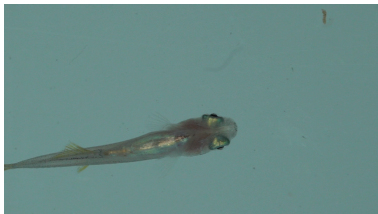


Statistical Analysis of Representation of Tracked Object

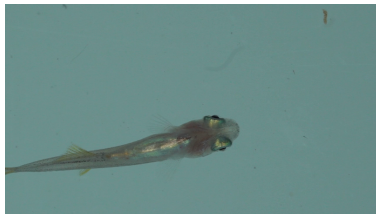


Region of Interest (ROI) Selection

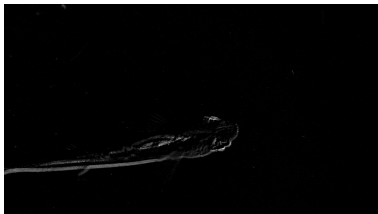
A Frame



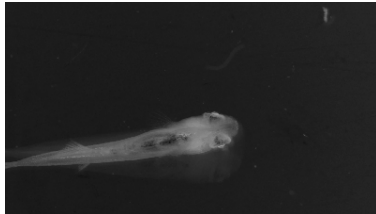
Next Frame



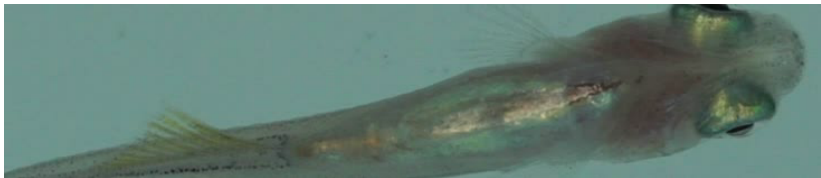
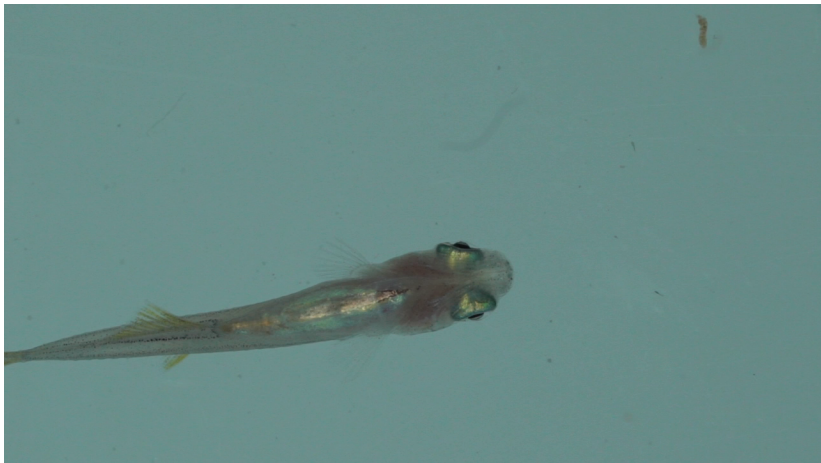
Frame Difference



Frame Difference of All Frames

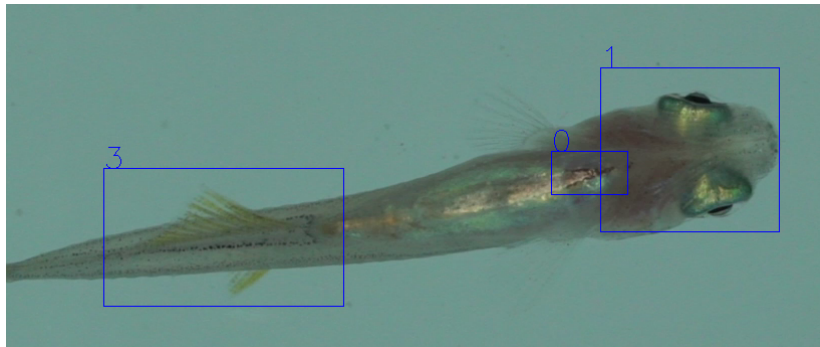


ROI



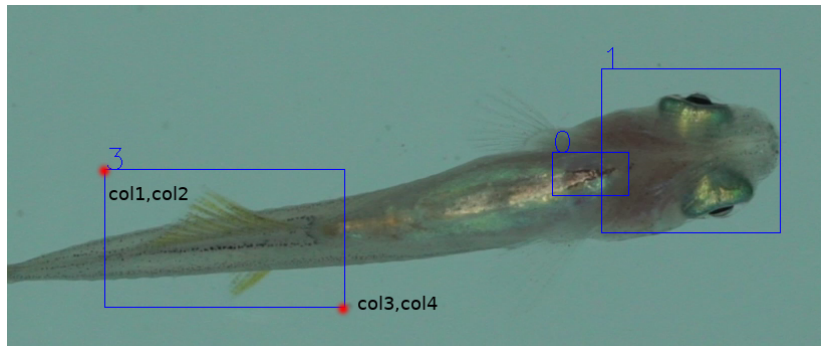
Multioobject Tracking

Tracking - Selection of Tracked Objects



Switch between trackers by 0 - 9, Q to Quit

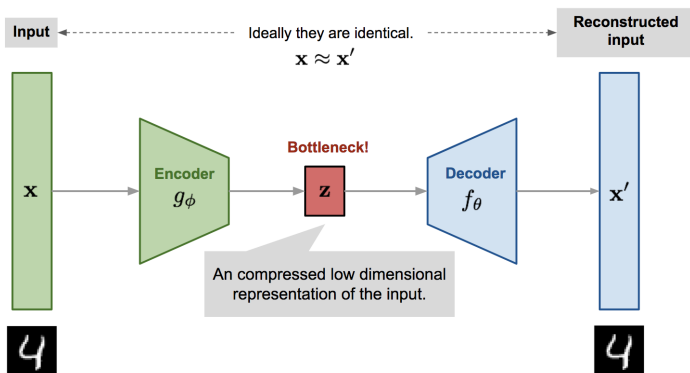
Tracking Outputs



- ▶ A numpy volume
- ▶ Bounding boxes
- ▶ Video with all tracklets visualised

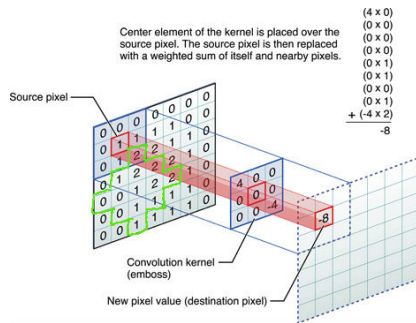
Finding Repetitive Patterns with Autoencoder

Autoencoder



<https://lilianweng.github.io/lil-log/2018/08/12/from-autoencoder-to-beta-vae.html>

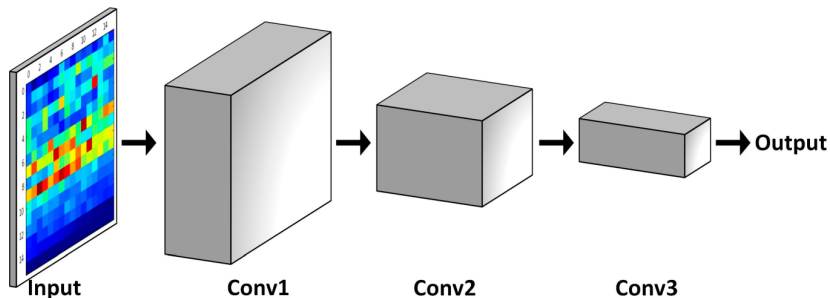
Convolution



<https://stats.stackexchange.com/questions/235032/any-use-of-non-rectangular-shaped-kernels-in-convolutional-neural-networks-espe>

- ▶ Images are matrices, videos are volumes
- ▶ Convolutions are fast trainable transformations of the images/volumes
- ▶ The convolved image usually gets smaller

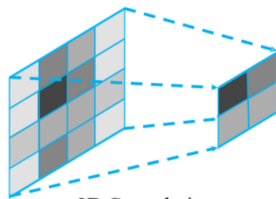
Convolutional Neural Net



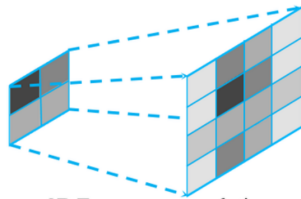
<https://www.mdpi.com/1424-8220/17/8/1729>

- ▶ The outputs from convolutions are called **feature maps**
- ▶ With each convolution the input image gets smaller spatially but thicker in feature dimension.

Deconvolution - Upsampling



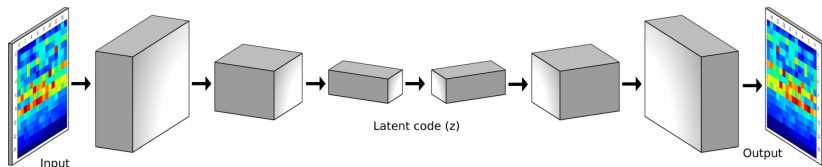
2D Convolution



2D Transpose convolution

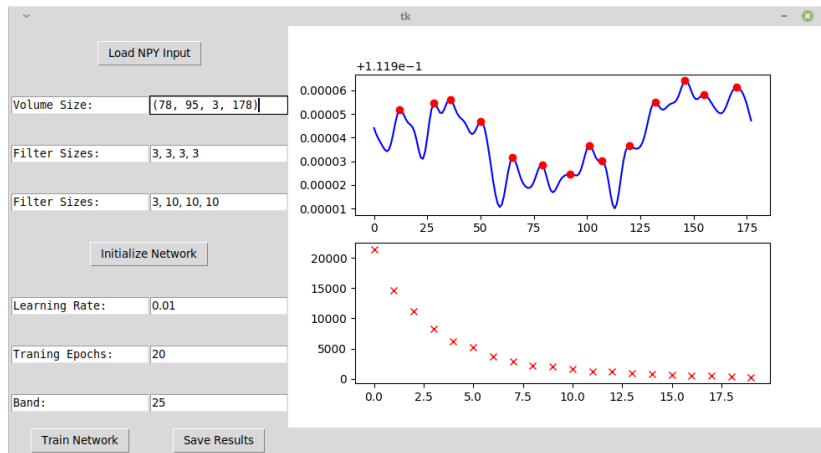
- ▶ The convolution downsamples the image.
- ▶ The transpose convolution **upsamples** the input feature map.

Autoencoder



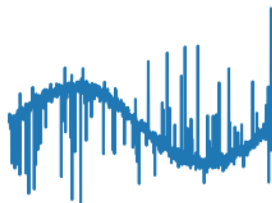
- ▶ After a series of convolutions, we get a **latent code** z .
- ▶ After a series of transposed convolutions, we reproduce the same size as input again.
- ▶ We train the convolutions and transposed convolutions to reproduce the same input.
- ▶ The latent code z is the bottleneck and should represent fine changes in the image

Autoencoder GUI



Bandwidth Selection

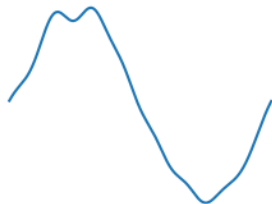
original



band 20



band 10



band 5

