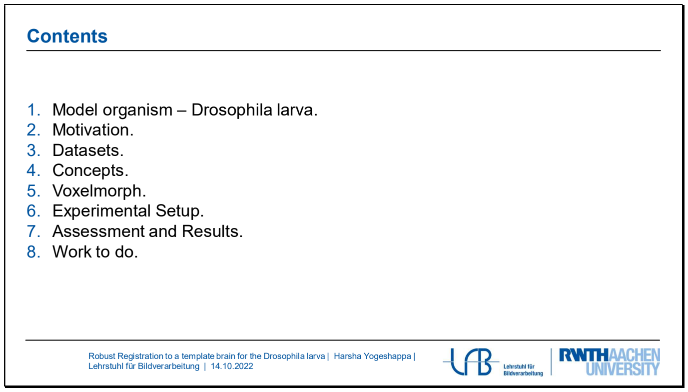
Slide 1

Hello everybody,

Welcome to the interim presentation of my master thesis.

The title of my master thesis is "Robust Registration to a template brain for the Drosophila larva", done under the supervision of Dr. Martin Strauch.

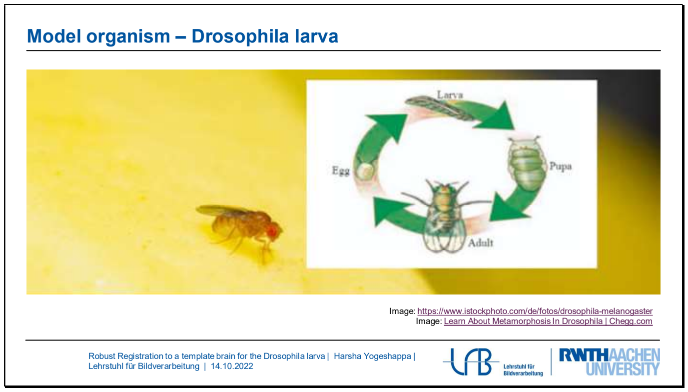
Slide 2



Before I begin to present the thesis and its current status, I would like to give you a brief overview of the sections I will be discussing.

And these are they.

Slide 3

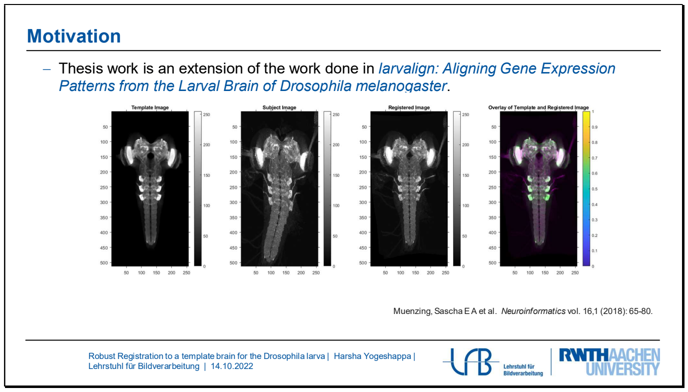


The Drosophila, commonly known as the fruit fly, has been used in scientific research and the study of neuroscience for about a century.

As you can see from the picture, and as most of you know, the fruit fly goes through different stages in its growth before becoming an adult: Egg - Larva - Pupa - Adult.

In this work we are working on 3D scans of the central nervous system of the Drosophila larva.

Slide 4



My thesis is an extension of the work in the Larvalign paper co-authored by Dr. Martin Strauch and Professor Dorit Mehrof.

Larvalign is a standard volume template (obtained by registration) and is also a registration method to register 3D volumes to it.

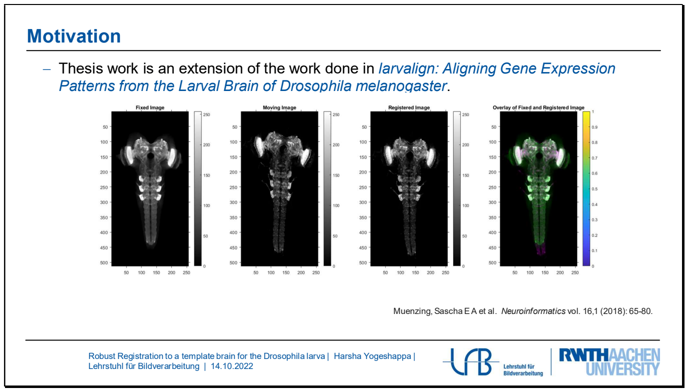
The following figure is an example of a Larvalign registration. While the registration itself is performed on the 3D volume, the figures shown here are the projection of the maximum intensity of the respective 3D volume data for visualization purposes.

On the left side you can see the template image against which the registration is performed and the subject image that needs to be registered.

In the image registration field, the template images are called fixed images and the subject images are called moving images. So, from now on, the template image will be referred to as a fixed image and the subject image will be referred to as a moving image.

On the right side you can see the registered image and an overlay between the registered image and the fixed image. The green represents the fixed image and the magenta represents the moving image. And wherever there is a perfect overlap, you see gray (or the original intensity values).Based on this overlap, we can visually inspect where the registration result has failed.

Slide 5



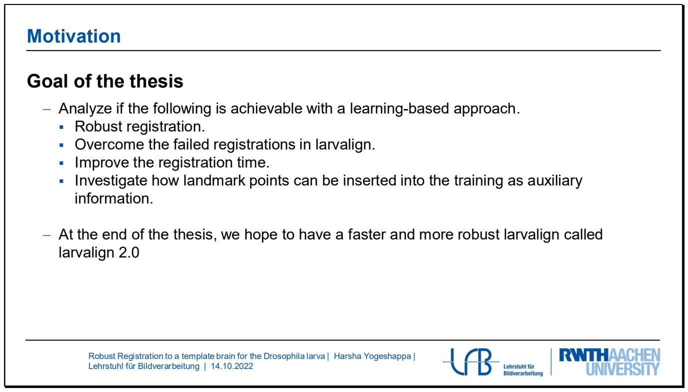
However, for a few volume scans, the registration output of larvalign is not perfect. This can be seen at the tip of the ventral nerve cord, as shown in this example.

larvalign is a parametric method of image registration. The success or failure of registration depends largely on the choice of parameters and the values selected for each of them. And there are several such parameters that can be set.

Therefore, by definition, Larvalign is non-learning and starts from scratch each time it registers new image pairs, repeating the same errors each time, as shown in this example.

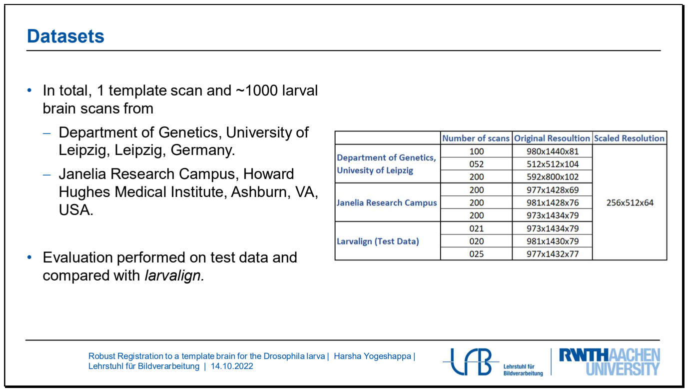
We hope to address this problem by moving from a non-learning type of registration to a learning type, in the hope that the network can learn to solve such registration problems at the tip of the ventral nerve cord with experience.

Slide 6



So, the goal of this work is to make the registration process more robust, to solve the problems that Larvalign could not solve, to improve the registration time, which is about 7 minutes on a 64 GB GPU machine, and most importantly to try to include the landmark points as auxiliary information in the training to assist the network.
And as mentioned on the slide, at the end of the thesis, we hope to have larvalign 2.0 which is faster and more robust.

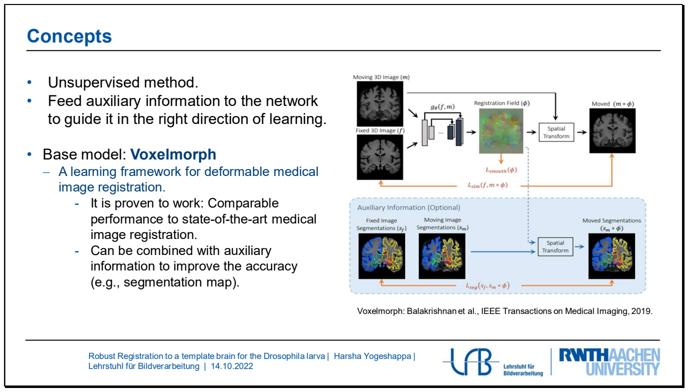
Slide 7



We have received about 1000 3D scans from the University of Leipzig and the Janelia Research Campus, and each of these images varies in resolution. However, for development purposes, the images are scaled to 256x512x64.
This is the largest resolution that will fit on a machine with a GPU capacity between 20GB and 30GB, even with a batch size of 1. Any larger resolution would cause OOM errors. However, the problem of the batch size being limited to 1 is mitigated by the gradient accumulation technique.

The third dataset we see here is a subset of the dataset that was used in the larvalign work. This will now be used as a test dataset to evaluate the performance of our network against larvalign.

Slide 8



We chose to use an unsupervised learning method because it is difficult to find a ground truth for all training images – be it a registered image or a deformation field that maps the transformation from a fixed image to a moving image.

Any information that could help steer the network in the right direction of learning is beneficial. And we would like to be able to provide the network with such additional information that could help us train it better.

We found that the model that meets these two requirements is Voxelmorph model. It has proven to be functional and offers the possibility to introduce additional information. In the voxelmorph paper, segmentation maps were used as auxiliary information, and it was found that this improved the accuracy of the registration.

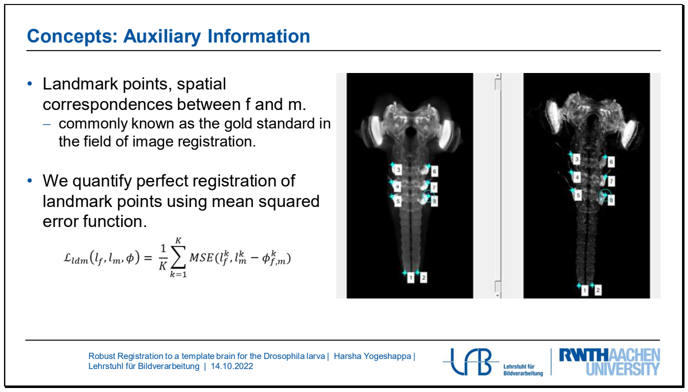
To briefly explain the Voxelmorph model:

The two images - one fixed and one moving - are concatenated at the input and fed into the network, which has an encoder-decoder UNet architecture with skip connections. The output of the network is a deformation field with three channels and in the size of the input images.

To calculate the unsupervised loss, the moving image is warped by the predicted deformation field using a spatial transformer network, and a similarity measurement is made between the warped image and the fixed image.

The supervised loss is the dice score between the ground truth segmentation map and the warped segmentation map.

Slide 9



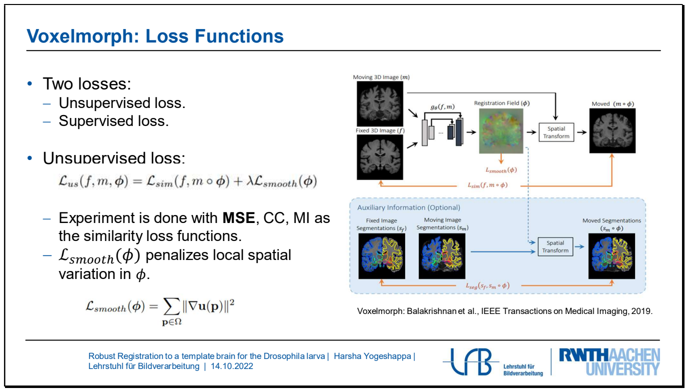
Just as segmentation maps were used in the Voxelmorph paper, we want to introduce our own auxiliary information, and those are landmarks.
So what exactly are landmarks?

A landmark is a spatial correspondence between f and m marked in the same place in both images, as shown in the figure.
In the field of image registration, these spatial correspondences are called gold standards.

We term perfect registration between landmarks when the distance between the landmarks in the registered image and in the fixed image is zero.

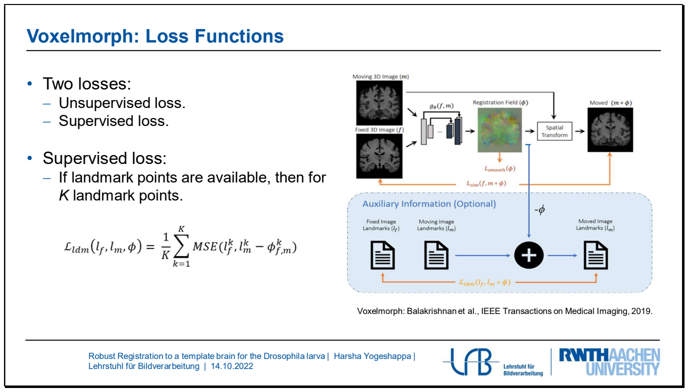
The following mathematical equation describes the same for K number of landmarks where phi represents the displacement or deformation field. This loss shall be termed as Landmark Registration Error.

Slide 10



As explained in the previous slide, unsupervised loss is nothing more than a measure of similarity between the fixed and registered images. The similarity measure can be a mean square error, a normalized cross correlation, or mutual information. However, it was empirically found that for our case the mean squared error performs better.

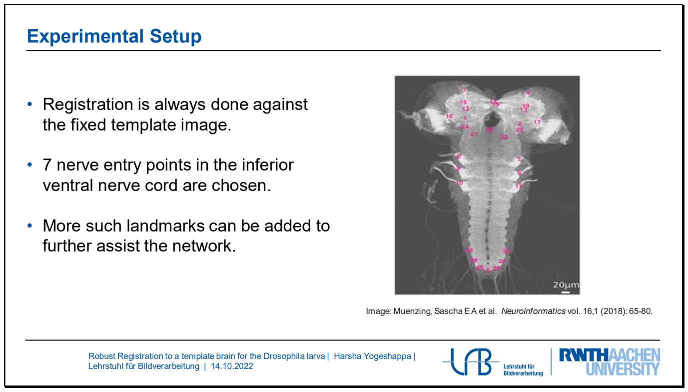
Slide 11



Each time a deformation field is predicted between the two images, in addition to the similarity loss, the landmark registration error is calculated and the total loss is updated.

The deformation field is nothing more than a displacement vector field that specifies by how many units a given pixel in the moving image must move in 3D space. The landmarks in the fixed and moving image are thus related to the displacement field by a simple subtraction operation.

Slide 12



The first and foremost: we always keep the fixed image constant, i.e. the registration is always done against the same image.

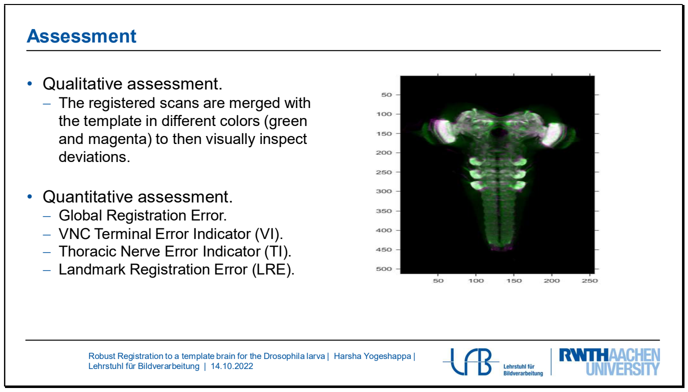
As for the selection of landmarks, the Larvalign paper suggests 30 such possible points, as shown here in the figure.

For training, not all of these 30 landmarks need to be provided as auxiliary information, but only a subset of them can be used (even 0). However, for the evaluation, the error in registering all 30 landmarks will be taken into account.

As shown in the first slides, we know that larvalign had problems registering the lower tip of the ventral nerve cord. For this reason, we decided to annotate only these 7 lower-end nerve inputs in the training examples where we thought the network might need help.

Adding more landmarks should not hurt either.

Slide 13



We plan to do two types of assessments. One is a qualitative assessment, where we overlay the registered image with the fixed image and look for areas where the registration failed. The second is quantitative, where we derive some numbers from the registered image in correspondence with the fixed image.

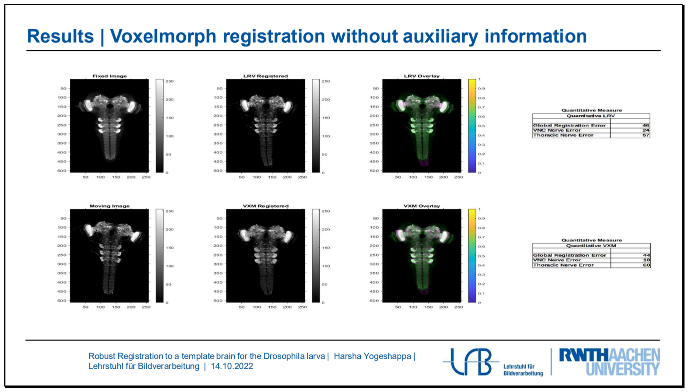
The global registration error is the correlation between the fixed image and the registered image as a whole. It is obvious that such a global measure may not be able to capture local errors. Therefore, we define two additional error values, the VNC error indicator and Thoracic error indicator, as proposed in larvalign paper.

VNC: On spherical regions with a radius of 10 µm at two terminal positions of the VNC landmarks defined in the template image, a mutual information score is calculated between the fixed image and the registered image.

Thoracic: Similarly, for spherical regions with a radius of 15 µm at all six entry points of the thoracic nerve defined in the template image, a mutual information value between the fixed image and the registered image is calculated.

LRE: Finally, LRE is the average of all eucledian distances between the 30 landmarks in the registered image and the fixed image.

Slide 14



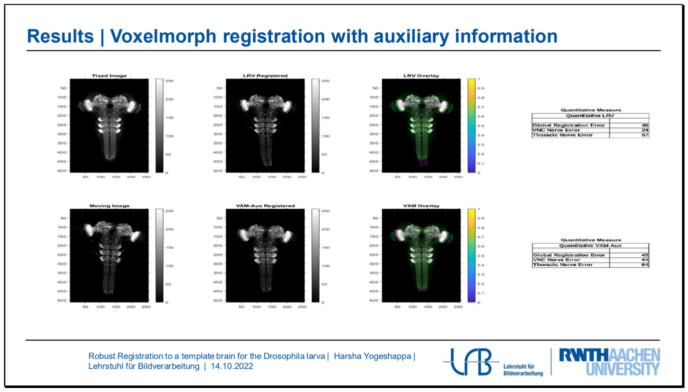
In the current slide and in the next slide, I want to show that a voxel morph network leads to improved registration output with the help of landmark points.

Here,
you can see in the first column the fixed image and the moving image.
In the second column, you can see the registered results with larvalign and voxelmorph without auxiliary information.
In the third column, you can see the respective overlay.
In the fourth column you see the quantitative results.

In this scenario of training without landmark information, voxelmorph actually performed almost as well as larvalign - both the visual inspection and the numbers are in agreement with it.

On the next slide, I show the result obtained with a network trained with landmark information.

Slide 15



As we can see, there is a significant correction at the VNC tip and mutual information scores also shows an improved result.

Slide 16

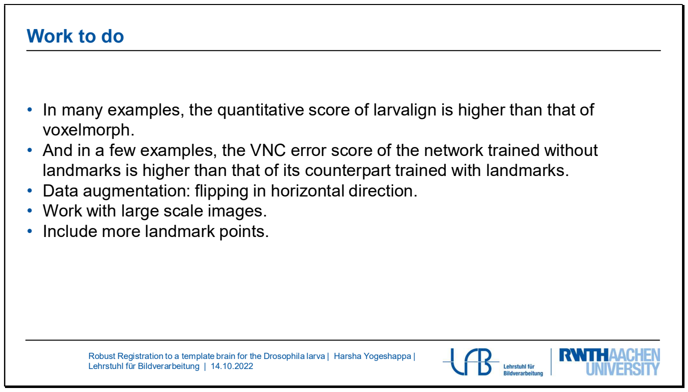
Results

Generalizability

- To evaluate the robustness of the network, the following test was performed.
  - Experimental configuration_1:
    - Train on larvalign dataset
    - Test on larvalgin dataset
  - Experimental configuration_2:
    - Train on janelia_dataset.
    - Test on larvalign_dataset

- The qualitative and quantitative assessment of configuration_1 is comparable with configuration_2 in both the respective scenarios of with and without auxiliary information.

Slide 17

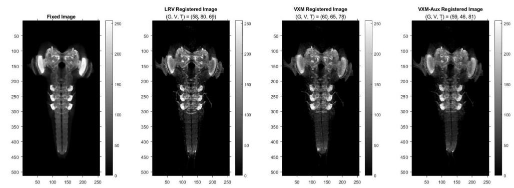


The next slide shows the plot of such an example, where the VNC error measure has a lower score in the network trained with landmark information. <That slide is hidden and will be shown only if needed>.

G: Global error
V: VNC error
T: Thoracic error

Slide 18

Work to do - Appendix



Slide 19

Slide 20

Model Architecture: Voxelmorph | Appendix

Model description

- The fixed image (f) and the moving image (m) is concatenated as two input channels and fed to the Unet.
- Number of features in each layer is mentioned in the figure.
- Output has 3 channels representing the amount of displacement that in each of 3 directions (u,v,w).

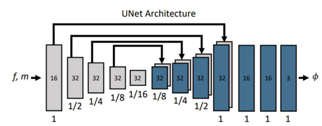


Image: Balakrishnan et al., IEEE Transactions on Medical Imaging, 2019.

Architectural constraint: Keep an eye on anisotropic nature of the image and make sure that smallest volume captures the maximum displacement expected in all 3 directions.

Slide 21

Model Architecture: Voxelmorph | Appendix

Unet architecture

- 3D Convolutions (Kernel=3, Stride=1, and ‚SAME‘ padding) with LeakyReLU activation function.
- Explicit MaxPooling to downsize the images in the encoder layer.
- Upsampling, 3D Convolution, and concatenating skip connections in the decoder layer.

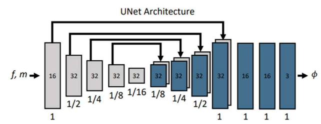
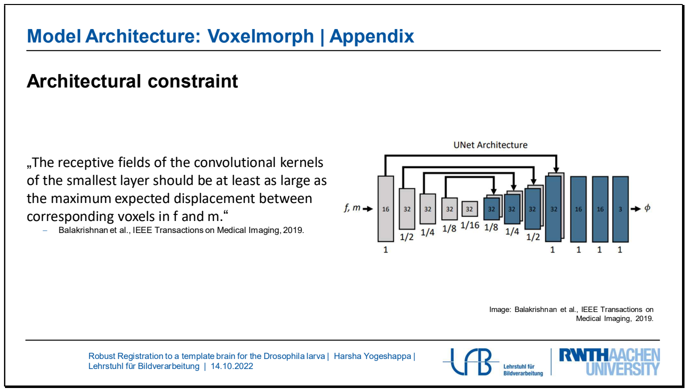


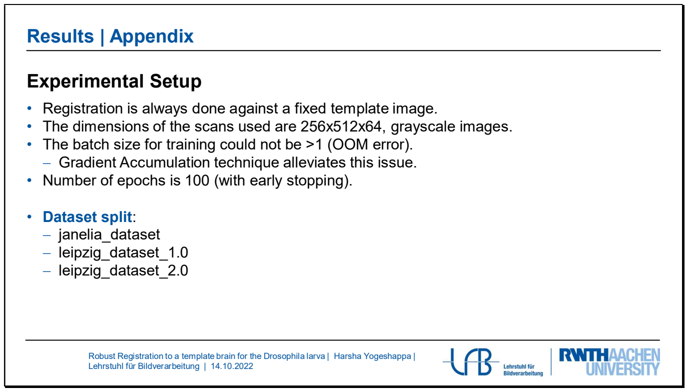
Image: Balakrishnan et al., IEEE Transactions on Medical Imaging, 2019.

Slide 22



Therefore, the pooling in z-direction is controlled.
- Give proof in the next slides.

Slide 23



leipzig_dataset_1.0 is the subset of data that larvalign was tested.
leipzig_dataset_2.0 is the new data recorded in the same laboratory but with a different microscope and registration is not good i.e., not robust.