## Comparative analysis of image datasets for hand development and regeneration in salamanders.

The analysis of salamander hand development and regeneration image dataset involves several steps. Firstly, image reading and preprocessing are performed, where the scikit-image Python package is used with the transform.resize function to adjust the size of images to a uniform size of (128, 128). Secondly, Resnet-18 is used for image embedding and features are extracted, followed by dimensionality reduction using UMAP method. The resnet(pretrained=True) function in the torchvision package is used to obtain the pre-trained Resnet-18 model on the imagenet dataset, and the umap-learn Python package is used for UMAP method.

To compare the differences between the development and regeneration image datasets, the mean pixel values of the two datasets are calculated, and the different areas between the two average images are visualized by subtracting one from the other. In addition, to quantitatively compare the variation in pixel intensity between the two datasets, per-pixel standard deviation is calculated, and the Mann-Whitney U test is performed to test for significant differences.

All the codes involved in the above analysis can be found here: <https://github.com/Nanguage/Regeneration>.

## Fig legend

**a.** Scatter plot of the dataset after Resnet embedding and UMAP dimensionality reduction, where each spot represents an image of salamander hand. The closer two points are, the more similar their features. It can be clearly seen that there are significant differences between development and regeneration images. The numbers annotated in the plot represent the image numbers in the sample images. **b.** Some sample images from the two sets of images (development and regeneration) can be displayed, which make it easier to observe the significant differences between them. **c.** The image on the left and right in the above picture show the average images of the salamander hand image dataset with normal development and tissue regeneration, respectively. It can be seen from the images that the tissues in normal development are more distinguishable between fingers, and overall changes are smaller. **d.** The difference image between the average images of development and regeneration shows the specific regions of the regeneration image in red (value > 0), while the specific regions of the developmental image are in blue (value < 0). **e.** The standard deviation images of the development and regeneration datasets show that the areas with darker colors have greater variation between the images. It can be observed that the average pixel standard deviation in the developmental images (26.79) is significantly smaller than that in the regeneration images (43.59). **f.** Violin plot of the pixel differences in standard deviation images between development and regeneration images shows that the pixel std in regeneration images is higher overall. The difference is extremely significant according to Mann-Whitney U test (p-value = 0.0).