Different from matching function, registration is transforming different set of data into one coordinate system. In medical image processing elastic registration need to respond to the object deformation. The deformable registration allowed an un-uniform mapping between images. The state of art ideas is to measure and correct small, varying discrepancies by deforming one image to match the other.

Scale-invariant feature transform (SIFT) is a traditional algorithm which detect and describe local features (interest points) in images. SIFT algorithm can implement matching task based on best-bin-first search algorithm which deal with the nearest neighbor search problem in high dimensional spaces. And filter out the useful pairs based on Euclidean distance.

Deep learning techniques is widely implemented in computer vision task. Many researchers implement deep learning techniques on image registration tasks.

Due to large image size, small datasets and histological image content problems, this project combined Scale-invariant feature transform (SIFT) features detection and deep learning techniques to retrieval paired landmarks from images.

(The challenge of the registration of the medical images is that the body might change posture even skeletal arrangement can change)

There are several methods can detect local interest points like SIFT, SURF, ORB and so on. This registration method combined the performance of the SIFT, SURF and ORB local interest point detection method. The SURF and ORB method is faster than SIFT. However, the quality and quantity of the detected points is not as good as the SIFT detector. (need a form and some picture)

The registration landmarks are retrieved based on paired images. After detected interest points from each image. We expand the landmarks to 16 x 16 image patches. And use these patches to find the real matching landmarks. The basic idea is to make use of the threshold based on k-Nearest Neighbors search and Euclidean distance to filter out feasible paired landmarks. But the performance is not good on the histological images. This registration method filtering is based on the ranking of the output of the convolution neural network (CNN). The CNN is used for determining the pair of the small patches score. The score is added in ranking. The k top of the pairs is the retrieved landmarks of each images.

We also implement a method that directly used CNN to get the landmarks. But the result implemented on this dataset is not good. The dataset including only 230 pairs of images. After training, the network is either underfitting or overfitting. And the landmarks points including negative number which is useless. To solve this problem, we get some ideas on the generative adversarial network (GAN). Which use a discrimination network to determine the generated image is fake or true. We added a discrimination network in training process. However, due to the large amount of the generated landmarks including negative number and the small dataset and the loss function is changed to not based on the error distance. The results are worse than the CNN method.

The main challenge of the non-rigid registration of the histological images is required to sensitive to the human tissue deformation and retrieval useful landmarks. Due to the big histological images size, some traditional methods are time consuming. And the textures in the images not homogeneous, so when make use of the SIFT or other method might get huge amount of the points on the texture area instead of the area required.

The device is also a considerable problem for this challenge. Due to the high quality of the image, the image size become a big issue when processing the image. Limited require of this method is the memory of running device is at least 8gb to ensure the image can be loaded and processed.

To solve the problem, this method will scale each image 50 times before using SIFT to detect local interest point from each image. And focus on the surrounding pixels of each landmark, instead of processing the whole images.

Others:

The deep learning method deal with the registration problem often generate the landmark directly from both images. However, when

Evaluation:

The performance is evaluated with relative Target Registration Error (rTRE). The evaluation method is also suggested to measure the performance in the registration task. This method measures the error between retrieved landmarks and the ground truth landmarks.

Target Registration Error is calculated as the Euclidean distance between the retrieved landmarks coordinates and the ground truth landmarks coordinates. And the Euclidean distance need be normalized by the image diagonal rTRE = TRE/(w^2 + h^2)^根号.

The Histology (CIMA) dataset is offered by Phd.Jiří Borovec. The images and landmarks data are free to use in adapted and share for any purpose. This dataset will be only used in this research and never be used in commercial.

State of art function

1. Optical Flow Assumptions (This alone is not a sufficient constraint!)

Challenge : Ambiguity within homogenous regions

1. Transient Quadratic (TQ) Approach瞬态二次（TQ）方法
   1. Enables better-constrained large deformations
   2. Uses Lagrangian regularization for specified time interval, followed by a re-gridding strategy使用拉格朗日正则化指定的时间间隔，然后是重新网格化策略
      1. After an interval’s deformation reaches a threshold, we begin a new interval for which the last deformation becomes the new starting point在间隔变形达到阈值后，我们开始一个新的间隔，最后一个变形成为新的起点
      2. TQ thus resets the coordinate system while permanently storing the past state of the algorithm因此，TQ重置坐标系，同时永久存储算法的过去状态
   3. Results in a hybrid E+L physical model, resembling soft, stretchable plastic
      1. Maintains the elastic regularization for a given time then takes on a new shape until new stresses are applied在给定时间内保持弹性正则化然后呈现新形状直到施加新应力

早期医学图像配准大部分是记录以不同形式获得的同一受试者的脑图像

Registration based on patient image content can be divided into geometric approaches and intensity approaches. 基于患者图像内容的登记可以分为几何方法和强度方法。

The most well-known example is the rigid or affine transformation

Using Random sample consensus (RANSAC) deal with the outliner

Training:

Expand each corresponding landmark point to 16 x 16 region pair with surrounded pixels. Then put this pairs of regions into the network and the expectation output is true (1) and train with un-matching points with expectation output is false (0).

Registration processing:

This registration method combined local interest points detection and global pair points matching discriminate methods. This method will using SIFT to detection local interest point from each image separately. And using the Brute-Force matcher to get the pairs of points. Then expand each point to 16 x 16 small image regions based on the around pixels. The pair sets of regions use the trained CNN model to discriminate each pair points is matching. Then return a pair of landmarks.

Train algorithms

For each pair images in training images set

Using the ground truth pair landmarks and expand the points to 16 x 16 small images

Put pairs of matching small images into the network the expectation output is true (1)

Put pairs of unmatching small images into the network the expectation output is false (0)

Calculate the loss and update the network parameters

Registration algorithms

For each pair images need registration

Using scale-invariant feature transform (SIFT) to detect features in each image

Expand the points to 16 x 16 small images

Put the images to the trained network and record each score

Ranking the score and take the top n pairs

Record the landmarks

Network

The CNN including 3 convolutional layers, 3 pooling layers and 3 fully connected layers. The activation function of each layer expects the last fully connected layer is linear relu. The last layer activation function is Sigmoid because the expectation output is a number between 0 and 1. The output number used for distinguishing the pair of small image region can satisfied the registration landmarks or not. Binary Cross Entropy (BCE loss) between the target and the output loss function is selected to use in training because the output is only one number. The optimizer is Stochastic gradient descent (SGD).

conv\_0 is a convolutional layer that the input channel is 3, output channel is 32, kernel size is 2 x 2, stride is 1 x 1. pool\_0 is a pooling layer that kernel size is 2, stride is 2. conv\_1 is a convolutional layer the input channel 32, output channel is 64, kernel size is 2 x 2, stride is 1 x 1. pool\_1 is same as pool\_0. conv\_2 is a convolutional layer that the input channel is 64, output channel is 128, kernel size 2 x 2, stride is 1 x 1. pool\_4 a pooling layer which is same as pool\_1 and pool\_0. Bottleneck is a linear layer (fully connected layer) that input features is 128, output features is 512. After the Bottle neck layer, the pair images features will added together (512 + 512 = 1024) and send to next layer. fc\_1 is a Linear that input features is all of the paired images features (1024), output features=256. fc\_2 is a Linear layer, the input features is 256, output features is 128. fc\_3 is a Linear layer which input features is 128, output features is 1.