

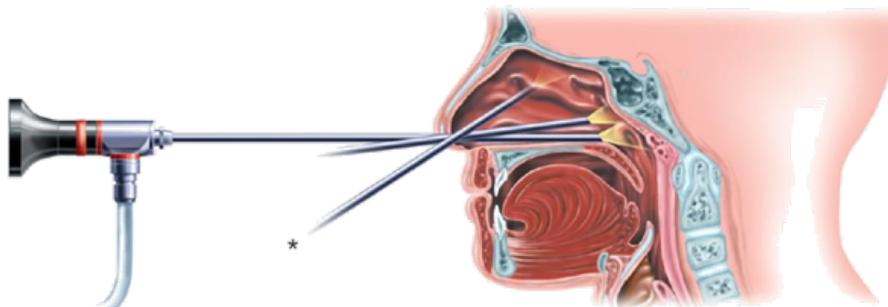
# Enhanced endoscopic navigation using shape statistics in the absence of CT imaging

Ayushi Sinha

Provost's Postdoctoral Fellow, LCSR

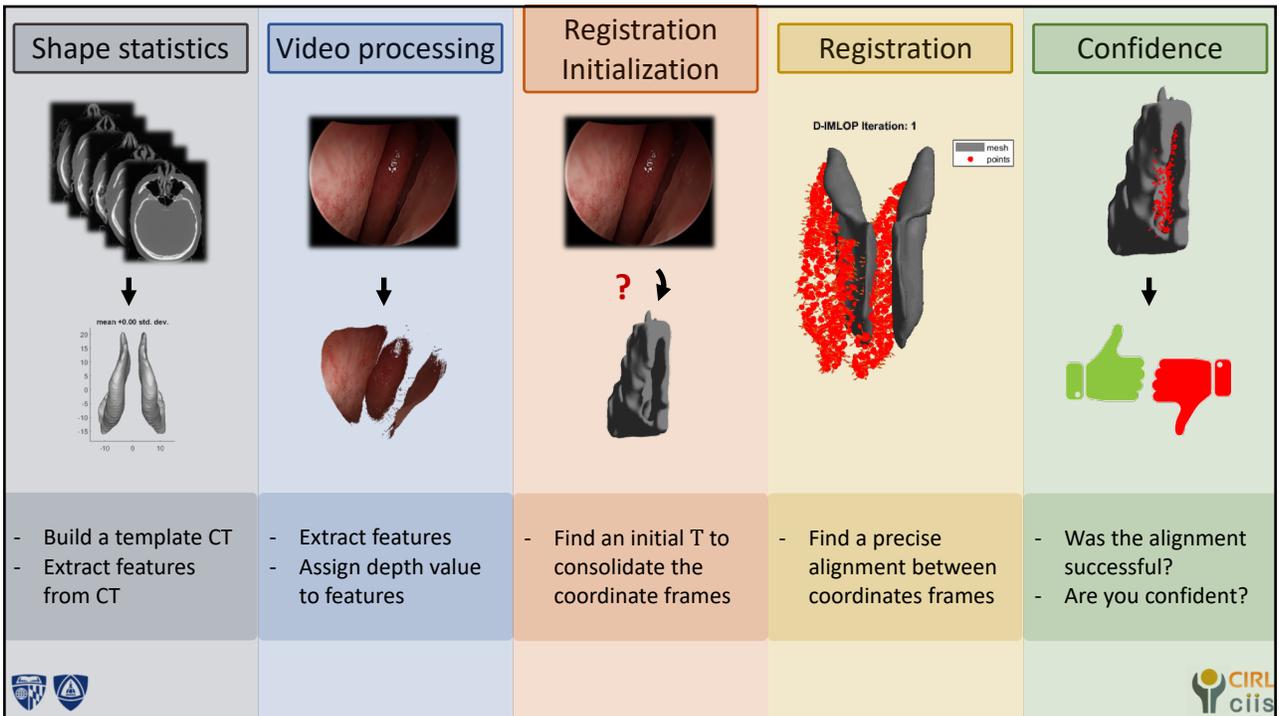
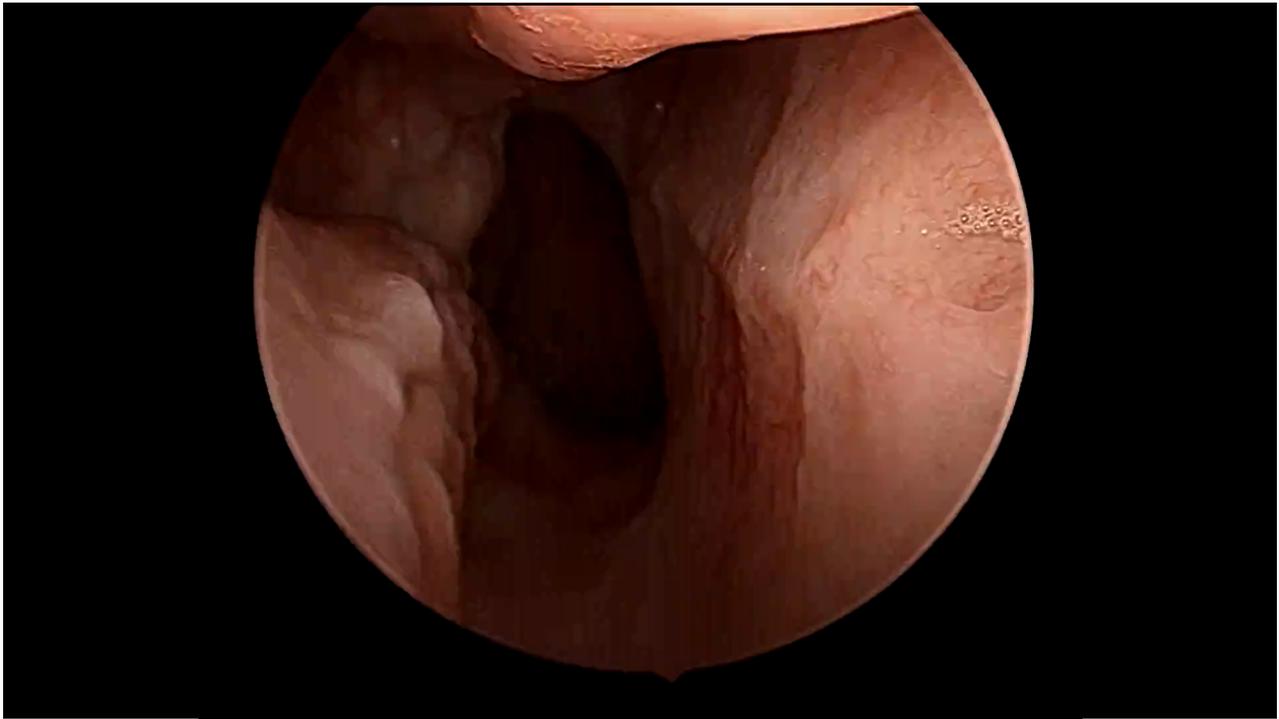


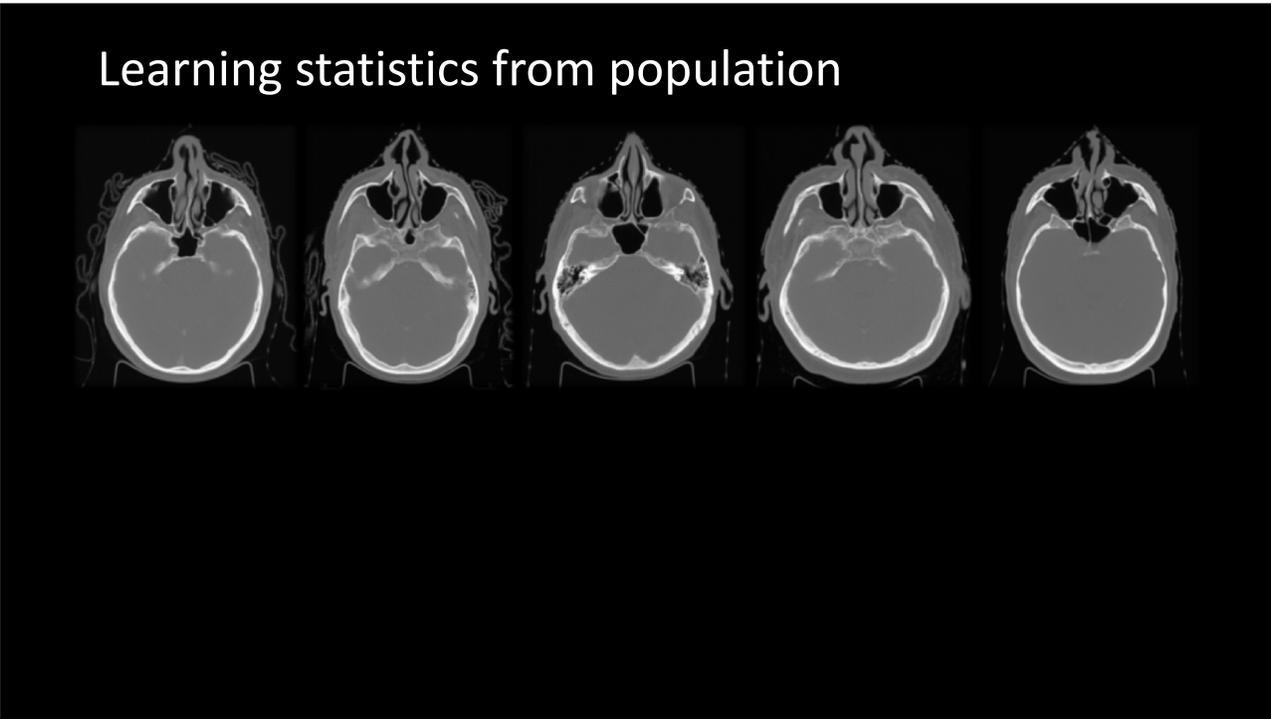
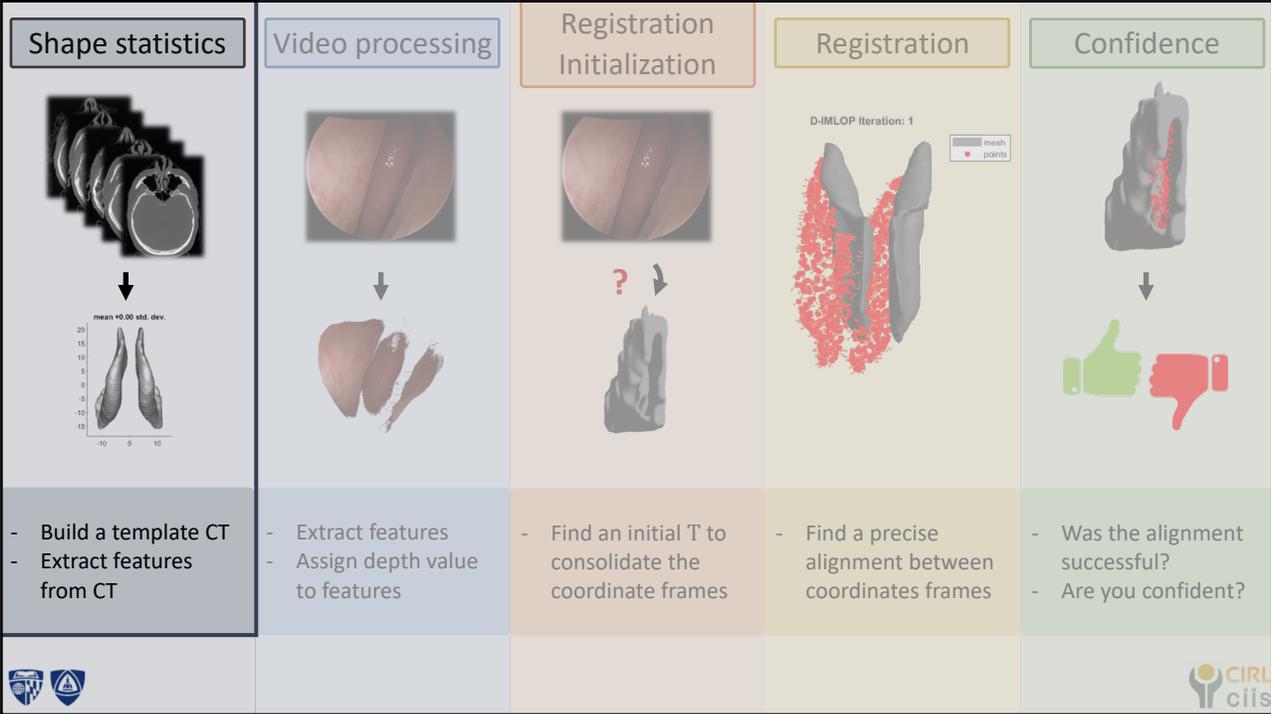
## Minimally invasive surgeries through the nasal cavity



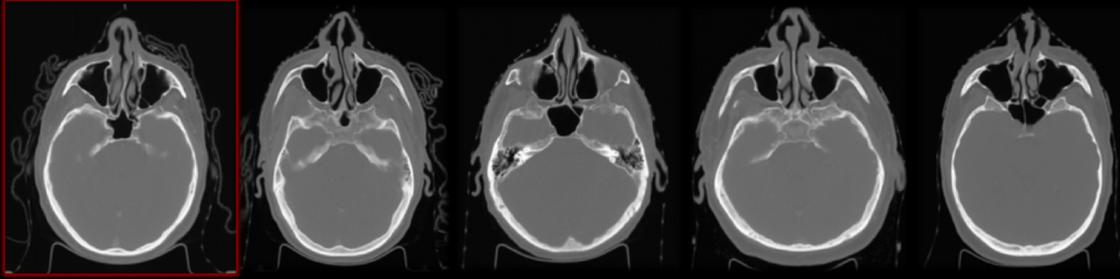
G. Scadding et al., *Diagnostic tools in Rhinology EAACI position paper*, Clinical and Translational Allergy, 1(2), 2011



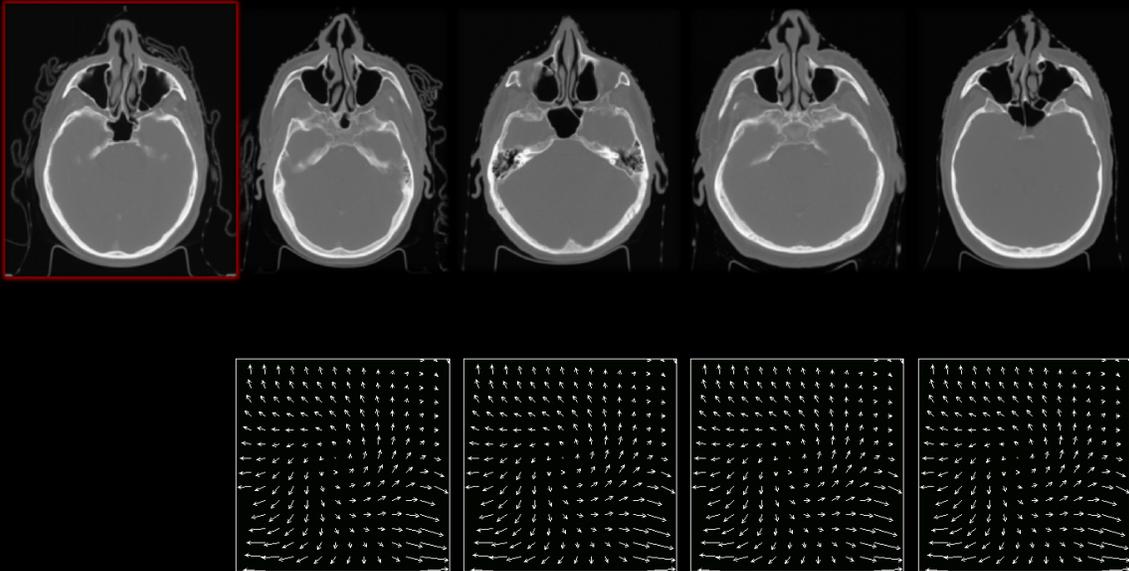


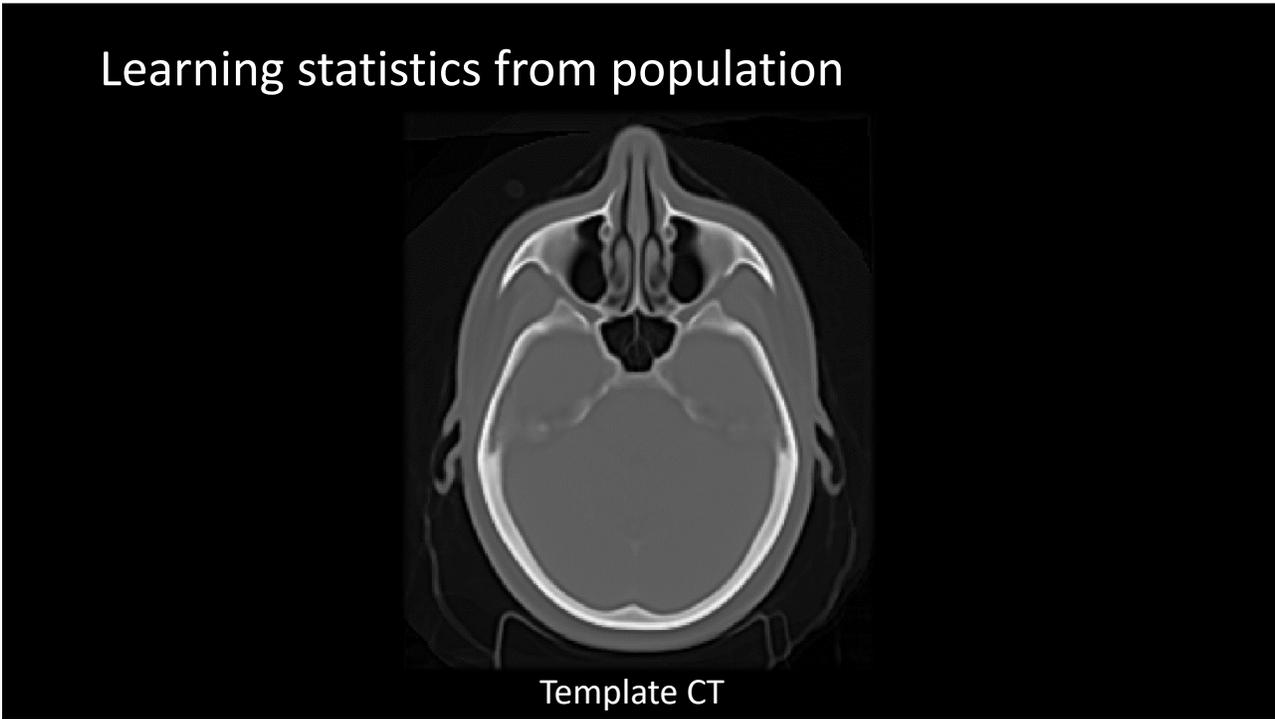
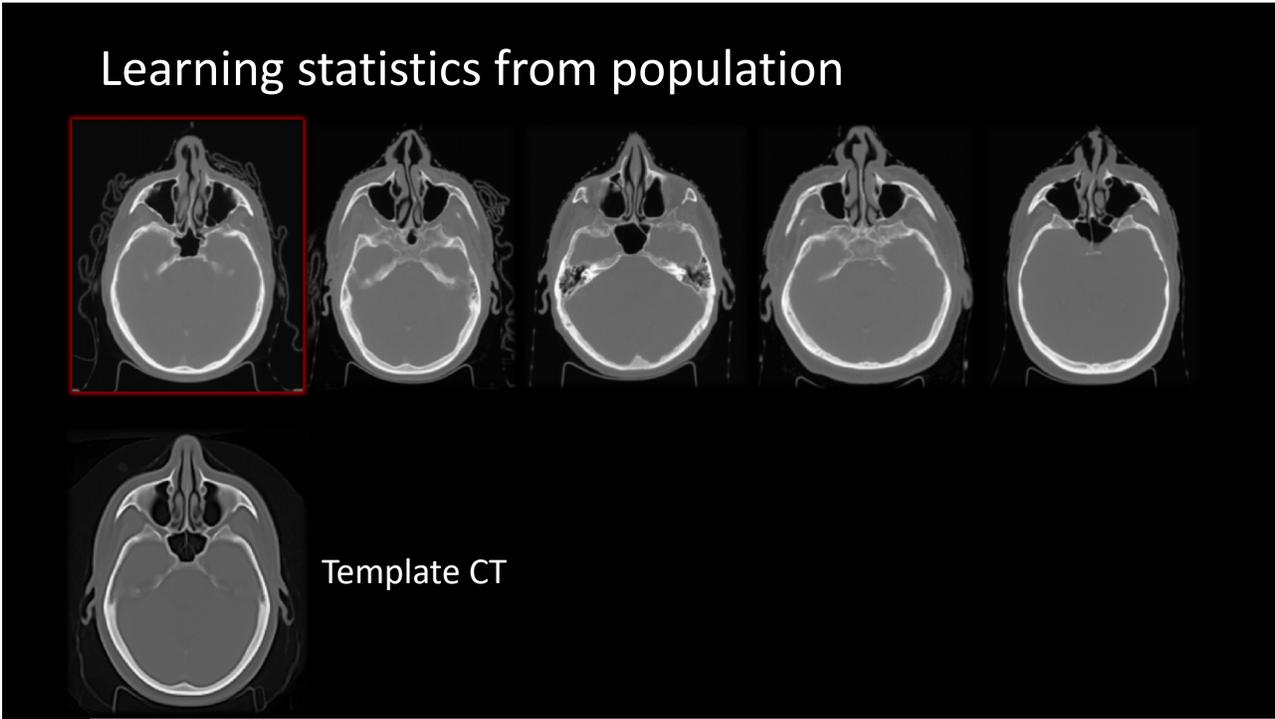


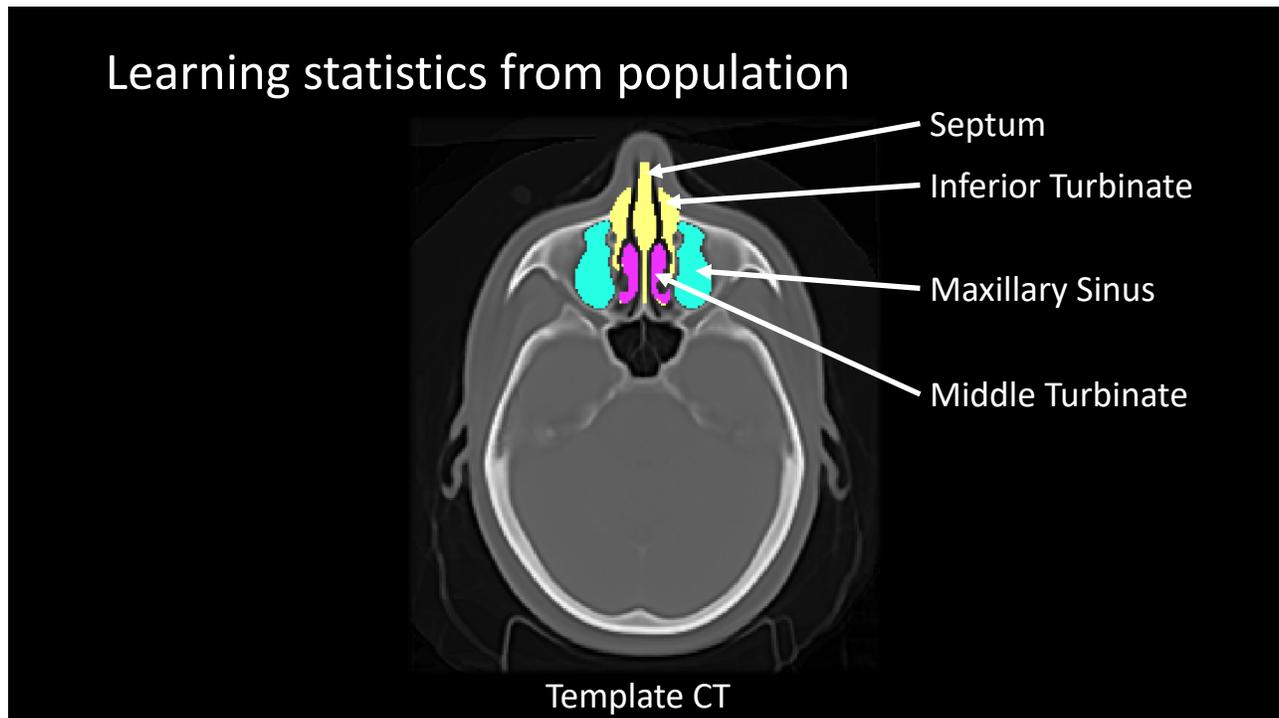
# Learning statistics from population



# Learning statistics from population







## Learning statistics from population

- Given shapes,  $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_{n_s}]^T$ , with correspondences, we can compute the principal components:

- Mean:

$$\bar{\mathbf{V}} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{V}_i$$

- Variance:

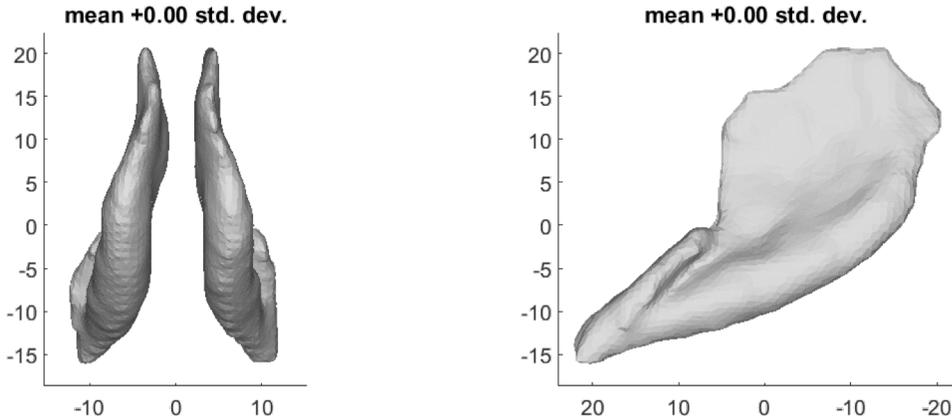
$$\Sigma = \frac{1}{n_s} \sum_{i=1}^{n_s} (\mathbf{V}_i - \bar{\mathbf{V}})(\mathbf{V}_i - \bar{\mathbf{V}})^T$$

$$\Sigma = [\mathbf{m}_1 \ \dots \ \mathbf{m}_{n_s}] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n_s} \end{bmatrix} [\mathbf{m}_1 \ \dots \ \mathbf{m}_{n_s}]^T$$



# Learning statistics from population

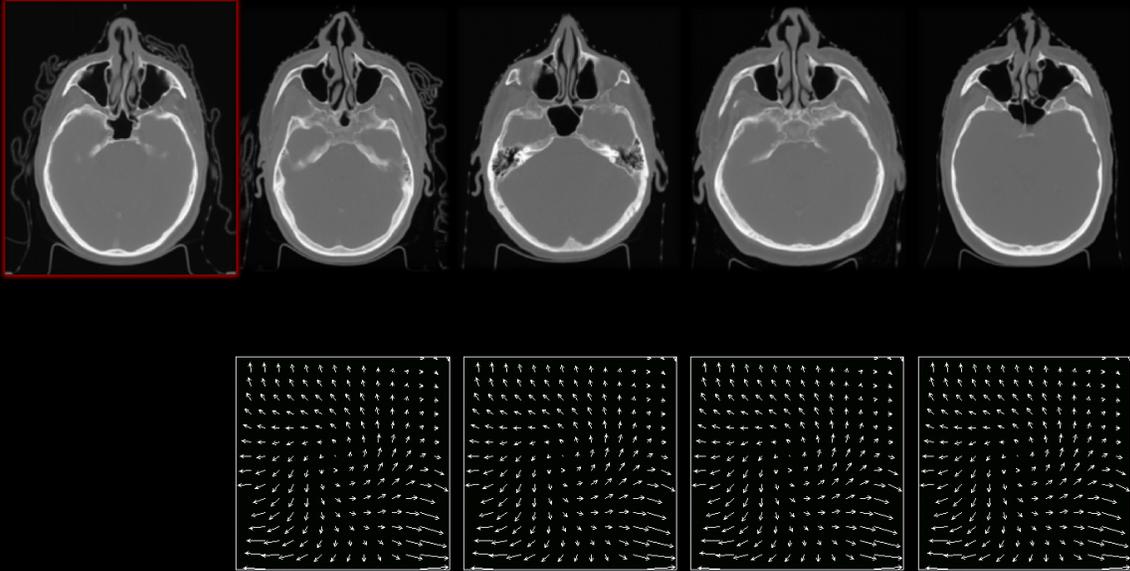
Built from ~50 CT scans

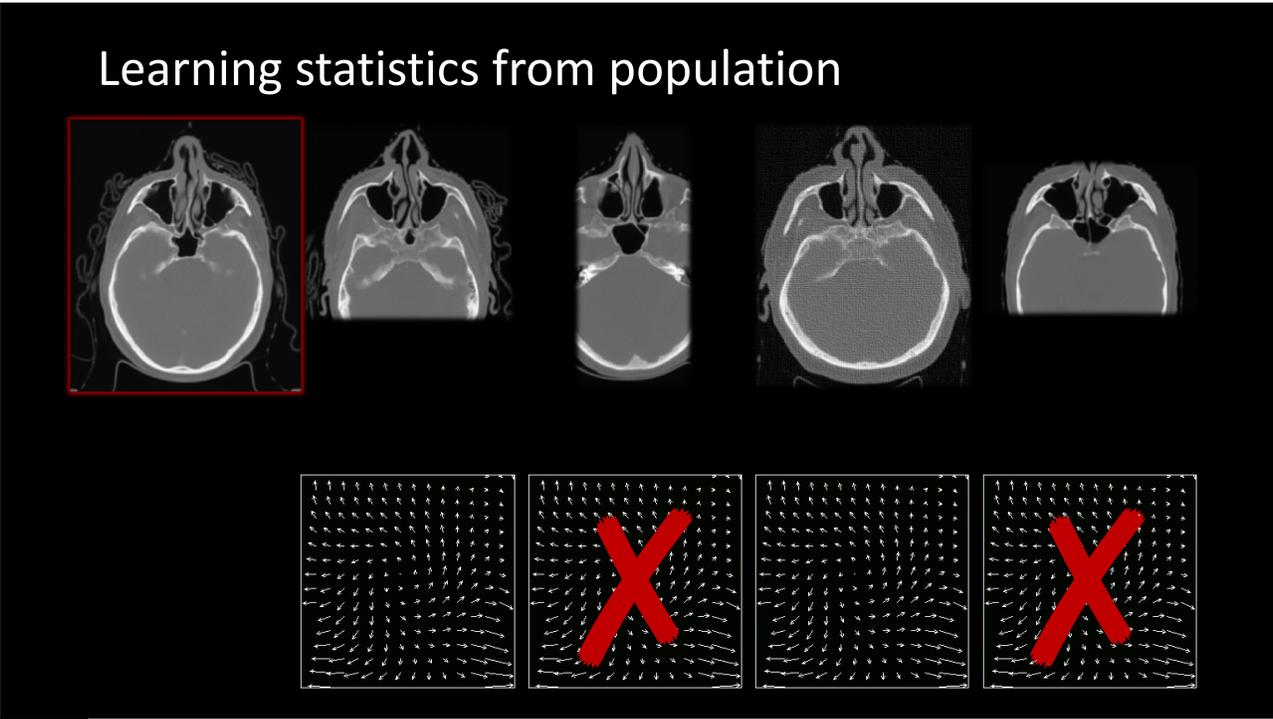


Can we scale up to ~5000 CT scans?



# Learning statistics from population





## Did the registration work?

This CVPR2015 paper is the Open Access version, provided by the Computer Vision Foundation. The authoritative version of this paper is available in IEEE Xplore.

### Learning to Compare Image Patches via Convolutional Neural Networks

Sergey Zagoruyko  
Universite Paris Est, Ecole des Ponts ParisTech  
sergey.zagoruyko@imagine.enpc.fr

Nikos Komodakis  
Universite Paris Est, Ecole des Ponts ParisTech  
nikos.komodakis@enpc.fr

Abstract

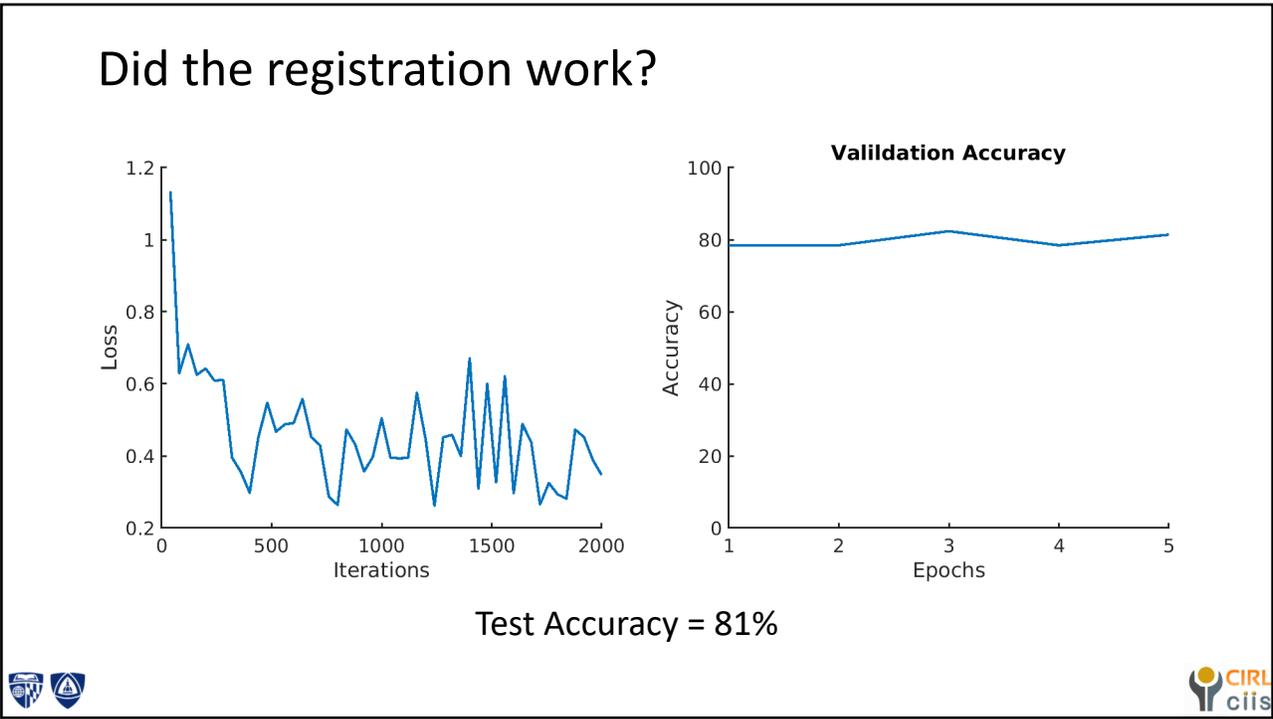
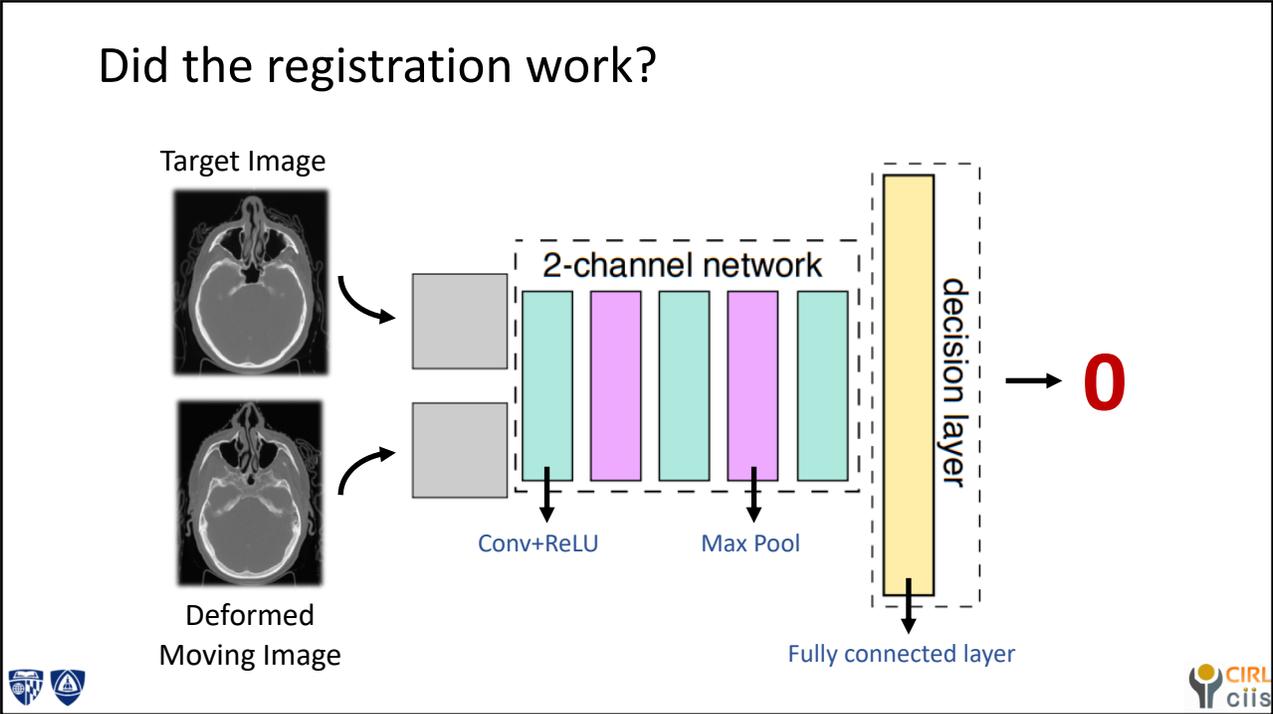
*In this paper we show how to learn directly from image data (i.e., without resorting to manually-designed features) a general similarity function for comparing image patches, which is a task of fundamental importance for many computer vision problems. To encode such a function, we opt for a CNN-based model that is trained to account for a wide variety of changes in image appearance. To that end, we explore and study multiple neural network architectures, which are specifically adapted to this task. We show that such an approach can significantly outperform the state-of-the-art on several problems and benchmark datasets.*

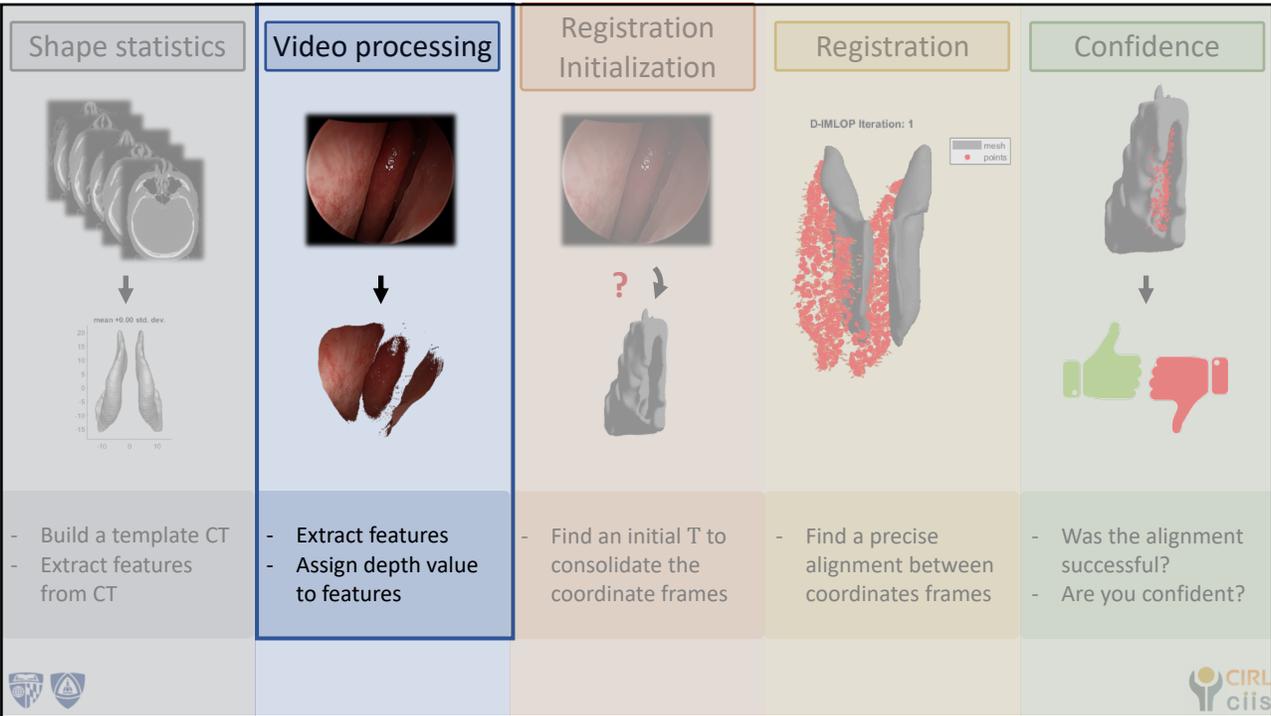
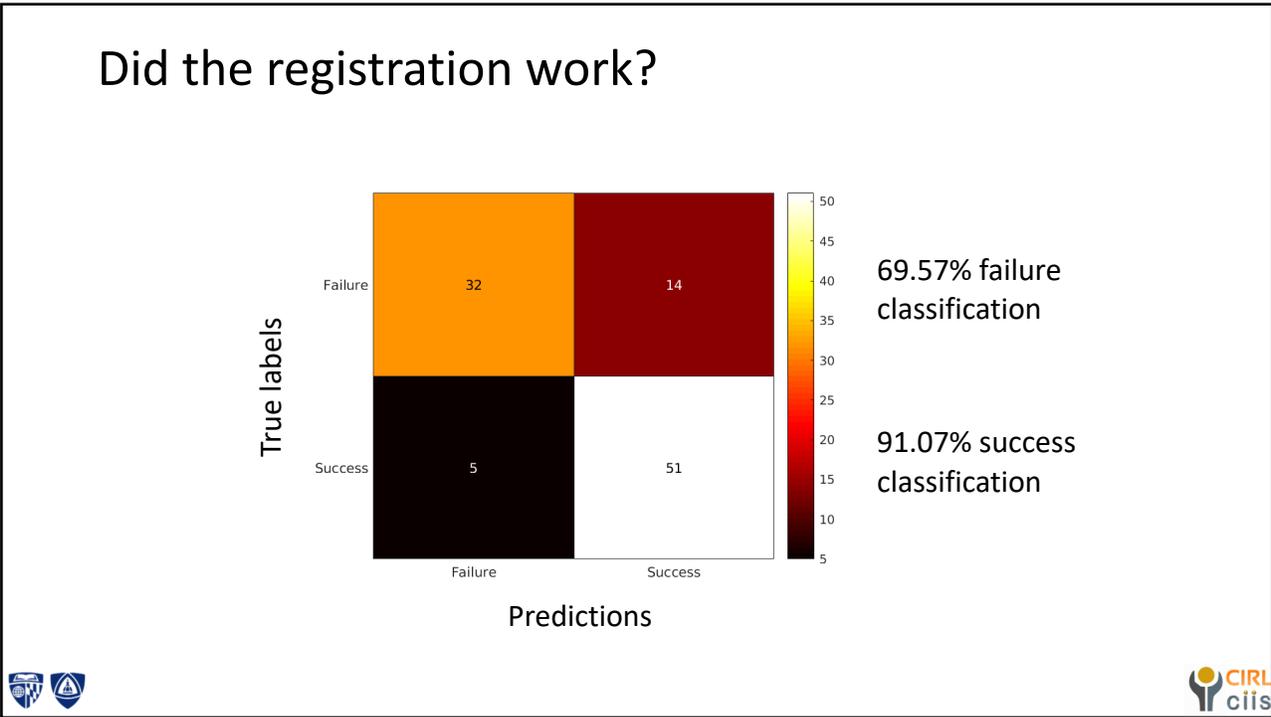
1. Introduction

Comparing patches across images is probably one of the most fundamental tasks in computer vision and image anal-

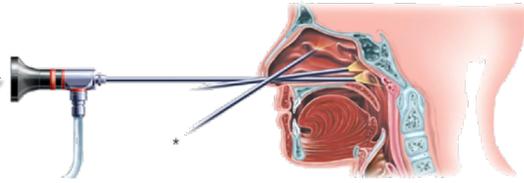
Figure 1. Our goal is to learn a general similarity function for image patches. To encode such a function, here we make use of and explore convolutional neural network architectures.

software) large datasets that contain patch correspondences between images [22]. This begs the following question: can





# Feature extraction from video



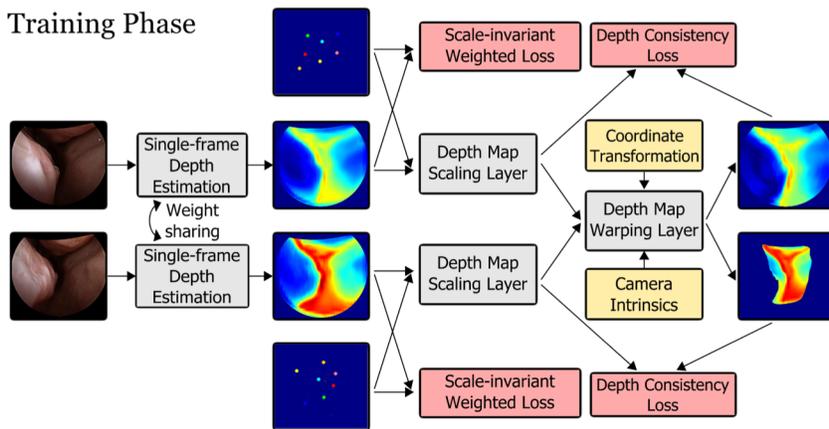
G. Scadding et al., *Diagnostic tools in Rhinology EAACI position paper*, Clinical and Translational Allergy, 1(2), 2011

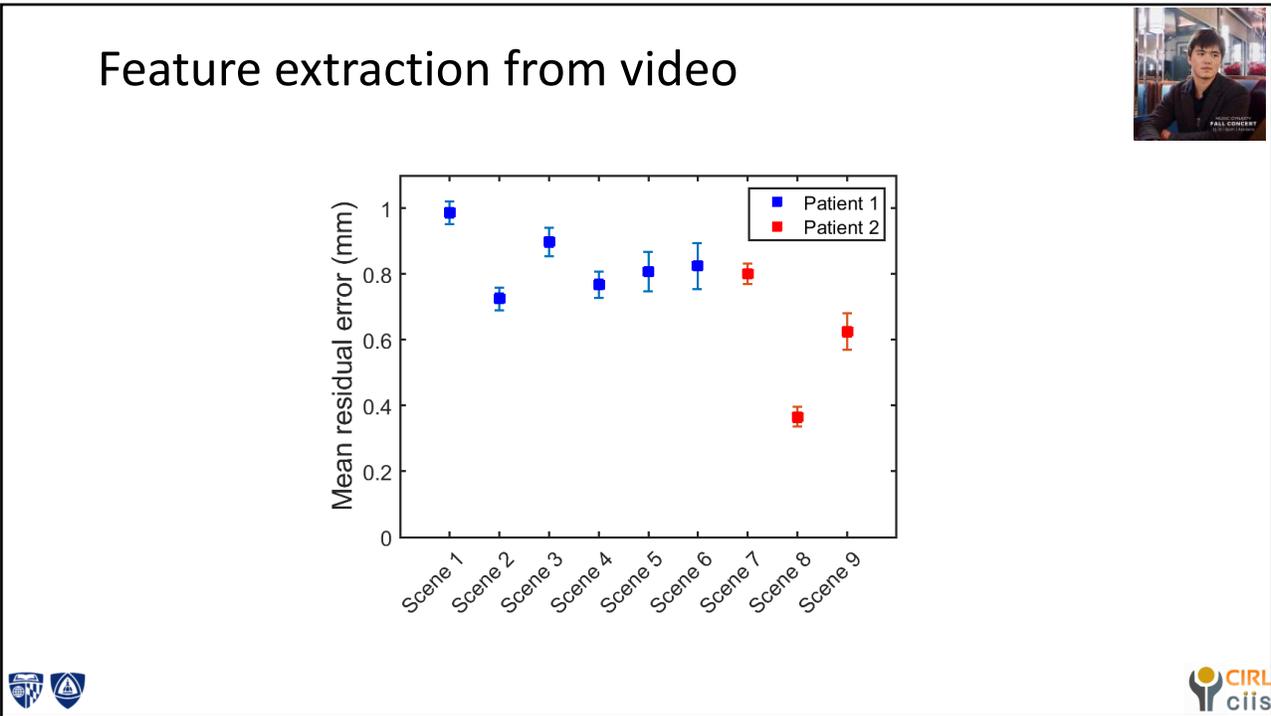
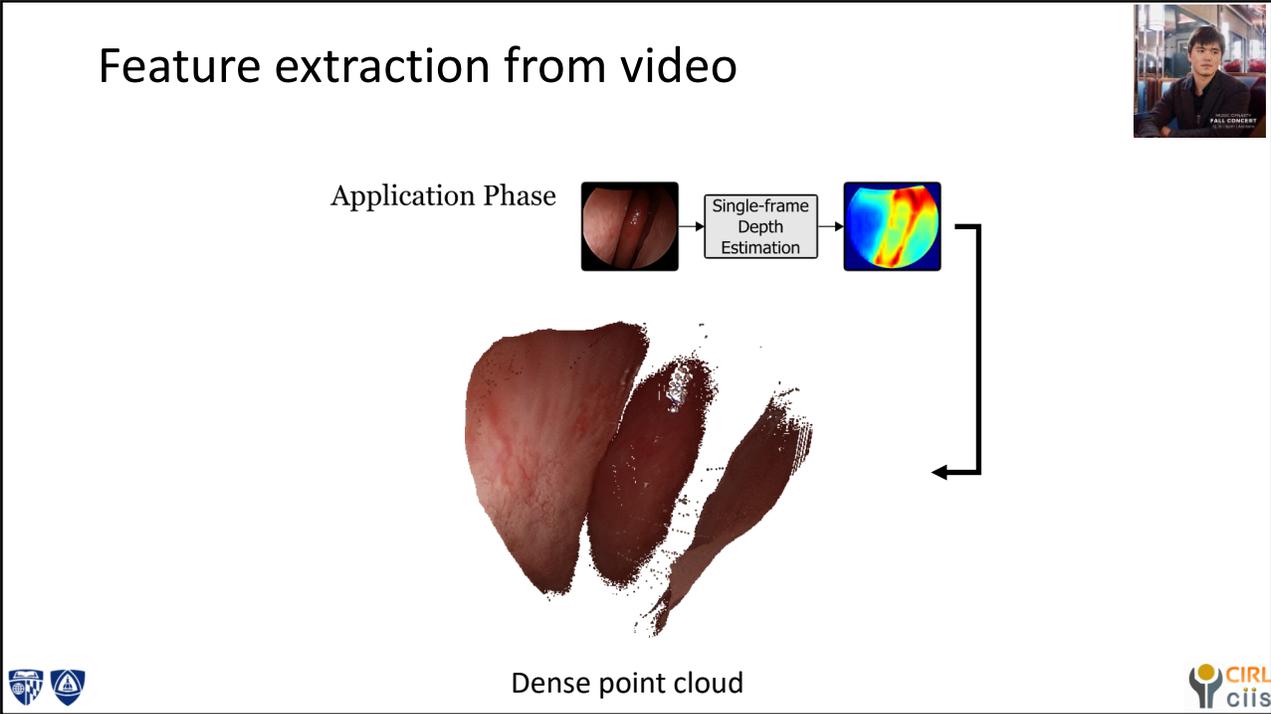


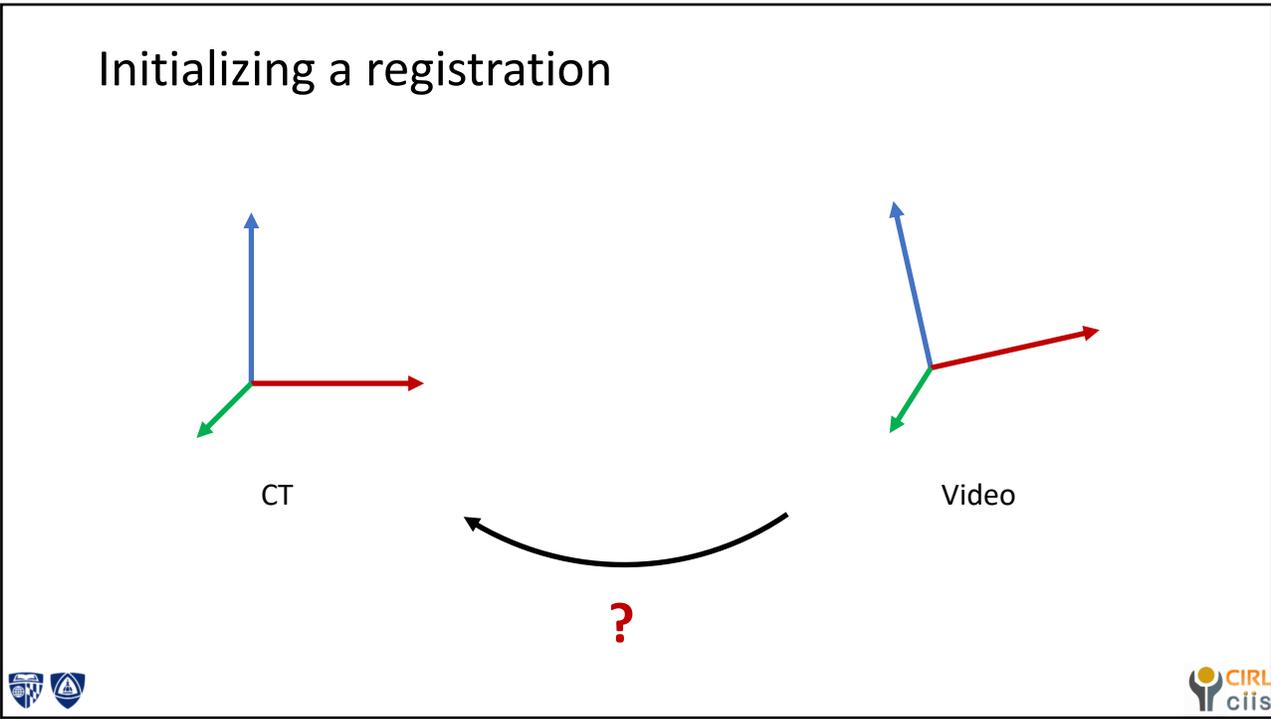
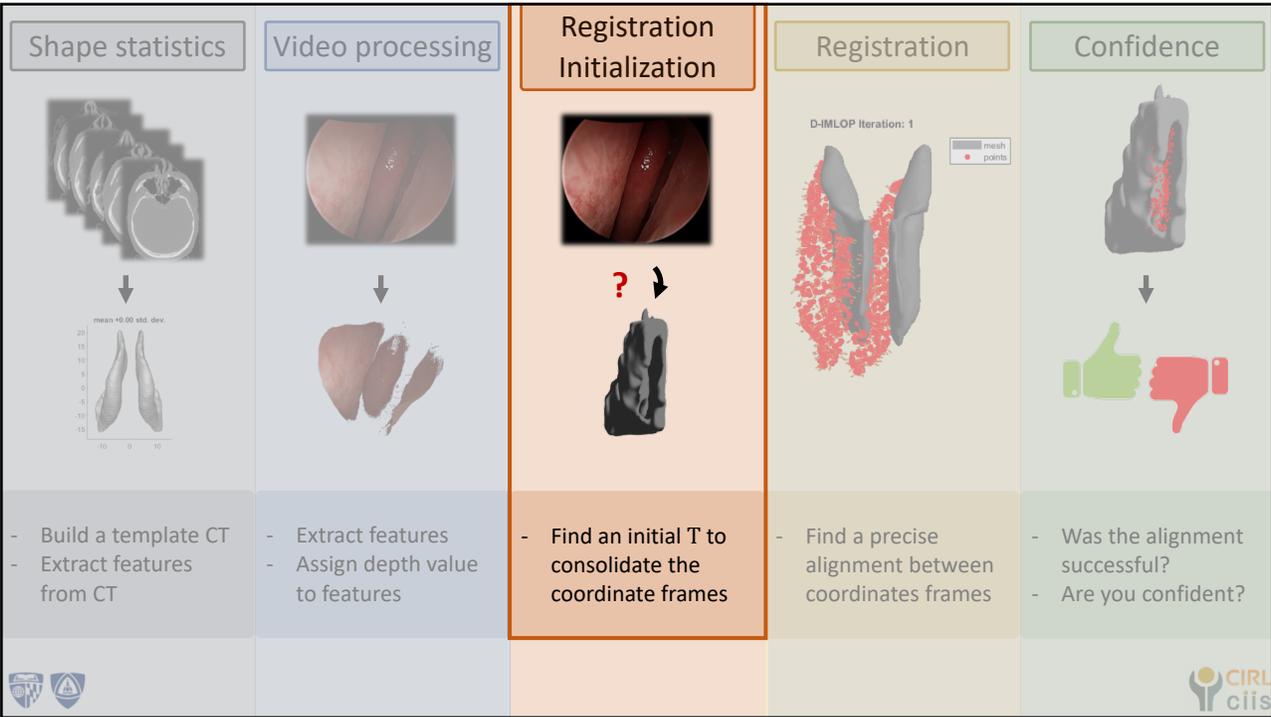
# Feature extraction from video



Training Phase







# Initializing a registration



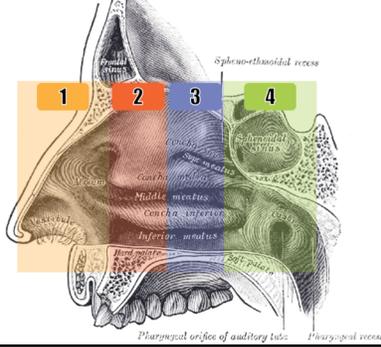
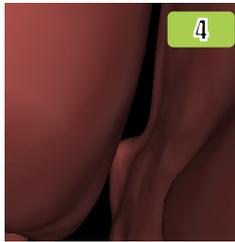
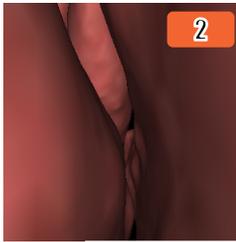
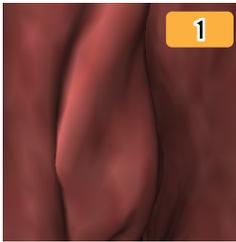
CT

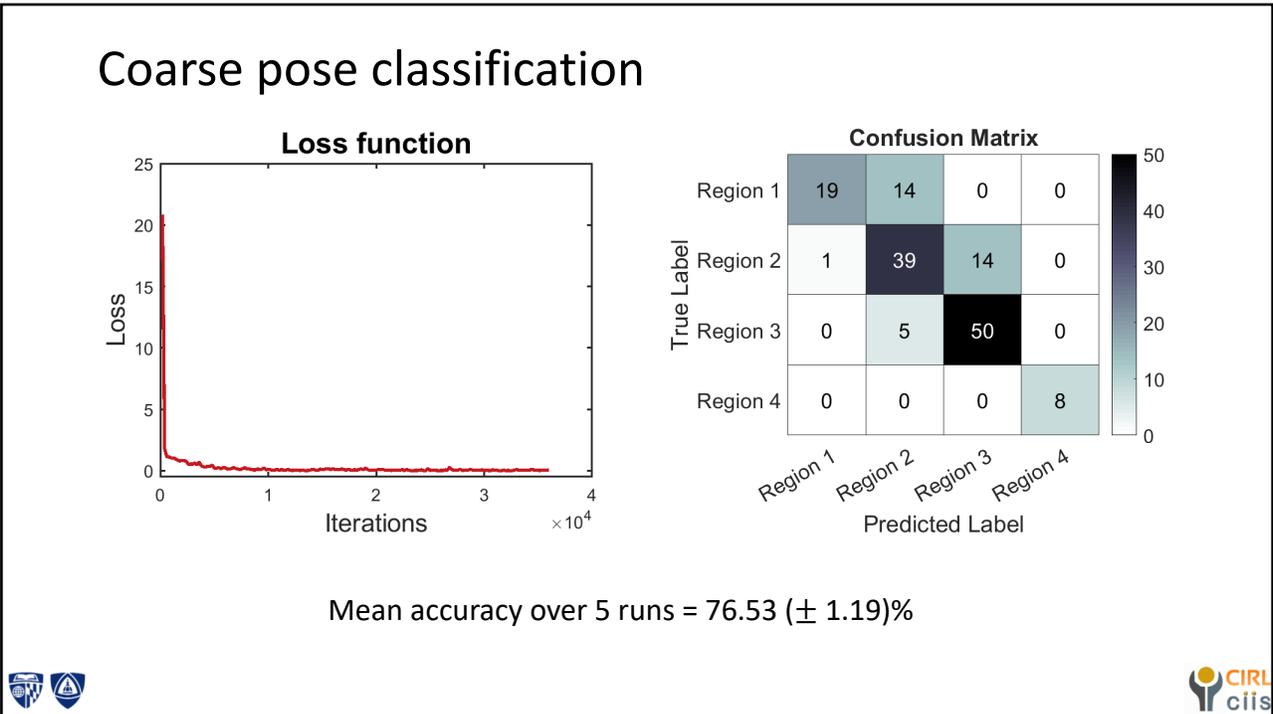
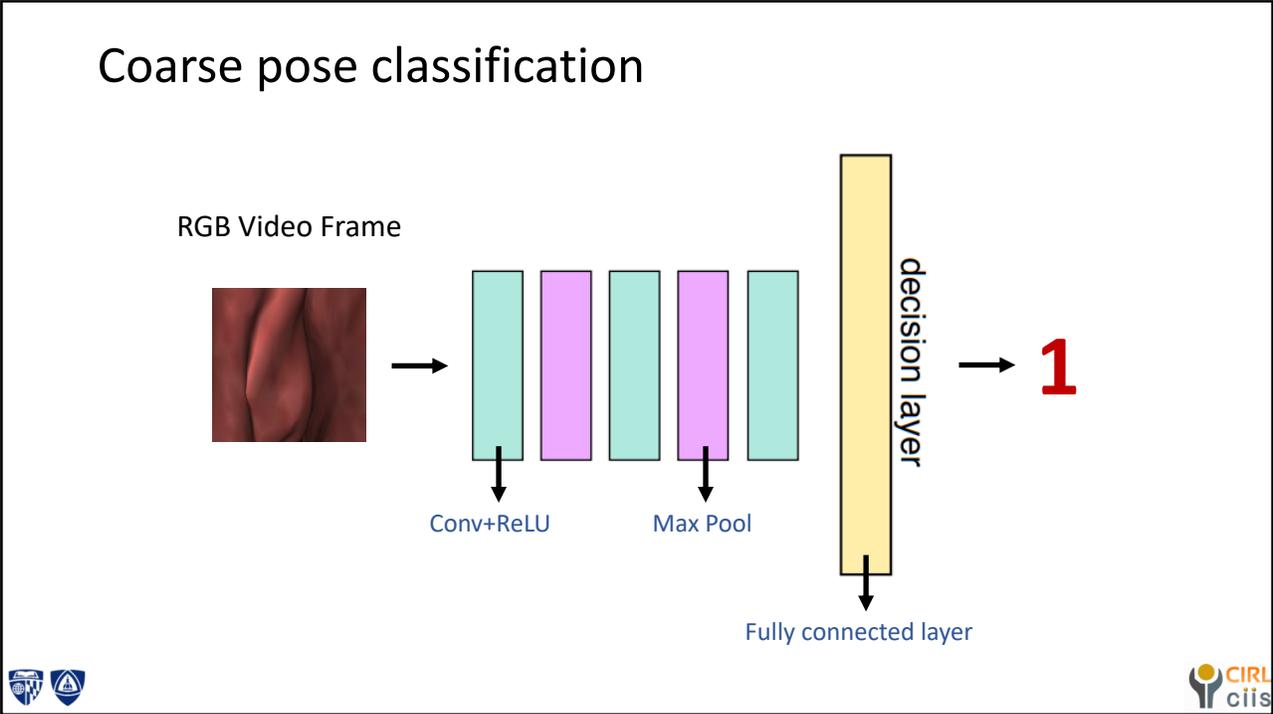


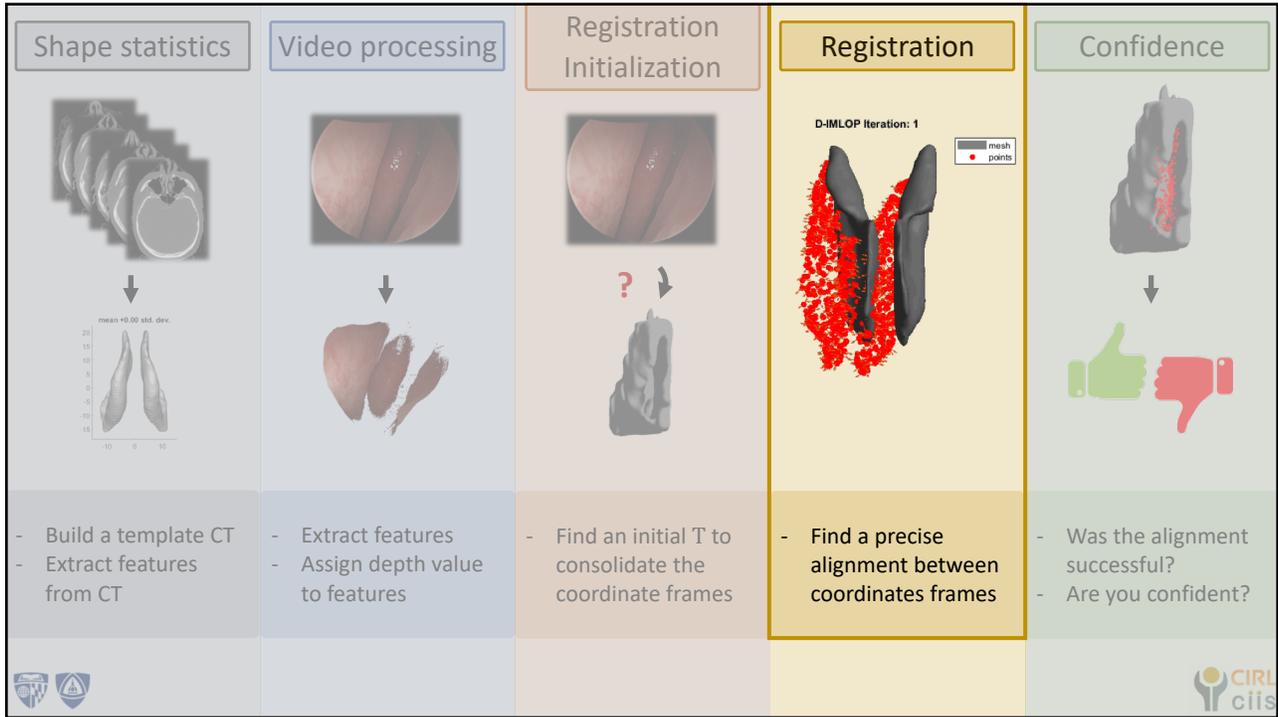
Video



# Coarse pose classification





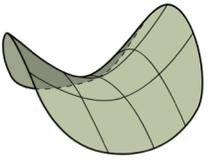


## Iterative **closest** point (ICP) algorithm

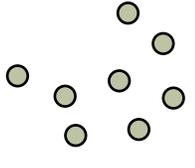
Find closest point match on  $\Psi$  for all  $X$



Compute transformation to align matches



$y_i \in \Psi$



$X = \{x_i\}$

$$\mathbf{T} = \operatorname{argmin}_{[\mathbf{R}, \mathbf{t}]} \sum_{i=1}^n \|y_i - \mathbf{R}x_i - \mathbf{t}\|_2^2$$

P. J. Besl, et al., A method for registration of 3-d shapes, Trans. PAMI, 1992

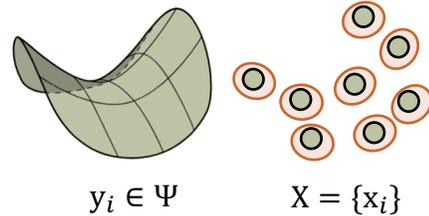



## Iterative **most likely** point (IMLP) algorithm

Find most likely point match on  $\Psi$  for all  $X$



Compute transformation to align matches



$$\mathbf{T} = \underset{[\mathbf{R}, \mathbf{t}]}{\operatorname{argmin}} \sum_{i=1}^n (\mathbf{y}_i - \mathbf{R}\mathbf{x}_i - \mathbf{t})^\top (\mathbf{R}\Sigma_{\mathbf{x}_i}\mathbf{R}^\top)^{-1} (\mathbf{y}_i - \mathbf{R}\mathbf{x}_i - \mathbf{t})$$



S. D. Billings, et al., *Iterative most-likely point registration (IMLP): A robust algorithm for computing optimal shape alignment*, PLOS ONE, 2015



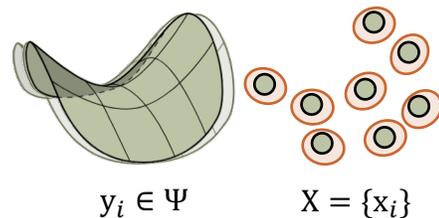
## **Deformable** iterative most likely point (D-IMLP) algorithm

Find most likely point match on  $\Psi$  for all  $X$



Compute transformation to align matches

**and deform  $\Psi$  to fit to  $X$**

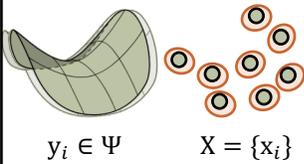


A Sinha, et al., *The deformable most-likely-point paradigm*, Medical Image Analysis, 2018 (in submission)



## Deformable iterative most likely point (D-IMLP) algorithm

$$T = \operatorname{argmin}_{[a, \mathbf{R}, \mathbf{t}], \mathbf{s}} \left( \frac{1}{2} \sum_{i=1}^{n_{\text{data}}} \left( (T_{\text{ssm}}(\mathbf{y}_i) - a\mathbf{R}\mathbf{x}_i - \mathbf{t})^T (\mathbf{R}\Sigma_{\mathbf{x}_i}\mathbf{R}^T)^{-1} (T_{\text{ssm}}(\mathbf{y}_i) - a\mathbf{R}\mathbf{x}_i - \mathbf{t}) \right) + \frac{1}{2} \sum_{j=1}^{n_m} \|s_j\|_2^2 \right)$$



$$T_{\text{ssm}}(\mathbf{y}_i) = f(\mathbf{y}_i, \mathbf{s})$$



A Sinha, et al., *The deformable most-likely-point paradigm*, Medical Image Analysis, 2018 (in submission)

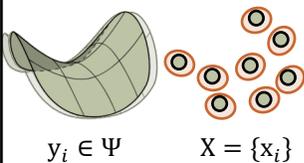


## Deformable iterative most likely point (D-IMLP) algorithm

$$T = \operatorname{argmin}_{[a, \mathbf{R}, \mathbf{t}], \mathbf{s}} \left( \frac{1}{2} \sum_{i=1}^{n_{\text{data}}} \left( (T_{\text{ssm}}(\mathbf{y}_i) - a\mathbf{R}\mathbf{x}_i - \mathbf{t})^T (\mathbf{R}\Sigma_{\mathbf{x}_i}\mathbf{R}^T)^{-1} (T_{\text{ssm}}(\mathbf{y}_i) - a\mathbf{R}\mathbf{x}_i - \mathbf{t}) \right) + \frac{1}{2} \sum_{j=1}^{n_m} \|s_j\|_2^2 \right)$$

Find  $\mathbf{R}$ ,  $\mathbf{t}$  and  $a$  such that  $\mathbf{x}$  is best aligned with a deformed  $\mathbf{y}$

Find  $\mathbf{s}$  such that  $\mathbf{y}$  deforms to fit  $\mathbf{x}$



$$T_{\text{ssm}}(\mathbf{y}_i) = f(\mathbf{y}_i, \mathbf{s})$$



A Sinha, et al., *The deformable most-likely-point paradigm*, Medical Image Analysis, 2018 (in submission)



## Deformable iterative most likely **oriented** point (D-IMLOP) algorithm

$$T = \operatorname{argmin}_{[a, \mathbf{R}, \mathbf{t}], \mathbf{s}} \left( \frac{1}{2} \sum_{i=1}^{n_{\text{data}}} (\mathbf{T}_{\text{ssm}}(\mathbf{y}_{\mathbf{p}_i}) - a\mathbf{R}\mathbf{x}_{i\mathbf{p}} - \mathbf{t})^T (\mathbf{R}\Sigma_x\mathbf{R}^T)^{-1} (\mathbf{T}_{\text{ssm}}(\mathbf{y}_{\mathbf{p}_i}) - a\mathbf{R}\mathbf{x}_{i\mathbf{p}} - \mathbf{t}) \right.$$

Find  $\mathbf{R}$ ,  $\mathbf{t}$  and  $a$  such that  $\mathbf{x}$  is best aligned with a deformed  $\mathbf{y}$ ...

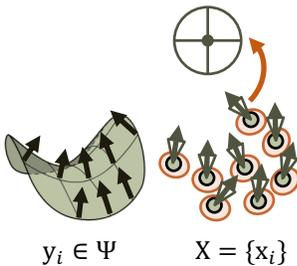
$$+ \kappa \sum_{i=1}^{n_{\text{data}}} (1 - \hat{\mathbf{y}}_{\mathbf{n}_i} \mathbf{R} \hat{\mathbf{x}}_{\mathbf{n}_i})$$

and such that the normal of  $\mathbf{y}$  aligns with that of  $\mathbf{x}$

$$+ \frac{1}{2} \sum_{j=1}^{n_m} \|s_j\|_2^2 \Bigg)$$

Find  $\mathbf{s}$  such that  $\mathbf{y}$  deforms to fit  $\mathbf{x}$

$$\mathbf{T}_{\text{ssm}}(\mathbf{y}_i) = f(\mathbf{y}_i, \mathbf{s}), \quad \kappa_0 = \frac{1}{\sigma_{\text{circ\_rad}}^2}$$



$\mathbf{y}_i \in \Psi$

$\mathbf{X} = \{\mathbf{x}_i\}$



A Sinha, et al., *The deformable most-likely-point paradigm*, Medical Image Analysis, 2018 (in submission)



## **Generalized** deformable iterative most likely oriented point (GD-IMLOP) algorithm

$$T = \operatorname{argmin}_{[a, \mathbf{R}, \mathbf{t}], \mathbf{s}} \left( \frac{1}{2} \sum_{i=1}^{n_{\text{data}}} (\mathbf{T}_{\text{ssm}}(\mathbf{y}_{\mathbf{p}_i}) - a\mathbf{R}\mathbf{x}_{i\mathbf{p}} - \mathbf{t})^T (\mathbf{R}\Sigma_x\mathbf{R}^T)^{-1} (\mathbf{T}_{\text{ssm}}(\mathbf{y}_{\mathbf{p}_i}) - a\mathbf{R}\mathbf{x}_{i\mathbf{p}} - \mathbf{t}) \right.$$

Find  $\mathbf{R}$ ,  $\mathbf{t}$  and  $a$  such that  $\mathbf{x}$  is best aligned with a deformed  $\mathbf{y}$ ...

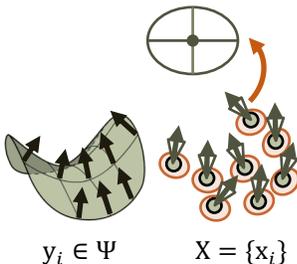
$$- \sum_{i=1}^{n_{\text{data}}} \left( \kappa_i \hat{\mathbf{y}}_{\mathbf{n}_i} \mathbf{R} \hat{\mathbf{x}}_{\mathbf{n}_i} + \beta_i \left( (\hat{\gamma}_{1i}^T \mathbf{R} \hat{\mathbf{y}}_{\mathbf{n}_i})^2 - (\hat{\gamma}_{2i}^T \mathbf{R} \hat{\mathbf{y}}_{\mathbf{n}_i})^2 \right) \right)$$

and such that the normal of  $\mathbf{y}$  aligns with that of  $\mathbf{x}$

$$+ \frac{1}{2} \sum_{j=1}^{n_m} \|s_j\|_2^2 \Bigg)$$

Find  $\mathbf{s}$  such that  $\mathbf{y}$  deforms to fit  $\mathbf{x}$

$$\mathbf{T}_{\text{ssm}}(\mathbf{y}_i) = f(\mathbf{y}_i, \mathbf{s}), \quad \kappa_0 = \frac{1}{\sigma_{\text{circ\_rad}}^2}, \quad \beta = e \frac{\kappa}{2}$$



$\mathbf{y}_i \in \Psi$

$\mathbf{X} = \{\mathbf{x}_i\}$



A Sinha, et al., *The deformable most-likely-point paradigm*, Medical Image Analysis, 2018 (in submission)



## What is $T_{\text{ssm}}(y_i)$ ?

- Given shapes,  $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_{n_v}]^T$ , with correspondences, we can compute the principal components:

- Mean:

$$\bar{\mathbf{V}} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{V}_i$$

- Variance:

$$\Sigma = \frac{1}{n_s} \sum_{i=1}^{n_s} (\mathbf{V}_i - \bar{\mathbf{V}})(\mathbf{V}_i - \bar{\mathbf{V}})^T$$

$$\Sigma = [\mathbf{m}_1 \ \dots \ \mathbf{m}_{n_s}] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n_s} \end{bmatrix} [\mathbf{m}_1 \ \dots \ \mathbf{m}_{n_s}]^T$$



## What is $T_{\text{ssm}}(y_i)$ ?

- Given a new shape,  $\mathbf{V}^*$ , we can compute:

- Mode weights:

$$s_i = \mathbf{w}_i^T (\mathbf{V}^* - \bar{\mathbf{V}})$$

$$\mathbf{w}_i = \sqrt{\lambda_i} \mathbf{m}_i$$

- Estimated shape:

$$\tilde{\mathbf{V}}^* = \bar{\mathbf{V}} + \sum_{i=1}^{n_m} s_i \mathbf{w}_i$$

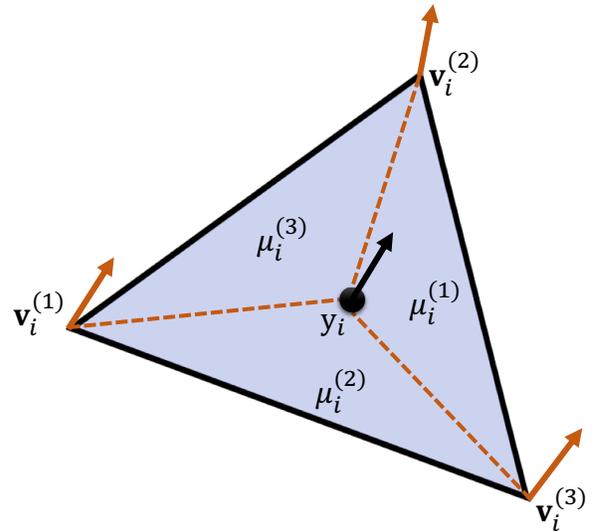


What is  $T_{\text{ssm}}(\mathbf{y}_i)$ ?

$$T_{\text{ssm}}(\mathbf{v}_i) = \bar{\mathbf{v}}_i + \sum_{j=1}^{n_m} s_j \mathbf{w}_j^{(i)}$$

$$T_{\text{ssm}}(\mathbf{y}_i) = \sum_{j=1}^3 \mu_i^{(j)} T_{\text{ssm}}(\mathbf{v}_i^{(j)})$$

$$\sum_{j=1}^3 \mu_i^{(j)} = 1$$



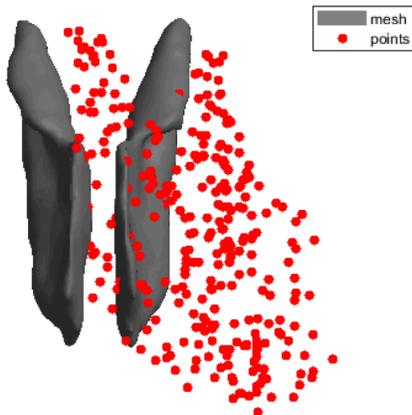
A Sinha, et al., *The deformable most-likely-point paradigm*, Medical Image Analysis, 2018 (in submission)



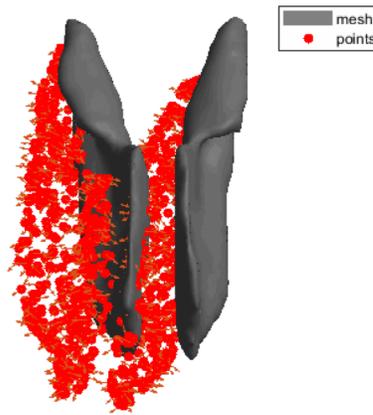
## Deformable most likely point paradigm

- Deformable most likely point paradigm

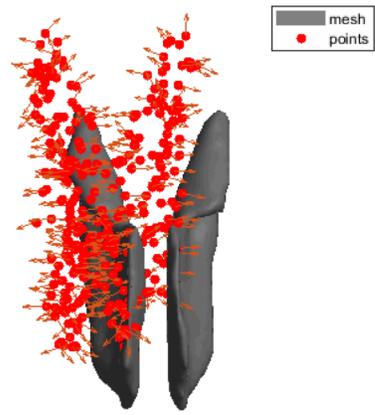
D-IMLP Iteration: 1

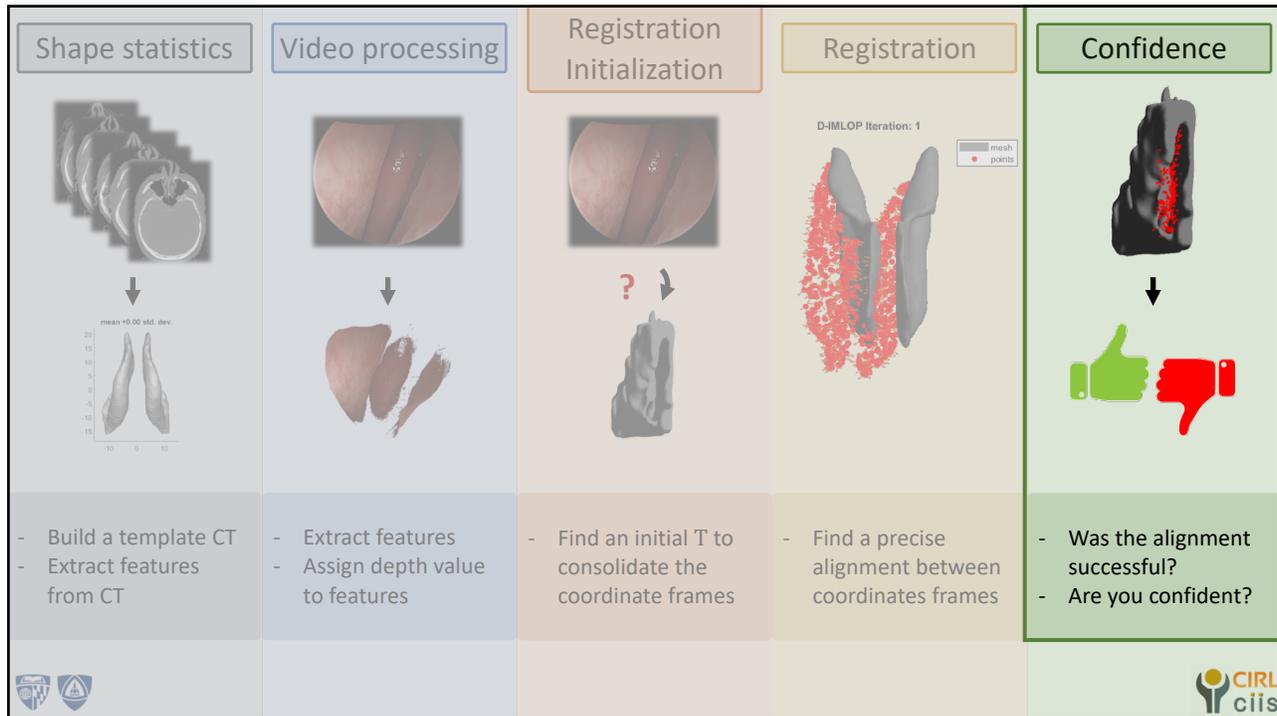


D-IMLOP Iteration: 1



GD-IMLOP Iteration: 1

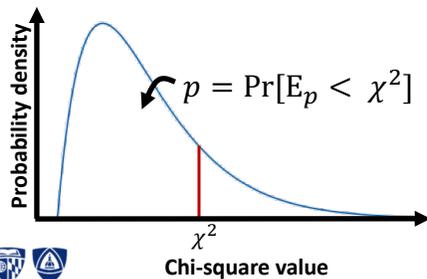




## Again, did the registration work?

$$E_p = \sum_{i=1}^{n_{data}} \left( (T_{ssm}(y_i) - aR x_i - t)^T (R \Sigma_{x_i} R^T)^{-1} (T_{ssm}(y_i) - aR x_i - t) \right) \approx \chi^2 \text{ distribution}$$

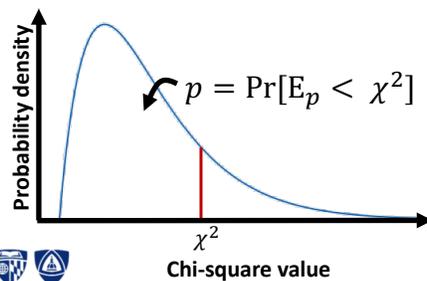
$$E_o = \sum_{i=1}^{n_{data}} \begin{bmatrix} \cos^{-1}(\hat{y}_{n_i}^T R \hat{x}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{1_i}^T R^T \hat{y}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{2_i}^T R^T \hat{y}_{n_i}) \end{bmatrix}^T \begin{bmatrix} \kappa_i & 0 & 0 \\ 0 & \kappa_i - 2\beta_i & 0 \\ 0 & 0 & \kappa_i + 2\beta_i \end{bmatrix} \begin{bmatrix} \cos^{-1}(\hat{y}_{n_i}^T R \hat{x}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{1_i}^T R^T \hat{y}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{2_i}^T R^T \hat{y}_{n_i}) \end{bmatrix} \approx \chi^2 \text{ distribution}$$



## Again, did the registration work?

$$E_p = \sum_{i=1}^{n_{\text{data}}} \left( (T_{\text{SSM}}(\mathbf{y}_i) - a\mathbf{R}\mathbf{x}_i - \mathbf{t})^T (\mathbf{R}\Sigma_{\mathbf{x}_i}\mathbf{R}^T)^{-1} (T_{\text{SSM}}(\mathbf{y}_i) - a\mathbf{R}\mathbf{x}_i - \mathbf{t}) \right) \leq \text{chi2inv}(p, 3n_{\text{data}})$$

$$E_o = \sum_{i=1}^{n_{\text{data}}} \begin{bmatrix} \cos^{-1}(\hat{\gamma}_{n_i}^T \mathbf{R} \hat{\mathbf{x}}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{1_i}^T \mathbf{R}^T \hat{\mathbf{y}}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{2_i}^T \mathbf{R}^T \hat{\mathbf{y}}_{n_i}) \end{bmatrix}^T \begin{bmatrix} \kappa_i & 0 & 0 \\ 0 & \kappa_i - 2\beta_i & 0 \\ 0 & 0 & \kappa_i + 2\beta_i \end{bmatrix} \begin{bmatrix} \cos^{-1}(\hat{\gamma}_{n_i}^T \mathbf{R} \hat{\mathbf{x}}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{1_i}^T \mathbf{R}^T \hat{\mathbf{y}}_{n_i}) \\ \sin^{-1}(\hat{\gamma}_{2_i}^T \mathbf{R}^T \hat{\mathbf{y}}_{n_i}) \end{bmatrix} \leq \text{chi2inv}(p, 2n_{\text{data}})$$



## Leave-one-out analysis

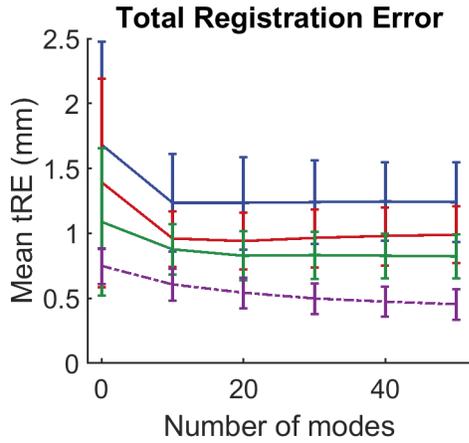
- # sample point: 3000
- Translational offset: [0, 10] mm
- Rotational offset: [0, 10] degrees
- Noise:
  - $0.5 \times 0.5 \times 0.75 \text{ mm}^3$
  - $10^\circ$  ( $e = 0.5$ )
- Noise assumed:
  - $1 \times 1 \times 2 \text{ mm}^3$
  - $30^\circ$  ( $e = 0.5$ )
- $n_m \in \{0, 10, 20, 30, 40, 50\}$



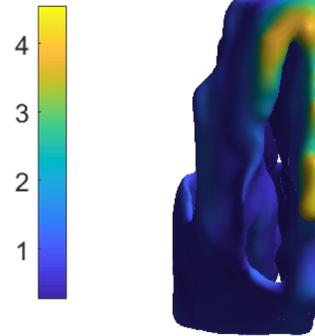
Right nasal airway



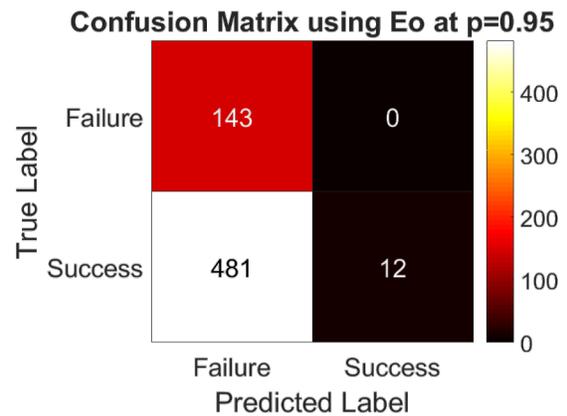
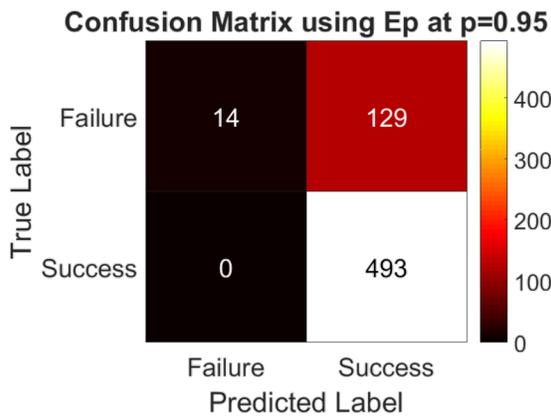
## Leave-one-out analysis



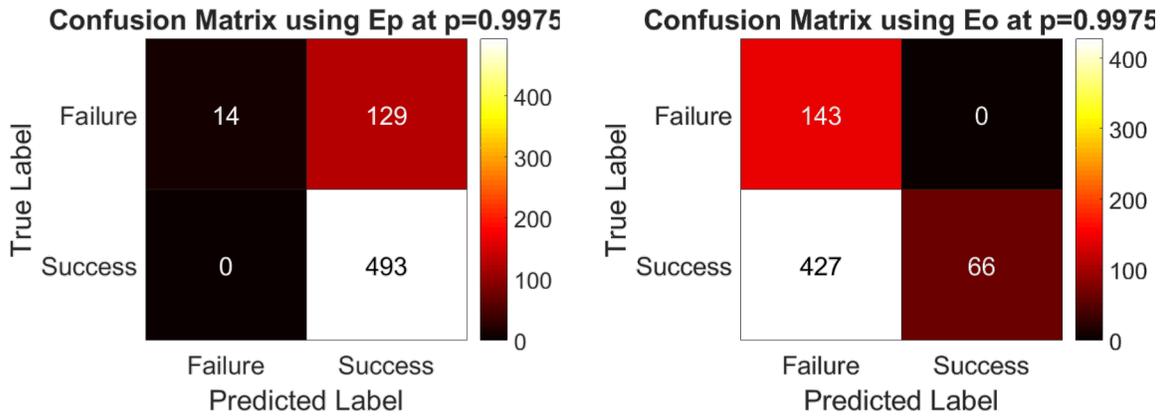
**Shape estimation error (50 modes)**



## Leave-one-out analysis

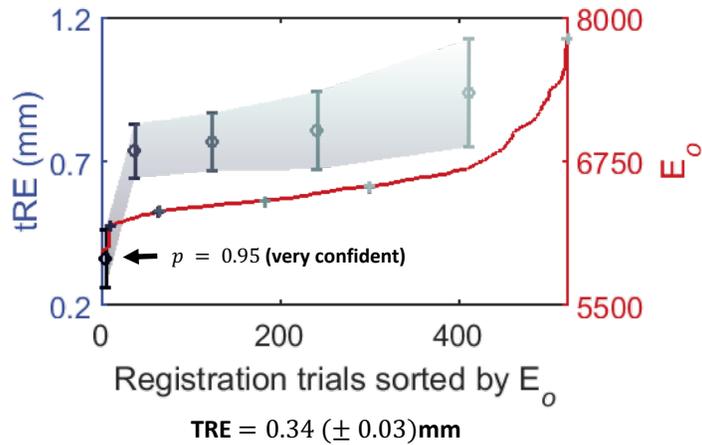


## Leave-one-out analysis



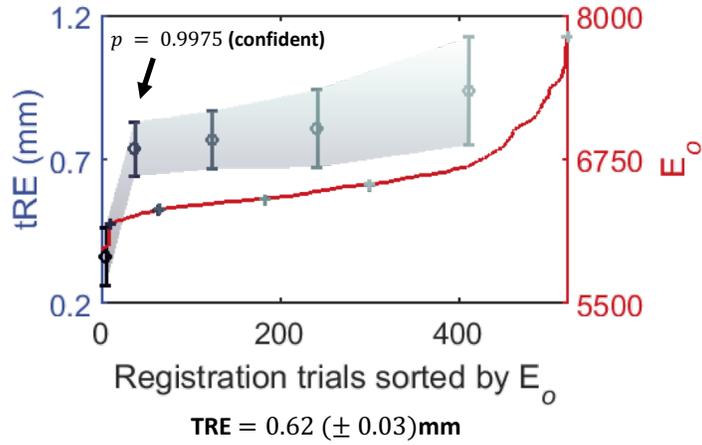
## Leave-one-out analysis

Decreasing confidence in registration accuracy with increasing  $p$



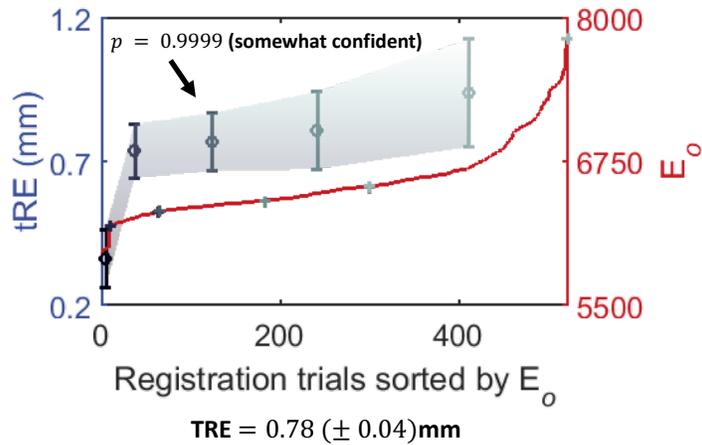
# Leave-one-out analysis

Decreasing confidence in registration accuracy with increasing  $p$



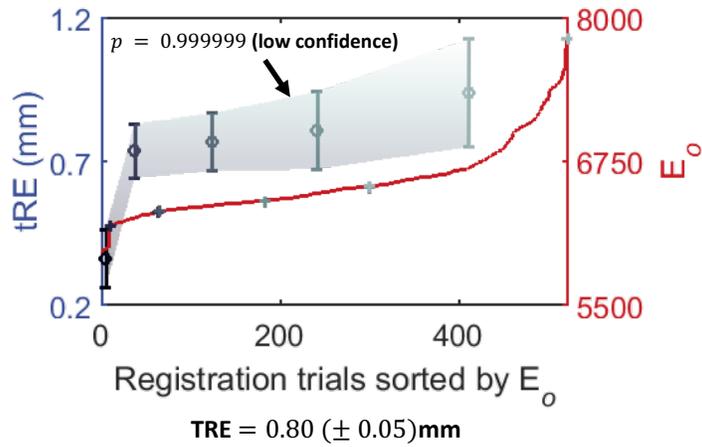
# Leave-one-out analysis

Decreasing confidence in registration accuracy with increasing  $p$



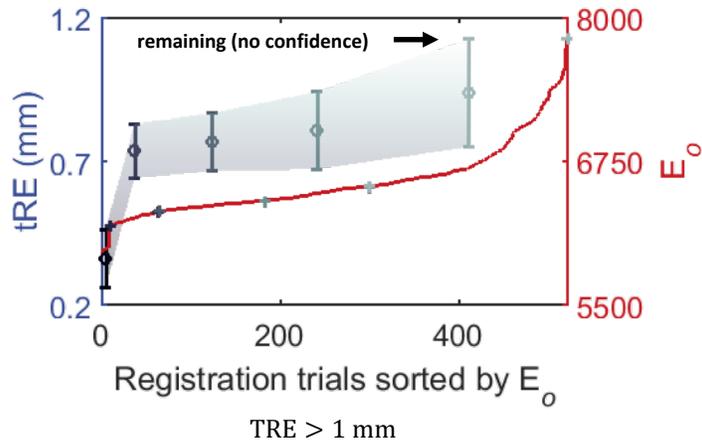
# Leave-one-out analysis

Decreasing confidence in registration accuracy with increasing p



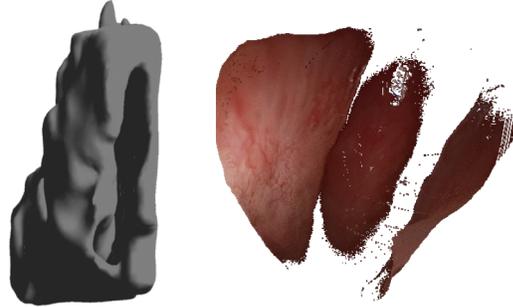
# Leave-one-out analysis

Decreasing confidence in registration accuracy with increasing p



## In vivo experiment

- 5 clinical sequences
- 3000 sample points
- Noise assumed:
  - $1 \times 1 \times 2 \text{mm}^3$
  - $30^\circ$  ( $e = 0.5$ )
- $n_m \in \{0, 10, 20, 30, 40, 50\}$



Right nasal airway

Dense reconstruction  
from video

## In vivo experiment

	# registrations	Residual error (mm)	Max error (mm)	Min error (mm)
All registrations	30/30	1.09 ( $\pm 1.03$ )	4.74	0.50
Registrations that pass $E_p$ test	27/30	0.76 ( $\pm 0.14$ )	0.99	0.50
Registrations that pass $E_p$ and $E_o$ tests	12/30	0.78 ( $\pm 0.07$ )	0.94	0.72



## Conclusion

- Steps towards fully automated navigation during clinical exploration and surgery
- In the process of building a comprehensive atlas of the nasal cavity and sinuses
- Confidence measures to allow clinicians to modulate trust in navigation system



Thank you!

Questions?

