Classification and Feature Selection for Craniosynostosis

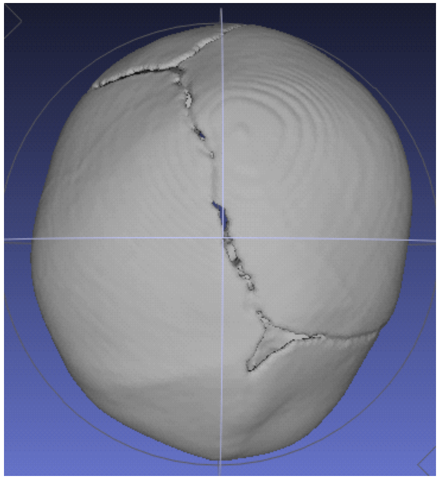
Shulin Yang, Linda G. Shapiro, Michael L. Cunningham, Matthew Speltz, Su-In Lee

University of Washington

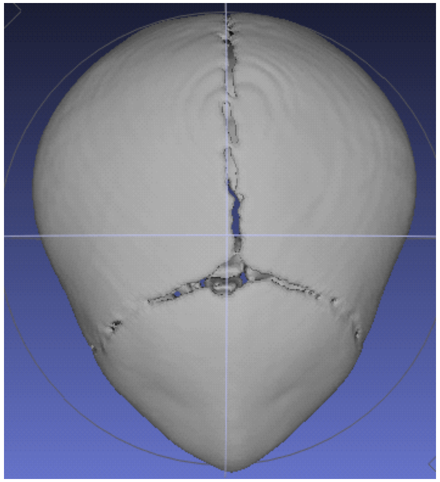
Craniosynostosis

- Craniosynostosis is a common congenital condition in which one or more of the fibrous sutures in an infant's calvaria fuse prematurely.

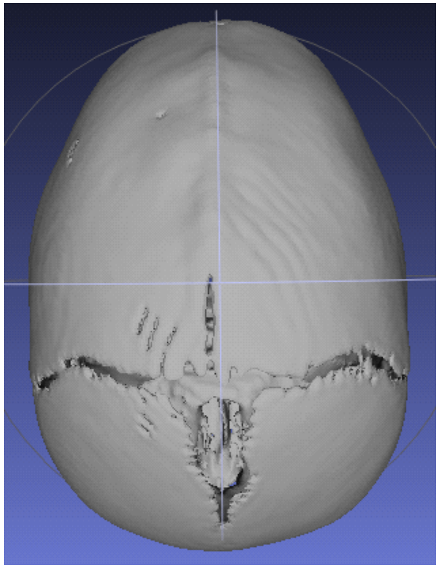
Coronal



Metopic



Sagittal



Goal of our Work

- To analyze 3D skull shapes for the purpose of medical research
  – Classification: which type of craniosynostosis
  – Region selection: which regions contribute most toward classification
  – Quantification: what is the degree of severity of the deformity

Related Work

- Previous work
  - Craniofacial descriptors that analyzed the shape of the mid-face and back of the head [ICIAP09]
  - Classification of two synostoses vs. normal using symbolic shape descriptors [ICCV05, CPCJournal06]

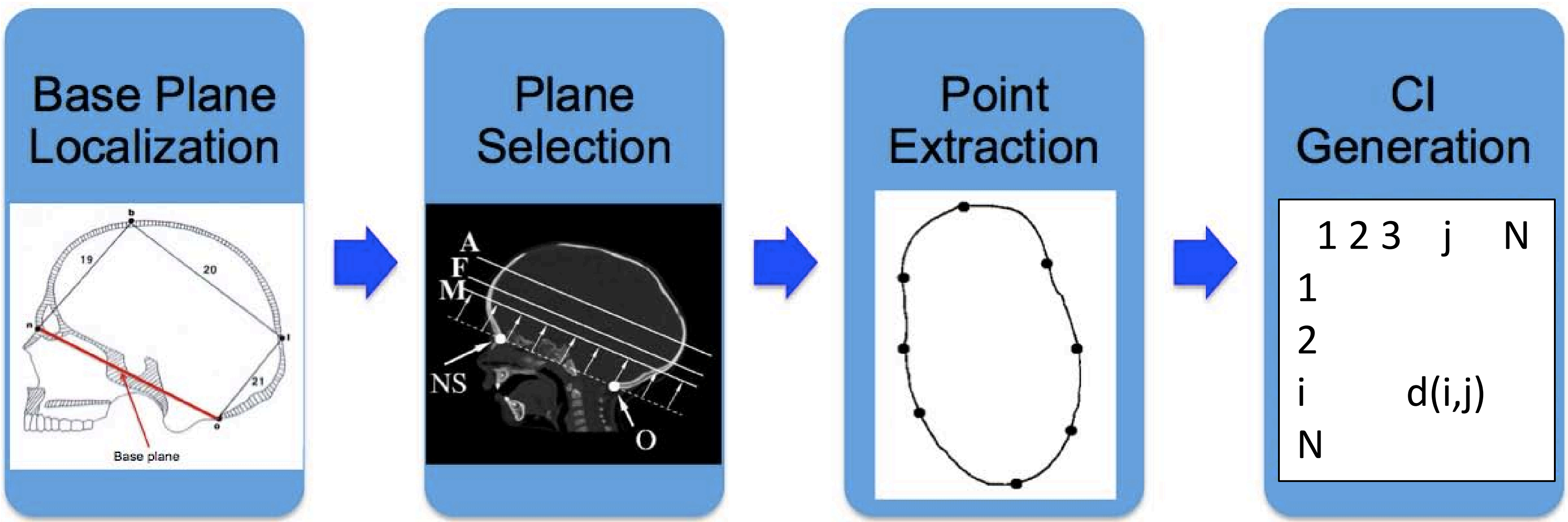
- Difference from ours
  - not fully automatic
  - doctors may not understand the methodology
  - focus on the whole skull

Overview of our Approach

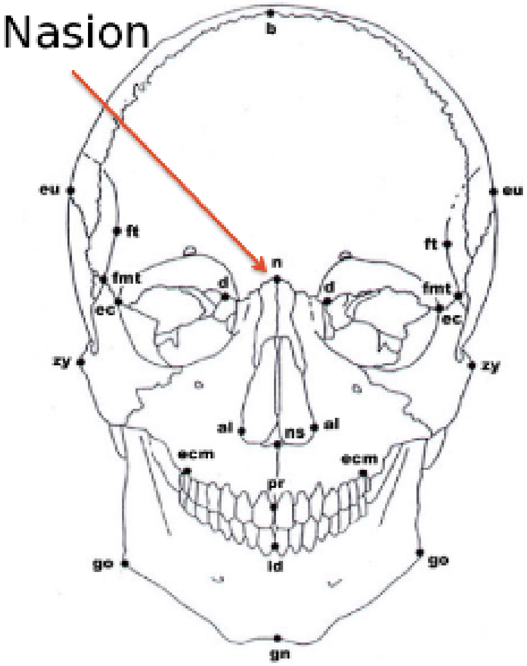
- Two step approach to shape analysis
  - Cranial image (CI) generation
    - A shape representation for 3D images
  - Shape analysis using CI
    - Classification
    - Localization of interest areas on skulls
    - Quantification of craniofacial abnormalities

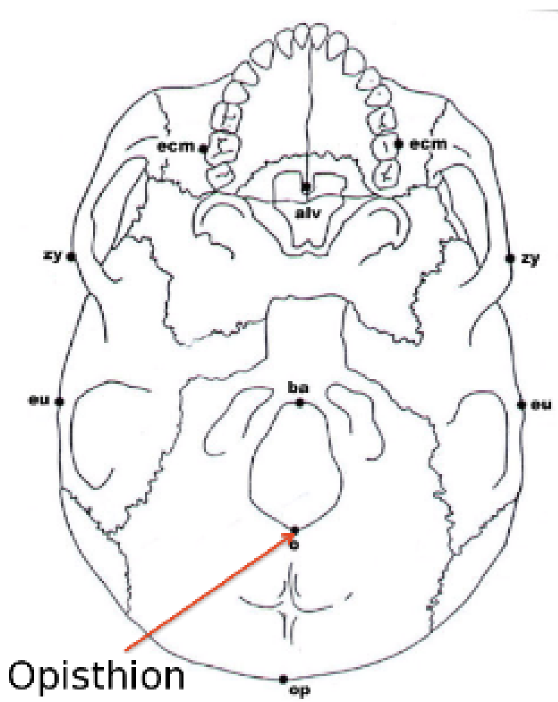
Cranial Image (CI) Generation

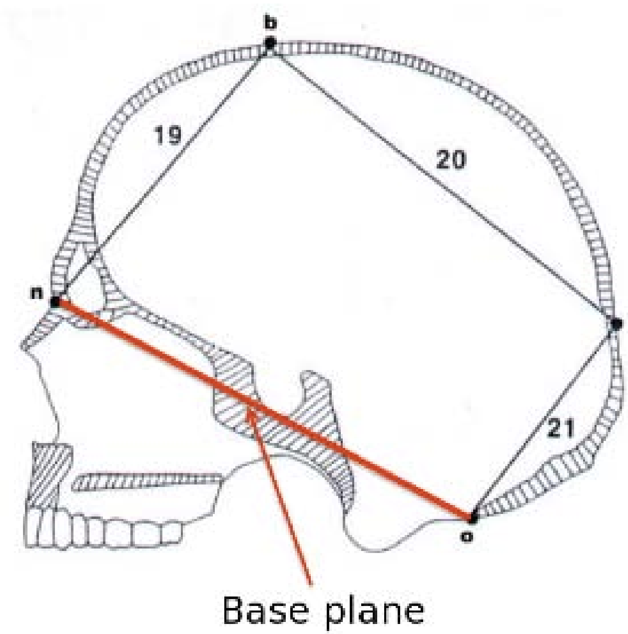
- Automatic system: process 3D CT images
  - Input: CT skull images
  - Output: a distance matrix – Cranial Image



Landmarks

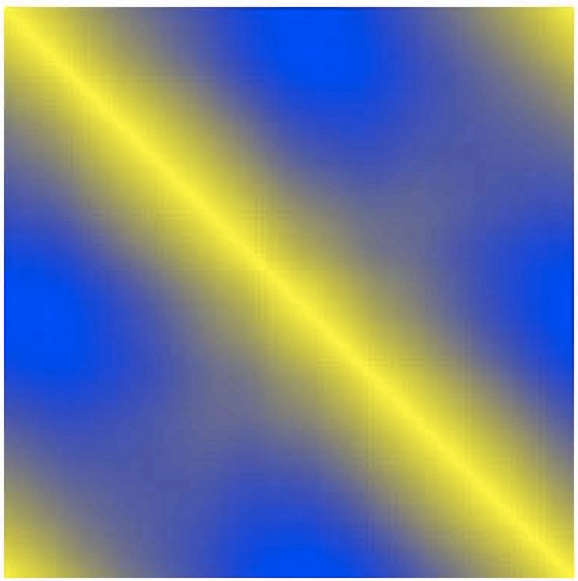

Nasion
b
eu
ft
fmt
ec
zy
n
d
al
ns
ecm
pr
id
go
gn


ecm
alv
zy
eu
ba
o
Opisthion
op

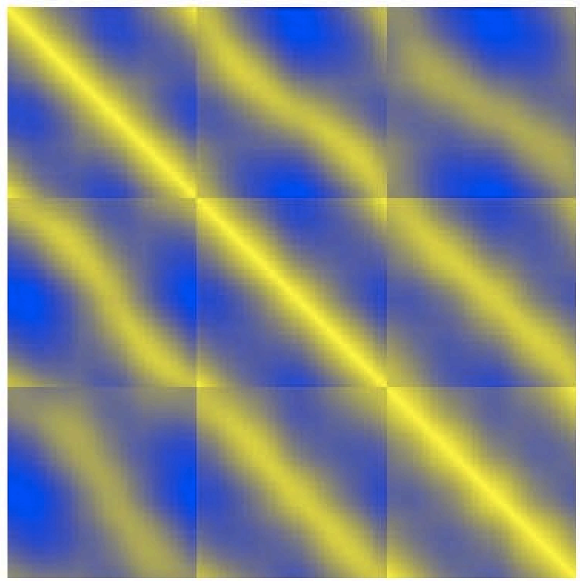

b
19
20
n
l
21
o
Base plane

Visualization of Cranial Image

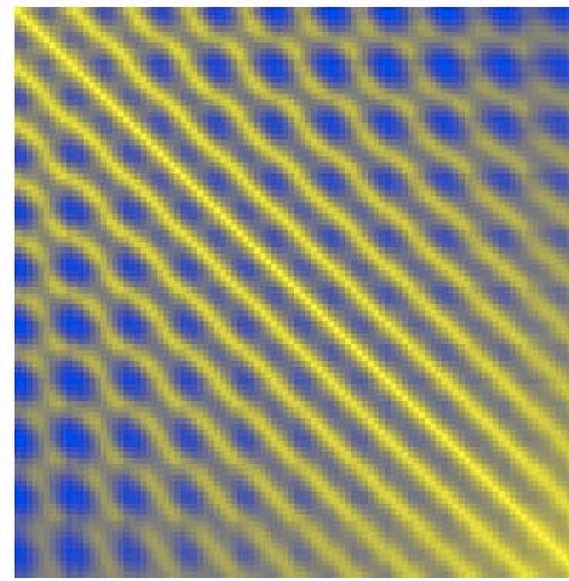
- Yellow represents 0 in the matrix;
- Blue represents 1 in the matrix



1 plane



3 planes



10 planes

Shape Analysis using CI

- Goal: classification and feature selection on CI

- Methodology: logistic regression model
  - x: features in a Cranial Image
  - y: classification result for a Cranial Image
  - w: weight assigned to each feature in a Cranial Image

$$p(y|\mathbf{x}, \mathbf{w}) = \frac{1}{1 + exp(-y(\mathbf{w}^T\mathbf{x} + w_0))}$$

Logistic Regression

- Find the "w" that minimize the loss function

$$l(w_0, \mathbf{w}) = \sum_{i=1}^{n} \log(1 + \exp(-y_i(\mathbf{w}^T \mathbf{x_i} + w_0)))$$

$$\{w_0, \mathbf{w}\} = \min_{w_0, \mathbf{w}} l(w_0, \mathbf{w})$$

Regularized Logistic Regression Models

- L_1 regularized logistic regression

$$l(w_0, \mathbf{w}) = \sum_{i=1}^{n} \log(1+\exp(-y_i(\mathbf{w}^T\mathbf{x_i}+w_0)))+\lambda\sum_{i=1}^{m}|w_i|$$

suppress the number of selected features

- Fused lasso

$$l(w_0, \mathbf{w}) = \sum_{i=1}^{n} \log(1 + \exp(-y_i(\mathbf{w}^T\mathbf{x_i} + w_0)))$$
$$+\lambda\sum_{i=1}^{m}|w_i| + \mu \sum_{\{w_i, w_j\}\in M} |w_i - w_j|$$

suppress the number of selected features

suppress weight differences for pairs of neighboring features

A New Form of Regularized Logistic Regression Models: cLasso

- cLasso – forming feature clusters
  - w^c: weights of the cluster centers of CI features
  - w: residual weights of the features

$$p(y|\mathbf{x}, \mathbf{w}, \mathbf{w^c}) = \frac{1}{1 + exp(-y(\mathbf{w}^T\mathbf{x} + \mathbf{w^c}^T + w_0))}$$

$$l(w_0, \mathbf{w}, \mathbf{w^c}) = \sum_{i=1}^{n} \log(1 + \exp(-y_i(\mathbf{w}^T\mathbf{x_i} + \mathbf{w^c}^T\mathbf{c_i} + w_0))) + \lambda \sum_{i=1}^{m} |w_i| + \nu \sum_{i=1}^{k} |w_i^c|$$

suppress the number of selected features

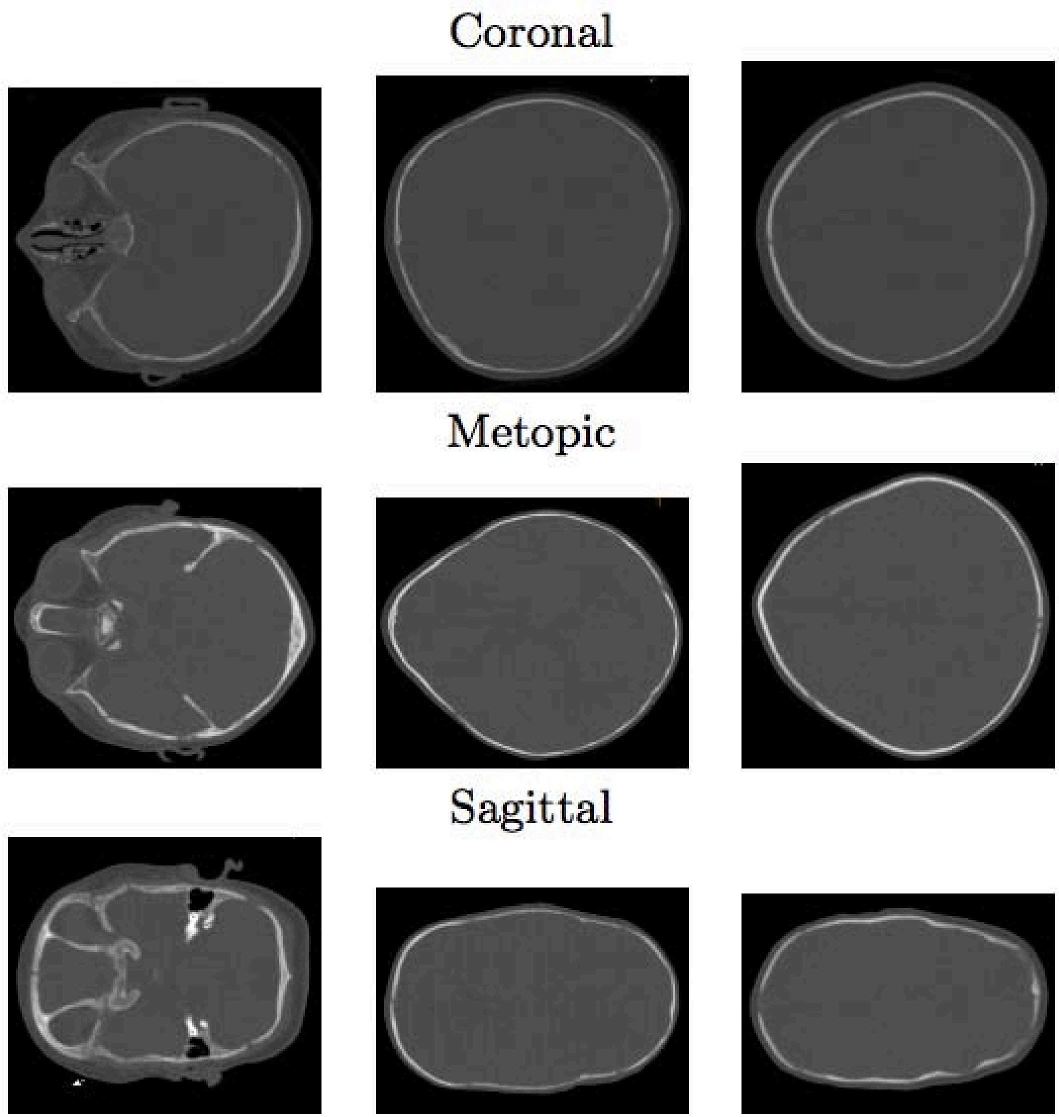
suppress the number of feature clusters

Experiments

- Medical data: 3D CT images of children’s heads from hospitals in four different cities in the US; 70 images in total; 3 types of craniosynostosis

- Parameter selection: the regularization parameters were found using 10-fold cross validation on the training set

CT data



Classification Results: Error Rates

Results using logistic regression only

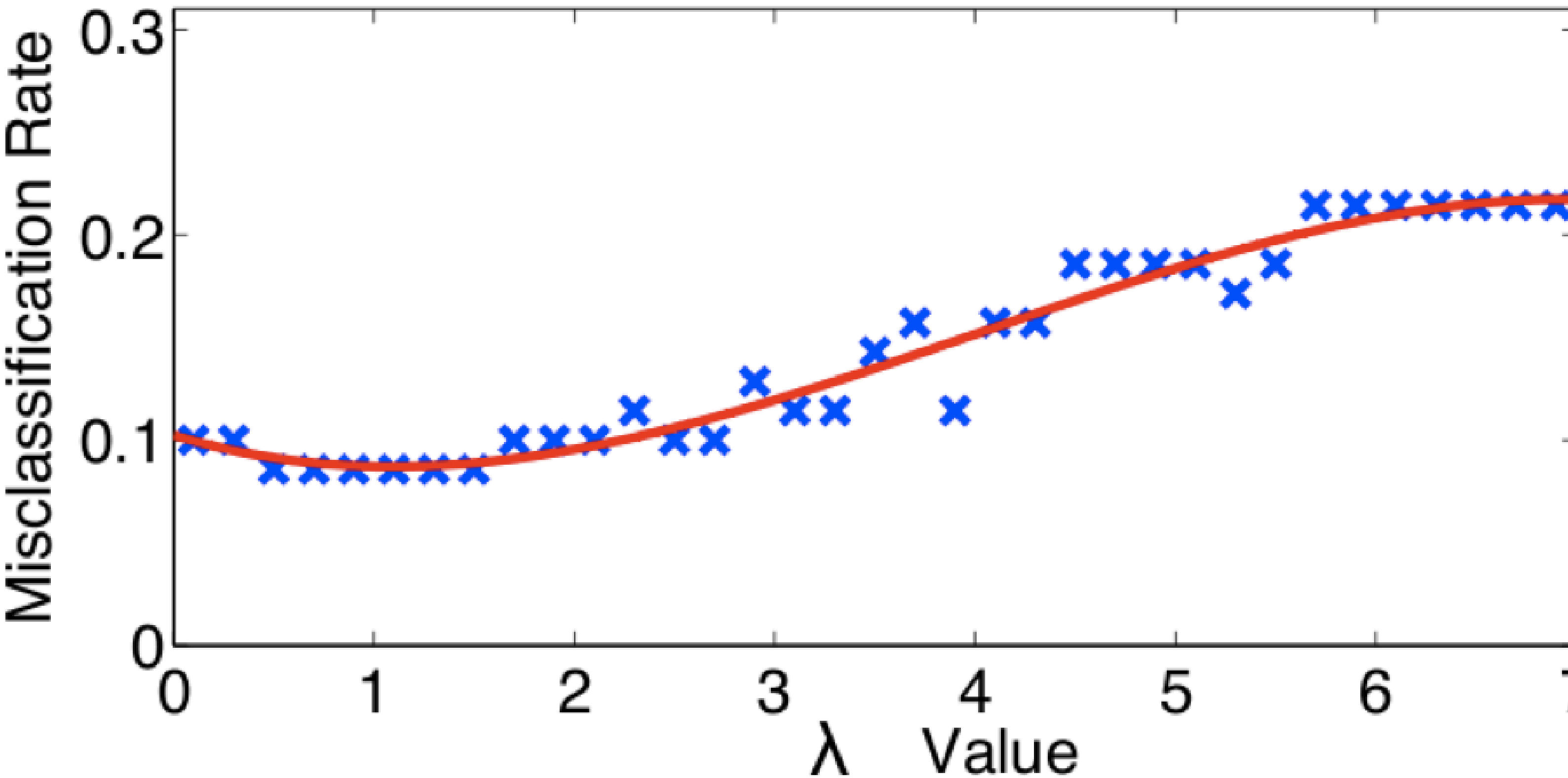
One Plane	**C vs M**	**M vs S**	**S vs C**	3-Classes
A-Plane	**3.29%**	**12.67%**	26.29%	10%
F-Plane	4.39%	17.57%	**25.57%**	10%
M-Plane	6.29%	17.14%	27.14%	10%

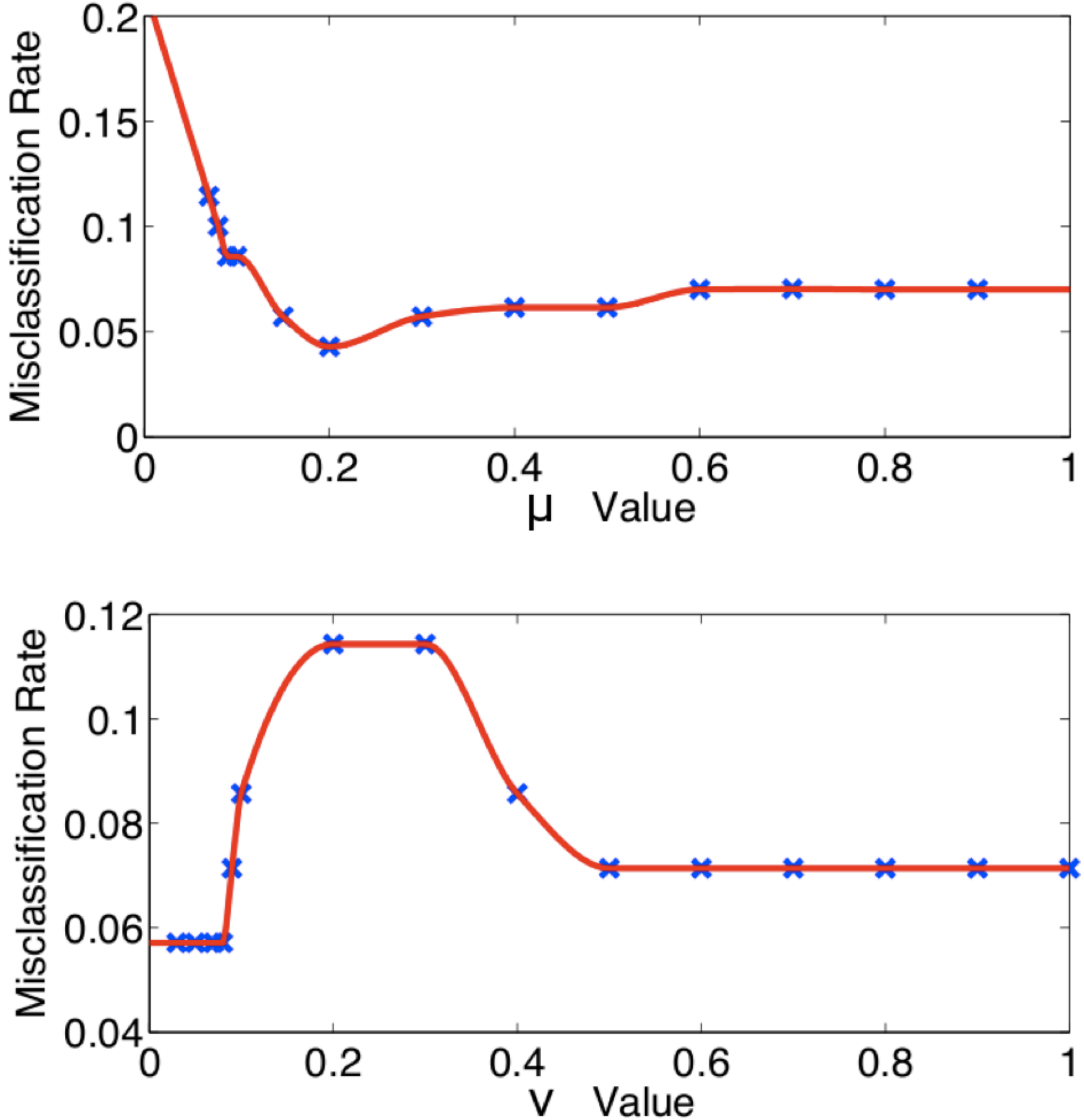
Results using four regression models

Multiple Planes	C vs M	M vs S	S vs C	3-Classes
Logistic regression	13.57%	13.57%	23.93%	10%
L_1 regression	7.14%	5%	6.43%	8.57%
Fused lasso	5.71%	5.71%	**4.29%**	18.57%
Clustering lasso	**4.29%**	**4.29%**	5.71%	**7.14%**

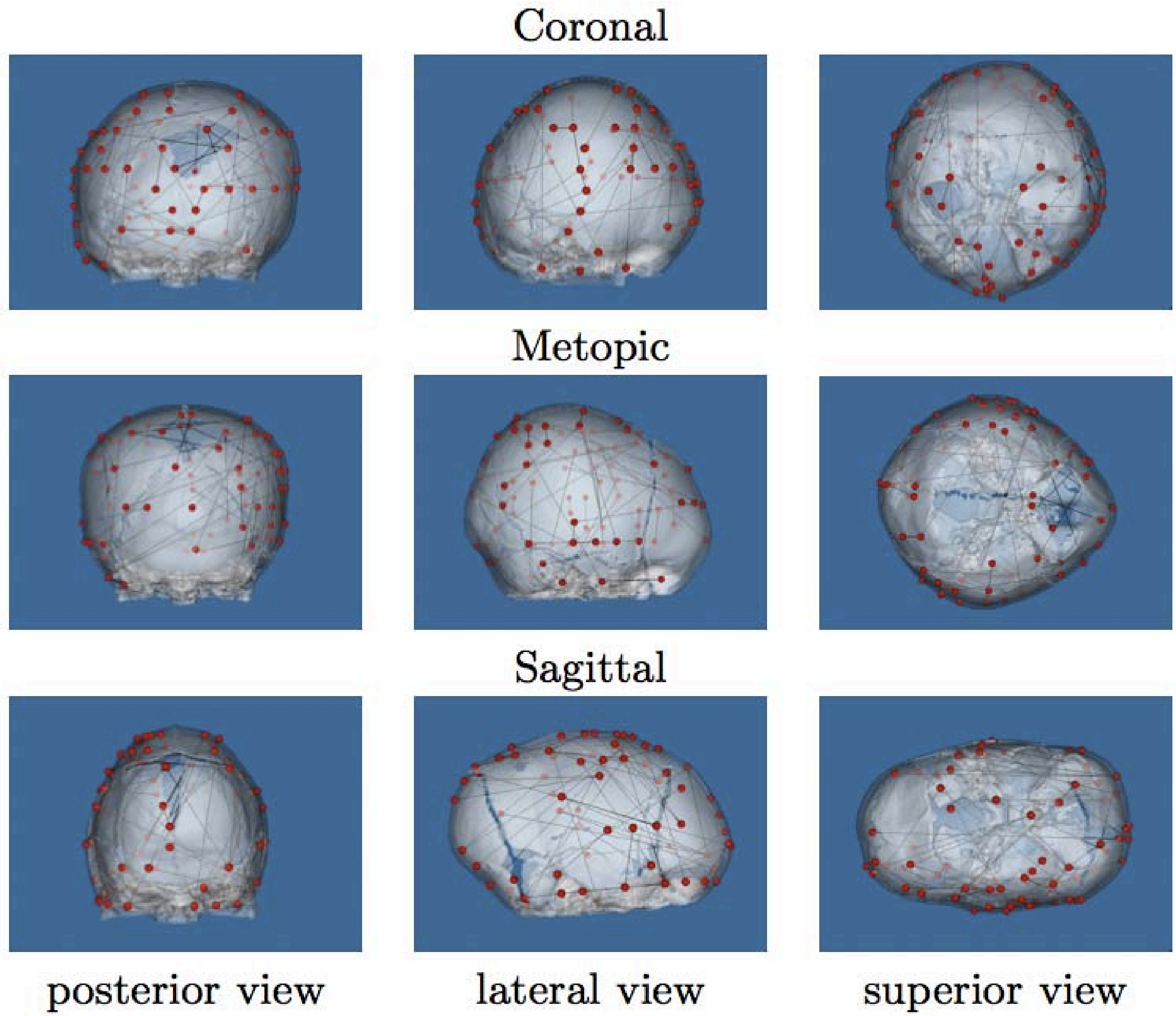
Parameter selection

Misclassification rate v.s. lamda value

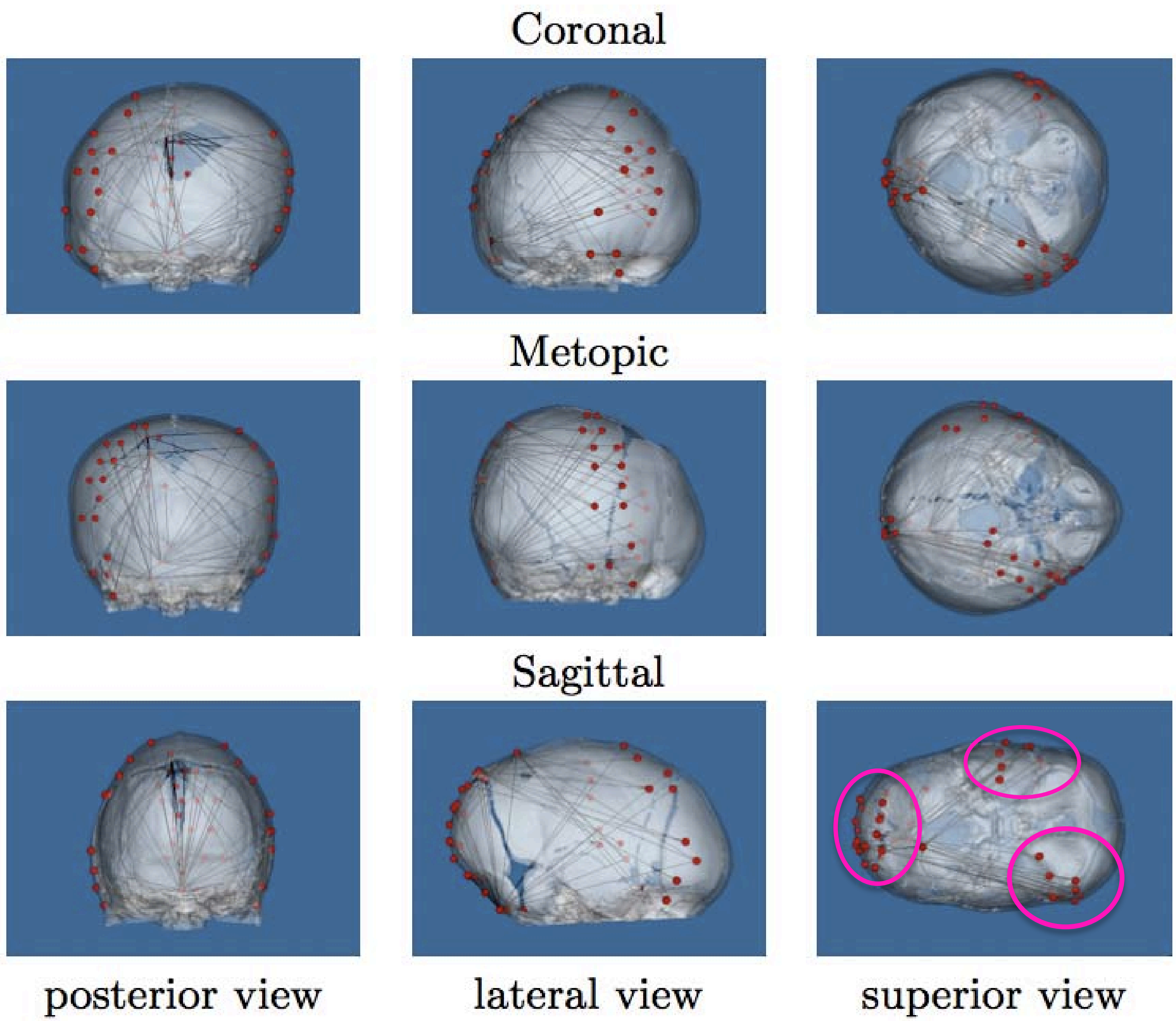



0.2
0.15
0.1
0.05
0
Misclassification Rate
0
0.2
0.4
0.6
0.8
1
μ Value
0.12
0.1
0.08
0.06
0.04
Misclassification Rate
0
0.2
0.4
0.6
0.8
1
ν Value

Visualization of Feature Selection using L_1 Regression

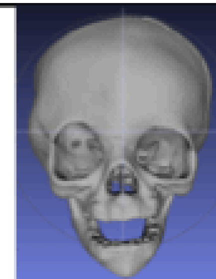
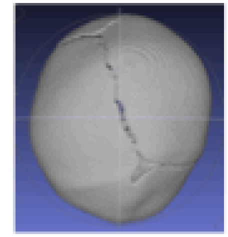
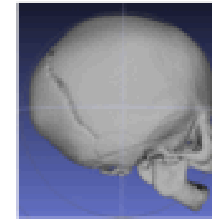
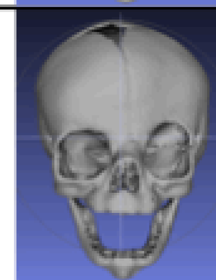
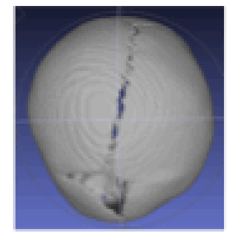
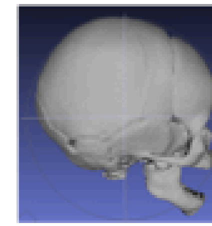
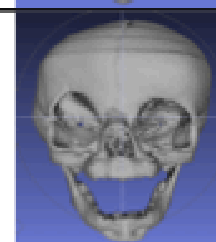
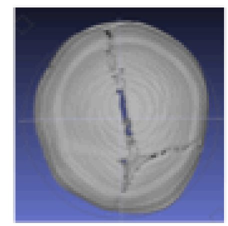
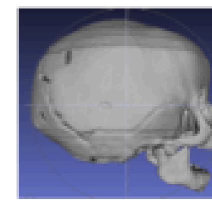
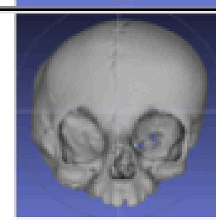
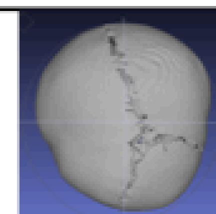
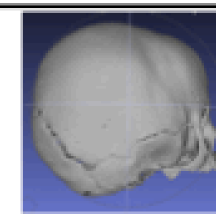
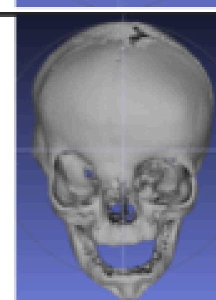
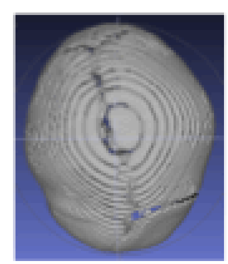
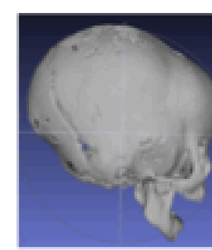


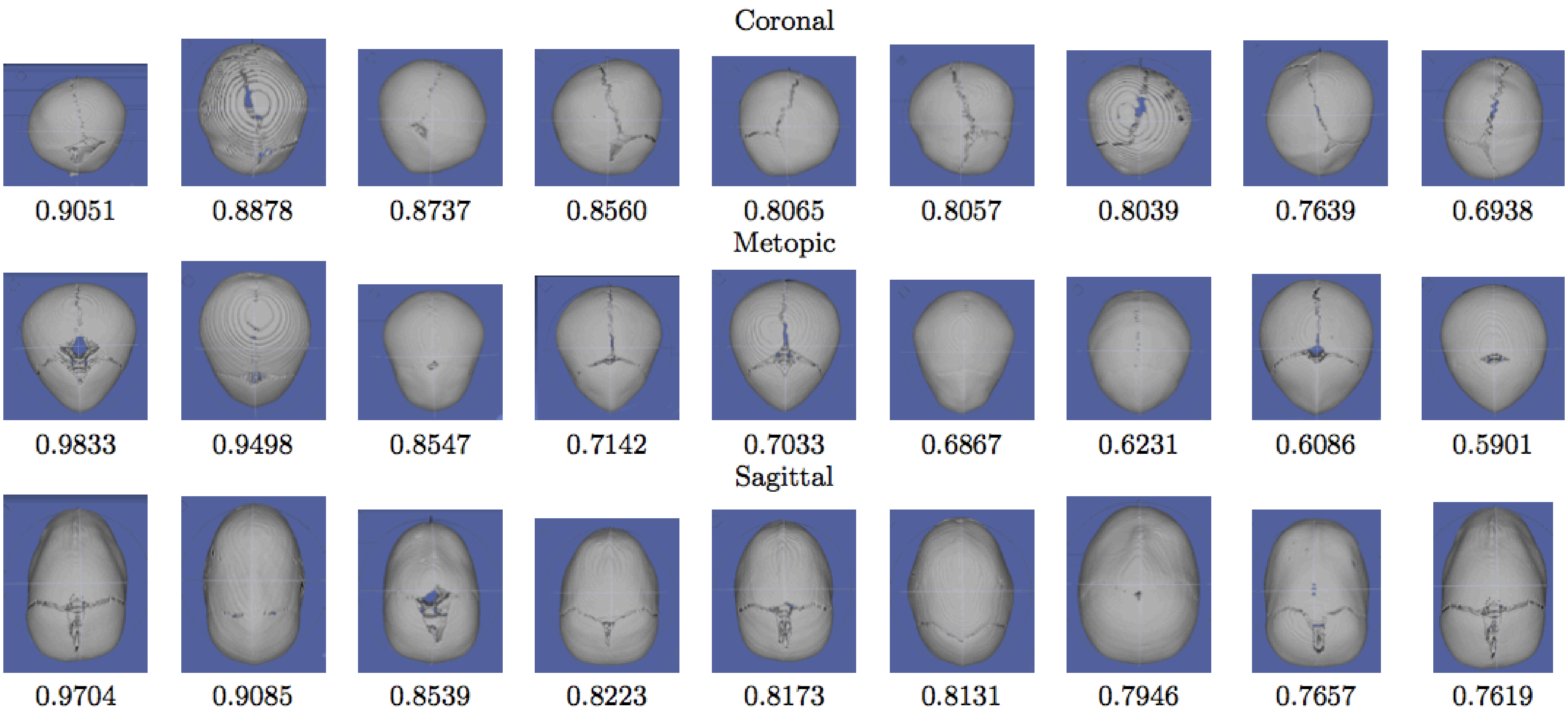
Visualization of Feature Selection using Fused Lasso



Quantification

- Use the same methodology as classification
  - Use the same training process
  - Replace the decision function: from sigmoid function to the linear combination before taking sigmoid
  - Function response: the severity of craniosynostosis
- Use different training data for each of the three craniosynostosis (coronal, metopic, sagittal)

ID	Type	M rank	O rank	O score	Front	Top	Side view
2086	c	1	1	0.9976			
3014	c	5	2	0.9971			
3026	c	4	3	0.9952			
1087	c	6	4	0.9909			
3003	c	3	5	0.9870			


Coronal
0.9051
0.8878
0.8737
0.8560
0.8065
0.8057
0.8039
0.7639
0.6938
Metopic
0.9833
0.9498
0.8547
0.7142
0.7033
0.6867
0.6231
0.6086
0.5901
Sagittal
0.9704
0.9085
0.8539
0.8223
0.8173
0.8131
0.7946
0.7657
0.7619

Summary and Future Work

- Contributions
  - A fully automatic system for skull shape analysis
  - New form of logistic regression for feature selection and interest region localization
- Future work
  - Extension to other 3D shapes, such as facial surfaces
  - Landmark detection on 3D surfaces
  - Run studies of controls vs abnormal for each class and use results to quantify the degree of abnormality