UCLA Applied Math Summer REU Final Report. Placenta Imaging Classification: At Risk Detector (PICARD)

Ben Jones
Department of Mathematics
Harvey Mudd College

Austin Adams
Department of Mathematics
CSU Long Beach

Anh Nguyen
Department of Mathematics
UC Los Angeles

August 6, 2010

Abstract

1 Introduction

A placenta is the organ connecting a developing fetus to its mother. The placenta is how the fetus absorbs nutrients and gets rid of waste. Until recently it was determined that placentas had no relevant information about a baby and were thrown away as a result.

2 Perinatal Pathology with Placenta

2.1 Data Set

We were provided around 1000 raw placenta images. About 200 of these images were given again to us but hand traced. The raw images were from many different studies of the placenta; there was no set standard for how the images were taken. This proved challenging later on to try and extract features from. The traced images were much easier to extract vessel features from.

2.2 Who is at risk

Each placenta image is accompanied by data about the mother. The only information accompanying the placenta concerning the child is the gestational age and birth weight. To find a correlation between placental features and child health further information about the true health of the child is required. Using the data set provided to the 2010 Summer REU the placentas are labeled as "at risk" or "not at risk" based solely upon the corresponding gestational age and birth weight of the child.

The labeling algorithm relies on a region of normal birth weights for any given gestational age. Since young gestational age contributes to a low birth weight, this region must be adjusted based upon the gestational age in weeks. The three labeling options are presented here.

First, Williams Obstetrics, a central text in the medical specialty of obstetrics, provides a table of percentiles for birth weights and gestational age [2]. The table is based on 3,134,879 single live births in the United States. It includes the fifth, tenth, fiftieth, ninetieth and ninety-fifth percentiles for birth weights of children born between the twentieth and forty-forth week of gestation. The table is used to create two possible labeling options using the 10 - 90th percentile range as normal and using the 5 - 95th percentile range as normal. Placentas corresponding to the children in the range are labeled "not at risk." Children above and below are considered to have "high birth weight" and "low birth weight," respectively. If the child has high or low birth weight, then the placenta is labeled "at risk" [Figures 1 & 2].

The data set provided to the 2010 Summer REU contains 614 placental images with supporting mother and baby data, but the study includes 1318 more cases without images. Therefore, there are 1932 birth weights and corresponding gestational ages ranging from 19 to 43 weeks. The third labeling option uses this data by splitting the data by gestational week and cosidering the mean birth weight and the standard deviation for each week. These numbers are used to form two smooth curves–one standard deviation above and below the mean– to create the normal region. Cubic functions (1) and (2) provide the best fit regression lines ($R^2 = 0.97$ and $R^2 = 0.99$, respectively) for the two curves.

$$f_{oneSDabove} = -0.4x^3 + 41.2x^2 - 1178.6x + 10663 \tag{1}$$

$$f_{oneSDbelow} = -0.2x^3 + 19.4x^2 - 587.2x + 5440.2 \tag{2}$$

A placenta corresponding to a point within the curves is labeled "not at risk." The placentas corresponding to points above equation (1) with high birth weights are labeled "at risk" along with the placentas corresponding to points below equation (2). [Figure 3]

Figures 1, 2 and 3 show the three options for labeling the placentas corresponding to the 614 images. While there are pros and cons to each labeling algorithm, the classification algorithm described in section 7 uses the labeling from the cubic curves derived from one standard deviation above and below the mean for each gestational week [Figure 3]. The pro of using the percentile table in Williams Obstetrics is that the percentiles in the table come from a large data set of over three million live births. The con, however, is that these labeling schemes label all children born before the 35^{th} week of gestation as "not at risk," which is simply not true. The 10 - 90^{th} percentile normal region labels 505 (or 82.25%) of the 614 images as "not at risk." The 5 - 95^{th} percentile normal region labels 559 (or 91%) of the 614 images as "not at risk." The chosen algorithm using standard deviations labels 418 (or 68%) of the 614 images as "not at risk." While there is significantly less data contributing to the chosen labeling algorithm, it does accurately allow low gestational ages to be more often classified as "at risk."

3 Preprocessing

Preprocessing is required in order for data to be relevant. Many of the raw images contain rulers, hands, etc. that are not of interest. In addition, the images often have high glare and low contrast. This produces data that

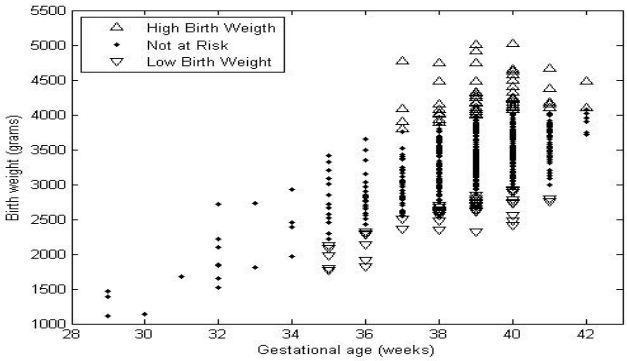


Figure 1: 10 - 90th Percentile Labeling

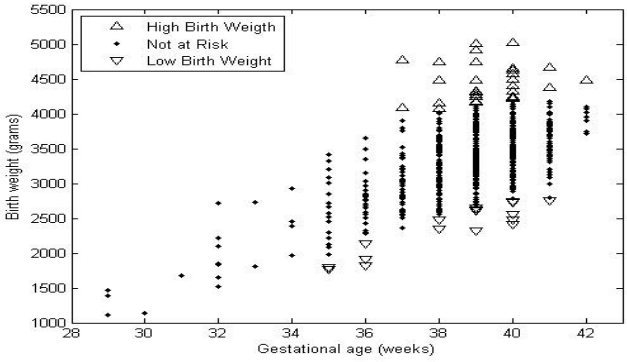


Figure 2: 5 - 95th Percentile Labeling

cannot be separated into classes.

In order to fix these problems we used the algorithm described in the following section.

3.1 Crop and Deglare

Cropping an image followed these steps.

1. Use only the green channel of the image.
2. Contrast stretch green channel.
3. Convert image to binary by thresholding about the mean intensity.
4. Detect the largest object.
5. Fill in the largest object.

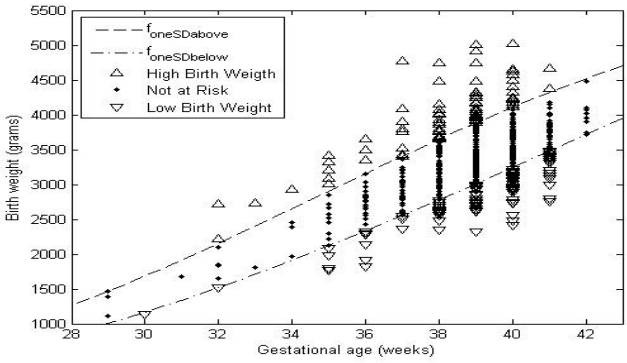


Figure 3: Standard Deviation Labeling: used for classification

6. Erode the image.
7. Use this mask on original image.
8. Deglare image.

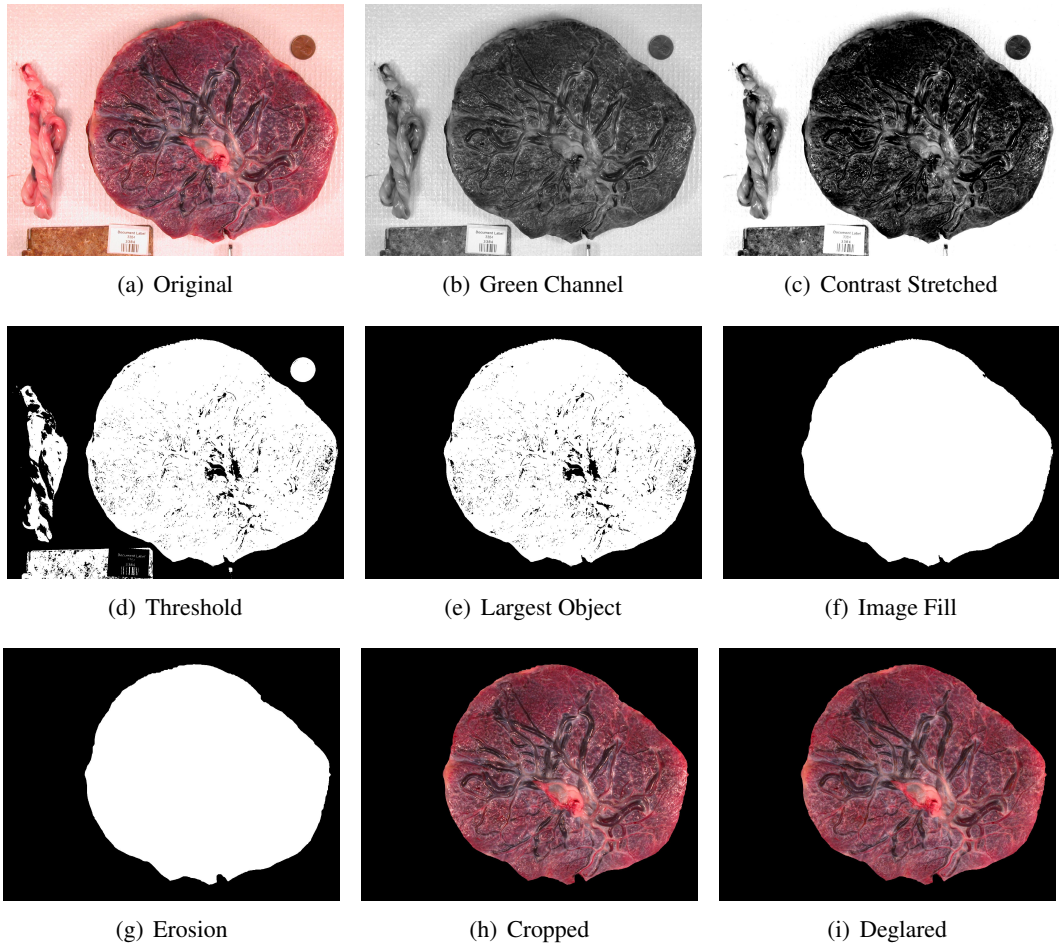
This cropping procedure can be seen visually in Figure 3.1. To deglare the image we used an algorithm written by Almoussa et al. [10].

3.2 Location of Cord Insertion

Automating location of the umbilical cord insertion proved to be too unreliable, so human observers were used to locate the umbilical cords in each image. To do so, four observers clicked on images of placentas where they thought the umbilical cord insertion point was. The three closest samples were then averaged to yield an average cord position. Three observers were then presented with the same placenta images with the previously chosen cord insertion locations and average highlighted, and allowed to choose again. The closest three of these plus the previous average were averaged to form the final insertion location. In cases where the variance of this final group of three in either the horizontal or vertical direction, whichever was higher, was greater than 40 pixels, the three observers re-evaluated the image to determine the true cord location, or whether the position was too unclear to be admissible into the data set. Placenta images in the later category were thrown out of the classification data set.

4 Placental Functions

The placenta is both the source of nutrients and the waste disposal system for the developing fetus. Thus if there are any problems with the placenta it is likely that the fetus will have long term problems as well. We think that the vascular network of the placenta reveals these problems long before the baby can be diagnosed



(a) Original (b) Green Channel (c) Contrast Stretched

(d) Threshold (e) Largest Object (f) Image Fill

(g) Erosion (h) Cropped (i) Deglared

Figure 4: Cropping and Deglaring Algorithm Process

for them. If we can properly analyze the vessel network and extract meaningful features from it we might be able to predict these future problems early on and get the baby better treatment.

4.1 Vessel Extraction

Vessel analysis in medical images is essential both for diagnostic and intervention planning purposes. The process of extracting placenta blood vessels is currently costly and labor-intensive; a human expert must hand-trace the vessels in each placenta image following a pre-defined protocol. Any misinterpretation of the protocol can result in a reiteration of the traces. In order to make the vascular network available independent of the tracer, an automated vasculature extraction algorithm is needed. However, the color information in blood vessel networks presented in placenta images appear to be low contrast and inconsistent. In addition, frequent vessel crossings and the inability to distinguish between arteries and veins also adds another layer of complication. This present a major challenge in the automated image segmentation and feature extraction architecture. To this end, we use a combination of a smoothing algorithm based on an anisotropic non-linear diffusion filter and an extraction algorithm based on road detection methods.

4.1.1 Smoothing

A vessel enhancing method based on second-order characteristics was first developed by Frangi [6] who used eigenvalues of the Hessian to determine locally the likelihood that a pixel is a vessel. For a 2D digital image $I(x, y)$, its Hessian matrix is

$$H(x,y) = \begin{pmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2}. \end{pmatrix} \tag{3}$$

Eigenvalues λ_1, λ_2 of H can then be used to define two *vesselness* measures suited for medical images. The first intensity-invariant measure accounts for the deviation from a blob-like structure but cannot distinguish between a line- and a plate-like pattern:

$$\text{(anisotropy)} \quad A = \frac{|\lambda_1|}{|\lambda_2|}, \quad \lambda_1 > \lambda_2 \tag{4}$$

This value is high if there is a big change in direction at the point of interest. The second intensity-invariant measure accounts for the fact that vessel structures are typically brighter than the background:

$$\text{(structureness)} \quad S = \sqrt{\lambda_1^2 + \lambda_2^2} \tag{5}$$

This measure will be low in the background where no structure is present and the eigenvalues are small for the lack of contrast. With these two measures, the probability of a pixel being a vessel is given by

$$\mathcal{V} = \begin{cases} \exp\left(\frac{-A^2}{2\beta^2}\right)\left(1 - \exp\left(\frac{-S^2}{2c^2}\right)\right) & \text{if } |\lambda_1| < |\lambda_2| \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where β and c are thresholds that control the sensitivity of the line filter to the measures. The Equation 6 is given for bright curvilinear structures. For dark objects the conditions of the eigenvalues (or the images) should be reversed.

An improved technique called Vessel Enhancing Diffusion algorithm (VED) was proposed by Manniesing et al. [9] to achieve better visualization results in vessel appearance. This method combines a smooth vessel filter based on a geometric analysis of the Hessian matrix with a non-linear anisotropic diffusion scheme and is essentially taking a continuous (smoothed) version of Frangi's work followed by a diffusion process. The basis of the method used is the one introduced by Weichert [12] and described in details by Kroon [8]. In short, the process takes on the following steps:

1. Calculate Hessian from every pixel of the Gaussian smoothed input image as described in Equation (3).

2. Gaussian smooth the Hessian to obtain

$$H_\sigma(x, y) = (G_\sigma \star I)(x, y), \tag{7}$$

   where G_σ is a Gaussian with standard deviation σ. Then calculate its eigenvectors and values: $(\mu_1, \mathbf{v}_1)$ and $(\mu_2, \mathbf{v}_2)$ with $\mu_1 \geq \mu_2$.

3. The eigenvectors are used to construct diffusion tensor

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} \tag{8}$$

   used in the diffusion tensor filtering equation [3]

$$L_t = \nabla \cdot (\mathbf{D} \nabla L) \tag{9}$$

   where

$$\begin{aligned} D_{xx} &= \lambda_1 v_{11}^2 + \lambda_2 v_{21}^2 && (10) \\ D_{xy} &= \lambda_1 v_{11} v_{12} + \lambda_2 v_{21} v_{22} && (11) \\ D_{yy} &= \lambda_1 v_{12}^2 + \lambda_2 v_{22}^2 && (12) \end{aligned}$$

   with

$$\begin{aligned} d &= \mu_1 - \mu_2 && (13) \\ \lambda_1 &= \alpha + (1 - \alpha) \exp\left(\frac{-C}{d^{2m}}\right) && (14) \\ \lambda_2 &= \alpha. && (15) \end{aligned}$$

   Constants which determine the amplitude of the diffusion smoothing in Weichert [12] equation are C, m, and α, which are defaulted to 10^{-10}, 1, and 0.001, respectively, in the MATLAB implementations given by [8]. Vascular structures will be enhanced by evolving the image according to the diffusion equation 9.

4. A finite difference scheme is used to do the diffusion. We adopted an optimized diffusion scheme proposed in [8].

5. Back to Step 1. until a certain diffusion time is reached. We used 20 iterations in our empirical calculations.

4.1.2 Extraction

Existing road detection techniques are limited to fixed-width rectangular regions in between two homogeneous regions of different intensity levels on either side. Almoussa et al. proposed a novel road detection technique to achieve better results on the placenta vessel structures [10] from Porikli's directional line filter [11]. This modified filter is capable of identifying vessel structures of varying width, which is commonly seen in blood vessel networks. The filter is given by

$$g_\theta(i,j) = \cos\left(\frac{\pi i'}{2M}\right)\cos\left(\frac{\pi j'}{M}\right)e^{\frac{-(i^2+j^2)}{2(cM)^2}}, \tag{16}$$

where M is a fixed window size, $i, j = -M, \ldots, M$, and i', j' are coordinates rotated by θ. It was shown in [10] that $c = 0.9$ is the most effective for reducing noise while still retaining all the vessel information. Other implementation details can also be found in [10]. The enhanced road detection filter works well on placenta images with darker blood vessels than their neighboring tissue and fails on blood vessels that are softer and brighter than their surrounding tissue. It is worth noting that no single detector could sufficiently and reliably identify all blood vessels exhibited in 2D color images produced under the current protocols. Ongoing efforts are devoted to standardize image acquisition protocols to provide researchers a consistent platform for testing their methods.

Figure 5(a) shows a sample placenta image while Figure 5(b) and (c) illustrate the result of the VED filter and enhanced road detection filter described above.

5 Graph Extraction

Weighted and unweighted graphs of the vessel structure were obtained for both raw and traced images. The initial extraction works similarly once both types of images are skeletonized, however additional refinements are possible on the traced images because a tree structure can be assumed.

5.1 Skeletonization

The first step in graph extraction is skeletonization, in which shapes are reduced to a single pixel thick centerline representation of themselves. This representation, the medial-axis, is the locus of the center-points of all circles tangent to the shape at two points. The medial axis was calculated using a brush-fire technique included in Matlab, which iteratively strips away a one pixel layer from all sides of the image until no more pixels can be removed without breaking connectivity.

The medial-axis transform leaves behind branch structures at the endpoints of elongated shapes such as vessels. These were removed by deleting N endpoint pixels from each endpoint of the resulting structure.

(a) Original placenta image (b) VED filtered image (c) Road detected image

Figure 5: Vessel enhancing diffusion and road detection results: (a) A sample placenta image preprocessed with our cropping and de-glaring algorithm; (b) result of running 20 iterations of VED filter; and (c) result of running enhanced road detection filter.

5.1.1 Traced Image Preprocessing

Traced images were preprocessed before skeletonization in order to correct for human error in tracing. A single pixel disk erosion was performed over all pixels except those belonging to branches of two pixels or less in thickness, as determined by tracing color in the tracing protocol. This separated individual vessels that overlap briefly due to thick pen sizes.

5.2 Initial Graph Extraction

After skeletonization, a weighted and an unweighted adjacency matrix on all on pixels was calculated. Adjacency was defined to be mixed adjacency.

Definition. Mixed Adjacency (Gonzalez and Wood) [7] Let V be the set of all on pixels in a binary image, I. Define the following three neighborhoods of a pixel, p at coordinates (x, y) in I:

$$N_4(p) = \{(x+1, y), (x-1, y), (x, y+1), (x, y-1)\} \quad (17)$$

$$N_D(p) = \{(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)\} \quad (18)$$

$$N_8(p) = N_4(p) \cup N_8(p). \quad (19)$$

Define a binary, symmetric relation, $\sim$, on V as follows:

$$\sim = \{(p, q) | q \in N_4(p) \text{ or } (q \in N_D(p) \text{ and } V \cap N_4(p) \cap N_4(q) = \emptyset)\} \quad (20)$$

From this definition arises weighted and unweighted adjacency matrices

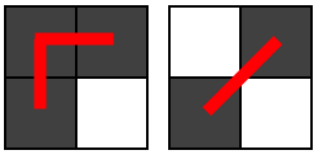


Figure 6: A visualization of mixed adjacency. On pixels connected by a red line are considered adjacent.

Definition. Adjacency Matrices

$$G_U(p,q) = \begin{cases} 0 & \text{if } p \not\sim q \\ 1 & \text{if } p \sim q \end{cases} \tag{21}$$

Weighted Adjacency Matrix

$$G_W(p,q) = \begin{cases} 0 & \text{if } p \not\sim q \\ 1 & \text{if } p \sim q \text{ and } q \in N_4(p) \\ \sqrt{2} & \text{if } p \sim q \text{ and } q \in N_D(p). \end{cases}$$

(22)

5.3 Refinement on Traced Images

Since the vessel structure is necessarily a tree it is possible to refine the graph by assuming the tree is binary, excepting one high-degree root vertex at the umbilical cord. This assumption is generally valid as arterial branching tends to occur on only one side of a vessel at any point.

5.3.1 Locating Crossing Pairs

Vertices in the adjacency matrix are classified by degree. If a vertex has degree 1 it is a leaf, degree 2 it is a trivial vertex, degree 3 it is a branching, and higher degrees are splittings. Trivial vertices are so named because they exist in the middle of vessels, between branches and endpoints. Some branchings are not actually part of the same vessel in the placenta, but rather the result of two vessels overlapping. This appear as two close together branchings in the graph. These are called crossing pairs. To find crossing pairs, pairs of branches within a threshold distance, T_c of each other were considered, and excluded if they were too close to the umbilical cord as marked on the tracing.

Definition. Crossing Pairs The set of all crossing pairs, P, is given by

$$P = \{(u,v) \in B^2 | R(u) \notin N(C), R(v) \notin N(C), ||R(u) - R(v)|| \leq T_c\} \tag{23}$$

where $N(C) = N_{\sqrt{A_C}}(C)$, the neighborhood of radius $\sqrt{A_C}$ about C, the centroid of the umbilical cord dot, A_C is the area of the umbilical cord dot, and T_c is the crossing pair threshold.

5.3.2 Bridging Crossing Pairs

In order to prepare the graph for further processing, crossing pairs must be removed in a process called bridging, in which the vertices of the crossing pair and all vertices between them are removed, and the new leaves created by this process joined in a manner consistent with how the actual vessels align in the image.

Let $\{p_n\}$ denote the vertices in the shortest path between u and v, excluding u and v, and L denote the weighted length of the entire path. Delete $\{p_n\}$. Denote the four remaining neighbors of u and v as u_1, u_2, v_1, and v_2. Define the following four angles and four distances:

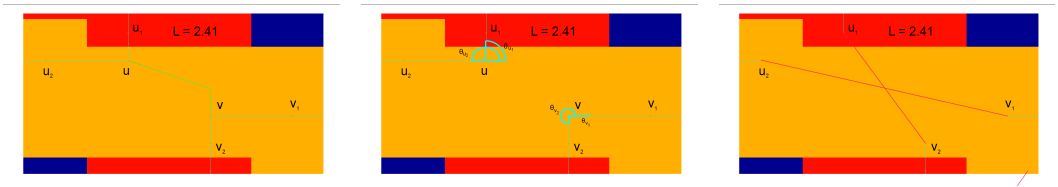
- $\theta_{u_1} = \measuredangle(u_1 - u)$
- $d(u_1) = ||R(u_1) - R(u)||$
- $\theta_{u_2} = \measuredangle(u_2 - u)$
- $d(u_2) = ||R(u_2) - R(u)||$
- $\theta_{v_1} = \measuredangle(v_1 - v)$
- $d(v_1) = ||R(v_1) - R(v)||$
- $\theta_{v_2} = \measuredangle(v_2 - v)$
- $d(v_2) = ||R(v_2) - R(v)||$

Now delete u and v. Compute the following two cost functions:

- Cis-Cost = $(\pi - |\theta_{u_1} - \theta_{j_1}|) + (\pi - |\theta_{u_2} - \theta_{j_2}|)$
- Trans-Cost = $(\pi - |\theta_{u_1} - \theta_{j_2}|) + (\pi - |\theta_{u_1} - \theta_{j_2}|)$

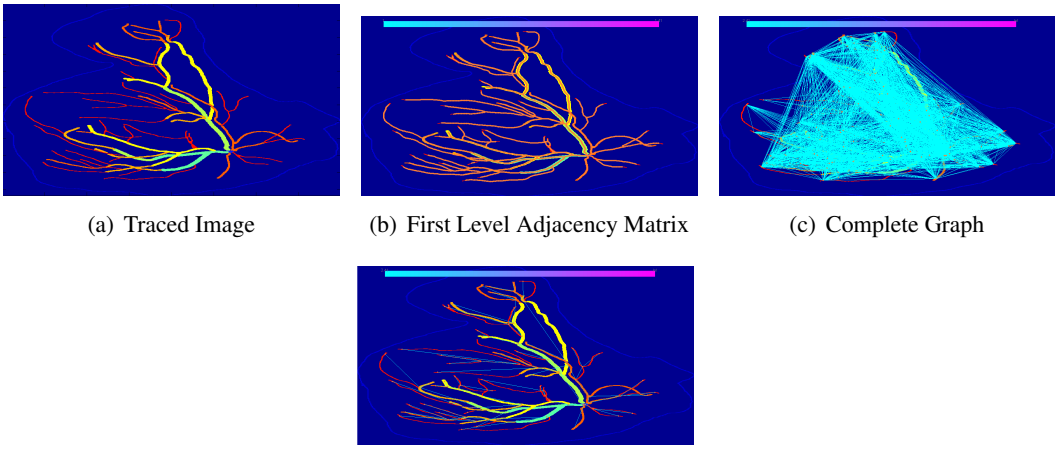
These are differences from a π radian difference in direction of the vectors pointing from u and v to the newly created leaves; how far off each of the two possible connection patters are from being straight lines across the crossing. The weighted adjacency matrix is then modified in one of two ways, depending on which cost function in smallest.

- If Cis-Cost $<$ Trans-Cost:
  - $G(u_1, v_1) = G(v_1, u_1) = d(u_1) + L + d(v_1)$
  - $G(u_2, v_2) = G(v_2, u_2) = d(u_2) + L + d(v_2)$
- If Trans-Cost $<$ Cis-Cost:
  - $G(u_1, v_2) = G(v_2, u_1) = d(u_1) + L + d(v_2)$
  - $G(u_2, v_1) = G(v_1, u_2) = d(u_2) + L + d(v_1)$



(a) Vertex Labeling (b) Angles Labeled (c) Bridged Crossing

Figure 7: Crossing pair bridging algorithm.



(a) Traced Image (b) First Level Adjacency Matrix (c) Complete Graph

(d) Final Graph

Figure 8: Graph Extraction Algorithm on a Traced Images

5.3.3 Enforcing Tree Structure

Once all crossing pairs have been bridged, compute the all shortest paths matrix for the weighted adjacency matrix using the Floyd-Warshall algorithm. Call this matrix S. Now define a new, complete graph on the non-trivial vertices of G as follows.

Definition. Complete Graph $G_c(u, v) = G_c(v, u) = S(u, v)$ where $\deg(u) \neq 2$ and $\deg(v) \neq 2$.

From here, find the minimum spanning tree of G_c using Kruskal's algorithm, and re-purpose the label G. This is valid because a tree structure for G was originally assumed, so there is only one path between any two vertices, which is the shortest path on the original full adjacency matrix. As a final step, merge the vertices in $V(G) \cap N(C)$, where $V(G)$ is the vertex set of G. This vertex represents the umbilical cord insertion.

6 Invariant Features

After converting all of the placental images into vessel images, we extracted features from them. Due to the images being different scales, we needed to extract features that were invariant.

6.0.4 Graph-based Feature

Definition. Vessel Length The total length of the vessels was calculated as the sum of all edges in the weighted adjacency matrix:

$$\sum_{(u,v)\in V\times V} \frac{||R(u)-R(v)||}{2} \tag{24}$$

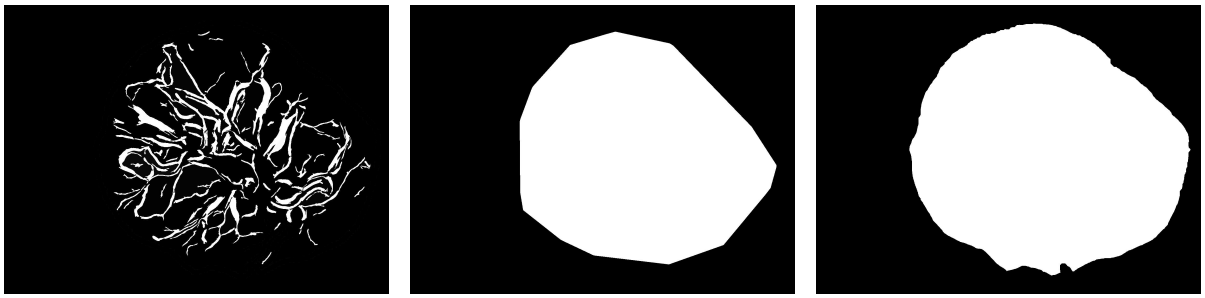
Definition. Number of Branches The number of branches in an image was defined as the number of vertices of degree 3 or greater in either the weighted or unweighted graph:

$$\sum_{\substack{v\in V,\\ \deg(v)\geq 3}} 1 \tag{25}$$

6.0.5 Vessel Features

We used several features related to just the vessels. Images of some of the areas we used to find these features can be seen in Figure 9

- Relative area of vessels $= \frac{\text{Vessel area}}{\text{Placenta area}}$.
- Area comparison $= \frac{\text{Convex hull area}}{\text{Placenta area}}$.
- Moments of distance to nearest vessel
  - We took the central moments of the distance from any pixel on the placenta to the nearest vessel.
  - Should tell us how spread out across the placenta all the vessels are.
- Images Moments
  - We found seven rotationally invariant moments as described in Flusser et al. [5] and Flusser [4].
  - These moments are based on the central image moments
  - Given central moments μ_{ij} we have new moments $\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{(1+\frac{i+j}{2})}}$
  - Here are all seven
    - $I_1 = \eta_{20} + \eta_{02}$
    - $I_2 = (\eta_{20} - \eta_{02})^2 + (2\eta_{11})^2$
    - $I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$



(a) Vessel Area (b) Convex Hull (c) Placenta Area

Figure 9: Some of the regions used for vessel features

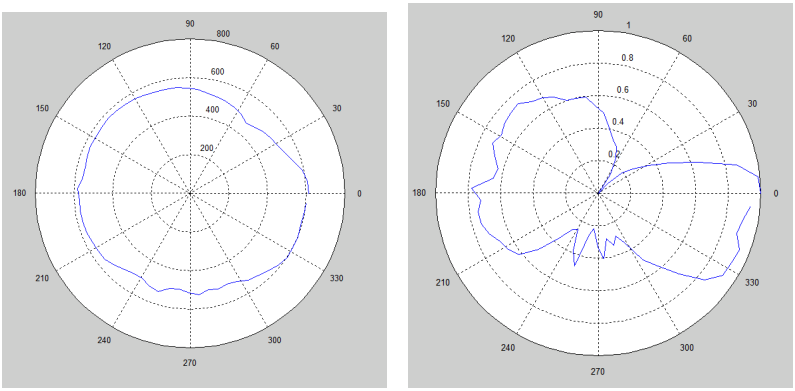
* $I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
* $I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$
* $I_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03}^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
* $I_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

- Length of vessels and number of nodes

6.0.6 Shape Features

Other features we based on the shape of the placenta edge.

- Vessel Area Roundness
  - $\frac{\text{Convex Hull Perimeter}}{\sqrt{\text{Convex Hull Area}}}$
  - Should measure how round the inner placenta is
- Placental Area Roundness
  - $\frac{\text{Placenta Perimeter}}{\sqrt{\text{Placenta Area}}}$
  - Measures how round the entire placenta is
- Boundary Features The boundaries of both the placenta and the convex hull of the vessels were parameterized as polar functions starting along the x-axis and moving counterclockwise is 0.1 radian increments. The origin point was chosen to be either the centroid of the placental area, or the nearest pixel on the shape to be parameterized to the umbilical cord location. For all of the following features, except where noted, the discrete function $x[n]$ is either the samples from the centroid or from the cord: most features were



(a) Un-Normalized (b) Normalized

Figure 10: Un-normalized and normalized placental boundaries, taken from the centroid.

calculated from both. In all but one case, which will be noted, the samples were normalized to lie in the range $[0, 1]$.

Definition. Fourier Coefficients The standard definition of the Discrete Fourier Tranform (DFT) was used to compute Fourier coefficients to represent the boundary of the placenta.

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-ik\frac{2\pi}{N}n} \tag{26}$$

Definition. Fourier Powers In order to retain rotation, scale, and translation invariance, only the real parts of the right-handed spectrum were considered. These Fourier Powers were given as follows:

$$c_k = \begin{cases} X(0) & \text{if } k = 0, \\ 2||X(k)|| & \text{if } k > 0 \end{cases}$$

Definition. Relevance Threshold Since cropping left many small oscillations in the boundaries, low amplitude oscillations were ignored. The threshold used to crop these out was taken to be the mode of the real amplitudes, rounded to 3 significant features.

$$T = \text{mode } \frac{\lceil 100 c_k + 0.5 \rceil}{100} \tag{27}$$

Definition. Complexity The number of relevant sinusoids needed to create the boundary was given as the placenta's complexity.

$$K = |\{k | c_k > T\}| \tag{28}$$

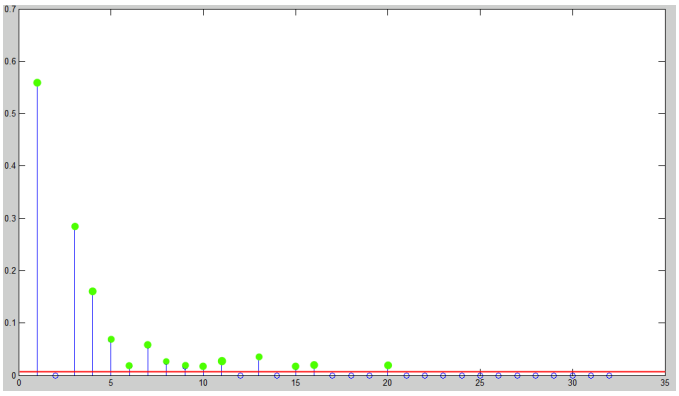


Figure 11: Right handed real spectrum with complexity and relevancy threshold marked.

Definition. Oscillation Feature The ratio of the period of the period of the boundary function (2π) to the period of the highest amplitude sinusoid in the boundary, is the oscillation feature, a measure of skewed-ness. This is computed as

$$F = \frac{2\pi k}{N} \tag{29}$$

where $k = \{n | c_n = \max\{c_m | m > 0\}\}$.

Definition. $|C|$ (Yampolsky et al.) In a paper by Yampolsky et al., the real amplitude of the first coefficient of the continuous Fourier transform of a hand created piecewise boundary function from the cord was found to correlate to birth weight.[13] The original implementation was given as

$$C = \frac{1}{\pi} \int_0^{2\pi} e^{-i\theta} r(\theta) \mathrm{d}\theta \tag{30}$$

We modified the equation to make a discrete version for quicker automation. It is given by

$$C = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-in} = c_{\lceil N/2\pi \rceil}$$

where $x[n]$ is the un-normalized boundary sampled from the nearest point to cord insertion. (31)

Definition. Dispersion Angle The angle at which the vessel network emerges from the umbilical cord is called the dispersion angle. This is the angle subtended in a circle around the umbilical cord by the convex hull of the vessels. The discrete implementation is given by

$$\sum_{\substack{n \\ v_c[n] > R}} \frac{2\pi}{N} \tag{32}$$

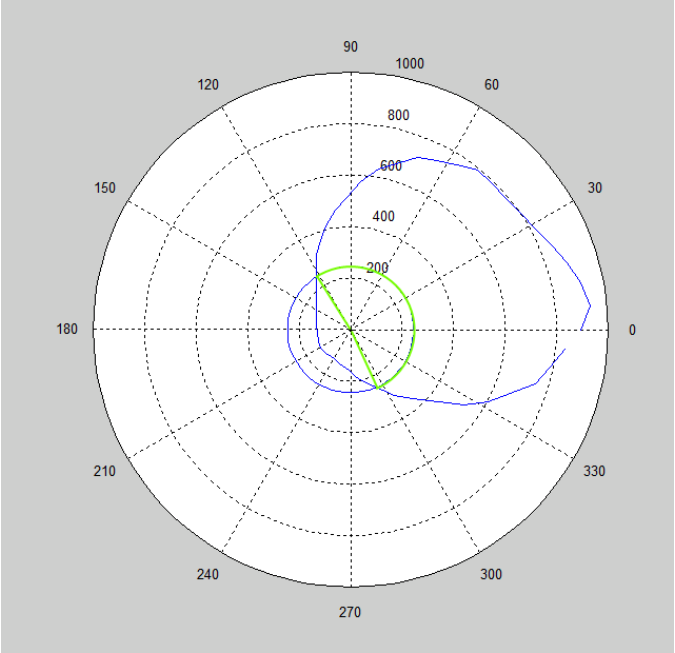


Figure 12: Geometric interpretation of the dispersion angle. The marked in green.

where $R = \frac{A_H}{2\sqrt{\pi}}$ is the radius around the cord, A_H is the area of vessel mask, v_c is the convex hull perimeter sampled radially from nearest point to cord insertion, and N is the number of samples.

7 Classifications and Results

After we extracted all of our features from the images, we wanted to try and classify the data into two groups.

7.1 Method

We attempted a fairly standard method of classification. We tried data reduction of several sorts, found a way to split the data into groups, and then did the same transformations to any new data. For data reduction we tried PCA, MDS, LLS and a few others. The following is a sample of the procedure used to classify.

1. Data reduction via PCA
2. Find optimal projection vector using LDA
3. Find a threshold value between the two classes

Look at Figures 13 and 14 for images of our data projected into two and three dimensions. Just looking at this told us our data was not going to classify well.

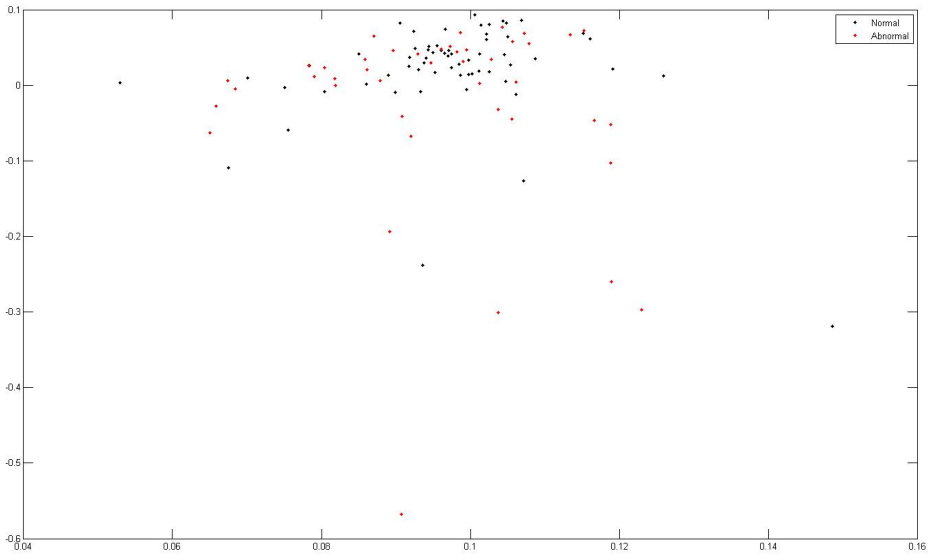


Figure 13: PCA Data projected into 2 Dimensions

7.2 Results

This did not end well. Running sample data through the classification resulted in a classification rate of 50%. Thus flipping a coin is a better option than trying to use our classification method. This may be do in part to the following reasons.

- It is possible that there is no difference between a placenta that gives rise to a healthy baby and one that does not. We sincerely hope this is not the case as otherwise the whole project is pointless.
- Our features might not have been the best ones to use.
- Our labeling of the data might have been in error. The results we got do seem to indicate a problem with the labeling of the data.

We still want to look into using Support Vector Machines to classify our data as it is a useful method and we have an excellent toolbox named LIBSVM [1].

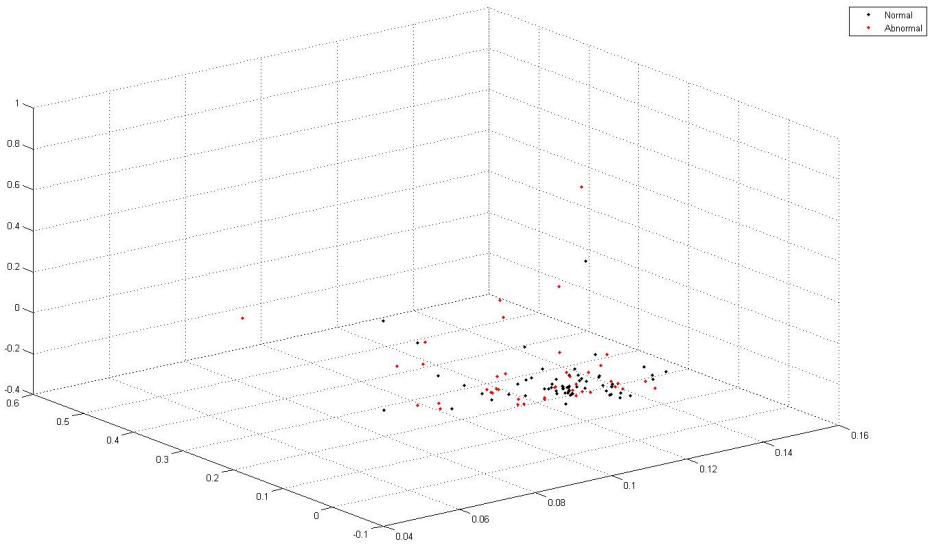


Figure 14: PCA Data projected into 3 Dimensions

8 Future Work

We would like to continue working on the following:

- Develop an interface for doctors to use.
- Improve method of extracting clear vessel images.
- Determine a better measure of 'at riskness'.
- More exploration into what the best features to use are.

Acknowledgement

We would like to thank Dr. Salafia of Placental Analytics for providing us our data. In addition we would like to thank the UCLA REU summer program for providing us the resources to do this study.

References

[1] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[2] Gary F. Cunningham. *Williams Obstetrics*. McGraw-Hill Publishing, 2005.

[3] Andinet Enquobahrie, Luis Ibanez, Elizabeth Bullitt, and Stephen Aylward. Vessel enhancing diffusion filter. Technical report, CASILab, The University of North Carolina and Kitware Inc., September 2007.

[4] Jan Flusser. On the independence of rotation moment invariants. *Pattern Recognition*, 33:1405–1410, 2000.

[5] Jan Flusser and Tomáš Suk. Rotation moment invariants for recognition of symmetric objects. *IEEE Transactions on Image Processing*, 15(12):3784–3790, December 2006.

[6] A. Frangi, W. Niessen, K. Vincken, and M. Viergever. Multiscale vessel enhancement filtering. In *Lecture Notes in Computer Science*, volume 1496, pages 130–137, Germany, 1998. Springer-Verlag.

[7] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Pearson Education, Inc., 2008.

[8] D. Kroon and C.H. Slump. Coherence filtering to enhance the mandibular canal in bone-beam ct data. In *IEEE-EMBS Benelux Chpater Symposium*, 2009.

[9] Rashindra Manniesing, Max Viergever, and Wiro Niessen. Vessel enhancing diffusion a scale space representation of vessel structures. *Medical Image Analysis*, 10:815–825, 2006.

[10] Nizar, Brittany Dutra, Bryce Lampe, Pascal Getreuer, Todd Wittman, Carolyn Salafia, and Luminita Vese. Automated vasculature extraction from placenta images. Technical report, UCLA, Department of Mathematics, August 2009. Applied Mathematics REU.

[11] Fatih Porikli. Road extraction by point-wise gaussian models. In *SPIE AeroSense Technologies and Systems for Defense and Security*, pages 758–764, 2003.

[12] Joachim Weichert. A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *Visual Communication and Image Representation*, 13:103–18, 2002.

[13] M. Yampolsky, C. Salafia, O. Shlakhter, D. Haas, B. Eucker, and J. Thorp. Centrality of the umbilical cord insertion in a human placenta influences the placental efficiency. *Placenta*, 30:1058–1064, 2009.