Contents

1 Introduction

The human skull is made from multiple plates (bones), connected together to form a solid structure. The connections between those plates are called cranial sutures and are located at the edges of the separate bones which form the skull. Throughout physical development, they function as bone growth sites in response to the separating forces from the growing brain requirements.

At an early stage of life, to allow for this growth, the sutures are wide and filled with a flexible soft tissue. As the individual grows up to adolescence, the plates of the skull fuse together, forming a much narrower suture path and providing the skull bones with a more solid connection over time. This change in width and shape of the sutures is directly dependent on aging, as it continues throughout the life of the developing human body.

Forensic scientists and anthropologists have been using this dependency: to assess the degree of fusion along the sutures and estimate the Age-At-Death (AAD) of the skull's holder. For many years, assessing the suture has mostly been conducted on the surface level only, due to the lack of non-invasive methods for accessing the suture cross-section.

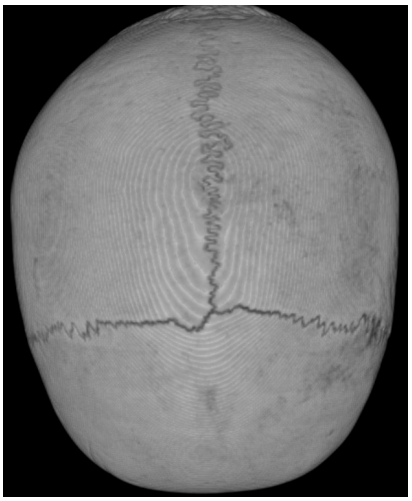
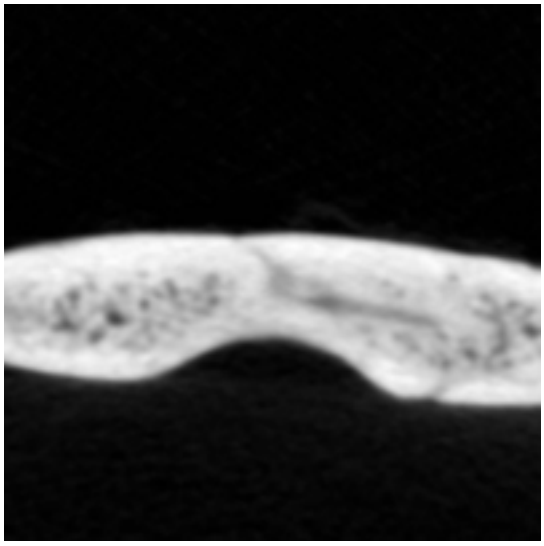


Figure 1: Skull surface view

This is a major limitation, as suture ossification often appears firstly in the core of the bone rather than the surface, therefore by not considering the cross-section of the suture, a lot of useful information about the degree of fusion can remain unrecorded, causing inaccuracy in the produced age estimates.

Nowadays, the increasing availability of CT scanners allows for an inside view of a skull's structure. This type of in-depth (cross-sectional) analysis of a suture is a new approach, and only three studies have been conducted on the topic. In two of the studies a human is responsible for evaluating the level of bone fusion along the suture, which is incredibly time consuming, subjective and therefore, sometimes inaccurate. There has been only one attempt to automatize this process with the use of a simple algorithm. The results were not able to prove any correlation between the degree of suture fusion and AAD, as the suture detection algorithm lacked the ability to consider many of the features which a suture can have. This flaw led to misclassification of the degree of fusion in many cases. This image is a good example of a miss-classified suture:



In the image, the suture can be seen as a crack passing through the bone cross-section. The algorithm's output, describing the degree of suture fusion in study [1] is only binary. On this image, the output would be '0', corresponding to a fused suture. While the suture is very vague, it can still be seen that the bones aren't completely fused, as the suture is still distinguishable. To determine how far from complete fusion the bones are, we need to consider how close to white the color intensity of the suture is. A simple binary output lacks the ability to describe such information, causing the previously used algorithm to perform poorly in the task of suture assessment.

Even if a professional were to analyze the degree of suture closure on this image, while distinguishing the suture, they could never tell exactly which shade of gray along the *1 - 255* grayscale color spectrum corresponds to the mean color of the particular suture. This is important information, as a small difference in this color intensity could mean a slightly higher or lower degree of bone ossification. Those bits of extra information on a certain image, like the exact color intensity of the suture or the average thickness of the suture, and many others, have been neglected by all currently known techniques for age estimation. Such measurements are impossible to do by the naked human eye, but relatively easy to calculate using modern computer algorithms. Considering this extra information can be key to improving the accuracy of age estimation based on cranial sutures, with real world applications in multiple fields including archaeology, criminology and anthropology.

This project addresses the lack of accuracy and up-to-date technology in the already existing methods, by elaborating an algorithm for more accurate

and objective AAD estimation, based on assessment of the suture cross-section with the help of mathematics and machine learning.

2 Project process

2.1 Image generation

Typically, methods for assessment of the suture cross- section in previous studies, extract slice data by manually adjusting the intersection of planes on the 3D skull model, using an imaging software. The purpose is to correct the skull incline, so that the intersecting plane cuts exactly perpendicular to the suture surface. This results in the most accurate cross- sectional view. It is important for the skull to be at a correct incline, otherwise a distorted and prolonged suture slice is produced, making the image unreliable for assessment. This is a comparison between intersecting planes producing a valid and a invalid cross- sectional suture slice. The images are produced by the intersection of the blue plane with the skull:

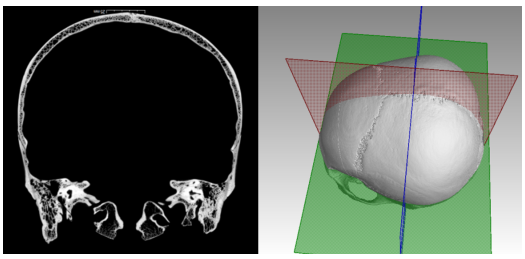


Figure 2: Perpendicular cross-sectional slice

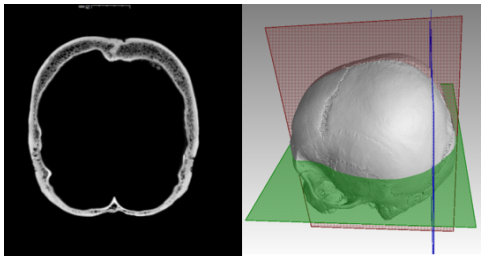


Figure 3: Prolonged cross-sectional slice

As can be seen, the suture on the right image appears to be very long, when in fact it is not. This is deceiving and can not be used for reliable assessment. Furthermore, there are several issues regarding this method for data extraction:

1. It is incredibly time consuming to do on bulk quantities of skull data, since the observer has to manually adjust the skull rotation for every image along the suture length.

2. It requires very expensive hardware. To simply load an entire CT scan of a skull into a 3D imaging software like this one, requires approximately 64 GB of RAM memory. Furthermore, to run the imaging

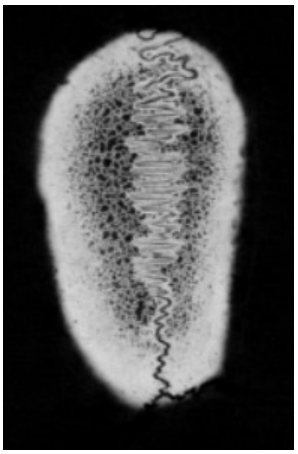
program without experiencing significant lagging, a very powerful video card is necessary.

3. The observer is entirely responsible for ensuring that the intersecting plane is perpendicular. This is subjective and can often lead to the plane not being truly perpendicular and therefore, can cause some distortion in the produced cross-sectional slice.

The first part of the project is focused on introducing an automatic and independent of the observer's opinion, method for cross-sectional data extraction. The introduced technique will be used on large quantities of skull data and produce the suture images needed for the AAD prediction further on.

2.1.1 Loading skull volume

A typical skull dataset consists of multiple horizontal slice images at different heights along the whole skull. This example shows how the suture appears on the skull surface, intersected by a horizontal slice:



Clearly, this image is not the cross-sectional view we're looking for. A cross-sectional slice of the bone at a certain point is defined as an image, perpendicular to the surface at that point. This is not the case with the top of the skull and the horizontal slices we are provided with.

To extract cross-sectional suture slices, the algorithm generates a 3D model of the skull, for the user to mark the path of the suture on. To generate the model, I first had to load the image volume in the Matlab workspace. The

workspace uses the RAM memory of the PC to store the loaded variables, in this case the loaded image volume. The first problem I encountered was that the RAM memory required to completely load a typical skull dataset was about 50 GB. My PC has only about 14 GB of usable RAM memory.

As a solution, I figured I should load a down-scaled version of the volume, instead of a full sized one.
Take every horizontal slice from the skull volume with dimensions:

$$Width \times Height$$

$z \in (0, 1]$, where z is a variable depending on user preferences

Resultant image with dimensions:
$$(Width \times z) \times (Height \times z)$$

I also took the scale factor into consideration with the number of images in the dataset. Instead of loading every single image, I chose to load every $\frac{1}{z}$ th image. This way, the dimensions ratio of the produced downscaled 3D model remain consistent with the full-scaled version.
Typical dimensions of a fully sized skull volume are 2190 x 1840 x 2130. The first two dimensions represent the size of each image, and the third is the number of images in the dataset. Displaying such large volumes as graphical object on the screen can be a computationally heavy task, so by down-scaling I improved the smoothness and significantly reduced the lag while working with the 3D model.

2.1.2 3D model generation

After having loaded the volume of images, I make a copy of the same volume and run it trough the *binarizeVolume()* function. This function binarizes and fills the holes in every image. Results of this process can be seen bellow.

Figure 4: Unprocessed image

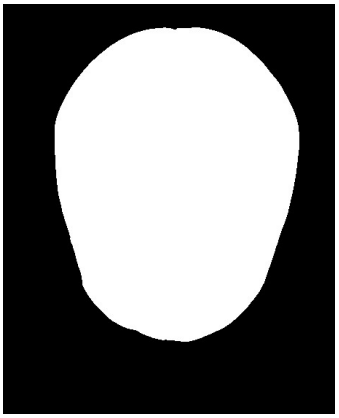


Figure 5: Processed image

The processed volume is then used to generate the 3D model in the *generateVolume()* function, by vertically stacking the details from each horizontal slice.

The advantage of processing the whole volume like this, is that unnecessary details are removed from each image. In this case, the only thing we're interested in seeing on the model, is the very outside surface of the skull. After processing, the very outside surface is the only detail preserved of the entire image; as seen above. This suits the needs of the project perfectly and significantly reduces the time taken to generate the 3D model, as the used volume is very simplified after processing. Having generated the 3D model of the skull, the program relies on the user to define the places on the surface through which a suture passes.

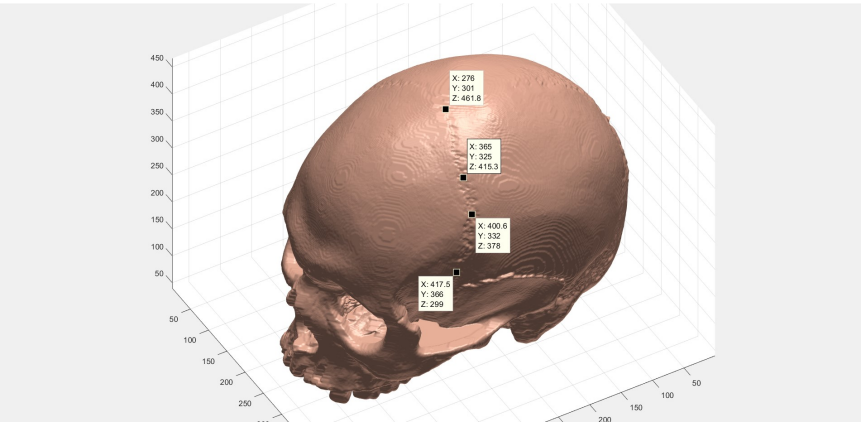
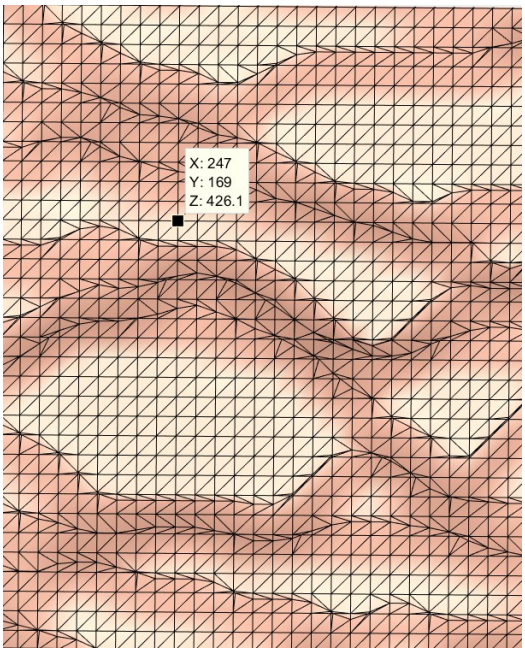


Figure 6: Suture path definition

2.1.3 Path finding

The next task is to generate the path that the suture follows from the given defining points.
Each point on the surface, on the very low level, is part of a triangular mesh.



The question is, which path of connected points on this mesh do we choose to get from the start to the end point? Now this is a question with multiple solutions. Ultimately, this triangular mesh can be taught of as a weighted graph with the weights corresponding to distance between connected vertices. A number of graph search algorithms can be used to find the smallest cost path between two points. Most algorithms like Djikstra or the A* search algorithm may need to go through a big part of the graph before finding an optimal solution. Considering the fact that the number of points on the triangulated mesh of the skull is in the millions range, this process can be computationally and time expensive. For this reason, a computationally cheaper, geometric approach has been used instead.

Let's first look at the simplest case with only three defining suture points. A line is draw between the start point A and the end point B. The foot of the normal (the blue line) to the vector $\vec{AB}$ from the middle point is found and recorded as the point M.

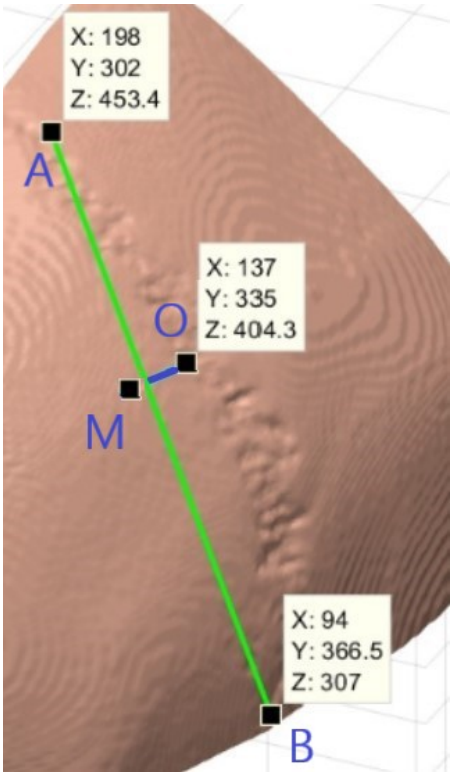


Figure 7: Suture path

Let's call the vector connecting the foot of the normal M and the middle point O a *compare vector*. This is represented by $\vec{MO}$

$$\overrightarrow{MO} \perp \overrightarrow{AB}$$
$$\overrightarrow{MO} = ?$$

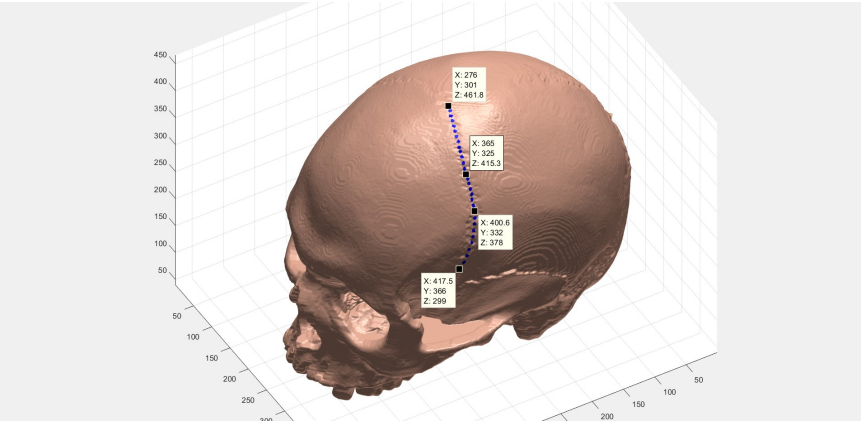
$$\left|\overrightarrow{AM}\right| = \frac{\overrightarrow{AO} \times \overrightarrow{AB}}{\left|\overrightarrow{AB}\right|} \qquad \frac{\left|\overrightarrow{AM}\right|}{\left|\overrightarrow{AB}\right|} = C$$

$$M = A + C \times \overrightarrow{AB}$$
$$\overrightarrow{MO} = O - M$$

Starting the path search from the top point A, we are now looking at which of the connected to that point vertices to go to next. We find the normal to the $\vec{AB}$ for each connected vertex. The next chosen point is the

point which has the normal making the smallest angle with the "compare vector". The process is repeated until the end point is reached.

For more complicated cases with multiple defining points, every three neighboring points are treated like separate cases of the described above. As the algorithm reaches the end point of one such triplet of points it starts to follow the normal of the next triplet. Each chosen path point is saved in an array, and after we find a path for the suture, we take j equidistant points from this array. j is the number of cross-sectional images we want to generate for the specific suture. The blue marks on the skull surface represent the points of image generation.



2.1.4 Surface normal calculation

The next step is to determine at what angle to the skull surface we want to generate an image. Let's call the point on the surface for which we're calculating the normal of the tangential plane, a ***desired point***. For the purpose, we take the N closest to the ***desired point*** points from the triangular mesh.

Next, we find all of the triangles from the mesh which the N points are a part of. The normal vector of each triangle from the mesh is calculated by taking the cross-product of the vertex vectors and multiplying it by the area of the triangle. Every normal vector is ensured to be pointing outwards to the surface, by comparing the angle that the two possible triangle surface normals make with the vector to the skull mid point. The vector with larger angle to the skull mid point is chosen as the normal vector for the triangle. This process is repeated for very triangle in the N nearest points radius. The mean value of all calculated normal vectors is taken as the best normal approximation of the tangential to the surface plane at the ***desired point***.

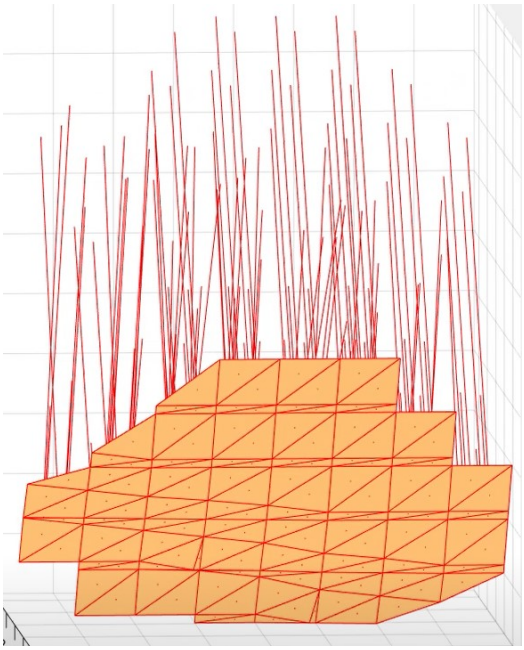


Figure 8: Triangular vector mesh for computing a single surface normal

The value of N depends on user preferences, but experimentally it has been determined that a value of

$$N = 20$$

is sufficient to produce an accurate surface normal estimation. The result, is an average vector which is a very close to being perpendicular to the skull's surface. The normal at each image point of image generation is calculated individually using the described method.

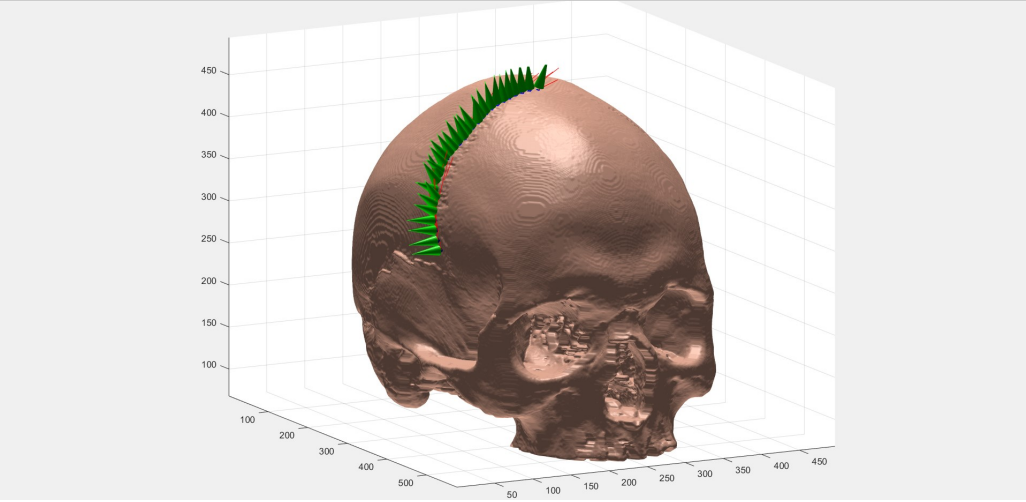
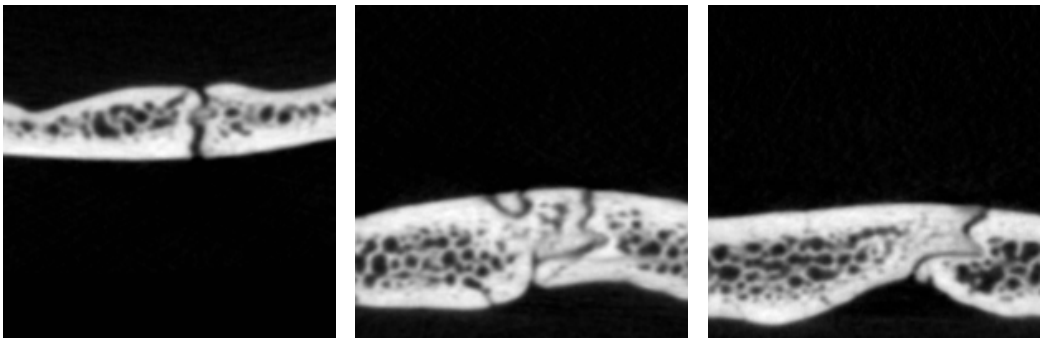


Figure 9: Calculated surface normals

A plane, parallel to the normal vector, will be used to generate each cross-sectional image. The plane extracts image data from a chunk of the fully sized volume, as if it extracted the data from the down-scaled version, image quality would degrade. The function *generateSlices()* automatically loads the volume in chunks, with a size respective to the currently available RAM memory.
To estimate color values at non-integer vertex points, cubic interpolation is used. The generated images are exported to a specified folder.

Examples of generated images



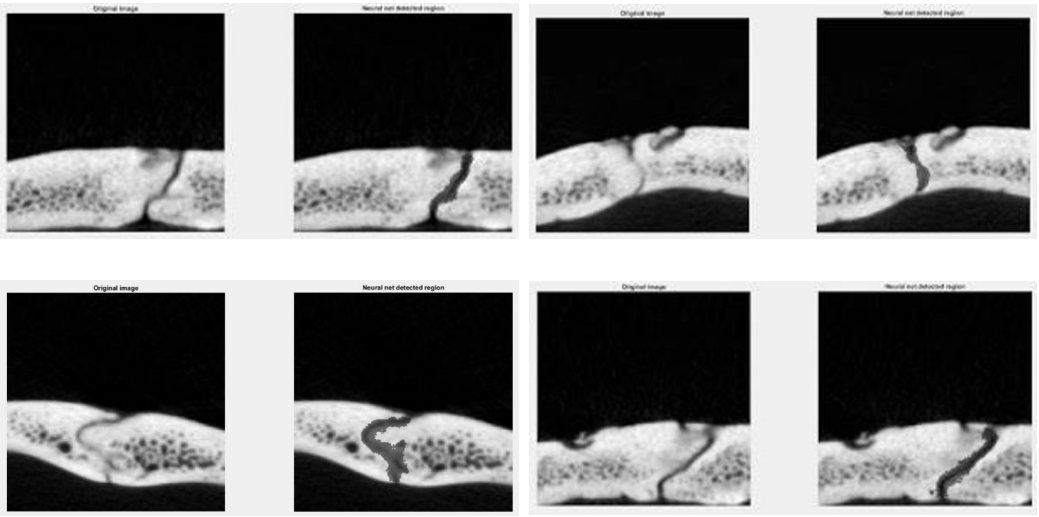
2.2 Suture assessment

This section will describe the process of assessing the degree of closure on the generated cross-sectional images.

2.2.1 Suture detection

In order to analyze the degree of suture fusion, we firstly need to determine the region of the image trough which the suture passes. For the purpose, I'm using a VGG-16 Semantic Segmentation network for precise suture detection. I've hand labeled about 7000 suture images to train the network, which achieves an accuracy of about 91% in the suture segmentation. The input size of the network is $192 \times 192 \times 3$

Examples of sutures detected by the neural network



2.2.2 Suture processing

As can be seen, some of the white pixels on the suture boundary are also classified as part of the suture by the neural network. To ensure that the detected area will only contain the darker suture region, I filter the lighter pixels using triple variable C-Means clustering. The variables considered are the XY coordinates of the pixels, and their color intensity as the third variable. The fuzzy C-Means algorithm splits the pixels into two cluster groups consisting of dark and white pixels. The *Gray Ratio* metric is measured with the help of this color separation.

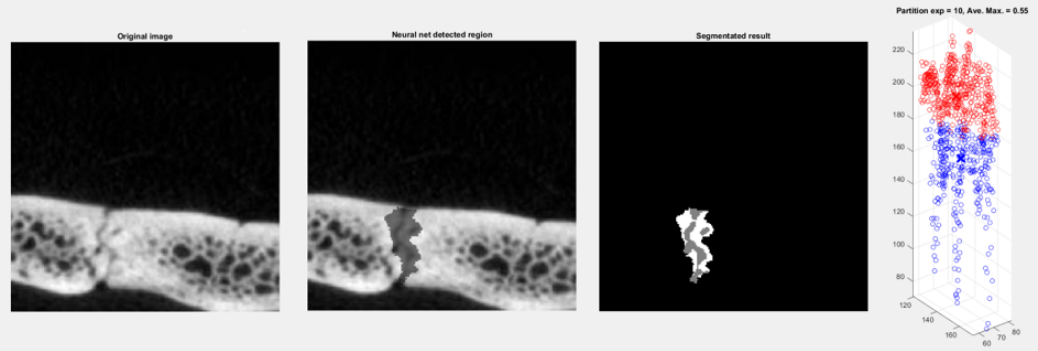


Figure 10: Suture region clustering

Number of gray pixels = *G*
Number of white pixels = *W*

$$\text{Gray Ratio} = \frac{\text{G}}{\text{G} + \text{W}}$$

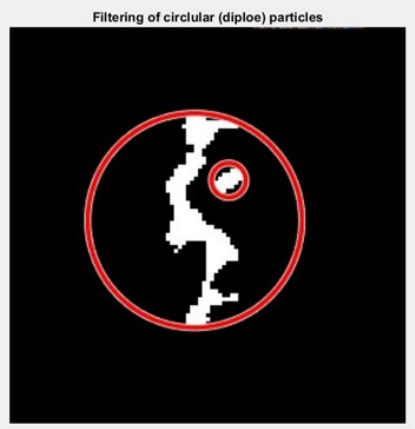
If the degree of suture fusion is low, the suture gap would be wide and result in the majority of pixels being classified as *gray*. This would make the *Gray ratio* closer to 1. Respectively, a very fused suture would result in this metric being closer to 0, due to the lower number of pixels being classified as *gray*.
This clustering is repeated twice, by clustering the *Gray* classification result from the previous time. This ensures a maximum filtration of white particles, and provides several readings for the *Gray ratio* metric.

2.2.3 Diploë filtering

Bone diploë particles are present in the entire skull cross-section. A characteristic feature of their color is that it's very similar to the suture color. This opens the possibility of diploë particles to be classified as being part of the suture, as they share a lot of common characteristics, if positioned near the suture path.

Figure 13 shows a good example of bone diploë being classifies as a suture (The gray colored particle to the right of the suture path is diploë). Our goal is to measure properties of the suture, so a dark particle in the segmentation can be a distraction for the produced results. To account for this error, I've taken advantage of the circular nature that the diploë particles have. For each of the segmented elements I calculate the minimum boundary circle

around it. The ratio of the area of the particle to the area of the circle is calculated.



$$\text{Particle Circularity} = \frac{\text{Particle area}}{\text{Minimum bound circle area}}$$

The more circular a particle is, the more likely it is to be diploe and not part of the suture. Particles with circularity larger than a certain threshold are tough of as diploe and discarded, like the particle in the smaller circle on this image.

2.2.4 Metrics measurement

After the clear segmentation of the suture and the removal of any miss-classified particles, the algorithm measures a number of important metrics used to describe the degree of suture fusion. To examine each skull I used 350 cross-sectional suture images along the length of the sagittal suture. Throughout the images, I measure the standard deviation, the median and the mean of each of the following metrics:

1. The average color intensity of the suture region marked as the *Color coefficient* metric.
   A fused suture consists of mostly solidified bone, which is white and has a high color intensity. This will increase the average, making this measurement a useful indicator of bone fusion

2. The average thickness of the suture, measured in pixel units and represented by the *suture thickness* metric.
   An ossified suture has a much narrower path than an open one, so this also gives useful information.

3. The depth of the suture, measured as a fraction from the whole bone cross-section. This metric is marked by *Cross ratio*. This image is a good representation of how the *Cross ratio* is measured.

$$\text{Cross Ratio} = \frac{\text{X} + \text{Y}}{\text{Z}}$$

2.3 Data filtering

I have collected and recorded the described above metrics for each of the skulls in the population, by assessing 350 cross-sectional suture images for every individual. To ensure a minimal error in the measurements, I filtered out any extreme metrics by using the interquartile range method for outlier filtering.

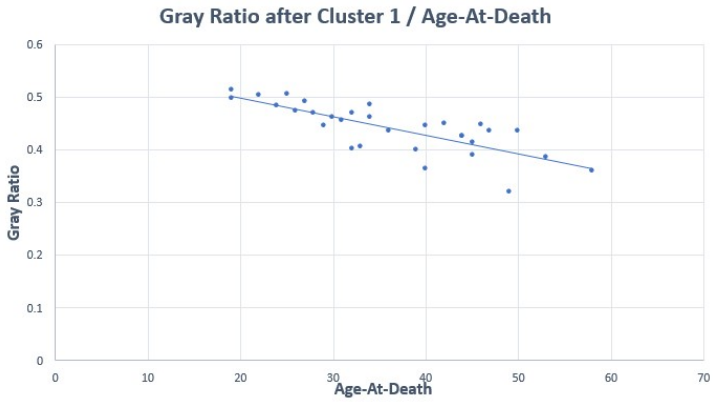
3 Project results

The skull sample is split into two groups - the analysis and the test group.

1. The analysis group is used for statistical analysis and the creation of a regression equation describing the AAD variable in terms of the algorithm's measurements. This group has a size of 31 individuals from various age groups.

2. The test group is used to test the regression equation independently and confirm it's accuracy. This groups has a size of 11 individuals from various age groups.

I've tested the correlation between the Age-At-Death (AAD) of individuals in the sample and the *Gray Ratio* metric. The following hypothesis test is conducted on the analysis group:



p – population correlation coefficient
r – Pearson's Product Moment Correlation Coefficient
H_0: $p = 0$
H_1: $p < 0$

Critical value for n=31 : -0.4421

1-tailed test, at the 0.5 % significance level

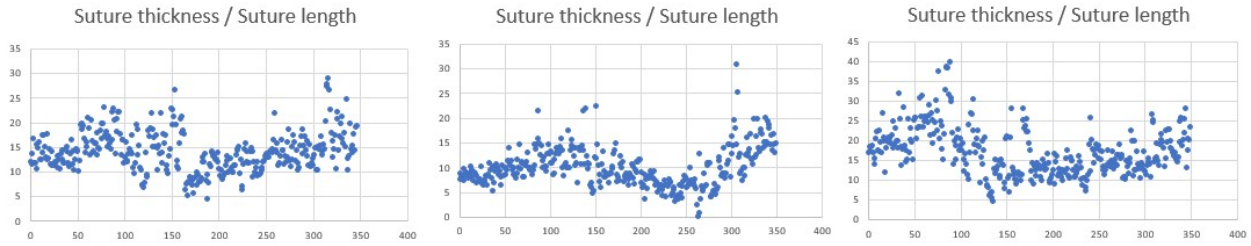
$$r = r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{(\sum x_i^2 - n\bar{x}^2)}\sqrt{(\sum y_i^2 - n\bar{y}^2)}}$$

$r = -0.78589$

$-0.7859 < -0.4421$, reject H_0, accept H_1

The currently produced results already show a statistically significant relationship between suture fusion and AAD, which has never been achieved by an algorithm before. Interpreting this result implies that by considering a sample of 31 individuals, we can still be 99.5 % certain, that there is a statistically significant correlation between the *Gray ratio* metric and AAD in the parent population of skulls.

This is a graph representation of the variation in the suture thickness metric along the length of the sagittal suture, measured for three different individuals from the sample. The Y axis corresponds to the Suture thickness value, and the X axis is the image index along the suture length:



The Suture thickness measurement along the length of the suture, produces a very consistent pattern. On all of the graphs, the "Suture Thickness" tends to increase towards the end of the suture length. This common graph pattern, further confirms the validity of the chosen segmentation technique, as this is exactly the type of patter one would expect the see on a manual suture inspection; the bone ossification is the lowest (corresponding at a high Suture thickness) towards the end of the suture length, as this is the place where the sagittal suture connects with the lambdoid suture. This graph comparison is an important analysis of the data, as it shows that the

segmentation technique is working in the right direction by confirming the known nature of bone ossification throughout the suture length.

Furthermore, I've used the analysis group to create a regression equation describing the AAD of an individual in terms of the algorithm's measurements.

Predicted age = 44.2 x Cross Ratio St. Dev - 118.3 x Cluster1 Gray Ratio St. Dev - 49.3 x Cluster 1 Median + + 194 x Cluster_C Gray Ratio Mean + 200.7 x Cluster_C Gray Ratio St. Dev - 134 x Cluster_C Median

The regression equation has been tested on the test group, which is entirely independent of the skull sample used to create the regression. This ensures the produced accuracy is not due to fitting the equation to the particular dataset. This is a summary of the generated predictions:

			Cross Ratio	Cluster1 Gray Ratio		Cluster_C gray ratio		
Actual Age-At-Death	Predicted Age-At-Death	95 % confidence Interval	St.Dev	St. Dev	Median	Mean	St.Dev	Median
49	47	(40.7405, 53.6866)	0.22739396	0.099844	0.380241	0.856502	0.099382	0.881188
44	45	(40.8906, 49.9918)	0.176391107	0.077687	0.42509	0.835647	0.101114	0.853123
37	40	(37.0005, 44.0575)	0.178441142	0.073199	0.426923	0.768825	0.083975	0.771503
54	52	(44.1675, 59.9964)	0.201911355	0.117068	0.358995	0.751608	0.173242	0.788094
36	39	(33.4420, 45.5139)	0.153593971	0.102141	0.365664	0.81217	0.073751	0.81548
45	44	(39.9980, 49.5191)	0.148925198	0.069709	0.422	0.83743	0.096184	0.852544
30	33	(17.3707, 27.1809)	0.128595576	0.090553	0.486642	0.827943	0.066691	0.827766
19	22	(17.3707, 27.1809)	0.03458675	0.050115	0.505019	0.701823	0.047405	0.700342
37	34	(29.9420, 38.6077)	0.068349185	0.062609	0.464559	0.78517	0.077483	0.791293
23	25	(18.9905, 31.5239)	0.1390023	0.055498	0.500464	0.680348	0.04877	0.680431
25	34	(21.6837, 47.1404)	0.218913757	0.237353	0.379032	0.43093	0.333347	0.588235

3.1 Outliers

Just like with anything, there always are outliers to the presented measurements, as can be seen on the last prediction of the presented table. Fortunately, the occurrence of outliers is rare (4 residuals in whole sample), which suggest that the method is capable of producing reliable predictions. The process of suture fusion, similarly to puberty, can occur earlier in some

individuals than in others, making it quite individual in some cases. Other causes for outliers can be diet and health condition for example.

If an individual is systematically exposed to starvation, it is likely that their bones are underdeveloped for their age, which can cause inaccuracy in the predicted age estimates. Also, conditions like osteoporosis can significantly affect the degree of suture fusion, as they can degrade bone matter even in the skull, regardless of age.

Such external factors can't always be taken into account, and can cause some inaccuracy, which is of course inevitable regardless of the age estimation method.

3.2 Results summary

Considering that the most accurate, already existing, methods for AAD estimation based on analysis of the cranial sutures have an accuracy of + / - 15 year, the achieved results are a significant improvement. The algorithm manages to reduce the error of the predicted AADs over 3 times compared to the already existing techniques.

In conclusion, the project completely automatizes the process of cranial suture assessment and provides a significantly more accurate age prediction than any of the existing methods in the field. The presented segmentation method can be used for analyzing skulls of any ethnicity after appropriate tuning of the regression equation, making the algorithm applicable in a wide variety of fields including criminology and arachaeology.

4 Technologies and resources

1. Source code of the project can be found at:
   https://github.com/andics/Cranio-analysis—Image-Extractor

2. The CT scans of the skulls for this project are provided by the Bulgarian Academy of Sciences. The scanning was performed using an industrial µCT system Nikon XT H 225, developed by Nikon Metrology.
   All of the skulls belonged to Bulgarian soldiers who died in the First and the Second Balkan Wars and the First World War. Their skeletal

remains were preserved in the Military Mausoleum Ossuary, at the National Museum of Military History (Bulgaria). The individuals were fit for service which means they were without severe disorders and malformations. Information about the AAD was taken from the museum's archive, where it was kept.
The skull population used for the project has a size of 42 individuals from various age groups.

3. For the implementation of the cross-sectional image generator, the image processing and the Semantic Segmentation network, I've used *MATLAB R2018a*.

4. For part of the data labeling, I've used *Labelbox* (https://labelbox.com/)

References

[1] High-resolution flat-panel volumetric CT images show no correlation between human age and sagittal suture obliteration—Independent of sex. `https://www.sciencedirect.com/science/article/pii/S0379073810001787?via%3Dihub`