Using Complex Computer Vision Systems and Machine Learning to Develop a Live Time Navigational and Surgical Aid for Spinal Reconstructions

Krithik Ramesh

Abstract

The existing navigational system for spinal reconstructions called fluoroscopy presents significant visual and physiological disadvantages. Fluoroscopy has a limited field of view of the spinal column and emits a significant amount of radiation that affects the patient and the surgical team. Implementation of a machine learning (ML) and computer vision based navigational system has the potential to eliminate existing risks. Microsoft HoloLens, an augmented reality headset, gives the surgeon the ability to see the real-world and pertinent medical data. These data were parsed through ML algorithms that mapped the spinal column, suggested the best approach, and would guide the surgeon in real time. The project was tested in two phases: algorithmic and real world viability. Algorithmically, the performance was measured by training and testing over 2,000 patient's publicly available CT and MRI scans. The data was used to build a spine model. The data was used to build a spine model. The real-world testing was conducted where 34 anonymized patient's spinal reconstruction was observed, and data was validated post-operatively. Algorithmic performance measured found that the algorithm created a map of the spine within 88 seconds with 98.6% accuracy. The real world testing found that the headset was able to map the vertebrae and suggest the correct approach 96.6% of the time within 1.33 mm accuracy of the true values. The data suggests that the developed navigational system would be a pragmatic and economically viable replacement for fluoroscopy.

Spinal Reconstructions - Patient Outcomes, Operating Time, and Navigational Systems:

Spinal reconstruction surgery may be necessary for patients who have a deformity or misalignment that affects a major portion of the spine. The procedure involves more than one level of the spine and corrects significant spinal deformities, stabilizes the newly shaped spine with rods and pins, and fuses the vertebrae together. In some cases, entire vertebrae are removed and replaced with artificial devices to replace the diseased segment. The most common conditions to be treated with spinal reconstruction are scoliosis, spondylolisthesis, and kyphosis. A study conducted by the National Institute of Health found that 63% of patients had significant post-operative pain more than three months after their surgery and physical therapy. In addition to post-operative complications, patients that undergo can have devastating sequelae, including hardware exposure, meningitis, and unplanned reoperation. The literature shows that wound complication rates in this patient population approach 19 percent and, in very high-risk patients (i.e., *prior* spinal surgery, existing spinal wound infection, cerebrospinal fluid leak, malignancy, or history of radiation therapy), as high as 40 percent and with reoperation rates as high as 12 percent (Mayo). The survival rate of patients that undergo spinal reconstructions is very high: however, the complexity of the surgery and dated medical imaging and navigational systems, prove to have relatively low long term post-operative patient success. Current navigational systems used for spinal reconstructions are called fluoroscopy machines. Fluoroscopy is an imaging technique that uses X-rays to obtain real-time moving images of the interior of an object. In its primary application of medical imaging, a fluoroscope allows a physician to see the internal structure and function of a patient's anatomy. While fluoroscopy provides live time imaging of a patient, it presents significant physiological harm to the patient and limited visual acuity for surgeons. Fluoroscopy has significant radiation emission that is absorbed by the patient and all proximal members in the operating room. In addition to the radiation emission, the field of vision received from fluoroscopy imaging is limited to two-dimensional view that doesn't reveal all the nervous and muscle systems in the region. This lack of anatomical detail and poor visual acuity leads to inaccuracies in the surgical placement of screws during the reconstructive surgeries that can either limit patient mobility or even cause paralysis. A study done by the Mayo Clinic found that existing navigational systems add anywhere between 30 minutes to 75 minutes to operating time of spinal and general surgery.

Figure 1: Fluoroscopy Image of lumbar spinal fusion with poor screw placement (American Association of Neurological Surgeon,15).

Project Goals:

There are three main objectives to effectively improve surgical reconstructions of spines.

1. Develop a comprehensive three-dimensional anatomical model – create a model of the entire spine and proximal thoracic and cervical regions. The anatomical model must include the peripheral nervous systems. This objective helps develop a comprehensive and customized anatomical model to optimize surgical approaches.

2. Creating an active computer vision system – make a computer vision system that uses existing surgical cameras and object identification Region-Based CNN machine learning algorithm to overlay a mask over anatomical regions and tumor areas. This would allow surgeons to use the anatomical model and the visual assistance to more accurately excise the tumors.

3. Get real world testing and actual test data– Use publicly available data from Stanford University's Radiology Clinic, University of Hawaii at Manoa, and Oxford University to train and validate my algorithm. A total of 34 patients from this given dataset are extensively documented allowing for end-to-end validation.

Hypothesis and Solution Approach:

Using a complex hybrid machine learning algorithm that uses CT or MRI (based on what medical imaging data is available) and develops a highly accurate nervous, vertebrae, muscle, and spinal cord map. This map would then be integrated with a computer vision region based convolutional neural network to identify and assist surgeons in reconstructive surgery. The system would be an application executable on Microsoft HoloLens as an augmented reality system that provides surgeons with feedback and instruction. By watching and training the machine learning algorithm on many past surgical cases in addition to real world surgical procedures, the algorithm can optimize screw placement, provide surgeons with enhanced visual acuity, and provide a highly accurate and computationally efficient anatomical map of the patient. If a novel machine learning algorithm is developed to map a patients' spinal column and proximal anatomical nervous and muscle systems, then surgeons will have an improved navigational system that reduces operating time, long-term patient outcomes, and reduce surgical error.

Region Based Convolutional Networks, Map Localization, and Rendering in 3D Methodology:

Region based Convolutional Neural Networks:

The region-based convolutional network is one of the main contemporary approaches to object detection. Consider the multi-class object detection problem. This method is called region-based convolution network or R-CNN. In R-CNN, a selective search algorithm was used to enhance computationally efficiency, specifically geared toward high confidence level for surgical instrumentation and anatomical components. Selective search is one of the generic object proposal generation methods (Lu, Zhang, et al.). Such methods can provide high recall rate but moderate precision. Thus, they cannot be used as object detectors themselves, but they can be used as a first step in the detection pipeline. The training procedure for the R-CNN consist of three steps. First, CNN is pre-trained on ImageNet of anatomical structures for image classification. Second, CNN is fine-tuned for object detection on a limited object detection data set. Third, linear classifier and bounding box regressors are trained on top of CNN features are extracted from the object proposal. Experimental evolution has demonstrated that R-CNN outperforms previous object detection methods, even if only pre-treated CNN is used for feature extraction. This algorithm serves an accurate method of identifying multiple objects with relatively high accuracy in a computationally efficient manner.

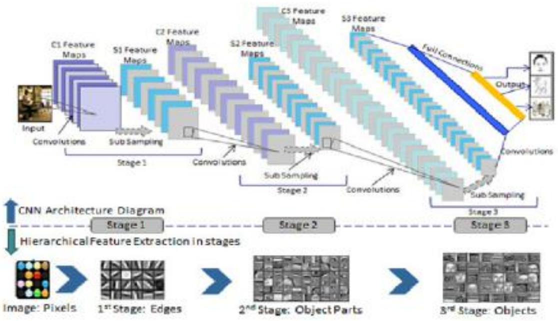


Figure 2: RCNN architecture break down of how an image model is learned and learned (Lu, Zhang, et al.)

Map Localization for Anatomical Structures:

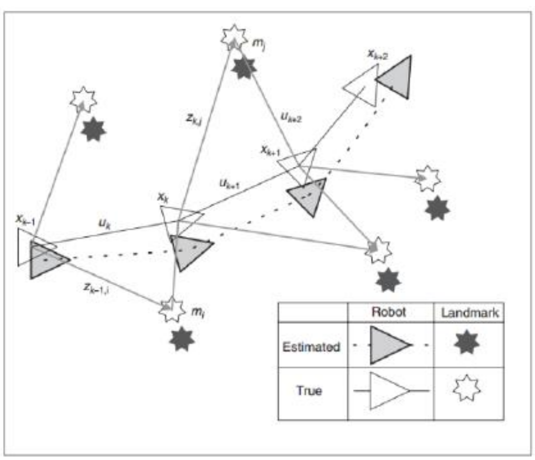


Figure 3: Map Localization diagram highlighting probabilistic methods of approximating anatomical of structures Zeller, Matt, et al.)

Simultaneous Localization and Mapping algorithm (SLAM) was used to perform three-dimensional spatial mapping of anatomical structure and operating environment in addition to predicting spinal behavior (Durrant-Whyte, Hugh, and Tim Bailey). The challenge is to place the HoloLens augmented reality headset at an unknown location in an unknown environment, and have the system incrementally build an anatomical map of the patient with the given medical imaging and determine optimal surgical approach for reconstruction (Zeller, Matt, et al.). The algorithm adds a layer of dimensionality to the region based convolutional neural network and volume rendering by translating computed anatomical data with real world data seen in the operating room. The application of this algorithm also serves as the primary factor in eliminating the need for fluoroscopy by using highly accurate location estimation, the software is effectively able to pinpoint significant features required for the performance of the surgery. Ultimately SLAM allows for all landmarks to be estimated locally without the need for any *a priori* knowledge of location or features.

Volume Rendering of Spinal Column Segmentation Approach:

As commonly implemented, 3D volume rendering takes the entire volume of data, sums the contributions of each voxel along a line from the viewer's eye through the data set, and displays the resulting composite for each pixel of the display (Figure 4). Incorporation of information from the entire volume can lead to greater fidelity to the data; however, much more powerful computers are required to perform volume rendering at a reasonable speed (Meißner, M, et al.). With wider availability and improved cost-to-performance ratios in computing, volume rendering is likely to enjoy widespread use in the medical community. This system is both computational efficient and highly accurate, allowing for map localization of nervous structures, spinal sections, and vertebrae using machine learning.

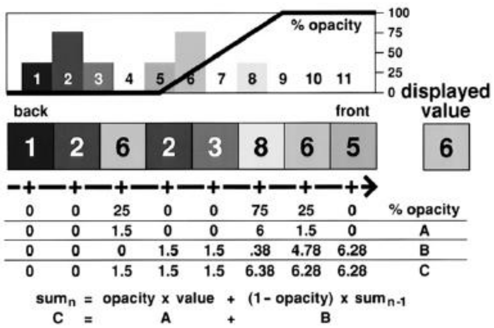


Figure 4: Diagram illustrates volume rendering technique (Meißner, M, et al.)..

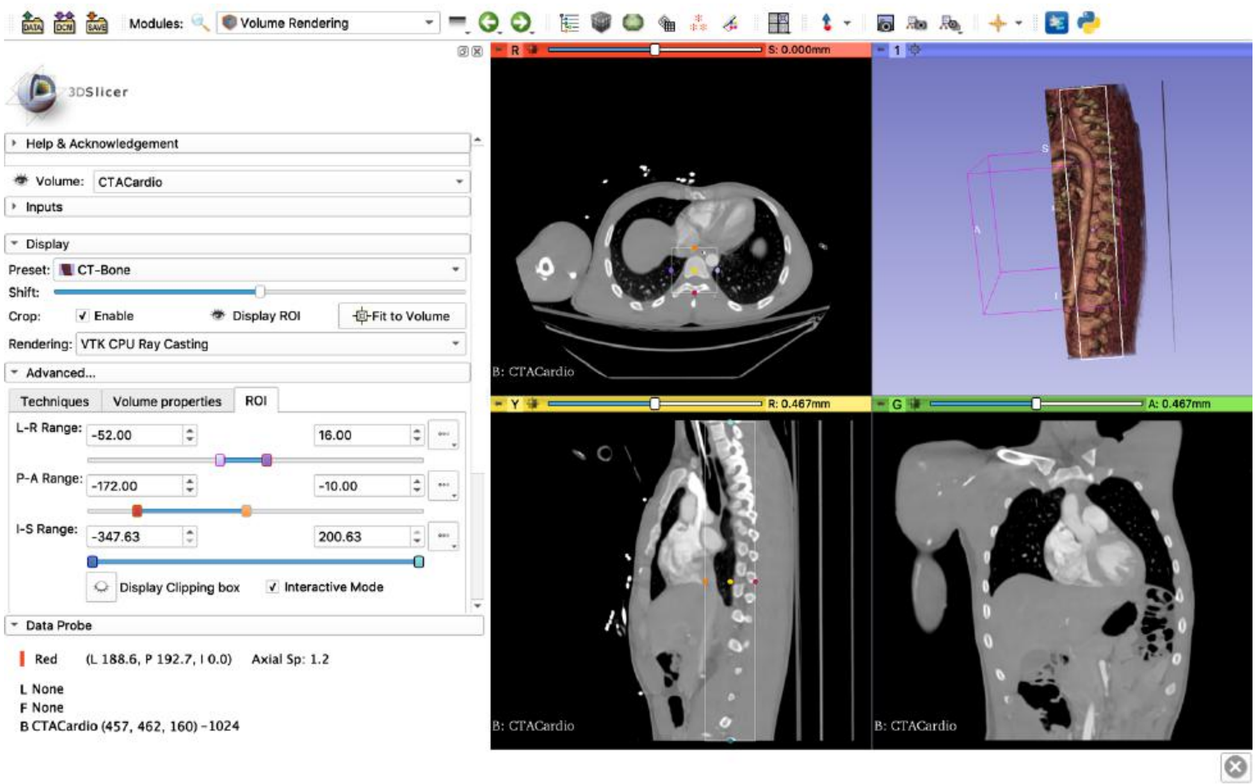


Figure 5: Muscle, nervous system, and spinal cord segmentation using map localization machine learning and volume rendering. Taken by Krithik Ramesh

Experimental Design:

Theoretically:
CT/MRI Scans and Operative Data from spinal reconstructions from the University of Hawaii were used as an accurate data set for training. The resulting medical imaging data was used to develop and refine the volume rendering and map localization algorithms. A wide array of medical equipment and OR room images. The data was built on different lighting conditions and OR orientations to increase overall accuracy and versality of the object detection algorithm.

Experimentally:
The developed algorithm will be implemented in a surgical room by providing surgeons with a HoloLens headset that provides live time object detection and surgical instructions for surgeons using the machine learning algorithm developed in the theoretical phase.

Implementation:
The HoloLens System is a versatile wearable computing system, the economic and practical viability as an alternative to fluoroscopy systems is very high. The system itself costs $5,000 per headset and lasts upwards of 5 hours of consistent use with both powerful internals and reliable hardware. A fluoroscopy machine in comparison costs upwards of $78,000. The integration of HoloLens would also reduce operating time for surgeons by integrating all medical imaging both pre-operative and live-time info in one field of view. HoloLens could be integrated into the surgical field in an effective manner due to its minimalistic and effective nature.

Practical Applications:
By implementing the given system, it would reduce operating time for reconstructive surgeries by approximately 16% and significantly lower radiation exposure for both the surgical team and the patient. By using machine learning for spinal behavior and object identification, it limits the need for multiple scans and expensive machinery to perform surgery. The system would also prove beneficial to trauma patients by providing time conscious and accurate data and in military/ disadvantaged regions. In war zones and rural areas, coming by expensive machinery is next to impossible and highly impractical; however, implementing compact headset systems increases chances of survival and increase overall medical capabilities in these situations.

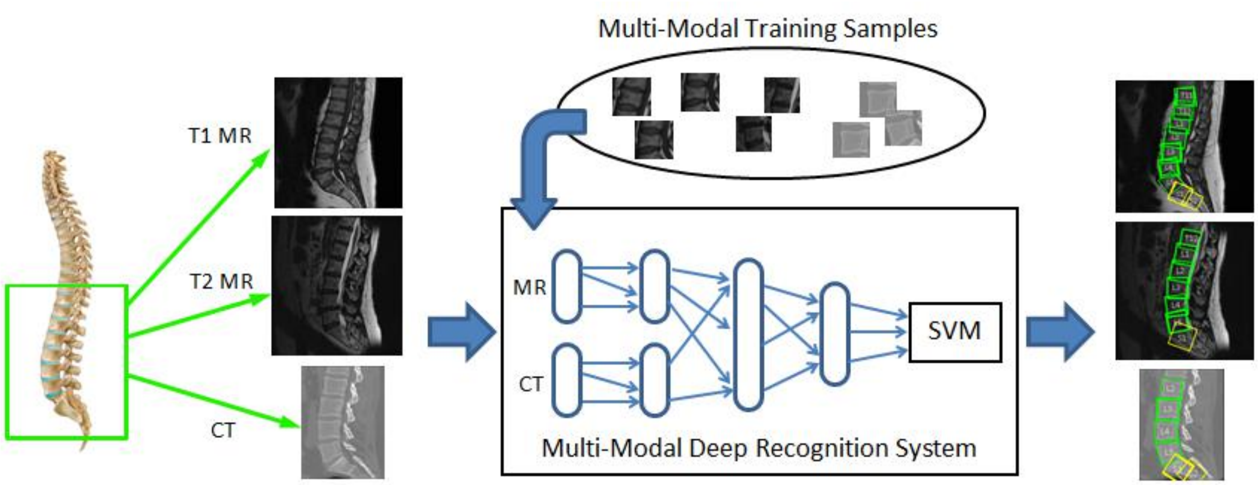
Methods and Materials

Algorithmic and Anatomical Model:

The algorithmic model developed was a homogenization of three machine learning networks. The model incorporated a hierarchal deformable model for back end anatomical development incorporated with map localization and a region based convolutional neural network to match computed features from medical imaging with real-time images collected from the augmented reality headset. The software set out to meet two main criteria: high accuracy anatomical feature recognition and be computationally efficient.

Hierarchical Deformable Model Program Methodology:
The deformable model is based on existing literature that approaches anatomical features from a global standpoint and then reconstructs the spinal model from provided medical and said features. Modifications made to the existing system includes identification of anomalous structures in the spinal coloumn along with approximation of vertebrae spacing and size measurements.

Step 1: Multi-modal Feature Extraction and Vetebraeresult detection
Part A: A deep-learning-based supervised detection approach. The leverages multiple medical



imaging results and a deep learning network to identify spinal column features and damaged vertebrae. The resulting data points from the medical imaging is parsed through a support vector

Figure 1: The multi-modal recognition for lumbar spine imaging. The modalities are uniformly trained and detected in one unified recognition system. In this system, features from different modalities are fused and enhanced by each other via a deep network (Cai et al.)

machine to classify individual vertebrae.
Part B: The mechanics behind the Transformed Deep Convolution Network (TDCN).

A TDCN is a variant of a deep convolutional neural network that utilizes feature attaction to enhance a given data set for classification. In the case of this algorithm the given Magnetic

Resonance (MR) and Computer Tomography (CT) Scans were amalgamated such that low-level features from MR Scans were transposed with CT features. The TDCN automatically extracts the

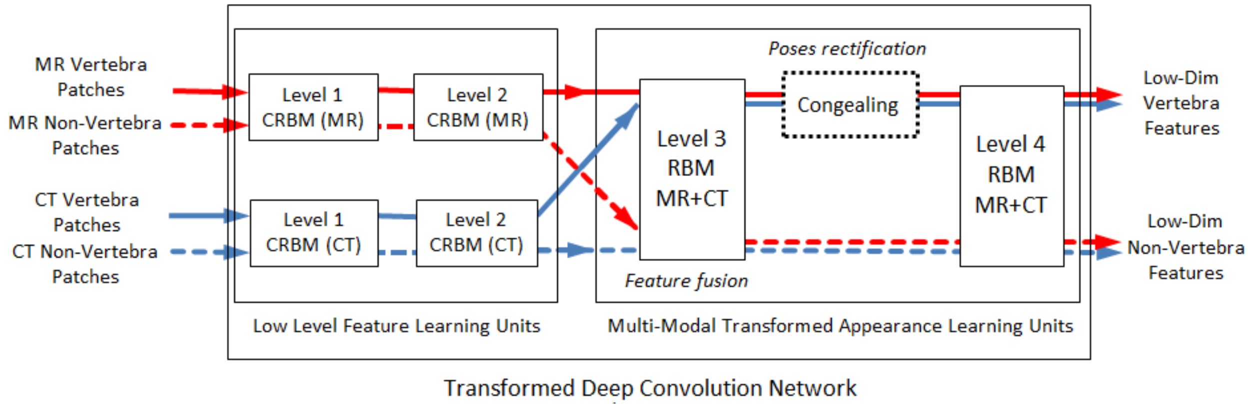


Figure 2: The structure of the Transformed Deep Convolution Network (TDCN). Two different types of layer, Convolution Restricted Boltzmann Machines (CRBM) and Restricted Boltzmann Machines (RBM), are used in the network. Note that the congealing component is applied in the training stage and will be bypassed in the testing stage. Congealing is utilized as a refinement method for improved data sets (referenced later) (Cai et al.).

best representative and invariant features for MR/CT. It employs MR–CT feature fusion to enhance the feature discriminativity, and applies alignment transforms for input data to generate invariant representation. This resolves the modality and pose variation problems in existing vertebra recognition systems.

Step 2: 2D and 3D Spine Structure Recognition and Modelling The most accurate and computational efficient anatomy approach is a Hierarchal Deformable Model (HDM). An HDM creates a vertebra-based compositional model that effectively simulates spine deformation. The

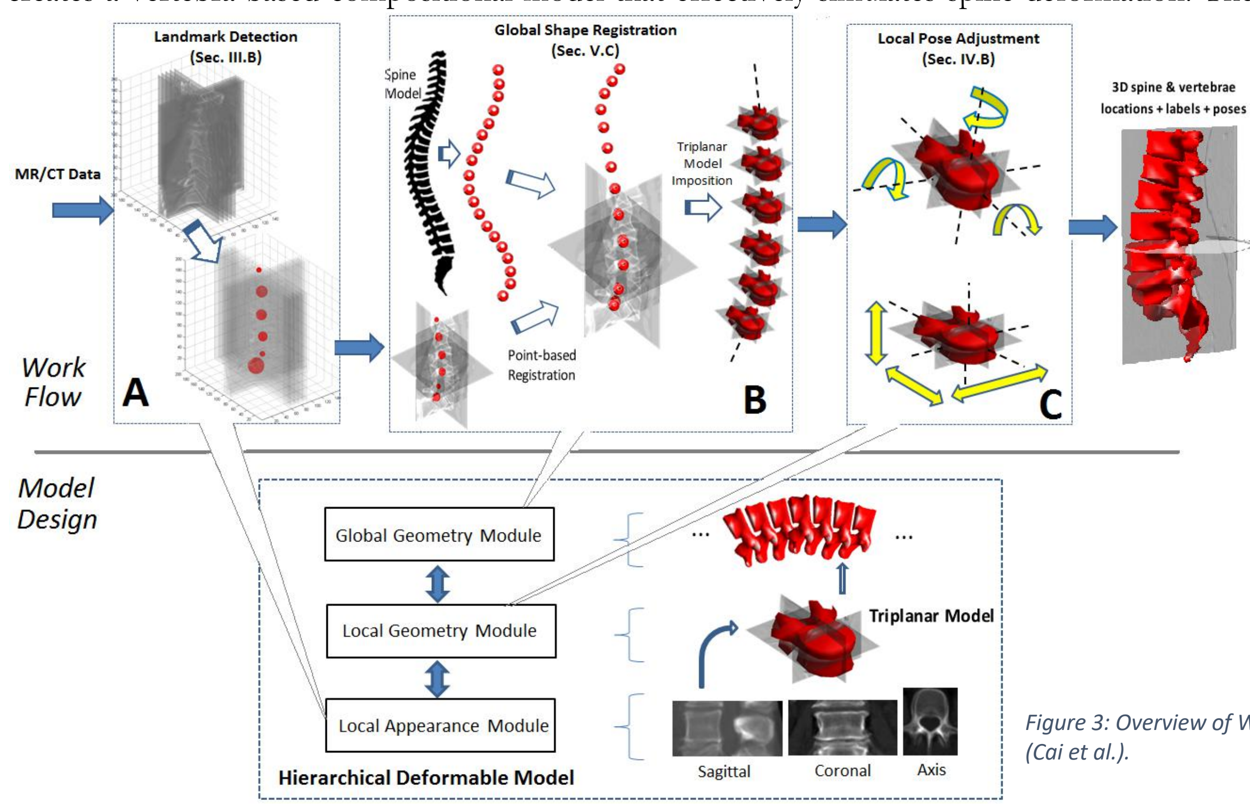
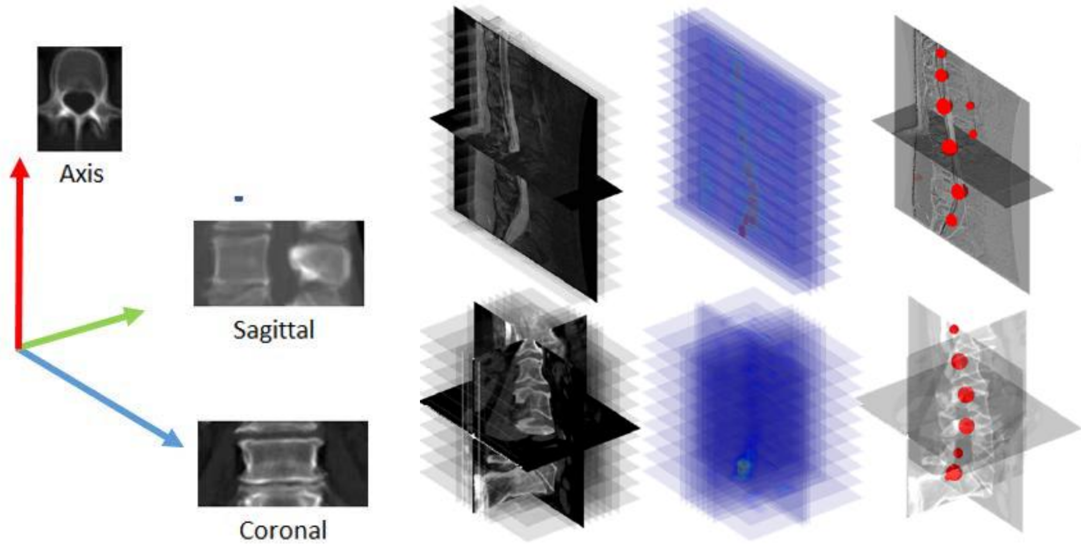


Figure 3: Overview of Workflow (Cai et al.).

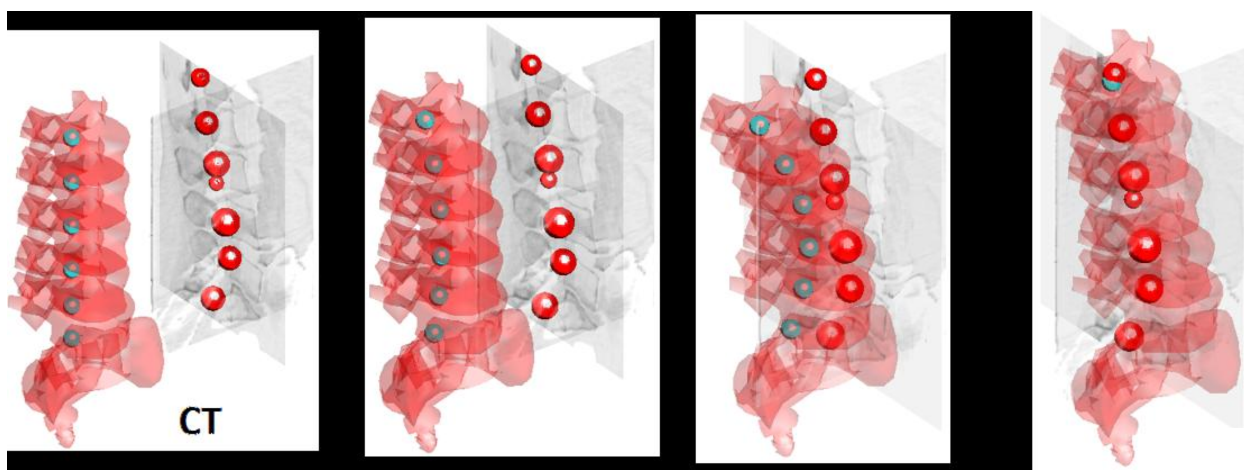
model takes a triplanar vertebrae (represented in far-left field in diagram below) and lines up the volumetric data from CT scans with the slices from MRI data. The final feature of the system is developing a model-based 3D reconstruction of the entire spine.

Part A: Triplanar template matching



Apply 2D template matching on sagittal, axial, coronal view using deep features. Also sets up the Center points of each vertebrae for global construction in part B

Part B: Global Shape Registration



This process ultimately provides the general structure and develops the initial spinal column. First, the system is intended to reduce the points from the tri-planar view and register anchor points for the vertebrae along the column. Finally, the vertebrae are anchored on the registered points which prevents translational mis-alignment of other vertebrae

Part C: Local Pose Alignment

Using all three views of the volumetric CT data and the congealing step from the deep neural network mentioned in step one. Finally, for refinement of angular placement of the individual vertebrae using back projection of the aligned two-dimensional planar poses from MR and CT data utilized previously.

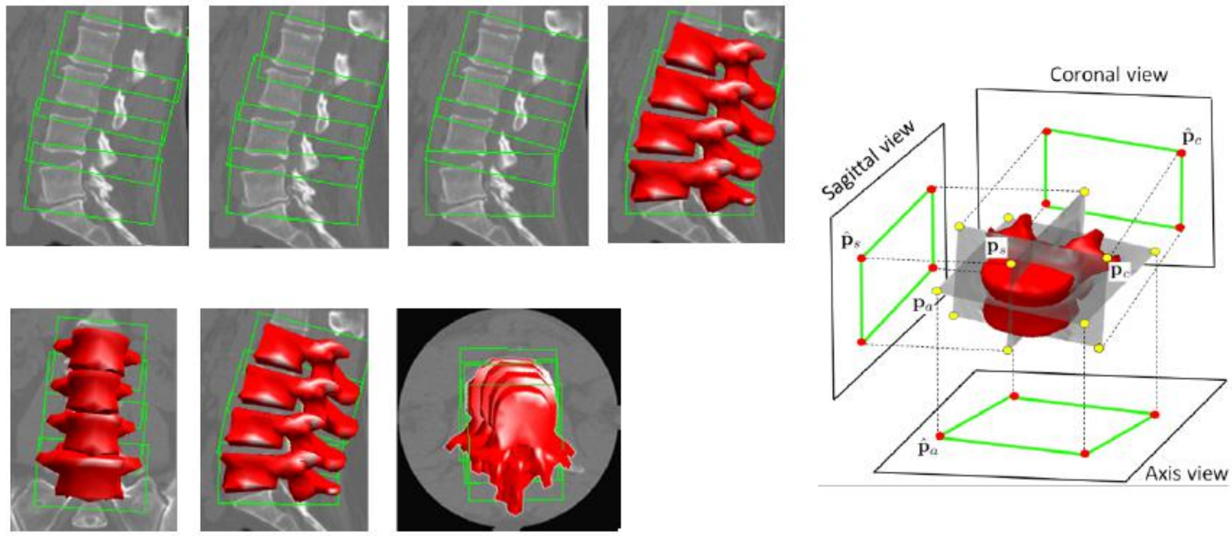


Figure 4: Triplanar registration and reconstruction

Simultaneous Localization and Mapping Algorithm:

The HoloLens system uses an array of infrared and motion sensors to accurately map an environment. By using the sensor array to do simultaneous localization and mapping of the operating environment, surgeons are allowed to interact with their environment and anatomical structures in 3D. The localization algorithm develops the entire basis of the spinal column by using interpolating images and sensor data with accelometer movements to piece together a map of the environment. In addition to the mapping based off of the sensor array, the detailed HDM from the backend medical imaging to connect the theoretical model with real world inputs.

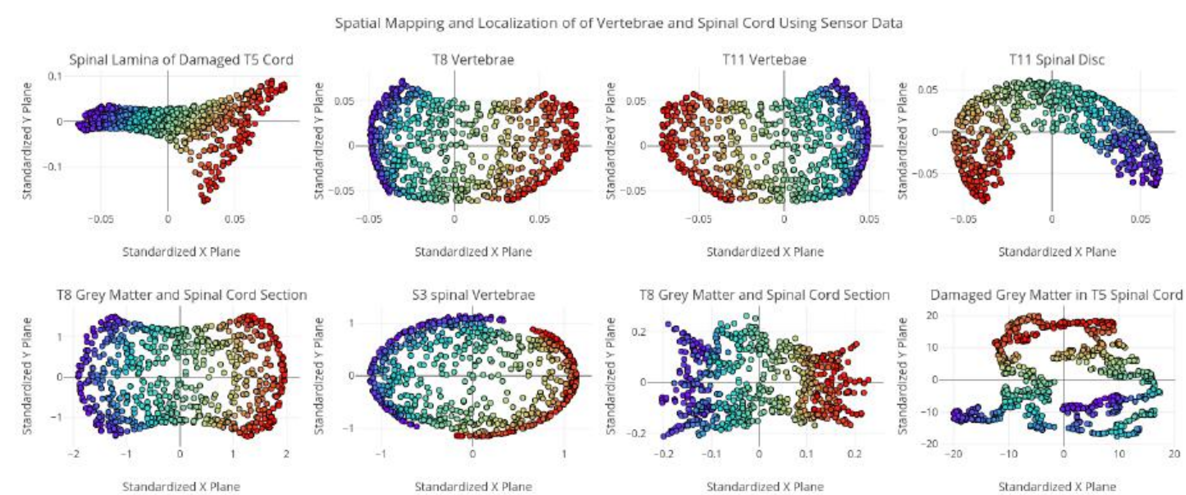


Figure 5: 3D Mapping and Localization in 3D of Vertebrae using sensor arrays from HoloLens. Explained Further in results section.

The Region-Based Convolutional Neural Network:

CNN is designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. This system allows computer vision to handle radiological and real-world tasks associated with spinal reconstruction. Two challenges in applying CNN to radiological tasks are small datasets and overfitting, given that spinal vertebrae are similar in structure, training to account for anatomical differences is key in developing a strong model. A large and accurate dataset is required for optimal results that takes into account various picture features (explained further in next section). Being familiar with the concepts and advantages, as well as limitations, of CNN is essential to leverage its potential in diagnostic radiology, with the goal of augmenting the performance of surgeons and improving patient care.

Data Set Modeling and Optimization:

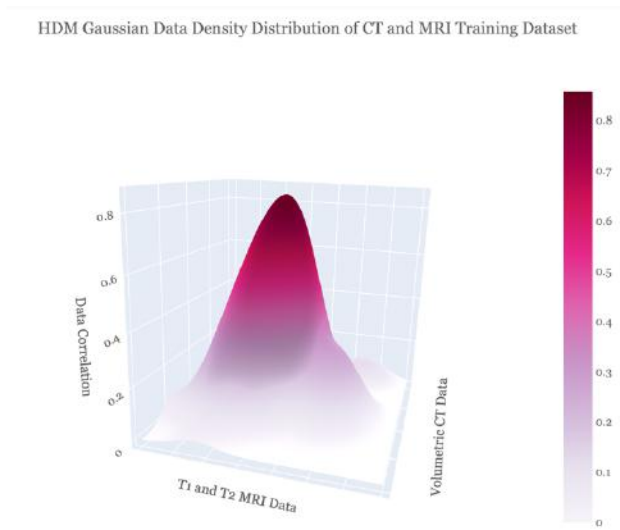


Figure 6: Network graph of mapped anatomical points. Nodes with greater connections are of greater significance to patient motor function and chance of survival. Map shows general trend of equal distributions except for few regions resulting in the Gaussian mean of 0.78.

To improve the versality of the RCNN and the HDM a varied and large dataset is required. Both the HDM and RCNN implement stochastic[1] data evaluation. In the case of the HDM both the TDCN and SVM and needed to be able to handle a wide variety of body types, genders, age, and deformities. To provide an optimal dataset that preferentializes accuracy and efficiency over memory usage, the dataset developed for the HDM used over 500 CT and MRI Scans collected from various medical institutions. The dataset was manually curated to identify the spinal column and individual vertebrae. The data set was split with 60% of the values being used to train the networks and the other 40% for testing and validation. The ideal dataset pertained a [15x15] with the density of the gaussian being 0. The data is vertically elongated due the complexity and range of the features, but ultimately the density of the medical dataset was approximately 0.78 resulting in near standard distribution. In the case of the RCNN the data was far more versatile as image identification required various lighting conditions, image resolutions, image angle, object light refraction, and object size. The resulting dataset for the RCNN included the entire vertebrae column as well as required surgical equipment for a spinal reconstruction. The dataset included a total of 25,000 images spanning over 125 anatomical structures and surgical instrumentations.

[1] The term stochastic means randomly determined or having probability distribution that can be statistically approximated.

Testing Procedures:

Testing Procedure:

Process:
The patients in the data set not vary throughout the experiment and the algorithm was trained on existing operation techniques and preoperative and postoperative care. If notes were provided on a given patient on surgical approach used, the algorithm was program with the same general operating parameters for the best validation metrics.

Patient Analysis:

Patients spanned across ages 22 – 68. Injuries tested from patients included herniated discs, fractured vertebrae (trauma and aging degradation), and scoliosis. According to patient reports in the datasets 96% of the patient pool went into the surgery with acute back pain, 42% of patients experiences some numbness in extremities. Subjects were asked to identify level of pain and numbness before surgery and 48 hours after care (to reduce the effects of painkilling medication).

Post algorithmic validation:

Computationally:

A regression metric was used to see if the screw placement determined by the algorithm was close to actual surgical placement. The distances were calculated using the Euclidean distances on a given image. The regression of the estimated vs real points will be used as one metric for validation of the data.

Additionally, analysis of the images and trajectories were done by using a cosine similarity index to ensure that the screw placement on a three-dimensional plane was computed correctly. This is also used a method of enduring the accuracy of the region based convolutional network

The region based convolutional neural network performance is also measured by looking at receiver operating characteristics and the precision vs. recall.

The Hierarchical deformable model is validated by cross referenced using a geometric similarity analysis from CT and MRI volume rending or 3D reconstruction.

The simultaneous localization and mapping algorithm did not be validated as this was implemented using an already validated by the Microsoft HoloLens developer kit internally.

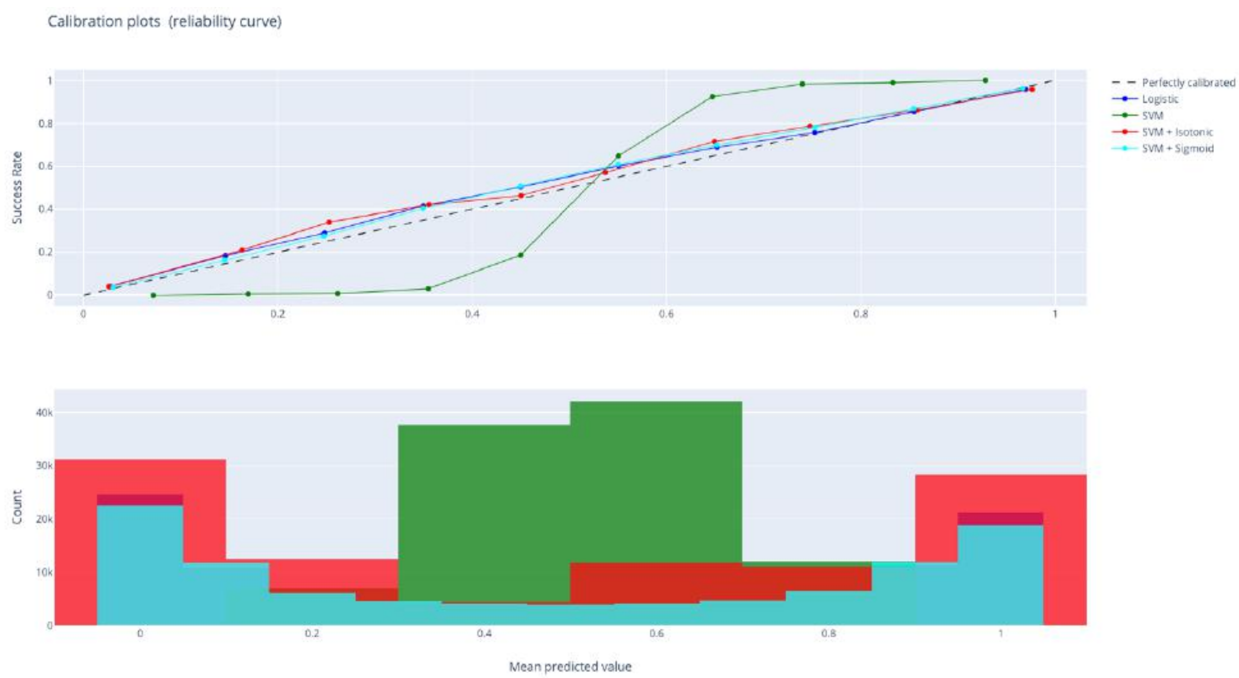
Results

Algorithmic Performance:

Hierarchical Deformable Model:

Training Performance:

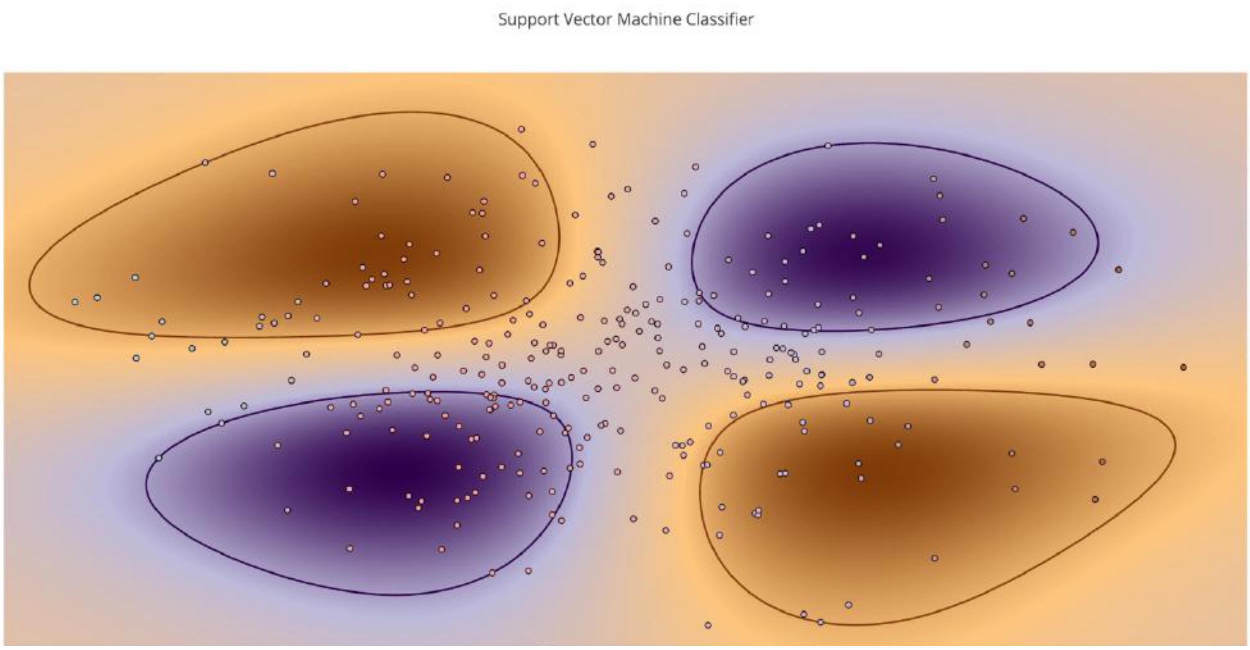
The training model was optimized to prefer accuracy over time, learn the greatest number of anatomical features, and categorize them into order of significance for identification. The data was analyzed over 125 CT and MRI scans from the University of Hawaiʻi at Mānoa. Each MRI and CT pain was trained up to 40,000 epochs[1]. The number of epochs for a given dataset was determined by overall confidence of the results. If the algorithm demonstrated a confidence level in images above 95.5% then training was completed. To remove the likelihood of overfitting due to the small dataset, the algorithm had a relatively high drop rate[2] of 21.7% (compared to median of 16.3%). Typical datasets for medical imaging ranges upward of thousands of medical images, but due to the specific requirements of patients the dataset was significantly limited. Thus, to reduce the likelihood of the algorithm learning trends to specific patients the drop rate was increased.



[1] An Epoch is when an algorithm has passed through the entire dataset forward and done backward propagation of the resulting values once.
[2] Drop rate is a function of a machine learning where random data points are dropped from final analysis a measure of reinforcing and validating derived data from dataset.

The first graph above represents the success rate in anatomical feature identification with respect to time training. The mode accuracy of the support vector machines was 98.6% with all datasets being fully trained in 32 hours. The weight towards accuracy ultimately proved beneficial in improving object identification time as well (explained further in real world performance). The second graph identifies the progression of confidence in terms of the number of epochs run by the algorithm. The data shows that as the algorithm was reaching maxima and medians of data (0%, 50%, and ~100%[3]) the epoch count went up significantly. This suggests that the SVM increased computational load as the algorithm learned very general features (spinal column location and proximal muscular nervous systems), and then more intricate features such as location of individual vertebrae and spinal cord segmentation, and finally very intricate systems such as anatomical curvatures of spinal vertebrae. This indicates that the SVM training took a holistic representation of the spinal column.

Support Vector Machine Visualization:

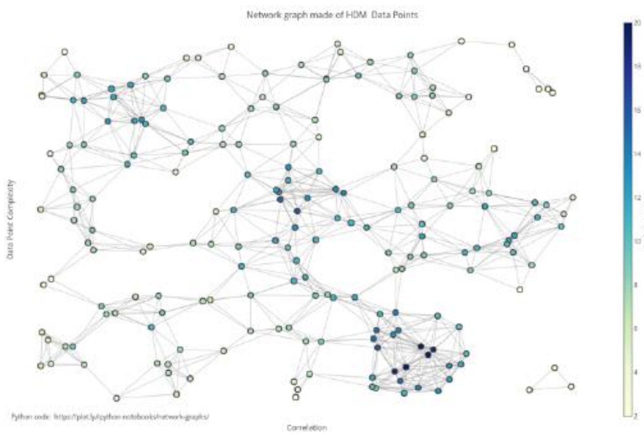


The graph above is a visual representation of the multi-class support vector machine, each clustering is represented by 4 main groupings of data. The batches include different regions of the spinal column. The top left class is the largest and is allocated for the Thoracic spinal group. The range in data is a result of significant curvature in global geometry[4] and number of vertebrae. The top right is cervical spinal group, the bottom right is the sacral region, and the bottom right is the miscellaneous structures. The intermittent data points between the groupings are anatomical structures near the spinal column that might be significant to the patient. This support vector

[3] Accuracy is never truly one hundred percent but comes very close to true value

[4] Global Geometry is looking at a section of vertebrae as a whole and identifying the curvature of the assembly as opposed to individual segments.

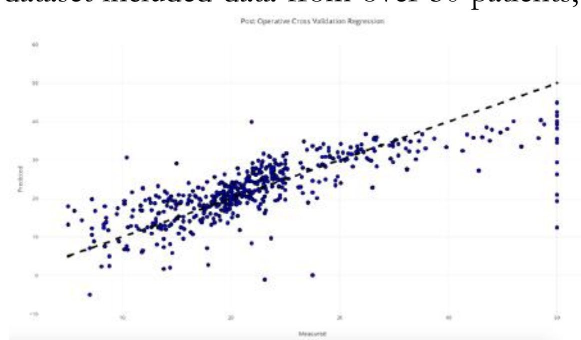
machine presented above allows for classification of the spinal column and sets up the framework for the network graph presented below. The network graph analyzes the spinal cord and vertebrae from the classification from the SVM and sets up a priority system. The priority system takes into account the level of damage each vertebra has endured and identify the risk associated with repairing a specific vertebrae or series.



Data Structure Analysis:

The graph above is a network of the significance of individual vertebrae and spinal cord regions to patient retention of motor function and safety. Nodes with darker colors have stronger correlation to overall patient wellbeing. In the case of the node clustering seen in the bottom right is the mapping of cervical region. The nodes at the center of that cluster are the Atlas, Axis, and C2 - C7 Vertebrae. These nodes have the greatest complexity as nerve roots and connection between the brain stem and the rest of the spinal cord exists in this section. The Thoracic region has the second greatest priority as prefaced in the previous section due to its size and connection to many body parts outside the spinal column. Conversely, regions such as the coccyx as represented at bottom right is isolated due to the negligible correlation to overall patient mobility. The correlation was derived from the existing data from the University of Hawaiʻi at Mānoa in addition to the colorized annotation[5] from the Imaging Clinic (Radiology) at Stanford University. The annotation included a color map of severity of damage based of location on spinal region. The dataset included data from over 50 patients, but the age and gender were anonymized.

Cross Validation Results:

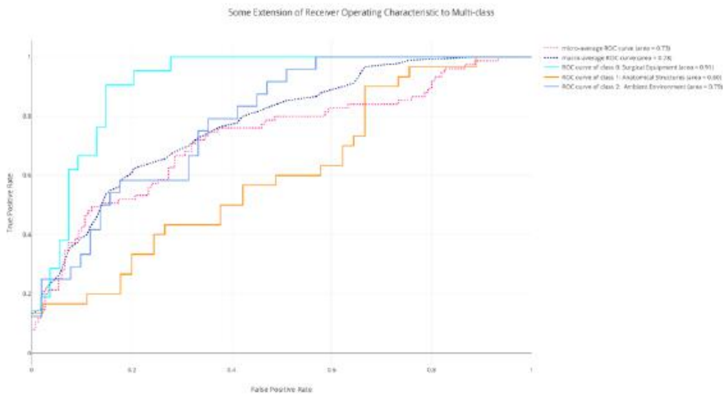


After surgeries were performed the values recorded from the HoloLens headset on anatomical points of reference and feature mapping were hand validated by surgical teams to improve analysis and run values back through network with backward propagation to correct previous analytical mistakes. The graph above shows the regression line produced from the data collected of anatomical structure features. The logistic line shows very high correlation and significant refinement was not needed. The r^2 value of the regression being ~0.973. This indicates that the accuracy of the model in addition to the calibration of the HoloLens was very accurate.

[5] The annotation was similar to a heatmap overlaid onto existing CT and MRI scans. The colors would get darker as the damaged increased.

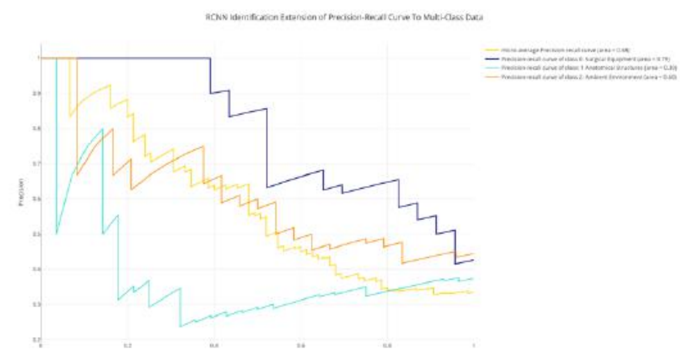
Region Based Convolution Neural Network:

Learning Results:



The RCNN was the visual object identification system used on the HoloLens Headset. Much like the Support Vector Machine the RCNN was separated into three main classes: Surgical Equipment, Anatomical Structures, and Ambient Environment. The graph presented on the left is visualization the Receiver Operating Characteristic (ROC) of the learning rate and accuracy of the training data. The greater the area of under the curve, the better the model. Class one had the best learning efficiency and accuracy with an area of 0.91, the characteristic of the graph sees a significant incline at the beginning of the graph and a levels of as true positive rate approaches one. Ambient environment performed the second best based with an area of 0.78 with a relatively sharp incline initially and a decreased rate of change in terms of the false positive rate. The Anatomical Structure class had the worst performance. The reasons behind the low performance is matter of the similarity of shape and form of the individual vertebrae. To counter act the low performance, the accelometer[6] data from the HoloLens and the backend data from the HDM was used to identify individual vertebrae that needed reconstruction.

Precision Vs. Recall:

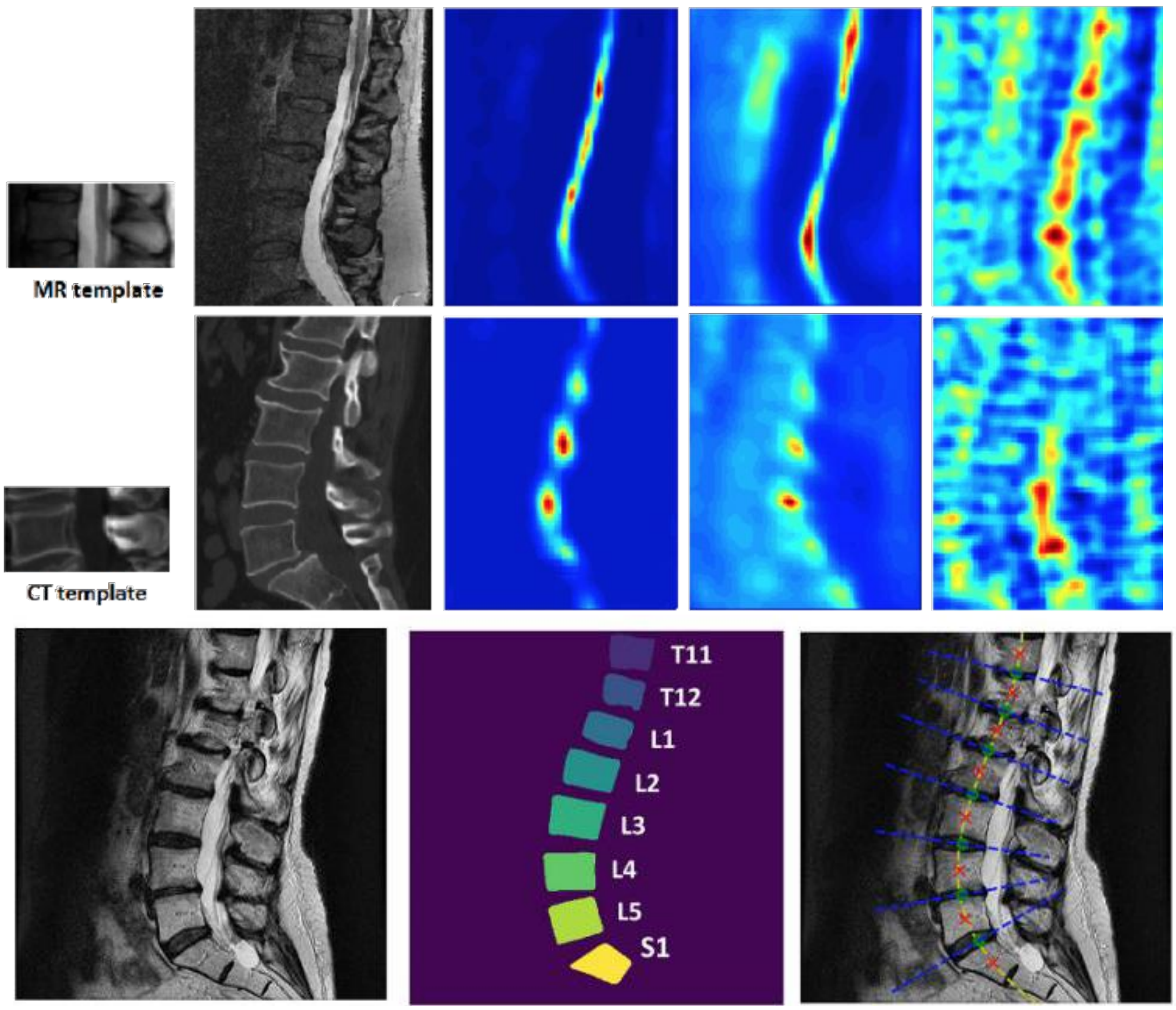


Precision and recall are two extremely important model evaluation metrics. Precision references the percentage of results which are relevant, and recall refers to the percentage of total relevant results correctly classified by the algorithm. The graph to the right assesses the results of recall and accuracy of the RCNN data of the three main classes. Analogous to the performance seen in the previous graph, Surgical equipment had the best recall and accuracy followed by Ambient Environment, and finally Anatomical structures. Despite the initial discrepancy, as the plots converge when the recall percentage increases significantly, the overall performance was approximately the same. In terms of runtime,

[6] The methodology behind the accelometer and spatial mapping will be explained in the Simultaneous Localization and Mapping section.

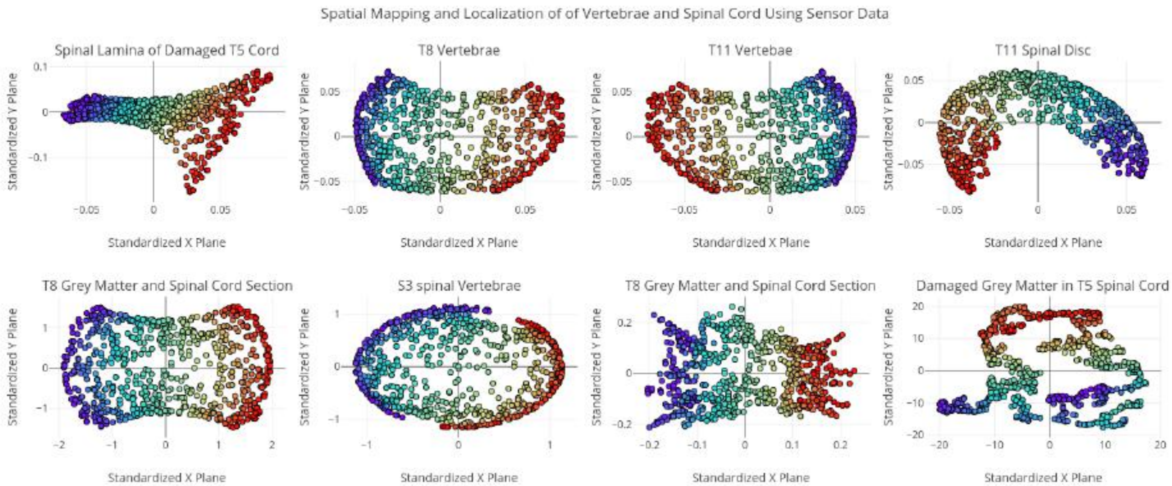
identification throughout all classes averaged at around 0.7 seconds for ~75 identifications a frame.

RCNN Analysis:



The resulting images are from the RCNN model with different volumetric slices being analyzed for segmentation and identification. The images above identify a variety of features over every iteration ultimately creating a comprehensive analysis of anatomical

Simultaneous Localization and Mapping:



Simultaneous localization and mapping was the most computational inexpensive algorithm of the three as it was already optimized by Microsoft with the HoloLens SDK. To reinforce the image identification done by the RCNN the mapping of anatomical structures and features using the depth camera and infrared sensors. The spatial data is matched with the HDM to identify the structures. The accuracy of the sensors was very high and was accurately able to trace edge and folding features of individual vertebrae and features.

Real World Performance:

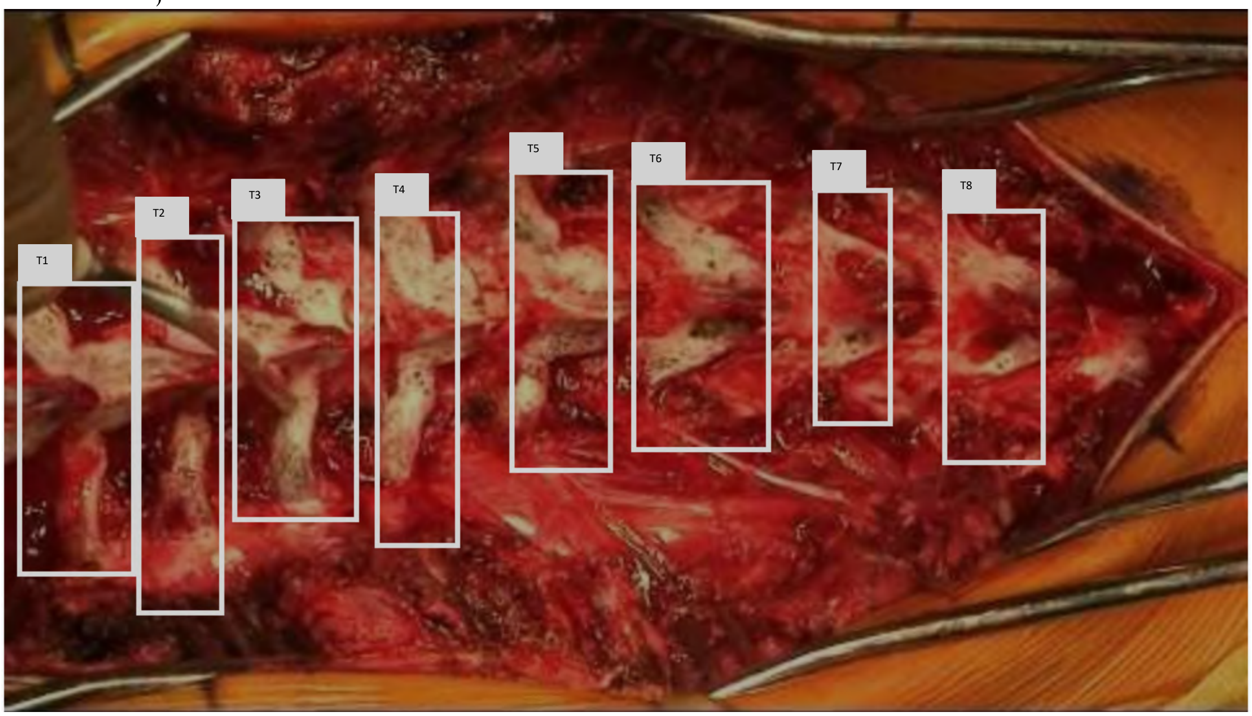
Spinal Cord Segmentation:

Figure 1: Spinal cord segmentation and mapping

Spinal cord segmentation was one of the primary objectives of the HDM algorithm. The segmentation is based off of the localization and identification of anatomical landmarks.

The identification was done in 3d and custom coordinates can be given to the software to view different angles. The data was cross-validated from real patient data during surgery and had a 98.6% accuracy, the segmentation of each spinal cord segment was within 1.33mm of the true value. The total processing time for segmentation and visual rendering took 1 minute and 18 seconds. Overall the performance of the spinal cord segmentation was both accurate and fast.

RCNN Object Identification:



The RCNN for object identification was accurately able to segment the field of view from the HoloLens. The headset was performing on average 27.3 frames per second. Some latency could be seen if the headset moved a significant amount but would normalize once image was partially stabilized. The algorithm was able to make 100 identifications a second with an average accuracy rate of 96.6%. The performance of the algorithm can be categorized as successful at identification and computationally efficient.

Final Words:
The overall performance of the collective algorithm was aptly able to develop a 3d model from the medical imaging provided, map the anatomical structures based off of IR data from the HoloLens headset, and perform object identification consistently with accuracy rates in the high 90s. The algorithms were computationally efficient enough to allow the HoloLens to computer all necessary components in ~88 seconds and battery life lasted the full four hours.

Discussion

The objective of this project was to develop an accurate navigational and diagnostic software system for spinal reconstructions. By implementing a novel combination of an HDM, SLAM, and R-CNN machine learning algorithms with implementation on the HoloLens augmented reality environment, the data suggests that the experimental diagnostic system was able to suggest the best surgical approach for reconstruction. The training and validation models were based of a series of medical scans from the University of Hawaiʻi at Mamoa, Oxford University, and the Imaging Clinic (Radiology) at Stanford University.

The control group (solely medical imaging data testing and virtual predictions) was tested first to analyze trends of spinal behavior with various structural forces being applied to it. The parameters that were analyzed through graphical data were spinal cord integrity, vertebrae integrity, nervous systems, and proximal muscle groups. Spinal cord integrity took into account regions of damage and patient's overall mobility with the given injuries. The spinal cord damage measurement served as the key indicator of surgical success, as the spinal cord is an integral part of patient mobility and outcome. Vertebrae damage allowed for three-dimensional understanding of anomalous global geometry, spinal behavior, segmentation nuances, and screw placement repercussions. To train the algorithm a powerful computational system with extensive memory to look through an extensive database of images and do complex analysis and manipulation of spinal data; for this purpose, I used *Google CoLaboratory* to offload all graphic intensive processing to googles off-site processing centers. In total, computational time took 32 hours of continuous training with a series of 5 GPUS each clocking at 8.9 teraflops each. Ultimately the dataset of over 500 patients with post-operative analysis proved beneficial to understanding the complications associated with spinal reconstruction and was more than adequate to develop the diagnostic device.

Once the algorithm was effectively trained with a confidence level of 98.6% and overall accuracy of over 96.6%, the next step was testing computationally efficiency to measure viability of software for implementation. The software was developed through a series of iterative design process that incorporated Autodesk Maya for layout design and Microsoft Visual Studios and Unity for code development. The software incorporated the HoloLens SDK for simultaneous localization and mapping and also used the unity windows development settings to create software to HoloLens. To fully develop the final version of the software that was both stable and computationally efficient took 26 design iterations. The beta programs had certain visual elements such as the RCNN identification bounding boxes, as well as the spatial domain visually present. As certain systems became more stable, they were eliminated from the development phase. The final software iteration included a 3D reconstruction of the spinal region with guiding lines and texting stating optimal method for operation. The algorithm on average used 89% of the HoloLens computing capability and thermals stayed within optimal conditions.

There were some practical and mathematical issues involved with predicting spinal behavior and segmentation of the spinal region. Since the morphology of every patient is different based off of weight, age, gender, condition, and other epigenetic factors. The RCNN faced difficulties with segmenting the spinal column as it has very similar visual and anatomical markers between each individual vertebra. As a result of the anatomical nature of the spine the RCNN on its own could not accurately identify the individual parts. As a result, both the SLAM and RCNN needed the HDM to perform the backend computations anatomical organization. Development issues included being able to develop machine learning algorithms to be accurate and computationally efficient to run on the Microsoft HoloLens. For the most part the algorithm was able to use the spinal template to precisely segment the spine into their regions with little inaccuracy. In addition to the algorithms, the dataset optimization was difficult as it included a variety of CT and MRI that required post processing to make the gaussian distribution mean zero. The algorithm is not very good at predicting highly skewed or damaged spinal columns.

Existing navigational systems and other novel solutions presented all require some form of radiology to guide the surgeon. The most common system employed in fluoroscopy is the O-arm and C-arm. The O-arm and C-arm are 2D / 3D intra-operative imaging system, that uses an X-ray system to navigate surgeons. The system is large and emits high dosage of radiation to operate. Other systems use anatomical cues along with radiology to create an effective navigation system. All of these systems require additional electronic accessories that are similar in profile to the mentioned fluoroscopy devices. No further studies have been done on novel navigation systems that don't require extensive radiology to operate. Most studies have been focused on the accuracy of medical imaging and improving resolutions

Applicable fields of this research include trauma patients, military applications, and general surgery. Most trauma patients related to the spine are in consistent acute pain, have some respiratory issues, and are usually not stable enough for highly detailed CT or MRI scans, but using simultaneous localization mapping and a low-resolution CT scan that only takes 5-10 minutes, an accurate anatomical map and surgical plan for reconstruction can be mapped in a reasonable amount of time. In case of military applications, most war zones do not have high level machinery available to preform highly accurate imaging or have navigational system presents. By using the Microsoft HoloLens, it makes the navigational system easily portable and transportable to different regions and its versatility to work under a variety of operating conditions. This also have applications to general spinal reconstruction and fusion surgeries by streamlining the navigation system and making it easily integrable with existing medical infrastructure.

The experimental diagnostic system was able to predict with 96.6% accuracy the correct spinal behavior and optimal screw placement with 1.33mm accuracy. A machine learning and computer vision based navigational and surgical aid for spinal reconstruction has significant implications of radiology and navigation for spinal surgeries. The data suggests that the implementation of the novel diagnostic system would decrease operating time, recovery time, and surgical costs.

Bibliography

American Association of Neurological Surgeons. “Spinal Reconstruction Analysis”. AANS, 17 May 2018

Durrant-Whyte, Hugh, and Tim Bailey. “Simultaneous Localization and Mapping: Part I .” *IEEE Robotics &Automation Magazine*, IEEE, 23 June 2006.

Meißner, M, et al. “Volume Visualization and Volume Rendering Techniques.” *Volume Visualization and Volume Rendering Techniques*, University of New Hampshire, 15 Feb. 2000.

Badalyn, Alex. “Unity Logo.” *Dribbble*, Dribble, 29 Aug. 2016.

Cai, Yunliang, et al. “Multi-Modal Vertebrae Recognition Using Transformed Deep Convolution Network.” *NeuroImage*, Academic Press, 8 Apr. 2016.

Cai, yunliang, et al. “Multi-Modality Vertebra Recognition in Arbitrary Views Using 3D Deformable Hierarchical Model.” *An Introduction to Biometric Recognition - IEEE Journals & Magazine*, Wiley-IEEE Press, 10 Nov. 2009.

Lu, Zhang, et al. “SpineNet: Automated Classification and Evidence Visualization in Spinal MRIs.” *Current Neurology and Neuroscience Reports.*, U.S. National Library of Medicine, 12 Oct. 2017.

Klinder, Tobias, et al. “Spine Segmentation Using Articulated Shape Models - Semantic Scholar.” *Semantic Scholar*, Semantic Scholar, 1 Jan. 1970.

Kumin, Michelle. “Disc-Degeneration and Vertebra-Deterioration Database.” *Research*, Oxford University, 1 Sept. 2011.

Leener, Benjamin De, et al. “SCT: Spinal Cord Toolbox, an Open-Source Software for Processing Spinal Cord MRI Data.” *NeuroImage*, Academic Press, 5 Oct. 2016.

MayaHowTos. “Maya Learning Channel.” *YouTube*, YouTube, 29 Jan. 2019.

Mayo Clinic. "Navigational Systems and Spinal Reconstruction” *Department of radiology*, Mayo Clinic, 11, July 2016.

Microsoft . “Visual Studio Products.” *Visual Studio*, Microsoft Corporation, 1 Jan. 2019.

Ng, Andrew. “Generative Learning Algorithms.” *CS229 Lecture Notes*, Stanford Univeristy, 6 Sept. 2017.

Stanford Unviersity. “Stanford MSK MRI Atlas (c) 2018.”Cervical *Neck - Shaft Angle*, Stanford (Radiology) Imaging Clinic, 5 Mar. 1997

University of Hawaii at Mamoa. “University of Hawai'i at Mamoa Radiology Database.” *Hawai`i Creole English*, English Language Institute, 30 Apr. 1999.

Zeller, Matt, et al. “Spatial Mapping - Mixed Reality.” *Microsoft Docs*, Microsoft Corporation, 20 Mar. 2018.