Jonathan Young, Damiano Marsilli, Arijit Nukala Computer Integrated Surgery

**Automated Spinal Segmentation and Remote Monitor Calibration for Surgical Assessment**

(Project 8)

Readings:

1. "SpotTune: Transfer Learning through Adaptive Fine-tuning"
2. “A novel IMU-based clinical assessment protocol for Axial Spondyloarthritis: a protocol validation study”
3. “A Cloud Scalable Platform for DICOM Image Analysis as a Tool for Remote Medical Support”

## Project Summary

Our project aims to develop a method to calibrate IMU sensors placed on the spine of patients utilizing computer vision. The collection of this data will result in a model approximating the behavior of the spine, giving the doctor the ability to assess whether or not a patient is eligible for a given spine surgery. Deployment of this process first entails utilizing monocular footage acquired from a camera to record a patient moving per a pre-defined set of movements. From this footage, spinal key points and curves will be approximated utilizing a deep-learning approach and interpolation. Afterward, the orientation of the IMUs, alongside their physical acceleration data, is combined via a transfer function. Once the calibration is complete, the sensor readings are acquired, and a subsequent model of the spine in a 3D volumetric model of the patient is generated for the doctor to make a prognosis. The last component of the project is migrating all components to the cloud and creating a pipeline to process that information. The rationale for doing cloud deployment is that the data of the patients will be collected at home. Therefore, for doctors to interact with the data, the data must be sent to the cloud, processed, and the subsequent result is sent to the doctor for visualization.

## Paper 1: "SpotTune: Transfer Learning through Adaptive Fine-tuning"

### Paper Selection

This paper was selected due to its novel approach to few-shot transfer learning, which is a problem constraint we will necessarily be working under for the spinal keypoint and curve estimation modules. The approach presented in the paper demonstrates improved accuracy in transfer learning for computer vision tasks compared to the standard, the static approach of freezing (leaving unchanged) the feature-extracting backbone, and training a separate classification head. The improvement in accuracy was particularly notable as the size of the transfer dataset decreased. In our case, we have a very limited dataset of annotated spinal keypoint videos and will need to train a keypoint estimation head through transfer learning. We will explore using the approach presented in the paper to improve the transfer learning of our spinal keypoint estimation function, which encapsulates our minimum, and one of our expected deliverables.

### Technical Summary

The paper presents a fine-tuning strategy that decides, for each sample, which layers of the pre-trained model should be frozen and which layers should be fine-tuned (adapted for the transfer task), to improve accuracy on the transfer task. A diagram of the approach is shown in Figure (5). To this end, the network is first segmented into a series of 'blocks'. The level of granularity of the blocks (neutron-level, layer-level, or multi-layer-level) is left as a hyperparameter and is not described.

Next, a separate policy network is trained to output routing decisions for each block of the network. At each block, there are two discrete routing decisions: freeze or finetune. This discrete decision is modeled by a Gumbel distribution, which is given by $G = -\log(-\log(U))$, where U is a uniform distribution. As this decision is discrete, the process is not differentiable and thus hard to optimize. To circumvent this issue, the policy network is trained to optimize the Gumbel-Softmax distribution, which is a differentiable function. Then, this policy model is jointly trained with the target fine-tuning task through standard backpropagation and gradient descent. By jointly training the policy network with the underlying network, the policy network is optimized to produce the best routing policy for the target task. At test time, the policy network determines, for a specific instance, if the instance should be routed through the pre-trained blocks or the fine-tuned blocks.

## Assessment

The method proposed achieves the above state-of-the-art performance on the target task of image recognition. Notably, the approach outperforms manually fine-tuned models, where previous researchers identified domain shifts in the distribution and manually selected layers to fine-tune. Thus, the model effectively presents the potential for autonomous fine-tuning, allowing us to focus on the overarching integration of our modules and not get bogged down in manual fine-tuning. Another key strength of the paper is that the approach is very well-described and should be simple to reproduce. However, one limitation of the paper is that the evaluation of the approach is limited to the target task of image recognition. Although one can expect an approach to transfer to other downstream tasks, as the focus is on tackling the domain shift, this is not empirically proven. Similarly, the method is image-specific, and a video analog will need to be developed for use in our approach.

## Relevance and Next Steps

The method presented in this paper will be used as the transfer-learning approach used to train the spinal keypoint estimation module from a pre-trained pose estimation module. This module will be responsible for estimating spinal key points on a patient from monocular video. These key points, paired with an interpolated spinal curve, will be used in our transfer function to calibrate the IMU. The calibration approach and background reading related is covered in Paper 2. As the approach relies on the spinal key points, the approach presented in SpotTune must allow for the learning of accurate spinal key point estimation.

## Paper 2: “A novel IMU-based clinical assessment protocol for Axial Spondyloarthritis: a protocol validation study”

### Paper Selection

This paper was selected due to its simple approach to computing spinal range of motion (ROM) angles from IMUs placed at key positions on the spine. The author’s simple approach to calculating spinal angles using quaternion mathematics and rotation matrices that we learned in CIS 1 is simple to understand and serves as an excellent reference for our team to use as we develop our transfer function approach. Additionally, the IMUs used in this study are of the same type and from the same company as the IMU our team is using.

### Technical Approach

The author’s goal was to collect accurate spinal angle data to asses Axial Spondyloarthritis. IMUs are a lower-cost and more accurate method of analyzing spinal ROM than objective clinician measures. This study differs from other studies that use IMUs to measure joint ROM by implementing simple functional movements for patients in a novel IMU-based protocol that validates IMU data against a gold-standard optical motion capture system.

In the experimental setup, healthy middle-aged participants between 18 and 40 years old had Delsys wireless IMUs attached with optical reference markers to the patient’s back at specific spinal segments (Figure 3). The IMUs used a proprietary Kalman filter provided by Delsys to output quaternion orientation vectors. Each participant had 5 IMUs attached using double-sided tape. The functional movement protocol included trunk flexion, trunk extension, trunk lateral flexion, cervical rotation, and cervical flexion/extension to collect IMU ROM data. These movements were repeated 3 times and performed slowly with the patient’s goal to reach the maximum ROM.

Spinal angles were calculated using the orientation of the spinal body IMU to the global reference system (Figure 4). Kinematic constraints were applied to correct vertical axis angular drift and to improve angle estimation. The primary constraint was the assumption that FG rotation as shown in Figure 4 constrained movements to lie only on the global frontal plane. This constraint assumes the sensor positions reflect the actual tilt of the spine and that the sensor is always aligned with the global reference system's X-axis. These constraints were applied at time zero. Additionally, infrared motion capture using clusters attached to the IMU served as a reference for the spinal angle calculations obtained because motion capture systems lack drift.

The joint angles were obtained using a quaternion expression and transformed into clinical angles. Lastly, the spinal ROM and kinematics were computed in MATLAB by subtracting the spinal angle at the beginning of the movement segment from the spinal angle at the end of the motion segment. The authors found that IMU measurements were significantly correlated with motion capture measurements and that segment angle error was low.

Assessment

This study's use of IMUs and motion capture to calculate spinal ROM shows that IMU results have a high degree of agreement with gold-standard motion capture systems. The results of this study outperform results from previously relevant literature. Meaningfully, the study describes measurement errors in each plane, showing that sagittal plane movements were more accurate than movements on the frontal and transverse planes. Additionally, the procedure in this paper is clearly described and could be readily replicated in future studies. However, this paper could be improved by increasing the patient sample size. Additionally, angle errors due to inaccuracies in vertical movement calculation and sensor drift from the IMUs could be better addressed with a specific solution rather than a simple description of the errors.

Relevance and Next Steps

The paper provides specific metrics of accuracy and consistency of measurement when compared to gold-standard measurement tools and provides suggestions on how IMU calculations can be improved given the set of IMUs and Kalman filters we plan on using. The basic calculations and kinematic constraints provided by the paper will allow us to calculate the final spinal ROM metrics from our IMU outputs as our team plans on implementing this paper's method of calculating spinal angles using global reference frames concerning the IMU locations. However, our method will provide an improvement over this study by enabling accurate angle measurement given the inaccurate placement of the IMU. Using the transfer learning approach described by “SpotTune”, we plan on creating a spinal keypoint detection algorithm using collected and annotated training data to shift the IMU’s orientation calculations to segmented spinal key points. This allows relatively close placement of the IMU without requiring the IMUs to be perfectly straight along the spine. While we plan on using the kinematic constraints for calibration given by the paper initially, our final plan is to improve upon these constraints using the same transfer learning outputs by calculating initial orientations of the spinal IMUs given the key points and an interpolated spline, allowing for increased accuracy of ROM calculations. This prevents the IMUs from having to zero their orientation during calibration, which can be problematic for patients with spinal issues where the spine during “normal” standing often is not straight.

## Paper 3: "A Cloud Scalable Platform for DICOM Image Analysis as a Tool for Remote Medical Support"

### Paper Selection

This paper was selected due to its approach to uploading, processing, and producing visual analysis of medical data on cloud-based platforms. The paper's approach enables our team to understand best how to link the various cloud systems in Amazon Web Services (AWS) together to enable a seamless upload of the data, to quickly process the data, and to create a method of visualizing the data on computationally weak devices by diverting the computationally intensive tasks to the cloud. Efforts in this area contribute to our maximum deliverable: a complete cloud pipeline of uploading, processing, and generating an output of the IMU and camera data via hosting the tasks related to the previous two papers on the cloud.

### Technical Approach

The author's work involves exemplifying a cloud architecture that processes DICOM data to generate results for the doctor to review. The authors created this architecture to demonstrate the viability of uploading medical data to the cloud, processing it, and producing output for doctors without reliance on heavy computational machinery.

The platform utilized to demonstrate this pipeline was Amazon Web Services (AWS) which is HIPPA-compliant for medical applications. The authors chose to upload the DICOM repository to an S3 bucket – AWS's storage system – to reduce the amount of data transferring involved since all processing occurs on the cloud. Moreover, an administration module called m3DicomAdmin is created to authenticate users and initiate tasks request, specifically those that move data between AWS services [3].

The process for utilizing DICOM data first begins with the user's authentication. Then, the user uploads a given DICOM file or utilizes one already within the S3 buckets – AWS storage systems – and a request by the m3DiacomAdmin to the EC2 instances – AWS's virtual machine – is sent to start processing the data. M3DicomAdmin then allocates the necessary resources based on a pre-defined configuration of CPU clusters defined by the user. Each type of task has a different CPU cluster configuration. The image analysis is then sent back to another S3 bucket for outputs as an X3D data file, and the user is subsequently notified that the results are ready. Afterward, a web client receives the information generated from the S3 output bucket and subsequently creates a 3D model and interface with HTML and JQuery mobile framework. To enable more flexibility, the interface also possesses the ability to retrieve additional information relating to the DICOM file to gain additional insights about the file's subject. This is also done through a request in the cloud. [3]

To show the real-world usage of their system, the authors chose two use cases to exemplify their system. The first is a mammography analysis system, in which a user uploads a mammogram DICOM file. The authors sought to exemplify using the cloud to process images and locate critical points and areas of interest. The system then proceeds to segment regions of interest (ROIs) and extracts features from them and attempts to locate micro-calcifications, which are typically associated with breast cancer (figure 1). The image is then returned to the user for medical assistance in a diagnosis. The second use case is a cloud-based system visualization of a patient's jaw.  In this use case, the authors sought to demonstrate how the cloud can be used for more complex tasks such as 3D visualization of an object. A DIACOM file of a patient's jaw is sent through the AWS pipeline, and an X3D data file is subsequently sent to the user for visualization through HTML5 (figure 2) [3].

<u>Assessment</u>

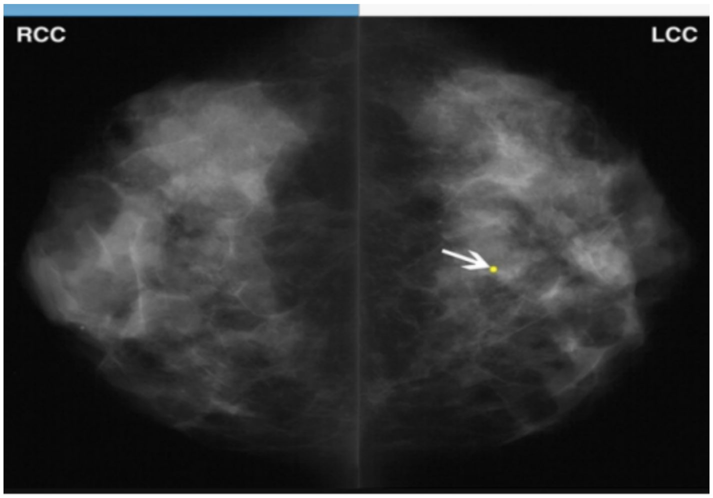
The paper provides an adequate overview and demonstration of the utilization of cloud services to produce an outcome that assists doctors in making a prognosis. The content of the paper is specific in detailing which services and systems are integrated and utilized to produce a coherent pipeline. Another strong point of the paper is that it provides examples of use cases of their system, specifically breast cancer and jaw visualization. However, the paper would benefit from detailing more information about how their cloud system is arranged and which services are used between S3 and EC2. It is not clear if the authors utilized lambda functions, queue systems - SNS and SQS, as well as other low-level cloud services. This is important when considering how to scale this pipeline or dynamically changed it for other use cases. Furthermore, it would be helpful to know how what devices are being used to view the output, as the output of the system requires a delicate balance between relying on a local device versus the cloud for computational needs.

## Relevance and Next Steps

This paper bears strong relevance to our project, as our team attempts to deploy our pipeline completely to the cloud - our maximum deliverable. Our configuration is similar to those of the paper. First, we seek to upload the data to the cloud and process it. Then, the cloud generates the output, a computer model, which can be viewed through a web-based application. Our team heavily relies on the cloud for a successful product, because our mentor's team product collects data at home.

Some next steps for this work, some of which are relevant to our project, largely focuses on enhancing the runtime and migrating the processed output to the cloud. Exploration of the former includes developing an algorithm that modifies the clusters to be more efficient when processing information or utilizing other AWS services instead of EC2 instances, potentially reducing costs. The latter involves exploring how to migrate visualization of data onto the cloud. The current paper visualizes medical data locally, but visualizing it on the cloud could prove beneficial in terms of enabling more computationally expensive visualizations and enhanced functions to be available for the user. However, cloud-based visualization could result in the choking of resources and may be slow due to connection speeds. Analysis of the drawbacks versus benefits must be done to decide how many cloud resources to use.

# Appendix



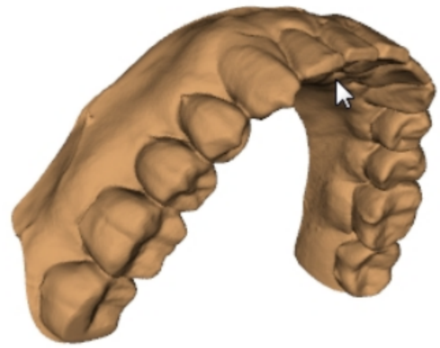



**Figure 1: Breast Cancer Visualization [3]** **Figure 2: Jaw Visualization [3]**

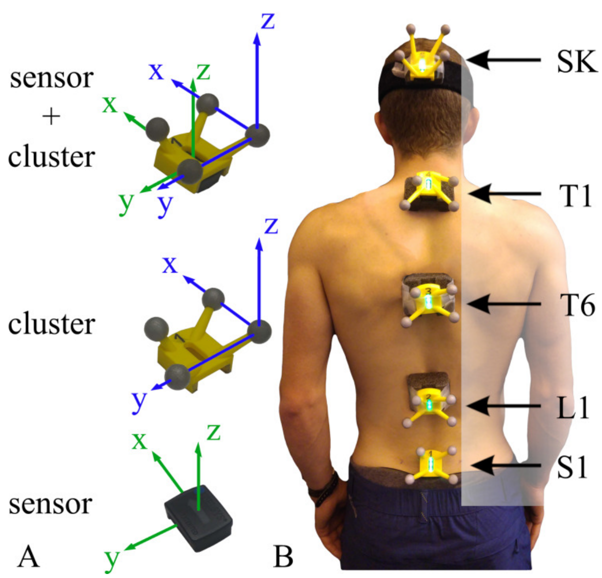



**Figure 3: IMU Sensor and cluster placement [5]**

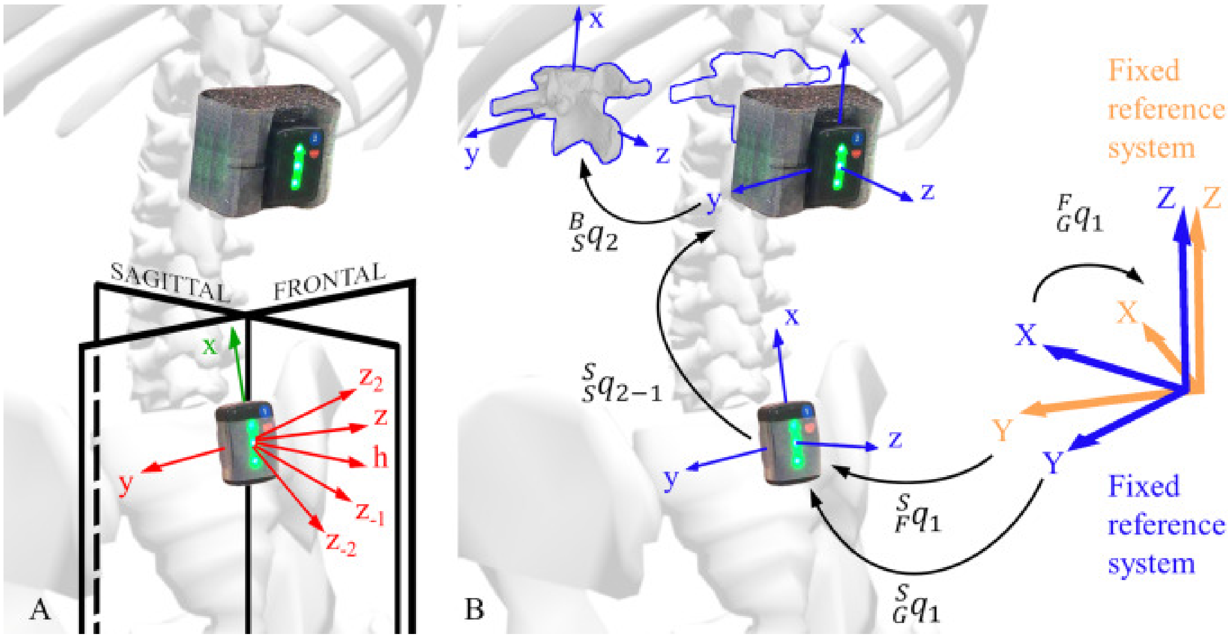



**Figure 4: Kinematic constraints and rotations between IMU reference systems [5]**

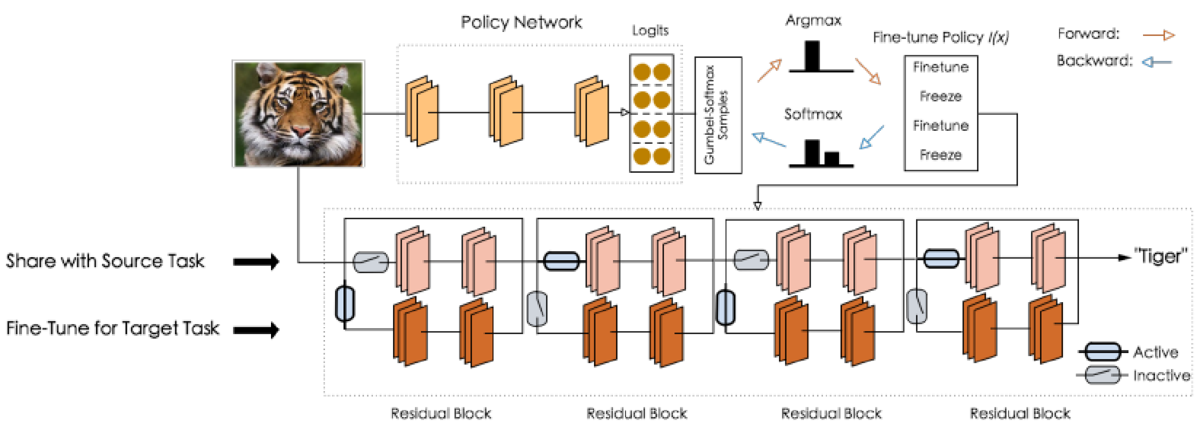



**Figure 5: SpotTune Approach Architecture [4]**

## References:

[1] C. O. Rolim, F. L. Koch, C. B. Westphall, J. Werner, A. Fracalossi and G. S. Salvador, "A Cloud Computing Solution for Patient's Data Collection in Health Care Institutions," *2010 Second International Conference on eHealth, Telemedicine, and Social Medicine*, Saint Maarten, Netherlands Antilles, 2010, pp. 95-99, doi: 10.1109/eTELEMED.2010.19.

[2] Bahga A, Madisetti VK. A cloud-based approach for interoperable electronic health records (EHRs) *IEEE J Biomed Health Inform.* 2013;17(5):894–906.

[3] Ojog, I., Arias-Estrada, M., Gonzalez, J. A., & Flores, B. (2013). A cloud scalable platform for DICOM image analysis as a tool for remote medical support. In *The Fifth International Conference on eHealth, Telemedicine, and Social Medicine. France*.

[4] Guo, Yunhui, et al. "Spottune: transfer learning through adaptive fine-tuning." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.

[5] Franco, Luca et al. "A novel IMU-based clinical assessment protocol for Axial Spondyloarthritis: a protocol validation study." PeerJ vol. 9 e10623. 26 Jan. 2021, doi:10.7717/peerj.10623