Virtual Rotator Cuff Arthroscopic Skill Trainer: Results and Analysis of a Preliminary Subject Study

Jake Farmer
University of Central Arkansas
201 Donaghey Ave
Conway, AR 72035
+1(501) 450-5000
jfarmer2@cub.uca.edu

Mustafa Tunc
University of Central Arkansas
201 Donaghey Ave
Conway, AR 72035
+1(501) 450-5000
mtunc1@cub.uca.edu

Daniel Ahmadi
Pulaski Academy
12701 Hinson Road
Little Rock, AR 72212
+1(501) 604-1910
danielahmadi2002@gmail.com

Doga Demirel
Florida Polytechnic University
4700 Research Way
Lakeland, FL 33805
+1(863) 583-9050
ddemirel@floridapoly.edu

Tansel Halic
University of Central Arkansas
201 Donaghey Ave
Conway, AR 72035
+1(501) 450-5000
tanselh@uca.edu

Sreekanth Arikatla
Kitware, Inc.
101 East Weaver Street
Carrboro, NC 27510
+1(919) 969-6990
sreekanth.arikatla@kitware.com

Sinan Kockara
University of Central Arkansas
201 Donaghey Ave
Conway, AR 72035
+1(501) 450-5000
SKockara@uca.edu

Shahryar Ahmadi
University of Arkansas for Medical Sciences
4301 W Markham Street
Little Rock, AR 72205
+1(501) 686-7000
SAhmadi@uams.edu

ABSTRACT

This works presents a design and development of Virtual Rotator Cuff Arthroscopic Skill Trainer (ViRCAST) and its preliminary subject study analysis using machine learning approach. Arthroscopy is a minimally invasive surgical intervention regarded as a part of orthopedic sub-specialty. The procedures are performed via small incisions in the patient's skin to examine, diagnose and repair the injuries inside a joint [1]. Surgeons insert tiny instruments and small lens and lighting (called arthroscope) into the joint. They perform surgical intervention seeing the anatomy on a 2D monitor screen streamed from arthroscope. Due to non-natural hand-eye coordination, narrow field-of-view and limited instrument control, training for arthroscopy is challenging and difficult to master. In this work, we developed a primarily ViRCAST platform for training the shoulder arthroscopy procedures. We performed initial validation study using 10 surgery resident subjects (Post-Graduate Year (PGY) 1-5) and performed statistical analysis to extract significant data features. This is followed with machine learning algorithms to cluster and classify the subject's expert level with training data. Our results show that we could successfully distinguish the expertise level.

CCS Concepts

ICISDM 2020, May 15–17, 2020, Hawaii, HI, USA

ACM ISBN 978-1-4503-7765-2/20/05…$15.00
https://doi.org/10.1145/3404663.3404673

Human centered computing → Human computer interaction (HCI) →Interactive systems and tools

Computing methodologies→ Modeling and simulation → Simulation evaluation.

Keywords

Surgical Simulator; Classification; Clustering; Validation Study

1. INTRODUCTION

Arthroscopy is a minimally invasive surgical intervention that is performed via small incisions at a joint [1]. Surgeons insert small surgical instruments and see the anatomy on a 2D monitor screen using arthroscope. Arthroscope is a fiber optic camera with accompanied with a rotatable light source. The arthroscopy has gained significant recognition over the years and become de facto and authentic procedure for the treatment of the various ailments such as bursitis, labral tears, repair and resection of torn cartilages (e.g. osteoarthritis) and ligaments, removal of inflamed synovial tissue, reconstruction of anterior cruciate ligament [2], [3]. It is commonly used especially in rotator cuff tear treatments. The rotator cuff is a group of muscles and tendons located in the shoulder that connects the humerus (upper arm) to the scapula (shoulder blade). The rotator cuff tendons and muscles provide stability and rotational motion of the shoulder. Each tendon of these muscles attaches to the humerus and extends to the scapula. These tendons create a cuff formation around the humerus. Rotator cuff tear is basically an injury of this cuff formation. Surgical solution is essential in order to eliminate the persistent symptoms [4].

In arthroscopy, the surgeons need to assess and approach the rotator cuff from several different angles to fully delineate the tear pattern and then repair it anatomically. However, due to unintuitive hand-eye coordination, narrow and confined field-of-view and confined space for instrument control, training for arthroscopy is challenging and difficult to master. Conventional training regimen in surgery practice methods such as cadaver, plastic mannequin, and apprenticeship training are not inadequate. The use of animals is unethical and anatomical differences reduce its fidelity. Cadavers are costly and mostly limited to single use. Plastic mannequins are unrealistic and can be used for only certain type of surgeries. Likewise, physical bench models can be useful for arthroscope navigation and instrument handling, but they do not possess the realism to teach the surgeons the joint anatomy nor aid surgical decision-making [5]–[7]. There are prior efforts from both academia and industry (e.g. Insightmist, ARTHRO Mentor, VirtaMed ArthroS [8]–[10]) that aim to fill the gap in arthroscopy procedure training. Although some of these works have undergone human subject validation, they have limitations in providing realistic physics-based interaction with arthroscopic instruments, physical instrument interfaces, various difficulty settings for proficiency, photo realistic rendering, comprehensive performance and validation studies with extensive analysis for the entire or each arthroscopic surgery steps. We therefore proposed our ViRCAST platform that aims to deliver highly realistic arthroscopy training for shoulder with rigorous validation study involving human subjects. We developed preliminary surgery simulation with haptic feedback and performed validation study for arthroscope navigation task and shaving task.

2. ViRCAST Design

ViRCAST was developed using Software Framework for Multimodal Interactive Simulations (SoFMIS) [11], a highly customizable, multithreaded simulation framework. SoFMIS allows for a quick and modular approach to creating visually realistic simulations with easy integration for haptic devices, external interfaces such as Arduino or any data acquisition (DAQ) devices, and simplistic data recording. This framework allows for easy extension of modules contained within it simply by extending classes and providing a custom implementation to fit the developer's current needs. Many aspects of the simulation, such as rendering, physics simulation, object properties, and scene environments are encapsulated, and in some respects abstracted from the developer for simplicity sake. We are in the process of porting the modules to the iMSTK (imstk.org) that is an open source and evolved and advanced version of SoFMIS that has been developed with multi-institutional efforts including Rensselaer Polytechnic Institute, University of Central Arkansas, and Kitware Inc.

In ViRCAST, we have used Physically Based Rendering (PBR) approach for creating realistic scene. We used specific PBR textures, such as albedo, roughness, normal, metallic, and ambient occlusion maps to provide surface characteristic. Those parameters are used to calculate realistic lighting based on the main spotlight of the scene, which is positioned at the virtual arthroscopic camera. The bidirectional reflective distribution function (BRDF) is used to calculate the final look of the scene (see Figure 1) [12].

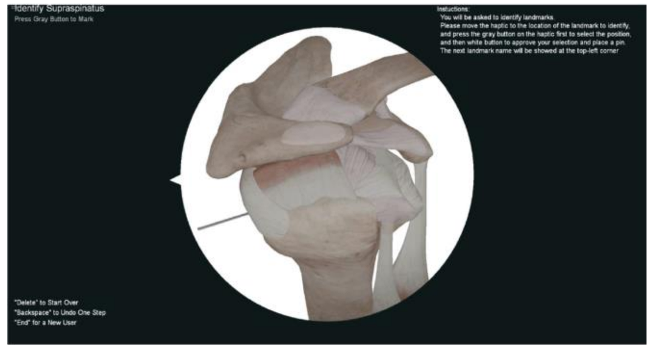

Figure 1 Arthroscopic view from ViRCAST.

It takes parameters such as the view direction, incoming light direction, and surface roughness and computes the light contribution to the final look of the model based on the surface roughness and reflectivity parameters defined in the roughness and Metallic map. The 3D Zygote shoulder and custom models (e.g for bursa, rotator cuff muscles, additional ligaments etc) are used for the geometry. The textures of the models are completely recreated for PBR. We used forward rendering mechanism. Once the light computations are performed, we perform post processing iteration that is to enable depth of field, arthroscopy lens effect other image-based effects such as bubble generation (seen due to fluid in the shoulder). The alterations of 3D models are generated with geometry language to support various tear shapes, sizes and length [13], [14].

In ViRCAST, the use of the actual arthroscope is important factor to mimic the real operator theater. We used 3DSystems Geomagic Touch haptic devices to control camera and tool movement in the virtual scene. To achieve true immersion, actual surgical tools that are used in the real surgery have been modified to work in conjunction with the haptic devices to give a better sense of realism for the user. An arthroscope was modified to house a Bournes EMS22A rotational encoder [15] to read the absolute rotation of the arthroscopic light source in order to accurately control the camera. The arthroscope is fixed to a haptic device with 3D printed adapters. A second encoder is also sued to control the camera rotation in the scene. This allows for natural movements and realistic control of the virtual arthroscope. Interactive physics simulation in ViRCAST is achieved using the NVidia PhysX library [16]. In ViRCAST, the tear model is created as a volumetric soft body. This allows manipulation of the tear structure in real time. The surgeons could able to assess the tear type and evaluate it for diagnostic and repair procedures. We fixed the tear model at the shoulder blade for limiting its movement and flexibility to simulate realistic behavior. Our ViRCAST surgical platform with haptic devices and attached instruments of can be seen in Figure 2. Our simulation car can be adjusted for two most common arthroscopy positions such as beach chair and lateral decubitus positions [17].

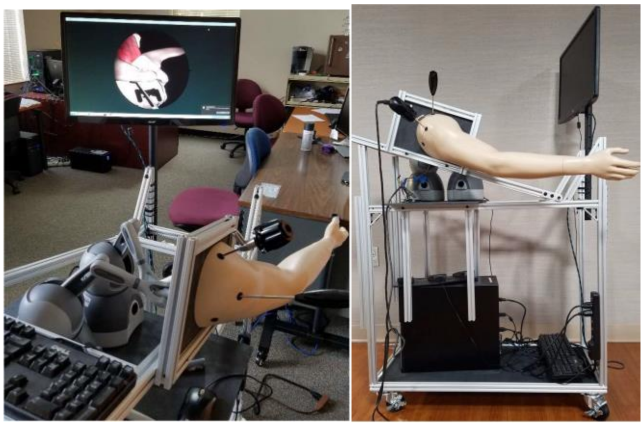

Figure 2 ViRCAST simulation cart and simulation view adjusted for beach chair position in arthroscopy.

3. VALIDATION STUDY

We validated our simulator with human subjects at the University of Arkansas for Medical Science (UAMS) with approved IRB protocols numbers; 205864 and 239772. The study involved a total of 10 participants. Prior to the beginning of designated task, the subjects were given a pre-questionnaire with questions about age, level of experience, hand dominance, etc. Following the questionnaire, the users were verbally told about the task. Each subject was given certain amount of time to get familiar with the task and ViRCAST environment. Then the subjects were expected to complete the landmark identification and shaving tasks without any certain time limit. Once the tasks were completed, the subjects were given a post-questionnaire with questions regarding the realism and effectiveness of the simulator as well as open-ended questions for feedback on the simulator.

3.1 Data sets

For each participant, we recorded instrument positions in virtual scene distance units (mm), forces exerted forces on the anatomical structures in newton, arthroscope motions, virtual pin locations placed on anatomical structures and distance to the actual locations of anatomical structures, velocity(mm/s) and acceleration (mm/s^2) of the tools (e.g. probe and shaver), shaved regions, shaving speed, and arthroscope motions, time to complete the task and its sub tasks (e.g. time for each landmark). We also post process data to compute measures for additional metrics [18]. These include average, mean, median velocity, jerk, turning angle in instrument motions, path length, soft body motion etc. We subdivide the subjects' data intro two expertise level; expert and novices. Since the number of the number subjects are limited, we classify them based on the number of years attending such as PGY1-3 and PGY 4-5 and as a criterion. This criterion was also determined after several runs of sample data clustering algorithms.

4. METHODS

In order to simplify the data analysis, reduce the computation time and increase the applicability of the well-studied clustering/classification algorithms, we first populate all the candidate features and then reduce feature size for further analysis. The initial feature set included the computed features as well. We then applied Welch's T-Test to determine the most significant features that might find the distinction among the expertise level. The overall steps of the selection process can be seen in Figure 3.

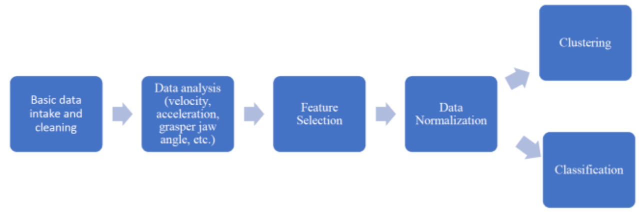

Figure 3. Feature reduction flow chart.

As an output of our feature selection process, we determine that Mean Tool Velocity (mm/s), Std. Dev. Jerk (Camera) (mm/s^3), Std. Dev. Acceleration (Camera) (mm/s^2), Std. Dev. Velocity (Camera) (mm/s), Mean Jerk (Camera) (mm/s^3), Mean Acceleration (Camera) (mm/s^2), Mean Velocity (Camera) (mm/s), Time Taken (seconds) (such as seen Figure 4) are the most important features ($p < 0.05$).

We normalized the data set in order to compare the features between each other and eliminate the bias due to saturation in data. This is to improve the accuracy of the analysis (e.g. removing the large skew in data). We used SciKit Learn and its preprocessing modules for normalization and machine learning [19]. We employed three normalization methods such as Z-score, Min-max, and absolute value.

We first tested to see if the clustering the data in two groups can be confirmed. We used several metrics to quantify the clustering results, such as internal and external indices; Silhouette Score, Adjusted Rand Index, Fowlkes—Mallows Index, the Jaccard Score, and the Mutual Information Index. We further want to determine that classification of the data is attainable with the known pre-determined clusters. We used K-Nearest Neighbors (n = 2), Logistic Regression, and SVM with both linear and radial basis function (RBF) kernels and looked into the at the precision, recall, F1 score, and the average accuracy. The data was then clustered and classified 100 times, with different test / train splits for the classification data, thus performing k-fold cross validation of the classifier performance.

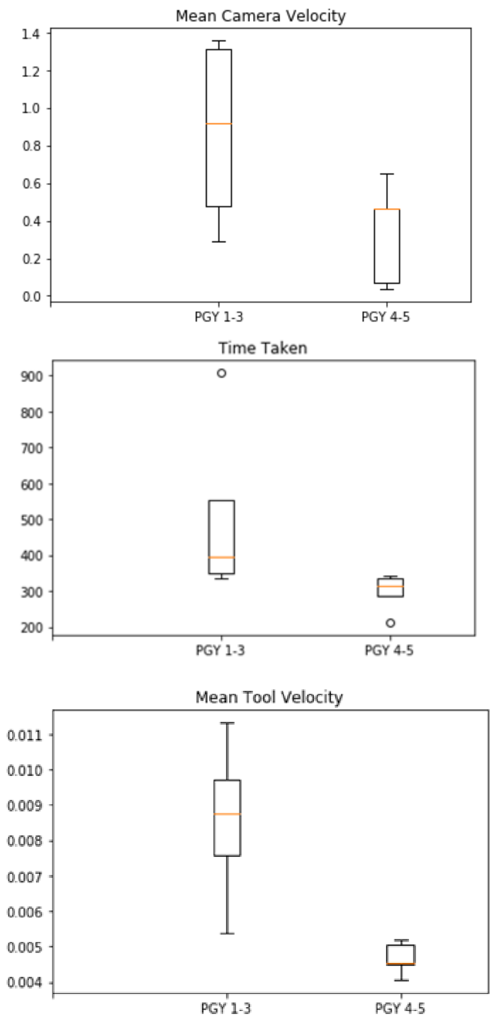

Figure 4. Boxplots of selected features for arthroscope velocity, time taken for the tasks and tool mean velocity.

5. RESUTS

Based on survey questionnaires, expert group rated the realism of the 3D anatomical models as 4.2 out of 5 (likert scale) in average. They scored 4.4 out of 5 for the realistic visual rendering which is very critical for the landmark identification task. The anatomical correctness of the tear was noted high 4.2 as well. These results are derived from the post-questionnaire.

Normalization of the data provided significant improvement over K-Means and Spectral clustering compared to data without normalization step. The highest accuracy is attained with the Agglomerative clustering. The best results are noted with Fowlkes-Mallows score index in all algorithms except the Agglomerative clustering. The score for each clustering algorithm can be seen in Figure 5 for absolute data. In the classification algorithms, we saw significant accuracy in the K-Nearest Neighbors algorithm. We were able to achieve up to 81% correct classification on the PGY 4-5, and up to 89% classification accuracy on the PGY 1-3, which can be seen in Figure 6.

All of the classification algorithms worked very well for our case. It is noted that the classifiers seemed to work better with Z-Score standardization. In datasets, we can see the weakness of the RBF kernel of the SVM, due to a small dataset that experience over- or under-fitted the data.

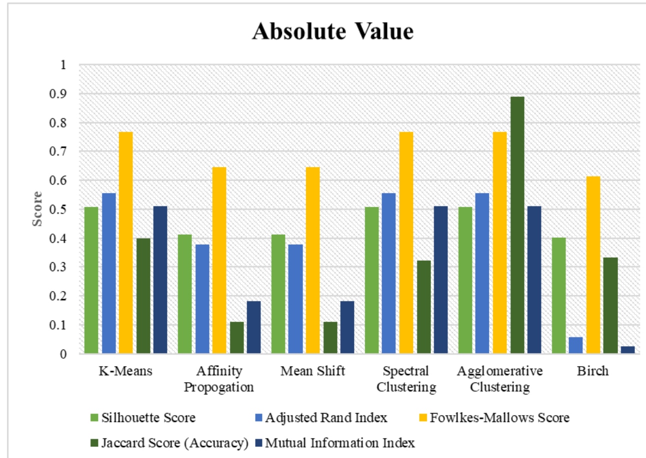

Figure 5. Clustering based on max absolute value normalized data

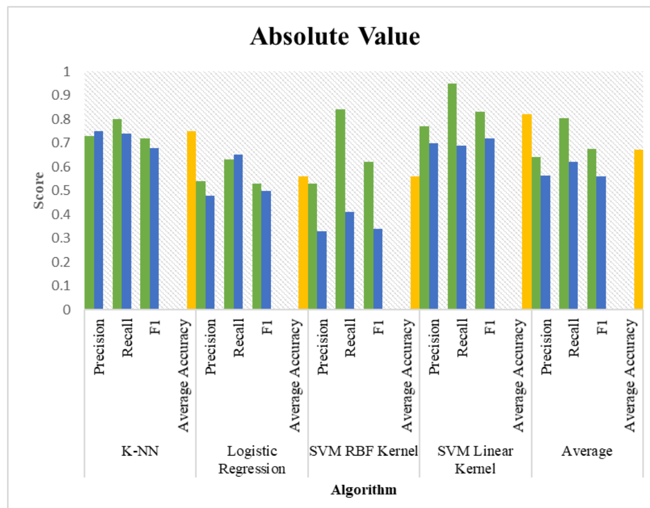

Figure 6. Classfication accuracy based on max absolute value normalized data.

The most prominent information that we derive is the difference in the movement between the left and right hand of the expert and novice groups. We identified significant differences in the tests on the use of the left hand. These results convey the experts have more mastery of ambidexterity than the novice surgeons. This is extremely important as the expectation from arthroscopist to use both hands efficiently. We also identified that experts were less concentrated on the movements of their non-dominant hand determined with turning angle and acceleration features. We might be related to the confidence. However, bigger sample size with more data is needed to verify this hypothesis,

6. CONCLUSION

In this work, we presented preliminary ViRCAST design and results of the case study which we performed to validate the effeteness of our simulator. Our subject study was carried out among orthopedic surgery residents ranging from PGY1-5. Based on the post-questionnaire in the study, we determined that our simulator was found to be effective for arthroscope navigation task and landmark identification. We employed clustering algorithms over the simulator data to first understand that splitting of groups in expert and novice is feasible. We later showed that the classification based on two clusters could provide successful results in understanding the skill level. As a future work, we plan to extend the study with more sample size and including additional classification techniques.

7. ACKNOWLEDGMENTS

This project was made possible by the Arkansas INBRE program, supported by a grant from the National Institute of General Medical Sciences, (NIGMS), P20 GM103429 from the National Institutes of Health (NIH). This project was also supported by NIH/NIAMS R44AR075481-01, NIH/NCI 5R01CA197491 and NIH/NHLBI NIH/NIBIB 1R01EB025241, R56EB026490.

8. REFERENCES

[1] R. Treuting, “Minimally Invasive Orthopedic Surgery: Arthroscopy,” *Ochsner J*, vol. 2, no. 3, pp. 158–163, Jul. 2000.

[2] A. C. Colvin, N. Egorova, A. K. Harrison, A. Moskowitz, and E. L. Flatow, “National Trends in Rotator Cuff Repair,” *J Bone Joint Surg Am*, vol. 94, no. 3, pp. 227–233, Feb. 2012, doi: 10.2106/JBJS.J.00739.

[3] M. P. Leathers, A. Merz, J. Wong, T. Scott, J. C. Wang, and S. L. Hame, “Trends and Demographics in Anterior Cruciate Ligament Reconstruction in the United States,” *J Knee Surg*, vol. 28, no. 5, pp. 390–394, Oct. 2015, doi: 10.1055/s-0035-1544193.

[4] J. P. Iannotti, “Full-thickness rotator cuff tears: factors affecting surgical outcome,” *Journal of the American Academy of Orthopaedic Surgeons*, vol. 2, no. 2, pp. 87–95, 1994.

[5] F. Aïm, G. Lonjon, D. Hannouche, and R. Nizard, “Effectiveness of Virtual Reality Training in Orthopaedic Surgery,” *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, vol. 32, no. 1, pp. 224–232, Jan. 2016, doi: 10.1016/j.arthro.2015.07.023.

[6] A. McCarthy, P. Harley, and R. Smallwood, “Virtual arthroscopy training: do the ‘virtual skills’ developed match the real skills required?,” *Stud Health Technol Inform*, vol. 62, pp. 221–227, 1999.

[7] “Fundamentals of Arthroscopy Surgery Training (FAST) Program,” 2017. http://www.aana.org/home/education/fundamentals-of-arthroscopy-surgery-training-(fast)-program (accessed May 19, 2017).

[8] J. M. Espadero, S. Bayona, J. M. Fernández, and M. García, “Advanced Athroscopy Training Simulator insightMIST,” *Avances en la Ciencia de la Computación*, p. 178, 2004.

[9] “ARTHRO Mentor | Simbionix.” https://simbionix.com/simulators/arthro-mentor/ (accessed Feb. 09, 2020).

[10] “VirtaMed | Medical education with virtual reality simulators.” https://www.virtamed.com/en/ (accessed Feb. 09, 2020).

[11] T. Halic, S. A. Venkata, G. Sankaranarayanan, Z. Lu, W. Ahn, and S. De, “A software framework for multimodal interactive simulations (SoFMIS),” *Stud Health Technol Inform*, vol. 163, pp. 213–217, 2011.

[12] M. Ashikmin, S. Premože, and P. Shirley, “A microfacet-based BRDF generator,” in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 65–74.

[13] D. Demirel, Y. U. Alexander, S. Cooper-Baer, T. Halic, S. Kockara, and S. Ahmadi, “Difficulty Scenario Modeling for Virtual Arthroscopic Rotator Cuff Surgery with GAML,” presented at the 5th International Conference on Computational and Mathematical Biomedical Engineering – CMBE2017, Pittsburgh, PA, Apr. 2017.

[14] D. Demirel, A. Yu, S. Baer, T. Halic, and C. Bayrak, “Generative Anatomy Modelling Language (GAML),” *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2017.

[15] “Bourns - Magnetic Encoders.” https://www.bourns.com/products/encoders/magnetic-encoders/product/EMS22A (accessed Feb. 09, 2020).

[16] A. Maciel, T. Halic, Z. Lu, L. P. Nedel, and S. De, “Using the PhysX engine for physics-based virtual surgery with force feedback,” *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 5, no. 3, 2009.

[17] C. M. Peruto, M. G. Ciccotti, and S. B. Cohen, “Shoulder arthroscopy positioning: lateral decubitus versus beach chair,” *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, vol. 25, no. 8, pp. 891–896, 2009.

[18] D. Demirel *et al.*, “A hierarchical task analysis of shoulder arthroscopy for a virtual arthroscopic tear diagnosis and evaluation platform (VATDEP),” *The International Journal of Medical Robotics and Computer Assisted Surgery*, 2016, Accessed: Jan. 28, 2017. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/rcs.1799/full.

[19] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.