B.A.Sc. Thesis

Division of Engineering Science
UNIVERSITY OF TORONTO

Reinforcement and Semi-Supervised Learning for Ultrasound Landmarking

Supervisor: *Dr. Pascal Tyrrell*

Student: *Diana Escoboza Salazar*

Student #: *1003650426*

Abstract

The detection of anatomical landmarks on ultrasounds is a significant step for medical assessment and diagnosis. Research in the area has shown Computer Vision models very successful in the automation of the task. However, acquiring and labelling enough data for training is well-know to be a very costly and impractical. An approach repeatedly studied on ultrasound landmarking is Reinforcement Learning (RL) that has shown successful results using a sensible amount of labelled data. In addition, Semi-supervised learning (SSL) is a methodology regularly applied to address the shortage of labelled data, and that is showing a decreasing performance gap between its supervised counterparts. There exists a gap on the comparison of both learning methodologies to address labelled data limitations in US landmarking. Here, we address this gap by comparing the performance of two state-of-the-art methods for the landmarking of joints' ultrasounds. Our results found RL to have a potential for more context-aware predictions and robustness against possible data outliers during training in comparison with the SSL method.

Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Premise

Ultrasound (US) landmarking is crucial to give medical staff a point of reference when performing medical assessments and diagnoses. [1] Automation of this task shortens the time for diagnosis and decreases human error. Research has found Computer Vision (CV) models to be very successful in US landmarking, real-time prostate gland localization in US images, liver tumour segmentation and more. [2], [3]. Nevertheless, obtaining enough labelled data to train these models is well-known for being time-consuming and costly, especially in medicine, where experts are required for the labelling process. Thus, further contribution in the area will bring advancements to solutions that address the lack of labelled data.

1.2 Background

The fast-growing use of CV techniques has brought new tools for solving real-life problems in various areas of healthcare. CV, a field in Artificial Intelligence (AI), consists of computers acquiring meaningful information from visual inputs and performing tasks based on that information. Medicine has benefited greatly from CV as image/video captures are widely used by staff daily. AI in medical research has previously incorporated the use of Deep Learning (DL) for predicting the probability of a patient showing up to their appointment, [4] as well as the classification of diseases using medical imaging, among others.

Radiology has become the field in medicine where DL is the most prominent. This came as a result of the potential it has demonstrated on the large amount of data in the area. [5] DL applied in radiology showed successful results in projects like liver tumour segmentation, prediction of Acute Ischemic Stroke and more. [3]

An area in radiology that has benefited from CV has been the landmarking in ultrasounds (US). There have been multiple studies addressing US landmark detection with strong performing frameworks. One of the first research studies used a combination of convolutional and recurrent networks to track the landmarks on US video frames. [6] Another study incorporated an adversarial component to train a more context-aware model for the detection of prostate glands on US images. [2]

Although state-of-the-art solutions have arrived at outstanding performance, most of the research has been fortunate to have access to a large amount of labelled data to perform their experiments. It is worth noting that the performance of DL networks has been directly linked to the amount and quality of the labelled data used for the training process. [7] Since obtaining large amounts of labelled data can be costly and impractical in specialized areas like medicine, there has been a big increase in the development of DL methods requiring less labelled data to achieve good performance.

One approach in US landmarking requiring a feasible amount of data is Reinforcement Learning (RL). The approach frames the problem as a sequential decision-making process, where an agent learns the optimal path toward the target landmark. A study on RL works on medical imaging showed the potential of these models while using a significantly smaller dataset, in comparison to standard solutions, for their experiments. [8], [9]

Moreover, semi-supervised learning is a methodology regularly applied to address the shortage of labelled data. In medical imaging, the approach has shown strong performance in classifying benign and malignant tumours and detecting and tracking landmarks in lung US videos, along with others. [10]–[12]

1.2.1 Deep Reinforcement Learning

RL is a sub-field of Machine Learning (ML) also lying under AI, and its theory is rooted in the psychology and neuroscience of behaviour where an agent directly interacts with its environment. The RL approach where the task is framed as a sequential decision-making process where an ‘agent’ interacts with the environment through observations, actions, and rewards. The agent learns to make decisions through trial and error with the goal of maximizing the future cumulative reward. [13] For US landmark detection, the 'agent' uses its current position in the image as the state to draw an action (direction to move) to achieve the final goal of arriving at the landmark. [8], [9]

One of the first papers proposing a RL solution for anatomical landmark detection was in 2016 by. [8] In [8] the inspiration came from the successful introduction of deep Q-network (DQN). This network combines RL with deep neural networks to eliminate the limitation RL used to have: only able to work in low-dimensional state spaces where the useful features were handcrafted. [13] The algorithm in [8] outperformed the then state-of-the-art solutions both in terms of accuracy and speed. [8] Continuing the innovation that [13] brought with DQNs, more studies came out applying such networks on CV tasks. In [14] they tackled the problem of Landmark Detection as image partitioning. Inspired by

DQNs and [8], they trained the network to construct a global action map to reach the target landmark from any location in the image. They based their approach on breaking out the agent optimal path of [8] into optimal action steps for every pixel, achieving a competitive performance against other state-of-the-art solutions on cardiac and obstetric US datasets. [14]

An evaluation paper on published RL works in medical imaging, experimented with different DQN-based agents for anatomical landmark detection. The results demonstrated high detection accuracy on three different datasets varying in size from 72 labelled US scans to 830 labelled MR scans. The results show the success of DQN models on the landmarking objective as well as the potential the method has for smaller labelled datasets. [9]

1.2.2 Semi-supervised Learning

Semi-supervised learning refers to using both labelled and unlabelled data for training. Most approaches in this area incorporate an unsupervised method into the training process to learn ingrained information in the data, creating a more robust model. [15]

A good example of today's state-of-the-art semi-supervised object detection can be seen in [7]. In [7] algorithm, a self-supervised module was added to learn proposal features from both labelled and unlabelled data, where the location loss helps to learn context-aware proposal features. Additionally, noise was added to the proposal features, and by ensuring consistency between the noisy and original predictions they created a more robust module. Their approach consistently improved the performance of fully-supervised baselines. [7]

A semi-supervised study in anatomical landmark detection used transfer learning by using a pre-trained object detection model on a different dataset to obtain feature maps. Then, the feature maps were fed into a scale discriminative correlation filter to obtain probability maps for the landmarking objective. The results showed the scale learning being able to alleviate influence from deformation when the images risk deformation. [16]

1.3 Gap

There exists a gap in research on the comparison of Reinforcement and Semi-supervised learning for limited labelled data in US landmarking.

1.4 Purpose

The purpose of this paper is to compare the performance between a Reinforcement and a Semi-Supervised Learning model on joints' ultrasounds for landmark detection. The distance between the predicted and the target landmark is to be considered as the evaluation metric for performance.

1.5 Hypothesis

It is hypothesized that the RL model will result in superior performance compared to the Semi-supervised model.

1.6 Objectives

1 The first objective of this project is to train the RL and Semi-supervised algorithms from [17], [18] on the US of the knee, ankle, and elbow joints. For each joint, the models will be trained to identify their corresponding landmarks.

2 The second objective is to evaluate the models on the test sets using the distance (L2 norm) between their predictions and the target landmarks as the evaluation criteria.

3 The third objective is to compare and analyze the models' performances on the test set.

Chapter 2

Methods

2.1 Datasets

The experiments relied on three datasets: ankle, elbow, and knee ultrasound images with labelled respective landmarks. Since the images were largely taken with a specific technique, the landmarks were chosen as the target structures on ultrasonography guidelines. [19]–[21]

Joint	Main Compartment on Ultrasound	Target Labelled Landmarks
Ankle	Anterior recess on the longitudinal axis	Tibia/Fibula Talus
Elbow	Lateral elbow on the longitudinal axis	Triceps Tendon Insertion Humerus
Knee	Medial knee on the longitudinal axis	Femur Quadriceps Tendon Patella

Table 1. Target landmarks for each joint US dataset and compartment in US.

2.1.1 Data Schematics

The following figures show the US image schematics for each joint with the annotated landmarks considered for this project.

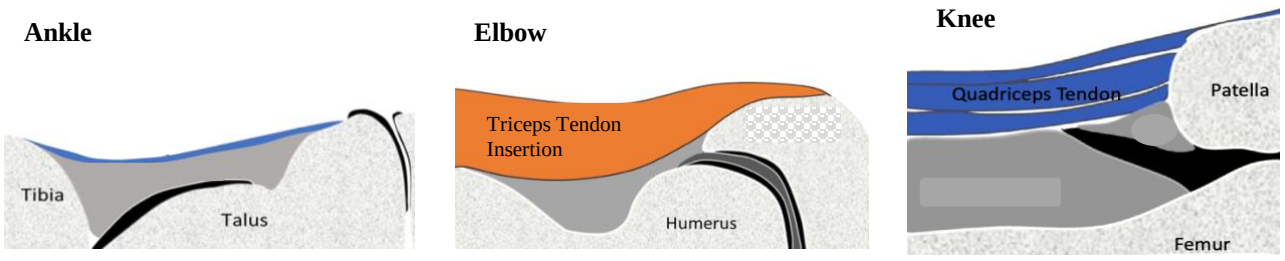

Figure 1. Schematics for ankle (left), elbow (centre), and knee (right) joints on US images used in project.

2.1.2 Data Labelling

The images were manually annotated using Label Studio application and then reviewed by experts in the area to ensure accurate labels. The annotations consisted of boxing each target structure and retrieving the centre coordinate.

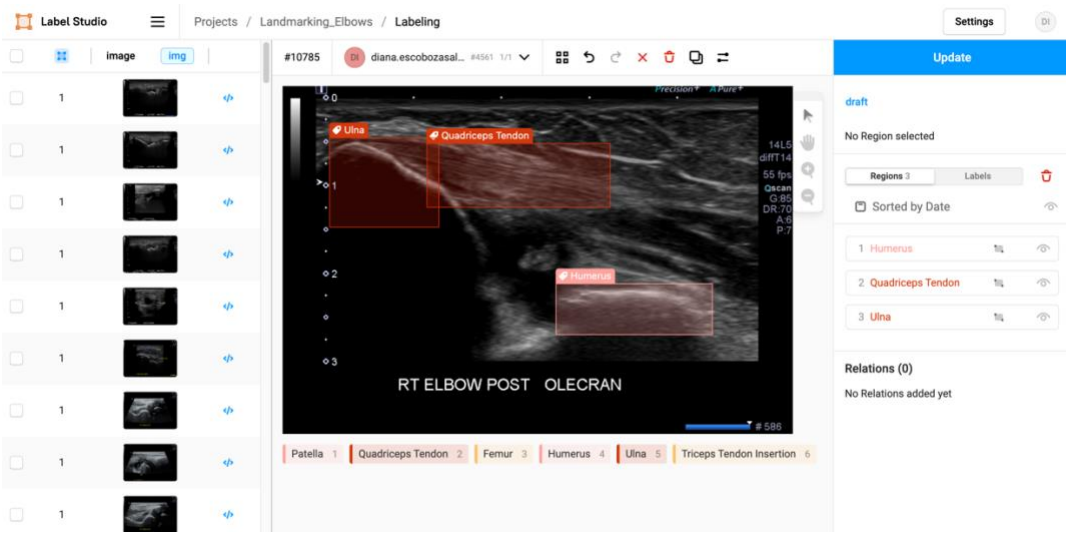

Figure 2. Label Studio platform used for labelling images in the dataset.

2.2 Data Pre-processing

Data pre-processing consisted of masking text on the images, data normalization and splitting each dataset into the training, validation, and testing sets.

2.2.1 Masking Labels

A significant number of the images in all three datasets contained text annotations of the target landmarks; as such, ***cleanup.pictures*** an inpainting AI application was used to remove them while maintaining the image structure. [22] These annotations were removed to prevent the models from learning correlations between text and the target, thus relying on the text annotations for the predictions. [23]

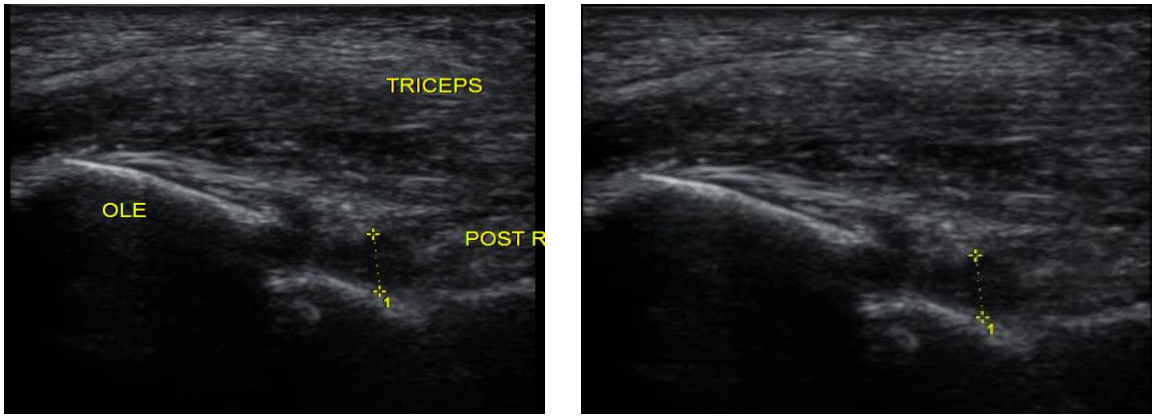

Figure 3. Display of images before (left) and after (right) text annotations were removed with cleanup.

2.2.2 Data Normalization

The images were normalized to improve and speed the model computations. As seen in [15], testing on medical imaging, image normalization had a great positive impact on prostate gland segmentation when the training and testing were both done with different site data.

2.2.3 Train-Validation-Test Data Split

Each dataset was split 70% for training, 20% for validation and 10% for testing.

# Labelled Images	Ankle		Elbow		Knee		
Total	425		320		323		
Train	302		231		232		
Validation	80		58		60		
Test	43		31		31		
# Landmark Instances	**Tibia/Fibula**	**Talus**	**Triceps Tendon Insertion**	**Humerus**	**Femur**	**Quadriceps Tendon**	**Patella**
Train	245	301	217	215	203	222	134
Validation	64	80	52	52	55	60	38
Test	37	41	30	29	27	33	19

Table 2. Number of images in each dataset before and after the train/val/test split, and number of instances of each landmark on the train/val/test split.

2.3 Architectures

Two models were used to gather results in this project: Multi-Agent Deep Reinforcement Learning from [17] and Unbiased Teacher for semi-supervised object detection from [18].

2.3.1 Multi-Agent Deep Reinforcement Learning

For this project, we lean on the project by G. Leroy [17] who built on and updated the work in [9]. The project involved finding landmarks in 3D medical images. Thus, multiple modifications were made to the open-source code of [17] to operate in our data composed of 2D images.

In the algorithm an agent navigates through an environment (the input image) and learns to make a sequence of decisions with the objective of arriving at the target landmark. The main elements composing the decision process are action space, states, and rewards. The action space is made up of four navigation actions: right, up, left, and down (or ± x and ± y directions). For the states, each state includes a history of $m = 4$ (changeable variable) past regions of interest within the image centered on its agent. After each step the oldest region of interest is replaced by the present one. Lastly, the reward at each action step is calculated as the difference between the past and current distance between the agent and target landmark.

Overview. The algorithm starts by initializing the Q-Network model with N number of parallel layers corresponding to the number of agents to be trained, where each agent is assigned a landmark. Every episode an image is passed on as input and the agents are initialized on a random location within the image. The agents then have a maximum number of steps to locate their target landmark. For each step of an episode their current and past 4 states' history are passed onto the Q-Network which outputs Q-values for each action (maximum expected reward) and for each agent. Each agent then choses their corresponding action to move to a new state and the reward is calculated according to the distance change to the landmark. Then the past state, action taken, and reward are saved into an experience buffer to be used for future steps. After every single step, all agents are checked for oscillations and to ensure their new state has not gone out of the image frame. Finally, the distance to the landmarks is used as the loss for the model's gradient backpropagation and optimization step.

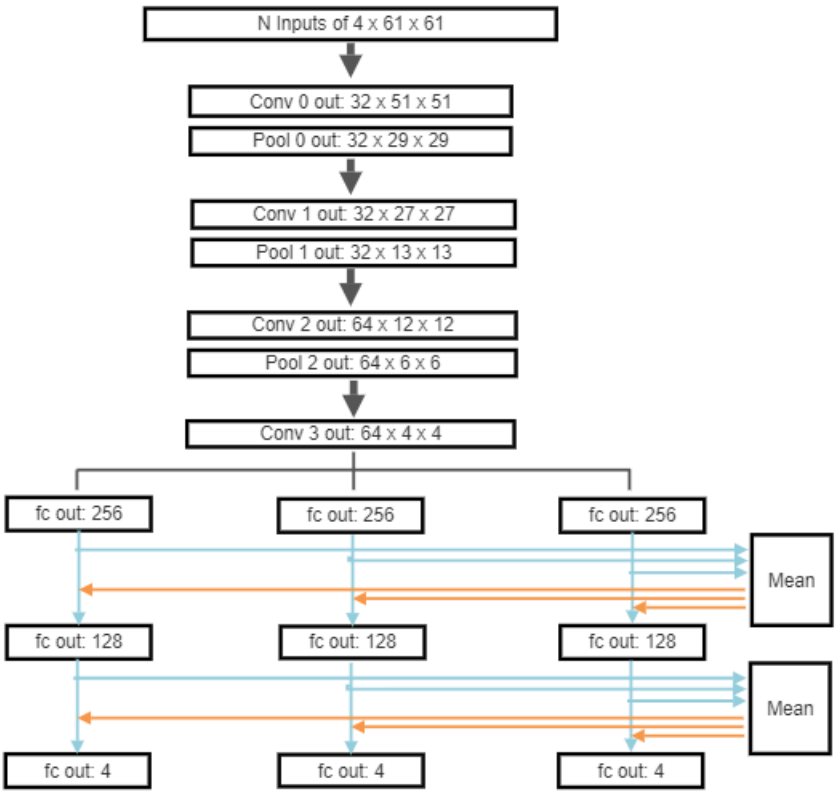

Figure 4. Q-Network architecture from [10] after being modified for experiments. The figure shows network structure for 3 agents. Number of parallel fully connected (FC) sections correspond to the number of landmarks being trained for (N, equal to number of agents). The architecture is first composed of a convolutional network shared by all agents, followed by FC layers. Each FC layer sends output to next layer and to a communication channel. In each step on the parallel FC sections, the communication channel calculates the mean of all sections (blue arrows) and concatenates them to the outputs (orange arrows) to pass onto the next FC layer.

2.3.2 Unbiased Teacher

Unbiased Teacher (UT) architecture was chosen as the semi-supervised network for this project's experiments. UT consists of two training stages; in the first stage (called Burn-In) an object detector is initialized by training on the labelled data. For the second stage (Teacher-student mutual learning) the initialized detector is duplicated as the 'student' and 'teacher' models. These models are then incorporated via a mutual learning mechanism utilizing all labelled and unlabelled data. [18]

Burn-In Stage. To initialize both 'student' and 'teacher' models, all supervised data is used for model optimization on object detection with a supervised loss.

Teacher-Student Mutual Learning Stage. The mutual learning stage aim is to leverage the unsupervised data, where the teacher generates pseudo-labels to optimize the student model, while gradually getting updated to improve the pseudo-labels.

After the models' initialization in stage one, the student training is initiated on all data by using the teacher's pseudo-labels for the unsupervised data. To prevent error accumulation of noisy pseudo-labels, they set a confidence threshold to filter-out possible false-positives. After a selected number of training iterations, the teacher model is gradually updated by applying Exponential Moving Average.

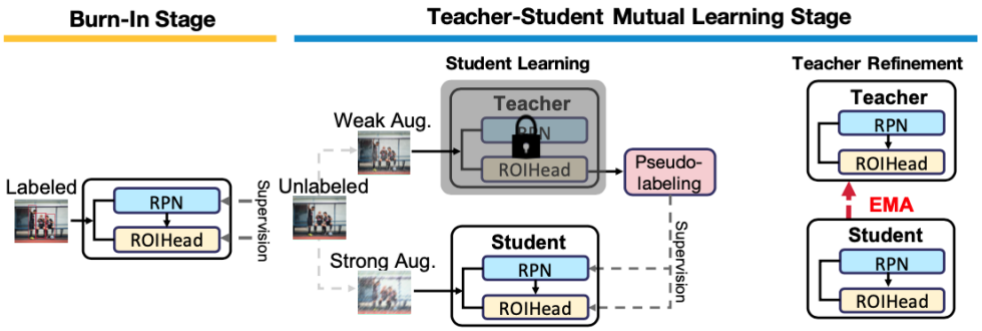

Figure 5. Unbiased Teacher architecture [18]

2.4 Experiments

2.4.1 Experimental Setups

The experiments were run on two Nvidia GeForce RTX 2080Ti and one GeForce TITAN X with CUDA version 11.7, and 12-core/24 thread Intel CPUs, GPU cards had 12GB RAM.

To arrive at more comparable results, each architecture was trained with a similar number of iterations/episodes (~30, 000). Due to a lack of resources, as the department machines were shared among other projects, both models had to be run with different GPUs and thus different computational power. The Unbiased Teacher (UT) model used the GeForce TITAN X while the Reinforcement Learning (RL) model run with both GeForce RTX GPUs concurrently.

2.4.1 Training and Evaluation

Both models were trained and evaluated using the same datasets split and were both trained a total of three times each, once per dataset (ankle, elbow, and knee). Open-source code links can be found in Appendix A.

For the UT algorithm each training run lasted about 1 day, ~5 hrs and while using around 9,000 MiB of GPU memory. On the other hand, the RL algorithm runs took 1 day, ~8 hrs (each) when training for datasets with two landmarks and almost 2 days for the dataset with three landmarks. In addition, the RL training process utilized less than 2,000 MiB GPU memory while training.

The difference in memory and time used during training is a direct result on their distinct model structures. RL algorithm takes a larger amount of time as every iteration (episode) is composed of multiple steps, and every step there is a forward pass through the model network for the agent to obtain the next action. Contrastingly, as the states passed into the model consists of only regions of interest and not the complete images, the data volume at once in the GPU for model operations is significantly less than for the UT model.

The trained models were used to draw inferences on the dataset's test sets. All test sets included between 30-40 images. The UT model return all inferences within 10 seconds per test set, whereas RL model took around 15 seconds. During the evaluation of the models, the distance (L2 norm) from the models' predictions and the target landmark was used as the evaluation criteria for performance comparison.

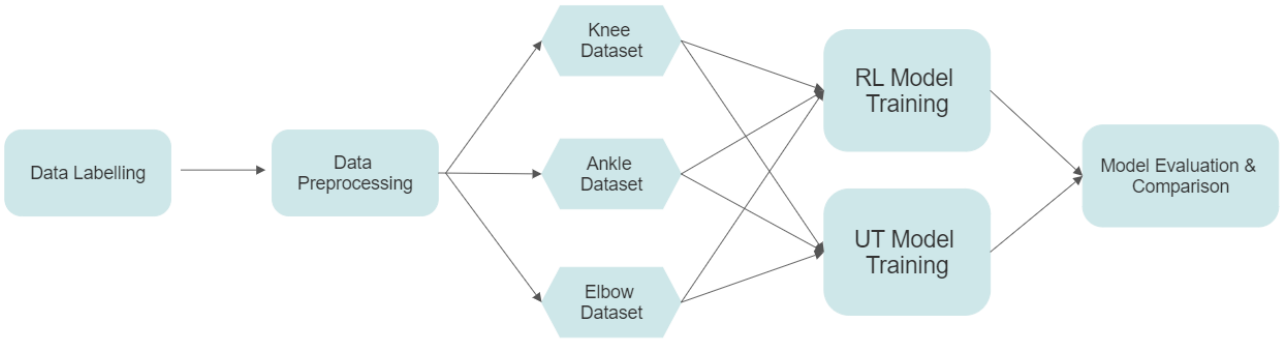

Figure 6. End-to-end pipeline of experiments. Steps taken from the data acquiring and labelling to the evaluation of the models after training on each joint US dataset.

Chapter 3

Results and Discussion

3.1 Test Set Results

Test set Distance to Target (mm)	Ankle		Elbow		Knee		
RL Avg	**41.41 ± 37.77**		**45.26 ± 36.85**		44.60 ± 36.17		
UT Avg	46.71 ± 40.63		47.62 ± 38.87		**39.80 ± 29.60**		
Distance to Target per Landmark (mm)	**Tibia/Fibula**	**Talus**	**Triceps Tendon Insertion**	**Humerus**	**Femur**	**Quadriceps Tendon**	**Patella**
RL Avg Distance	47.2 ± 44.96	**38.82 ± 31.98**	36.43 ± 31.28	**57.23 ± 42.14**	**41.6 ± 17.41**	60.25 ± 40.16	32.93 ± 14.24
UT Avg Distance	**33.86 ± 29.59**	56.59 ± 45.33	**32.62 ± 25.64**	66.57 ± 44.79	53.45 ± 41.22	**37.21 ± 21.45**	**25.35 ± 13.70**
RL Min Distance	8.48	**1**	**2.82**	3.6	**6.4**	6.4	**3.6**
UT Min Distance	**3.33**	4.01	3.07	**3.26**	7.46	**5.29**	3.82
RL Max Distance	175.41	**117.83**	128.16	173.14	**135.32**	184.48	199.54
UT Max Distance	**100.97**	166.96	**96.42**	**163.83**	136.61	**88.73**	**50.63**

Table 3. Distance to target results from test set for each algorithm.

	Ankle		Elbow		Knee		
	Tibia/Fibula	**Talus**	**Triceps Tendon Insertion**	**Humerus**	**Femur**	**Quadriceps Tendon**	**Patella**
T-Test	t(29) = 1.311 p = 0.1	t(39) = 2.042 p = 0.025	t(24) = 1.31 p = 0.1	t(19) = 1.33 p = 0.1	t(16) = 1.753 p = 0.05	t(23) = 2.819 p = 0.005	t(11) = 2.074 p = 0.025

Table 4. T-test on RL and UT distributions on distance results.

Table 3 shows the comparative results of the performance of both models on the test sets for each joint and respective landmarks. In general, the RL model reached an average distance smaller than UT's for two out of the three datasets. While the RL model achieves the minimal distance for all joints, it also arrives at the maximum.

When looking at the results on the particular landmarks, the average shortest distance is achieved by both models in almost an equal number each. Considering Table 4, for Tibia/Fibula, and both elbow landmarks, there was no significant difference between the two model's results distributions. The results

showed the biggest significant difference on the Quadriceps Tendon landmark, where UT model showed superior results than the RL model. Meanwhile, on the same joint, RL attained a notably better performance than that of UT on the Femur landmark.
Looking just at Tables 3 and 4 both models seem to take on a ‘preference’ to different landmarks for each joint, making their overall performance appear similar.

3.2 Displaying Worst/Best Models Performances

To get a better look at what the distance measurements represent on the US images, the next few sections will show the models worst/best performing predictions within the images.
On the images, the red dot corresponds to the target landmark, the blue for the RL model prediction and the green for the UT prediction.

3.2.1 Ankle US Worst/Best Results

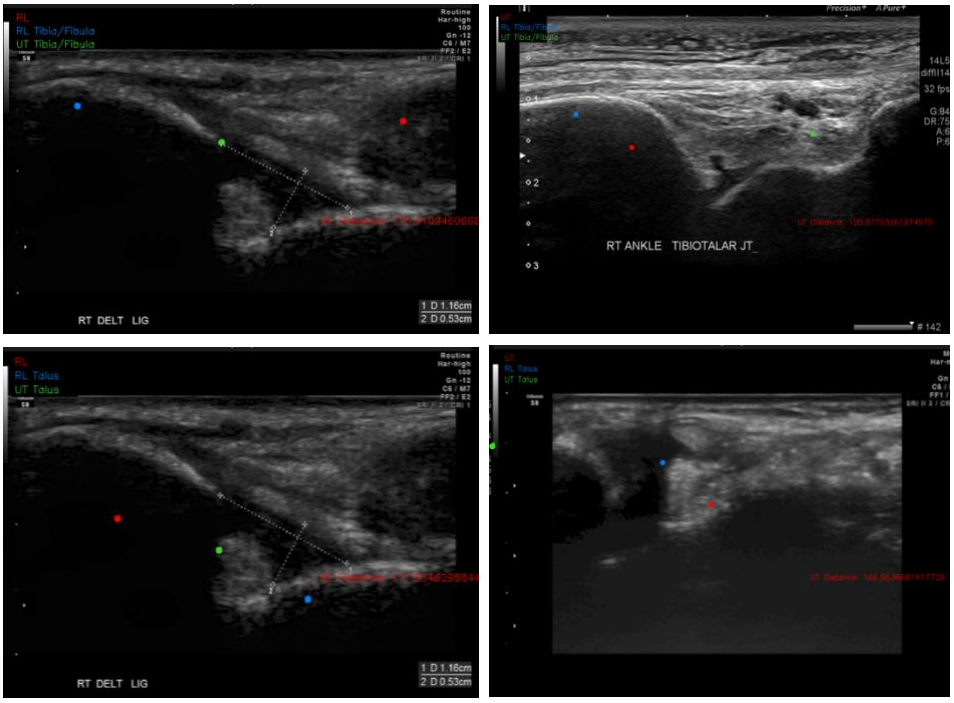

Figure 7. Ankle test images RL (left) and UT (right) model’s worst performance on Tibia/Fibula (top) and Talus (bottom) landmarks.

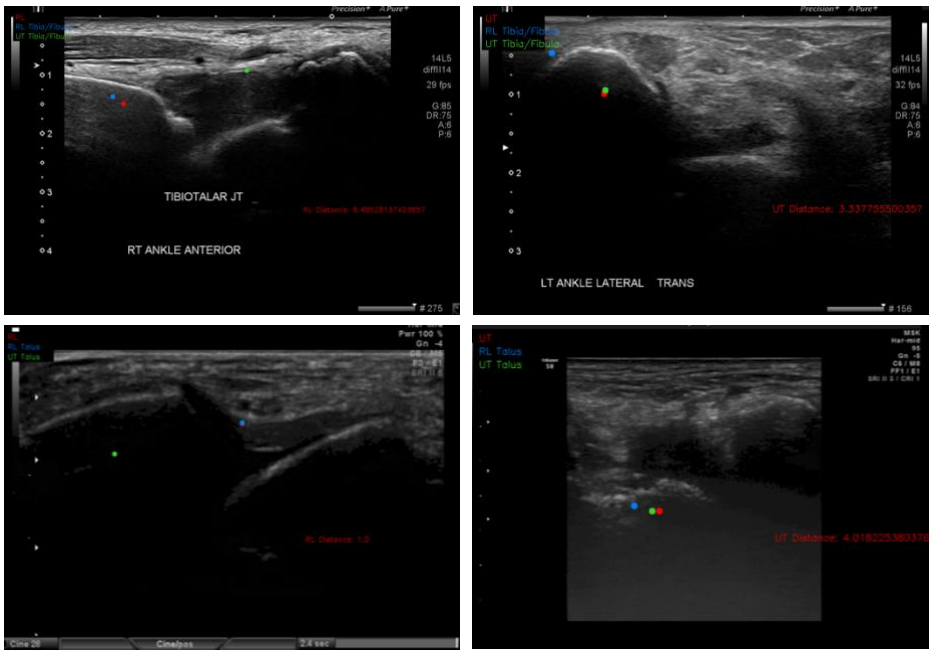

Figure 8. Ankle test images RL (left) and UT (right) model's best performance on Tibia/Fibula (top) and Talus (bottom) landmarks.

Looking at Figure 7, the RL model worst predictions for both ankle landmarks were produced in the same image. Meanwhile, for the same image, UT model's predictions of both landmarks stayed in the same x coordinates while just changing a short distance on the y-axis.

Further analysing the image and comparing it with the ankle diagram in Figure 1 (left), we can see the image is flipped horizontally from the US standard photo setup. Image transformations, like this one, were not included for this project. Thus, this flip came from the user taking the photography by using the probe inverted. Since this technique is not a standard practice, these cases composed an extremely small portion of the dataset. With this context in mind, the UT model's predictions show us that it could have been affected by outliers during training, more so than RL model. This could explain its static behaviour on the x-axis when predicting both landmarks in this image, despite that having these landmarks on the same x-coordinate to be unfeasible. On the other hand, the RL model predictions show to be more consistent to the joint schematics, which could indicate the model is likely less generalizable.

Another thing to note in Figure 7 is the worst performance of UT when predicting the Talus landmark. The prediction can be found in the left side of the image almost out of frame and located outside of the actual ultrasound photography.

3.2.2 Elbow US Worst/Best Results

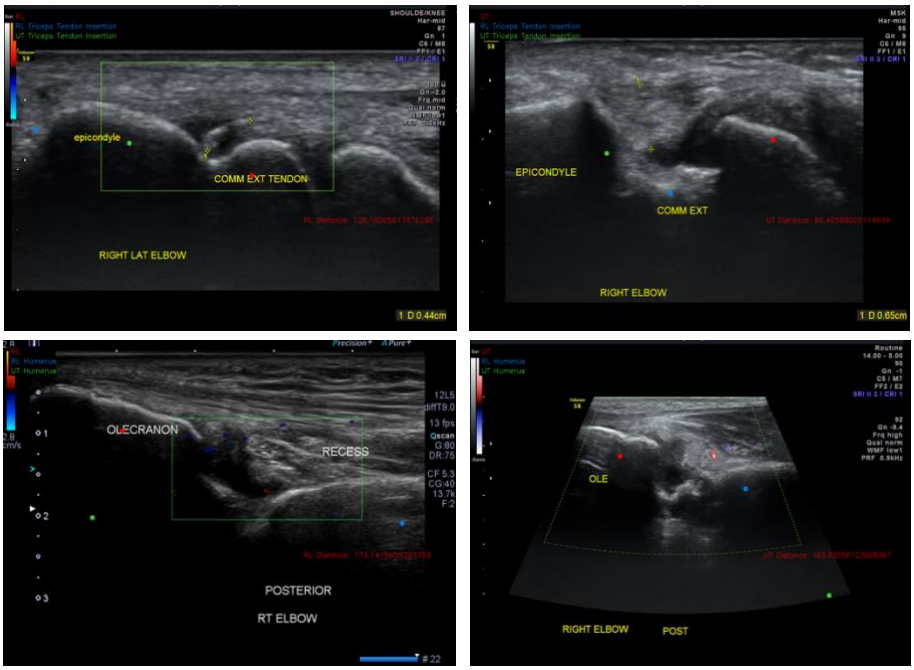

Figure 9. Ankle test images RL (left) and UT (right) model's worst performance on Triceps Tendon Insertion (top) and Humerus (bottom) landmarks.

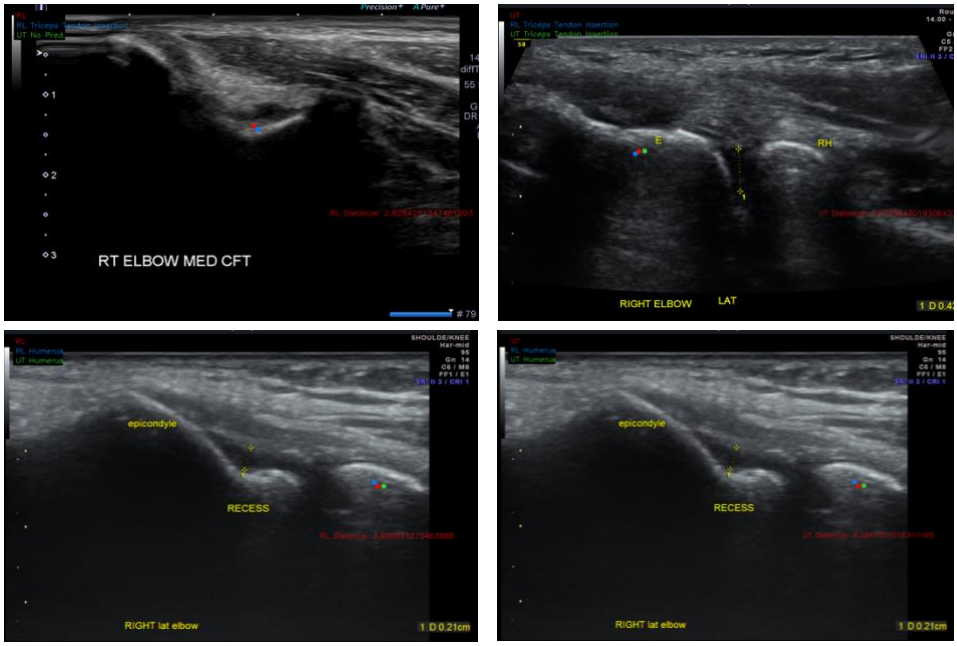

Figure 10. Ankle test images RL (left) and UT (right) model's best performance on Triceps Tendon Insertion (top) and Humerus (bottom) landmarks.

On Figure 9 bottom right, we see a similar situation as in Figure 7 with the UT prediction being located at the edge/outside the true ultrasound image.

In the results of Figure 10, both UT and RL model are seen to achieve their best performance on their prediction of humerus within the same image. Continuing in Figure 10 top left image, for the predictions of Triceps Tendon Insertion (TTI), there is no prediction from the UT model since the model's confident threshold was not met in this instance. On the other hand, RL model performed just as well as UT model in its TTI best performance.

3.2.3 Knee US Worst/Best Results

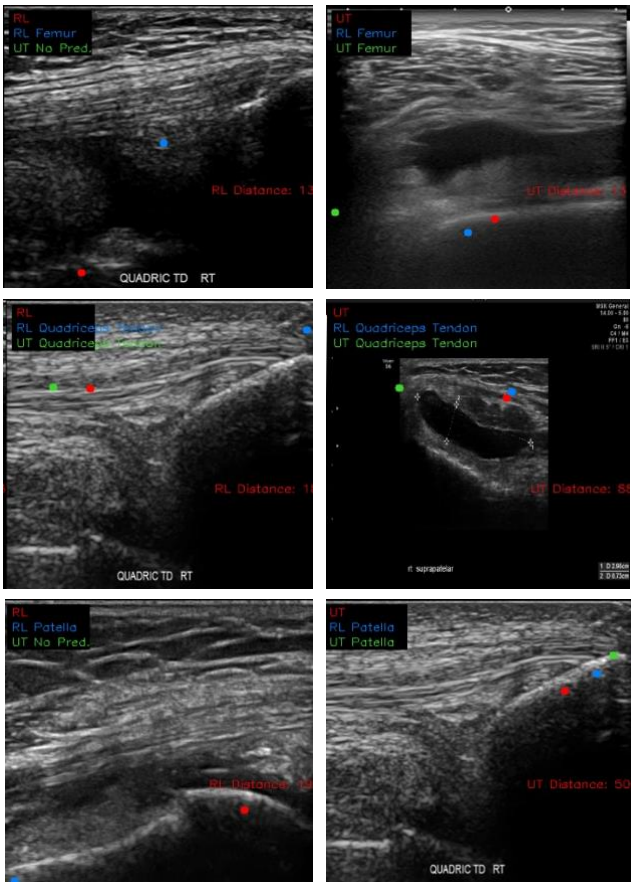

Figure 11. Knee test images RL (left) and UT (right) model's worst performance on Femur (top), Quadriceps Tendon (centre) and Patella (bottom) landmarks.

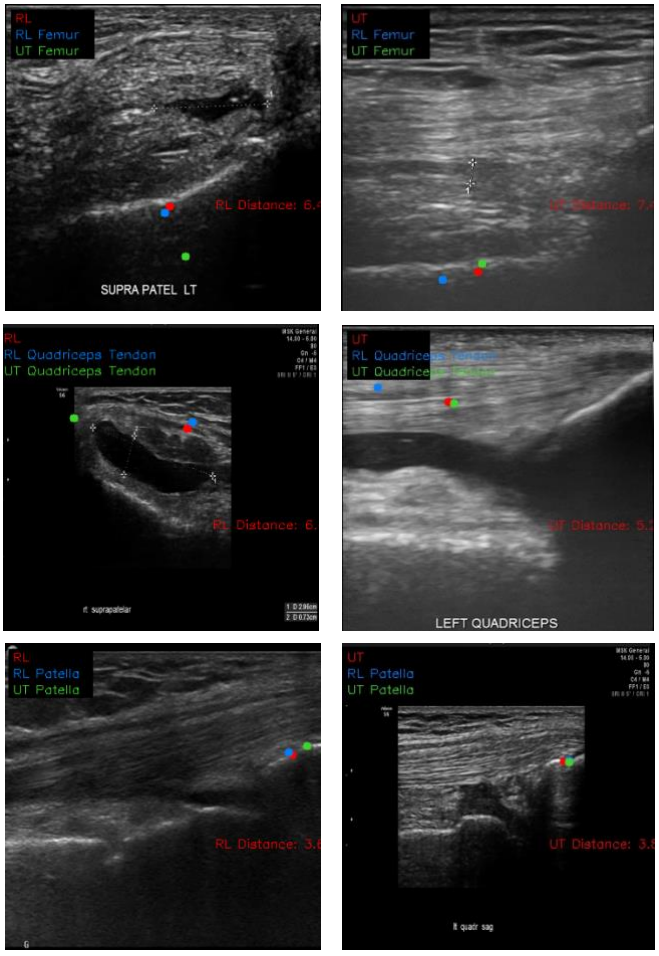

Figure 12. Knee test images RL (left) and UT (right) model's best performance on Femur (top), Quadriceps Tendon (centre) and Patella (bottom) landmarks.

Quickly noticing the worst-case on patella prediction for UT in Figure 11, it is clear the prediction lies inside of landmark and could seem relatively close to target. Although this could be considered a surprisingly good worst-case, it is worth noting that RL worst performance in patella landmark lied in an image where UT had no prediction.

Analysing both Figure 11 and Figure 12 on Quadriceps Tendon landmark, we see UT worst prediction lands at the edge/out the concrete US image. This is a similar case seen above in both elbow and ankle joints landmarks. In addition, this image ends up being the same where the RL model achieves its best prediction.

Looking at all instances in Figure 11 and Figure 12, every RL prediction either lie at a very close proximity of the target, lie within the landmark, or are in an image where UT model made no prediction.

This evidence gives more information when considering results shown in Table 3. The visual analysis on their prediction allows a deeper view needed to compare their behaviour and performance.

By the analysis on Table 3 and sections 3.2.1 through 3.2.3, it seems the RL model was able to learn more context-awareness as its' predictions did not suffer from exiting the US image bounds as much as the UT model did. Meanwhile, UT might be more likely to give significance to outliers in data during training, possibly by prioritizing generalization. For our main objective, RL model appears as the more promising approach to take in our environment consisting of limited labelled data.

Chapter 4
Conclusion

4.1 Conclusion

In this project, we compared the performance of two state-of-the-art methodologies (reinforcement and semi-supervised learning) on landmark detection in joints ultrasounds (US) with limited labelled data. These methods involve a multi-agent deep Q-learning and a mutual learning mechanism using pseudo-label generation (Unbiased-Teacher model). In particular, we explored how these models compare in performance when we are limited by the amount of labelled data. This is to address the costly process that comes with labelling data in the medical field, where a large amount of time from experts is needed. When trained on a moderate length training run and with limited data, we found both models to display similar successful results at this early stage. Additionally, we have found that across all joint US datasets depending on the landmark, the optimal model approach may change. Nevertheless, we show that reinforcement learning (RL) methodology shows a great potential for more context-aware predictions as well as, possible robustness against outliers in data during training compared to the semi-supervised UT approach.

4.2 Future Work

There are still further experiments to explore in this area. Longer training of both models is a certain next step for further performance enhancement and optimization. One may also want to data augmentation with the guidance of a professional in the area to prevent new data from disrupting helpful features learning. For the RL architecture, it is also possible to use the different datasets in conjunction with the aim of promoting additional learning for the shared convolutional network between agents. While some of these propositions may not increase performance significantly, incorporating new ideas may lead to future advancements in this or in other areas of AI.

Bibliography

[1] S. Hoeber, A.-R. Aly, N. Ashworth, and S. Rajasekaran, 'Ultrasound-guided hip joint injections are more accurate than landmark-guided injections: a systematic review and meta-analysis', *Br. J. Sports Med.*, vol. 50, no. 7, pp. 392–396, Apr. 2016.

[2] A. Tuysuzoglu, J. Tan, K. Eissa, A. P. Kiraly, M. Diallo, and A. Kamen, 'Deep adversarial context-aware landmark detection for ultrasound imaging', *arXiv [cs.CV]*, 27-May-2018.

[3] L. Saba *et al.*, 'The present and future of deep learning in radiology', *Eur. J. Radiol.*, vol. 114, pp. 14–24, May 2019.

[4] D. Liu *et al.*, 'Machine learning approaches to predicting no-shows in pediatric medical appointment', *NPJ Digit Med*, vol. 5, no. 1, p. 50, Apr. 2022.

[5] A. Esteva *et al.*, 'Deep learning-enabled medical computer vision', *NPJ Digit Med*, vol. 4, no. 1, p. 5, Jan. 2021.

[6] A. Rangamani and T. Xiong, 'Landmark Detection and Tracking in Ultrasound using a CNN-RNN Framework'.

[7] P. Tang, C. Ramaiah, Y. Wang, R. Xu, and C. Xiong, 'Proposal Learning for Semi-Supervised Object Detection', *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2021.

[8] F. C. Ghesu, B. Georgescu, T. Mansi, D. Neumann, J. Hornegger, and D. Comaniciu, 'An Artificial Agent for Anatomical Landmark Detection in Medical Images', *Lecture Notes in Computer Science*. pp. 229–237, 2016.

[9] A. Alansary *et al.*, 'Evaluating reinforcement learning agents for anatomical landmark detection', *Med. Image Anal.*, vol. 53, pp. 156–164, Apr. 2019.

[10] W.-C. Shia, L.-S. Lin, and D.-R. Chen, 'Classification of malignant tumours in breast ultrasound using unsupervised machine learning approaches', *Sci. Rep.*, vol. 11, no. 1, p. 1418, Jan. 2021.

[11] T. Pang, J. H. D. Wong, W. L. Ng, and C. S. Chan, 'Semi-supervised GAN-based Radiomics Model for Data Augmentation in Breast Ultrasound Mass Classification', *Comput. Methods Programs Biomed.*, vol. 203, p. 106018, May 2021.

[12] A. Tripathi *et al.*, 'Learning the Imaging Landmarks: Unsupervised Key point Detection in Lung Ultrasound Videos', *arXiv [eess.IV]*, 13-Jun-2021.

[13] V. Mnih *et al.*, 'Human-level control through deep reinforcement learning', *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[14] Z. Xu *et al.*, 'Supervised Action Classifier: Approaching Landmark Detection as Image Partitioning', in *Medical Image Computing and Computer Assisted Intervention − MICCAI 2017*, 2017, pp. 338–346.

[15] L. Schmarje, M. Santarossa, S.-M. Schröder, and R. Koch, 'A Survey on Semi-, Self- and Unsupervised Learning for Image Classification', *IEEE Access*, vol. 9, pp. 82146–82168, 2021.

[16] J. A. Onofrey *et al.*, 'GENERALIZABLE MULTI-SITE TRAINING AND TESTING OF DEEP NEURAL NETWORKS USING IMAGE NORMALIZATION', *Proc. IEEE Int. Symp. Biomed. Imaging*, vol. 2019, pp. 348–351, Apr. 2019.

[17] G. Leroy, 'Multi-agent Deep Reinforcement Learning for Anatomical Landmark Detection'. 06-Jul-2020.

[18] Y.-C. Liu *et al.*, *unbiased-teacher: PyTorch code for ICLR 2021 paper Unbiased Teacher for Semi-Supervised Object Detection*. Github.

[19] Ian Beggs, Stefano Bianchi, Angel Bueno, Michel Cohen,Michel Court-Payen,Andrew Grainger,Franz Kainberger,Andrea Klauser,Carlo Martinoli,Eugene McNally,Philip J.

O'Connor,Philippe Peetrons,Monique Reijnierse,Philipp Remplik,Enzo Silvestri, 'Musculoskeletal Ultrasound Technical Guidelines II. Elbow', European Society of MusculoSkeletal Radiology.
[20] Ian Beggs, Stefano Bianchi, Angel Bueno, Michel Cohen,Michel Court-Payen,Andrew Grainger,Franz Kainberger,Andrea Klauser,Carlo Martinoli,Eugene McNally,Philip J. O'Connor,Philippe Peetrons,Monique Reijnierse,Philipp Remplik,Enzo Silvestri, 'Musculoskeletal Ultrasound Technical Guidelines VI. Ankle', European Society of MusculoSkeletal Radiology.
[21] Ian Beggs, Stefano Bianchi, Angel Bueno, Michel Cohen,Michel Court-Payen,Andrew Grainger,Franz Kainberger,Andrea Klauser,Carlo Martinoli,Eugene McNally,Philip J. O'Connor,Philippe Peetrons,Monique Reijnierse,Philipp Remplik,Enzo Silvestri, 'Musculoskeletal Ultrasound Technical Guidelines V. Knee', European Society of MusculoSkeletal Radiology.
[22] 'Cleanup.Pictures'. [Online]. Available: https://cleanup.pictures/. [Accessed: 16-Aug-2022].
[23] K. Xiao, L. Engstrom, A. Ilyas, and A. Madry, 'Noise or signal: The role of image backgrounds in object recognition', *arXiv [cs.CV]*, 17-Jun-2020.

Appendix A

Source-Codes GitHub Links

A.1 Multi-Agent Reinforcement Learning

- GitHub: https://github.com/gml16/rl-medical

A.2 Unbiased Teacher

- GitHub: https://github.com/facebookresearch/unbiased-teacher