Final Report

A study into the feasibility of using machine learning techniques in diagnostic vetiranary imaging

Robert Clark

2020

# Abstract

within veterinary medicine. This report investigates two different aspects of this; diagnostic CT for elbow dysplasia and MRI for analysing abnormal neuro-morphology. Preliminary work was performed on elbow dysplasia, while full results were obtained for the study into neuro-morphology.

Canine elbow dysplasia is a disease afflicting many species of dogs, which has been hypothesised to be related to joint incongruity. Due to the malformed joint, an uneven distribution of forces results in contact between the different surfaces of the joint, leading to damage to bone and cartilage tissue. A possible way of characterising joint incongruity within Labradors suffering from elbow dysplasia is described. This is implemented via three-dimensional models from X-ray CT of the radial-ulnar joint and deforming healthy joints to quantify the difference between them and a joint of dogs presenting symptoms of elbow dysplasia.The preliminary stages of an AI-based tool designed to achieve this is described, with the bone structure of the joint being read from x-ray images and translated into three dimensional models which can then be mapped onto each other via rigid registration. However, due to issues of data availability and formal data sharing it was not possible to proceed with this aspect of the project. Therefore, the goal of the project was changed to AI-based analysis of the canine neuromorphological disease Chiari-like malformation. (CLM).

Chiari-like malformation is a condition afflicting canine and bears sufficient resemblance to the human condition of Chiari malformation for the two to be treated and diagnosed using similar techniques. The exact cause is disputed within the medical community, but it is largely attributed to the brain of the subject being too large for the skull. In this paper, attempts are made to identify the characteristics of chiari-like malformation responsible for the development of clinical symptoms and different machine learning approaches are used to diagnose the condition within MRI scans. The pre-processing of the images necessary for this is described, in addition to the retraining of a neural network via transfer learning and the development of a Support Vector Machine for classification.

A transfer learning approach was employed, resulting in a peak accuracy of 0.7368 but a specificity of only 0.2, which can likely be attributed to issues with the dataset used. Attempts to classify the data using an SVM based on the affine transformations used to map the central slices of the MRI scans onto each other showed no correlation between this feature and CLM.

Contents

[Abstract 1](#_Toc40461422)

[List of Figures 4](#_Toc40461423)

[List of Tables 5](#_Toc40461424)

[Chapter 1: Introduction 6](#_Toc40461425)

[1.1. Canine Elbow Dysplasia 6](#_Toc40461426)

[1.2. Canine Chiari-Like Malformation 7](#_Toc40461427)

[1.3. Current Approaches to Diagnosis and Treatment of Canine Elbow Dysplasia 8](#_Toc40461428)

[1.4. Current Approaches to Diagnosis and Treatment of Canine Chiari-Like Malformation 8](#_Toc40461429)

[1.5. X-Ray Computerised Tomography 10](#_Toc40461430)

[1.6. Magnetic Resonance Imagery (MRI) 10](#_Toc40461431)

[1.7. Aims and Objectives 11](#_Toc40461432)

[Chapter 2: Literature Review 12](#_Toc40461433)

[2.1. Prior Analysis Work on Canine Elbow Dysplasia 12](#_Toc40461434)

[2.2. Prior Analysis Work on Canine Chiari-Like Malformation 12](#_Toc40461435)

[2.3. Machine Learning as a Diagnosis Aid for CLM 13](#_Toc40461436)

[3.4. Machine Learning as a Diagnosis Aid for Other Conditions 14](#_Toc40461441)

[3.5. Summary 15](#_Toc40461442)

[Chapter 3: Preliminary work for investigation into Canine Elbow Dysplasia 16](#_Toc40461443)

[3.1. Data Description 16](#_Toc40461444)

[3.2. Data Processing 16](#_Toc40461445)

[3.2.1. ImageJ 16](#_Toc40461446)

[3.2.2. MATLAB 17](#_Toc40461447)

[3.3. Summary 21](#_Toc40461448)

[Chapter 4: Methodology of Investigation into CLM 22](#_Toc40461449)

[3.1. Data Description 22](#_Toc40461450)

[3.2. Data Processing 22](#_Toc40461451)

[3.2.1. Transfer Learning with a Convolutional Neural Network (CNN) 22](#_Toc40461452)

[3.2.2. Affine Transformations as Feature 24](#_Toc40461453)

[3.3. Summary 25](#_Toc40461454)

[Chapter 4: Results and Discussion 26](#_Toc40461455)

[4.1. Transfer Learning 26](#_Toc40461456)

[4.2. Affine Transformation as a Feature 32](#_Toc40461457)

[Chapter 5: Conclusions and Future Work 33](#_Toc40461458)

[Appendix A: Project Management 34](#_Toc40461459)

[Appendix B: Unabridged Results of Classification through Transfer Learning 35](#_Toc40461460)

[References 36](#_Toc40461461)

# List of Figures

[Figure 1 - A dog elbow displaying signs of elbow dysplasia. Arrow 1 indicates a step between the radius and ulna, arrow 2 an unfused upper anconeal process and 3 shows a fragmented medial coronoid [43] 6](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461462)

[Figure 2 - A cyst forming within a cervical spine, characteristic of Syringomyelia. [48] 7](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461463)

[Figure 3 - The flow of cerebrospinal fluid around the brain, with the perivascular space shown within the inset red box. [12] Licensed under Creative Commons Attribution 4.0 International. 7](#_Toc40461464)

[Figure 4 - A CT scan of a Labrador’s elbow, showing the humerus connecting to the radius and ulna at the joint. 8](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461465)

[Figure 5 - A central slice of an MRI scan depicting a healthy Cavalier King Charles Spaniel (left) and one affected by both CM and Syringomyelia (right), with the syrinx indicated by the red rectangle. 9](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461466)

[Figure 6 - The posterior fossa of the human skull, shown in green. Licensed under the Creative Commons Attribution-Share Alike 2.1 Japan. [49] 9](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461467)

[Figure 7 - A representation of the operation of a CT scanner, with the image subject at 1, the x-ray emitter at 2, the receiving sensor at 3, transmission beam at 4, the path of travel for the projector and sensor at 5, the origin at 6 and an image at 7. [45] 10](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461468)

[Figure 8 - A sinogram (left) and corresponding CT slice for a human jawbone. (right) [18]. 10](#_Toc40461469)

[Figure 9 - A cross section of an MRI scanner with labelled components. [50] Licensed under the Creative Commons Attribution-ShareAlike 3.0. 10](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461470)

[Figure 10 - Spheres fitted to congruent (left), positive RU incongruent (middle) and negative RU incongruent (right) joints. [24] 12](#_Toc40461471)

[Figure 11 - The Atlantooccipital joint and atlantoaxial ligament, located at the base of the skull. [51] 12](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461472)

[Figure 12 - A healthy CKCS (left) with a continuous craniocervical junction and a CLM/SM affected CKCS (right) with a noticeably elevated caudal medulla oblongata. Both junctions are highlighted with a red line. 13](#_Toc40461473)

[Figure 13 - A labelled cross section of a brain with ventricles exposed. [52] Licensed under the Creative Commons Attribution 4.0 International. 13](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461474)

[Figure 14 - A "healthy" hip joint (top) and one presenting symptoms of hip osteoarthritis with arrows at regions indicating this (bottom) [34]. 14](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461475)

[Figure 15 - A three-dimensional reconstruction of a CT of a Labrador joint. 16](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461476)

[Figure 16 - A 3D reconstruction of an elbow joint from a series of CT scans, from the same perspective as Figure 15. 17](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461477)

[Figure 17 - Histogram showing distribution of values across all slices of a CT scan of an elbow joint. 17](#_Toc40461478)

[Figure 18 - The original CT scans (left), the same CT scans thresholded to show only the cortical bone tissue (middle) and then the "cleaned up" binary masks (right). 18](#_Toc40461479)

[Figure 19 - Visualisation of three-dimensional point cloud A, describing Elbow A. 19](#_Toc40461480)

[Figure 20 - Visualisation of three-dimensional point cloud B, describing Elbow B. 19](#_Toc40461481)

[Figure 21 - The results of rigid transformations via ICP of point clouds A (green) onto A' (purple). 21](#_Toc40461482)

[Figure 22 - Summary of planned process to investigate the potential for artificial intelligence in understanding canine elbow dysplasia. 21](#_Toc40461483)

[Figure 23 - The structure of the VGG19 network. The final Classification layer is highlighted in cyan, and the final fully connected layer is highlighted in black. 22](#_Toc40461484)

[Figure 24 - Distribution of intensity levels over the dataset consisting of central slices from the MRI scans of nineteen CKCS. 23](#_Toc40461485)

[Figure 25 - The average edge map of the 5 control images generated using a Sobel filter, to be used as a stationary image during the ICP phase. 23](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461486)

[Figure 12 - A binary mask created by thresholding a central slice of an MRI scan of CKCS. 24](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461487)

[Figure 13 - The objective function model generated for the 9th iteration of the leave-one-out cross-validation phase, indicating the optimal fit for the three-dimensional hyperplane. 24](C:\\Users\\robbi\\Downloads\\Final Report Week 10_kw.docx" \l "_Toc40461488)

[Figure 28 - A block image describing the methodology of this project. 25](#_Toc40461489)

[Figure 29 - The training progress for the first through fifth iterations of the leave-one-out cross-validation process. 26](#_Toc40461490)

[Figure 30 - Results for iterations six through thirteen, with the same colour scheme as Figure 11 and an overall learning rate of 0.0003. 27](#_Toc40461491)

[Figure 31 - Results for iterations fourteen through nineteen, with the same colour scheme as and an overall learning rate of 0.0003. 28](#_Toc40461492)

[Figure 32 - Four Class Activation Maps showing a significant weighting granted to the craniocervical junction. The top three images are from the “affected” group and the bottom from the “control” group. All affected were correctly diagnosed, while the control was falsely positively diagnosed. 30](#_Toc40461493)

[Figure 33 - Six Class Activation Maps showing moderate weighting granted to the craniocervical junction. All six were affected by the disease and were correctly diagnosed. 30](#_Toc40461494)

[Figure 34 - A member of the control group which also indicated the craniocervical junction as an area of interest. This patient was correctly diagnosed. 31](#_Toc40461495)

[Figure 35 - Five Class Activation Maps for members of the Affected group with areas of interest apparently irrelevant to Chiari-Like Malformation. Only the top right patient was incorrectly negatively diagnosed. 31](#_Toc40461496)

[Figure 36 - Three Class Activation Maps for members of the Control group with no discernible pattern to areas of interest. All three were falsely diagnosed. 31](#_Toc40461497)

# List of Tables

[Table 1 - The expected values for radiodensity in Hounsfield Units for the different substances and tissues expected in a CT scan [19]. 16](#_Toc40461498)

[Table 2 - The transformation matrices found by ICP mapping with various levels of downsampling alongside the corresponding RMSE and the time taken for the operation. 20](#_Toc40461499)

[Table 3 - Summary of patient data within dataset 22](#_Toc40461500)

[Table 4 - Comparison of results from classification via transfer learning for an unaligned and aligned dataset. 29](#_Toc40461501)

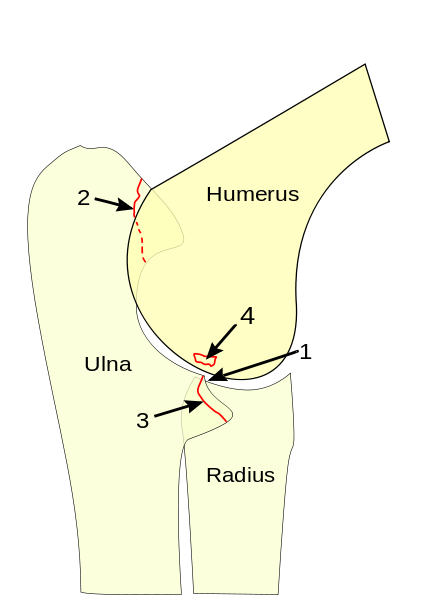
[Table 5 -Results of classifying dataset using SVM based off transforms required to map images onto average of controls, after Principal Component Analysis and Bayesian Optimisation. 32](#_Toc40461502)

[Table 6 - Gantt chart for the entirety of the project, updated to reflect the change in aim for the project. Deadlines are indicated with a "!". 34](#_Toc40461503)

[Table 7 - Classifications and scores for each of the 19 unaligned central slices classified by a retrained VGG19 CNN. 35](#_Toc40461504)

[Table 8 - Classifications and scores for each of the 19 aligned central slices classified by a retraining VGG19 CNN. 35](#_Toc40461505)

# Chapter 1: Introduction

This report attempts to investigate the potential for Artificial Intelligence in veterinary medicine. The original ambition of the project involved research into Canine Elbow Dysplasia, a condition affecting the bones and cartilage of dogs. To explore this avenue of research, it was anticipated that I would have access to a dataset of over 80 CT scans of Labrador but due to unforeseen circumstance this dataset was unavailable and so the project changed to research into Chiari-Like Malformation. This is a neuromorphological conditions primary affected Cavalier King Charles Spaniels, and was investigated using MRI slices obtained as part of a previous research project.

This change is reflected in the Gantt Chart (Table 6) of Appendix A: Project Management.

## Canine Elbow Dysplasia

In recent years, there has been an increasing awareness among the veterinary community of limb failure within dogs. This condition is specifically known as canine elbow dysplasia and is characterised by the bones of the elbow and surrounding cartilage showing signs of developmental abnormalities. [1] The diagnosis is typically used to refer to a condition in which the anconeal process, normally found at the tip of the ulna as shown in Figure 1, is unfused with the upper ulna bone [2]. Over time this can cause the erosion of the cartilage between the joint, which when left untreated can result in severe joint pain and lameness.

While elbow dysplasia is an umbrella term for many issues that can arise with the medial compartment of the elbow in dogs, the focus of this project will be on joint failure due to abnormal distribution of forces on the joint causing micro fractures, also known as Fragmented Coronoid Process (FCP). These abnormal forces are often attributed to either soft tissue forces such as bicep forces pulling on the ulna [3]. However, it has also been theorised that incongruity of the surface of the joint could also be responsible [4].

Figure - A dog elbow displaying signs of elbow dysplasia. Arrow 1 indicates a step between the radius and ulna, arrow 2 an unfused upper anconeal process and 3 shows a fragmented medial coronoid [43]

This incongruity can typically be characterised in three ways [5] [6]: Sagittal R-U Incongruity where the radial bone pistons away from the ulna when the joint is extended, H-U incongruity where the semi-lunar notch (the concave region of the Ulna shown in Figure 1) does not contour to the shape of the humerus and Transverse R- U incongruity where the radial head does not fit to the ulna. The three are not mutually exclusive, with it being possible for a joint to exhibit characteristics of multiple types of incongruity.

If elbow dysplasia such as that potentially caused by joint incongruity is left untreated, it can often deteriorate into Medial Compartmental Disease (MComD) or osteoarthritis [7] and as such it is imperative to detect the disease early.

## Canine Chiari-Like Malformation

Syringomyelia (SM) is a term used by the medical community to refer to the formation of cavities or cysts known as “syrinxes” within the spinal cord, resulting in discomfort, paralysis and loss of sensation throughout the body [8]. The most common cause of Syringomyelia is Chiari-Like Malformation (CLM), which is typically characterised as an incongruity between the size of the brain and the size of the skull [9] and is believed to be present in up to 95% of the world wide population of Cavalier King Charles Spaniels (CKCS) [10].

The discrepancies between the shape of the skull and brain in patients suffering from Canine CLM cause an obstruction for cerebrospinal fluid and tissue compression within craniocervical junction where the skull meets the spinal cord [11]. It is believed the resulting irregular flow of fluid then results in a mismatch of timing between arterial blood flow and cerebrospinal fluid flow. The perivascular space (shown in Figure 3) widens during the lull of the cardiac cycle, resulting in the cerebrospinal fluid essentially “leaking” through while its own pressure is high. This could result in cerebrospinal fluid entering the central canal of the spinal cord, and the eventual formation of syrinxes as shown in Figure 2.

Figure - A cyst forming within a cervical spine, characteristic of Syringomyelia. [48]

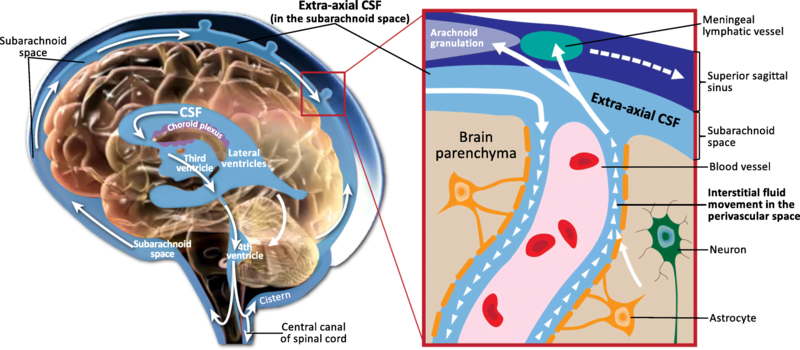


Figure - The flow of cerebrospinal fluid around the brain, with the perivascular space shown within the inset red box. [12] Licensed under Creative Commons Attribution 4.0 International.

In addition to Cavalier King Charles Spaniels, both Syringomyelia and Chiari-Like Malformation are known to present in humans. [13] Canine Chiari-Like Malformation is sufficiently analogous to its human counterpart of types 0 and 1 Chiari Malformation for diagnostic and surgical techniques used to treat one to be effective on the other, resulting in extensive veterinary research being performed upon the Cavalier King Charles Spaniel breed as well as similarly affected dogs. This is an example of “One Health” science [14], an interdisciplinary initiative to combine the results of research in various fields to benefit other research areas.

## Current Approaches to Diagnosis and Treatment of Canine Elbow Dysplasia

As elbow dysplasia is inherently an issue with the hard tissue and surrounding cartilage of the joint, diagnosis often requires x-ray computerised tomography (CT) (see Figure 4) and radiography to produce an image of the joint, which can be checked for fragmentation and incongruence [6].

Figure - A CT scan of a Labrador’s elbow, showing the humerus connecting to the radius and ulna at the joint.

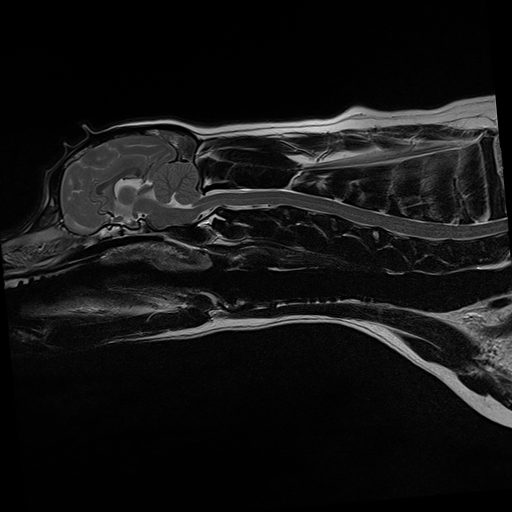
After diagnosis, there are various possible treatment options depending on the nuances of the situation. In some cases, the bone itself can be reshaped via surgery but this carries a high rate of morbidity so is often avoided. Bone fragments can also be surgically removed to attempt to prevent further cartilage damage, but this does not entirely prevent future wear and so can often require continuous treatment. In severe cases with large amounts of cartilage erosion, it may be necessary for partial or total joint replacement [6].

The success rates of these operations are greatly increased if the disease is detected early and at a young age, and in some cases it is even possible to perform non-invasive treatments such as hormone treatments to prevent growth spurts and careful diet management to prevent increased strain on the joints [15]. Artificial intelligence can potentially benefit this by reducing images and models to quantifiable data and then parsing through it to observe for trends which a human eye may not detect. This could allow for earlier detection of the condition if these trends are identified in patients not yet presenting with clinical symptoms, or by observing the trends their causes can be better understood. Better treatments could then be devised which address the failure of the joint at its cause rather than merely treating symptoms or slowing the rate of deterioration.

## Current Approaches to Diagnosis and Treatment of Canine Chiari-Like Malformation

A further application which has significant potential for AI to assist in the veterinary diagnostic clinical pathway is in the area of Chiari-like Malformation. Prior to the advent of Magnetic Resonance Imaging (MRI), only the behavioural signs of Chiari-like Malformation could be used in diagnosis. These would include disruption to motor skills, altered emotional state due to pain and excessive head rubbing and scratching and as such the condition was often misdiagnosed as epilepsy or an allergic reaction. [13]

After Chiari-like Malformation was observed in humans, the increased understanding of the condition allowed for the condition to be more easily diagnosed within animals. [13] Syringomyelia can be diagnosed via imaging of the spinal cord, where syrinxes will appear as anomalous regions along the central canal as shown in Figure 5.

A pair of black shoes

Description automatically generated

Figure - A central slice of an MRI scan depicting a healthy Cavalier King Charles Spaniel (left) and one affected by both CM and Syringomyelia (right), with the syrinx indicated by the red rectangle.

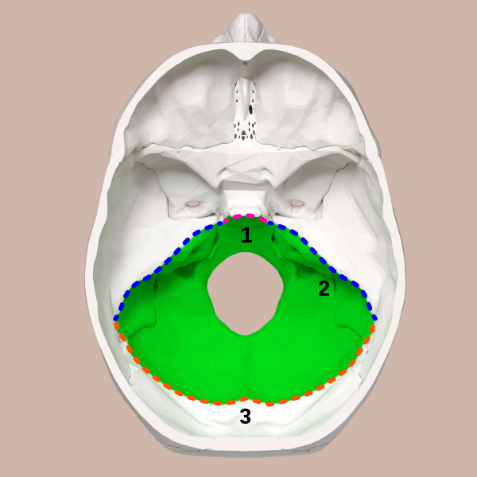
However, Chiari-like Malformation occurs independently of Syringomyelia and so syrinxes may not form until a considerable amount of time after the former condition develops. Diagnosis of Chiari-like Malformation alone is hence often done via the ruling out of other conditions and study of anatomical characteristics of the patient.

Figure - The posterior fossa of the human skull, shown in green. Licensed under the Creative Commons Attribution-Share Alike 2.1 Japan. [49]

Once Syringomyelia secondary to Chiari-like Malformation has been diagnosed, the treatment usually depends on the age of the patient. Younger canines typically have a higher recovery rate after surgery, and so removal of the syrinx is often recommended to prevent further development as the dog ages. Older dogs with fewer clinical symptoms are instead treated medically with opioids or antiepileptics to limit neuropathic pain. [16]

To limit production of further syrinxes and treat Chiari-like malformation directly, the most common treatment is Foramen Magnum Decompression (FMD). This involves enlarging the posterior fossa (Figure 6), the region of the skull which houses the brain stem and cerebellum, to alleviate pressure on the craniocervical junction. [17] This approach has an approximately 80% success rate, but has a 25% to 50% chance of relapse due to scar tissue formation at the decompression site.

Artificial intelligence could again be used to reduce this chance of surgical complications and relapse by allowing for biomarkers associated with the condition to be detected early while the patient is in optimal condition for surgery. By observing the trends associated with the condition, it may also be possible to increase understanding of the characteristics of the disease such as the level of compression required upon the brain and the specific areas which directly lead to the presentation of clinical symptoms. Less intrusive surgical measures could then potentially be devised, or potential non-surgical methods of intervention introduced.

## X-Ray Computerised Tomography

The dataset used for this project is composed of x-ray CT scans of canine elbows such as Figure 4. To obtain these CT scans, x- rays are passed through the subject by an x-ray source which rotates around the subject, illustrated in Figure 7. Both the x-ray emitter and detector are physically moved in a spiral fashion, resulting in a series of absorption patterns from multiple angles.

The initial output of a CT scan is a sinogram, shown in Figure 8. This is applied to a tomographic reconstruction algorithm to produce a series of cross sections [18]. Hard tissues, such as bone, are denser than the surrounding soft tissue and so absorb greater amounts of the x-ray energy. The level of attenuation is assigned a pixel value, and hence hard tissues appear in each layer of the slice as brighter.

Figure - A representation of the operation of a CT scanner, with the image subject at 1, the x-ray emitter at 2, the receiving sensor at 3, transmission beam at 4, the path of travel for the projector and sensor at 5, the origin at 6 and an image at 7. [45]

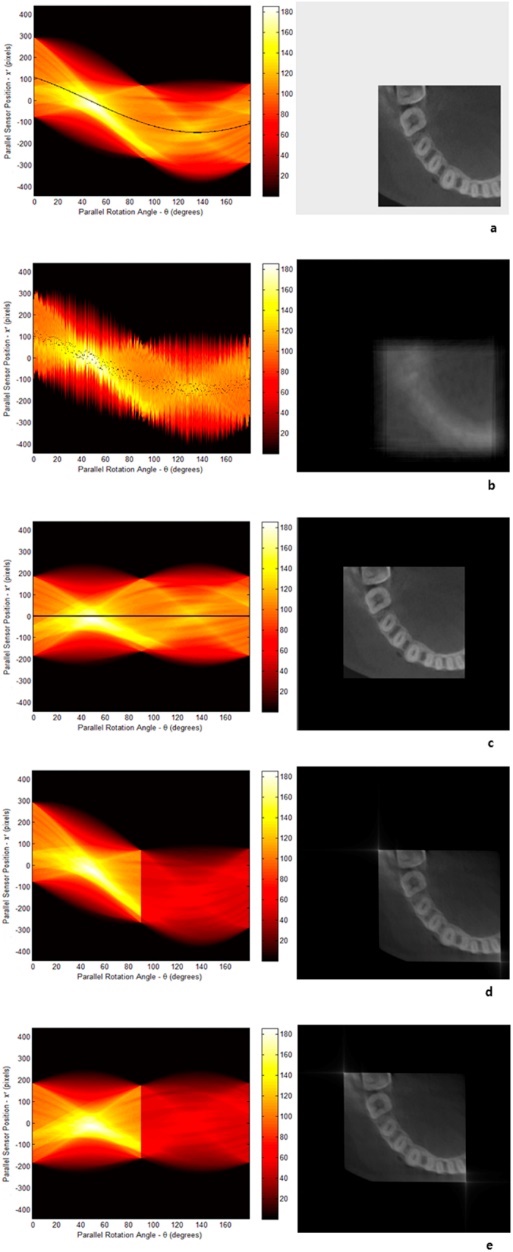


Figure - A sinogram (left) and corresponding CT slice for a human jawbone. (right) [18].

These pixel values are based off the Hounsfield scale, where a region of space populated only by air has a Hounsfield Unit (HU) value of -1000 and a region containing high density metals such as gold has a HU value of 30,000. [19] These regions correspond to a grayscale pixel value of 0 and 255 respectively. A change of one HU is equivalent to a change of 0.1% of the attenuation coefficient of water, which has a HU value of 0.

## Magnetic Resonance Imagery (MRI)

The dataset used for this project is composed of MRI scans of Cavalier King Charles Spaniel skulls, as shown in Figure 5. To obtain these images, the subjects were anesthetized to reduce movement and placed within a strong oscillating magnetic field. This aligns the positively charged protons within the water molecules of the subject. Targeted bursts of RF energy are used to “knock” the water molecules out of alignment. As they realign, they emit RF signals of their own which are detected by the receiving coil with different tissues aligning at different speeds producing varying signals. The three gradient coils within the MRI machine allow the oscillating magnetic field to be moved within the uniform field generated by the primary coil, allowing for data to be obtained for a number of different perspectives which can be combined to obtain an overall image.

The “frequency content” of the signal emitted as the hydrogen nuclei realign can then be analysed through use of Fourier Transforms, and the different tissues associated with the frequencies identified alongside their spatial location. In order to image the brain for this project, T2-weighted imaging was used in order to produce a higher contrast between the soft tissues of the brain and the hard tissue of the bone which will appear as light and dark regions respectively. This involved using comparatively long intervals between repeated radio bursts in order to allow magnetization to decay. [20]

## Aims and Objectives

As explained in Section 1.4, early diagnosis of Chiari-like Malformation can allow for surgical intervention while the Cavalier King Charles Spaniel is still young enough to have a high rate of recovery and before Syringomyelia can develop and require addition surgery to remove syrinxes. Currently, this is difficult due to the ambiguity as to what directly qualifies as Chiari-like Malformation and the process of diagnosis often requiring specialist knowledge which may not be available to all patients.

If accurate diagnosis of the condition through machine learning is possible, access to an MRI machine would be the only obstacle to diagnosis of the condition. This would allow for the condition to be identified earlier, before it can develop into Syringomyelia or intrusive surgery becomes too damaging for the patient and so result in a significant quality of life improvement for dogs presenting with the disease. Evidence also suggests that Chiari-like malformation is a hereditary condition [21], meaning that dogs which suffer from the condition but which do not present many clinical conditions could be identified before they produce less fortunate descendants and hence provide better breeding values for healthy dogs.

The similarity of canine Chiari-like Malformation to its human equivalent could also potentially mean that any machine learning based approach to diagnosis could be used on humans as well, hence reducing the time needed to diagnose the condition and expedite pain alleviating surgery. Therefore, the aims of this project are to investigate the potential for artificial intelligence within the study of Chiari-like Malformation and to attempt to increase the understanding of potential causes of the disease through characterisation.

# Chapter 2: Literature Review

## 2.1. Prior Analysis Work on Canine Elbow Dysplasia

As radiography and CT are necessary steps in diagnosing elbow dysplasia in dogs, it is common to attempt construction of three-dimensional models from the layers produced by CT scans to use as diagnostic tools. These have lent credence to the idea of the disease being characterised by elbow incongruence. Previous research into elbow dysplasia have supported this hypothesis, with studies into quantifying joint incongruity by measuring the distance between the humerus-ulnar joint and the radio-ulnar joint indicating a clear correlation between the level of incongruity at the coronoid apex and the probability of FCP [22] [23].

By fitting a sphere to three dimensional renderings of the ulnar trochlear notch to determine sensitivity of radioulnar incongruence, two specific types of radio-ulnar (RU) incongruity could be identified. As shown by Figure 6, a “congruent” joint accommodated the sphere perfectly, while “positive” RU incongruence presented a wider curve where the surface of the joint was not parallel to the surface of the sphere. “Negative” RU incongruent joints instead had to narrow a curve, resulting in an overlap between the volume of the sphere and the volume of the joint. By searching for these forms of incongruity, accuracy of diagnosis for elbow dysplasia could be seen to improve to the point where the technique can be “considered safe and highly accurate for clinical application” [24]. This project focused upon H-U incongruity (see section 1.1), whereas we will instead be investigating Transverse R-U incongruity, but the principle of comparing an incongruent joint shape to a “normal” shape (characterised in their study by a simple sphere) remains the same.

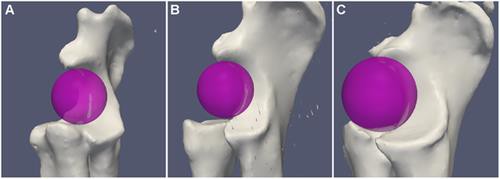


Figure - Spheres fitted to congruent (left), positive RU incongruent (middle) and negative RU incongruent (right) joints. [24]

## Prior Analysis Work on Canine Chiari-Like Malformation

A picture containing text, map

Description automatically generatedThough the definition of chiari-like malformation as a product of an underdeveloped skull compressing an overdeveloped brain is well understood, the exact dimensions needed to produce clinical symptoms are still debated. Cerebellar compression as a result of overlapping of the atlas and occipital bones (Figure 11) so that the atlas bone is partly within the cranium has a stronger incidence rate of CLM than other causes [25], meaning that the nature of the occipital bone may be one of the deciding factors of the condition. However, atlantooccipital overlapping also resulted in more severe incidences of cerebellar compression and so the correlation may instead be to that rather than any one particular cause.

Figure - The Atlantooccipital joint and atlantoaxial ligament, located at the base of the skull. [51]

Alternative studies have instead linked CLM developing into Syringomyelia to lesions on the atlantoaxial ligament known as “bands” [26] or “medullary kinking” of the craniocervical junction [27] where the spinal cord meets the brain stem at a non-continuous angle as illustrated in Figure 12.

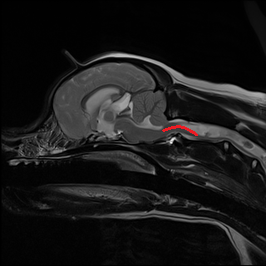
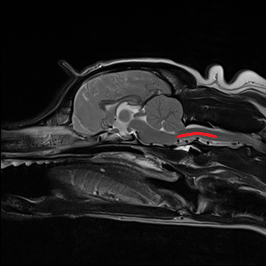


Figure - A healthy CKCS (left) with a continuous craniocervical junction and a CLM/SM affected CKCS (right) with a noticeably elevated caudal medulla oblongata. Both junctions are highlighted with a red line.

These investigations are hampered by the sheer prevalence of Chiari-like Malformation, with control groups being hard to establish when the condition simply could not have presented itself yet.

## Machine Learning as a Diagnosis Aid for CLM

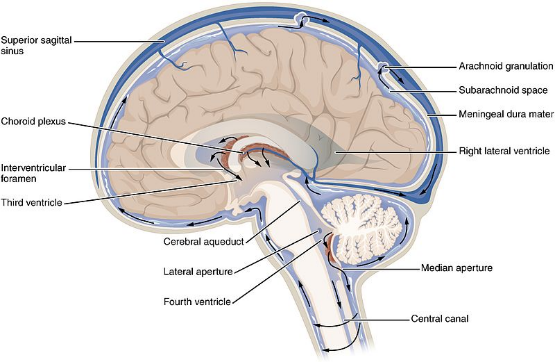
Some investigation into machine learnings use diagnosing Chiari-like Malformation has already been performed. The floor of the third ventricle (shown in Figure 13) and a region in the sphenoid bone as potential biomarkers for CLM by locating morphological differences between MRI scans of affected and healthy dogs. [28] By quantifying these morphological differences and using them as features within a Support Vector Machine (SVM), researchers were able to produce a binary classification system with an Area Under the Curve (AUC) of 77.77. This is sufficient to lend credence to these biomarkers being related to CLM and eventual pain but lacks the reliability necessary for consistent use within medicine.

Figure - A labelled cross section of a brain with ventricles exposed. [52] Licensed under the Creative Commons Attribution 4.0 International.

Additional supervised learning approaches have also been tried on MRI scans of human Chiari Malformation patients, producing mixed results. Distance from the Foranem Magnum (FM) to the peak of the fourth ventricle, distance from the FM to the brain stem and the angle of the brain stem were identified as key characteristics through use of SVMs and produced sensitivity and specificity rates of above 90% when testing for type 1 Chiari Malformation [29]. This condition is analogous to CLM in canines, but structural differences between human and canine brains such as the enlarged cerebellum may decrease the importance of these specific measurements in the diagnosis of CLM.



## Machine Learning as a Diagnosis Aid for Other Conditions

In addition to research into CLM, machine learning has also proved to be invaluable in the study of other conditions. By obtaining MRI data from Alzheimer’s patients and providing it to a Convolutional Neural Network (CNN), researchers were able to create a binary classifier capable of diagnosing the condition with an average accuracy 96.8588% [30]. The dataset used here consisted of a “stack” of images obtained from each patient, with non-brain tissue removed from each layer and the eight outermost layers discarded. The high accuracy rate obtained here was accredited by researchers to the use of the LeNet CNN.

SVMs have also been used in the study of Parkinson’s Disease, with a dataset of MRI scans of Parkinson’s afflicted brains reduced via Principle Component Analysis (PCA) to identify regions within the 3D space which expressed significant variation and then treated these as features. This resulted in an overall accuracy of 92.7% when classifying Parkinson’s afflicted and healthy brains [31], a significant improvement over the 39.53% obtained by a similar study performed without dimensionality reduction via PCA [32]. This could potentially be attributed to differences in the dataset, with the former study having an approximately 27% larger dataset which may have allowed for more generalisations but should still be considered a testament to the potential use of PCA within this investigation.

Machine learning has also shown potential within orthopaedic medicine. Hip dysplasia is a condition in which the socket portion of the hip joint imperfectly maps to the ball section, bearing some similarity to the H-U incongruity mentioned in Section 1.1**.** If left untreated, this condition can develop into hip osteoarthritis and result in significant lameness and discomfort [33]. By supplying a CNN with 420 CTs of human hip joints, it was shown to be possible to diagnose hip osteoarthritis with an accuracy of 92.8% when compared to chief physicians. These results were obtained by constructing a dataset composed of both “healthy” hips and hips presenting symptoms of osteoarthritis and using these as training set for the CNN by selecting the regions of the hip joint indicative of osteoarthritis (see Figure 7) and comparing the shape of the ball and socket at these points [34].

Figure - A "healthy" hip joint (top) and one presenting symptoms of hip osteoarthritis with arrows at regions indicating this (bottom) [34].

Machine learning has also shown potential for diagnosing fractures within human vertebra, again using CT scans of the affected regions. By extracting features from slices in a two-dimensional CT scan and then producing a set of lower dimension features through feature aggregation, it has been shown to be possible to detect osteoporotic vertebral fractures with an accuracy of 89.2% by feeding these aggregated features again into a CNN [35].

Of note is the fact that the uses of CNNs in diagnostic medicine typically only use two-dimensional cross-sectional CT scans, rather than utilising the intrinsic three-dimensional geometry required for this particular problem. CNNs capable of efficiently processing three-dimensional images are still in the early stages and volumetric representations of data can easily grow “*unwieldy and computationally intractable*” [17]. However, work using three dimensional CNNs has been performed in order to aid in medical diagnosis, specifically for detecting early warning signs of Alzheimer’s in Magnetic Resonance Imaging (MRI) scans of human brains [36]. This work was similar for the aspirations for this project in that it used a Class Activation Map (CAM) to identify key regions within a three-dimensional model which acted as early indicators for a disease and compared effectiveness of both CNNs and different methods of analysing their performance. Of the three 3D- CNNs used, 3D-VGGNet had the highest classification accuracy and largest area under the Receiver Operating Characteristic (ROC) curve indicating it may be best suited for processing three- dimensional medical images.

## Summary

A large part of this investigation will involve deforming an incongruent Labrador joint and looking for mismatches against “healthy” joints which do not present signs of incongruence or elbow dysplasia. Through this, we will be able to quantify the magnitude of incongruence. Attempts to quantify congruity between the ulnar trochlear notch and humeral trochlea have previously been successful, with both computed tomography and radiographic assessment of the shapes and peaks along the joint being consistent with each other. This suggests that if we can quantify the incongruity between the radial head and the ulnar notch in digital form, then said data should be a sufficient approximation for data obtained by observing the physical limb. However, the computerised three- dimensional models did typically detect a larger curvature which we may witness within our own work and must compensate for [37].

Little research appears to have been undertaken into attempting to specifically characterise joint incongruity via deformation patterns against healthy joints, and so this paper could potentially have been the first of its kind.

Though research into Chiari-Like Malformation has already been performed to some success, they have typically been performed using larger datasets and required a greater deal of pre-processing before classification. Regions of interest within the brain have been identified through both machine learning and clinical studies, but there does not appear to one definitive agreed cause and so research performed here will have to pursue multiple avenues.

Machine learning approaches to diagnosis for other neurological conditions, such as transfer learning and SVMs have yielded significant results. Reproducing these results within this investigation however may not be feasible; these studies have mostly been results driven and have been performed on conditions with clearly understood origins. As such, identification has been easier than it will be for this hypothesis driven study.

# Chapter 3: Preliminary work for investigation into Canine Elbow Dysplasia

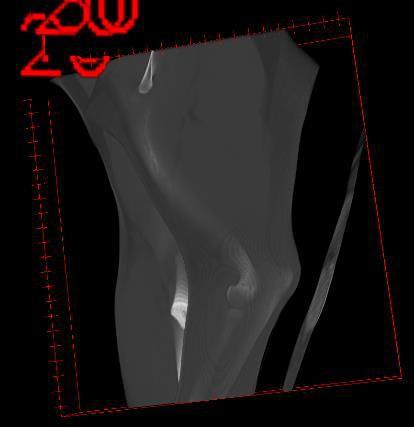
Further investigation into this avenue of research was abandoned after time constraints and circumstances beyond my control resulted in the requisite legal work necessary for my use of the dataset could not be completed in time. The preliminary work described in this chapter were performed on two scans with which I was provided early on in the project. An exemplar slice from the first scan can be found in Figure 4, and the three-dimensional reconstruction of the joint in Figure 29. The shape of the joint depicted in this scan has also been characterised by thresholding the pixel brightness of each slice to obtain the brighter regions indicating cortical (exterior) bone.

Figure - A three-dimensional reconstruction of a CT of a Labrador joint.

## Data Description

The data set with which we anticipated to perform our research would have been obtained from live subjects presenting with various degrees of limb lameness, with the majority of the data indicating some degree of incongruity alongside 80 “normal” joints which we would have used as our control.

Each joint image will be a CT file comprised of approximately 90 slices obtained via CT scans of the afflicted limb, obtained using a 164 slice clockwise spiral scanner with a slice thickness of 500µm. From the attached DICOM data, it is possible to view information such as the patient sex, age and weight.

The value of each pixel’s intensity in the image is based off the Hounsfield scale, the standard for quantifying radiodensity of different tissues. The expected data ranges for different tissues can be seen in Table 1.

Table - The expected values for radiodensity in Hounsfield Units for the different substances and tissues expected in a CT scan [19].

|  |  |
| --- | --- |
| **Substance** | **Hounsfield Units (HU)** |
| Air | -1,000 |
| Muscle | +35 to +55 |
| Soft tissue | +100 to +300 |
| Cancellous (inner) bone | +300 to +400 |
| Cortical (outer) bone | +1,800 to +1,900 |
| Foreign Metals | >14,000 |

For the preliminary work, the dataset available was composed of only two joints: The left and right elbows of a Labrador retriever which will be hence forth referred to as Elbow A and Elbow B. Elbow A consisted of 92 slices, and Elbow B of 90.

## Data Processing

### ImageJ

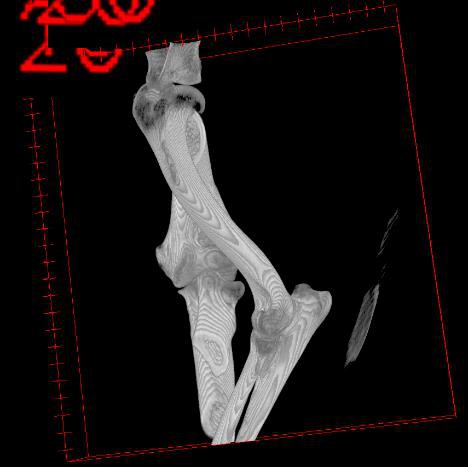
To obtain a three-dimensional model, the CT file for the original example joint supplied to us was imported as a sequence of images to ImageJ. Each slice was converted to an 8-bit grayscale image and the voxel depth set to match the slice thickness at 500µm. After conversion, the image sequence could then be converted into a three-dimensional model using the inbuilt 3D viewer plug in. The result of this can be seen in Figure 15, with the darker grey soft tissue surrounding the lighter bones.

Figure - A 3D reconstruction of an elbow joint from a series of CT scans, from the same perspective as Figure 15.

By then thresholding the image, the soft tissue regions could be removed in order to reveal the hard bone tissue beneath (see Figure 16). After empirical experimentation, it was found that a threshold at a grayscale pixel value of 99 provided the results which offered the clearest depiction of the bone structure while preserving the smallest amount of surrounding soft tissue.

### MATLAB

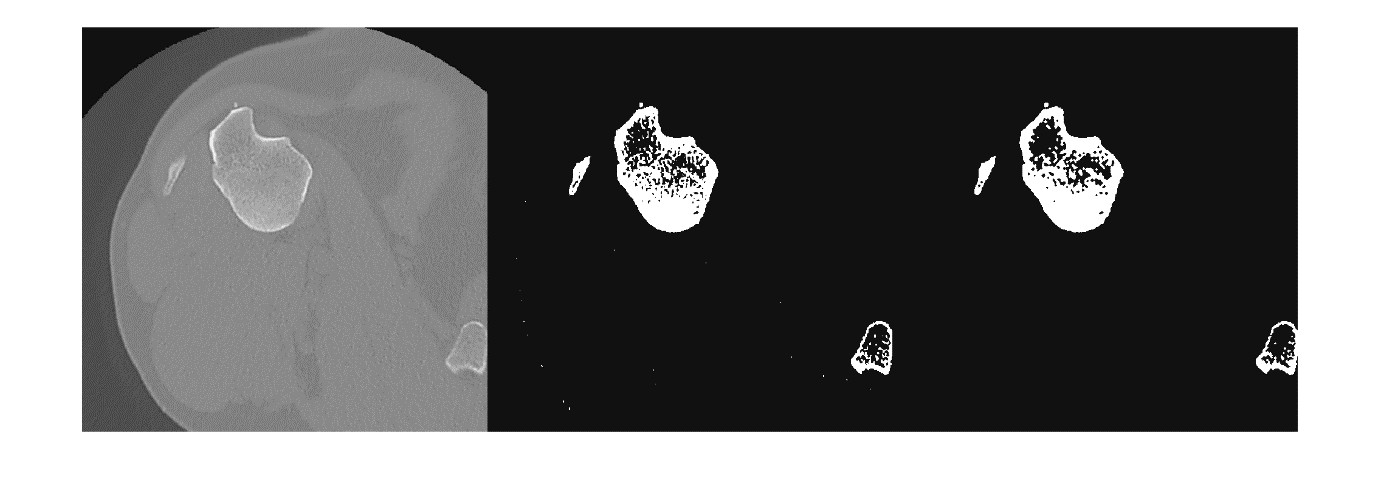
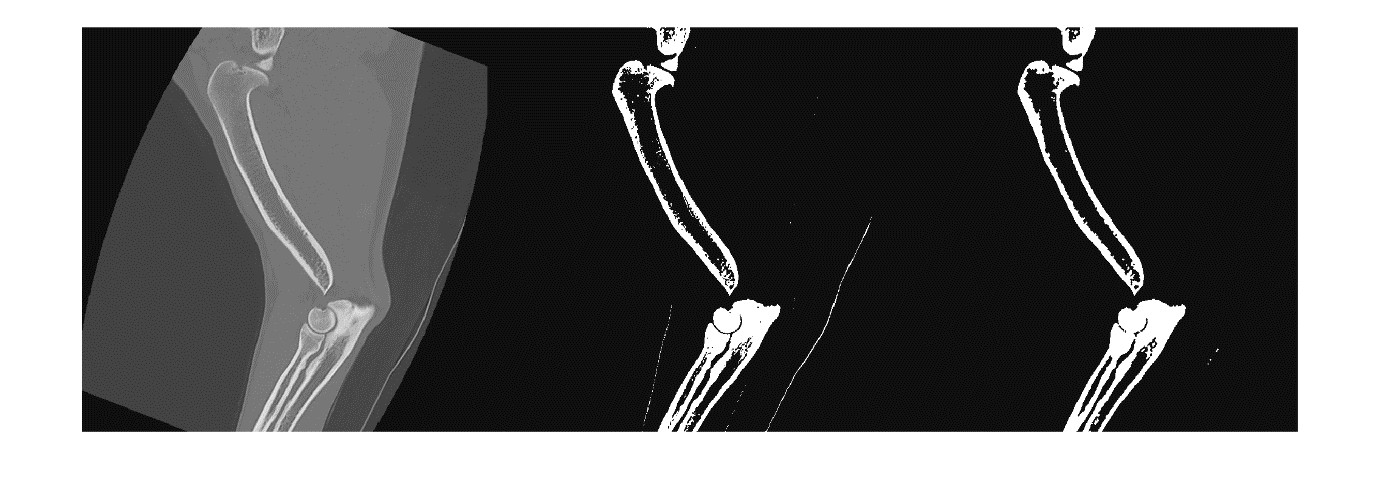
Each slice of the original CT files could also be opened in MATLAB to access the HU values associated with each pixel and segment the image. Through the “dicomread” command it is possible to access the HU values for each pixel and use it as the grayscale intensity value. As shown by the peaks in Figure 17, each slice in the CT scan is primarily composed of three different materials (air, soft tissue and bone) with different HU values.

A close up of text on a black background

Description automatically generated

Figure - Histogram showing distribution of values across all slices of a CT scan of an elbow joint.

By selecting a threshold at 300 HU to eliminate all soft tissue, it is possible to produce a binary mask of each layer containing only the cortical bone tissue, shown in the CT scans as a bright white area. These contained errant regions of high intensity, as shown by Figure 18. These could be eliminated using the inbuilt MATLAB “*imopen*” and then “*imclose*” commands to first erode and then dilate the image in order to form a clean mask.



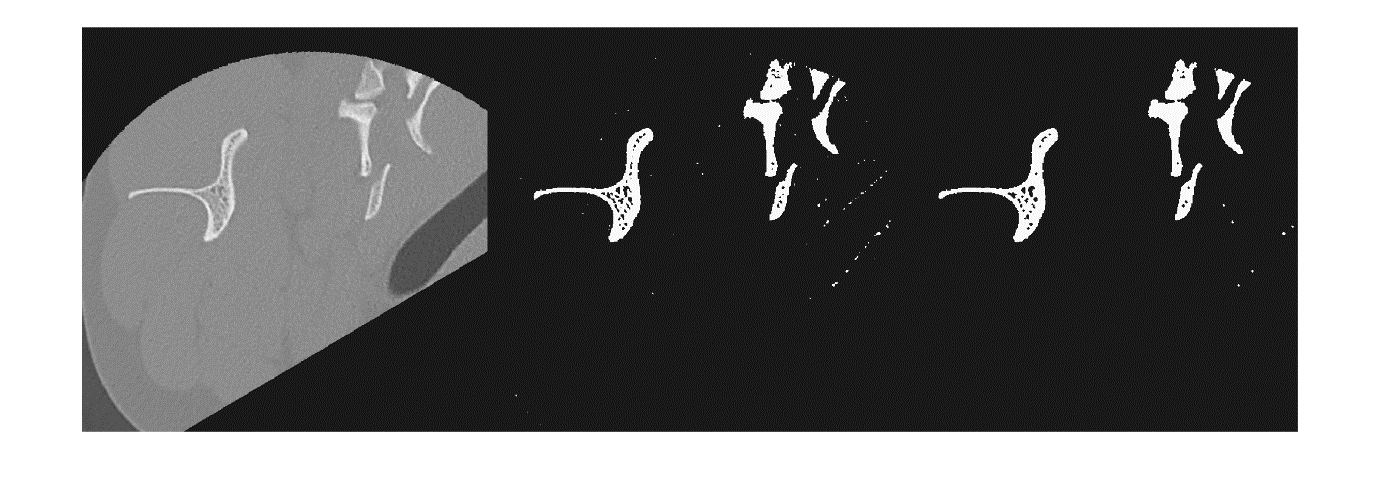


Figure - The original CT scans (left), the same CT scans thresholded to show only the cortical bone tissue (middle) and then the "cleaned up" binary masks (right).

From these binary masks, a three-dimensional point cloud could then be formed by plotting a point at each positive pixel, with the z-coordinate being calculated from the product of the number of the slice, the x-y scale factor and the slice thickness (500µm).

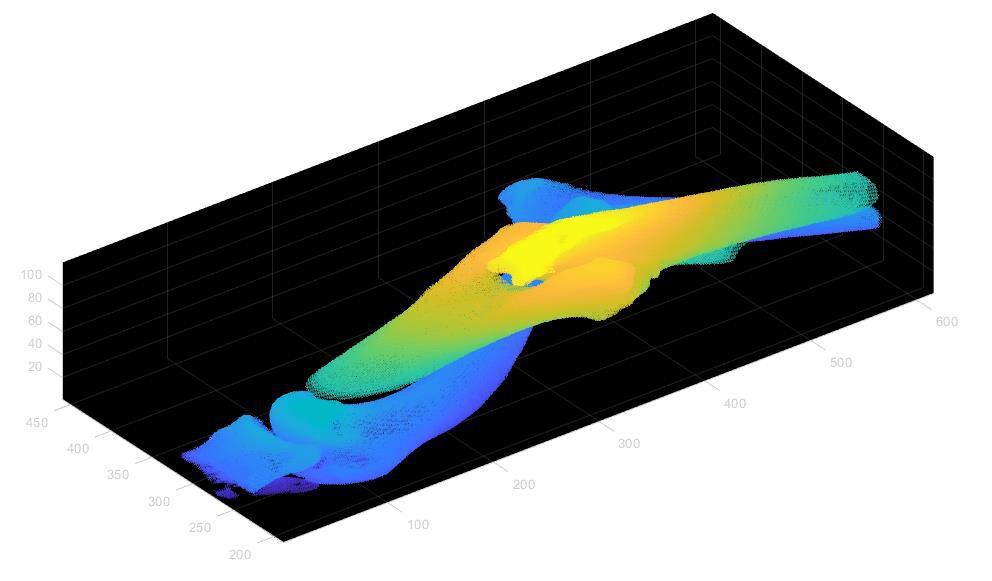


Figure - Visualisation of three-dimensional point cloud A, describing Elbow A.

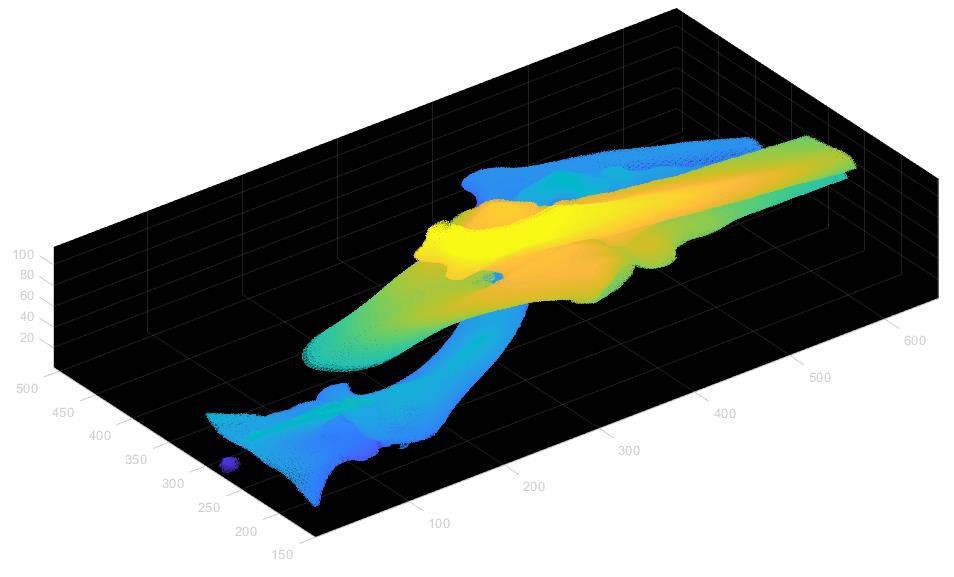


Figure - Visualisation of three-dimensional point cloud B, describing Elbow B.

As shown in Figure 19 and Figure 20, three-dimensional point clouds were formed for both images. To quantify the difference between the two joints, MATLAB’s inbuilt Iterative Closest Point (ICP) functionality could then be used to map the two point clouds onto each other and return a rigid transformation matrix describing the mapping process and the root mean squared error (RMSE) of the Euclidean distance between the two point clouds.

Equation - Known transformation matrix used to rotate Point Cloud A to Point Cloud A'.

To decrease the time cost of the ICP operation with the larger dataset, it may be advisable to reduce the number of operations required by downsampling using a Box grid filter. In order to identify the appropriate size of the box filter, an additional point cloud A’ was generated by rotating point cloud A 45° along the z-axis using known transformation matrix 𝑇𝐴 (Equation 1). A was mapped onto A’ using various different step sizes and the optimal choice selected based on the lowest RMSE. The times shown in Table 2 were obtained using a NVidia GeForce GTX 1080 graphics card with compute capability 6.1 and clock speed 1607 MHz and may be improved with access to more powerful hardware.

Table - The transformation matrices found by ICP mapping with various levels of downsampling alongside the corresponding RMSE and the time taken for the operation.

|  |  |  |  |
| --- | --- | --- | --- |
| Box filter | Transformation matrix | RMSE | Execution Time (s) |
| None |  | 1.1551 | 698.80 |
| 0.5 |  | 1.1550 | 603.08 |
| 1 |  | 1.2112 | 186.50 |
| 2 |  | 1.3861 | 16.58 |
| 5 |  | 2.3361 | 0.97 |
| 10 |  | 4.5663 | 0.42 |

As can be seen in , the magnitude and dimension of the rotation for mapping A onto A’ is similar for all levels of downsampling, with only the original number of points returning the original transformation matrix 𝑇𝐴. However, the RMSE value remains unchanged once downsampled using a step size of 0.5 and sees only a minimal increase at a step size of 1 in exchange for a significantly decreased computation time. As such, for the larger dataset it may be advisable to use a down

sampling step size of 1 when attempting to find the correct orientation of images. The product of reorienting A onto A’ using the transformation matrix found using ICP with a downsampling step of 5 is shown in .

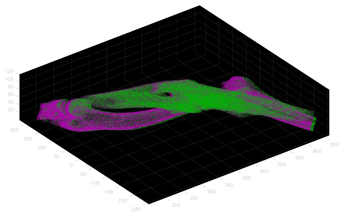


Figure - The results of rigid transformations via ICP of point clouds A (green) onto A' (purple).

This down sampling would not have affected later processing steps, as the down sampled models should only be used during the ICP process and the originals preserved.

## 3.3. Summary

Though the full work planned could not be completed, the software required to produce data for processing by a convolutional neural network was and experimentation was done to identify the optimal process for this. Had the full dataset been available, classification of the three-dimensional models by the CNN could then have been attempted and areas of interest identified by examining the Class Activation Maps (CAMs) produced. An “idealised” diseased joint could then have been envisaged through Google’s DeepDream software [38].

Figure - Summary of planned process to investigate the potential for artificial intelligence in understanding canine elbow dysplasia.

# Chapter 4: Methodology of Investigation into CLM

## Data Description

The dataset used for this project originated from an agreement between the University of Surrey and Fitzpatrick Referrals, and consisted of sagittal T2-weighted MRI scans of 19 anaesthetised CKCS. Of those 19, 14 had been previously diagnosed as CLM afflicted by a trained veterinarian, leaving a control group of 5.

Of the 14 affected by CLM, 11 also presented with syringomyelia and this will be reflected within the MRI scans. A brief description of the variation amongst the 19 patients can be found in Table 3.

|  |  |  |
| --- | --- | --- |
| Group | Affected | Control |
| Sex (M:F) | 5:9 | 5:0 |
| Age Range (Years) | 0.92 - 7 | 2 - 8 |
| Age Median (Years) | 4 | 7 |
| Weight Range (kg) | 6.4 – 12.5 | 9.2 - 16.15 |
| Weight Median (kg) | 8.6 | 11.95 |

Table - Summary of patient data within dataset

For this project, only the central slices of each sagittal scan were used. These were selected by taking the median value slice within each sagittal scan provided within the dataset.

## Data Processing

### Transfer Learning with a Convolutional Neural Network (CNN)

Due to the limited size of the dataset and limited scope of the project, creating and training an original CNN specifically for recognition of CLM symptoms was considered impractical. Instead, a pre-existing CNN was retrained through bottleneck feature extraction. Due to its versatility in regard to transfer learning [39] and ability to generalise well [40], VGG-19 was chosen for this investigation. VGG-19 is a forty-seven-layer network containing nineteen learnable layers. The pre-existing instance of VGG19 used in this investigation was originally trained on a set of over a million images from the ImageNet database [41] to recognise everyday objects such as stationery and animals. This original network would sort data into thousands of potential categories, so it was necessary to modify the network through replacement of the final Classification layer, as shown in Figure 23, with one with only two outputs. This produced a binary classifier.

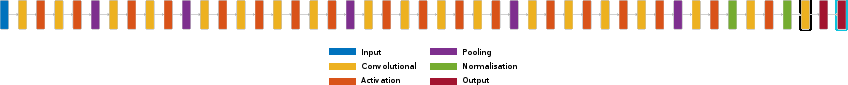


Figure 23 - The structure of the VGG19 network. The final Classification layer is highlighted in cyan, and the final fully connected layer is highlighted in black.

The weights within the nodes of the network would originally be set so as to aid in classification of the images it was trained on rather than MRI slices, so to modify these values the final Fully Connected Layer, shown in Figure 23, was replaced with one of significantly higher learn rate. The first ten layers were then “frozen” by setting their learn weight to zero and a low value used for the initial learn rate when training the data. This forced the vast majority of the learning performed by the modified network to occur within the final layers and so focus on superficial regions of the image rather than the more general shapes the network had previously been trained to recognise. The epoch number and batch size were selected through trial and error to produce the highest potential accuracy.

To obtain a lower bound result, the classifier was first run on a dataset with minimal pre-processing. For each subject, the central slice of the MRI scan was isolated and for scans which included the cervical spine in addition to the skull the images were cropped to form a dataset of uniform dimensions. Due to the small amount of data available, leave-one-out cross-validation was deemed suitable for testing the performance of the model and so the dataset was partitioned so that the Test Set contained only a single image. The model was then repeatedly trained using a different test image each time until each image in the dataset had been used, and the results averaged to produce values for overall accuracy, specificity and sensitivity.

Figure 24 - Distribution of intensity levels over the dataset consisting of central slices from the MRI scans of nineteen CKCS.

A picture containing nature, water, sitting

Description automatically generatedThese values left room for improvement, with the specificity in particular being low, so optimisation through alignment of the images prior to classification was then attempted. A histogram of the pixel values was formed (), and from the peaks within this it was possible to identify the values associated with brain and bone tissues by their prevalence. By thresholding the image so as to leave only binary masks describing these tissues, a “master” stationary image was selected from the control group and noise and inconsistencies removed to create a template. Each of the 19 individual masks were then mapped onto this template through an Iterative Closest Point (ICP) algorithm. The affine transformations returned by the ICP algorithm could then be applied to the original central slice images, and the now homogenous dataset supplied to the retrained network for improved results after manually cropping the images to a uniform size where necessary.

Figure - The average edge map of the 5 control images generated using a Sobel filter, to be used as a stationary image during the ICP phase.

Once the training of the network had been completed, it was then possible to identify the regions within the canine brains which the retrained CNN associated with CLM. The dot product of the feature map of the model and the extracted “weights” from the final layer was then calculated to produce a Class Activation Map (CAM).

### A close up of a logo Description automatically generatedAffine Transformations as Feature

In addition to being used to produce a homogenous dataset for the CNN, the affine transformations obtained through the ICP algorithm were used as a feature in their own right. Here, the stationary image was generated by producing edge maps of all nineteen images and averaging the controls to produce Figure 25. A binary mask, shown in Figure 12, was then produced for each of the nineteen central slices, and a two-dimensional point cloud created from this within three-dimensional space. The ICP algorithm could then be applied to map the masks onto the average shown in Figure 25, producing a 4x4 transformation matrix which acted as a 16-dimensional feature.

Figure - A binary mask created by thresholding a central slice of an MRI scan of CKCS.

A close up of a map

Description automatically generatedThis data was then used to train a Support Vector Machine (SVM). As in Section 3.2.1, a leave-one-out cross validation system was used so that eighteen points were used to establish the geometry of the separating hyperplane dividing the two categories and then one additional point placed within the feature space and classified based on its position relative to the hyperplane. This was then repeated eighteen times to again obtain average accuracy, specificity and sensitivity.

To improve these values, the dimensionality of the features was then reduced via Principle Component Analysis (PCA). As could be expected, the dimensions relating to rotation and translation showed little variation leaving the focus on skew. The cross-validation loss was then minimised through Bayesian optimisation, meaning that repeated trials of the training of each network were performed during each stage in the leave-one-out cross validation and the optimal fit for the sixteen-dimensional and three-dimensional hyperplanes was found by random assignment of points within the plane until optimal accuracy was achieved and discovered via gradient descent.

Figure - The objective function model generated for the 9th iteration of the leave-one-out cross-validation phase, indicating the optimal fit for the three-dimensional hyperplane.

Attempting to construct a three-dimensional point cloud using the entire MRI scan for this area would yield unproductive results, due to the dataset available containing scans of varying volume and hence incomparable models.

## Summary

This project explores two approaches to using machine learning for diagnosis of Chiari-like malformation, with similar pre-processing steps for both. Both used the information contained within the central slice of an MRI scan of the head of a Cavalier King Charles Spaniel, and both will require the central slices to be aligned using an ICP algorithm before the final result can be produced. A block image of the process can be found below in Figure 28.

Figure - A block image describing the methodology of this project.

Chapter 4: Results and Discussion

## Transfer Learning

During the training phase of the network, it can be seen that the area under the curve is typically lower than one would desire.



A close up of a map

Description automatically generated

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

Figure 29 - The training progress for the first through fifth iterations of the leave-one-out cross-validation process.

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

Figure 30 - Results for iterations six through thirteen, with the same colour scheme as Figure 11 and an overall learning rate of 0.0003.

A close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generatedA close up of a map

Description automatically generated

Figure - Results for iterations fourteen through nineteen, with the same colour scheme as and an overall learning rate of 0.0003.

Validation often remains at 50% for the entire duration of the training phase, while training loss and validation loss mostly follow similar curves showing that neither underfitting nor overfitting are occurring. The latter implies that the low training rate of 0.0003 should be sufficient for transfer learning, meaning the former issue should purely be a result of an inability to form generalisations.

While the accuracy of the classification performed by the retrained neural network upon completion was above the 0.5 value you would expect of purely random classification, the results in show that this result may not be as good as it appears. An overall accuracy of 0.6842 and 0.7368 was achieved for classification of unaligned and aligned images respectively, but this appears to have been heavily weighted by the comparatively low number of control images.

An explanation for this high sensitivity yet low specificity may be the small and uneven nature of the dataset. For the fourteen affected images used to train the network, there would be only four control while the final fifth was held back for testing. This likely resulted in the network being unable to form accurate generalisations about control images, and this behaviour not being penalised as overspecialisation and an increased likelihood to classify as *Affected* would be rewarded during the validation phase. Full results for this experiment can be found in Appendix B.

Table 4 - Comparison of results from classification via transfer learning for an unaligned and aligned dataset.

|  |  |  |
| --- | --- | --- |
|  | With Unaligned Dataset | With Aligned Dataset |
| Rate of True Positive | 0.786 | 0.928 |
| Rate of False Positive | 0.6 | 0.8 |
| Rate of True Negative | 0.214 | 0.071 |
| Rate of False Negative | 0.4 | 0.2 |
| Average Certainty | 0.921671053 | 0.906621579 |
| Standard Deviation of Certainty | 0.121330492 | 0.089782302 |
| Maximum Certainty | 1 | 1 |
| Minimum Certainty | 0.58716 | 0.72568 |

Further evidence for is provided by the CAMs on the next two pages, which show little consistent trend for areas of interest between patients and do not conclusively prove any of the hypothesises proposed by prior studies [29] [9] [28]. Of the nineteen class activation maps produced, eleven appear to show some focus on the medullary kinking at the craniocervical junction, lending credence to that as a potential cause but this may be influenced by the development of syrinxes within the same region causing the CNN to inadvertently consider symptoms of Syringomyelia rather than the desired Chiari-Like Malformation. The other CAM’s tendency to focus on irrelevant areas such as the jaw and surrounding black regions suggests that the CNN was unable to form accurate generalisations, while the leave-one-out validation method would have meant that even if a single functioning model was produced it would only provide a single result. As such, a larger and more even dataset which can be validated without a leave-one-out method may produce more accurate and useful results.

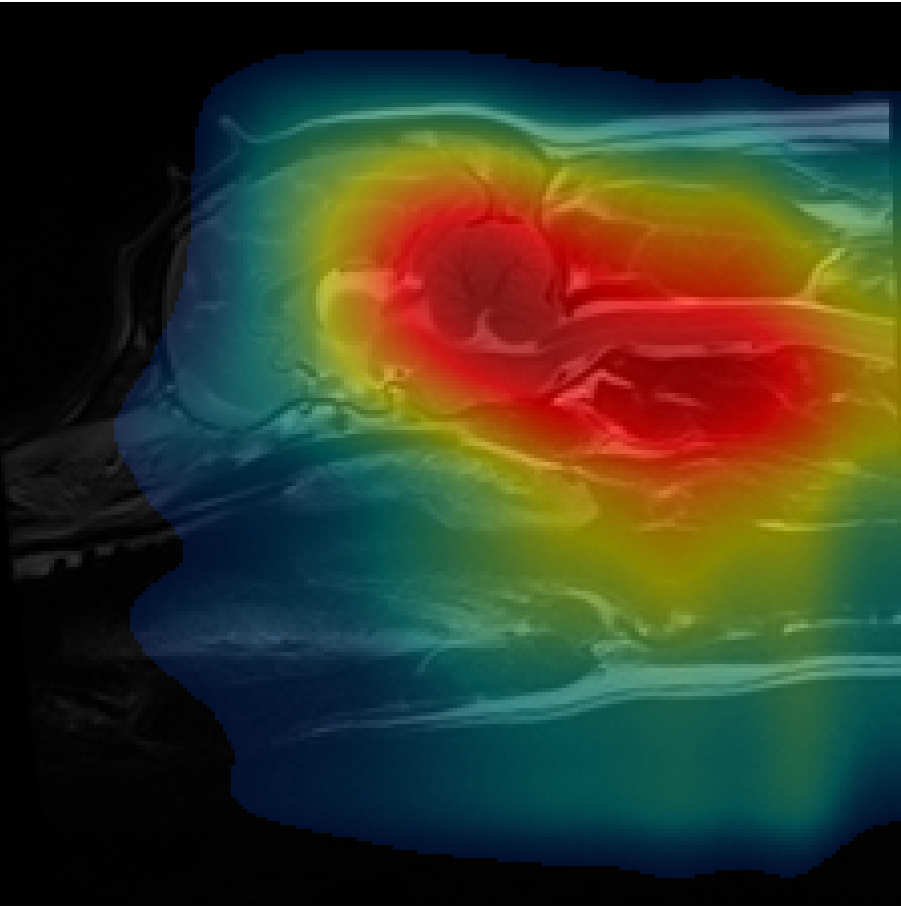
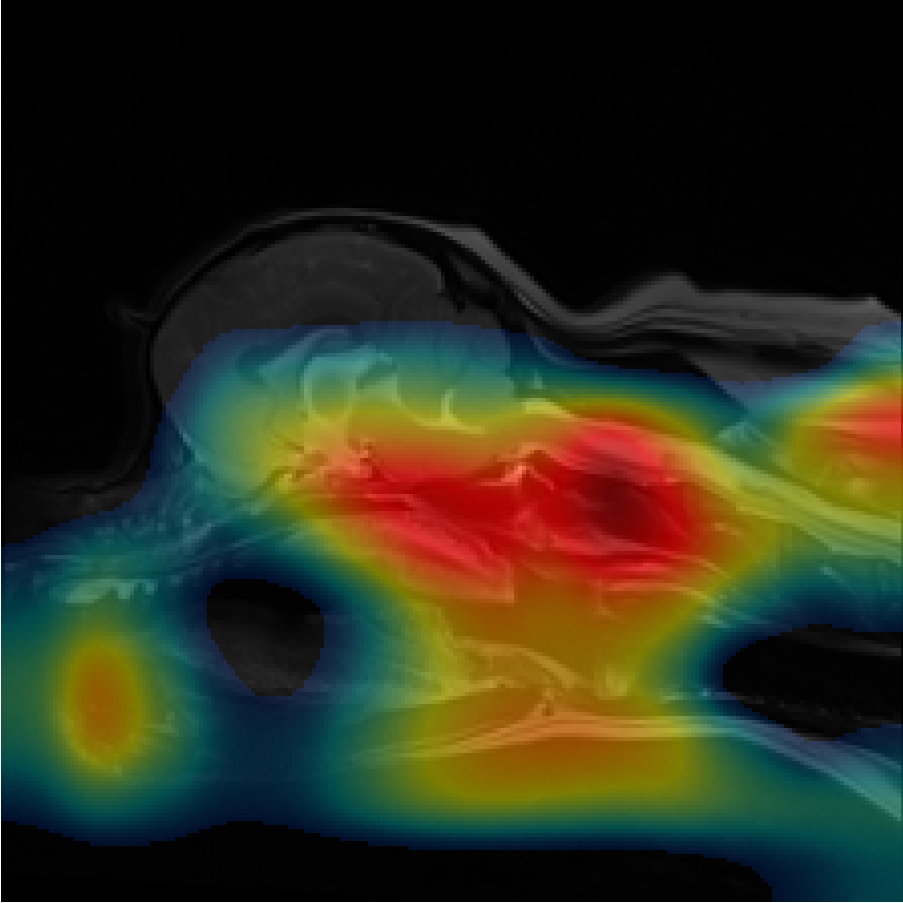
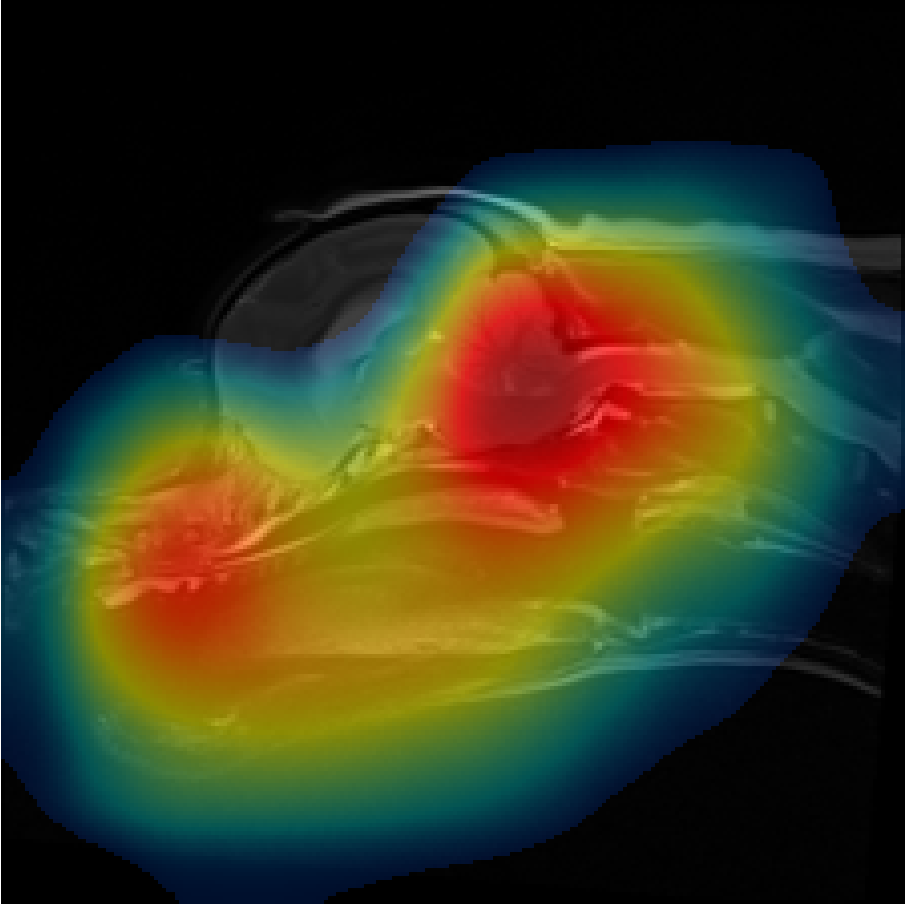
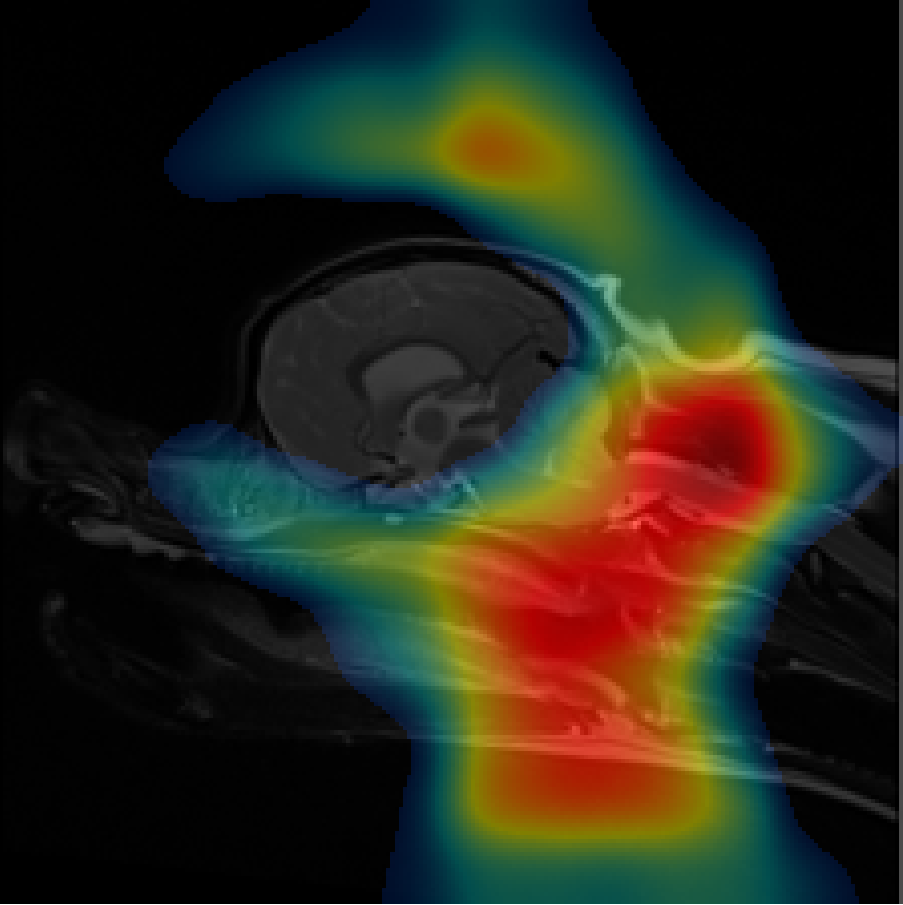


Figure 32 - Four Class Activation Maps showing a significant weighting granted to the craniocervical junction. The top three images are from the “affected” group and the bottom from the “control” group. All affected were correctly diagnosed, while the control was falsely positively diagnosed.

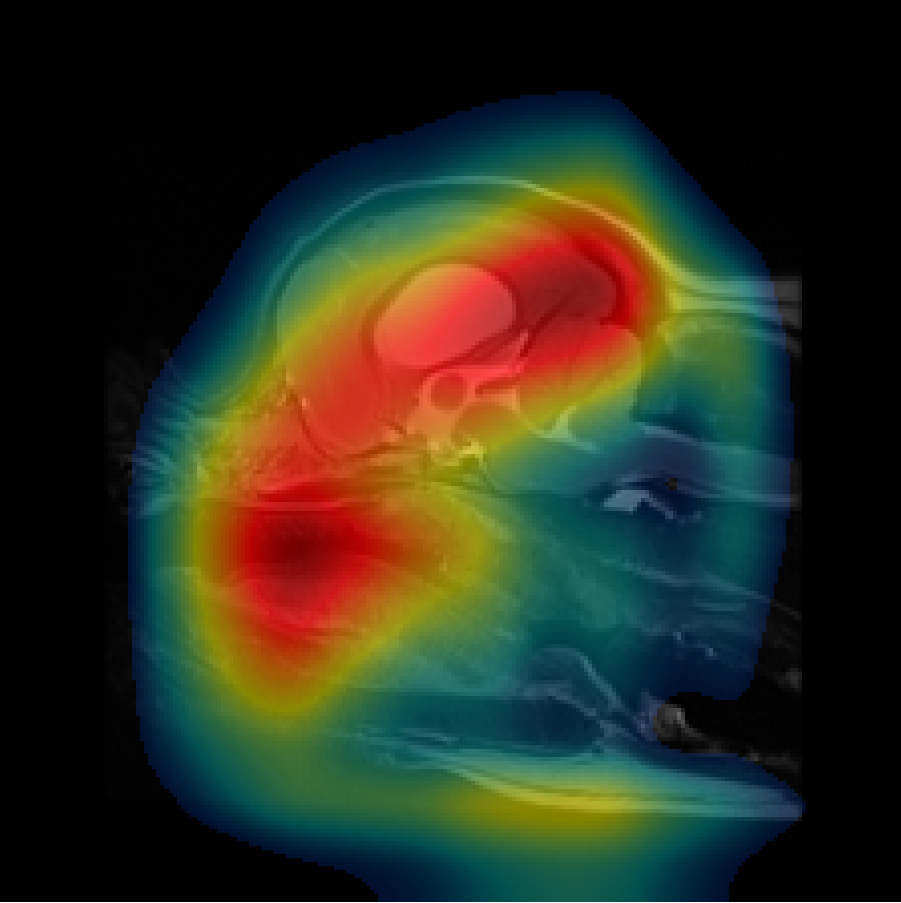
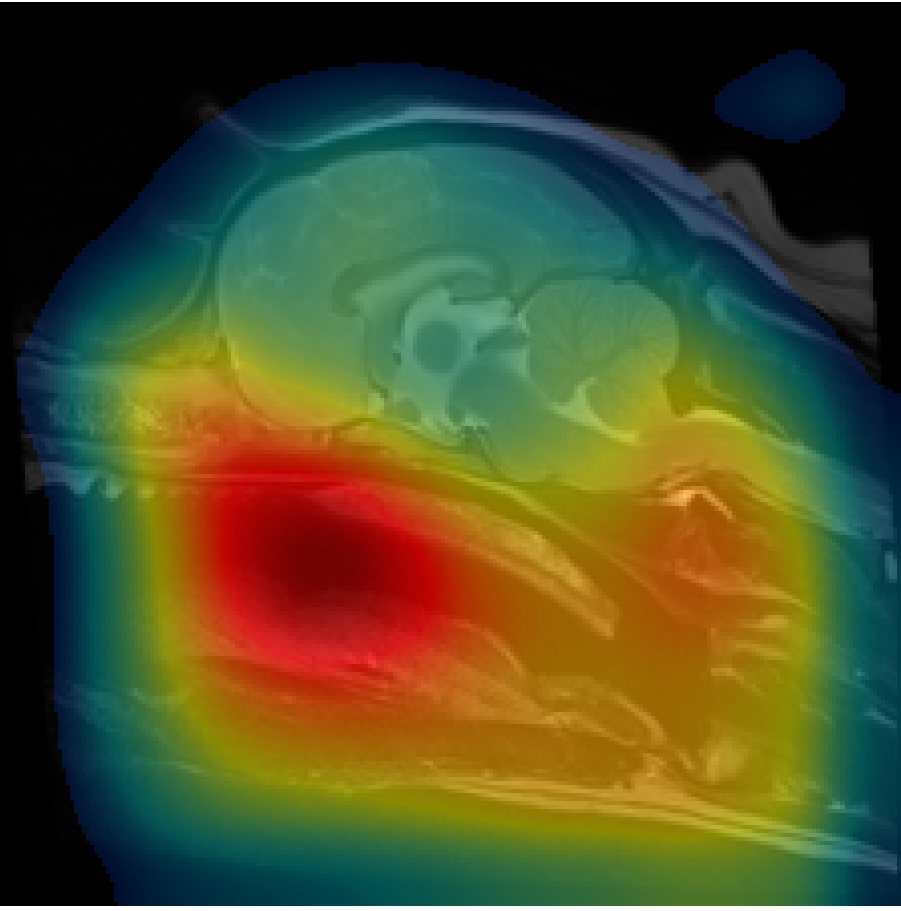
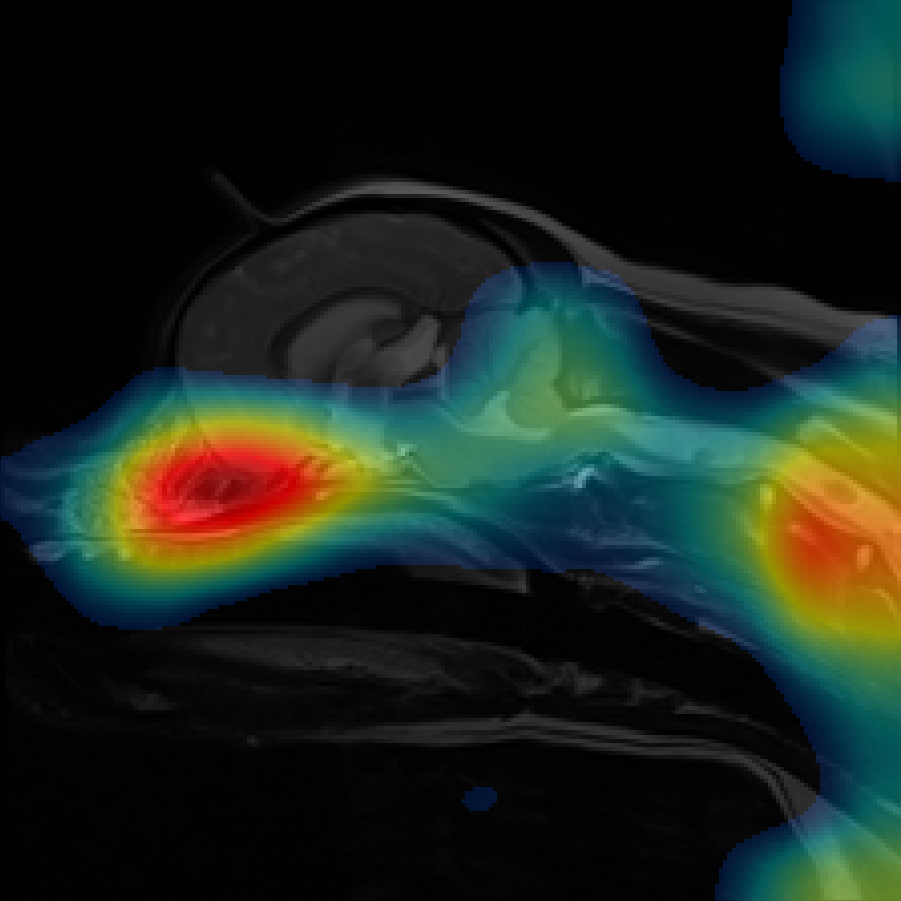
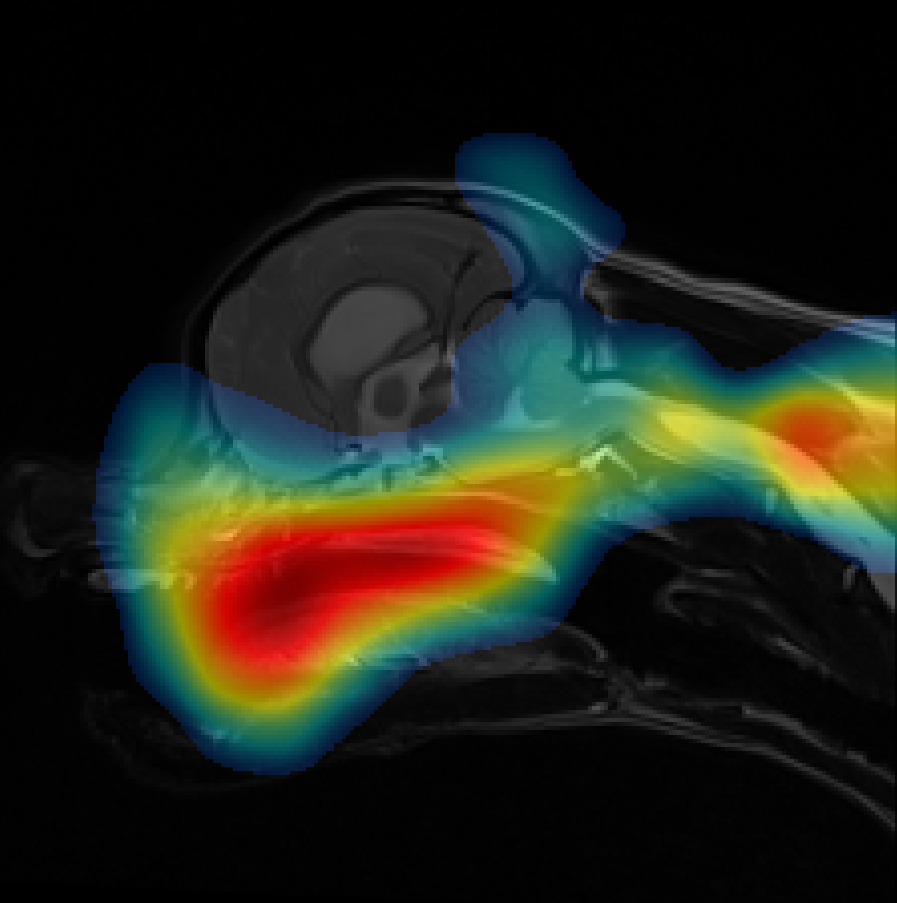
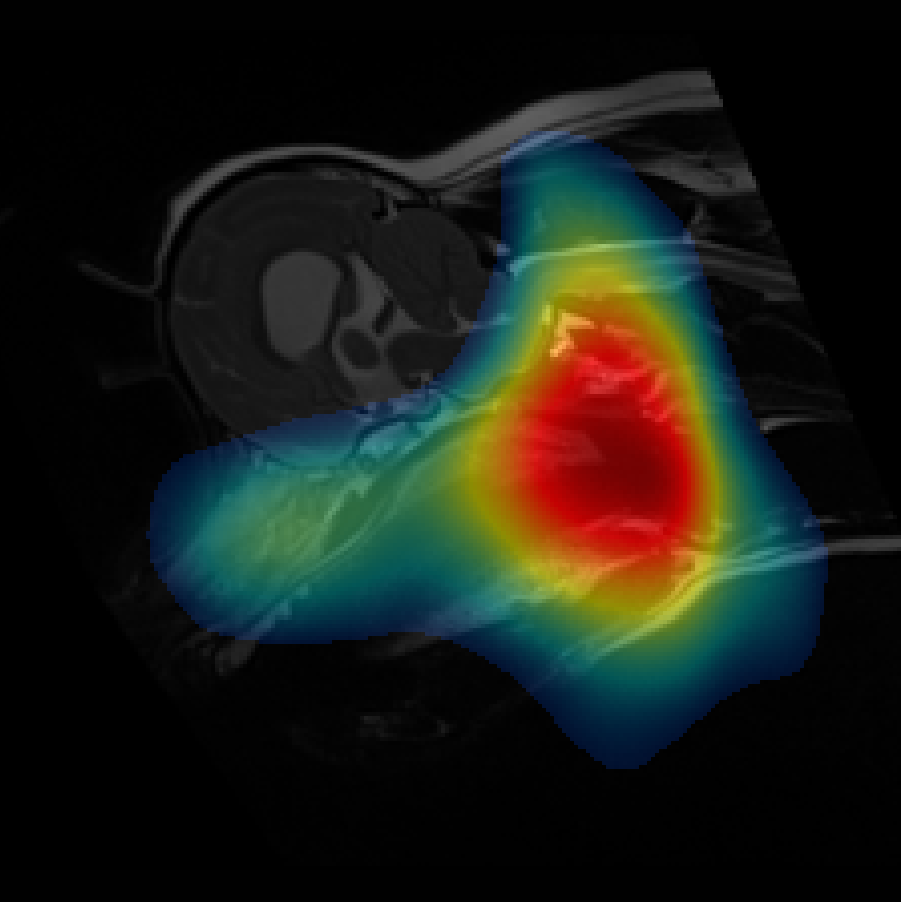
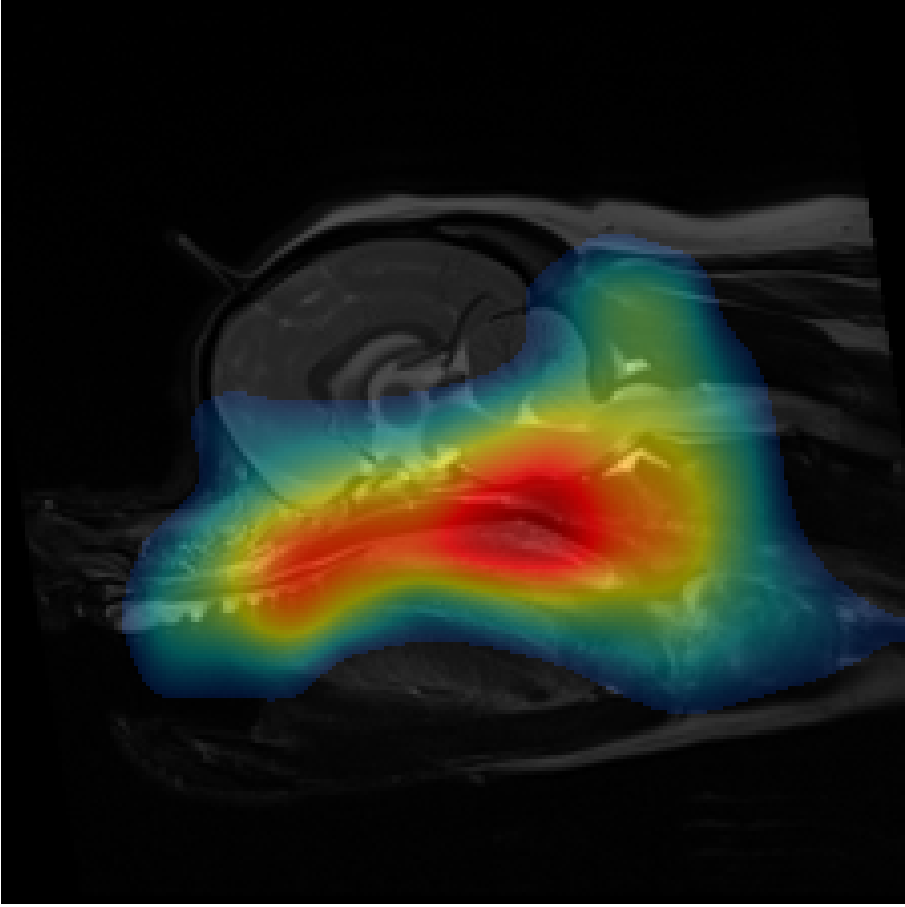


Figure 33 - Six Class Activation Maps showing moderate weighting granted to the craniocervical junction. All six were affected by the disease and were correctly diagnosed.

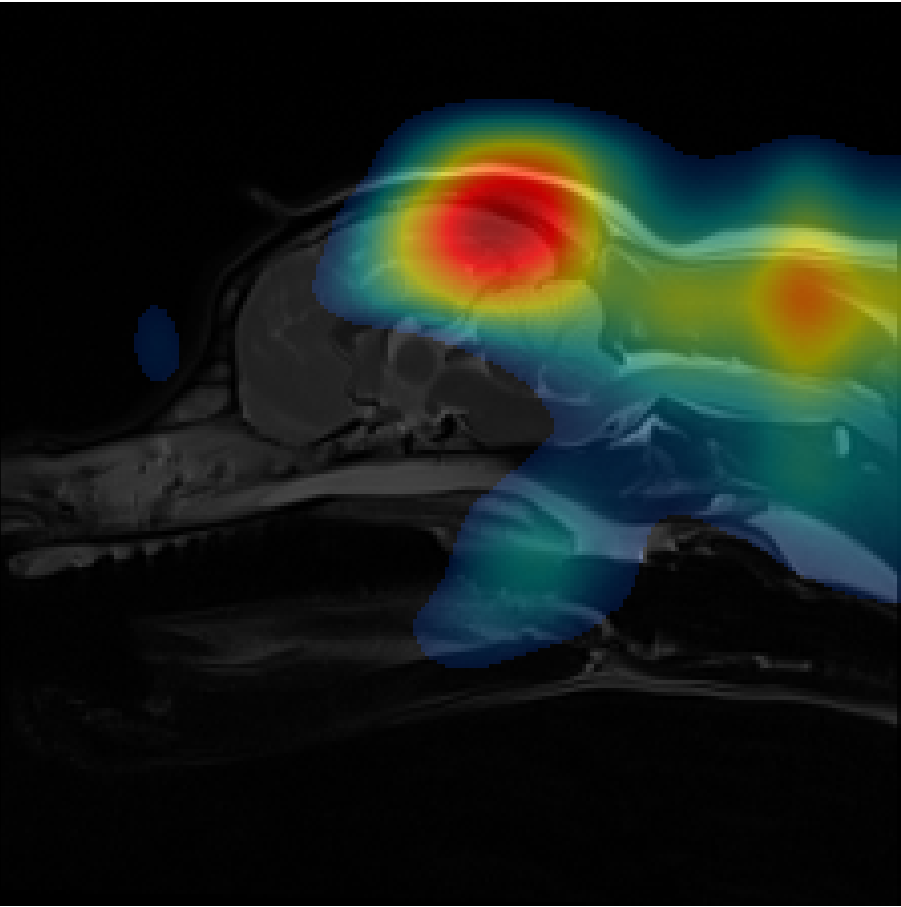


Figure 34 - A member of the control group which also indicated the craniocervical junction as an area of interest. This patient was correctly diagnosed.

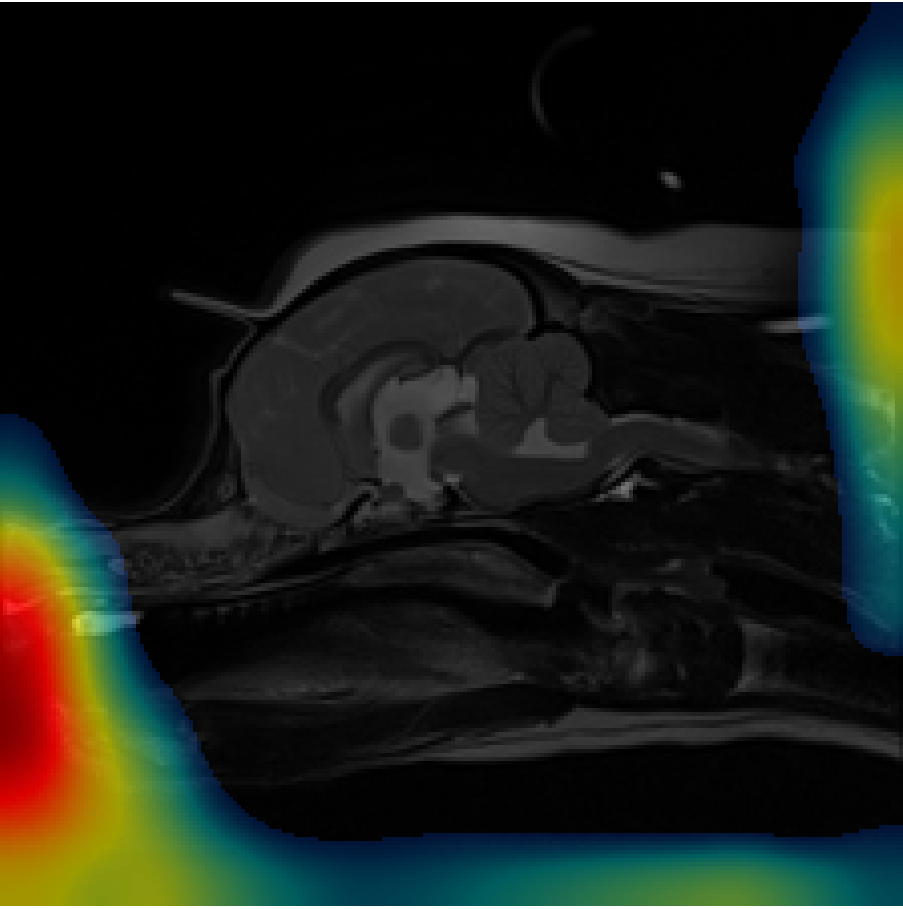
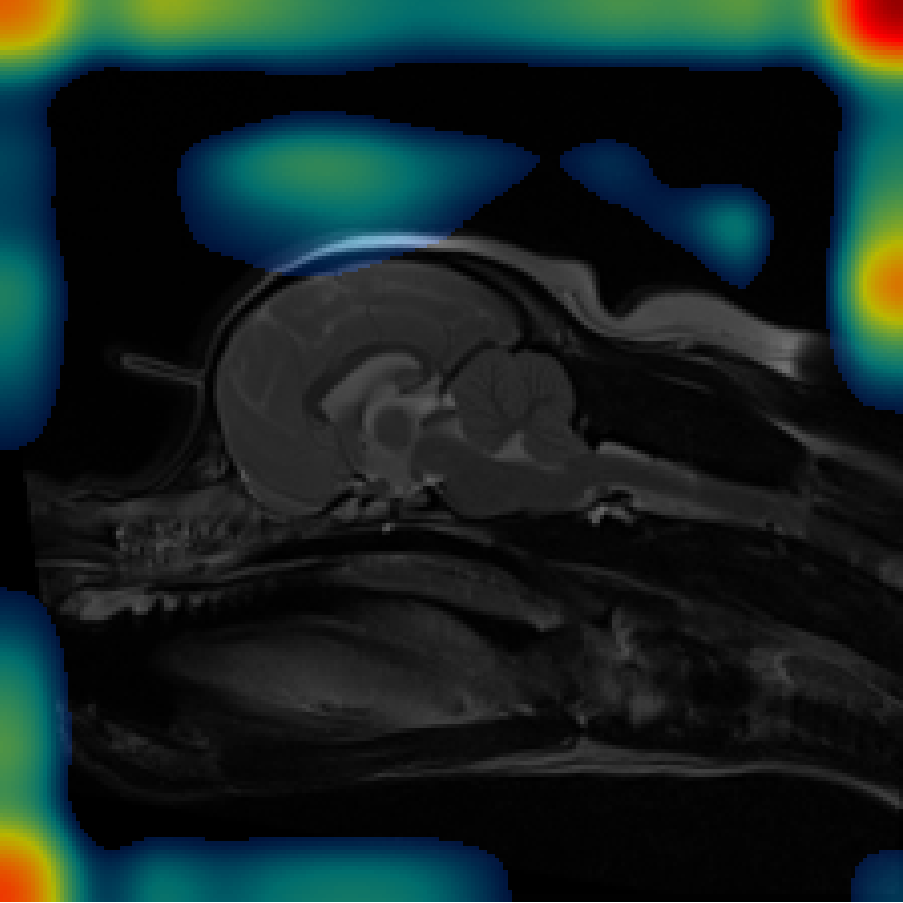
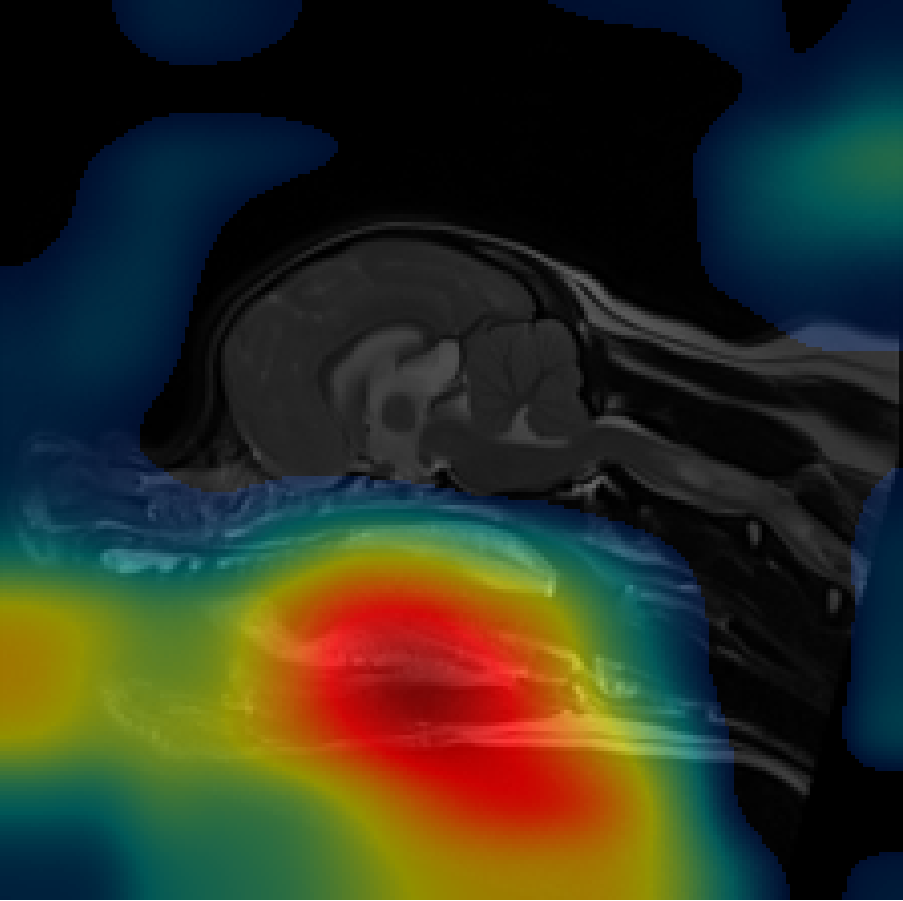
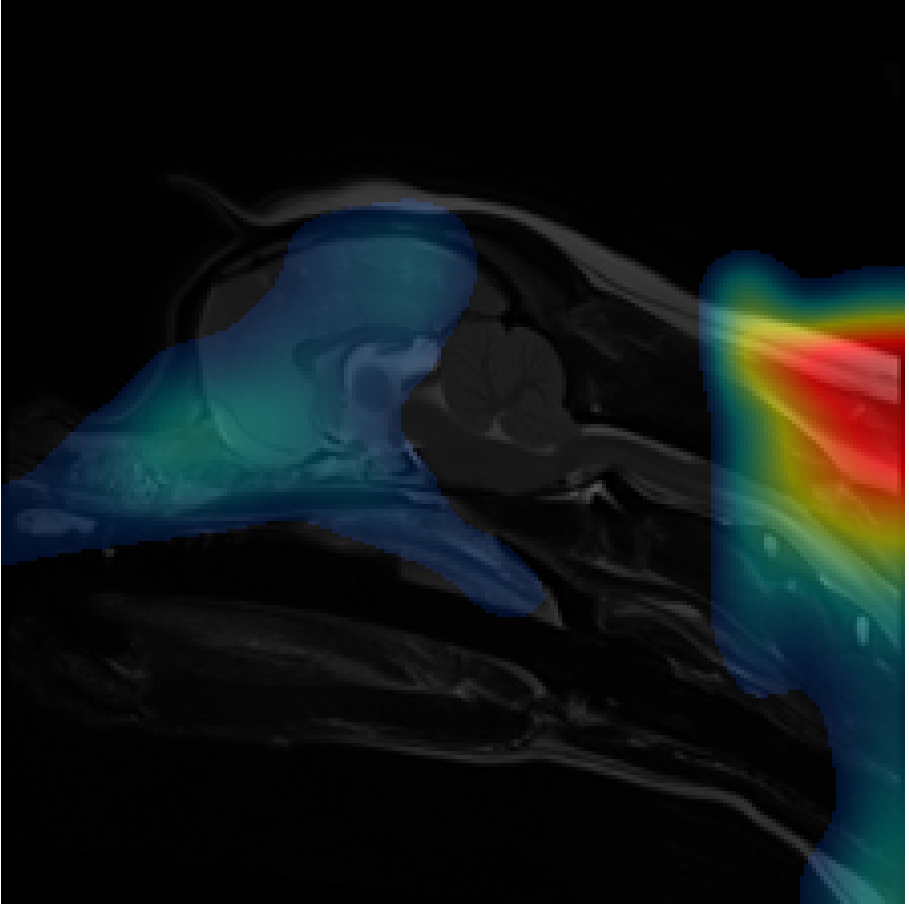


Figure 35 - Five Class Activation Maps for members of the Affected group with areas of interest apparently irrelevant to Chiari-Like Malformation. Only the top right patient was incorrectly negatively diagnosed.

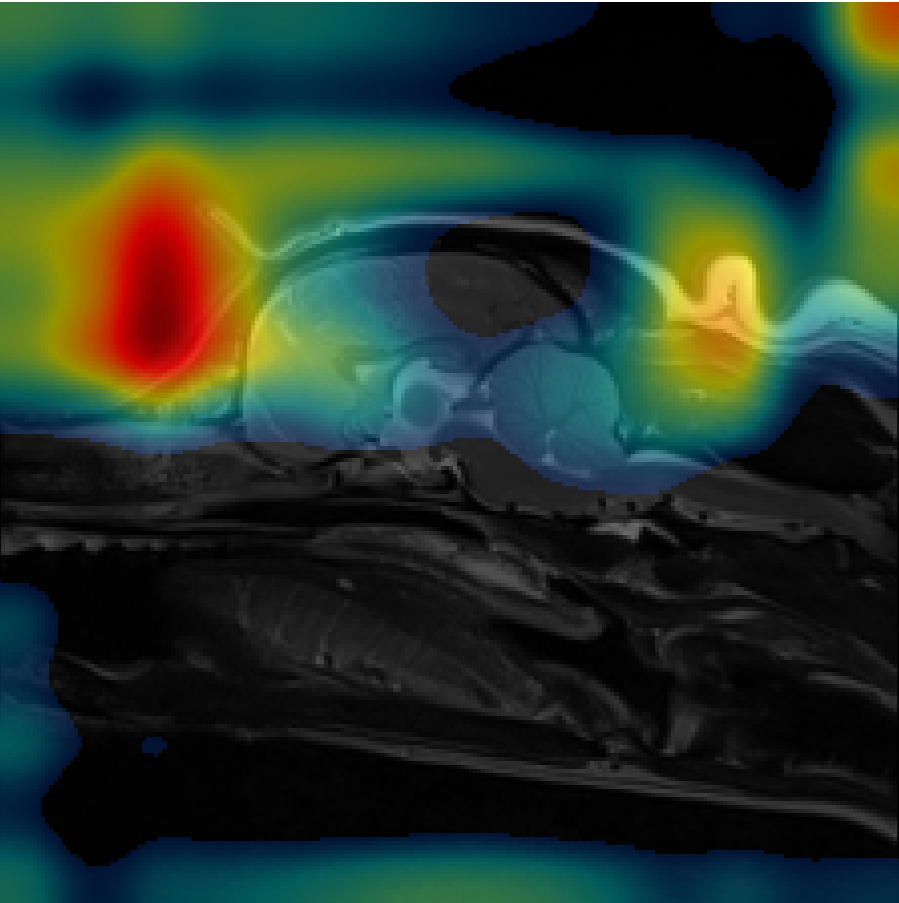
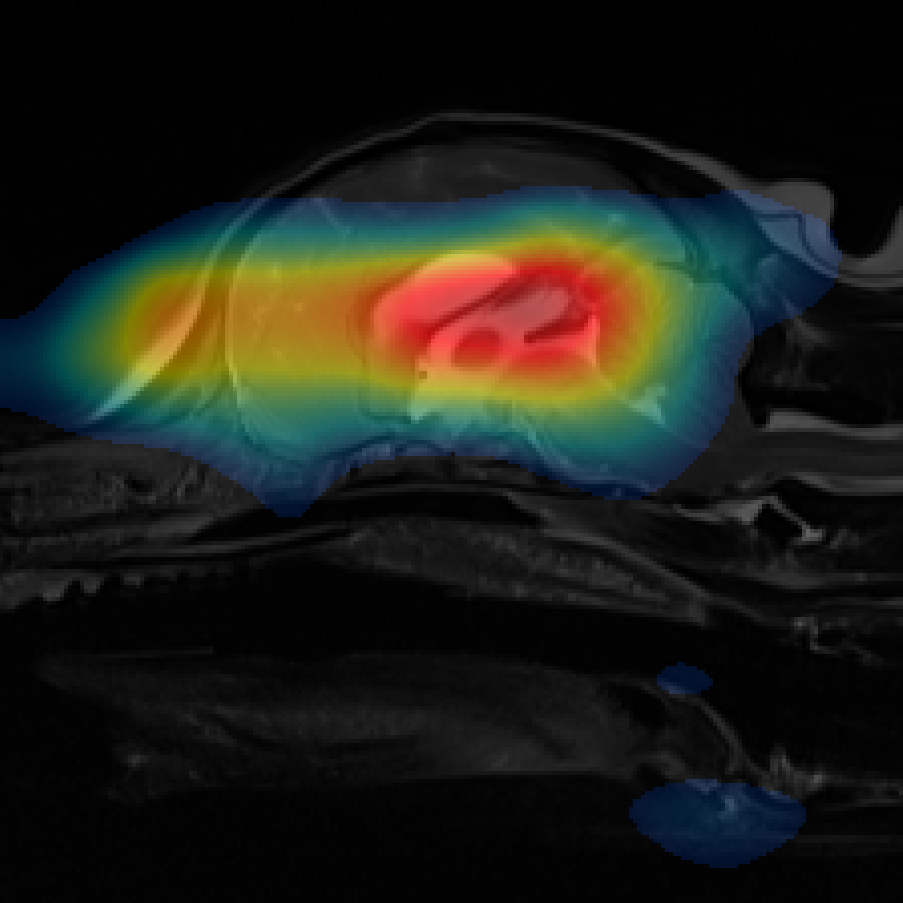
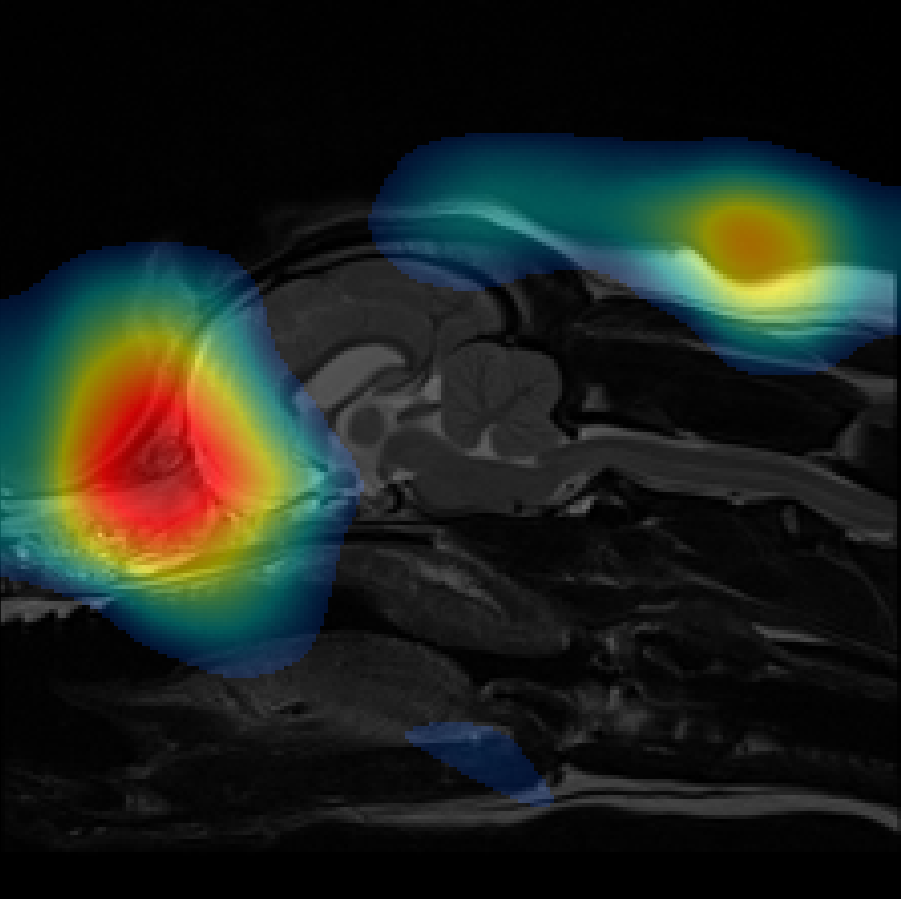


Figure 36 - Three Class Activation Maps for members of the Control group with no discernible pattern to areas of interest. All three were falsely diagnosed.

## Affine Transformation as a Feature

After attempting to plot the affine transformations generated by the ICP process as points within the feature space, it would appear that the level of skew needed to map the images onto the base has no correlation with CLM. All nineteen images were diagnosed with the condition during the cross-validation phase. Attempting to refine the results through PCA resulted in no significant change, with their being no practical way for a three-dimensional hyperplane to separate the two groups due to their similarity.

|  |  |  |
| --- | --- | --- |
| True Label | Predicted Label | Confidence Interval (CI) |
| Affected | Affected | 0.99994 |
| Affected | Affected | 0.99978 |
| Affected | Affected | 1.3826 |
| Affected | Affected | 0.99894 |
| Affected | Affected | 1.0176 |
| Affected | Affected | 0.86 |
| Affected | Affected | 0.99972 |
| Affected | Affected | 1.0051 |
| Affected | Affected | 1.002 |
| Affected | Affected | 1.0002 |
| Affected | Affected | 1.0029 |
| Affected | Affected | 0.94462 |
| Affected | Affected | 1.0013 |
| Affected | Affected | 1.3967 |
| Control | Affected | 1.0003 |
| Control | Affected | 1.0056 |
| Control | Affected | 0.99905 |
| Control | Affected | 1.0009 |
| Control | Affected | 0.9999 |

Table -Results of classifying dataset using SVM based off transforms required to map images onto average of controls, after Principal Component Analysis and Bayesian Optimisation.

These results may have been somewhat affected by the non-uniform nature of the dataset, with variation in the angle of the heads during the scan potentially affecting the effectiveness of the ICP algorithm. If a full scan of the brain could be obtained and the affine transformation for the three-dimensional model used instead then the results may be more productive.

# Chapter 5: Conclusions and Future Work

Over the course of this project, multiple approaches to diagnosing Chiari-Like Malformation within canines through machine learning were attempted with various degrees of success. Transfer Learning with a Convolutional Neural Network provided the best results, with a peak accuracy of 0.7368. Classification based on the affine transformations needed to map edge maps of each image onto a generated average of the control images proved fruitless, with no correlation being found between this descriptor and the presence of CLM.

Though none of the approaches explored over the course of this project met the high level of accuracy needed to be used within medicine, they were not without merit or potential. Based off the low specificity of this approach, it is possible that this may still result in the transfer learning approach being limited as a diagnostic tool, but it may still have use when attempting to understand the condition’s causes. The highlighting of regions identified by prior studies within the Class Activation Maps generated by the Transfer Learning approach indicate that further research with a larger dataset may potential provide weight to existing hypothesises, particularly that of medullary kinking.

The dataset was one of the great limitations for this project, with the small overall size and unequal distribution resulting in leave-one-out cross-validation to be used despite its unreliability. Few conclusions can be made about the ability to recognise control images due to the lack of results, and the neural networks inability to form accurate generalisations likely stems from the unequal split rather than any software-based limitation.

Aside from repeating the investigation with a larger dataset, further work could involve attempting to create a three-dimensional model of the brain within the skull and attempting to retrain an existing network with that rather than the two-dimensional image of the central slice. This would eliminate the discrepancies which came as a result of varying head positioning within the scanner, and would allow for a deeper investigation of the correlation between the shape of the brain and the skull and CLM. The investigation performed here could even be attempted with central slices extrapolated from the generated three-dimensional model for more accurate results. If more consistent Class Activation Maps can then be generated, Google’s Deepdream could then be used to envisage an idealised version of a CLM afflicted head.

In addition to the issues with the dataset, this project was limited in scope by both time constraints and issues beyond the university which caused the aim of the project to be changed from investing canine elbow dysplasia to investigating Chiari-like malformation. These issues were largely unavoidable and could not have been foreseen when the project first began. As a result of this change, there was a considerable gap where no work could be performed other than fine tuning preliminary work which would later be revealed to be irrelevant. This has been reflected within the Gantt chart for this project shown in

Overall, while no definite conclusions can be made about CLM there have been multiple interesting avenues for potential future research revealed. The low specificity of transfer learning approach employed here mean that despite it’s high accuracy, the model developed would be of limited use to current veterinary work, but the high sensitivity indicates that the use of machine learning in diagnosing CLM should be further investigated.

# Appendix A: Project Management



Table - Gantt chart for the entirety of the project, updated to reflect the change in aim for the project. Deadlines are indicated with a "!".

# Appendix B: Unabridged Results of Classification through Transfer Learning

|  |  |  |  |
| --- | --- | --- | --- |
| True Label | Classification | Affected Score | Control Score |
| Affected | Control | 0.00091011 | 0.99909 |
| Affected | Affected | 0.98229 | 0.017708 |
| Affected | Affected | 0.62008 | 0.37992 |
| Affected | Control | 0.033639 | 0.96636 |
| Affected | Control | 0.22736 | 0.77264 |
| Affected | Affected | 0.96106 | 0.038938 |
| Affected | Affected | 0.91918 | 0.08082 |
| Affected | Affected | 0.8943 | 0.1057 |
| Affected | Affected | 0.9593 | 0.040696 |
| Affected | Affected | 0.99944 | 0.00056119 |
| Affected | Affected | 0.99762 | 0.0023837 |
| Affected | Affected | 0.58716 | 0.41284 |
| Affected | Affected | 0.96921 | 0.030791 |
| Affected | Affected | 0.99877 | 0.0012266 |
| Control | Affected | 1 | 2.6329e-06 |
| Control | Control | 0.051473 | 0.94853 |
| Control | Affected | 0.99911 | 0.00089427 |
| Control | Control | 0.00013338 | 0.99987 |
| Control | Affected | 0.93774 | 0.062257 |

Table - Classifications and scores for each of the 19 unaligned central slices classified by a retrained VGG19 CNN.

|  |  |  |  |
| --- | --- | --- | --- |
| True Label | Classification | Affected Score | Control Score |
| Affected | Affected | 0.95951 | 0.040494 |
| Affected | Affected | 0.97798 | 0.022023 |
| Affected | Affected | 0.97973 | 0.020274 |
| Affected | Affected | 1 | 4.8532e-07 |
| Affected | Affected | 0.80494 | 0.19506 |
| Affected | Affected | 0.96783 | 0.032171 |
| Affected | Control | 0.27432 | 0.72568 |
| Affected | Affected | 0.96542 | 0.034583 |
| Affected | Affected | 0.76618 | 0.23382 |
| Affected | Affected | 0.84741 | 0.15259 |
| Affected | Affected | 0.98453 | 0.015259 |
| Affected | Affected | 0.99717 | 0.0028332 |
| Affected | Affected | 0.98542 | 0.014581 |
| Affected | Affected | 0.82044 | 0.17956 |
| Control | Affected | 0.88552 | 0.11448 |
| Control | Control | 0.14675 | 0.85325 |
| Control | Affected | 0.77376 | 0.22624 |
| Control | Affected | 0.93469 | 0.065309 |
| Control | Affected | 0.99635 | 0.0036514 |

Table - Classifications and scores for each of the 19 aligned central slices classified by a retraining VGG19 CNN.

# References

|  |  |
| --- | --- |
| [1] | M. J. Pead and S. Guthrie, “Elbow dysplasia in dogs - a new scheme,” British Veterinary Association (BVA), London, 2011. |
| [2] | A & C Black, Black's Veterinary Dictionary, London: A & C Black, 2014. |
| [3] | J. Temwichitr, A. P. A. Leegwater and A. H. A. Hazewinkel, “Fragmented coronoid process in the dog: A heritable disease,” *The Veterinary Journal,* vol. 185, no. 2, pp. 123-129, 2010. |
| [4] | J. Michelsen, “Canine elbow dysplasia: Aetiopathogenesis and current treatment recommendations,” *The Veterinary Journal,* vol. 196, no. 1, pp. 12-19, 2013. |
| [5] | T. Gemmill and D. Clements, “Fragmented coronoid process in the dog: is there a role for incongruency?,” *Journal of Small Animal Practice,* no. 48, pp. 361-368, 2007. |
| [6] | N. M. C. C. Fitzpatrick and R. M. V. Yeadon, “Working Algorithm for Treatment Decision Making for Developmental Disease of the Medial Compartment of the Elbow in Dogs,” *Veterinary Surgery,* vol. 38, no. 2, pp. 285-300, 2009. |
| [7] | J. D. Demko and R. D. D. McLaughlin, “Developmental Orthopedic Disease,” *Veterinary Clinics of North America: Small Animal Practice,* vol. 35, no. 5, pp. 1111-1135, 2005. |
| [8] | E. L. Hilton and L. J. Henderson, “Neurosurgical considerations in posttraumatic syringomyelia,” *AORN Journal,* vol. 77, no. 1, pp. 135-150, 2003. |
| [9] | “Use of Morphometric Mapping to Characterise Symptomatic Chiari-Like Malformation, Secondary Syringomyelia and Associated Brachycephaly in the Cavalier King Charles Spaniel,” *PLoS One,* vol. 12, no. 1, 2017. |
| [10] | C. A. Loughin, “Chiari-like Malformation,” *Veterinary Clinics of North America: Small Animal Practice,* vol. 46, no. 2, pp. 231-242, 2016. |
| [11] | C. Rusbridge, “Chiari–like malformation and syringomyelia,” *European Journal of Companion Animal Practice,* vol. 23, no. 3, p. 70, 2013. |
| [12] | M. D. Shen, “Wikimedia Commons,” 13 December 2018. [Online]. Available: https://commons.wikimedia.org/wiki/File:CSF\_circulation.png. [Accessed 20 April 2020]. |
| [13] | C. Rusbridge, F. Stringer and S. P. Knowler, “Clinical Application of Diagnostic Imaging of Chiari-Like Malformation and Syringomyelia,” *Front Vet Sci,* vol. 5, no. 280, 2018. |
| [14] | American Veterinary Medical Association, “One Health: A New Professional Imperative,” American Veterinary Medical Association, 2008. |
| [15] | E. Chandler, Ettinger Feldman Textbook of Veterinary Internal Medicine 4th Edn, London: W. B. Saunders, 1996. |
| [16] | A. C. Hechler and S. A. Moore, “Understanding and Treating Chiari-like Malformation and Syringomyelia in Dogs,” *Topics in Companion Animal Medicine,* vol. 33, no. 1, pp. 1-11, 2018. |
| [17] | P. Salunke, M. Karthigeyan and P. Malik, “Foramen magnum decompression without bone removal: C1–C2 posterior fixation for Chiari with congenital atlantoaxial dislocation/basilar invagination,” Department of Neurosurgery, Postgraduate Institute of Medical Education & Research (PGIMER), Chandigarh, 2019. |
| [18] | K. Jun and S. Yoon, “Alignment Solution for CT Image Reconstruction using Fixed Point and Virtual Rotation Axis,” *Sci Rep,* vol. 7, no. 41218, 2017. |
| [19] | J. Broder, Diagnostic Imaging for the Emergency Physician, Philadelphia: Saunders, 2011. |
| [20] | D. W. McRobbie, E. A. Moore, M. J. Graves and M. R. Prince, MRI from Picture to Proton, Cambridge: Cambridge University Press, 2006, pp. 33-34. |
| [21] | P. Lemay, S. P. Knowler, S. Bouasker, Y. Nédélec, S. Platt, C. Freeman, G. Child, L. B. Barreiro, R. G. A., C. Rusbridge and Z. Kibar, “Quantitative Trait Loci (QTL) Study Identifies Novel Genomic Regions Associated to Chiari-Like Malformation in Griffon Bruxellois Dogs,” *PLoS One,* vol. 9, no. 4, 2014. |
| [22] | T. J. Gemmill, D. J. Mellor, D. N. Clements, S. P. Clarke, M. Farrell, D. Bennett and S. Carmichael, “Evaluation of elbow incongruencyusing reconstructed CT in dogssuffering fragmented coronoid process,” *Journal of Small Animal Practice,* vol. 46, pp. 327-333, 2005. |
| [23] | P. Proks, L. Stehlík, K. Irova, M. Dvořák, R. Srnec and A. Nečas, “Relationship between Radioulnar Incongruity of Elbow Joints and the Type ofFragmented Processus Coronoideus Medialis,” *Acta Veterinaria Brno,* vol. 79, pp. 307-312, 2010. |
| [24] | H. Eljack, H. Werner and P. Bottcher, “Sensitivity and Specificity of 3D Models of the Radioulnar Joint Cup in Combination With a Sphere Fitted to the Ulnar Trochlear North for Estimation of Radioulnar Incongruence In Vitro,” *Veterinary Surgery,* vol. 42, pp. 365-370, 2013. |
| [25] | J. J. Sackman, D. J. Marino, C. A. Loughin, C. W. Dewey, L. J. Marino, M. L. Lesser and M. B. Akerman, “Morphometric features of the craniocervical junction region in dogs with suspected Chiari-like malformation determined by combined use of magnetic resonance imaging and computed tomography,” *American Journal of Veterinary Research,* vol. 73, no. 1, pp. 105-111, 2012. |
| [26] | S. Cerda‐Gonzalez, N. J. Olby and E. H. Griffith, “Dorsal Compressive Atlantoaxial Bands and the Craniocervical Junction Syndrome: Association with Clinical Signs and Syringomyelia in Mature Cavalier King Charles Spaniels,” *Journal of Veterinary Internal Medicine,* vol. 29, no. 3, pp. 887-892, 2015. |
| [27] | S. Cerda‐Gonzalez, N. J. Olby and E. H. Griffith, “Medullary Position at the Craniocervical Junction in Mature Cavalier King Charles Spaniels: Relationship with Neurologic Signs and Syringomyelia,” *Journal of Veterinary Internal Medicine,* vol. 29, no. 3, p. 882–886, 2015. |
| [28] | S. Micheal, S. P. Knowler, C. Rusbridge and K. Wells, “Using machine learning to understand neuromorphological change and image‐based biomarker identification in Cavalier King Charles Spaniels with Chiari‐like malformation‐associated pain and syringomyelia,” *Journal of Veterinary Internal Medicine,* vol. 33, no. 6, pp. 2665-2674, 2019. |
| [29] | A. Urbizu, B. A. Martin and D. Moncho, “Machine learning applied to neuroimaging for diagnosis of adult classic Chiari malformation: role of the basion as a key morphometric indicator,” *Journal of Neurosurgery,* vol. 129, no. 3, pp. 567-851, 2018. |
| [30] | S. Sarraf and G. Tofighi, “Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data,” in *2016 Future Technologies Conference (FTC)*, San Francisco, 2016. |
| [31] | C. Salvatore, A. Cerasa, I. Castiglioni, F. Gallivanone, A. Augimeri, M. M. López, G. Arabia, M. Morelli, M. C. Gilardi and A. Quattrone, “Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy,” *Journal of Neuroscience Methods,* vol. 222, no. 1, pp. 230-237, 2014. |
| [32] | N. K. Focke, G. Helms, S. Scheewe, P. M. Pantel, C. G. Bachmann, J. Ebentheuer, P. Dechent, A. Mohr, W. Paulus and C. Trenkwalder, “Individual voxel‐based subtype prediction can differentiate progressive supranuclear palsy from idiopathic parkinson syndrome and healthy controls,” *Human Brain Mapping,* vol. 32, no. 11, pp. 1905-1915, 2011. |
| [33] | C. Dezateux, “Developmental dysplasia of the hip,” *The Lancet,* vol. 369, no. 9572, pp. 1541-1552, 2007. |
| [34] | Y. Xue, R. Zhang, Y. Deng, K. Chen and T. Jiang, “A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis,” *PLoS One,* 2017. |
| [35] | Y. Y. C. S. H. N. Tomitaa, “Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans,” *Computers in Biology and Medicine,* vol. 98, no. 1, pp. 8-15, 2018. |
| [36] | C. Yang, A. Rangarajan and S. Ranka, “Visual Explanations From Deep 3D Convolutional Neural Networks for Alzheimer's Disease Classification,” Dept. of Computer & Information Science & Engineering, University of Florida, Gainsville, 2018. |
| [37] | S. Alves-Pimenta, M. M. Ginja, A. M. Fernandes, A. J. Ferreira, P. Melo-Pinto and B. Colaco, “Computed Tomography and Radiographic Assessment of Congruity Between the Ulnar Trochlear Notch and Humeral Trochle in Large Breed Dogs,” *Veterinary and Comparative Orthopaedics and Traumatology,* vol. 1, no. 30, pp. 8-14, 2017. |
| [38] | C. O. a. M. T. A. Mordvintsev, “Google AI Blog,” Google, 15 June 2015. [Online]. Available: https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html. [Accessed 2 January 2020]. |
| [39] | A. S. Lundervold and A. Lundervold, “An overview of deep learning in medical imaging focusing on MRI,” *Zeitschrift für Medizinische Physik,* vol. 29, no. 2, pp. 102-127, 2019. |
| [40] | K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-scale Image Recognition,” in *ICLR*, Ithaca, 2015. |
| [41] | ImageNet, “ImageNet,” [Online]. Available: http://www.image-net.org/. [Accessed 12 May 2020]. |
| [42] | J. M. M. Pérez and J. Pascau, Image processing with ImageJ : discover the incredible possibilities of ImageJ, from basic image procesing to macro and plugin development, Birmingham: Packt Publishing, 2013. |
| [43] | U. Gille, “Wikimedia Commons,” 13 June 2006. [Online]. Available: https://commons.wikimedia.org/wiki/File:Elbow-Dysplasia-Manifestations.svg. [Accessed 3 December 2019]. |
| [44] | S. M., “CT Scan,” in *Encyclopedia of Cancer*, Springer, Berlin, Heidelberg, 2011, p. 78. |
| [45] | T. Toyaski, “Wikicommons,” 2 September 2012. [Online]. Available: https://commons.wikimedia.org/wiki/File:CT\_PRINCI\_PB.jpg. [Accessed 3 December 2019]. |
| [46] | I. Kononeko, “Machine learning for medical diagnosis: history, state of the art and perspective,” *Artificial Intelligence in Medicine,* vol. 23, no. 1, pp. 89-109, 2001. |
| [47] | D. Maturana and S. Scherer, “VoxNet: A 3D Convolutional Neural Network for real-time object recognition,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Hamburg, IEEE, 2015, pp. 922-928. |
| [48] | E. Linsey, “Wikimedia Commons,” 19 February 2007. [Online]. Available: https://commons.wikimedia.org/wiki/File:Syringomyelia.jpg. [Accessed 20 April 2020]. |
| [49] | BodyParts3D, “Wikimedia Commons,” 21 July 2019. [Online]. Available: https://commons.wikimedia.org/wiki/File:Posterior\_cranial\_fossa\_boundaries.svg. [Accessed 20 April 2020]. |
| [50] | Wikipedia, “Wikipedia,” 10 November 2017. [Online]. Available: https://en.wikipedia.org/wiki/File:Mri\_scanner\_schematic\_labelled.svg. [Accessed 21 April 2020]. |
| [51] | H. V. Carter and H. Gray, Anatomy of the Human Body, Philadelphia: Lea and Febiger, 1918. |
| [52] | OpenStax, “Wikimedia Commons,” 18 May 2016. [Online]. Available: https://commons.wikimedia.org/wiki/File:1317\_CFS\_Circulation.jpg. [Accessed 22 April 2020]. |