Final Report

A study into the feasibility of using machine learning techniques to process MRI scans of Cavalier king charles spaniels skulls and detect signs of chiari-like malformation.

Robert Clark

2020

# Abstract

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.

The aim of this dissertation project was to investigate the potential of artificial intelligence, specifically deep learning, could be used to better understand chiari-like malformation. 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](#_Toc39449807)

[List of Figures 3](#_Toc39449808)

[List of Tables 3](#_Toc39449809)

[Chapter 1: Introduction 4](#_Toc39449810)

[1.1. Canine Chiari-Like Malformation 4](#_Toc39449811)

[1.2. Current Approaches to Diagnosis and Treatment of Canine Chiari-Like Malformation 4](#_Toc39449812)

[1.3. Magnetic Resonance Imagery (MRI) 5](#_Toc39449813)

[1.4. Aims and Objectives 6](#_Toc39449814)

[Chapter 2: Literature Review 7](#_Toc39449815)

[2.1 Prior Analysis Work on Canine Chiari-Like Malformation 7](#_Toc39449816)

[2.2. Machine Learning as a Diagnosis Aid for CLM 8](#_Toc39449817)

[2.3. Machine Learning as a Diagnosis Aid for Other Neurological Conditions 8](#_Toc39449818)

[2.4. Summary 8](#_Toc39449819)

[Chapter 3: Methodology 10](#_Toc39449820)

[3.1. Data Description 10](#_Toc39449821)

[3.2. Data Processing 10](#_Toc39449822)

[3.2.1. Transfer Learning with a Convolutional Neural Network (CNN) 10](#_Toc39449823)

[3.2.2. Affine Transformations as Feature 11](#_Toc39449824)

[3.3. Summary 12](#_Toc39449825)

[Chapter 4: Results and Discussion 13](#_Toc39449826)

[4.1. Transfer Learning 13](#_Toc39449827)

[4.2. Affine Transformation as a Feature 17](#_Toc39449828)

[Chapter 5: Conclusions and Future Work 18](#_Toc39449829)

[Appendix A: Project Management 19](#_Toc39449830)

[Appendix B: Unabridged Results of Classification through Transfer Learning 20](#_Toc39449831)

[References 21](#_Toc39449832)

# List of Figures

[Figure 1 - A cyst forming within a cervical spine, characteristic of Syringomyelia. [42] 5](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449833)

[Figure 2 - The flow of cerebrospinal fluid around the brain, with the perivascular space shown within the inset red box. [5] Licensed under Creative Commons Attribution 4.0 International. 5](#_Toc39449834)

[Figure 3 - 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. 6](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449835)

[Figure 4 - The posterior fossa of the human skull, shown in green. Licensed under the Creative Commons Attribution-Share Alike 2.1 Japan. [43] 6](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449836)

[Figure 5 - A cross section of an MRI scanner with labelled components. [44] Licensed under the Creative Commons Attribution-ShareAlike 3.0. 7](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449837)

[Figure 6 - The Atlantooccipital joint and atlantoaxial ligament, located at the base of the skull. [45] 8](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449838)

[Figure 7 - 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. 8](#_Toc39449839)

[Figure 8 - A labelled cross section of a brain with ventricles exposed. [46] Licensed under the Creative Commons Attribution 4.0 International. 9](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449840)

[Figure 9 - Distribution of intensity levels over the dataset consisting of central slices from the MRI scans of nineteen CKCS. 12](#_Toc39449841)

[Figure 10 - 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. 12](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449842)

[Figure 11 - A binary mask created by thresholding a central slice of an MRI scan of CKCS. 12](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449843)

[Figure 12 - 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. 13](file:///D:\Coursework\Final-Year-Project-2\Reports\Final%20Report.docx#_Toc39449844)

[Figure 13 - A block image describing the methodology of this project. 13](#_Toc39449845)

[Figure 14 - The training progress for the first iteration through the leave-one-out cross-validation process. 14](#_Toc39449846)

[Figure 15 - Results for iterations 2 through 19, with the same colour scheme as Figure 11 and an overall learning rate of 0.0003. 16](#_Toc39449847)

[Figure 16 - Class Activation Map for a patient in the control group, which was correctly cleared, indicating an area of interest near the Parietal lobe. 17](#_Toc39449848)

[Figure 17 - A class activation map for a member of the control group, which was falsely diagnosed, indicating an area of interest near the fourth ventricle. 17](#_Toc39449849)

[Figure 18 - A class activation map for a member of the affected group, indicating no correlation between CLM and the shape of the skull or brain. 17](#_Toc39449850)

# List of Tables

[Table 1 - Summary of patient data within dataset 11](#_Toc39449851)

[Table 2 - Confusion table for classification of unaligned central slices 14](#_Toc39449852)

[Table 3 - Confusion table for classification of aligned central slices 14](#_Toc39449853)

[Table 4 -Results of classifying dataset using SVM based off transforms required to map images onto average of controls, after Principal Component Analysis and Bayesian Optimisation. 18](#_Toc39449854)

[Table 5 - Gantt chart for the entirety of the project, updated to reflect the change in aim for the project. Deadlines are indicated with a "!". 20](#_Toc39449855)

[Table 6 - Classifications and scores for each of the 19 unaligned central slices classified by a retrained VGG19 CNN. 21](#_Toc39449856)

[Table 7 - Classifications and scores for each of the 19 aligned central slices classified by a retraining VGG19 CNN. 21](#_Toc39449857)

# Chapter 1: Introduction

## 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 [1]. 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 [2] and is believed to be present in up to 95% of the world wide population of Cavalier King Charles Spaniels (CKCS) [3].

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 [4]. 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 2) 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 1.

Figure - A cyst forming within a cervical spine, characteristic of Syringomyelia. [42]

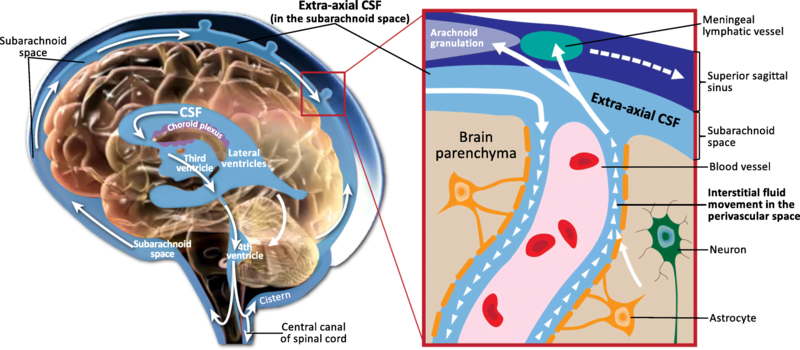


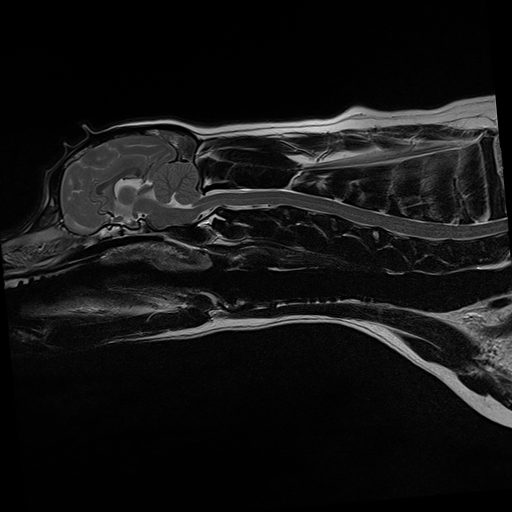
Figure - The flow of cerebrospinal fluid around the brain, with the perivascular space shown within the inset red box. [5] Licensed under Creative Commons Attribution 4.0 International.

In addition to Cavalier King Charles Spaniels, both Syringomyelia and Chiari-Like Malformation are known to present in humans. [6] Canine Chiari-Like Malformation is sufficiently analogous to its human counterpart 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.

## Current Approaches to Diagnosis and Treatment of Canine 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. [6]

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. [6] 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 3.

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.

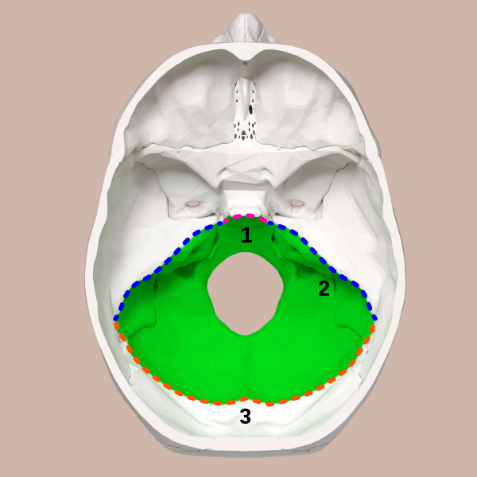
However, Chiari-like Malformation occurs independently of Syringomyelia and so syrinxes may not form until a considerable amount of time after the former condition develops. Diagnosis of Chiari-like Malformation alone is hence often done via the ruling out of other conditions and study of anatomical 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. [43]

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. [7]

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 4), the region of the skull which houses the brain stem and cerebellum, to alleviate pressure on the craniocervical junction. [8] 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.

## 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 3. To obtain these images, the subjects were athetized 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 radio waves were then fired at the subjects in A picture containing device

Description automatically generatedorder to “knock” the water molecules out of alignment. As they realign, they emit radio 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.

Figure - A cross section of an MRI scanner with labelled components. [44] Licensed under the Creative Commons Attribution-ShareAlike 3.0.

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. [9]

## Aims and Objectives

As explained in Section 1.2, 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 [10], 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.

# Chapter 2: Literature Review

## 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 6) so that the atlas bone is partly within the cranium has a stronger incidence rate of CLM than other causes [11], 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 the that rather than any particular cause.

Figure - The Atlantooccipital joint and atlantoaxial ligament, located at the base of the skull. [45]

Alternative studies have instead linked CLM developing into Syringomyelia to lesions on the atlantoaxial ligament known as “bands” [12] or “medullary kinking” of the craniocervical junction [13] where the spinal cord meets the brain stem at a non-continuous angle as illustrated in Figure 7.

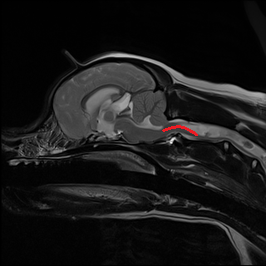
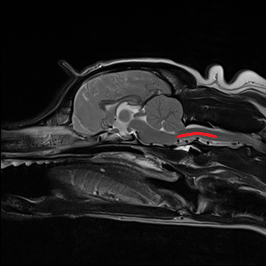


Figure - A healthy CKCS (left) with a continuous craniocervical junction and a CLM/SM affected CKCS (right) with a noticeably elevated caudal medulla oblongata. Both junctions are highlighted with a red line.

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

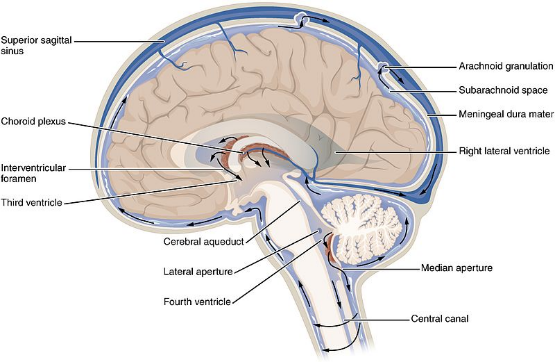
Some investigation into machine learnings use diagnosing chiari-like malformation has already been performed. The floor of the third ventricle (shown in Figure 8) and a region in the sphenoid bone as potential biomarkers for CLM by locating morphological differences between MRI scans of affected and healthy dogs. [14] 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. [46] 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 [15]. 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 Neurological Conditions

In addition to research into CLM, machine learning has also proved to be invaluable in the study of other neurological 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% [16]. 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 [17], a significant improvement over the 39.53% obtained by a similar study performed without dimensionality reduction via PCA [18]. 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.

## Summary

Though research into this area 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: Methodology

## 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 1.

|  |  |  |
| --- | --- | --- |
| 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 to act as a binary classifier through replacement of the final Fully Connected layer with one of a significantly higher learn rate and the replacement of the final Classification layer with one with only two outputs.

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. Due to its versatility in regard to transfer learning [19], VGG19 was chosen for this investigation and the epoch number and batch size selected through trial and error.

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.

These 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 (Figure 9), 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.

A picture containing comb

Description automatically generated

Figure - 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 generatedOnce 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).

### Affine Transformations as Feature

A close up of a logo

Description automatically generatedIn 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 10. A binary mask, shown in Figure 11, 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 10, producing a 4x4 transformation matrix which acted as a 16-dimensional feature.

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.

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

This 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 A close up of a map

Description automatically generatedthe 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.

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

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.

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 12.

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

Chapter 4: Results and Discussion

## Transfer Learning

While the accuracy of the classification performed by the retrained neural network was above the 0.5 value you would expect of purely random classification, the confusion tables shown in Table 2 and Table 3 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.

|  |  |  |
| --- | --- | --- |
|  | Affected | Control |
| Affected | 0.786 | 0.6 |
| Control | 0.214 | 0.4 |

Table - Confusion table for classification of unaligned central slices

|  |  |  |
| --- | --- | --- |
|  | Affected | Control |
| Affected | 0.928 | 0.8 |
| Control | 0.071 | 0.2 |

Table - Confusion table for classification of aligned central slices

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.

This theory is supported by the values provided for accuracy and loss during the training phase as shown in Figure 11 and Figure 12, where the area under the curve is typically lower than one would desire. There is rarely a consistent increase in accuracy over each epoch, with loss similarly being unpredictable and often remaining high.

A close up of a map

Description automatically generated

Figure - The training progress for the first iteration through 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 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 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 2 through 19, with the same colour scheme as Figure 11 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.

Further evidence for is provided by the CAMs, which show little consistent trend for areas of interest between patients and do not indicate a similar conclusion to that reached by prior studies [15] [2] [14]. Their 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.

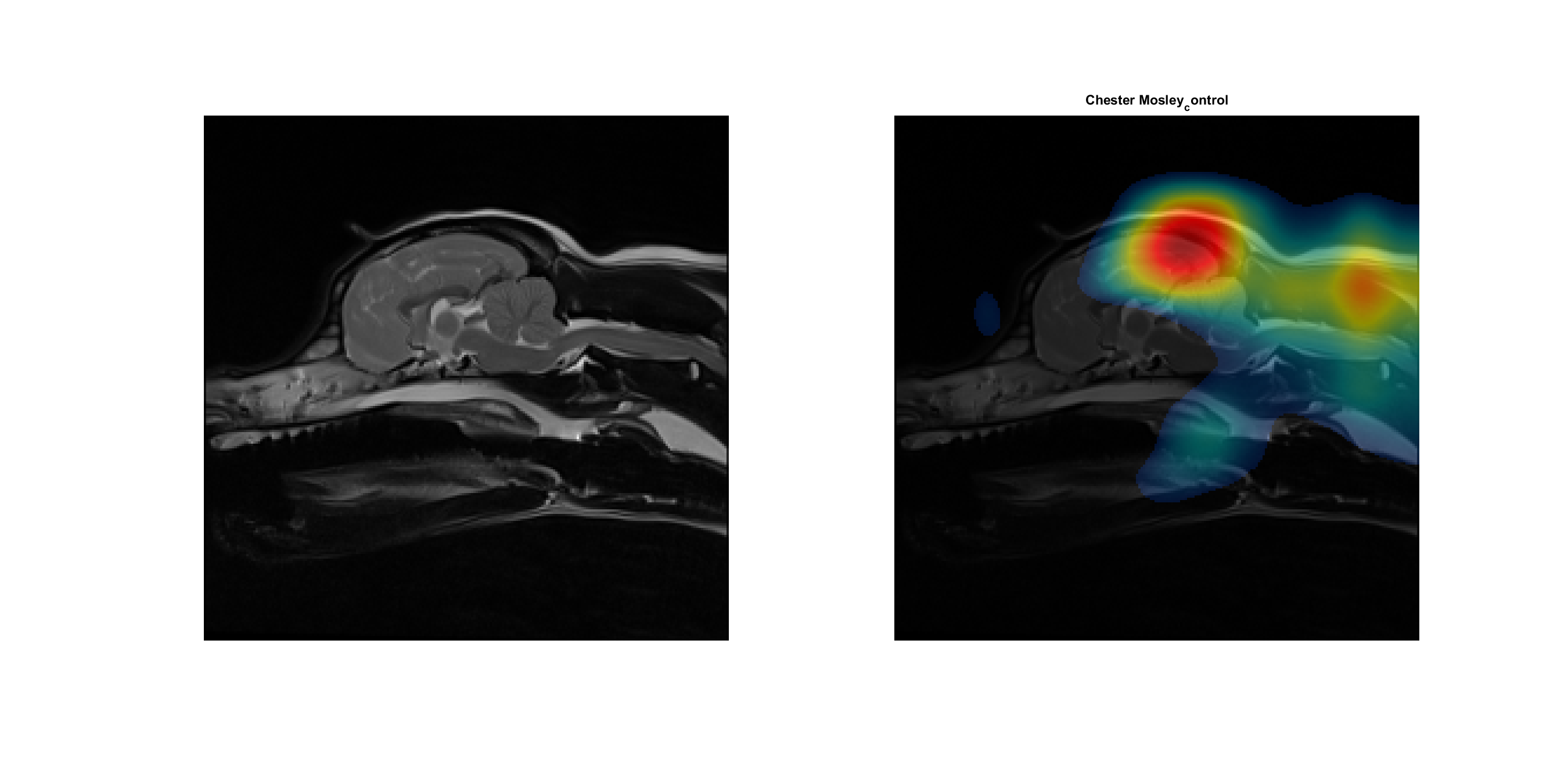


Figure - Class Activation Map for a patient in the control group, which was correctly cleared, indicating an area of interest near the Parietal lobe.

A picture containing sitting, colorful, food

Description automatically generated

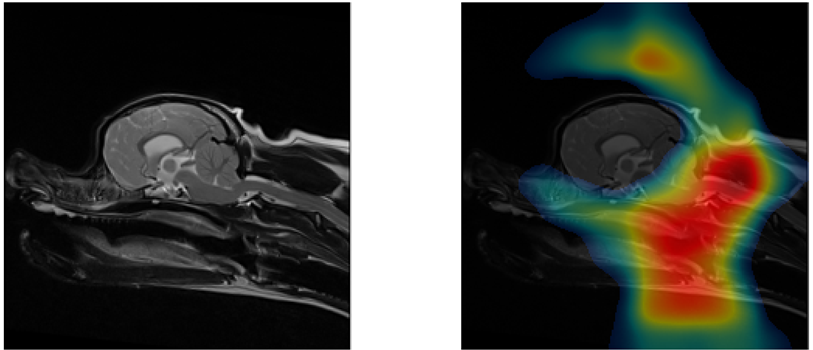
Figure - A class activation map for a member of the control group, which was falsely diagnosed, indicating an area of interest near the fourth ventricle.

Figure - A class activation map for a member of the affected group, indicating no correlation between CLM and the shape of the skull or brain.

## 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.

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] | E. L. Hilton and L. J. Henderson, “Neurosurgical considerations in posttraumatic syringomyelia,” *AORN Journal,* vol. 77, no. 1, pp. 135-150, 2003. |
| [2] | “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. |
| [3] | C. A. Loughin, “Chiari-like Malformation,” *Veterinary Clinics of North America: Small Animal Practice,* vol. 46, no. 2, pp. 231-242, 2016. |
| [4] | C. Rusbridge, “Chiari–like malformation and syringomyelia,” *European Journal of Companion Animal Practice,* vol. 23, no. 3, p. 70, 2013. |
| [5] | M. D. Shen, “Wikimedia Commons,” 13 December 2018. [Online]. Available: https://commons.wikimedia.org/wiki/File:CSF\_circulation.png. [Accessed 20 April 2020]. |
| [6] | 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. |
| [7] | 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. |
| [8] | 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. |
| [9] | 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. |
| [10] | 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. |
| [11] | 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. |
| [12] | 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. |
| [13] | 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. |
| [14] | 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. |
| [15] | 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. |
| [16] | 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. |
| [17] | 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. |
| [18] | 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. |
| [19] | 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. |
| [20] | M. J. Pead and S. Guthrie, “Elbow dysplasia in dogs - a new scheme,” British Veterinary Association (BVA), London, 2011. |
| [21] | A & C Black, Black's Veterinary Dictionary, London: A & C Black, 2014. |
| [22] | 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. |
| [23] | J. Michelsen, “Canine elbow dysplasia: Aetiopathogenesis and current treatment recommendations,” *The Veterinary Journal,* vol. 196, no. 1, pp. 12-19, 2013. |
| [24] | 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. |
| [25] | 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. |
| [26] | 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. |
| [27] | E. Chandler, Ettinger Feldman Textbook of Veterinary Internal Medicine 4th Edn, London: W. B. Saunders, 1996. |
| [28] | 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. |
| [29] | 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. |
| [30] | 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. |
| [31] | 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. |
| [32] | 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. |
| [33] | U. Gille, “Wikimedia Commons,” 13 June 2006. [Online]. Available: https://commons.wikimedia.org/wiki/File:Elbow-Dysplasia-Manifestations.svg. [Accessed 3 December 2019]. |
| [34] | S. M., “CT Scan,” in *Encyclopedia of Cancer*, Springer, Berlin, Heidelberg, 2011, p. 78. |
| [35] | T. Toyaski, “Wikicommons,” 2 September 2012. [Online]. Available: https://commons.wikimedia.org/wiki/File:CT\_PRINCI\_PB.jpg. [Accessed 3 December 2019]. |
| [36] | 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. |
| [37] | 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. |
| [38] | C. Dezateux, “Developmental dysplasia of the hip,” *The Lancet,* vol. 369, no. 9572, pp. 1541-1552, 2007. |
| [39] | 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. |
| [40] | 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. |
| [41] | J. Broder, Diagnostic Imaging for the Emergency Physician, Philadelphia: Saunders, 2011. |
| [42] | E. Linsey, “Wikimedia Commons,” 19 February 2007. [Online]. Available: https://commons.wikimedia.org/wiki/File:Syringomyelia.jpg. [Accessed 20 April 2020]. |
| [43] | BodyParts3D, “Wikimedia Commons,” 21 July 2019. [Online]. Available: https://commons.wikimedia.org/wiki/File:Posterior\_cranial\_fossa\_boundaries.svg. [Accessed 20 April 2020]. |
| [44] | Wikipedia, “Wikipedia,” 10 November 2017. [Online]. Available: https://en.wikipedia.org/wiki/File:Mri\_scanner\_schematic\_labelled.svg. [Accessed 21 April 2020]. |
| [45] | H. V. Carter and H. Gray, Anatomy of the Human Body, Philadelphia: Lea and Febiger, 1918. |
| [46] | OpenStax, “Wikimedia Commons,” 18 May 2016. [Online]. Available: https://commons.wikimedia.org/wiki/File:1317\_CFS\_Circulation.jpg. [Accessed 22 April 2020]. |