final report

Project code: P.PIP.0765

Prepared by: Scott Nicholson, Jonathan Cook, Merv Shirazi
Scott Automation and Robotics

Date published: 29 March 2019

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

Investigating Neural Network algorithms for imaging points of interest identification

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Executive Summary

The aim of this project was to construct training data and create a neural network to accurately calculate beef rib scribing cut locations from X-ray images. Previously, algorithms were developed to calculate these cuts without the use of neural networks. However, these algorithms were prone to robustness issues especially with the deterioration of the detectors and X-ray image quality over time.

To achieve accurate and robust results from a supervised neural network large amounts of labelled data is needed. Fortunately, the system had been operational for a number of years before the upgrade and thus, a large amount of X-ray images existed of the same product in the same configuration of the system. The labelling process is a tedious, manual task that must be completed to provide the training set with a ground truth to compare its performance against. Due to this, the right trade-off between minimal time spent labelling and the performance of the network must be considered.

Overall, the neural networks greatly improved the accuracy of determining the side orientation, locating the rib 1 junction and determining the rib 2 and 8 intersections between the rib 1 junction and the Aitch bone.

Table of contents

1 Milestone description

The aim of project was to implement the neural network software platform on a beef processing site. The system should be commissioned and returned to production with this newly modified software and returned to full production.

2 Project objectives

The project aimed to achieve the following objectives:

1. Obtain and setup suitable hardware, and create neural network training and verification datasets
2. Develop and train a neural network to identify the rib 1 junction
3. Report on accuracy of the neural network result
4. Implement the neural network algorithms into the existing scribe system and evaluate the improvement in accuracy achieved in a production environment

3 Methodology

3.1 Background

Neural networks and machine learning have existed for a long time however, in recent years there has been a rapid increase in usage across a multitude of industries and applications although primarily in image processing. This is due to increasing computational power along with falling costs; standard desktop computers can now be used to train and develop complex models without the use of supercomputers.

Many types of machine learning and artificial algorithms exist, these can be categorised into unsupervised networks or supervised networks based on the way the models learn. Supervised learning is the most common and requires labels (where the desired output is known and identified) so that the network can evaluate its performance throughout the training. Unsupervised learning does not require any labels and instead can identify groups or trends with a dataset. The networks discussed in this report are all supervised however even they can be diverse. For example they can perform classification (e.g. left or right) or linear regression (e.g. object is located at a specific coordinate).

If applied correctly, these machine learning tools have the potential to offer significant gain to the red meat industry. Benefits include, but are not limited to, more accurate and more robust cutting with the possibility to perform cuts not previously possible.

Specifically relating to this project, there was a significant challenge to identify the rib 1 junction since this point lies in one of the thickest parts of a carcass thus making it difficult to obtain clear X-ray images for large cattle. Additionally, the junction itself varies greatly in shape, size and location along with clarity due to varying levels of ossification in the joint. The previous algorithms used other clearer features to fixture the location from however now with machine learning the rib and therefore junction can be identified directly.

3.2 Neural Network Architectures

This project implemented many separate neural networks with two main architecture types. Firstly, a decreased resolution X-ray image is passed into a left or right classifier to determine the orientation of the side on the conveyor. If the image contained a right side a mirror was performed so it appears as a left. This reduces variability in the networks that follow which helps greatly as it one less feature they have to learn.

To determine the location of the rib 1 junction the image is fed into a segmentation network to segment the whole of rib 1.

The final network uses a similar architecture to the rib 1 segmentation but instead segments all the ribs from 2 to 10. The result of this image is used to determine the start and end points for the rib scribing cut.

3.2.1 Left or Right Classification

The left or right classifier utilises a convolutional neural network (CNN). These networks contain many layers with three main operations. Firstly, the convolution layer uses a filter matrix that slides over the previous layer's output to produce a feature map.

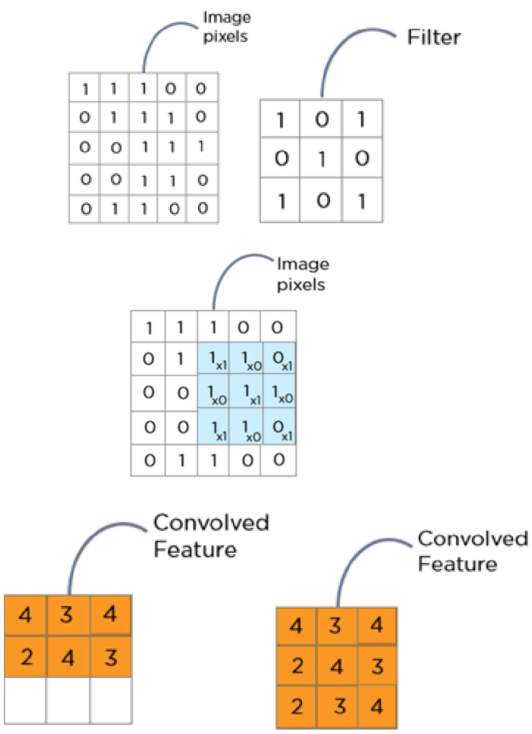

Figure 1 Convolution Example

Secondly, a pooling layer reduces the dimensionality of the feature map. This allows the network to learn higher level features as it gets deeper. Typically, more features are learnt the deeper the network goes which can be seen in Figure 3 with the number of feature maps increasing while the size of each feature map decreases through the pooling layer. The convolutional layer generates and learns these feature maps which is followed by a pooling layer which feeds into another convolutional layer, etcetera.

The final stage of the network is known as fully connected layers. This is where all the pixels from the previous layer are flattened into one dimension. Then each element is fully connected to each element in the next layer with the network learning the weights of each connection to produce a final result which is a probability shown in Figure 2

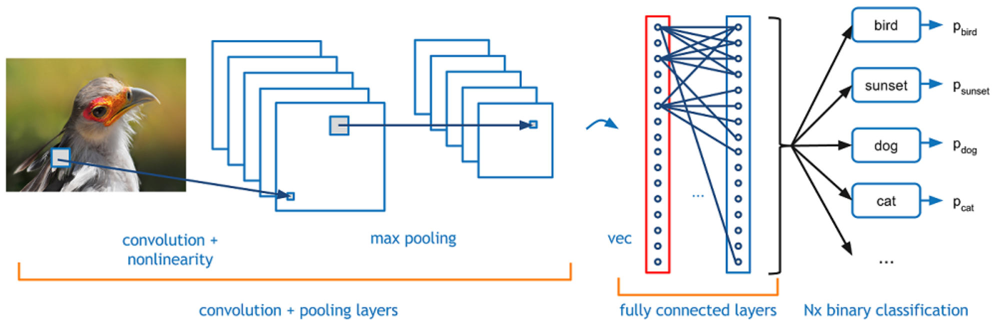

Figure 2 Convolutional Neural Network Example

The initial layers of the network will learn features such as edges, dots and curves. Then it may progress to learning higher level features such as legs, ribs and neck. The advantage of a convolutional neural network is that the features to learn do not need to be pre-defined. After each training iteration the left or right prediction is compared to the manually labelled orientation and then the network learns from any wrong predictions and modifies itself.

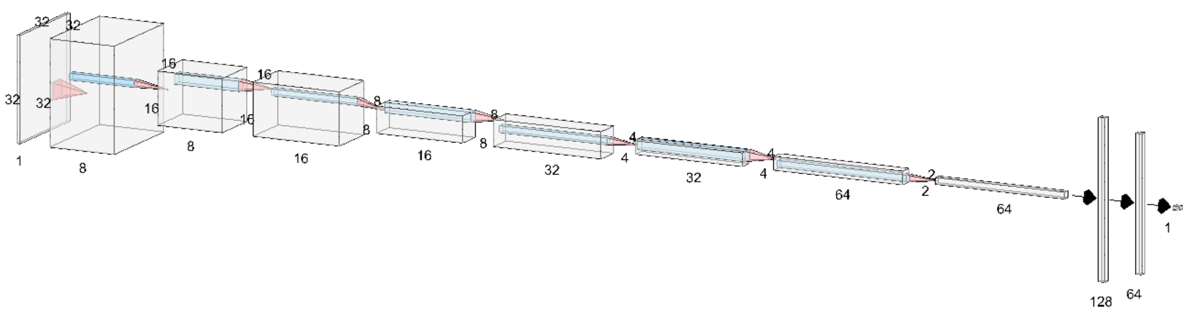

Figure 3 Convolutional Neural Network Architecture

3.2.2 Rib Segmentation

All of the rib segmentation networks are based off of the same architecture, the main differences are the input and output image sizes. These networks follow a U-Net architecture, given its name from the architecture diagram looking similar to the letter U seen in Figure 4. Firstly the image is reduced in a similar manner to that of a convolutional neural network using convolutions and pooling to learn high level features. The second half of the network essentially performs the opposite to that of a CNN. It takes lots of very small feature maps and increases their resolution while learning how to reconstruct the image so it is similar to the labelled output. These networks are often used to reduce noise in images but can also be used for segmentation.

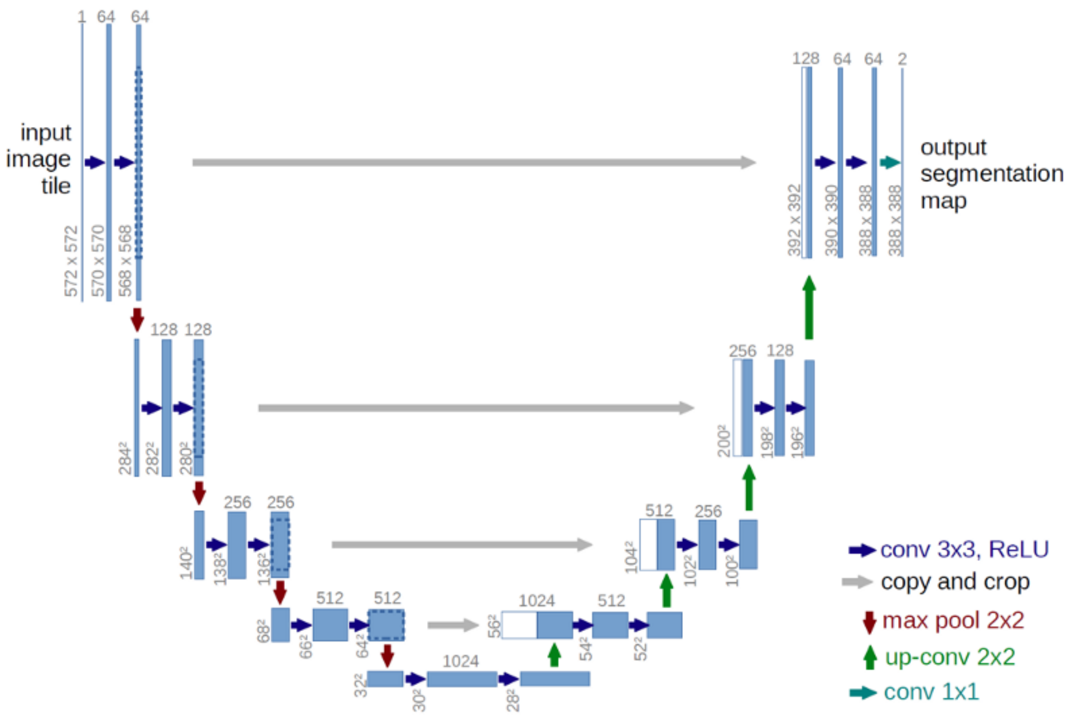

Figure 4 U-Net Architecture

A loss function in neural networks compares the predicted output and the labelled ground truth and then modifies itself slightly depending on the result.
Image Augmentation
Due to the manual labelling process along with hardware memory restrictions it is necessary to augment the images so the network is more robust and does not overfit.

4 Success in meeting the milestone

Constructing training data, creating neural networks and integration into the software platform constitute the three main components of this milestone and were completed in that order.

4.1 Training Data Construction

In order to train a supervised neural network, training labels or ground-truth images need to be provided so that the network can learn by comparing its result to the label. A small labelling tool was developed to aid the manual process of segmenting each rib. This tool provided six control points to manipulate two splines which outline the rib of interest.

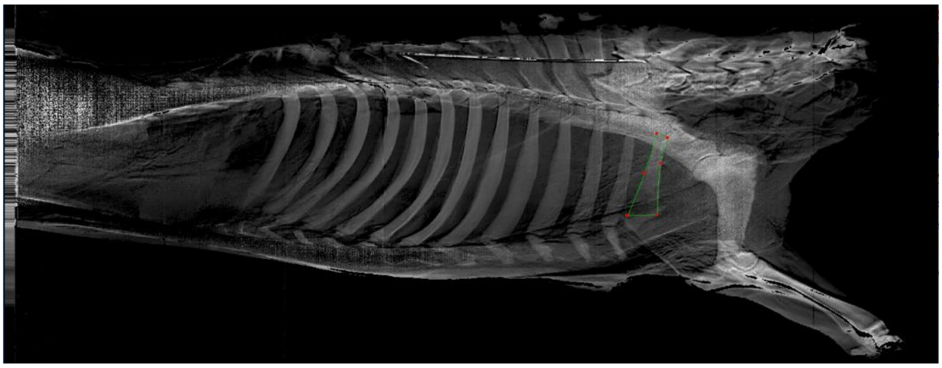

Figure 5 Manual Labelling Control Points

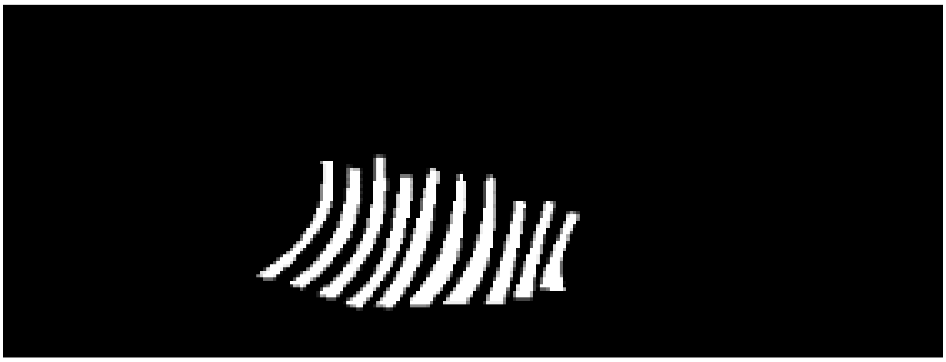

Figure 6 Labelled Ground Truth Image

4.2 Neural Network Creation

As previously discussed, there are three neural networks implemented in the software upgrade. These were all developed separately to each other.

4.2.1 Left or Right Classification

The left or right classification network was developed with great success. It achieved 100% accuracy for in-spec sides (that is, in-tact carcase beef sides meeting the product specifications for the rib cutting system). Due to the simplicity of the task and its high accuracy, it was implemented ahead of the RFID left or right label.

4.2.2 Rib 1 Segmentation

The rib 1 segmentation was trained separately to the rest of the ribs since a higher level of accuracy is needed in finding the rib 1 junction. The success of this network can be seen below with the accurate segmentation of a very variable dataset.

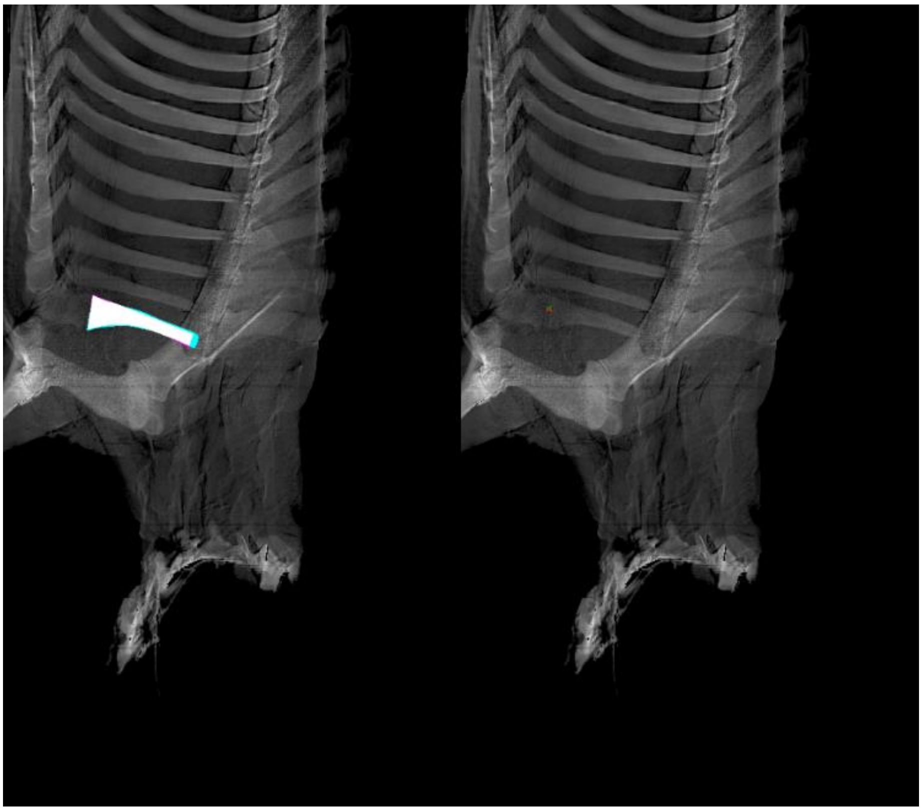

Figure 7 Rib 1 Segmentation Prediction. Left - full rib segmentation. Right – Original Image

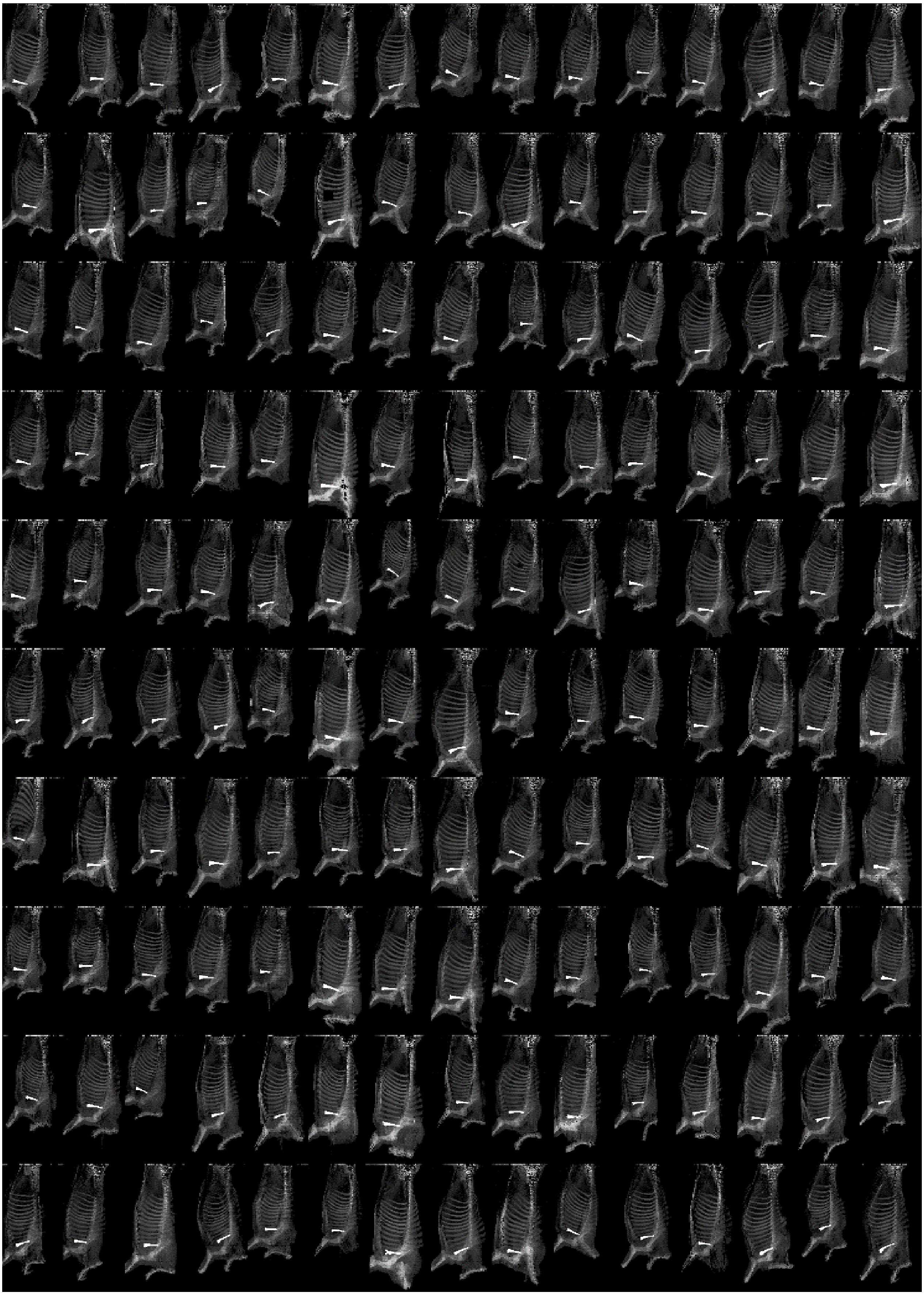

Figure 8 Rib 1 Predictions over Variable Dataset

4.2.3 Rib 2-10 Segmentation

The segmentation accuracy of ribs two to ten can be seen in Figure 9. Cleanly segmented ribs allows for the ability to count the number of ribs to be cut and the location of these ribs. If ribs are missing

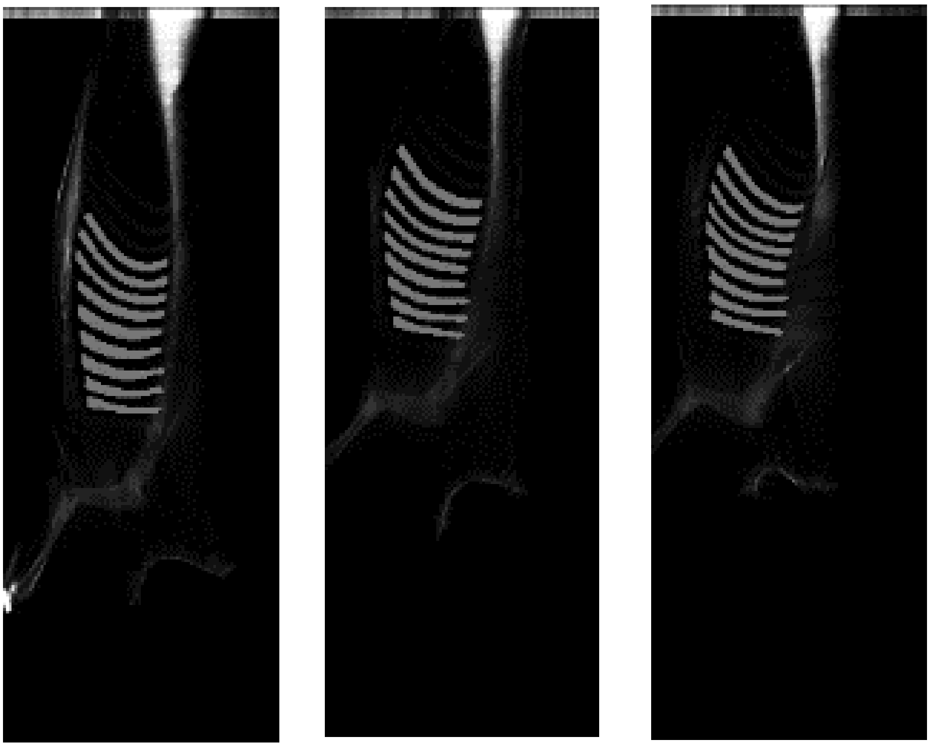

Figure 9 Rib 2 to 10 Predictions Overlayed on Original Images

or broken it could cause errors in the counting of these ribs as well as the performance of the network as it is trained to expect a full, intact rib cage.

4.3 Software Platform Integration

Comparison to the original algorithm was used to determine the success of the software platform integration. Firstly, a line along the rib 1 junction was manually labelled for 200 sides from when the old algorithm was performing at its best. Then a line is extended from the Aitch bone through the predicted rib 1 junction and the minimum distance between the two lines is calculated. In the example below the minimum distance would be 0 as the two lines intersect.

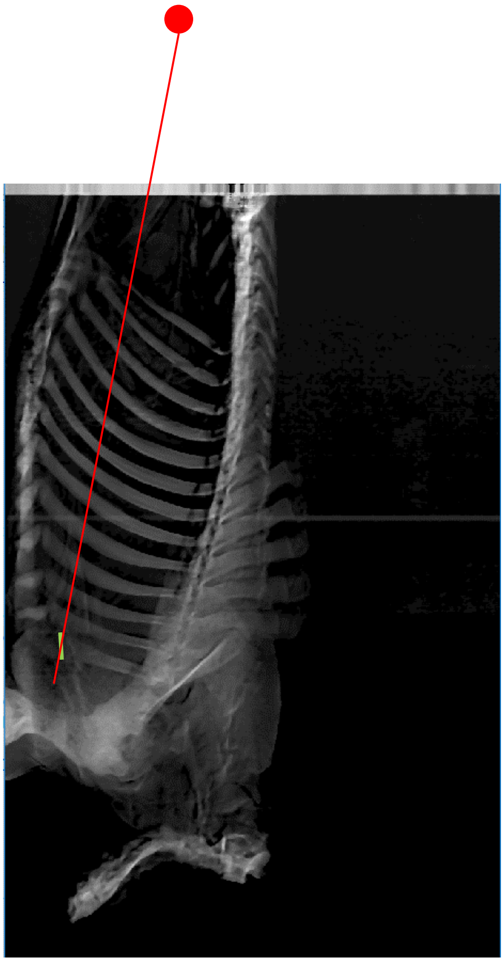

Figure 10 Accuracy Measurement Method

The table below shows a significant improvement in accurately locating the rib 1 junction compared to the original algorithm used in this system. The 95th percentile means that 95% of the dataset has a minimum distance of less than the number of pixels shown. The largest minimum distance is the 100th percentile of the results or the worst case result in this dataset. It can be seen that errors in the original algorithm have been reduced by over 400%.

	95th percentile of minimum distances (pixels)	Largest minimum distance (pixels)
Old Algorithm	28	58.2
New Neural Networks	6.4	8

Table 1 Overall Accuracy Comparison

5 Conclusions/Recommendations

Overall this milestone was completed successfully with the return to production of the rib scribing system with very successful rate of average 96.6% cell cut and potential cut of **99.7%** if there were no longus colli, flipped sides, double ups or other issues outside scope of project. It can be seen through statistics gathered over multiple weeks and months that the new algorithms vastly outperform the previous solution in both accuracy and cut rate. Although there were some difficulties in commissioning the system these were not due to the upgrades made on the software platform. Various sensor and hardware issues arose but were subsequently solved. This shows that although the neural network solution can greatly increase performance regular maintenance is still required to benefit from these gains.