Subcutaneous Fat Depth Regression Using Hyperspectral and Depth Imaging

Fernando Peña[†,‡*], Jasprabhjit Mehami[†,*], Raphael Falque[†], Timothy Patten[†], Alen Alempijevic[†] and Teresa Vidal-Calleja[†]
[†]Robotics Institute, University of Technology Sydney, NSW, Australia
[‡] Universidad de Zaragoza, Spain
fpena@unizar.es, jasprabhjit.mehami@student.uts.edu.au
{raphael.guenot-falque, timothy.patten, alen.alempijevic, teresa.vidalcalleja}@uts.edu.au

Abstract

Robotic perception is becoming an important component for automation in the meat processing industry. Whether for contaminant detection or automatic cutting, multimodal perception systems, in particular, based on hyperspectral imaging have the ability to provide information that goes beyond the texture and colour of a surface. In this paper, we present a learning-based method to estimate subcutaneous fat depth in meat cuts by leveraging hyperspectral data models that rely on the knowledge of modelled light sources and surface shape information. Data from a fully calibrated hyperspectral and colour depth (RGB-D) camera system is used as input. Fat depth ground truth is recovered via a novel systematic approach that ray casts a computed tomography (CT) mesh of the meat cuts, which is non-rigidly aligned with a depth reconstruction captured by the RGB-D camera. We thus evaluate machine learning methods that can handle small datasets, by employing dimensionality reduction and data augmentation to address the limited amount of imbalanced data that is acquired. Our results show that leveraging shape and light models, coupled with machine learning methods that capture nonlinearities and spatial correlations produces the most accurate results.

1 Introduction

Intelligent robotics and automation in the Australian meat processing industry is gaining interest for applications like contaminant detection, bruising or animal body composition estimation or even carcass handling, cutting, deboning, trimming, and packaging [de Medeiros Esper *et al.*, 2020]. Robotics perception systems, in particular comprising multiple sensing modalities, are key to providing autonomy and intelligence to these processing plants.

Figure 1: Top: Surface reconstruction of the muscle (in red) and fat (in white) from the computed tomography (CT) scans by thresholding the CT data and applying marching cubes [Adrian, 2020]. Bottom: Reprojected hyperspectral composite image overlaid onto the RGB image.

Perception systems that capture the visible range of the electromagnetic spectrum typically limit the information to the surface level. Meat and other food product applications might require information under the surface such as to detect the presence of contaminants, defects or bruising. A promising sensing modality for food processing applications is hyperspectral imaging. Hyperspectral cameras capture information in the near-infrared (NIR) range of the electromagnetic spectrum that cannot be seen by conventional colour cameras. In

[*]Authors contributed equally to this work.

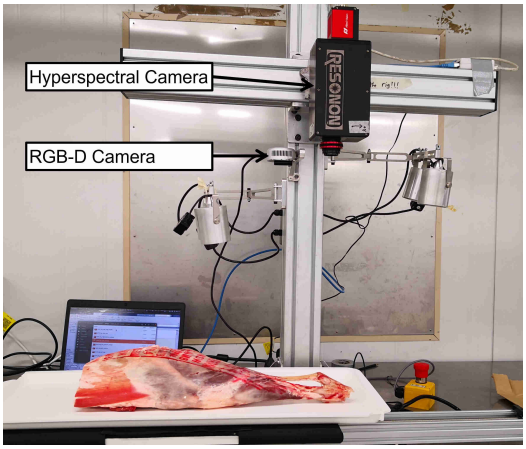


Figure 2: Acquisition rig, rigidly-mounted camera system comprised of a Resonon Pika NIR-320 hyperspectral line-scan camera and an Intel Realsense L515 RGB-D camera. The specimen (lamb leg), moves linearly under the camera system.

particular, through hyperspectral imaging, it is possible to gather internal (under the surface) information about certain meat or food products and correlate it with the spectrum of reflected light [Wendel *et al.*, 2018; Yuwono *et al.*, 2021].

In this paper, we are interested in using a multimodal perception system based on hyperspectral imaging to estimate internal information about meat products. Specifically, our focus is on estimating subcutaneous fat depth in lamb cuts given hyperspectral, surface shape and light source models. Our perception system, shown in Figure 2, consists of a line-scan hyperspectral camera, a depth camera and a light source, all fully calibrated.

The main challenge of the task at hand is the lack of data available to evaluate and, more importantly, to develop machine learning methods to regress fat depth values based on imagery data. Note that ground truth fat depth values such as the one shown in Figure 1 for a full meat cut can only be obtained via ultrasound or CT.

Thus, we propose a new systematic approach to obtain subcutaneous fat depth by ray casting the CT scans of the lamb cuts and registering them via non-rigid alignment with a 3D reconstruction obtained from our camera system.

Given our perception system and the ground truth data of the meat cuts, we also present a pipeline following our previous work [Mehami *et al.*, 2022] that leverages light source modelling, shape reconstruction and hyperspectral data together, and classical and deep learning models for fat depth regression. We evaluate regression methods that have the ability to deal with small datasets and use dimensionality reduction for the hyperspectral data based on [Peña *et al.*, 2022].

In summary, the contributions of this paper are,

- A subcutaneous fat depth estimation methodology for lamb cuts given hyperspectral, shape and modelled light source information and limited ground truth data.
- An approach to obtain subcutaneous fat depth ground truth with CT scanned lamb cuts.
- A benchmarking study that analyses the performance of different regression models.

2 Related Work

Hyperspectral cameras measure radiance, which typically depends on a number of unknown factors such as surface shape, material properties, viewing direction of the camera, and incident illumination. A material property that is invariant to these changes is the reflectance [Rahman *et al.*, 2018; Robles-Kelly and Huynh, 2012; Krebs *et al.*, 2020; Mehami *et al.*, 2022], which can be estimated from the radiance through either a calibration or an optimisation process. The calibration process performs a normalisation of the radiance between dark and white references, where the result is referred to as calibrated reflectance. The optimisation process assumes that the reflections of a material can be described by a model, such as the dichromatic reflectance model (DRM), where a cost function can be constructed in order to recover the reflectance. This is referred to as the estimated reflectance. The calibrated reflectance has shown to be less invariant to changes in lighting and across materials with changing surface shape [Mehami *et al.*, 2022; Rahman *et al.*, 2018; Robles-Kelly and Huynh, 2012].

Once calibrated or estimated reflectance values are obtained, this data is analysed to infer the properties of the phenomena that have been observed. Most commonly, machine learning is used to understand the relationship between reflectance values and the type or proportion of specific materials in the hyperspectral images. While the classification of land types from aerial or satellite-borne hyperspectral cameras has been the main driver due to the establishment of a number of standard datasets [Paoletti *et al.*, 2019; Li *et al.*, 2019], the diversity of applications ranges from the study of tumours in medicine [Weijtmans *et al.*, 2019; Trajanovski *et al.*, 2021] to the detection of weeds amongst crops [Wendel and Underwood, 2016; Eddy *et al.*, 2014]. For these applications, the hyperspectral data is used to semantically segment classes with similar approaches developed for RGB images [Ronneberger *et al.*, 2015], either on individual pixel level or

densely for all pixels in a hypercube. Learning continuous values through regression from hyperspectral data is less commonly studied but still spans a broad range of applications. Torres and Amigo [Torres and Amigo, 2019] investigate the capability of different models in food and pharmaceutical production.

In [Maier and Keller, 2018], multiple regression models are applied to learn the content and parameters in samples of inland water. Similarly, the performance of multiple regression models is evaluated for bathymetry estimation [Alevizos, 2020]. More recently, Xin et al. [Xin *et al.*, 2020] incorporate deep learning to estimate the depth of cadmium residue on lettuce leaves. A deep-stacked autoencoder is employed to extract spectral features that are passed to a least squares support vector machine to learn the regression targets. Finally, and most similar to our work, end-to-end neural network architectures are employed to quantify specific properties of objects or phenomena scanned with a hyperspectral camera. For example, convolutional neural networks (CNNs) are used to learn the moisture content of sea cucumbers [Yuwono *et al.*, 2021], the dry matter content of mangoes as a means of estimating their maturity [Wendel *et al.*, 2018] or the concentrations of cyanobacteria pigments in water systems [Pyo *et al.*, 2019].

The influence of subcutaneous fat depth in humans and diffuse reflectance in the NIR bands was studied by Yang et. al. [Yang *et al.*, 2005], which showed that a nonlinear relationship exists. This was further validated for lamb short loins [Rahman *et al.*, 2018]. In both cases, the surfaces examined are flat, and the sensor and light source are fixed. We aim for an approach that can deal with curved surfaces, such as a whole carcass, and a changing spatial configuration w.r.t the surface during scanning.

3 Problem Statement

Given a calibrated camera system of a line-scan hyperspectral camera HS, a 2D frame colour depth (RGB-D) camera F, and a non-isotropic near-field disk light source LS as shown in Figure 3, together with a limited dataset with ground truth labels, the aim of this work is to estimate fat depth in lamb cuts. The fat depth estimation is formulated as a regression problem. The hyperspectral measurement is defined as a hypercube of dimension $n \times m \times \lambda \in \mathbb{R}^3$, where n and m are spatial components, and λ is a spectral component. The predicted fat depth is a $n \times m \in \mathbb{R}^2$ matrix and the regression of the fat depth is defined as the estimation of a function $f : \mathbb{R}^3 \mapsto \mathbb{R}^2$.

4 Hyperspectral Reflectance

Following Figure 3, I_{HS} is the line image of HS and I_F represents both the 2D colour and depth images of F,

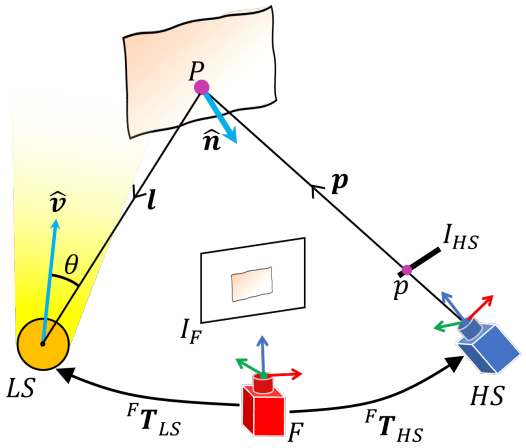


Figure 3: The frame camera (F), line-scan hyperspectral camera (HS), and light source (LS). The light vector carries the radiant flux, which is then reflected off the surface and measured by the hyperspectral camera.

where the depth image is aligned to the colour image. The camera system is fully calibrated using the active method outlined by [Mehami *et al.*, 2020]. $^F\boldsymbol{T}_{HS} \in \mathbb{R}^{4\times4}$ is the fixed homogeneous transformation of HS w.r.t to F which comes from the extrinsic parameters. A non-isotropic near-field disk light source LS is rigidly attached to this camera system via the known transformation $^F\boldsymbol{T}_{LS} \in \mathbb{R}^{4\times4}$.

The non-isotropic distribution of irradiance produced by the light source is modelled using a Gaussian Process (GP) following the methods of [Mehami *et al.*, 2022]. $\hat{\boldsymbol{v}} \in \mathbb{R}^3$ represents the principal direction of LS which is also inline with its Z-axis. The light vector $\boldsymbol{l} \in \mathbb{R}^3$ represents the ray that is emitted by LS and strikes a 3D point P on a sample with the surface normal $\hat{\boldsymbol{n}} \in \mathbb{R}^3$. $\boldsymbol{l}$ is defined as originating from P in the direction towards LS to ensure the dot product between $\boldsymbol{l}$ and $\hat{\boldsymbol{n}}$ is positive [Ma *et al.*, 2019]. The radiance from P is measured by HS via pixel p along its pixel ray $\boldsymbol{p}$ which originates from the camera origin.

4.1 Hyperspectral Ray Casting

For the radiance measurement of p by HS, the 3D coordinate P must be determined in order to estimate surface normals and the irradiance from LS. P can be determined by the intersection of $\boldsymbol{p}$ with the sample that is being viewed. Given that there is a depth image of the sample from F, the corresponding point cloud can be easily acquired. A 3D occupancy grid map is then constructed using the point cloud, which can then be used for ray casting as shown in Figure 4. Once P is acquired, the camera calibration parameters can be used to reproject the hyperspectral measurements to I_F as shown in Figure 1.

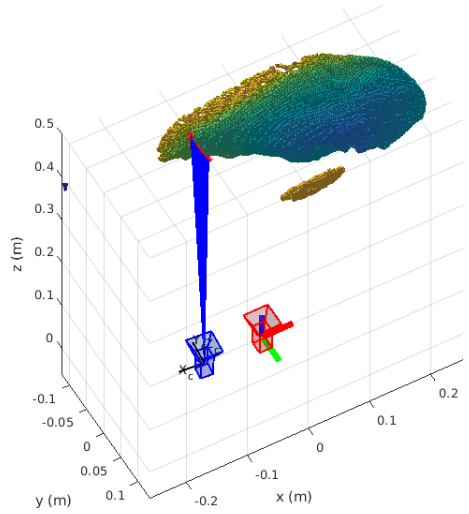


Figure 4: Ray casting 3D occupancy grid map of lamb cut. The camera in red is the frame camera which is the world coordinate system. The blue camera is the hyperspectral camera. The rays originating from the hyperspectral are shown in blue and the intersection points on the cut are shown as red crosses.

4.2 Estimating the Irradiance

The irradiance E arriving on the surface of the sample is assumed to only come from LS and is estimated using a modelled Gaussian process (GP) distribution of LS as follows:

$$E(\mathbf{x}) \sim GP(m(\mathbf{x}), K(\mathbf{x}, \mathbf{x}')) \,. \tag{1}$$

We use the light-src-mean function described by Mehami [Mehami *et al.*, 2022] as the mean function $m(\cdot)$, and a standard squared exponential (SE) as the covariance function $K(\cdot)$. The coordinate $\mathbf{x} = [r, \theta]$ describes the 2D coordinate of $P_{LS}(x, y, z)$, which is P in the coordinate system of LS. The 2D coordinate system is defined as follows:

$$\begin{aligned} r &= \|\boldsymbol{l}\| \\ \theta &= \arctan2\left(y, \sqrt{x^2 + z^2}\right) \,. \end{aligned} \tag{2}$$

The irradiance arriving at all points P on the sample can be estimated from this GP model once converted to the 2D coordinate system. A map of the normalised irradiance for a lamb cut sample is shown in Figure 5 for a single spectral band. Alongside the irradiance, the light source direction vectors $\hat{\boldsymbol{l}}$ can be estimated for all P using the known pose of LS. By calculating the dot product between $\hat{\boldsymbol{l}}$ and $\hat{\boldsymbol{n}}$, a property that is proportional to the shading can be acquired, where the shading describes the interaction of the irradiance with the surface geometry. Maps of the $\hat{\boldsymbol{l}}$, $\hat{\boldsymbol{n}}$ and their respective dot product are shown in Figure 6.

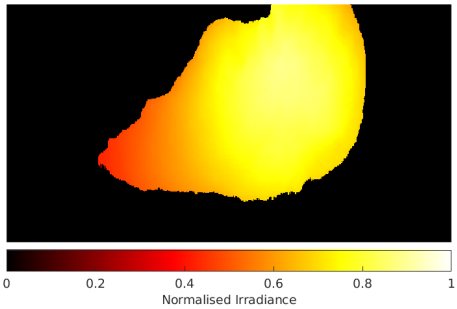


Figure 5: Normalised irradiance for a lamb cut sample estimated from the GP light-source model for a single spectral band.

(a) (b) (c)

Figure 6: The variation in shape and lighting for a lamb cut sample. (a) The surface normals, (b) the light source direction map, and (c) the resulting dot product of (a) and (b) which relates to the shading.

4.3 Estimating the Reflectance

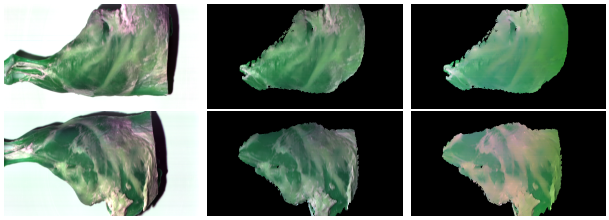
The reflectance is estimated using three different methods: calibration, Krebs [Krebs *et al.*, 2020], and Mehami [Mehami *et al.*, 2022]. The last two methods are formulated as an optimisation problem by assuming the reflections of the sample material are described by the DRM. Krebs only requires irradiance, and Mehami requires both irradiance, surface normals, and light source direction vectors. The resulting reflectance is calculated for two different lamb cut samples using the three methods. The false colour images of the reflectances are shown in Figure 7. The reflectance estimated by Mehami's method was found to show the least texture which is expected from the reflectance property.

5 Fat Depth Ground Truth

The ground truth fat depth of the lamb cut samples is obtained through CT scans. A CT measures the X-rays attenuation and is an ideal modality for distinguishing muscle from fat components [Borkan *et al.*, 1982; Goodpaster *et al.*, 2000]. The CT measurements are in the form of a series of sequential 2D slices. They are stacked into a 3D tensor to build a 3D grid of the X-ray absorption. The voxels have a size of $0.56 \times 0.56 \times 0.7$ millimetres and are stored as DICOM images.

5.1 CT Reconstructions

To acquire the subcutaneous fat depth of the lamb cut samples and align CT scans to the data of the camera



(a) (b) (c)

Figure 7: Outputs for the reflectance estimation of both lamb cut samples, top row is for the left cut and bottom row is for the right cut. Columns: (a) calibrated, (b) Krebs [Krebs *et al.*, 2020], and (c) Mehami [Mehami *et al.*, 2022]. Note that the reflectance is presented as a false colour image.

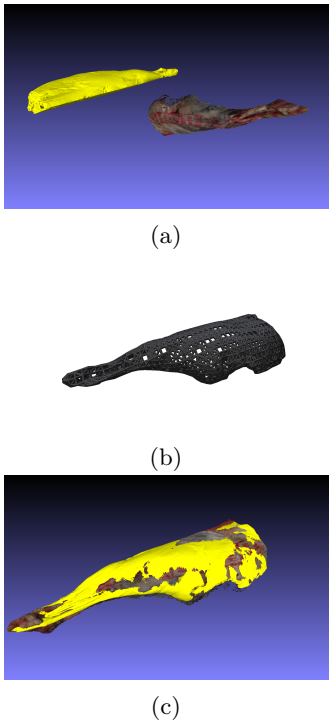
system shown earlier, 3D reconstructed meshes were created from the 2D CT scan slices. This is achieved using marching cubes algorithm with thresholds on the voxel absorption [Adrian, 2020]. Three meshes are created which isolate the fat, muscle and outer contour of the lamb cut. The fat and muscle are shown in Figure 1.

5.2 CT and Depth Alignment

The same sample has been measured separately using a CT scanner and the camera system. The CT scans are reconstructed into meshes and the RGB-D camera from the camera system allows us to create a 3D depth reconstruction. Ideally, these two separate 3D models of the sample should perfectly align in a rigid manner. However, this is not what happens in practice, as slight deformations occur when transporting the sample between the hyperspectral rig and the CT scanner. Further sagging is caused by changes in temperature and moisture. To solve this problem, the outer contour is aligned to the depth reconstruction in a non-rigid manner by using embedded deformation [Sumner *et al.*, 2007] with manually selected 3D correspondences. The deformation is then applied to the fat and muscle meshes to align them to the depth reconstruction. The alignment of the CT fat mesh to the depth reconstruction using the deformation is shown in Figure 8.

5.3 Fat Depth Ray Casting

The fat depth must be acquired for all radiance measurements at p. This is estimated by ray casting the intersection of $\boldsymbol{p}$ with the two aligned meshes of fat and muscle resulting in the intersection points P_f and P_m. If the intersections are valid, the distances between HS and the fat and muscle intersection points are determined as d_f and d_m respectively. There are five outcomes that can occur based on the combinations of intersections and the calculated distances. These are summarised in Table 1. For a $\boldsymbol{p}$, where both muscle and fat intersections occur



(a) (b) (c)

Figure 8: Non-rigid alignment of CT mesh to depth reconstruction using embedded deformation. (a) misaligned CT fat mesh shown in yellow and depth reconstruction, (b) deformation graph of CT fat mesh, and (c) aligned CT fat mesh to depth reconstruction.

and $d_f < d_m$, the fat depth is calculated by performing a nearest-neighbour search on the muscle mesh at P_f as shown in Figure 9.

6 Experiments

In this section, we evaluate the capability of various machine learning models for performing fat depth regression on hyperspectral data, and analyse how they are able to leverage the proposed light modelling.

6.1 Evaluation Protocol

The evaluation uses the data of the two cuts of lamb shown in Figure 7. One cut is used for training and

Table 1: Different outcomes of ray casting onto the CT meshes.

P_f Found	P_m Found	$d_f < d_m$	Outcome
✗	✗	-	No intersection
✗	✓	-	Only muscle
✓	✗	-	Only fat
✓	✓	False	Fat under muscle
✓	✓	True	Fat depth

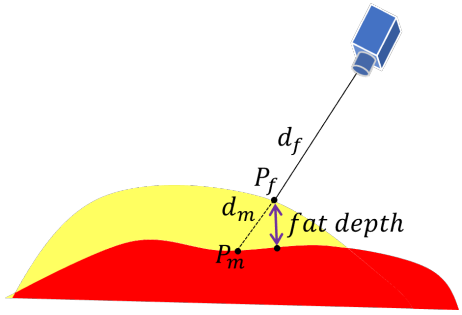


Figure 9: Ray casting from the hyperspectral camera to the fat (yellow) and muscle (red) meshes showing the intersection points and their corresponding distances. The actual fat depth is the distance highlighted by the purple arrow.

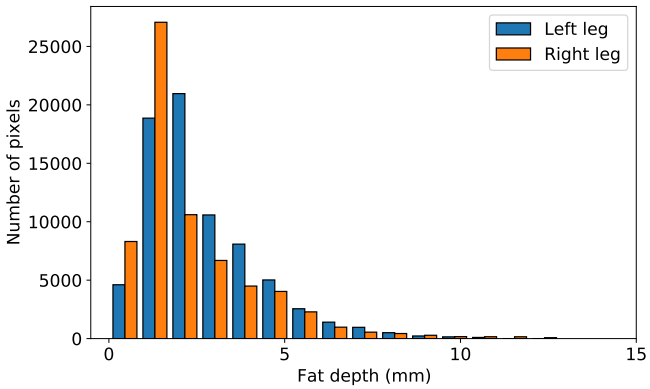


Figure 10: Histogram of fat depth in millimetres for both lamb leg samples showing the imbalance in the data.

the other for testing. The different models are evaluated with the same pair of training and testing hypercubes. The hypercubes for each cut have a size of $320 \times 599 \times 164$, where each has approximately 66000 valid pixels for which reflectance and fat depth are known.

Performance is measured using two metrics. First, we consider the mean absolute error (MAE), which quantifies the overall prediction error by averaging the error for each pixel. In addition, we measure the normalised area under the curve (AUC) of the error with respect to the ground truth fat depth, which is computed by averaging the absolute error at different depth intervals. As such, the AUC quantifies the error on binarised samples, allowing to balance the different depth thicknesses of the test samples. This is especially important to consider for our dataset because the fat depth distribution has a long tail as shown in Figure 10.

6.2 Fat Depth Regression Methods

As a baseline, the fat depth is learned by linear regression. A significant portion of the observed fat depth is very shallow (approximately 1–2 mm) with only a few regions exhibiting deeper fat content (up to 14 mm). It is expected that linear regression can model the depth well for the majority of shallow measurements but at the expense of incorrect predictions for the deeper fat regions. A multilayer perceptron (MLP) is also evaluated as this provides higher capabilities to model the nonlinearities of the deep fat. The MLP consists of three hidden layers with output sizes $16 \times 10 \times 1$. Each layer employs rectified linear unit (ReLU) as activation function. Both linear regression and MLP learn from pixel reflectance inputs where each sample is a λ-dimensional normalised feature vector.

To process a full image and estimate the fat depth for each pixel simultaneously and leverage local spatial correlations, we use MiniNet-v2 [Alonso *et al.*, 2020], an efficient CNN-based architecture. MiniNet-v2 is originally developed for semantic segmentation of RGB images and has been shown to perform well on autonomous driving datasets but with a significant reduction in memory (parameters) and computation time compared to other state-of-the-art architectures. It makes use of multi-dilation depthwise convolutional layers and an additional convolutional branch instead of skip connections, which offer a good trade-off between accuracy and efficiency. Recently, MiniNet-v2 was applied to hyperspectral images for a remote sensing dataset in [Peña *et al.*, 2022].

We used the MiniNetv2-cpu variant because its lower complexity is appropriate for training with a reduced set of data. This model can also be used in CPU-only systems, which are common in robotics settings. For performing fat depth regression, we modify the architecture by replacing the final classification layer with a 1D convolutional layer, in order to predict a single continuous value instead of a discrete class. Both the MLP and the modified MiniNet-v2 are trained with L2 loss and a learning rate of 0.01 that decays at each epoch with a rate of 0.9.

MiniNet-v2 takes a hypercube as input instead of individual pixels. Since only one hypercube is available for training, we apply data augmentation to aid model generalisation. First, the image is rotated randomly, then a random crop is extracted and resized to the original dimensions, followed by random horizontal and vertical flips. Details of the data augmentation are provided in Table 2.

6.3 Dimensionality Reduction

Due to the high dimensionality of hyperspectral data, it is common to apply band selection or reduction techniques. These methods not only compress the data and reduce its overall size, but can also remove redundancies that are detrimental when training machine learning al-

Table 2: Parameters of the data augmentation pipeline used for training MiniNet-v2.

Augmentation type	Parameter	Value
Random rotation	Angle	$[-180°, 180°]$
	Interpolation	Bilinear
	Fill	Zeros
Random crop and resize	Crop scale	$[0.2, 1.0]$
	Interpolation	Bilinear
Random vertical flip	Probability	0.5
Random horizontal flip	Probability	0.5

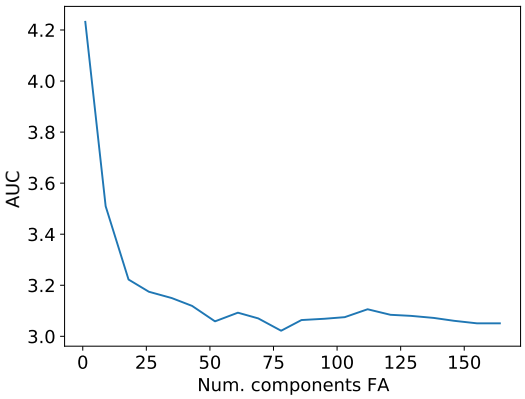


Figure 11: Linear regression using reflectance from Mehami's method with a varying number of factor analysis components.

gorithms. In line with [Peña *et al.*, 2022], this work applies factor analysis to reduce the data. Figure 11 shows the performance of the linear regression model trained with a varying number of components. It is clear the performance gain beyond 20 components is small, thus we use this value for all experiments as it is a good trade-off between performance and data size.

6.4 Learning Benchmarking

Tables 3 and 4 present quantitative results for fat depth regression on the tested lamb cut for the different regression models as well as for different reflectance data. The absolute error in fat depth across the lamb cut is shown in Figure 12 for the different reflectance data and regression models, with the ground truth fat depth. The distribution of the MAE was plotted against fat depth with 1-σ STD bounds as shown in Figure 13. In general, using the reflectance data from Mehami's method showed the least error in fat depth, regardless of the regression model. This is expected as it uses both lighting and shape information to estimate the reflectance. The learning-based approaches are shown to have the least error regardless of the reflectance data, and were able to best capture the regions with high fat depth, given the limited data samples. Interestingly, the linear regression model also has an increasing linear error with fat depth, highlighting its inability to capture the complex interactions between features.

7 Conclusion

This paper investigates the regression of subcutaneous fat depth from lamb cut samples using hyperspectral reflectance. The reflectance is estimated using three methods, where one involves radiance normalisation, and the other two are optimisation-based that assume that material reflections can be described by the DRM. One of the optimisation methods also incorporates shape and lighting information. For each hyperspectral reflectance pixel, the fat depth ground truth is obtained by ray casting on meshes reconstructed from CT data that are aligned to a depth reconstruction. After applying a dimensionality reduction to the reflectance data, three fat depth regression models are trained on a full lamb cut and tested on a separate cut. It was found that, overall, the learning-based model best captured the imbalanced fat depth distribution. The optimisation-based reflectance method that incorporated shape and lighting, has the least error for fat depth estimation.

The estimated reflectance is shown to be invariant to shape and lighting changes, but acquiring it is an expensive multi-step process. We will investigate the potential to learn invariance to shape and lighting by synthesising new hyperspectral measurements. This could be achieved by assuming that the material always obeys the DRM, where new radiance measurements can be acquired using the estimated reflectance of one captured dataset, modelled light sources and known surface shape.

Acknowledgments

This work was supported in part by an Australian Government Research Training Program (RTP) Scholarship, the University of Technology Sydney and by the Australian Government Department of Agriculture and Water Resources as part of its Rural R&D for Profit program under MLA grant number V.RDP.2005.

References

[Adrian, 2020] Schneider Adrian. DicomToMesh, 2020. https://github.com/AOT-AG/DicomToMesh.

[Alevizos, 2020] Evangelos Alevizos. A combined machine learning and residual analysis approach for improved retrieval of shallow bathymetry from hyperspectral imagery and sparse ground truth data. *Remote Sensing*, 12(21), 2020.

[Alonso *et al.*, 2020] Inigo Alonso, Luis Riazuelo, and Ana C Murillo. MiniNet: An efficient semantic

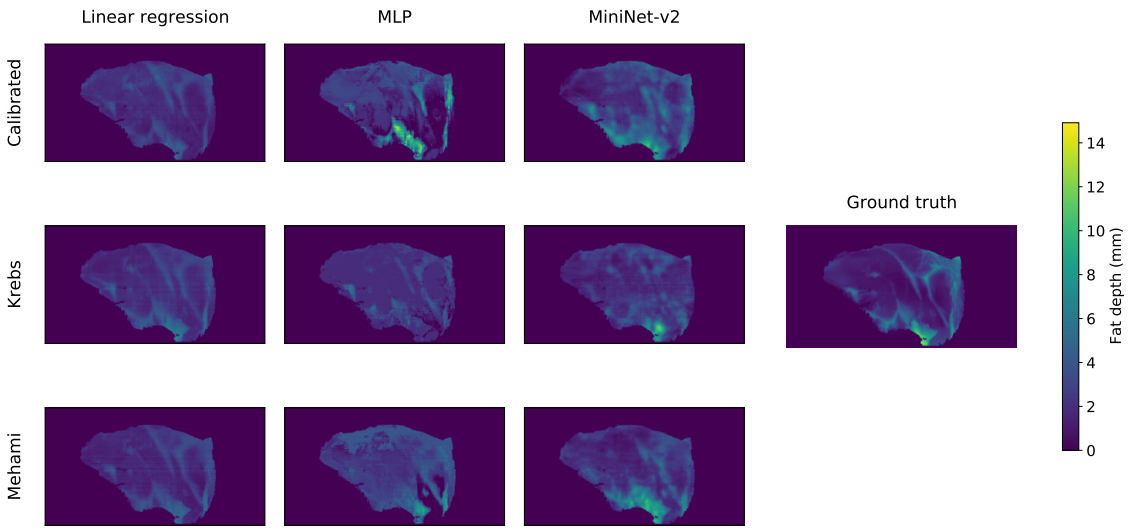


Figure 12: The absolute error in fat depth for the tested lamb cut sample. Rows from top to bottom represent the three different reflectance methods of Calibrated, Krebs, and Mehami. Columns from left to right represent the three regression models of linear regression, MLP, and MiniNet-v2.

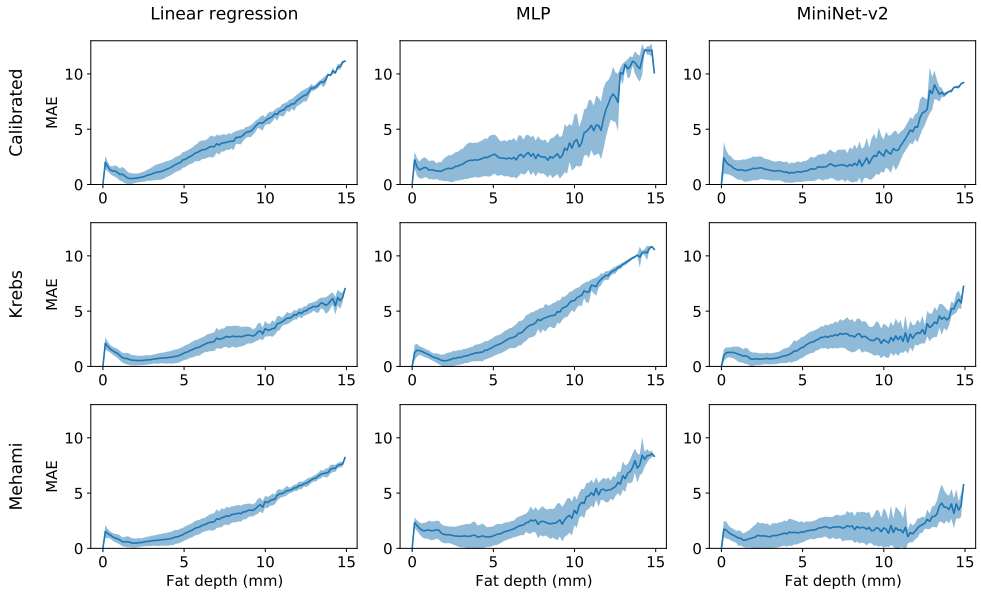


Figure 13: The plots of the MAE vs fat depth with 1-σ STD bounds. Rows from top to bottom represent the three different reflectance methods of Calibrated, Krebs, and Mehami. Columns from left to right represent the three regression models of linear regression, MLP, and MiniNet-v2.

Table 3: Mean absolute error (MAE) in millimetres for the three reflectance methods and three regression models.

	Irradiance	Normals	Shading	Linear regression	MLP	MiniNet [Alonso *et al.*, 2020]
Calibrated	-	-	-	1.16	1.64	1.40
Krebs [Krebs *et al.*, 2020]	✓	-	-	0.99	**1.24**	1.15
Mehami [Mehami *et al.*, 2022]	✓	✓	✓	**0.93**	1.52	**1.13**

Table 4: Normalised area under the curve (AUC) of the error vs. fat depth plots shown in Figure 13. This computes the error on binarised samples to account for the imbalance in fat depth.

	Irradiance	Normals	Shading	Linear regression	MLP	MiniNet [Alonso *et al.*, 2020]
Calibrated	-	-	-	3.53	4.16	2.10
Krebs [Krebs *et al.*, 2020]	✓	-	-	**2.74**	4.48	2.42
Mehami [Mehami *et al.*, 2022]	✓	✓	✓	3.18	**3.18**	**1.87**

segmentation convnet for real-time robotic applications. *IEEE Transactions on Robotics*, 36(4):1340–1347, 2020.

[Borkan *et al.*, 1982] Gary A Borkan, Stephen G Gerzof, Alan H Robbins, DE Hults, CK Silbert, and JE Silbert. Assessment of abdominal fat content by computed tomography. *The American journal of clinical nutrition*, 36(1):172–177, 1982.

[de Medeiros Esper *et al.*, 2020] Ian de Medeiros Esper, Alex Mason, et al. Robotisation and intelligent systems in abattoirs. *Trends in Food Science & Technology*, 2020.

[Eddy *et al.*, 2014] P.R. Eddy, A.M. Smith, B.D. Hill, D.R. Peddle, C.A. Coburn, and R.E. Blackshaw. Weed and crop discrimination using hyperspectral image data and reduced bandsets. *Canadian Journal of Remote Sensing*, 39(6):481–490, 2014.

[Goodpaster *et al.*, 2000] Bret H Goodpaster, F Leland Thaete, and David E Kelley. Composition of skeletal muscle evaluated with computed tomography. *Annals of the New York Academy of Sciences*, 904(1):18–24, 2000.

[Krebs *et al.*, 2020] Alexandre Krebs, Yannick Benezeth, and Franck Marzani. Intrinsic RGB and multispectral images recovery by independent quadratic programming. *PeerJ Computer Science*, 6, 2020.

[Li *et al.*, 2019] Shutao Li, Weiwei Song, Leyuan Fang, Yushi Chen, Pedram Ghamisi, and Jón Atli Benediktsson. Deep learning for hyperspectral image classification: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9):6690–6709, 2019.

[Ma *et al.*, 2019] Long Ma, Jirui Liu, Xin Pei, Yanmin Hu, and Fengming Sun. Calibration of position and orientation for point light source synchronously with single image in photometric stereo. *Optics Express*, 27(4):4024, 2019.

[Maier and Keller, 2018] Philipp M. Maier and Sina Keller. Machine learning regression on hyperspectral data to estimate multiple water parameters. In *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pages 1–5, 2018.

[Mehami *et al.*, 2020] Jasprabhjit Mehami, Teresa Vidal-Calleja, and Alen Alempijevic. Observability driven multi-modal line-scan camera calibration. *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 285–290, 2020.

[Mehami *et al.*, 2022] Jasprabhjit Mehami, Raphael Falque, Teresa Vidal-Calleja, and Alen Alempijevic. Multi-modal non-isotropic light source modelling for reflectance estimation in hyperspectral imaging. *IEEE Robotics and Automation Letters*, 7(4):10336–10343, 2022.

[Paoletti *et al.*, 2019] M.E. Paoletti, J.M. Haut, J. Plaza, and A. Plaza. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158:279–317, 2019.

[Peña *et al.*, 2022] Fernando Peña, Pilar Vidal Aguilar, Darío Suárez Gracia, and Ana C. Murillo. Efficient semantic segmentation with hyperspectral images. In *Proceedings of the Iberian Robotics Conference*, 2022. (to appear).

[Pyo *et al.*, 2019] JongCheol Pyo, Hongtao Duan, Sangsoo Baek, Moon Sung Kim, Taegyun Jeon, Yong Sung Kwon, Hyuk Lee, and Kyung Hwa Cho. A convolutional neural network regression for quantifying

cyanobacteria using hyperspectral imagery. *Remote Sensing of Environment*, 233:111350, 2019.

[Rahman *et al.*, 2018] S. Rahman, P. Quin, T. Walsh, T. Vidal-Calleja, M. J. McPhee, E. Toohey, and A. Alempijevic. Preliminary estimation of fat depth in the lamb short loin using a hyperspectral camera. *Animal Production Science*, 58(8):1488–1496, 2018.

[Robles-Kelly and Huynh, 2012] Antonio Robles-Kelly and Cong Phuoc Huynh. *Imaging spectroscopy for scene analysis*. Springer Science & Business Media, 2012.

[Ronneberger *et al.*, 2015] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In *Proceedings of International Conference on Medical Image Computing and Computer Assisted Intervention,*, pages 234–241, 2015.

[Sumner *et al.*, 2007] Robert W. Sumner, Johannes Schmid, and Mark Pauly. Embedded deformation for shape manipulation. In *ACM siggraph 2007 papers*, pages 80–es, 2007.

[Torres and Amigo, 2019] Irina Torres and José Manuel Amigo. An overview of regression methods in hyperspectral and multispectral imaging. In José Manuel Amigo, editor, *Hyperspectral Imaging*, volume 32 of *Data Handling in Science and Technology*, pages 205–230. Elsevier, 2019.

[Trajanovski *et al.*, 2021] Stojan Trajanovski, Caifeng Shan, Pim J. C. Weijtmans, Susan G. Brouwer de Koning, and Theo J. M. Ruers. Tongue tumor detection in hyperspectral images using deep learning semantic segmentation. *IEEE Transactions on Biomedical Engineering*, 68(4), 2021.

[Weijtmans *et al.*, 2019] P.J.C. Weijtmans, C. Shan, T. Tan, S.G. Brouwer de Koning, and T.J.M. Ruers. A dual stream network for tumor detection in hyperspectral images. In *IEEE 16th International Symposium on Biomedical Imaging*, pages 1256–1259, 2019.

[Wendel and Underwood, 2016] A. Wendel and J. Underwood. Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In *Proceedings of the International Conference on Robotics and Automation*, pages 5128–5135, 2016.

[Wendel *et al.*, 2018] Alexander Wendel, James Underwood, and Kerry Walsh. Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. *Computers and Electronics in Agriculture*, 155:298–313, 2018.

[Xin *et al.*, 2020] Zhou Xin, Sun Jun, Tian Yan, Chen Quansheng, Wu Xiaohong, and Hang Yingying. A deep learning based regression method on hyperspectral data for rapid prediction of cadmium residue in lettuce leaves. *Chemometrics and Intelligent Laboratory Systems*, 200:103996, 2020.

[Yang *et al.*, 2005] Ye Yang, Olusola O. Soyemi, Michelle R. Landry, and Babs R. Soller. Influence of a fat layer on the near infrared spectra of human muscle: quantitative analysis based on two-layered Monte Carlo simulations and phantom experiments. *Optics express*, 13(5):1570–1579, 2005.

[Yuwono *et al.*, 2021] Hendra Angga Yuwono, Adhi Harmoko Saputro, and Sabar. Hyperspectral and deep learning-based regression model to estimate moisture content in sea cucumbers. In *Proceedings of the International Conference on Electrical Engineering, Computer Science and Informatics*, pages 283–287, 2021.